(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0074104
(43) 공개일자 2020년06월24일

(51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) *C12N 15/10* (2017.01)
C12N 9/78 (2006.01)
(52) CPC특허분류
C12N 9/22 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2020-7010211
(22) 출원일자(국제) 2018년10월04일
심사청구일자 없음
(85) 번역문제출일자 2020년04월08일
(86) 국제출원번호 PCT/US2018/054469
(87) 국제공개번호 WO 2019/071048
국제공개일자 2019년04월11일
(30) 우선권주장
62/568,313 2017년10월04일 미국(US)

(71) 출원인
더 브로드 인스티튜트, 인코퍼레이티드
미국 02142 매사 추세츠주 캠브리지 메인 스트리트 415
프레지던트 앤드 펠로우즈 오브 하바드 칼리지
미합중국, 메사추세츠 02138, 캠브리지, 퀸시스트리트17
매사추세츠 인스티튜트 오브 테크놀로지
미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 77
(72) 발명자
장 펑
미국 02142 매사추세츠주 캠브리지 메인 스트리트 415 씨/오
구텐버그 조나단
미국 02138 매사추세츠주 캠브리지 퀸시 스트리트 17 씨/오
아부다예 오마르
미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 77 씨/오
(74) 대리인
특허법인코리아나

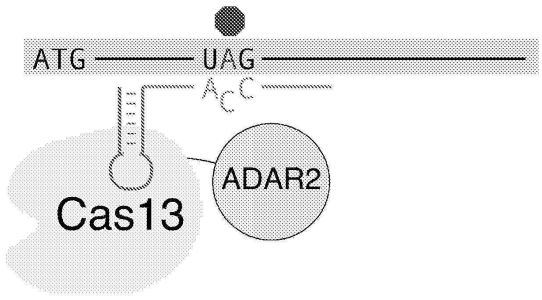
전체 청구항 수 : 총 92 항

(54) 발명의 명칭 **표적화된 핵산 편집을 위한 시스템, 방법, 및 조성물**

(57) 요 약

본 발명은 핵산을 표적화 및 편집하기 위한 시스템, 방법, 및 조성물을 제공한다. 특히, 본 발명은 RNA-표적화 Cas13 단백질, 적어도 하나의 가이드 분자, 및 적어도 하나의 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 포함하는 비천연 발생 또는 조작된 RNA-표적화 시스템을 제공한다.

대 표 도 - 도1



(52) CPC특허분류
C12N 9/78 (2013.01)
C12Y 305/04004 (2013.01)
C12Y 305/04005 (2013.01)
C12N 2310/20 (2017.05)

명 세 서

청구범위

청구항 1

부위 지정된 염기 편집을 위한 조작된 조성물로서,

a. 표적화 도메인; 및

b. 아데노신 데아미나제 또는 이의 촉매 도메인으로서, 아데노신 데아미나제는 활성이 시티딘 데아미나제로 전환되도록 변형된 것인 아데노신 데아미나제 또는 이의 촉매 도메인

을 포함하는, 조성물.

청구항 2

제1항에 있어서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376, S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, V525 및 I520으로부터 선택되는 하나 이상의 위치에서 하나 이상의 돌연변이에 의해 변형되는 것인 조성물.

청구항 3

제1항에 있어서, 아데노신 데아미나제는 E488, V351, S486, T375, S370, P462, 및 N597로부터 선택되는 하나 이상의 위치에서 돌연변이되는 것인 조성물.

청구항 4

제1항에 있어서, 아데노신 데아미나제는 E488Q, V351G, S486A, T375S, S370C, P462A, 및 N597I로부터 선택되는 하나 이상의 돌연변이를 포함하는 것인 조성물.

청구항 5

제1항에 있어서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 인간, 두족류 (cephalopod), 또는 초파리 (Drosophila) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인인 조성물.

청구항 6

제1항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 hADAR2-D 아미노산 서열의 글루탐산^{488}, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함하도록 변형된 것인 조성물.

청구항 7

제6항에 있어서, 위치 488 또는 상동성 ADAR 단백질 내 상응하는 위치의 상기 글루탐산 잔기는 글루타민 잔기 (E488Q)로 치환되는 것인 조성물.

청구항 8

제6항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 돌연변이 E488Q를 포함하는 돌연변이된 hADAR2d 또는 돌연변이 E1008Q를 포함하는 돌연변이된 hADAR1d인 조성물.

청구항 9

제1항 내지 제4항 중 어느 한 항에 있어서, 표적화 도메인은 촉매적 불활성 Cas13 단백질이거나, 또는 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열인 조성물.

청구항 10

제9항에 있어서, 촉매적 불활성 Cas13 단백질은 촉매적 불활성 Cas13a, 촉매적 불활성 Cas13b, 또는 촉매적 불

활성 Cas13c인 조성물.

청구항 11

제9항에 있어서, 촉매적 불활성 Cas13 단백질은 임의의 표 1, 2, 3, 또는 4에 열거된 박테리아 종으로 이루어진 군으로부터 선택되는 박테리아 종에서 유래된 Cas13 뉴클레아제로부터 수득되는 것인 조성물.

청구항 12

제1항 내지 제4항 중 어느 한 항에 있어서, 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자, 또는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열을 더 포함하는 것인 조성물.

청구항 13

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 표적화 도메인에 공유적으로 또는 비공유적으로 연결되는 것인 조성물.

청구항 14

관심 표적 유전자좌 내 아데닌을 변형시키기에 적합한 조작된, 비천연 발생 시스템으로서,

(a) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자, 또는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열;

(b) 촉매적 불활성 Cas13 단백질, 또는 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열; 및

(c) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인, 또는 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열로서, 아데노신 데아미나제는 활성이 시티딘 데아미나제로 전환되도록 변형되는 것인, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인, 또는 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열

을 포함하고,

상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후에 그에 연결되도록 적합화되고;

상기 가이드 서열은 아데닌을 포함하는 표적 RNA 서열과 하이브리드화되어 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여, 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시키는 것인, 조작된, 비천연 발생 시스템.

청구항 15

제14항에 있어서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376, S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, V525 및 I520으로부터 선택되는 하나 이상의 돌연변이에 의해 변형되는 것인 조작된, 비천연 발생 시스템.

청구항 16

제14항에 있어서, 아데노신 데아미나제는 E488, V351, S486, T375, S370, P462, 및 N597로부터 선택되는 하나 이상의 위치에서 돌연변이되는 것인 방법.

청구항 17

관심 표적 RNA 서열 내 아데닌을 변형시키는 방법으로서, 상기 표적 RNA에,

(a) 촉매적 불활성 (데드) Cas13 단백질;

(b) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자; 및

(c) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인

을 전달하는 단계를 포함하고,

상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후에 그에 연결되도록 적합화되고;

상기 가이드 분자는 상기 데드 Cas13 단백질과 복합체를 형성하고 상기 복합체가 상기 관심 표적 RNA 서열에 결합하도록 유도시키며, 상기 가이드 서열은 상기 아데닌을 포함하는 표적 서열과 하이브리드화하여 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여서 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시키고;

상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 RNA 듀플렉스 내 상기 아데닌을 탈아미노화시키는 것인, 변형 방법.

청구항 18

제17항에 있어서, 상기 Cas13 단백질은 Cas13a, Cas13b, 또는 Cas13c인 변형 방법.

청구항 19

제17항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질의 N-말단 또는 C-말단에 융합되는 것인 변형 방법.

청구항 20

제19항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 링커에 의해 상기 데드 Cas13 단백질에 융합되는 것인 변형 방법.

청구항 21

제20항에 있어서, 상기 링커는 $(GGGGS)_{3-11}$, GSG_5 또는 LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR이거나, 또는 상기 링커는 XTEN 링커인 변형 방법.

청구항 22

제17항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 어댑터 단백질 및 상기 가이드 분자에 연결되거나 또는 상기 데드 Cas13 단백질은 상기 어댑터 단백질에 결합할 수 있는 압타머 서열을 포함하는 것인 변형 방법.

청구항 23

제22항에 있어서, 상기 어댑터 서열은 MS2, PP7, Qβ, F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ϕCb5, ϕCb8r, ϕCb12r, ϕCb23r, 7s, 및 PRR1로부터 선택되는 것인 변형 방법.

청구항 24

제23항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질의 내부 루프에 삽입되는 것인 변형 방법.

청구항 25

제24항에 있어서, 상기 Cas13 단백질은 Cas13a 단백질이고 상기 Cas13a는 2개 HEPN 도메인 내에, 특히 렙토트리키아 웨이데이 (*Leptotrichia wadei*)로부터 기원하는 Cas13a 단백질의 위치 R474 및 R1046, 또는 Cas13a 오솔로그의 이에 상응하는 아미노산 위치에 하나 이상의 돌연변이를 포함하는 것인 변형 방법.

청구항 26

제25항에 있어서, 상기 Cas13 단백질은 Cas13b 단백질이고 상기 Cas13b는 베르게이엘라 주헬컴 (*Bergeyella zoohelcum*) ATCC 43767로부터 기원하는 Cas13b 단백질의 위치 R116, H121, R1177, 또는 H1182, 또는 Cas13b 오솔로그의 이에 상응하는 아미노산 위치 중 하나 이상에 돌연변이를 포함하는 것인 변형 방법.

청구항 27

제26항에 있어서, 상기 돌연변이는 베르게이엘라 주헬컴 ATCC 43767로부터 기원하는 Cas13b 단백질의 R116A, H121A, R1177A, 또는 H1182A, 또는 Cas13b 오솔로그의 이에 상응하는 아미노산 위치의 돌연변이 중 하나 이상인 변형 방법.

청구항 28

제17항 내지 제27항 중 어느 한 항에 있어서, 상기 가이드 서열은 상기 표적 서열과 상기 RNA 듀플렉스를 형성할 수 있는 약 20 내지 약 53 nt, 약 25 내지 약 53 nt, 또는 약 29 내지 약 53 nt의 길이를 갖는 것인 변형 방법.

청구항 29

제28항에 있어서, 상기 가이드 서열은 상기 표적 서열과 상기 RNA 듀플렉스를 형성할 수 있는 약 40 내지 약 50 nt의 길이를 갖는 것인 변형 방법.

청구항 30

제17항에 있어서, 상기 비-쌍형성 C 및 상기 가이드 서열의 5' 말단 간 거리는 20-30개 뉴클레오티드인 변형 방법.

청구항 31

제17항 내지 제30항 중 어느 한 항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 인간, 두족류, 또는 초파리 아데노신 데아미나제 단백질 또는 이의 촉매 도메인인 변형 방법.

청구항 32

제17항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 hADAR2-D 아미노산 서열의 글루탐산^{488}, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함하도록 변형된 것인 변형 방법.

청구항 33

제32항에 있어서, 위치 488 또는 상동성 ADAR 단백질 내 상응하는 위치의 상기 글루탐산 잔기는 글루타민 잔기(E488Q)로 치환되는 것인 변형 방법.

청구항 34

제32항 또는 제33항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 돌연변이 E488Q를 포함하는 돌연변이된 hADAR2d 또는 돌연변이 E1008Q를 포함하는 돌연변이된 hADAR1d인 변형 방법.

청구항 35

제17항 내지 제34항 중 어느 한 항에 있어서, 가이드 서열은 표적 RNA 서열 내 상이한 아데노신 부위에 상응하는 하나 초과의 미스매치를 포함하거나 또는 2종의 가이드 분자가 사용되고, 그 각각은 표적 RNA 서열 내 상이한 아데노신 부위에 상응하는 미스매치를 포함하는 것인 변형 방법.

청구항 36

제17항 내지 제35항 중 어느 한 항에 있어서, 상기 Cas13 단백질 및 임의로상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 하나 이상의 이종성 핵 국재화 신호(들) (NLS(들))을 포함하는 것인 변형 방법.

청구항 37

제17항 내지 제36항 중 어느 한 항에 있어서, 상기 방법은 상기 관심 표적 서열을 결정하는 단계 및 상기 표적 서열에 존재하는 상기 아데닌을 가장 효율적으로 탈아미노화시키는 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 선택하는 단계를 포함하는 것인 변형 방법.

청구항 38

제17항 내지 제37항 중 어느 한 항에 있어서, 상기 촉매적 불활성 Cas13 단백질은 임의의 표 1, 2, 3, 또는 4에 열거된 박테리아 종으로 이루어진 군으로부터 선택되는 박테리아 종으로부터 유래된 Cas13 뉴클레아제로부터 수득되는 것인 변형 방법.

청구항 39

제17항 내지 제38항 중 어느 한 항에 있어서, 상기 관심 표적 RNA 서열은 세포 내에 존재하는 것인 변형 방법.

청구항 40

제39항에 있어서, 상기 세포는 진핵생물 세포인 변형 방법.

청구항 41

제40항에 있어서, 상기 세포는 인간이외의 동물 세포인 변형 방법.

청구항 42

제41항에 있어서, 상기 세포는 인간 세포인 변형 방법.

청구항 43

제42항에 있어서, 상기 세포는 식물 세포인 변형 방법.

청구항 44

제17항 내지 제43항 중 어느 한 항에 있어서, 상기 관심 표적 유전자좌는 동물 내에 존재하는 것인 변형 방법.

청구항 45

제17항 내지 제44항 중 어느 한 항에 있어서, 상기 관심 표적 유전자좌는 식물 내에 존재하는 것인 변형 방법.

청구항 46

제17항 내지 제45항 중 어느 한 항에 있어서, 상기 관심 표적 RNA 서열은 시험관내 RNA 폴리뉴클레오티드에 포함되는 것인 변형 방법.

청구항 47

제17항 내지 제46항 중 어느 한 항에 있어서, 상기 성분 (a), (b) 및 (c)는 상기 세포에 리보핵단백질 복합체로서 전달되는 것인 변형 방법.

청구항 48

제17항 내지 제47항 중 어느 한 항에 있어서, 상기 성분 (a), (b) 및 (c)는 상기 세포에 하나 이상의 폴리뉴클레오티드 분자로서 전달되는 것인 변형 방법.

청구항 49

제48항에 있어서, 상기 하나 이상의 폴리뉴클레오티드 분자는 성분 (a) 및/또는 (c)를 코딩하는 하나 이상의 mRNA 분자를 포함하는 것인 변형 방법.

청구항 50

제49항에 있어서, 상기 하나 이상의 폴리뉴클레오티드 분자는 하나 이상의 벡터 내에 포함되는 것인 변형 방법.

청구항 51

제50항에 있어서, 상기 하나 이상의 폴리뉴클레오티드 분자는 상기 Cas13 단백질, 상기 가이드 분자, 및 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 발현하도록 작동적으로 구성된 하나 이상의 조절 구성요

소를 포함하고, 임의로 상기 하나 이상의 조절 구성요소는 유도성 프로모터를 포함하는 것인 변형 방법.

청구항 52

제48항에 있어서, 상기 하나 이상의 폴리뉴클레오티드 분자 또는 상기 리보핵단백질 복합체는 입자, 소포체, 또는 하나 이상의 바이러스 벡터를 통해서 전달되는 것인 변형 방법.

청구항 53

제52항에 있어서, 상기 입자는 지질, 당, 금속 또는 단백질을 포함하는 것인 변형 방법.

청구항 54

제53항에 있어서, 상기 입자는 지질 나노입자를 포함하는 것인 변형 방법.

청구항 55

제52항에 있어서, 상기 소포체는 엑소솜 또는 리포솜을 포함하는 것인 변형 방법.

청구항 56

제50항에 있어서, 상기 하나 이상의 바이러스 벡터는 하나 이상의 아데노바이러스, 하나 이상의 렌티바이러스 또는 하나 이상의 아데노-연합 바이러스를 포함하는 것인 변형 방법.

청구항 57

제17항 내지 제56항 중 어느 한 항에 있어서, 상기 방법은 하나 이상의 표적 RNA 서열의 조작에 의해 세포, 세포주 또는 유기체를 변형시키는 것인 변형 방법.

청구항 58

제57항에 있어서, 상기 관심 표적 RNA 내 상기 아데닌의 상기 탈아미노화는 병원성 G→A 또는 C→T 점 돌연변이를 함유하는 전사물에 의해 초래되는 질환을 치유하는 것인 변형 방법.

청구항 59

제58항에 있어서, 상기 질환은 메이어-골린 증후군, 섹켈 증후군 4, 주버트 증후군 5, 레버 선천성 흑암시 10, 샤르코-마리-투스 질환, 2형; 샤르코-마리-투스 질환, 2형; 어셔 증후군, 2C형; 척수소뇌성 실조증 28; 척수소뇌성 실조증 28; 척수소뇌성 실조증 28; 긴 QT 증후군 2; 쇠그렌-라슨 증후군; 유전성 프룩토스뇨증; 유전성 프룩토스뇨증; 신경아세포종; 신경아세포종; 칼만 증후군 1; 칼만 증후군 1; 칼만 증후군 1; 이염색성 백질이영양증, 레트 증후군, 근위축성 측삭 경화증 10형, 리-프라우메니 증후군, 또는 표 5에 열거된 질환으로부터 선택되는 것인 변형 방법.

청구항 60

제58항에 있어서, 상기 질환은 미성숙 종결 질환인 변형 방법.

청구항 61

제57항에 있어서, 상기 변형은 유기체의 생식력에 영향을 미치는 것인 변형 방법.

청구항 62

제57항에 있어서, 상기 변형은 상기 표적 RNA 서열의 스플라이싱에 영향을 미치는 것인 변형 방법.

청구항 63

제57항에 있어서, 상기 변형은 아미노산 변화를 도입시키고 암 세포에서 새로운 항원의 발현을 초래하는 전사물 내 돌연변이를 도입시키는 것인 변형 방법.

청구항 64

제57항에 있어서, 상기 표적 RNA는 마이크로RNA 내에 포함되는 것인 변형 방법.

청구항 65

제57항에 있어서, 상기 관심 표적 RNA 내 상기 아데닌의 상기 탈아미노화는 유전자의 기능 획득 또는 기능 상실을 초래하는 것인 변형 방법.

청구항 66

제65항에 있어서, 상기 유전자는 암 세포에 의해 발현되는 유전자인 변형 방법.

청구항 67

제17항 내지 제66항 중 어느 한 항의 방법으로 수득된 변형된 세포, 또는 상기 변형된 세포의 자손으로서, 상기 세포는 방법이 수행되지 않은 대응 세포와 비교하여 상기 관심 표적 RNA 내 상기 아데닌 대신에 하이포잔틴 또는 구아닌을 포함하는 것인 변형된 세포 또는 이의 자손.

청구항 68

제67항에 있어서, 상기 세포는 진핵생물 세포인 변형된 세포 또는 이의 자손.

청구항 69

제67항에 있어서, 상기 세포는 동물 세포인 변형된 세포 또는 이의 자손.

청구항 70

제67항에 있어서, 상기 세포는 인간 세포인 변형된 세포 또는 이의 자손.

청구항 71

제67항에 있어서, 상기 세포는 치료적 T 세포인 변형된 세포 또는 이의 자손.

청구항 72

제67항에 있어서, 상기 세포는 항체-생산 B 세포인 변형된 세포 또는 이의 자손.

청구항 73

제67항에 있어서, 상기 세포는 식물 세포인 변형된 세포 또는 이의 자손.

청구항 74

제67항의 상기 변형된 세포를 포함하는 인간이외의 동물.

청구항 75

제74항의 상기 변형된 세포를 포함하는 식물.

청구항 76

세포 요법을 위한 방법으로서, 이러한 요법을 필요로 하는 환자에게 제67항 내지 제75항 중 어느 한 항의 상기 변형된 세포를 투여하는 단계를 포함하고, 상기 변형된 세포의 존재는 환자에서 질환을 치유하는 것인 방법.

청구항 77

관심 표적 유전자좌 내 아데닌을 변형시키기에 적합한 조작된, 비천연 발생 시스템으로서,

a) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자, 또는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열;

b) 촉매적 불활성 Cas13 단백질, 또는 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열;

c) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인, 또는 상기 아데노신 데아미나제 단백질 또는 이의 촉매

도메인을 코딩하는 뉴클레오티드 서열

을 포함하고,

상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후 그에 연결되도록 적합화되고;

상기 가이드 서열은 아데닌을 포함하는 표적 RNA 서열과 하이브리드화하여 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여, 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시키는 것인, 조작된, 비천연 발생 시스템.

청구항 78

제77항의 a), b) 및 c)의 뉴클레오티드 서열을 포함하는, 관심 표적 유전자좌 내 아데닌을 변형시키기에 적합한, 조작된, 비천연 발생 벡터 시스템.

청구항 79

제78항에 있어서,

a) 상기 가이드 서열을 포함하는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열에 작동적으로 연결된 제1 조절 구성요소,

b) 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열에 작동적으로 연결된 제2 조절 구성요소; 및

c) 상기 제1 또는 제2 조절 구성요소의 제어 하에 있거나 또는 제3 조절 구성요소에 작동적으로 연결된 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열

을 포함하는 하나 이상의 벡터를 포함하고,

아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 상기 뉴클레오티드 서열이 제3 조절 구성요소에 작동적으로 연결되면, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 발현 후에 상기 가이드 분자 또는 상기 Cas13 단백질에 연결되도록 적합화되고,

성분 (a), (b) 및 (c)는 시스템의 동일하거나 또는 상이한 벡터에 위치되는 것인, 조작된, 비천연 발생 벡터 시스템.

청구항 80

제77항 내지 제79항 중 어느 한 항의 시스템을 포함하는 시험관내 또는 생체외 숙주 세포 또는 이의 자손 또는 세포주 또는 이의 자손.

청구항 81

제80항에 있어서, 상기 세포는 진핵생물 세포인, 숙주 세포 또는 이의 자손 또는 세포주 또는 이의 자손.

청구항 82

제80항에 있어서, 상기 세포는 동물 세포인, 숙주 세포 또는 이의 자손 또는 세포주 또는 이의 자손.

청구항 83

제80항에 있어서, 상기 세포는 인간 세포인, 숙주 세포 또는 이의 자손 또는 세포주 또는 이의 자손.

청구항 84

제80항에 있어서, 상기 세포는 식물 세포인, 숙주 세포 또는 이의 자손 또는 세포주 또는 이의 자손.

청구항 85

제117항에 있어서, 상기 시토신은 구아노신이 5'에 측접되지 않는 것인 변형 방법.

청구항 86

제17항에 있어서, 상기 아데노신 데아미나제는 ADAR, 임의로는 huADAR, 임의로는 (hu)ADAR1 또는 (hu)ADAR2인 변형 방법.

청구항 87

제17항에 있어서, 상기 Cas13, 바람직하게 Cas13b는 절단되고, 바람직하게 C-말단에서 절단되며, 바람직하게 상기 Cas 13은 상응하는 야생형 Cas13의 절단된 기능성 변이체인 변형 방법.

청구항 88

제17항에 있어서, 상기 아데노신 데아미나제는 RNA 특이적 아데노신 데아미나제인 변형 방법.

청구항 89

제17항 또는 제32항에 있어서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 ADAR의 하나 이상의 돌연변이, 바람직하게 본 명세서에 기술된 바와 같은 돌연변이, 예를 들어 임의의 도 43A-43D, 44, 45, 46A-46B, 47A-47B에 제공되는 바와 같은 돌연변이, 또는 ADAR 상동체 또는 오솔로그 내 상응하는 돌연변이를 포함하도록 변형된 것인 변형 방법.

청구항 90

관심 표적 RNA 서열 내 아데닌을 변형시키는 방법으로서, 상기 표적 RNA에,

(a) 촉매적 불활성 (데드) Cas13 단백질;

(b) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자; 및

(c) 활성이 시티딘 데아미나제로 전환되도록 돌연변이된 아데노신 데아미나제 단백질 또는 이의 촉매 도메인

을 전달하는 단계를 포함하고,

상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후 그에 연결되도록 적합화되고;

상기 가이드 분자는 상기 데드 Cas13 단백질과 복합체를 형성하고 상기 복합체가 상기 관심 표적 RNA 서열에 결합하도록 유도시키고, 상기 가이드 서열은 상기 아데닌을 포함하는 표적 서열과 하이브리드화하여 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여, 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시키고;

상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 RNA 듀플렉스 내 상기 아데닌을 탈아미노화시키는 것인 변형 방법.

청구항 91

제90항에 있어서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376, S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, V525 및 I520으로부터 선택된 하나 이상의 위치에서 돌연변이되는 것인 변형 방법.

청구항 92

제91항에 있어서, 아데노신 데아미나제는 E488, V351, S486, T375, S370, P462, 및 N597로부터 선택되는 하나 이상의 위치에서 돌연변이되는 것인 변형 방법.

발명의 설명

기 술 분 야

[0001] **관련 출원에 관한 교차 참조**

[0002] 본 출원은 2017년 10월 4일 출원된 미국 가출원 제62/568,313호의 우선권을 청구한다. 상기 식별된 출원의

전체 내용은 참조로 본 명세서에 완전히 편입된다.

[0003] **주정부 지원 연구에 관한 진술**

[0004] 본 발명은 미국 국립 보건원이 수여하는 승인 번호 MH100706 및 MH110049 하의 정부 지원으로 수행되었다. 정부는 본 발명에 대해 일정 권리를 갖는다.

[0005] **전자 서열 목록에 관한 참조**

[0006] 전자 서열 목록의 내용 ("BROD-2345WP_ST25_FOR_FILING.txt"; 크기는 1,427,232 바이트이고 2018년 9월 25일 생성되었음)은 그 전문이 참조로 본 명세서에 편입된다.

[0007] **발명의 분야**

[0008] 본 발명은 일반적으로 핵산을 표적화하고 편집하기 위한, 특히 관심 표적 유전자좌에서 아데닌의 프로그램가능한 탈아미노화를 위한, 시스템, 방법 및 조성물에 관한 것이다.

배 경 기 술

[0009] 게놈 시퀀싱 기술 및 분석 방법에서 최근의 진보는 다양한 범위의 생물학적 기능 및 질환과 연관된 유전 인자를 분류하고 맵핑하는 능력을 상당히 가속화시켰다. 정밀한 게놈 표적화 기술은 합성 생물학, 생명공학 및 의학적 적용 분야를 진보시키기 위해서 뿐만 아니라, 개별 유전자 구성요소의 선택적 교란을 허용하여 인과적인 유전자 변이의 체계적인 역조작을 가능하게 하기 위해 필요하다. 게놈-편집 기술 예컨대 디자이너 징크 핑거, 전사 활성인자-유사 이펙터 (TALE; transcription activator-like effector), 또는 귀소성 메가뉴클레아제가 표적화된 게놈 교란을 일으키기 위해 이용가능하지만, 새로운 전략 및 분자 기전을 적용하고, 감당할만하고, 셋업이 용이하고, 규모확장가능하며, 진핵생물 게놈 내 복수개 위치의 표적화를 처리할 수 있는 신규한 게놈 조작 기술에 대한 필요성이 여전히 남아있다. 이것은 게놈 조작 및 생명공학의 새로운 적용분야에 주요한 자원을 제공하게 될 것이다.

[0010] 시토신의 프로그램가능한 탈아미노화가 보고되었고 A→G 및 T→C 점 돌연변이의 교정에 사용될 수 있다. 예를 들어, [Komor et al., Nature (2016) 533:420-424]는 Cas9-가이드 RNA 복합체와 표적화된 DNA 가닥의 결합에 의해서 대체된 비표적화 DNA 가닥에서 APOBEC1 시티딘 데아미나제에 의한 시토신의 표적화 디아민화가, 시토신의 우라실로의 전환을 야기시킨다는 것을 보고한다. 또한, 문헌 [Kim et al., Nature Biotechnology (2017) 35:371-376]; [Shimatani et al., Nature Biotechnology (2017) doi:10.1038/nbt.3833]; [Zong et al., Nature Biotechnology (2017) doi:10.1038/nbt.3811]; [Yang Nature Communication (2016) doi:10.1038/ncomms13330]을 참조한다.

발명의 내용

[0011] 본 출원은 관심 표적 RNA 서열을 변형시키는 것에 관한 것이다. DNA 표적화 대신에 RNA-표적화의 사용은 치료제 개발에 적절한 몇몇 장점을 제공한다. 첫번째로, RNA 표적화는 실질적으로 안전성의 이득이 존재한다: 전사체 내 이용가능한 서열 공간이 게놈에 비해서 상당히 적기 때문에 오프-표적 사건이 더 적을 것이고, 오프-표적 사건이 일어나면, 그것은 일시적일 것이고 아마도 부정적인 부작용을 유도할 가능성이 더 적을 것이다. 두번째로, RNA-표적화 치료제는 그들이 세포-유형 독립적이고 핵으로 진입하지 않아서, 그들의 전달을 보다 용이하게 하기 때문에 더 효율적일 것이다.

[0012] 일 양상에서, 본 개시는 표적화 도메인; 및 아데노신 데아미나제 또는 이의 촉매 도메인을 포함하는 부위 지정 염기 편집을 위한 조작된 조성물을 포함하고, 여기서 아데노신 데아미나제는 활성을 시티딘 데아미나제로 전환시키도록 변형된다.

[0013] 일부 실시형태에서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376, S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, V525 및 I520으로부터 선택되는 하나 이상의 위치에서 하나 이상의 돌연변이에 의해 변형된다. 일부 실시형태에서, 아데노신 데아미나제는 E488, V351, S486, T375, S370, P462 및 N597로부터 선택되는 하나 이상의 위치에서 돌연변이된다. 일부 실시형태에서, 아데노신 데아미나제는 E488Q, V351G, S486A, T375S, S370C, P462A 및 N597I로부터 선택되는 하나 이상의 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 인간, 두족류 (cephalopod), 또는 초파리 (Drosophila) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인이다. 일부 실시형태에서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 hADAR2-D 아미노

산 서열의 글루탐산488, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함하도록 변형되었다. 일부 실시형태에서, 위치 488 또는 상동성 ADAR 단백질 내 상응하는 위치의 상기 글루탐산 잔기는 글루타민 잔기로 치환된다 (E488Q). 일부 실시형태에서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 돌연변이 E488Q를 포함하는 돌연변이된 hADAR2d 또는 돌연변이 E1008Q를 포함하는 돌연변이된 hADAR1d이다. 일부 실시형태에서, 표적화 도메인은 촉매적 불활성 Cas13 단백질이거나, 또는 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열이다. 일부 실시형태에서, 촉매적 불활성 Cas13 단백질은 촉매적 불활성 Cas13a, 촉매적 불활성 Cas13b, 또는 촉매적 불활성 Cas13c이다. 일부 실시형태에서, 상기 촉매적 불활성 Cas13 단백질은 임의의 표 1, 2, 3, 또는 4에 열거된 박테리아 종으로 이루어진 군으로부터 선택되는 박테리아 종으로부터 유래되는 Cas13 뉴클레아제로부터 수득된다. 일부 실시형태에서, 조성물은 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자, 또는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열을 더 포함한다. 일부 실시형태에서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 표적화 도메인에 공유적으로 또는 비공유적으로 연결된다.

[0014] 다른 양상에서, 본 발명은 관심 표적 RNA 서열 내 아데닌을 변형시키는 방법에 관한 것이다. 특정한 실시형태에서, 방법은 상기 표적 RNA에 (a) 촉매적 불활성 (데드) Cas13 단백질; (b) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자; 및 (c) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 전달하는 단계를 포함하고, 여기서 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후에 그에 연결되도록 적합화되고; 가이드 분자는 상기 데드 Cas13 단백질과 복합체를 형성하고 상기 복합체는 상기 관심 표적 RNA 서열와 결합되도록 유도되며, 상기 가이드 서열은 상기 아데닌을 포함하는 표적 서열과 하이브리드화되어 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여서 형성된 RNA 듀플렉스에 A-C 미스매치가 생성되며, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 RNA 듀플렉스 내 상기 아데닌을 탈아미노화시킨다.

[0015] 일정 예의 실시형태에서, Cas13 단백질은 Cas13a, Cas13b 또는 Cas 13c이다.

[0016] 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질의 N-말단 또는 C-말단에 융합된다. 일정 예의 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 링커에 의해서 상기 데드 Cas13 단백질에 융합된다. 링커는 $(GGGGS)_{3-11}$ (SEQ ID Nos. 1-9) GSG_5 (SEQ ID No. 10) 또는 LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID No. 11)일 수 있다.

[0017] 일정 예의 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 어댑터 단백질 및 상기 가이드 분자에 연결되거나 또는 상기 데드 Cas13 단백질은 상기 어댑터 단백질에 결합할 수 있는 압타머 서열을 포함한다. 어댑터 서열은 MS2, PP7, Qβ, F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ϕCb5, ϕCb8r, ϕCb12r, ϕCb23r, 7s 및 PRR1로부터 선택될 수 있다.

[0018] 일정 예의 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질의 내부 루프에 삽입된다. 일정 예의 실시형태에서, Cas13a 단백질은 2개 HEPN 도메인 내에, 특히 렙토트리키아 웨이데이 (*Leptotrichia wadei*)로부터 기원하는 Cas13a 단백질의 위치 R474 및 R1046 또는 Cas13a 오솔로그의 이의 상응하는 아미노산 위치에 하나 이상의 돌연변이를 포함한다.

[0019] 일정 예의 실시형태에서, Cas 13 단백질은 Cas13b 단백질이고, Cas13b는 베르게이엘라 주헬컴 (*Bergeyella zoohelcum*) ATCC 43767로부터 기원하는 Cas13b 단백질의 위치 R116, H121, R1177, H1182 중 하나 이상 또는 Cas13b 오솔로그의 이의 상응하는 아미노산 위치에 돌연변이를 포함한다. 일정한 다른 예의 실시형태에서, 돌연변이는 베르게이엘라 주헬컴 ATCC 43767로부터 기원하는 Cas13b 단백질의 R116A, H121A, R1177A, H1182A 중 하나 이상이거나 또는 Cas13b 오솔로그의 이의 상응하는 아미노산 위치에 있다.

[0020] 일정 예의 실시형태에서, 가이드 서열은 상기 표적 서열과 상기 RNA 듀플렉스를 형성할 수 있는 약 29-53 nt의 길이를 갖는다. 일정한 다른 예의 실시형태에서, 가이드 서열은 상기 표적 서열과 상기 RNA 듀플렉스를 형성할 수 있는 약 40-50 nt의 길이를 갖는다. 일정 예의 실시형태에서, 상기 가이드 서열의 5' 말단과 상기 비-쌍형성 C 사이의 거리는 20-30개 뉴클레오티드이다.

[0021] 일정 예의 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 인간, 두족류, 또는 초파리 아데노신 데아미나제 단백질 또는 이의 촉매 도메인이다. 일정 예의 실시형태에서, 아데노신 데아미나제 단백

질 또는 이의 촉매 도메인은 hADAR2-D 아미노산 서열의 글루탐산[488], 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함하도록 변형되었다. 일정 예의 실시형태에서, 글루탐산 잔기는 위치 488에 존재할 수 있거나 또는 상동성 ADAR 단백질 내 상응하는 위치는 글루타민 잔기로 치환된다 (E488Q).

[0022] 일정한 다른 예의 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 돌연변이 E488Q를 포함하는 돌연변이된 hADAR2d이거나 또는 돌연변이 E1008Q를 포함하는 돌연변이된 hADAR1d이다.

[0023] 일정 예의 실시형태에서, 가이드 서열은 표적 RNA 서열 내에 상이한 아데노신 부위에 상응하는 하나 초과의 미스매치를 포함하거나 또는 각각이 표적 RNA 서열 내에 상이한 아데노신 부위에 상응하는 미스매치를 포함하는 것인 2개 가이드 분자가 사용된다.

[0024] 일정 예의 실시형태에서, Cas13 단백질 및 임의로는 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 하나 이상의 이종성 핵 국재화 신호(들) (NLS(들))를 포함한다.

[0025] 일정 예의 실시형태에서, 방법은 관심 표적 서열을 결정하는 단계 및 이어서 표적 서열에 존재하는 상기 아데닌을 가장 효율적으로 탈아미노화시키는 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 선택하는 단계를 더 포함한다.

[0026] 관심 표적 RNA 서열은 세포 내에 존재할 수 있다. 세포는 진핵생물 세포, 인간이외 동물 세포, 인간 세포, 식물 세포일 수 있다. 관심 표적 유전자좌는 동물 또는 식물 내에 존재할 수 있다.

[0027] 관심 표적 RNA 서열은 시험관 내 RNA 폴리뉴클레오티드에 포함될 수 있다.

[0028] 본 명세서에 기술된 시스템의 성분은 상기 세포에 리보핵단백질 복합체로서 또는 하나 이상의 폴리뉴클레오티드 분자로서 전달될 수 있다. 하나 이상의 폴리뉴클레오티드 분자는 성분을 코딩하는 하나 이상의 mRNA 분자를 포함할 수 있다. 하나 이상의 폴리뉴클레오티드 분자는 하나 이상의 벡터 내에 포함될 수 있다. 하나 이상의 폴리뉴클레오티드 분자는 상기 Cas13 단백질, 상기 가이드 분자, 및 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 발현하도록 작동적으로 구성된 하나 이상의 조절 구성요소를 더 포함할 수 있고, 임의로는 상기 하나 이상의 조절 구성요소는 유도성 프로모터를 포함한다. 하나 이상의 폴리뉴클레오티드 분자 또는 상기 리보핵단백질 복합체는 입자, 소포체, 또는 하나 이상의 바이러스 벡터를 통해서 전달될 수 있다. 입자는 지질, 당, 금속 또는 단백질을 포함할 수 있다. 입자는 지질 나노입자를 포함할 수 있다. 소포체는 엑소솜 또는 리포솜을 포함할 수 있다. 하나 이상의 바이러스 벡터는 하나 이상의 아데노바이러스, 하나 이상의 렌티바이러스 또는 하나 이상의 아데노-연합 바이러스를 포함할 수 있다.

[0029] 본 명세서에 개시된 방법은 하나 이상의 표적 RNA 서열의 조작에 의해서 세포, 세포주 또는 유기체를 변형시키는데 사용될 수 있다.

[0030] 일정 예의 실시형태에서, 상기 관심 표적 RNA 내 상기 아데닌의 탈아미노화는 병원성 G→A 또는 C→T 점 돌연변이를 함유하는 전사물에 의해 초래되는 질환을 치유한다.

[0031] 방법은 질환을 치료하는데 사용될 수 있다. 일정 예의 실시형태에서, 질환은 메이어-골린 증후군, 섹켈 증후군 4, 주버트 증후군 5, 레버 선천성 흑암시 10, 샤르코-마리-투스 질환, 2형; 샤르코-마리-투스 질환, 2형; 어셔 증후군, 2C형; 척수소뇌성 실조증 28; 척수소뇌성 실조증 28; 척수소뇌성 실조증 28; 긴 QT 증후군 2; 쇠그렌-라슨 증후군; 유전성 프룩토스뇨증; 유전성 프룩토스뇨증; 신경아세포종; 신경아세포종; 칼만 증후군 1; 칼만 증후군 1; 칼만 증후군 1; 이염색성 백질이영양증, 레트 증후군, 근위축성 측삭 경화증 10형, 리-프라우메니 증후군, 또는 표 5에 열거된 질환으로부터 선택된다. 질환은 미성숙 종결 질환일 수 있다.

[0032] 본 명세서에 개시된 방법은 유기체의 생식력에 영향을 미치는 변형을 만드는데 사용될 수 있다. 변형은 상기 표적 RNA 서열의 스플라이싱에 영향을 미칠 수 있다. 변형은 아미노산 변화를 도입시켜서 암 세포에서 새로운 항원의 발현을 야기시키는 전사물 내 돌연변이를 도입시킬 수 있다.

[0033] 일정 예의 실시형태에서, 표적 RNA는 마이크로RNA일 수 있거나 또는 마이크로RNA 내에 포함될 수 있다. 일정 예의 실시형태에서, 상기 관심 표적 RNA 내 상기 아데닌의 탈아미노화는 유전자의 기능 획득 또는 기능 상실을 초래한다. 일정 예의 실시형태에서, 유전자는 암 세포에 의해 발현되는 유전자이다.

[0034] 일정 예의 실시형태에서, 상기 시토신은 구아노신이 5'에 측접되지 않는다. 일정 예의 실시형태에서, 상기 아데노신 데아미나제는 ADAR, 임의로는 huADAR, 임의로는 (hu)ADAR1 또는 (hu)ADAR2이다. 일정 예의 실시형태에서, 상기 Cas13, 바람직하게 Cas13b는 절단되고, 바람직하게는 C-말단에서 절단되며, 바람직하게 상기

Cas 13은 상응하는 야생형 Cas13의 절단된 기능성 변이체이다. 일정 예의 실시형태에서, 상기 아데노신 데아미나제는 RNA 특이적 아데노신 데아미나제이다. 일정 예의 실시형태에서, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 ADAR의 하나 이상의 돌연변이, 바람직하게는 본 명세서에 기술된 바와 같은 돌연변이, 예를 들어 임의의 도 43A-43D, 44, 45, 46A-46B, 47A-47B,에 제공된 바와 같은 돌연변이 또는 ADAR 상동체 또는 오솔로그 내 상응하는 돌연변이를 포함하도록 변형되었다.

[0035] 다른 양상에서, 본 발명은 본 명세서에 개시된 방법을 사용하여 수득되는 변형된 세포 또는 이의 자손을 포함하고, 상기 세포는 방법이 수행되지 않은 대응 세포와 비교하여 상기 관심 표적 RNA 내 상기 아데닌 대신에 하이포잔틴 또는 구아닌을 포함한다. 변형된 세포 또는 이의 자손은 진핵생물 세포, 동물 세포, 인간 세포, 치료적 T 세포, 항체-생산 B 세포, 식물 세포일 수 있다.

[0036] 다른 양상에서, 본 발명은 상기 변형된 세포 또는 이의 자손을 포함하는 인간이외의 동물을 포함한다. 변형된 세포는 식물 세포일 수 있다.

[0037] 다른 양상에서, 본 발명은 치료를 필요로 하는 환자에게 본 명세서에 개시된 변형된 세포를 투여하는 단계를 포함하는, 세포 요법을 위한 방법을 포함하고, 여기서 상기 변형된 세포의 존재는 환자에서 질환을 치유한다.

[0038] 다른 양상에서, 본 발명은 A) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자, 또는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열; B) 촉매적 불활성 Cas13 단백질, 또는 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열; C) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인, 또는 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열을 포함하는, 관심 표적 유전자좌 내 아데닌을 변형시키는데 적합한 조작된, 비천연 발생 시스템에 관한 것이고; 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 이후에 이에 연결되도록 적합화되고; 상기 가이드 서열은 아데닌을 포함하는 표적 RNA 서열과 하이브리드화되어 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여서 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시킨다.

[0039] 다른 양상에서, 본 발명은 a), b) 및 c)의 뉴클레오티드 서열을 포함하는, 관심 표적 유전자좌 내 아데닌을 변형시키는데 적합한 조작된, 비천연 발생 벡터 시스템에 관한 것이다.

[0040] 다른 양상에서, 본 발명은 상기 가이드 서열을 포함하는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열에 작동적으로 연결된 제1 조절 구성요소, 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열에 작동적으로 연결된 제2 조절 구성요소; 및 상기 제1 또는 제2 조절 구성요소의 제어 하에 있거나 또는 제3 조절 구성요소에 작동적으로 연결된 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열을 포함하는 하나 이상의 벡터를 포함하는, 조작된, 비천연 발생 벡터 시스템에 관한 것이고, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 상기 뉴클레오티드 서열이 제3 조절 구성요소에 작동적으로 연결되면, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 발현 후에 상기 가이드 분자 또는 상기 Cas13 단백질에 연결되도록 적합화되고, 성분 A), B) 및 C)는 시스템의 동일하거나 또는 상이한 벡터에 위치된다.

[0041] 본 명세서에 개시된 방법이 절단 RNA의 결합 및 특이성을 위해 포유동물 세포에서 기능하는 Cas13 단백질의 능력을 입증하므로, 추가의 확대된 적용분야는 스플라이스 변이체의 편집, 및 RNA-결합 단백질이 RNA와 상호작용하는 방법의 측정을 포함한다.

[0042] 다른 양상에서, 본 발명은 본 명세서에 개시된 시스템을 포함하는 시험관내 또는 생체외 숙주 세포 또는 이의 자손 또는 세포주 또는 이의 자손에 관한 것이다. 숙주 세포 또는 이의 자손은 진핵생물 세포, 동물 세포, 인간 세포, 또는 식물 세포일 수 있다.

[0043] 다른 양상에서, 본 개시는 상기 표적 RNA에, (a) 촉매적 불활성 (데드) Cas13 단백질; (b) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자; 및 (c) 활성이 시티딘 데아미나제로 전환되도록 돌연변이된 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 전달하는 단계를 포함하는, 관심 표적 RNA 서열 내 아데닌을 변형시키는 방법을 포함하고, 여기서 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후에 그에 연결되도록 적합화되고, 상기 가이드 분자는 상기 데드 Cas13 단백질과 복합체를 형성하고 상기 복합체가 상기 관심 표적 RNA 서열에 결합되도록 유도시키며, 상기 가이드 서열은 상기 아데닌을 포함하는 표적 서열과 하이브리드화되어 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신

을 포함하여서 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시키며, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 RNA 듀플렉스 내 상기 아데닌을 탈아미노화시킨다.

[0044] 일부 예의 실시형태에서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376, S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, V525 및 I520으로부터 선택되는 하나 이상의 위치에서 돌연변이된다. 일부 예의 실시형태에서, 아데노신 데아미나제는 E488, V351, S486, T375, S370, P462 및 N597로부터 선택되는 하나 이상의 위치에서 돌연변이된다.

[0045] 다른 양상에서, 본 개시는 (a) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자, 또는 상기 가이드 분자를 코딩하는 뉴클레오티드 서열; (b) 촉매적 불활성 Cas13 단백질, 또는 상기 촉매적 불활성 Cas13 단백질을 코딩하는 뉴클레오티드 서열; (c) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인, 또는 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열을 포함하는, 관심 표적 유전자좌 내 아데닌을 변형시키는데 적합한 조작된, 비천연 발생 시스템을 포함하고, 여기서 아데노신 데아미나제는 활성이 시티딘 데아미나제로 전환되도록 변형되고, 상기 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후에 그에 연결되도록 적합화되고, 상기 가이드 서열은 아데닌을 포함하는 표적 RNA 서열과 하이브리드화되어 RNA 듀플렉스를 형성할 수 있고, 상기 가이드 서열은 상기 아데닌에 상응하는 위치에 비-쌍형성 시토신을 포함하여서 형성된 RNA 듀플렉스 내에 A-C 미스매치를 생성시킨다.

[0046] 일부 예의 실시형태에서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376, S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, V525 및 I520으로부터 선택되는 하나 이상의 돌연변이에 의해 변형된다. 일부 예의 실시형태에서, 아데노신 데아미나제는 E488, V351, S486, T375, S370, P462, 및 N597로부터 선택되는 하나 이상의 위치에서 돌연변이된다.

[0047] 예로의 실시형태의 이들 및 다른 양상, 목적, 특성 및 장점은 예시된 예로서이 실시형태의 하기 상세한 설명을 고려할때 당업자에게 자명하게 될 것이다.

도면의 간단한 설명

[0048] 본 발명의 특성 및 장점의 이해는 본 발명의 원리를 이용할 수 있는, 예시적인 실시형태를 기재한 하기 상세한 설명, 및 첨부된 도면을 참조하여 얻게 될 것이다:

도 1은 Cas13b 단백질과 함께, 본 명세서에서 예시된 관심 표적 RNA 서열에서 아데닌의 표적화된 탈아미노화를 위한 본 발명의 예시적 실시형태를 예시한다.

도 2는 Cas13b를 사용할 때와 같이 전사물 수준에서 인간 질환을 치료하기 위한 치료적 전략으로서 RNA 편집의 개발을 예시한다. 루시퍼라제 전사물 내 조작된 사전-종결 정지 코돈을 표적화하는 Cas13-ADAR2 융합체에 의한 RNA 염기 편집의 개략도.

도 3: 루시퍼라제 발현을 복구시키기 위한 가이드 위치 및 길이 최적화.

도 4: 아데닌 데아미나제 단백질의 예시적인 서열을 도시한 도면 (SEQ ID No. 650-656).

도 5: 예시적인 실시형태에서 사용된 가이드 (뉴클레오티드 서열의 경우 SEQ ID No. 657-660, 그리고 아미노산 서열의 경우 SEQ ID No. 696에 표시됨).

도 6: 편집 효율은 가이드 길이를 연장하여 완성된, 긴 RNA 듀플렉스를 가지고 DR로부터 더 멀리 떨어져 있는 편집된 염기와 상관성이 있다.

도 7: 편집 효율이 클수록 편집 부위는 DR/단백질 결합 영역으로부터 더욱 더 멀어진다.

도 8: DR로부터의 편집된 부위의 거리.

도 9A 및 도 9B: N 말단 또는 C-말단 상에서 Cas13b12(이중 R HEPN 돌연변이체)에 융합된 ADAR1 또는 ADAR2. 가이드는 루시퍼라제 내 정지 코돈과 완벽하게 매치된다. 신호는 편집 염기와 가이드의 5' 말단부 간 거리와 상관성이 있는 것으로 나타났고, 거리가 짧을수록 보다 양호한 편집성을 제공한다.

도 10: Cas13a/Cas13b 간섭을 위한 Cluc/Gluc 타일링 (tiling).

도 11A-11C: NGS에 의한 ADAR 편집 정량 (루시퍼라제 리포터).

도 12: NGS에 의한 ADAR 편집 정량 (KRAS 및 PPIB).

도 13A-13C: RNA Seq로부터의 Cas13a/b + shRNA 특이성.

도 14: 오프 표적 (off target)(A:A 또는 A:G)을 감소시키기 위한 미스매치 특이성 (SEQ ID No. 661-668).

도 15: 온-표적 (on-target) 활성에 대한 미스매치.

도 16: ADAR 모티프 선호도.

도 17: RNA 편집 효율을 증강시키기 위한 더 큰 버블.

도 18: 전사물 내 다수 A의 편집 (SEQ ID No. 669-672)

도 19는 RNA 편집을 위한 가이드 길이 적정.

도 20A-20F: 포유동물 코돈-최적화된 Cas13b 오솔로그는 고도로 효율적인 RNA 녹다운을 매개한다. (**도 20A**) 대표적인 Cas13a, Cas13b 및 Cas13c 유전자좌 및 연관된 crRNA의 개략도. (**도 20B**) HEK293FT 세포에서 Cas13a 절단 활성을 측정하기 위한 루시퍼라제 분석의 개략도. (**도 20C**) 19종 Cas13a, 15종 Cas13b 및 5종 Cas13c 오솔로그에 의해 Cluc를 표적화하는 2종의 상이한 가이드를 사용한 RNA 녹다운 효율. 루시퍼라제 발현은 비표적화 (non-targeting) 가이드 제어 조건의 발현에 대해 정규화된다. (**도 20D**) 파트 C에서 수행되는 상위 7종 오솔로그는 Cluc를 표적화하는 2종의 상이한 가이드 RNA와 3종의 상이한 NLS 및 NES 태그 존재에서 활성에 대해 분석된다. (**도 20E**) Cas13b12 및 Cas13a2 (LwCas13a)는 Gluc 및 Cluc에 대해서 녹다운 활성이 비교된다. 가이드는 전사물을 따라서 타일링되고, Cas13b12와 Cas13a2 사이의 가이드는 위치 매칭된다. (**도 20F**) 내생성 KRAS 전사물에 대한 Cas13a2, Cas13b6, Cas13b11 및 Cas13b12의 가이드 녹다운은 상응하는 shRNA와 비교된다.

도 21A-21G: Cas13 효소는 포유동물 세포에서 특이적 RNA 녹다운을 매개한다. (**도 21A**) Cas13a/b 미스매치 특이성 시험에 대한 반-축퇴성 표적 서열의 개략도 (SEQ ID No. 673-694). (도 21B) Cas13 a/b에 대한 단일 미스매치 녹다운 데이터의 히트맵. 녹다운은 각각의 효소에 대한 비표적화 (NT) 가이드에 대해 정규화된다. (**도 21C**) Cas13a에 대한 이중 미스매치 녹다운 데이터. 각각의 미스매치의 위치는 X 및 Y 축 상에 표시된다. 녹다운 데이터는 소정 위치 세트에 대한 모든 이중 미스매치의 합이다. 데이터는 각각의 효소에 대해 NT 가이드에 대해 정규화된다. (**도 21D**) Cas13b에 대한 이중 미스매치 녹다운 데이터. 설명은 C를 참조한다. (**도 21E**) Cas13 a/b에 대한 전사체-와이드 특이성 및 위치-매칭된 가이드에 대한 shRNA를 비교하는 RNA-seq 데이터. Y 축은 표적화 조건에 대한 판독 계측치를 나타내고, X 축은 비표적화 조건에 대한 계측치를 나타낸다. (**도 21F**) Cas13 a/b 및 shRNA에 대한 RNA-seq 데이터로부터 계산된 RNA 발현. (**도 21G**) RNA-seq 데이터로부터 Cas13 a/b 및 shRNA에 대한 유의한 오프-표적. FDR <0.05를 이용하여 유의한 오프-표적이 계산되었다.

도 22A-22F: 촉매적 불활성 Cas13b-ADAR 융합이 포유동물 세포에서 표적화된 RNA 편집을 가능하게 한다. (**도 22A**) 시프리디나 (Cypridina) 루시퍼라제 전사물 상의 정지 코돈을 제거하기 위한 Cas13b-ADAR 융합 단백질에 의한 RNA 편집의 개략도. (**도 22B**) 다수 가이드 위치에 대한 야생형 ADAR2와 융합된 Cas13b 및 과활성 ADAR2 E488Q 돌연변이체와 융합된 Cas13b 간 RNA 편집 비교. 루시퍼라제 발현은 가우시아 (Gaussia) 루시퍼라제 대조군 값에 대해 정규화된다. (**도 22C**) 편집 부위 주위의 다양한 위치를 표적화하도록 디자인된 30, 50, 70 및 84 nt 가이드 간 RNA 편집 비교. (**도 22D**) 시프리디나 루시퍼라제 전사물 상의 정지 코돈에 대한 시퀀싱 정량에 대해 사용된 가이드의 위치 및 길이를 도시한 개략도 (SEQ ID No. 695). (**도 22E**) 시퀀싱으로 정량된 시프리디나 루시퍼라제 전사물 상의 상응하는 아데닌 염기에서 각각의 가이드 디자인에 대한 온-표적 및 오프-표적 편집 효율. (**도 22F**) 표적화된 아데닌 반대쪽의 다양한 염기에 따른 가이드의 루시퍼라제 판독.

도 23: Cas13b-ADAR 융합체에 의한 내생성 RNA 편집. 내생성 KRAS 및 PPIB 유전자좌의 내생성 Cas13b12-ADAR 편집의 차세대 시퀀싱. 전사물 당 2개의 상이한 영역을 표적화하였고 A->G 편집은 표적화된 아데닌 근처의 모든 아데닌에서 정량되었다. A549 세포의 STK11 전사물 내 사전-종결 부위는 Cas13b12-ADAR가 표적화하여, STK11 단백질의 발현이 복원된다.

도 24: 최적의 가이드 위치를 결정하기 위한 전략.

도 25A-25C: (**도 25A**) Cas13b-huADAR2는 돌연변이된 루시퍼라제 전사물의 복구를 촉진시킨다. (**도 25B**)

Cas13b-huADAR1은 돌연변이된 루시퍼라제 전사물의 복구를 촉진시킨다. (**도 25C**) 인간 ADAR1 및 인간 ADAR2의 비교.

도 26: E488Q 대 wt dADAR2 편집의 비교. E488Q는 dADAR2의 과활성 돌연변이체이다.

도 27A-27B: Cas13b-huADAR2-E488Q에 의해 표적화된 전사물은 예상되는 A-G 편집을 함유한다. (**도 27A**) 히트맵. (도 27B) 주형 내 위치. 히트맵에서와 같이 G로의 편집율에 따라 A 부위만이 도시된다.

도 28A-28B: 가이드의 내생성 타일링. (**도 28A**) KRAS: 히트맵 (위). 주형내 위치 (아래). 히트맵에서와 같이 G로의 편집율에 따라 A 부위만이 도시된다. (**도 28B**) PPIB: 히트맵 (위). 주형 내 위치 (아래). 히트맵에서와 같이 G로의 편집율에 따라 A 부위만이 도시된다.

도 29: 비표적화 편집.

도 30: 링커 최적화.

도 31: Cas13b ADAR은 발현된 cDNA에서 환자 유래 병원성 A>G 돌연변이를 교정하기 위해 사용될 수 있다.

도 32: Cas13b-ADAR은 5' G 모티프 상에 약간의 제한성을 갖는다.

도 33A-33D: 편집 효율에 대한 효과에 대한 축퇴성 PFS 위치 스크리닝. 모든 PFS(4-N) 정체는 비-표적화보다 더 높은 편집성을 갖는다. 패널 A의 SEQ ID No. 697, 698, 699, 및 703에 표시되어 있다.

도 34: 표적 전사물 내에서 오프-표적 편집 감소.

도 35: 표적 전사물 내에서 오프-표적 편집 감소.

도 36A-36B: Cas13b-ADAR 전사체 특이성. 온-표적 편집은 71%이다. (**도 36A**) 표적화 가이드; 482개의 유의한 부위. (**도 36B**) 비-표적화 가이드; 949개의 유의한 부위. 염색체 0은 Gluc이고, 염색체 1은 Cluc이며; 인간 염색체는 그 이후의 순서임의 유의한다.

도 37A-37B: Cas13b-ADAR 전사체 특이성. (**도 37A**) 표적화 가이드. (**도 37B**) 비-표적화 가이드.

도 38: Cas13b는 경쟁적 ADAR 편집 전략과 비교해 최고 효율을 갖는다.

도 39A-39D: 경쟁적 RNA 편집 시스템. (A-B) BoxB; 온-표적 편집은 63%이다; (**도 39A**) 표적화 가이드 - 2020개의 유의한 부위; (**도 39B**) 비-표적화 가이드 - 1805개의 유의한 부위. (**도 39C-도 39D**) Stafforst; 온-표적 편집은 36%이다; (**도 39C**) 표적화 가이드 - 176개의 유의한 부위; (**도 39D**) 비-표적화 가이드 - 186개의 유의한 부위.

도 40: ADAR의 용량 적정. crRNA 양은 일정하다.

도 41A-41D: 특이성에 대한 용량 반응 효과. (**도 41A-도 41B**) 150 ng의 Cas13-ADAR; 온-표적 편집은 83%이다; (**도 41A**) 표적화 가이드 - 1231개의 유의한 부위; (**도 41B**) 비-표적화 가이드 - 520개의 유의한 부위. (**도 41C-도 41D**) 10 ng의 Cas13-ADAR; 온-표적 편집은 80%이다; (**도 41C**) 표적화 가이드 - 347개의 유의한 부위; (**도 41D**) 비-표적화 가이드 - 223개의 유의한 부위.

도 42A-42B: ADAR1이 ADAR2에 비해 더 특이적인 듯 하다. 온-표적 편집은 29%이다. (**도 42A**) 표적화 가이드; 11개의 유의한 부위. (**도 42B**) 비-표적화 가이드; 6개의 유의한 부위. 염색체 0은 Gluc이고, 염색체 1은 Cluc이며; 인간 염색체는 그 이후의 순서라는 것을 유의한다.

도 43A-43D: ADAR 특이성 돌연변이체는 증강된 특이성을 갖는다. (**도 43A**) 표적화 가이드. (**도 43B**) 비표적화 가이드. (**도 43C**) 표적화 대 비표적화 비율. (**도 43D**) 표적화 및 비-표적화 가이드.

도 44: RNA 표적과 각 잔기의 접촉점을 따라서 그래프화시킨 ADAR 돌연변이체 루시퍼라제 결과. SEQ ID No. 704 및 735로 표시되어 있다.

도 45: ADAR 특이성 돌연변이체는 증강된 특이성을 갖는다. 보라색 점은 전체 전사체 오프-표적 NGS 분석에 선택된 돌연변이체이다. 적색 점은 시작점(즉, E488Q 돌연변이체)이다. 모든 추가 돌연변이체는 또한 E488Q 돌연변이를 갖는다는 것을 유의한다.

도 46A-46B: ADAR 돌연변이체는 NGS에 따라서 더 특이적이다. (**도 46A**) 온-표적. (**도 46B**) 오프-표적.

도 47A-47B: ADAR 특이성 돌연변이체에 대한 루시퍼라제 데이타는 NGS와 일치한다. (**도 47A**) NGS를 위해 선

택된 표적화 가이드. (**도 47B**) NGS를 위해 선택된 비-표적화 가이드. 루시퍼라제 데이터는 **도 46A-도 46B**의 NGS 데이터와 일치된다. 비-표적화 가이드에 의해 더 적은 활성을 갖는 오솔로그는 전사체 전반에서 소수의 오프-표적을 가지며 그들의 온-표적 편집 효율은 표적화 가이드 루시퍼라제 조건에 따라 예측할 수 있다.

도 48A-48C: Cas13b 12의 C-말단 절단은 ADAR 편집에서 여전히 활성이 높다.

도 49A-49G: RNA 녹다운을 위한 고활성 Cas13b 오솔로그의 특징규명. **도 49A**) 전형적인 Cas13 유전자좌 및 상응하는 crRNA 구조의 개략도. **도 49B**) 2종의 상이한 가이드를 사용한 루시퍼라제 녹다운에 대한 19종 Cas13a, 15종 Cas13b 및 7종 Cas13c 오솔로그의 평가. 양쪽 가이드를 사용하여 효율적으로 녹다운되는 오솔로그는 그들 숙주 유기체 명칭으로 표지된다. **도 49C**) PspCas13b 및 LwaCas13a 녹다운 활성은 Gluc에 대해 가이드를 타일링하고 루시퍼라제 발현을 측정하여 비교된다. **도 49D**) PspCas13b 및 LwaCas13a 녹다운 활성은 Cluc에 대해 가이드를 타일링하고 루시퍼라제 발현을 특정하여 비교된다. **도 49E**) LwaCas13a (적색) 및 shRNA (검정색)에 대한 Gluc-표적화 조건 (y-축)에 비해 비표적화 대조군 (x-축)의 RNA-seq 라이브러리에서 검출된 모든 유전자의 log2(백만개당 전사물 (transcripts per million: TPM)) 값의 발현 수준. 3개 생물학적 복제물의 평균으로 나타낸다. Gluc 전사물 데이터 점이 표지된다. **도 49F**) PspCas13b (파란색) 및 shRNA (검정색)에 대한 Gluc-표적화 조건 (y-축)에 비해 비표적화 대조군 (x-축)의 RNA-seq 라이브러리에서 검출된 모든 유전자의 log2 (백만개당 전사물 (TPM)) 값의 발현 수준. 3개 생물학적 복제물의 평균으로 나타낸다. Gluc 전사물 데이터 점이 표지된다. **도 49G**) E 및 F의 전사체 와이드 분석에서 LwaCas13a, PspCas13b, 및 shRNA에 대한 Gluc 녹다운으로부터의 유의한 오프-표적의 개수.

도 50A-50E: RNA 편집을 위한 dCas13b-ADAR 융합 조작. **도 50A**) dCas13b-ADAR 융합 단백질에 의한 RNA 편집 개략도. **도 50B**) 시프리디나 루시퍼라제 W85X 표적 및 표적화 가이드 디자인의 개략도 (SEQ ID No. 700 및 701). **도 50C**) 길이 30, 50, 70 또는 84 nt의 타일링 가이드에 의한 Cas13b-dADAR1 (좌) 및 Cas13b-ADAR2-cd (우)에 대한 루시퍼라제 활성 복원의 정량. **도 50D**) 시프리디니아 루시퍼라제 W85X를 표적화하기 위한 표적 부위의 개략도 (SEQ ID No. 702). **도 50E**) 시프리디니아 루시퍼라제 W85X를 표적화하는 50 nt 가이드에 대한 A->I 편집의 시퀀싱 정량.

도 51A-51D: 도 51A) REPAIRv1에 의한 RNA 편집에 대한 서열 탄력성 측정. REPAIRv1에 의한 RNA 편집의 프로토스페이서 측접 부위 (Protospacer Flanking Site: PFS) 선호도를 결정하기 위한 스크리닝의 개략도. SEQ ID No. 737로 표시되어 있다. **도 51B**) 2개의 상이한 편집 부위에서 모든 4-N PFS 조합에 대한 RNA 편집 효율의 분포. **도 51C**) 모든 가능한 3개 염기 모티 전반의 Cluc W85에서 REPAIRv1의 편집율의 정량. **도 51D**) 모든 가능한 3개의 염기 모티프에 대해 Cluc W85에서 RNA 편집의 5' 및 3' 염기 선호도의 히트맵.

도 52A-52G: REPAIRv1에 의한 질환-관련 돌연변이의 교정. **도 52A**) *AVPR2* 878G>A의 표적화를 위한 표적 및 가이드 디자인의 개략도 (SEQ ID No. 705-708). **도 52B**) *AVPR2*에서 878G>A 돌연변이는 3종의 상이한 가이드 디자인의 REPAIRv1을 사용하여 다양한 백분율로 교정된다. **도 52C**) *FANCC* 1517G>A의 표적화를 위한 표적 및 가이드 디자인의 개략도 (SEQ ID No. 709-712). **도 52D**) *FANCC*에서 1517G>A 돌연변이는 3종 상이한 가이드 디자인의 REPAIRv1을 사용하여 다양한 백분율로 교정된다. **도 52E**) REPAIRv1을 사용한 34종의 상이한 질환-관련 G>A 돌연변이의 편집율의 정량. **도 52F**) ClinVar 데이터베이스에서 주석 달린 바와 같이 교정될 수 있는 모든 가능한 G>A 돌연변이의 분석. **도 52G**) ClinVar에서 모든 G>A 돌연변이에 대한 편집 모티프의 분포는 Gluc 전사물 상에서 정량된 모티프 당 REPAIRv1에 의한 편집 효율에 대비해서 도시된다.

도 53A-53D: REPAIRv1의 특이성 특징규명. **도 53A**) *KRAS* 표적 부위 및 가이드 디자인의 개략도 (SEQ ID No. 713-720). **도 53B**) 타일링된 *KRAS*-표적화 가이드에 대한 편집 백분율의 정량. 편집 백분율은 온-표적 및 이웃 아데노신 부위에서 도시된다. 각 가이드의 경우에, 듀플렉스 RNA의 영역은 적색 직사각형으로 표시된다. **도 53C**) Cluc 표적화 가이드 존재의 REPAIRv1에 의한 유의한 RNA 편집의 전사체-와이드 부위. 온-표적 부위 Cluc 부위 (254 A>G)는 오렌지색으로 강조된다. **도 53D**) 비-표적화 가이드 존재의 REPAIRv1에 의한 유의한 RNA 편집의 전사체-와이드 부위.

도 54A-54F: REPAIRv1의 특이성을 개선시키기 위한 ADAR2의 합리적 돌연변이유발법. **도 54A**) 핵심 ADAR2 데아미나제 잔기 및 dsRNA 표적 간 접촉부에 대한 개략도를 따라 그래프화된 그들 특이성 점수를 비롯하여 다양한 dCas13-ADAR2 돌연변이체에 의한 루시퍼라제 신호 복원의 정량. 특이성 점수는 표적화 가이드 및 비표적화 가이드 조건 간 루시퍼라제 신호의 비율로서 정의된다. SEQ ID No. 738 및 739로 표시되어 있다. **도 54B**) 그들 특이성 점수 대비 다양한 dCas13-ADAR2 돌연변이체에 의한 루시퍼라제 신호 복원의 정량. **도 54C**) mRNA의 전사체 와이드 시퀀싱에 의한 각 dCas13-ADAR2 돌연변이체에 대한 유의한 오프-표적의 개수를 비롯

하여 온-표적 편집 분율의 측정. **도 54D**) Cluc 내 사전종결 부위를 표적화하는 가이드 존재의 REPAIRv1 및 REPAIRv2에 의한 유의한 RNA 편집의 전사체-와이드 부위. 온-표적 Cluc 부위 (254 A>G)는 오렌지 색으로 강조된다. **도 54E**) REPAIRv1 및 REPAIRv2 간 오프-표적 편집 차이를 강조하는 온-표적 Cluc 편집 부위 (SEQ ID No. 721, 740-745, 및 763-792) (254 A>G) 주변에서 RNA 시퀀싱 판독. 모든 A>G 편집은 적색으로 강조되는 한편 시퀀싱 오차는 파란색으로 강조된다. **도 54F**) 내생성 *KRAS* 및 *PPIB* 전사물에서 프레임 밖 UAG 부위를 표적화하는 가이드 존재의 REPAIRv1 및 REPAIRv2에 의한 RNA 편집. 온-표적 편집 분율은 각각의 조건 행에 대해 우측의 옆쪽 막대 도표로서 도시된다. 가이드 RNA에 의해 형성된 듀플렉스 영역은 적색 윤곽의 박스로 도시된다.

도 55A-55C: 생체내 효율 및 PFS 결정을 위한 Cas13b 오솔로그의 박테리아 스크리닝. **도 55A**) Cas13b 오솔로그의 PFS를 결정하기 위한 박테리아 분석의 개략도. 베타-락타마제 표적화 스페이서 (SEQ ID No. 722)를 갖는 Cas13b 오솔로그를 베타-락타마제 발현 플라스미드와 함께 공동 형질전환시키고 이중 선별을 수행한다. **도 55B**) 콜로니 형성 유닛 (cfu)으로 측정된 베타-락타마제를 표적화하는 Cas13b 오솔로그의 간섭 활성의 정량. **도 55C**) 박테리아 분석으로부터 감손 서열을 통해 결정되는 Cas13b 오솔로그에 대한 PFS 로고 (logo).

도 56A-56E: Cas13b 녹다운의 최적화 및 미스매치 특이성의 추가 특징규명. **도 56A**) 2종의 상이한 가이드에 의한 Gluc 녹다운은 다양한 핵 국재화 및 핵 유출 태그에 융합된 상위 2종의 Cas13a 및 상위 4종의 Cas13b 오솔로그를 사용하여 측정된다. **도 56B**) *KRAS*의 녹다운은 4종의 상이한 가이드로 LwaCas13a, RanCas13b, PguCas13b 및 PspCas13b에 대해 측정되고, 4종의 위치-매칭된 shRNA 대조군과 비교된다. **도 56C**) LwaCas13a 및 PspCas13b 녹다운의 특이성을 평가하기 위해 사용된 단일 및 이중 미스매치 플라스미드 라이브러리의 개략도. 모든 가능한 단일 및 이중 미스매치는 표적 서열을 비롯하여 표적 부위의 5' 및 3' 말단부에 직접 측접된 3 위치에 존재한다 (SEQ ID No. 723-734). **도 56D**) 표시된 단일 미스매치를 갖는 전사물의 감손 수준은 LwaCas13a와 PspCas13b 조건 둘 모두에 대한 히트맵으로서 그래프화된다. (SEQ ID No. 723 및 736) **도 56E**) 표시된 이중 매치를 갖는 전사체의 감손 수준은 LwaCas13a와 PspCas13b 조건 둘 모두에 대한 히트맵으로서 플롯팅된다 (SEQ ID No. 723 및 736).

도 57A-57F: dCas13-ADAR2 RNA 편집에 대한 디자인 매개변수의 특징규명. **도 57A**) 야생형 Cas13b 및 촉매적 불활성 H133A/H1058A Cas13b (dCas13b)에 대한 Gluc 표적화의 녹다운 효율. **도 57B**) 타일링 Cluc 표적화 가이드로 시험된, 야생형 ADAR2 촉매적 도메인 또는 과활성 E488Q 돌연변이체 ADAR2 촉매적 도메인 중 하나에 융합된 dCas13b에 의한 루시퍼라제 활성 복원의 정량. **도 57C**) 시프리디니아 루시퍼라제 W85X를 표적화하는 30 nt 가이드에 대한 A->I 편집의 가이드 디자인 및 시퀀싱 정량 (SEQ ID No. 793-801). **도 57D**) *PPIB*를 표적화하는 50 nt 가이드에 대한 A->I 편집의 가이드 디자인 및 시퀀싱 정량 (SEQ ID No. 746-753). **도 57E**) REPAIRv1에 의한 루시퍼라제 활성 복원에 대한 링커 선택의 영향. **도 57F**) REPAIRv1에 의한 루시퍼라제 활성 복원에 대한 표적화된 아데노신 반대쪽 염기 정체의 영향 (SEQ ID No. 754 및 755).

도 58: G>A 돌연변이에 대한 ClinVar 모티프 분포. 모든 G>A 돌연변이에 대한 ClinVar 데이터베이스에서 관찰된 각각의 가능한 삼중 모티프의 개수.

도 59: dCas13b의 절단은 여전히 기능성 RNA 편집성을 갖는다. dCas13b의 다양한 N-말단 및 C-말단 절단은 루시퍼라제 신호의 복원으로 측정되는 RNA 편집을 가능하게 한다.

도 60A-60F: 다른 프로그램 가능한 ADAR 시스템과 dCas13-ADAR2 편집자의 비교. **도 60A**) 2종의 프로그램 가능한 ADAR 전략의 개략도: BoxB-기반 표적화 및 전체 길이 ADAR2 표적화. BoxB 개략도 (위)에서, ADAR2 데아미나제 도메인 (ADAR2$_{DD}$(E488Q))은 BoxB-λ로 불리는 소형 RNA 서열에 특이적으로 결합하는 람다 N (λN)이라고 하는 소형 박테리아 바이러스 단백질에 융합된다. 2개의 BoxB-λ 헤어핀을 함유하는 가이드 RNA는 부위 특이적 편집을 위해서, ADAR2 $_{DD}$(E488Q), -λN을 가이드할 수 있다. 전체 길이 ADAR2 개략도 (아래)에서, ADAR2의 dsRNA 결합 도메인은 가이드 RNA 내 헤어핀에 결합하여, 프로그램가능한 ADAR2 편집을 가능하게 한다 (SEQ ID No. 756-760). **도 60B**) Cluc 표적화 가이드 및 비-표적화 가이드 존재의 BoxB-ADAR2 $_{DD}$(E488Q)에 의한 유의한 RNA 편집의 전사체-와이드 부위. 온-표적 Cluc 부위 (254 A>G)는 오렌지색으로 강조된다. **도 60C**) Cluc 표적화 가이드 및 비-표적화 가이드 존재의 ADAR2에 의한 유의한 RNA 편집의 전사체-와이드 부위. 온-표적 Cluc 부위 (254 A>G)는 오렌지색으로 강조된다. **도 60D**) Cluc 표적화 가이드 및 비-표적화 가이드 존재의 REPAIRv1에 의한 유의한 RNA 편집의 전사체-와이드 부위. 온-표적 Cluc 부위 (254 A>G)는 오렌지색으로 강조된다. **도 60E**) Cluc 표적화 가이드에 대한 BoxB-ADAR2 $_{DD}$(E488Q), ADAR2 및 REPAIRv1의 온-표적

편집율%의 정량. **도 60F**) 프로그램가능한 ADAR 시스템에 대한 상이한 표적화 조건 및 비표적화 조건 간 오프-표적 부위의 중복.

도 61A-61C: dCas13b-ADAR2 돌연변이체의 효율 및 특이성. **도 61A**) Cluc-표적화 및 비-표적화 가이드에 대한 dCas13b-ADAR2$_{DD}$(E488Q) 돌연변이체에 의한 루시퍼라제 활성 복원의 정량. **도 61B**) 전사체-와이드 시퀀싱으로 정량된 RNA-편집 오프-표적의 개수 및 표적화 및 비표적화 가이드의 비율 간 관련성. **도 61C**) 전사체-와이드 오프-표적 RNA 편집 부위의 개수 대 dCas13b-ADAR2$_{DD}$(E488Q) 돌연변이체에 대한 온-표적 Cluc 편집 효율의 정량.

도 62A-62B: dCas13b-ADAR2$_{DD}$(E488Q) 돌연변이체에 의한 RNA 편집의 전사체-와이드 특이성. **도 62A**) Cluc 표적화 가이드 존재의 dCas13b-ADAR2$_{DD}$(E488Q) 돌연변이체에 의한 유의한 RNA 편집의 전사체-와이드 부위. 온-표적 Cluc 부위 (254 A>G)는 오렌지색으로 강조된다. **도 62B**) 비-표적화 가이드 존재의 dCas13b-ADAR2$_{DD}$(E488Q) 돌연변이체에 의한 유의한 RNA 편집의 전사체-와이드 부위.

도 63A-63C: dCas13b-ADAR2$_{DD}$(E488Q) 편집의 오프-표적에서 모티프 편향성의 특징규명. **도 63A**) 각각의 dCas13b-ADAR2$_{DD}$(E488Q) 돌연변이체의 경우, 전사체 내 모든 A>G 오프-표적 편집에 걸쳐 존재하는 모티프가 도시된다. **도 63B**) 모티프 정체 당 오프-표적 A>I 편집의 분포가 표적화 및 비-표적화 가이드 존재의 REPAIRv1에 대해 도시된다. **도 63C**) 모티프 정체 당 오프-표적 A>I 편집의 분포가 표적화 및 비-표적화 가이드 존재의 REPAIRv2에 대해 도시된다.

도 64A-64F: REPAIRv1 및 REPAIRv2 오프-표적의 추가 특징규명. **도 64A**) REPAIRv1에 대한 전사물 당 오프-표적 개수의 히스토그램. **도 64B**) REPAIRv2에 대한 전사물 당 오프-표적 개수의 히스토그램. **도 64C**) REPAIRv1 오프 표적의 변이체 효과 예측. **도 64D**) REPAIRv1 오프 표적의 잠재적인 종양원성 효과의 분포. **도 64E**) REPAIRv2 오프 표적의 잠재적인 종양원성 효과의 분포. **도 64F**) REPAIRv2 오프 표적의 잠재적인 종양유전자 효과의 분포.

도 65A-65C: REPAIRv1 및 REPAIRv2의 RNA 편집 효율 및 특이성. **도 65A**) REPAIRv1 및 REPAIRv2에 대한 표적화된 아데노신 및 이웃 부위에서 *KRAS*-표적화 가이드 1에 의한 *KRAS*의 편집%의 정량. **도 65B**) REPAIRv1 및 REPAIRv2에 대한 표적화된 아데노신 및 이웃 부위에서 *KRAS*-표적화 가이드 3에 의한 *KRAS*의 편집%의 정량. **도 65C**) REPAIRv1 및 REPAIRv2에 대한 표적화된 아데노신 및 이웃 부위에서 *PPIB*-표적화 가이드 2에 의한 *PPIB*의 편집%의 정량.

도 66A-66G: A>G RNA 편집자에 의한 모든 잠재적인 코돈 변화의 입증. **도 66A**) A>I 편집에 의해 가능한 모든 잠재적인 코돈 전이의 표. **도 66B**) A>I 편집에 의해 가능한 모든 잠재적인 코돈 전이를 입증하는 코돈 표. **도 66C** 는 ADAR 및 dCas13b에 의한 유전자 편집을 도시한다. **도 66D**) 는 ADAR에 의한 멘델 질환을 교정하는 일례를 도시한다. **도 66E**) 는 질환-보호적 대립유전자의 다중복합적인 생성의 일례를 도시한다. **도 66F**) 는 단백질 기능 조절의 일례를 도시한다. **도 66G**) 는 스플라이싱 조절 일례를 도시한다.

도 67: Psp dCas13b의 추가 절단.

도 68: 오프 표적 활성에 대한 용량의 잠재적 효과.

도 69 는 스크리닝된 ADAR 돌연변이의 활성을 도시한다.

도 70 은 상이한 모티프에 대한 돌연변이 P462A의 활성을 도시한다.

도 71 은 돌연변이 P462가 모티프를 개선시킨 것을 도시한다.

도 72 는 N579I를 동정한 ADAR 돌연변이 스크리닝의 활성을 도시한다.

도 73 은 N579I의 검증을 도시한다.

도 74 는 30 bp 가이드 존재의 RESCUEv3의 성능을 도시한다.

도 75 는 50 bp 가이드 및 30 bp 가이드와 다양한 ADAR 돌연변이체의 비교를 도시한다.

도 76 은 ADAR 돌연변이의 활성이 Cas13 의존적이었음을 도시한다.

본 명세서의 도면은 오직 예시의 목적이며 반드시 일정 비율로 도시되는 것은 아니다.

발명을 실시하기 위한 구체적인 내용

[0049] **일반 정의**

[0050] 달리 정의하지 않으면, 본 명세서에서 사용되는 기술 및 과학 용어는 본 개시가 속하는 분야의 당업자가 통상적으로 이해하는 바와 동일한 의미를 갖는다. 분자 생물학의 공통 용어 및 기술의 정의는 다음의 문헌들에서 확인할 수 있다: [Molecular Cloning: A Laboratory Manual, 2nd edition (1989) (Sambrook, Fritsch, and Maniatis)]; [Molecular Cloning: A Laboratory Manual, 4th edition (2012) (Green and Sambrook)]; [Current Protocols in Molecular Biology (1987) (F.M. Ausubel et al. eds.)]; [the series Methods in Enzymology (Academic Press, Inc.): PCR 2: A Practical Approach (1995) (M.J. MacPherson, B.D. Hames, and G.R. Taylor eds.): Antibodies, A Laboratory Manual (1988) (Harlow and Lane, eds.): Antibodies A Laboratory Manual, 2nd edition 2013 (E.A. Greenfield ed.)]; [Animal Cell Culture (1987) (R.I. Freshney, ed.)]; [Benjamin Lewin, Genes IX, published by Jones and Bartlet, 2008 (ISBN 0763752223)]; [Kendrew et al. (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0632021829)]; [Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 9780471185710)]; [Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, N.Y. 1992)]; 및 [Marten H. Hofker and Jan van Deursen, Transgenic Mouse Methods and Protocols, 2nd edition (2011)].

[0051] 2016년 6월 17일 출원된 US 가출원 제62/351,662호 및 제62/351,803호, 2016년 8월 17일 출원된 US 가출원 제62/376,377호, 2016년 10월 19일 출원된 US 가출원 제62/410,366호, 2016년 12월 9일 출원된 US 가출원 제62/432,240호, 2017년 3월 15일 출원된 US 가출원 제62/471,792호, 및 2017년 4월 12일 출원된 US 가출원 제62/484,786호를 참조한다. 2017년 6월 19일 출원된 PCT 국제 특허 출원 번호 PCT/US2017/038154를 참조한다. 2017년 3월 15일 출원된 US 가출원 제 62/471,710호 (발명의 명칭: "Novel Cas13B Orthologues CRISPR Enzymes and Systems," 대리인 참조 번호: BI-10157 VP 47627.04.2149)을 참조한다. 또한, 2016년 12월 9일 출원된 US 가출원 제62/432,553허, 2017년 2월 8일 출원된 US 가출원 제62/456,645호, 및 2017년 3월 15일 출원된 US 가출원 제62/471,930호 (발명의 명칭: "CRISPR Effector System Based Diagnostics," 대리인 참조 번호: BI-10121 BROD 0842P) 및 2017년 4월 12일 출원되어 양도될 것인 US 가출원 (발명의 명칭: "CRISPR Effector System Based Diagnostics," 대리인 참조 번호: BI-10121 BROD 0843P)을 참조한다.

[0052] 본 명세서에서 사용되는 단수 형태 "한", "하나", 및 "그"는 문맥에서 달리 명확하게 명시하지 않으면, 단수 및 복수 대상 둘 모두를 포함한다.

[0053] 용어 "임의의" 또는 "임의로는"은 후술되는 사건, 상황 또는 치환기가 존재하지 않을 수도 있거나 또는 존재할 수도 있고, 그 설명은 사건 또는 상황이 일어나는 예 및 일어나지 않는 예를 포함한다는 것을 의미한다.

[0054] 종료점에 의한 수치 범위의 설명은 언급된 종료점을 비롯하여, 각 범위 내에 포함된 모든 수 및 분수를 포함한다.

[0055] 측정가능한 값 예컨대 매개변수, 양, 시간적 지속기간 등을 언급할 때 본 명세서에서 사용되는 용어 "약" 또는 "대략"은 명시된 값과 그로부터의 변동, 예컨대 그러한 변동이 개시된 발명에서 수행하기에 적절하다면, 명시된 값과 그로부터의 +/-10% 이하, +/-5% 이하, +/-1% 이하, 및 +/-0.1% 이하의 변동을 포괄한다는 것을 의미한다.
수식어 "약" 또는 "대략"이 언급되는 값은 그 자체로 또한 특별히, 그리고 바람직하게 개시된다는 것을 이해해야 한다.

[0056] 본 명세서 전반에서 "하나의 실시형태", "일 실시형태", "일례의 실시형태"에 대한 언급은 실시형태와 함께 기술된 특정한 특성, 구조 또는 특징이 본 발명의 적어도 하나의 실시형태에 포함된다는 것을 의미한다. 따라서, 본 명세서 전반의 다양한 위치에서 어구 "하나의 실시형태에서", "일 실시형태에서", 또는 "일례의 실시형태"의 출현은 반드시 모두 동일한 실시형태를 언급하지 않지만, 그럴 수도 있다. 더 나아가서, 특정한 특성, 구조 또는 특징은 하나 이상의 실시형태에서, 본 개시로부터 당업자에게 자명하게 되는 바와 같이, 임의

의 적합한 방식으로 조합될 수 있다. 더 나아가서, 본 명세서에 기술된 일부 실시형태가 다른 실시형태에 포함된 다른 특성이 아닌 일부 특성을 포함하지만, 상이한 실시형태의 특성의 조합은 본 발명의 범주 내에 두고자 한다. 예를 들어, 첨부된 청구항에서, 임의의 청구된 실시형태들은 임의 조합으로 사용될 수 있다.

[0057] C2c2는 Cas13a로서 공지되어 있다. 본 명세서에서 용어 "C2c2"는 "Cas13a"와 상호교환적으로 사용된다는 것을 이해하게 될 것이다.

[0058] 본 명세서에서 인용되는 모든 출판물, 공개 특허 문서, 및 특허 출원은 각각의 개별 출판물, 공개 특허 문서, 또는 특허 출원이 참조로 편입된다고 특별히 개별적으로 표시한 바와 동일한 정도로 참조로 본 명세서에 편입된다.

[0059] 이하 다양한 실시형태를 기술한다. 특별한 실시형태들이 배타적인 설명으로서 또는 본 명세서에서 논의되는 더 넓은 양상에 대한 제한으로서 의도되지 않는다는 것을 주의해야 한다. 특정 실시형태와 함께 기술되는 일 양상은 반드시 그 실시형태로 국한되는 것이 아니며 임의의 다른 실시형태(들)로 실시될 수 있다.

[0060] **개요**

[0061] 본 명세서에 개시된 실시형태는 표적화된 염기 편집을 위한 시스템, 구성체, 및 방법을 제공한다. 일반적으로, 본 명세서에 개시된 시스템은 표적화 성분 및 염기 편집 성분을 포함한다. 표적화 성분은 하나 이상의 뉴클레오티드를 편집하려는 표적 뉴클레오티드 서열에 대해 염기 편집 성분을 특이적으로 표적화하는 기능을 한다. 다음으로 염기 편집 성분은 표적 서열 중 제1 뉴클레오티드를 제2 뉴클레오티드로 전환시키는 화학 반응을 촉매할 수 있다. 예를 들어, 염기 편집자는 세포의 전사 또는 번역 기구가 구아닌으로서 인식하도록 아데닌의 전환을 촉매할 수 있거나, 또는 그 반대일 수도 있다. 유사하게, 염기 편집 성분은 우라실로 시티딘의 전환을 촉매할 수 있거나, 또는 그 반대일 수도 있다. 일정 예의 실시형태에서, 염기 편집자는 기지의 염기 편집자, 예컨대 아데닌 데아미나제 또는 시티딘 데아미나제로부터 출발하여 유도될 수 있고, 새로운 기능성을 유도시키기 위해 유도 진화와 같은 방법을 사용해 변형될 수 있다. 유도 진화 기술은 당분야에 공지되어 있고 WO 2015/184016 "High-Throughput Assembly of Genetic Permuatations"에 기술된 것을 포함할 수 있다.

[0062] 일정 예의 실시형태에서, 아데노신 데아미나제 단백질은 RNA-결합 단백질 또는 이의 RNA-결합 기능성 도메인에 융합 또는 연결되어 관심 표적 유전자좌로 동원된다. RNA-결합 단백질은 표적 유전자좌에서 또는 그에 인접한 서열, 모티프, 또는 구조적 특성에 결합할 수 있다. 구조적 특성은 헤어핀, 테트라루프, 또는 핵산의 다른 2차 구조적 특성을 포함할 수 있다. 본 명세서에서 사용되는 "인접한"은 아데노신 데아미나제가 이의 염기 편집 기능을 완료할 수 있는 표적 유전자좌의 거리 및/또는 배향 내인 것을 의미한다. 일정 예의 실시형태에서, RNA-결합 단백질 또는 이의 기능성 도메인은 가이드 분자의 서열, 모티프, 또는 구조적 서열에 결합할 수 있다. 가이드 분자는 관심 표적 유전자좌에 하이브리드화하는 서열을 포함하고, 그리하여 표적 유전자좌로 아데노신 데아미나제를 유도시킨다.

[0063] 일정 예의 실시형태에서, RNA-결합 단백질 또는 이의 기능성 도메인은 RNA 인식 모티프를 포함한다. RNA 인식 모티프를 포함하는 예시적인 RNA-결합 단백질은 제한없이, 다음을 포함한다: A2BP1; ACF; BOLL; BRUNOL4; BRUNOL5; BRUNOL6; CCBL2; CGI96; CIRBP; CNOT4; CPEB2; CPEB3; CPEB4; CPSF7; CSTF2; CSTF2T; CUGBP1; CUGBP2; D10S102; DAZ1; DAZ2; DAZ3; DAZ4; DAZAP1; DAZL; DNAJC17; DND1; EIF3S4; EIF3S9; EIF4B; EIF4H; ELAVL1; ELAVL2; ELAVL3; ELAVL4; ENOX1; ENOX2; EWSR1; FUS; FUSIP1; G3BP; G3BP1; G3BP2; GRSF1; HNRNPL; HNRPA0; HNRPA1; HNRPA2B1; HNRPA3; HNRPAB; HNRPC; HNRPCL1; HNRPD; HNRPDL; HNRPF; HNRPH1; HNRPH2; HNRPH3; HNRPL; HNRPLL; HNRPM; HNRPR; HRNBP1; HSU53209; HTATSF1; IGF2BP1; IGF2BP2; IGF2BP3; LARP7; MKI67IP; MSI1; MSI2; MSSP2; MTHFSD; MYEF2; NCBP2; NCL; NOL8; NONO; P14; PABPC1; PABPC1L; PABPC3; PABPC4; PABPC5; PABPN1; POLDIP3; PPARGC1; PPARGC1A; PPARGC1B; PPIE; PPIL4; PPRC1; PSPC1; PTBP1; PTBP2; PUF60; RALY; RALYL; RAVER1; RAVER2; RBM10; RBM11; RBM12; RBM12B; RBM14; RBM15; RBM15B; RBM16; RBM17; RBM18; RBM19; RBM22; RBM23; RBM24; RBM25; RBM26; RBM27; RBM28; RBM3; RBM32B; RBM33; RBM34; RBM35A; RBM35B; RBM38; RBM39; RBM4; RBM41; RBM42; RBM44; RBM45; RBM46; RBM47; RBM4B; RBM5; RBM7; RBM8A; RBM9; RBMS1; RBMS2; RBMS3; RBMX; RBMX2; RBMXL2; RBMY1A1; RBMY1B; RBMY1E; RBMY1F; RBMY2FP; RBPMS; RBPMS2; RDBP; RNPC3; RNPC4; RNPS1; ROD1; SAFB; SAFB2; SART3; SETD1A; SF3B14; SF3B4; SFPQ; SFRS1; SFRS10; SFRS11; SFRS12; SFRS15; SFRS2; SFRS2B; SFRS3; SFRS4; SFRS5; SFRS6; SFRS7; SFRS9; SLIRP; SLTM; SNRP70; SNRPA; SNRPB2; SPEN; SR140; SRRP35; SSB; SYNCRIP; TAF15; TARDBP; THOC4; TIA1; TIAL1; TNRC4; TNRC6C; TRA2A; TRSPAP1; TUT1; U1SNRNPBP; U2AF1; U2AF2; UHMK1; ZCRB1; ZNF638; ZRSR1; 및 ZRSR2.

[0064] 일정 예의 실시형태에서, RNA-결합 단백질 또는 이의 기능 도메인은 K 상동성 도메인을 포함할 수 있다. K 상동성 도메인을 포함하는 예시적인 RNA-결합 단백질은 제한없이, 다음을 포함한다: AKAP1; ANKHD1; ANKRD17; ASCC1; BICC1; DDX43; DDX53; DPPA5; FMR1; FUBP1; FUBP3; FXR1; FXR2; GLD1; HDLBP; HNRPK; IGF2BP1; IGF2BP2; IGF2BP3; KHDRBS1; KHDRBS2; KHDRBS3; KHSRP; KRR1; MEX3A; MEX3B; MEX3C; MEX3D; NOVA1; NOVA2; PCBP1; PCBP2; PCBP3; PCBP4; PNO1; PNPT1; QKI; SF1; 및 TDRKH.

[0065] 일정 예의 실시형태에서, RNA-결합 단백질은 징크 핑거 모티프를 포함한다. RNA-결합 단백질 또는 이의 기능성 도메인은 Cys2-His2, Gag-knuckle, Treble-clet, 징크 리본, Zn2/Cys6 클래스 모티프를 포함할 수 있다.

[0066] 일정 예의 실시형태에서, RNA-결합 단백질은 푸밀리오 (Pumilio) 상동성 도메인을 포함할 수 있다.

[0067] 일 양상에서 본 발명은 RNA, 보다 특히 관심 RNA 서열에서 아데닌의 표적화된 탈아미노화를 위한 방법을 제공한다. 본 발명의 방법에 따라서, 아데노신 데아미나제 (AD) 단백질은 표적 서열에 특이적으로 결합할 수 있는 CRISPR-Cas 복합체에 의해 관심 RNA 서열 내 관련 아데닌으로 특이적으로 동원된다. 이를 달성하기 위해서, 아데노신 데아미나제 단백질은 CRISPR-Cas 효소에 공유적으로 연결될 수 있거나 또는 별개 단백질로서 제공되지만, CRISPR-Cas 복합체로 이의 동원을 보장하기 위해서 적합화된다.

[0068] 본 발명의 방법의 특정 실시형태에서, 표적 유전자좌로 아데노신 데아미나제의 동원은 Cas13 단백질인 CRISPR-Cas 단백질에 아데노신 데아미나제 또는 이의 촉매 도메인을 융합시켜 보장된다. 2종의 별개 단백질로부터 융합 단백질을 생성시키는 방법은 당분야에 공지되어 있고 전형적으로 스페이서 또는 링커의 사용을 포함한다. Cas13 단백질은 아데노신 데아미나제 단백질 또는 이의 촉매 도메인과 이의 N-말단 또는 C-말단 상에서 융합될 수 있다. 특정 실시형태에서, CRISPR-Cas 단백질은 불활성화되거나 또는 데드 Cas13 단백질이고 데아미나제 단백질 또는 이의 촉매 도메인의 N-말단에 연결된다.

[0069] 융합 단백질에 대해서 사용되는 용어 "링커"는 융합 단백질을 형성하도록 단백질을 연결시키는 분자를 의미한다. 일반적으로, 이러한 분자는 단백질 간 일부 최소 거리 또는 다른 공간적 관련성을 보존하거나 또는 연결시키는 것 이외에 특별한 생물학적 활성을 갖지는 않는다. 그러나, 일정 실시형태에서, 링커는 링커 및/또는 융합 단백질의 일부 속성 예컨대 링커의 폴딩, 순전하, 또는 소수성에 영향을 미치도록 선택될 수도 있다.

[0070] 본 발명의 방법에서 사용을 위해 적합한 링커는 당업자에게 충분히 공지되어 있고 제한없이, 직쇄 또는 분지쇄 탄소 링커, 복소환 탄소 링커, 또는 펩티드 링커를 포함한다. 그러나, 본 명세서에서 사용되는 링커는 또한 공유 결합 (탄소-탄소 결합 또는 탄소-이종원자 결합)일 수 있다. 특정 실시형태에서, 링커는 각 단백질이 이의 필요한 기능적 성질을 유지하도록 보장하기 위해 충분한 거리만큼 CRISPR-Cas 단백질 및 아데노신 데아미나제를 분리시키는데 사용된다. 바람직한 펩티드 링커 서열은 가요성의 연장된 입체형태를 채택하고 정렬된 2차 구조를 발생시키는 경향을 보이지 않는다. 일정 실시형태에서, 링커는 단량체, 이량체, 다량체 또는 중합체일 수 있는 화학적 모이어티일 수 있다. 바람직하게, 링커는 아미노산을 포함한다. 가요성 링커 내 전형적인 아미노산은 Gly, Asn 및 Ser을 포함한다. 따라서, 특정 실시형태에서, 링커는 Gly, Asn 및 Ser 아미노산 중 하나 이상의 조합을 포함한다. 다른 근처의 중성 아미노산, 예컨대 Thr 및 Ala이 또한 링커 서열에서 사용될 수 있다. 예시적인 링커는 [Maratea et al. (1985), Gene 40: 39-46]; [Murphy et al. (1986) Proc. Nat'l. Acad. Sci. USA 83: 8258-62]; U.S. 특허 제4,935,233호; 및 U.S. 특허 제4,751,180호에 개시되어 있다. 예를 들어, GlySer 링커 GGS, GGGS 또는 GSG가 사용될 수 있다. GGS, GSG, GGGS 또는 GGGGS 링커는 적합한 길이를 제공하기 위해서, 3 (예컨대 (GGS)3 (SEQ ID No. 12), (GGGGS)3) 또는 5, 6, 7, 9 또는 심지어 12 (SEQ ID No. 13) 또는 그 이상으로 반복되어 사용될 수 있다. 특정한 실시형태에서, (GGGGS)3과 같은 링커가 바람직하게 본 명세서에서 사용된다. (GGGGS)6 (GGGGS)9 또는 (GGGGS)12가 바람직하게는 대안으로서 사용될 수 있다. 다른 바람직한 대안은 (GGGGS)1 (SEQ ID No 14), (GGGGS)2 (SEQ ID No. 15), (GGGGS)4, (GGGGS)5, (GGGGS)7, (GGGGS)8, (GGGGS)10, 또는 (GGGGS)11이다. 역시 추가의 실시형태에서, LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID No.11)이 링커로서 사용된다. 역시 추가의 실시형태에서, 링커는 XTEN 링커이다. 특정한 실시형태에서, CRISPR-cas 단백질은 Cas13 단백질이고 LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID No. 11) 링커에 의해서 데아미나제 단백질 또는 이의 촉매 도메인에 연결된다. 추가의 특정 실시형태에서, Cas13 단백질은 LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID No. 11) 링커를 통해서 데아미나제 단백질 또는 이의 촉매 도메인의 N-말단과 C-말단에서 연결된다. 또한, N-말단 및 C-말단 NLS가 역시 링커 (예를 들어, PKKKRKVEASSPKKRKVEAS (SEQ ID No. 16))로서 기능할 수 있다.

[0071] 본 발명의 방법의 특정 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 세포로 전달되거나

또는 별개 단백질로서 세포 내에서 발현되지만, Cas13 단백질 또는 가이드 분자에 연결될 수 있도록 변형된다.

특정한 실시형태에서, 이것은 다양한 박테리오파지 외피 단백질 내에 존재하는 직교성 RNA-결합 단백질 또는 어댑터 단백질/압타머 조합의 사용에 의해 보장된다. 이러한 외피 단백질의 예는 제한없이 MS2, Qβ, F2, GA, fr, JP501, M12, R17, BZ13, JP34, JP500, KU1, M11, MX1, TW18, VK, SP, FI, ID2, NL95, TW19, AP205, ΦCb5, ΦCb8r, ΦCb12r, ΦCb23r, 7s 및 PRR1을 포함한다. 압타머는 천연 발생일 수 있거나 또는 특이적 표적에 결합하도록 시험관내 선별 또는 SELEX (systematic evolution of ligands by exponential enrichment)의 반복 라운드를 통해서 조작된 합성 올리고뉴클레오티드일 수 있다.

[0072] 본 발명의 방법 및 시스템의 특정한 실시형태에서, 가이드 분자는 어댑터 단백질을 동원할 수 있는 하나 이상의 별개 RNA 루프(들) 또는 별개 서열(들)이 제공된다. 가이드 분자는 별개 RNA 루프(들) 또는 별개 서열(들)에 결합할 수 있는 어댑터 단백질을 동원할 수 있는 별개 RNA 루프(들) 또는 별개 서열(들)의 삽입에 의해 Cas13 단백질과 충돌없이 연장될 수 있다. 변형된 가이드 및 CRISPR-Cas 복합체로 이펙터 도메인을 동원시키는데서 그들의 용도가 문헌 [Konermann (Nature 2015, 517(7536): 583-588)]에 제공된다. 특정 실시형태에서, 압타머는 포유동물에서 이량체화된 MS2 박테리아파지 외피 단백질에 선택적으로 결합하고, 가이드 분자, 예컨대 스템루프 및/또는 테트라루프 내에 도입되는 최소 헤어핀 압타머이다. 이들 실시형태에서, 아데노신 데아미나제 단백질은 MS2에 융합된다. 아데노신 데아미나제 단백질은 CRISPR-Cas 단백질 및 상응하는 가이드 RNA와 함께 공동-전달된다.

[0073] 본 명세서에서 사용되는 용어 "AD-기능화된 CRISPR 시스템"은 (a) 촉매적 불활성인 CRISPR-Cas 단백질, 보다 특히 Cas13 단백질; (b) 가이드 서열을 포함하는 가이드 분자; 및 (c) 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 포함하는 핵산 표적화 및 편집 시스템을 의미하고; 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 CRISPR-Cas 단백질 또는 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 이후에 그에 연결되도록 적합화되고; 여기서 가이드 서열은 표적 서열과 실질적으로 상동성이지만, 탈아미노화를 위해 표적화되는 A에 대응하는 쌍형성하지 않는 C를 포함하여서, 가이드 서열과 표적 서열에 의해 형성된 RNA 듀플렉스 내에 A-C 미스매치를 야기시킨다. 진핵생물 세포에서 적용을 위해, CRISPR-Cas 단백질 및/또는 아데노신 데아미나제는 바람직하게 NLS-태그화된다.

[0074] 일부 실시형태에서, 성분 (a), (b) 및 (c)는 리보핵단백질 복합체로서 세포로 전달된다. 리보핵단백질 복합체는 하나 이상의 지질 나노입자를 통해 전달될 수 있다.

[0075] 일부 실시형태에서, 성분 (a), (b) 및 (c)는 하나 이상의 RNA 분자, 예컨대 하나 이상의 가이드 RNA 및 CRISPR-Cas 단백질, 아데노신 데아미나제 단백질, 및 임의로는 어댑터 단백질을 코딩하는 하나 이상의 mRNA 분자로서 세포로 전달된다. RNA 분자는 하나 이상의 지질 나노입자를 통해서 전달될 수 있다.

[0076] 일부 실시형태에서, 성분 (a), (b) 및 (c)는 하나 이상의 DNA 분자로서 세포로 전달된다. 일부 실시형태에서, 하나 이상의 DNA 분자는 하나 이상의 벡터 예컨대 바이러스 벡터 (예를 들어, AAV) 내에 포함된다. 일부 실시형태에서, 하나 이상의 DNA 분자는 CRISPR-Cas 단백질, 가이드 분자, 및 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 발현하도록 작동적으로 구성된 하나 이상의 조절 구성요소를 포함하고, 임의로는 하나 이상의 조절 구성요소는 유도성 프로모터를 포함한다.

[0077] 일부 실시형태에서, CRISPR-Cas 단백질은 데드 Cas13이다. 일부 실시형태에서, 데드 Cas13은 HEPN 도메인 내에 하나 이상의 돌연변이를 포함하는 데드 Cas13a 단백질이다. 일부 실시형태에서, 데드 Cas13a는 렙토트리키아 웨이데이 (LwaCas13a) 내 R474A 및 R1046A에 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 데드 Cas13은 베르게이엘라 주헬컴 ATCC 43767로부터 기원된 Cas13b 단백질의 R116A, H121A, R1177A, H1182A 중 하나 이상, 또는 Cas13b 오솔로그의 이에 상응하는 아미노산 위치에서의 돌연변이 중 하나 이상을 포함하는 데드 Cas13b 단백질이다.

[0078] 일부 실시형태에서 가이드 분자는 RNA 서열 내에서 탈아미노화하려는 아데닌을 포함하는 표적 서열과 하이브리드화할 수 있어서 상기 아데닌 반대쪽에 쌍형성되지 않는 시토신을 포함하는 RNA 듀플렉스를 형성할 수 있다.

RNA 듀플렉스 형성 시에, 가이드 분자는 Cas13 단백질과의 복합체를 형성하고 관심 표적 RNA 서열에서 RNA 폴리뉴클레오티드에 결합하도록 복합체를 유도한다. AD-기능화된 CRISPR-Cas 시스템의 가이드의 양상에 대한 상세 설명은 하기 본 명세서에서 제공된다.

[0079] 일부 실시형태에서, 예를 들어 LawCas13의 정규 길이를 갖는 Cas13 가이드 RNA는 표적 DNA와 RNA 듀플렉스를 형성시키는데 사용된다. 일부 실시형태에서, 예를 들어 LawCas13a에 대한 정규 길이보다 긴 Cas13 가이드 분

자는 Cas13-가이드 RNA-표적 DNA 복합체의 외부를 포함한 표적 DNA와 RNA 듀플렉스를 형성하는데 사용된다.

[0080] 적어도 제1 디자인에서, AD-기능화된 CRISPR 시스템은 (a) CRISPR-Cas 단백질에 융합되거나 또는 연결된 아데노신 데아미나제로서, CRISPR-Cas 단백질은 촉매적 불활성인 것인 아데노 데아미나제, 및 (b) 가이드 서열과 표적 서열 간에 형성된 RNA 듀플렉스에 A-C 미스매치를 도입하기 위해 디자인된 가이드 서열을 포함하는 가이드 분자를 포함한다. 일부 실시형태에서, CRISPR-Cas 단백질 및/또는 아데노신 데아미나제는 N-말단 또는 C-말단 또는 둘 모두 상에서, NLS-태그화된다.

[0081] 적어도 제2 디자인에서, AD-기능화된 CRISPR 시스템은 (a) 촉매적 불활성의 CRISPR-Cas 단백질, (b) 가이드 서열 및 표적 서열 간에 형성된 RNA 듀플렉스에 A-C 미스매치를 도입시키기 위해 디자인된 가이드 서열, 및 어댑터 단백질 (예를 들어, MS2 외피 단백질 또는 PP7 외피 단백질)에 결합할 수 있는 압타머 서열 (예를 들어, MS2 RNA 모티프 또는 PP7 RNA 모티프)을 포함하는 가이드 분자, 및 (c) 어댑터 단백질에 융합 또는 연결된 아데노신 데아미나제로서, 압타머와 어댑터 단백질의 결합이 A-C 미스매치의 A에서 표적화된 탈아미노화를 위해 가이드 서열과 표적 서열 간에 형성된 RNA 듀플렉스로 아데노신 데아미나제를 동원하는 것인 아데노신 데아미나제를 포함한다. 일부 실시형태에서, 어댑터 단백질 및/또는 아데노신 데아미나제는 N-말단 또는 C-말단 또는 둘 모두 상에서, NLS-태그화된다. CRISPR-Cas 단백질은 또한 NLS-태그화될 수 있다.

[0082] 상이한 압타머 및 상응하는 어댑터 단백질의 사용은 또한 직교성 유전자 편집을 제공할 수 있게 한다. 아데노신 데아미나제를 직교성 유전자 편집/탈아미노화를 위해 시티딘 데아미나제와 조합하여 사용하는 일례에서, 상이한 유전자좌를 표적화하는 sgRNA는 각각 MS2-아데노신 데아미나제 및 PP7-시티딘 데아미나제 (또는 PP7-아데노신 데아미나제 및 MS2-시티딘 데아미나제)를 동원하기 위해 별개의 RNA 루프에 의해 변형되어서, 각각 관심 표적 유전자좌에서 A 또는 C의 직교성 탈아미노화를 일으킨다. PP7은 박테리오파지 슈도모나스의 RNA-결합 외피 단백질이다. MS2 처럼, 이것은 특별한 RNA 서열 및 2차 구조에 결합된다. PP7 RNA-인식 모티프는 MS2의 것과 상이하다. 결론적으로, PP7 및 MS2는 동시에 상이한 게놈 유전자좌에서 별개 효과를 매개하도록 다중복합화될 수 있다. 예를 들어, 유전자좌 A를 표적화하는 sgRNA는 MS2 루프에 의해 변형되어서, MS2-아데노신 데아미나제를 동원하는 한편, 유전자좌 B를 표적화하는 다른 sgRNA는 PP7 루프에 의해 변형되어, PP7-시티딘 데아미나제를 동원한다. 따라서 동일한 세포에서, 직교성, 유전자좌-특이적 변형을 인지한다. 이러한 원리는 다른 직교성 RNA-결합 단백질을 도입시키기 위해 확대될 수 있다.

[0083] 적어도 제3 디자인에서, AD-기능화된 CRISPR 시스템은 (a) CRISPR-Cas 단백질의 내부 루프 또는 비구조화된 영역에 삽입된 아데노신 데아미나제로서, CRISPR-Cas 단백질은 촉매적 불활성 또는 닉카제인 아데노신 데아미나제, 및 (b) 가이드 서열과 표적 서열 간에 형성된 RNA 듀플렉스에 A-C 미스매치를 도입시키기 위해 디자인된 가이드 서열을 포함하는 가이드 분자를 포함한다.

[0084] 아데노신 데아미나제의 삽입에 적합한 CRISPR-Cas 단백질 분할 부위는 결정 구조의 도움으로 확인될 수 있다. 오솔로그와 의도하는 CRISPR-Cas 단백질 간에 비교적 높은 정도의 상동성이 존재하면 오솔로그의 결정 구조를 사용할 수 있다.

[0085] 분할 위치는 영역 또는 루프 내에 위치될 수 있다. 바람직하게, 분할 위치는 아미노산 서열의 파괴가 구조적 특성 (예를 들어, 알파-나선 또는 베타-시트)의 부분 또는 완전 파괴를 초래하지 않는 곳에서 발생된다. 비구조화된 영역 (이들 영역이 결정 내에서 "동결"되기에 충분하게 구조화되지 않기 때문에 결정 구조에서 나타나지 않는 영역)이 종종 바람직한 선택사항이다. 비구조화된 영역 또는 외부 루프 내 위치는 정확하게 상기 제공된 번호일 필요는 없지만, 분할 위치가 루프 외부의 비구조화된 영역 내에 여전히 위치되는 한, 루프의 크기에 따라서, 상기 제공된 위치의 양쪽 측면에서 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 아미노산에 따라 다양할 수 있다.

[0086] 본 명세서에 기술된 AD-기능화된 CRISPR 시스템은 탈아미노화를 위해 RNA 폴리뉴클레오티드 서열 내 특이적 아데닌을 표적화하는데 사용될 수 있다. 예를 들어, 가이드 분자는 CRISPR-Cas 단백질과 복합체를 형성할 수 있고 관심 RNA 폴리뉴클레오티드 내 표적 RNA 서열에 결합하도록 복합체를 유도시킨다. 가이드 서열은 비-쌍형성 C를 갖도록 디자인되므로, 가이드 서열과 표적 서열 간에 형성된 RNA 듀플렉스는 A-C 미스매치를 포함하고, 이것은 아데노신 데아미나제가 비-쌍형성 C 반대쪽 A와 접촉하여 탈아미노화시키도록 유도시켜서, 이것을 이노신 (I)으로 전환시킨다. 이노신 (I) 염기가 C와 쌍형성하여 세포 과정에서 G처럼 기능하므로, 본 명세서에 기술된 A의 표적화된 탈아미노화는 비바람직한 G-A 및 C-T 돌연변이의 교정을 비롯하여, 바람직한 A-G 및 T-C 돌연변이의 수득에 유용하다.

[0087] 일부 실시형태에서, AD-기능화된 CRISPR 시스템은 시험관내에서 RNA 폴리뉴클레오티드 분자의 표적화된 탈아미노화에 사용된다. 일부 실시형태에서, AD-기능화된 CRISPR 시스템은 세포 내에서 DNA 분자의 표적화된 탈아미노화에 사용된다. 세포는 진핵생물 세포, 예컨대 동물 세포, 포유동물 세포, 인간, 또는 식물 세포일 수 있다.

[0088] 본 발명은 또한 AD-기능화된 CRISPR 시스템을 사용하여 표적화된 탈아미노화에 의해 질환을 치료 또는 예방하기 위한 방법에 관한 것으로서, A의 탈아미노화는 병원성 G→A 또는 C→T 점 돌연변이를 함유하는 전사물에 의해 야기되는 질환을 치유한다. 본 발명으로 치료 또는 예방할 수 있는 질환의 예는 암, 메이어-골린 증후군, 섹켈 증후군 4, 주버트 증후군 5, 레버 선천성 흑암시 10; 샤르코-마리-투스 질환, 2형; 샤르코-마리-투스 질환, 2형; 어셔 증후군, 2C형; 척수소뇌성 실조증 28; 척수소뇌성 실조증 28; 척수소뇌성 실조증 28; 긴 QT 증후군 2; 쇠그렌-라슨 증후군; 유전성 프룩토스뇨증; 유전성 프룩토스뇨증; 신경아세포종; 신경아세포종; 칼만 증후군 1; 칼만 증후군 1; 칼만 증후군 1; 이염색성 백질이영양증을 포함한다.

[0089] 특정한 실시형태에서, 따라서 본 발명은 요법에서 사용을 위한 위한 조성물을 포함한다. 이것은 방법을 생체내, 생체외, 또는 시험관내에서 수행할 수 있다는 것을 암시한다. 특정한 실시형태에서, 방법은 동물 또는 인간 신체의 치료 방법 또는 인간 세포의 생식선의 유전적 정체를 변형시키는 방법이 아니다. 특정 실시형태에서, 방법을 수행할 때, 표적 RNA는 인간 또는 동물 세포 내에 포함되지 않는다. 특정한 실시형태에서, 표적이 인간 또는 동물 표적일 때, 방법은 생체외 또는 시험관내에서 수행된다.

[0090] 본 발명은 또한 유전자의 비바람직한 활성을 녹아웃 또는 녹다운시키기 위한 방법에 관한 것으로서, 여기서 유전자의 전사물에서 A의 탈아미노화는 기능 상실을 야기시킨다. 예를 들어, 일 실시형태에서, AD-기능화된 CRISPR 시스템에 의한 표적화된 탈아미노화는 내생성 유전자 내 미성숙 정지 코돈을 야기시키는 넌센스 돌연변이를 초래할 수 있다. 이것은 내생성 유전자의 발현을 변경시킬 수 있고 편집된 세포 내에서 바람직한 특성을 야기시킬 수 있다. 다른 실시형태에서, AD-기능화된 CRISPR 시스템에 의한 표적화된 탈아미노화는 비보존성 미스센스 돌연변이를 초래하여 그 결과 내생성 유전자 내 상이한 아미노산 잔기를 코딩하게 한다. 이것은 발현되는 내생성 유전자의 기능을 변경시킬 수 있고 또한 편집된 세포에서 바람직한 특성을 야기시킬 수 있다.

[0091] 본 발명은 또한 AD-기능화된 CRISPR 시스템을 사용하여 표적화된 탈아미노화에 의해 수득된 변형된 세포, 또는 이의 자손에 관한 것으로서, 변형된 세포는 표적화된 탈아미노화 이전 상응하는 세포와 비교하여 관심 표적 RNA 서열에서 A 대신에 I 또는 G를 포함한다. 변형된 세포는 진핵생물 세포, 예컨대 동물 세포, 식물 세포, 포유동물 세포, 또는 인간 세포일 수 있다.

[0092] 일부 실시형태에서, 변형된 세포는 치료적 T 세포, 예컨대 CAR-T 요법에 적합한 T 세포이다. 변형은 제한없이, 면역 체크포인트 수용체 (예를 들어, PDA, CTLA4)의 감소된 발현, HLA 단백질 (예를 들어, B2M, HLA-A)의 감소된 발현, 및 내생성 TCR의 감소된 발현을 포함하여, 치료적 T 세포에서 하나 이상의 바람직한 특성을 일으킬 수 있다.

[0093] 일부 실시형태에서, 변형된 세포는 항체-생산 B 세포이다. 변형은 제한없이 증강된 항체 생산을 포함하여, B 세포에서 하나 이상의 바람직한 특성을 일으킬 수 있다.

[0094] 본 발명은 또한 변형된 인간 이외의 동물 또는 변형된 식물에 관한 것이다. 변형된 인간이외의 동물은 가축일 수 있다. 변형된 식물은 농작물일 수 있다.

[0095] 본 발명은 또한 본 명세서에 기술된 변형된 세포를 치료를 필요로 하는 환자에게 투여하는 단계를 포함하는, 세포 요법 방법에 관한 것으로서, 변형된 세포의 존재는 환자의 질환을 치유한다. 일 실시형태에서, 세포 요법을 위한 변형된 세포는 종양 세포를 인식 및/또는 공격할 수 있는 CAR-T 세포이다. 다른 실시형태에서, 세포 요법을 위한 변형된 세포는 줄기 세포, 예컨대 신경 줄기 세포, 중간엽 줄기 세포, 조혈 줄기 세포, 또는 iPSC 세포이다.

[0096] 본 발명은 추가로, 관심 표적 유전자좌에서 아데닌을 변형시키기에 적합한 조작된, 비천연 발생 시스템에 관한 것으로서, 가이드 서열, 또는 가이드 분자를 코딩하는 뉴클레오티드 서열을 포함하는 가이드 분자; CRISPR-Cas 단백질, 또는 CRISPR-Cas 단백질을 코딩하는 하나 이상의 뉴클레오티드 서열; 아데노신 데아미나제 단백질 또는 이의 촉매 도메인, 또는 이를 코딩하는 하나 이상의 뉴클레오티드 서열을 포함하고; 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 CRISPR-Cas 단백질 또는 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 이후 그에 연결되도록 적합화되고; 가이드 서열은 관심 RNA 폴리뉴클레오티드 내에 아데닌을 포함하

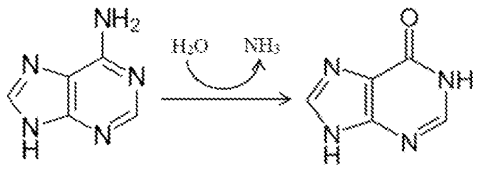
는 표적 서열과 하이브리드화될 수 있지만, 아데닌에 상응하는 위치에 시토신을 포함한다.

[0097] 본 발명은 추가로 관심 표적 유전자좌 내에 아데닌을 변형시키기에 적합한 조작된, 비천연 발생 벡터 시스템에 관한 것으로서, 가이드 서열을 포함하는 가이드 분자를 코딩하는 하나 이상의 뉴클레오티드 서열에 작동적으로 연결된 제1 조절 구성요소; CRISPR-Cas 단백질을 코딩하는 뉴클레오티드 서열에 작동적으로 연결된 제1 조절 구성요소; 및 임의로는 제1 또는 제2 조절 구성요소의 제어 하에 있거나 또는 제3 조절 구성요소에 작동적으로 연결된 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열을 포함하는 하나 이상의 벡터를 포함하고; 아데노신 데아미나제 단백질 또는 이의 촉매 도메인을 코딩하는 뉴클레오티드 서열이 제3 조절 구성요소에 작동적으로 연결되면, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 발현 이후에 가이드 분자 또는 CRISPR-Cas 단백질에 연결되도록 적합화되고; 가이드 서열은 표적 유전자좌 내 아데닌을 포함하는 표적 서열과 하이브리드화될 수 있지만, 아데닌에 상응하는 위치에 시토신을 포함하고; 성분 (a), (b) 및 (c)는 시스템의 동일하거나 또는 상이한 벡터에 위치된다.

[0098] 본 발명은 추가로 본 명세서에 기술된 조작된, 비천연 발생 시스템 또는 벡터 시스템을 포함하는 시험관내, 생체외 또는 생체내 숙주 세포 또는 세포주 또는 이의 자손에 관한 것이다. 숙주 세포는 진핵생물 세포, 예컨대 동물 세포, 식물 세포, 포유동물 세포, 또는 인간 세포일 수 있다.

[0099] **아데노신 데아미나제**

[0100] 본 명세서에서 사용되는 용어 "아데노신 데아미나제" 또는 "아데노신 데아미나제 단백질"은 이하에 표시된 바와 같이, 아데닌 (또는 분자의 아데닌 모이어티)을 하이포잔틴 (또는 분자의 하이포잔틴 모이어티)으로 전환시키는 가수분해 탈아미노화 반응을 촉매할 수 있는 단백질, 폴리펩티드, 또는 단백질 또는 폴리펩티드의 하나 이상의 기능성 도메인(들)에 관한 것이다. 일부 실시형태에서, 아데닌-함유 분자는 아데노신(A)이고, 하이포잔틴-함유 분자는 이노신(I)이다. 아데닌-함유 분자는 데옥시리보핵산 (DNA) 또는 리보핵산 (RNA)일 수 있다.



아데닌 하이포잔틴

[0101]

[0102] 본 개시내용에 따르면, 본 개시내용과 관련하여 사용될 수 있는 아데노신 데아미나제는 RNA 상에서 작용하는 아데노신 데아미나제 (ADAR)로서 알려진 효소 패밀리의 구성원, tRNA 상에서 작용하는 아데노신 데아미나제 (ADAT)로서 알려진 효소 패밀리의 구성원, 및 다른 아데노신 데아미나제 도메인-함유 (ADAD) 패밀리 구성원을 포함하지만, 이들로 제한되지 않는다. 본 개시내용에 따르면, 아데노신 데아미나제는 RNA/DNA 및 RNA 듀플렉스에서 아데닌을 표적화할 수 있다. 실제로, Zheng 등 (Nucleic Acids Res. 2017, 45(6): 3369-3377)은 ADAR이 RNA/DNA 및 RNA/RNA 듀플렉스 상에서 아데노신의 이노신 편집 반응을 수행할 수 있다는 것을 입증한다.
특정 실시형태에서, 아데노신 데아미나제는 본 명세서에서 이하에 상술하는 바와 같이 RNA/DNAn RNA 듀플렉스에서 DNA를 편집하는 그의 능력을 증가시키도록 변형되었다.

[0103] 일부 실시형태에서, 아데노신 데아미나제는 포유동물, 조류, 개구리, 오징어, 어류, 파리 및 벌레를 포함하지만, 제한없이, 1종 이상의 후생동물 종으로부터 유래된다. 일부 실시형태에서, 아데노신 데아미나제는 인간, 오징어 또는 초파리 아데노신 데아미나제이다.

[0104] 일부 실시형태에서, 아데노신 데아미나제는 hADAR1, hADAR2, hADAR3을 비롯한 인간 ADAR이다. 일부 실시형태에서, 아데노신 데아미나제는 ADR-1 및 ADR-2를 비롯한 예쁜꼬마선충 (*Caenorhabditis elegans*) *ADAR* 단백질이다. 일부 실시형태에서, 아데노신 데아미나제는 dAdar을 비롯한 초파리 ADAR 단백질이다. 일부 실시형태에서, 아데노신 데아미나제는 sqADAR2a 및 sqADAR2b를 비롯한 오징어 롤리고 페알레이 (*Loligo pealeii*) ADAR 단백질이다. 일부 실시형태에서, 아데노신 데아미나제는 인간 ADAT 단백질이다. 일부 실시형태에서, 아데노신 데아미나제는 초파리 ADAT 단백질이다. 일부 실시형태에서, 아데노신 데아미나제는 TENR(hADAD1) 및 TENRL (hADAD2)를 비롯한 인간 ADAD 단백질이다.

[0105] 일부 실시형태에서, 아데노신 데아미나제 단백질은 듀플렉스 핵산 기질 내 하나 이상의 표적 아데노신 잔기(들)를 인식하고 이를 이노신 잔기(들)로 전환시킨다. 일부 실시형태에서, 듀플렉스 핵산 기질은 RNA-DNA 하이브리드 듀플렉스이다. 일부 실시형태에서, 아데노신 데아미나제 단백질은 듀플렉스 기질 상에서 결합창

을 인식한다. 일부 실시형태에서, 결합창은 적어도 하나의 표적 아데노신 잔기(들)를 함유한다. 일부 실시형태에서, 결합창은 약 3 bp 내지 약 100 bp 범위이다. 일부 실시형태에서, 결합창은 약 5 bp 내지 약 50 bp범위이다. 일부 실시형태에서, 결합창은 약 10 bp 내지 약 30 bp 범위이다. 일부 실시형태에서, 결합창은 약 1 bp, 2 bp, 3 bp, 5 bp, 7 bp, 10 bp, 15 bp, 20 bp, 25 bp, 30 bp, 40 bp, 45 bp, 50 bp, 55 bp, 60 bp, 65 bp, 70 bp, 75 bp, 80 bp, 85 bp, 90 bp, 95 bp 또는 100 bp 범위이다.

[0106] 일부 실시형태에서, 아데노신 데아미나제 단백질은 하나 이상의 데아미나제 도메인을 포함한다. 특정 이론에 의해 구속되는 것으로 의도하지는 않지만, 데아미나제 도메인은 듀플렉스 핵산 기질에 함유된 하나 이상의 표적 아데노신(A) 잔기(들)를 인식하고, 이를 이노신(I) 잔기(들)로 전환시키는 작용을 하는 것으로 고려된다.

일부 실시형태에서, 데아미나제 도메인은 활성 중심을 포함한다. 일부 실시형태에서, 활성 중심은 아연 이온을 포함한다. 일부 실시형태에서, A>I 편집 과정 동안, 표적 아데노신 잔기에서 염기쌍이 붕괴되며, 표적 아데노신 잔기는 아데노신 데아미나제에 의해 접근 가능하게 되도록 이중 나선 밖으로 "플립(flipped)"된다.

일부 실시형태에서, 활성 중심 내 또는 근처의 아미노산 잔기는 표적 아데노신 잔기에 대해 하나 이상의 뉴클레오티드(들) 5'과 상호작용한다. 일부 실시형태에서, 활성 중심 내 또는 근처의 아미노산 잔기는 표적 아데노신 잔기에 대해 하나 이상의 뉴클레오티드(들) 3'과 상호작용한다. 일부 실시형태에서, 활성 중심 내 또는 근처의 아미노산 잔기는 마주보는 가닥 상의 표적 아데노신 잔기에 상보성인 뉴클레오티드와 추가로 상호작용한다. 일부 실시형태에서, 아미노산 잔기는 뉴클레오티드의 2' 히드록실기와 수소 결합을 형성한다.

[0107] 일부 실시형태에서, 아데노신 데아미나제는 인간 ADAR2 완전 단백질(hADAR2) 또는 이의 데아미나제 도메인(hADAR2-D)을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 hADAR2 또는 hADAR2-D에 상동성인 ADAR 패밀리 구성원이다.

[0108] 특히, 일부 실시형태에서, 상동성 ADAR 단백질은 인간 ADAR1 (hADAR1) 또는 이의 데아미나제 도메인 (hADAR1-D)이다. 일부 실시형태에서, hADAR1-D의 글리신1007은 글리신487 hADAR2-D에 상응하고, hADAR1-D의 글루탐산1008은 hADAR2-D의 글루탐산488에 상응한다.

[0109] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D의 야생형 아미노산 서열을 포함한다. 일부 실시형태에서, hADAR2-D의 편집 효율, 및/또는 기질 편집 선호도가 특정 필요에 따라 변화되도록, 아데노신 데아미나제는 hADAR2-D 서열에서 하나 이상의 돌연변이를 포함한다.

[0110] hADAR1 및 hADAR2 단백질의 일정 돌연변이는 문헌 [Kuttan *et al.*, Proc Natl Acad Sci U S A. (2012) 109(48):E3295-304], [Want *et al.* ACS Chem Biol. (2015) 10(11):2512-9] 및 [Zheng *et al.* Nucleic Acids Res. (2017) 45(6):3369-337]에 기술되어 있으며, 이들 각각은 본 명세서에 그 전문이 참조로 편입된다.

[0111] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 글리신336 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 336번 위치의 글리신 잔기는 아스파르트산 잔기로 치환된다 (G336D).

[0112] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 글리신487, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 487번 위치에서 글리신 잔기는 상대적으로 작은 측쇄를 갖는 비극성 아미노산 잔기로 치환된다. 예를 들어, 일부 실시형태에서, 487번 위치에서 글리신 잔기는 알라닌 잔기로 치환된다 (G487A). 일부 실시형태에서, 487번 위치에서 글리신 잔기는 발린 잔기로 치환된다 (G487V). 일부 실시형태에서, 487번 위치에서 글리신 잔기는 상대적으로 큰 측쇄를 갖는 아미노산 잔기로 치환된다. 일부 실시형태에서, 487번 위치에서 글리신 잔기는 아르기닌 잔기로 치환된다 (G487R).

일부 실시형태에서, 487번 위치에서 글리신 잔기는 리신 잔기로 치환된다 (G487K). 일부 실시형태에서, 487번 위치에서 글리신 잔기는 트립토판 잔기로 치환된다 (G487W). 일부 실시형태에서, 487번 위치에서 글리신 잔기는 티로신 잔기로 치환된다 (G487Y).

[0113] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 글루탐산488, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 글루타민 잔기로 치환된다 (E488Q). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 히스티딘 잔기로 치환된다 (E488H). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 아르기닌 잔기로 치환된다 (E488R). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 리신 잔기로 치환된다 (E488K). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 아스파라긴 잔기로 치환된다 (E488N). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 알라닌 잔기로 치환된다 (E488A). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 메티

오닌 잔기로 치환된다 (E488M). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 세린 잔기로 치환된다 (E488S). 일부 실시형태에서, 488번 위상응하는치에서 글루탐산 잔기는 페닐알라닌 잔기로 치환된다 (E488F). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 리신 잔기로 치환된다 (E488L). 일부 실시형태에서, 488번 위치에서 글루탐산 잔기는 트립토판 잔기로 치환된다 (E488W).

[0114] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 트레오닌490, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 시스테인 잔기로 치환된다 (T490C). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 세린 잔기로 치환된다 (T490S). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 알라닌 잔기로 치환된다 (T490A). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 페닐알라닌 잔기로 치환된다 (T490F). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 티로신 잔기로 치환된다 (T490Y). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 세린 잔기로 치환된다 (T490R). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 알라닌 잔기로 치환된다 (T490K). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 페닐알라닌 잔기로 치환된다 (T490P). 일부 실시형태에서, 490번 위치에서 트레오닌 잔기는 티로신 잔기로 치환된다 (T490E).

[0115] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 발린493, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 493번 위치에서 발린 잔기는 알라닌 잔기로 치환된다 (V493A). 일부 실시형태에서, 493번 위치에서 발린 잔기는 세린 잔기로 치환된다 (V493S). 일부 실시형태에서, 493번 위치에서 발린 잔기는 트레오닌 잔기로 치환된다 (V493T). 일부 실시형태에서, 493번 위치에서 발린 잔기는 아르기닌 잔기로 치환된다 (V493R). 일부 실시형태에서, 493번 위치에서 발린 잔기는 아스파르트산 잔기로 치환된다 (V493D). 일부 실시형태에서, 493번 위치에서 발린 잔기는 프롤린 잔기로 치환된다 (V493P). 일부 실시형태에서, 493번 위치에서 발린 잔기는 글리신 잔기로 치환된다 (V493G).

[0116] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 알라닌589, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 589번 위치에서 알라닌 잔기는 발린 잔기로 치환된다 (A589V).

[0117] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아스파라긴597, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 리신 잔기로 치환된다 (N597K). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 597번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 아르기닌 잔기로 치환된다 (N597R). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 597번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 알라닌 잔기로 치환된다 (N597A). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 597번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 글루탐산 잔기로 치환된다 (N597E). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 597번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 히스티딘 잔기로 치환된다 (N597H). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 597번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 글리신 잔기로 치환된다 (N597G). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 597번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 티로신 잔기로 치환된다 (N597Y). 일부 실시형태에서, 597번 위치에서 아스파라긴 잔기는 페닐알라닌 잔기로 대체된다 (N597F).

[0118] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 세린599, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 599번 위치에서 세린 잔기는 트레오닌 잔기로 치환된다 (S599T).

[0119] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아스파라긴613, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 613번 위치에서 아스파라긴 잔기는 리신 잔기 (N613K)로 치환된다. 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 613번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 613번 위치에서 아스파라긴 잔기는 아르기닌 잔기로 치환된다 (N613R). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 613번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 613번 위치에

서 아스파라긴 잔기는 알라닌 잔기로 치환된다 (N613A). 일부 실시형태에서, 아데노신 데아미나제는 야생형 서열에서 아스파라긴 잔기를 갖는 아미노산 서열의 613번 위치에 돌연변이를 포함한다. 일부 실시형태에서, 613번 위치에서 아스파라긴 잔기는 글루탐산 잔기로 치환된다 (N613E).

[0120] 일부 실시형태에서, 편집 효율을 개선시키기 위해서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, 돌연변이: G336D, G487A, G487V, E488Q, E488H, E488R, E488N, E488A, E488S, E488M, T490C, T490S, V493T, V493S, V493A, V493R, V493D, V493P, V493G, N597K, N597R, N597A, N597E, N597H, N597G, N597Y, A589V, S599T, N613K, N613R, N613A, N613E, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다.

[0121] 일부 실시형태에서, 편집 효율을 감소시키기 위해서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, 돌연변이: E488F, E488L, E488W, T490A, T490F, T490Y, T490R, T490K, T490P, T490E, N597F, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다. 특정한 실시형태에서, 오프-표적 효과를 감소시키기 위해서 효율이 감소된 아데노신 데아미나제 효소를 사용하는 것이 흥미로울 수 있다.

[0122] 일부 실시형태에서, 오프-표적 효과를 감소시키기 위해서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, R348, V351, T375, K376, E396, C451, R455, N473, R474, K475, R477, R481, S486, E488, T490, S495, R510에서의 돌연변이, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 E488, 및 R348, V351, T375, K376, E396, C451, R455, N473, R474, K475, R477, R481, S486, T490, S495, R510으로부터 선택된 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 T375, 및 임의로는 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 N473, 및 임의로는 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 V351, 및 임의로는 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 E488 및 T375, 및 임의로는 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 E488 및 N473, 및 임의로는 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 E488 및 V351, 및 임의로는 하나 이상의 추가 위치에 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 E488, 및 T375, N473, 및 V351 중 하나 이상에 돌연변이를 포함한다.

[0123] 일부 실시형태에서, 오프-표적 효과를 감소시키기 위해서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, R348E, V351L, T375G, T375S, R455G, R455S, R455E, N473D, R474E, K475Q, R477E, R481E, S486T, E488Q, T490A, T490S, S495T, 및 R510E로부터 선택되는 돌연변이, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 E488Q, 및 R348E, V351L, T375G, T375S, R455G, R455S, R455E, N473D, R474E, K475Q, R477E, R481E, S486T, T490A, T490S, S495T 및 R510E로부터 선택되는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 T375G 또는 T375S, 및 임의로는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 N473D, 및 임의로는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 V351L, 및 임의로는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 E488Q, 및 T375G 또는 T375G, 및 임의로는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 E488Q 및 N473D, 및 임의로는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 E488Q 및 V351L, 및 임의로는 하나 이상의 추가 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 돌연변이 E488Q, 및 T375G/S, N473D 및 V351L 중 하나 이상을 포함한다.

[0124] 듀플렉스 RNA에 결합된 인간 ADAR2 데아미나제 도메인의 결정 구조는 변형 부위의 5' 측면 상에서 RNA에 결합하는 단백질 루프를 밝혀준다. 이러한 5' 결합 루프는 ADAR 패밀리 구성원 간 기질 특이성 편차에 대한 하나의 원인 인자이다. [Wang et al., Nucleic Acids Res., 44(20):9872-9880 (2016)]을 참조하고, 이의 내용은 그 전문이 참조로 본 명세서에 편입된다. 또한, ADAR2-특이적RNA-결합 루프가 효소 활성 부위 근처에서 확인되었다. [Mathews et al., Nat. Struct. Mol. Biol., 23(5):426-33 (2016)]을 참조하고, 이의 내용은 그 전문이 참조로 본 명세서에 편입된다. 일부 실시형태에서, 아데노신 데아미나제는 편집 특이성 및/또는 효율을 개선시키도록 RNA 결합 루프 내에 하나 이상의 돌연변이를 포함한다.

[0125] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 알라닌454, 또는 상동성 ADAR 단백질 내

상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 454의 알라닌은 세린 잔기로 치환된다 (A454S). 일부 실시형태에서, 위치 454의 알라닌 잔기는 시스테인 잔기로 치환된다 (A454C). 일부 실시형태에서, 위치 454의 알라닌 잔기는 아스파르트산 잔기로 치환된다 (A454D).

[0126] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌455, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 455의 아르기닌 잔기는 알라닌 잔기로 치환된다 (R455A). 일부 실시형태에서, 위치 455의 아르기닌 잔기는 발린 잔기로 치환된다 (R455V). 일부 실시형태에서, 위치 455의 아르기닌 잔기는 히스티딘 잔기로 치환된다 (R455H). 일부 실시형태에서, 위치 455의 아르기닌 잔기는 글리신 잔기로 치환된다 (R455G). 일부 실시형태에서, 위치 455의 아르기닌 잔기는 세린 잔기로 치환된다 (R455S). 일부 실시형태에서, 위치 455의 아르기닌 잔기는 글루탐산 잔기로 치환된다 (R455E).

[0127] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 이소류신456, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 456의 이소류신 잔기는 발린 잔기로 치환된다 (I456V). 일부 실시형태에서, 위치 456의 이소류신 잔기는 류신 잔기로 치환된다 (I456L). 일부 실시형태에서, 위치 456의 이소류신 잔기는 아스파르트산 잔기로 치환된다 (I456D).

[0128] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 페닐알라닌457, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 457의 페닐알라닌 잔기는 티로신 잔기로 치환된다 (F457Y). 일부 실시형태에서, 위치 457의 페닐알라닌 잔기는 아르기닌 잔기로 치환된다 (F457R). 일부 실시형태에서, 위치 457의 페닐알라닌 잔기는 글루탐산 잔기로 치환된다 (F457E).

[0129] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 세린458, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 458의 세린 잔기는 발린 잔기로 치환된다 (S458V). 일부 실시형태에서, 위치 458의 세린 잔기는 페닐알라닌 잔기로 치환된다 (S458F). 일부 실시형태에서, 위치 458의 세린 잔기는 프롤린 잔기로 치환된다 (S458P).

[0130] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 프롤린459, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 459의 프롤린 잔기는 시스테인 잔기로 치환된다 (P459C). 일부 실시형태에서, 위치 459의 프롤린 잔기는 히스티딘 잔기로 치환된다 (P459H). 일부 실시형태에서, 위치 459의 프롤린 잔기는 트립토판 잔기로 치환된다 (P459W).

[0131] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 히스티딘460, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 460의 히스티딘 잔기는 아르기닌 잔기로 치환된다 (H460R). 일부 실시형태에서, 위치 460의 히스티딘 잔기는 이소류신 잔기로 치환된다 (H460I). 일부 실시형태에서, 위치 460의 히스티딘 잔기는 프롤린 잔기로 치환된다 (H460P).

[0132] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 프롤린462, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 462의 프롤린 잔기는 세린 잔기로 치환된다 (P462S). 일부 실시형태에서, 위치 462의 프롤린 잔기는 트립토판 잔기로 치환된다 (P462W). 일부 실시형태에서, 위치 462의 프롤린 잔기는 글루탐산 잔기로 치환된다 (P462E).

[0133] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아스파르트산469, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 469의 아스파르트산 잔기는 글루타민 잔기로 치환된다 (D469Q). 일부 실시형태에서, 위치 469의 아스파르트산 잔기는 세린 잔기로 치환된다 (D469S). 일부 실시형태에서, 위치 469의 아스파르트산 잔기는 티로신 잔기로 치환된다 (D469Y).

[0134] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌470, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 470의 아르기닌 잔기는 알라닌 잔기로 치환된다 (R470A). 일부 실시형태에서, 위치 470의 아르기닌 잔기는 이소류신 잔기로 치환된다 (R470I). 일부 실시형태에서, 위치 470의 아르기닌 잔기는 아스파르트산 잔기로 치환된다 (R470D).

[0135] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 히스티딘471, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 471의 히스티딘 잔기는 리신 잔기로 치환된다 (H471K). 일부 실시형태에서, 위치 471의 히스티딘 잔기는 트레오닌 잔기로 치환된다 (H471T). 일부 실시형태에서, 위치 471의 히스티딘 잔기는 발린 잔기로 치환된다 (H471V).

[0136] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 프롤린472, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 472의 프롤린 잔기는 리신 잔기로 치환된다 (P472K). 일부 실시형태에서, 위치 472의 프롤린 잔기는 트레오닌 잔기로 치환된다 (P472T). 일부 실시형태에서, 위치 472의 프롤린 잔기는 아스파르트산 잔기로 치환된다 (P472D).

[0137] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아스파라긴473, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 473의 아스파라긴 잔기는 아르기닌 잔기로 치환된다 (N473R). 일부 실시형태에서, 위치 473의 아스파라긴 잔기는 트립토판 잔기로 치환된다 (N473W). 일부 실시형태에서, 위치 473의 아스파라긴 잔기는 프롤린 잔기로 치환된다 (N473P). 일부 실시형태에서, 위치 473의 아스파라긴 잔기는 아스파르트산 잔기로 치환된다 (N473D).

[0138] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌474, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 474의 아르기닌 잔기는 리신 잔기로 치환된다 (R474K). 일부 실시형태에서, 위치 474의 아르기닌 잔기는 글리신 잔기로 치환된다 (R474G). 일부 실시형태에서, 위치 474의 아르기닌 잔기는 아스파르트산 잔기로 치환된다 (R474D). 일부 실시형태에서, 위치 474의 아르기닌 잔기는 글루탐산 잔기로 치환된다 (R474E).

[0139] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 리신475, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 475의 리신 잔기는 글루타민 잔기로 치환된다 (K475Q). 일부 실시형태에서, 위치 475의 리신 잔기는 아스파라긴 잔기로 치환된다 (K475N). 일부 실시형태에서, 위치 475의 리신 잔기는 아스파르트산 잔기로 치환된다 (K475D).

[0140] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 알라닌476, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 476의 알라닌 잔기는 세린 잔기로 치환된다 (A476S). 일부 실시형태에서, 위치 476의 알라닌 잔기는 아르기닌 잔기로 치환된다 (A476R). 일부 실시형태에서, 위치 476의 알라닌 잔기는 글루탐산 잔기로 치환된다 (A476E).

[0141] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌477, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 477의 아르기닌 잔기는 리신 잔기로 치환된다 (R477K). 일부 실시형태에서, 위치 477의 아르기닌 잔기는 트레오닌 잔기로 치환된다 (R477T). 일부 실시형태에서, 위치 477의 아르기닌 잔기는 페닐알라닌 잔기로 치환된다 (R477F). 일부 실시형태에서, 위치 474의 아르기닌 잔기는 글루탐산 잔기로 치환된다 (R477E).

[0142] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 글리신478, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 478의 글리신 잔기는 알라닌 잔기로 치환된다 (G478A). 일부 실시형태에서, 위치 478의 글리신 잔기는 아르기닌 잔기로 치환된다 (G478R). 일부 실시형태에서, 위치 478의 글리신 잔기는 티로신 잔기로 치환된다 (G478Y).

[0143] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 글루타민479, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 479의 글루타민 잔기는 아스파라긴 잔기로 치환된다 (Q479N). 일부 실시형태에서, 위치 479의 글루타민 잔기는 세린 잔기로 치환된다 (Q479S). 일부 실시형태에서, 위치 479의 글루타민 잔기는 프롤린 잔기로 치환된다 (Q479P).

[0144] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌348, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 348의 아르기닌 잔기느느 알라닌 잔기로 치환된다 (R348A). 일부 실시형태에서, 위치 348의 아르기닌 잔기는 글루탐산 잔기로 치환된다 (R348E).

[0145] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 발린351, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 351의 발린 잔기는 류신 잔기로 치환된다 (V351L).

[0146] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 트레오닌375, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 375의 트레오닌 잔기는 글리신 잔기로 치환된다 (T375G). 일부 실시형태에서, 위치 375의 트레오닌 잔기는 세린 잔기로 치환된다 (T375S).

[0147] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌481, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 481의 아르기닌 잔기는 글루탐산 잔기

로 치환된다 (R481E).

[0148] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 세린486, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 486의 세린 잔기는 트레오닌 잔기로 치환된다 (S486T).

[0149] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 트레오닌490, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 490의 트레오닌 잔기는 알라닌 잔기로 치환된다 (T490A). 일부 실시형태에서, 위치 490의 트레오닌 잔기는 세린 잔기로 치환된다 (T490S).

[0150] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 세린495, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 495의 세린 잔기는 트레오닌 잔기로 치환된다 (S495T).

[0151] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 아르기닌510, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 510의 아르기닌 잔기는 글루타민 잔기로 치환된다 (R510Q). 일부 실시형태에서, 위치 510의 아르기닌 잔기는 알라닌 잔기로 치환된다 (R510A). 일부 실시형태에서, 위치 510의 아르기닌 잔기는 글루탐산 잔기로 치환된다 (R510E).

[0152] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 글리신593, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 593의 글리신 잔기는 알라닌 잔기로 치환된다 (G593A). 일부 실시형태에서, 위치 593의 글리신 잔기는 글루탐산 잔기로 치환된다 (G593E).

[0153] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 리신594, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 위치 594의 리신 잔기는 알라닌 잔기로 치환된다 (K594A).

[0154] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 위치 A454, R455, I456, F457, S458, P459, H460, P462, D469, R470, H471, P472, N473, R474, K475, A476, R477, G478, Q479, R348, R510, G593, K594, 또는 상동성 ADAR 단백질 내 상응하는 위치 중 임의의 하나 이상에 돌연변이를 포함한다.

[0155] 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D 아미노산 서열의 돌연변이 A454S, A454C, A454D, R455A, R455V, R455H, I456V, I456L, I456D, F457Y, F457R, F457E, S458V, S458F, S458P, P459C, P459H, P459W, H460R, H460I, H460P, P462S, P462W, P462E, D469Q, D469S, D469Y, R470A, R470I, R470D, H471K, H471T, H471V, P472K, P472T, P472D, N473R, N473W, N473P, R474K, R474G, R474D, K475Q, K475N, K475D, A476S, A476R, A476E, R477K, R477T, R477F, G478A, G478R, G478Y, Q479N, Q479S, Q479P, R348A, R510Q, R510A, G593A, G593E, K594A, 또는 상동성 ADAR 단백질 내 상응하는 위치 내 돌연변이 중 임의의 하나 이상을 포함한다.

[0156] 일정 실시형태에서, 아데노신 데아미나제는 활성이 시티딘 데아미나제로 전환되도록 돌연변이된다. 따라서, 일부 실시형태에서, 아데노신 데아미나제는 E396, C451, V351, R455, T375, K376? S486, Q488, R510, K594, R348, G593, S397, H443, L444, Y445, F442, E438, T448, A353, V355, T339, P539, T339, P539, V525 I520, P462 및 N579로부터 선택되는 위치에 하나 이상의 돌연변이를 포함한다. 특정한 실시형태에서, 아데노신 데아미나제는 V351, L444, V355, V525 및 I520으로부터 선택되는 위치에 하나 이상의 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, E488, V351, S486, T375, S370, P462, N597에서의 돌연변이, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다. 일부 실시형태에서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, 돌연변이: E488Q, V351G, S486A, T375S, S370C, P462A, N597I, 및 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다. 이들 변형은 상기 표적 RNA에 (a) 촉매적 불활성 (데드) Cas13 단백질; (b) 직접 반복부 서열에 연결된 가이드 서열을 포함하는 가이드 분자; 및 (c) 시토신 데아미나제 활성을 보장하기 위해서 상기 본 명세서에 기술된 바와 같이 변형된 아데노신 데아미나제인 시토신 데아미나제 단백질 또는 이의 촉매 도메인을 전달하는 단계를 포함하는, 관심 표적 RNA 내 시토신을 변형시키기 위한 방법에서의 사용을 위해 흥미로우며, 여기서 상기 시토신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 데드 Cas13 단백질 또는 상기 가이드 분자에 공유적으로 또는 비공유적으로 연결되거나 또는 전달 후에 이에 연결되도록 적합화되고, 가이드 분자는 상기 데드 Cas13 단백질과 복합체를 형성하고 상기 관심 표적 RNA 서열에 결합하도록 상기 복합체를 유도시키며, 상기 가이드 서열은 상기 시토신을 포함하는 표적 서열과 하이브리드화하여 RNA 듀플렉스를

형성할 수 있고, 상기 가이드 서열은 상기 시토신에 상응하는 위치에 비-쌍형성 아데닌 또는 우라실을 포함하여서 형성된 RNA 듀플렉스 내에 C-A/U 미스매치가 생성되고, 상기 시토신 데아미나제 단백질 또는 이의 촉매 도메인은 상기 RNA 듀플렉스 내 상기 시토신을 탈아미노화시킨다. 이들 방법은 T→C 또는 A→G 점 돌연변이 또는 병원성 SNP에 의해 초래되는 질환의 치료 또는 예방에서 흥미롭다.

[0157] 일정 실시형태에서, 오프-표적 변형의 감소 및 편집의 개선은 gRNA의 화학적 변형에 의해 달성된다. 문헌 [Vogel et al. (2014), Angew Chem Int Ed, 53:6267-6271, doi:10.1002/anie.201402634 (본 명세서에 전문이 참고로 편입됨)]에 예시된 바와 같이 화학적으로 변형된 gRNA는 오프-표적 활성을 감소시키고, 온-표적 효율을 개선시킨다. 2'-O-메틸 및 포스포티오에이트 변형된 가이드 RNA는 일반적으로 세포에서 편집 효율을 개선시킨다.

[0158] ADAR은 편집된 A의 양쪽 측면 상의 이웃 뉴클레오티드에 대한 선호도를 입증한 것이 알려져 있다 (www.nature.com/nsmb/journal/v23/n5/full/nsmb.3203.html, Matthews et al. (2017), Nature Structural Mol Biol, 23(5): 426-433, 본 명세서에 전문이 참고로 편입됨). 따라서, 일정 실시형태에서, gRNA, 표적 및/또는 ADAR은 모티프 선호도에 대해 최적으로 선택된다.

[0159] 의도적 미스매치는 비선호 모티프의 편집을 허용하는 것이 시험관내에서 입증되었다 (https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gku272; Schneider et al (2014), Nucleic Acid Res, 42(10):e87); Fukuda et al. (2017), Scientific Reports, 7, doi:10.1038/srep41478, 본 명세서에 전문이 참고로 편입됨). 따라서, 일정 실시형태에서, 비선호 5' 또는 3' 이웃 염기 상에서 RNA 편집 효율을 증강시키기 위해, 이웃하는 염기에서의 의도적 미스매치가 도입된다.

[0160] 결과는 ADAR 데아미나제 도메인의 표적 창에서 C의 반대쪽 A는 다른 염기보다 우선적으로 편집될 수 있다는 것을 시사한다. 추가적으로, 표적화된 염기의 소수 염기 내에서 U와 염기쌍 형성하는 A의 염기는 Cas13b-ADAR 융합에 의해 낮은 편집 수준을 나타내는데, 이는 다중 A를 편집하는 효소에 대해 유연성이 있다는 것을 시사한다. 예를 들어, 도 18을 참조한다. 이들 두 관찰은 Cas13b-ADAR 융합의 활성 창에서의 다중 A는 C와 편집될 모든 A를 미스매칭시킴으로써 편집에 대해 구체화될 수 있다는 것을 시사한다. 따라서, 일정 실시형태에서, 활성 창에서 다중 A:C 미스매치는 다중 A:I 편집을 생성하도록 디자인된다. 일정 실시형태에서, 활성 창에서 잠재적 오프-표적 편집을 억제하기 위해, 비-표적 A는 A 또는 G와 쌍형성한다.

[0161] 용어 "편집 특이성" 및 "편집 선호도"는 듀플렉스 기질 내 특정 아데노신 부위에서 A>I 편집 정도를 지칭하기 위해 본 명세서에서 상호교환 가능하게 사용된다. 일부 실시형태에서, 기질 편집 선호도는 표적 아데노신 잔기의 5' 가장 가까운 이웃 및/또는 3' 가장 가까운 이웃에 의해 결정된다. 일부 실시형태에서, 아데노신 데아미나제는 U>A>C>G(">"는 더 큰 선호도를 나타냄)으로 순위매겨지는, 기질의 5' 가장 가까운 이웃에 대한 선호도를 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 G>C~A>U(">"는 더 큰 선호도를 나타냄; "~"는 유사한 선호도를 나타냄)로서 순위매겨지는, 기질의 3' 가장 가까운 이웃에 대해 선호도를 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 G>C>U~A (">"는 더 큰 선호도를 나타냄; "~"는 유사한 선호도를 나타냄)로서 순위매겨지는, 기질의 3' 가장 가까운 이웃에 대해 선호도를 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 G>C>A>U (">"는 더 큰 선호도를 나타냄)로서 순위매겨지는, 기질의 3' 가장 가까운 이웃에 대해 선호도를 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 C~G~A>U (">"는 더 큰 선호도를 나타냄; "~"는 유사한 선호도를 나타냄)로서 순위매겨지는, 기질의 3' 가장 가까운 이웃에 대해 선호도를 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 TAG>AAG>CAC>AAT>GAA>GAC (">"는 더 큰 선호도를 나타냄)(중심 A는 표적 아데노신 잔기임)로서 순위매겨지는, 표적 아데노신 잔기를 함유하는 삼중 서열에 대한 선호도를 갖는다.

[0162] 일부 실시형태에서, 아데노신 데아미나제의 기질 편집 선호도는 아데노신 데아미나제 단백질에서 핵산 결합 도메인의 존재 또는 부재에 의해 영향받는다. 일부 실시형태에서, 기질 편집 선호도를 변형시키기 위해, 데아미나제 도메인은 이중-가닥 RNA 결합 도메인 (dsRBD) 또는 이중-가닥 RNA 결합 모티프 (dsRBM)와 결합된다. 일부 실시형태에서, dsRBD 또는 dsRBM은 ADAR 단백질, 예컨대 hADAR1 또는 hADAR2로부터 유래될 수 있다. 일부 실시형태에서, 적어도 하나의 dsRBD 및 데아미나제 도메인을 포함하는 전체 길이 ADAR 단백질이 사용된다. 일부 실시형태에서, 하나 이상의 dsRBM 또는 dsRBD는 데아미나제 도메인의 N-말단에 있다. 다른 실시형태에서, 하나 이상의 dsRBM 또는 dsRBD는 데아미나제 도메인의 C-말단에 있다.

[0163] 일부 실시형태에서, 아데노신 데아미나제의 기질 편집 선호도는 효소의 활성 중심부 내 또는 그 근처 아미노산 잔기에 의해 영향받는다. 일부 실시형태에서, 기질 편집 선호도를 변형시키기 위해서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, G336D, G487R, G487K, G487W, G487Y, E488Q, E488N, T490A,

V493A, V493T, V493S, N597K, N597R, A589V, S599T, N613K, N613R의 돌연변이, 및 상기 상응하는 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다.

[0164] 특히, 일부 실시형태에서, 편집 특이성을 감소시키기 위해서, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, E488Q, V493A, N597K, N613K의 돌연변이, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다. 일부 실시형태에서, 편집 특이성을 증가시키기 위해서, 아데노신 데아미나제는 돌연변이 T490A를 포함할 수 있다.

[0165] 일부 실시형태에서, 바로 5' G를 갖는 표적 아데노신(A), 예컨대, 삼중 서열 GAC를 포함하는 기질 (중심 A는 표적 아데노신 잔기임)에 대한 편집 선호도를 증가시키기 위해, 아데노신 데아미나제는 hADAR2-D의 아미노산 서열 위치를 기반으로, G336D, E488Q, E488N, V493T, V493S, V493A, A589V, N597K, N597R, S599T, N613K, N613R의 돌연변이, 및 상기에 상응하는 상동성 ADAR 단백질에서의 돌연변이 중 하나 이상을 포함할 수 있다.

[0166] 특히, 일부 실시형태에서, 아데노신 데아미나제는 하기의 삼중 서열을 포함하는 기질을 편집하기 위해서, 돌연변이 E488Q 또는 상동성 ADAR 단백질 내 해당 돌연변이를 포함한다: GAC, GAA, GAU, GAG, CAU, AAU, UAC, 중심 A는 표적 아데노신 잔기이다.

[0167] 일부 실시형태에서, 아데노신 데아미나제는 SEQ ID No. 761로 정의된 바와 같은 hADAR1-D의 야생형 아미노산 서열을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 hADAR1-D의 편집 효율, 및/또는 기질 편집 선호도를 특별한 요구도에 따라 변화시키도록, hADAR1-D 서열 내에 하나 이상의 돌연변이를 포함한다.

[0168] 일부 실시형태에서, 아데노신 데아미나제는 hADAR1-D 아미노산 서열의 글리신1007에서, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 상대적으로 작은 측쇄를 갖는 비극성 아미노산 잔기로 치환된다. 예를 들어, 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 알라닌 잔기로 치환된다 (G1007A). 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 발린 잔기로 치환된다 (G1007V). 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 상대적으로 큰 측쇄를 갖는 아미노산 잔기로 치환된다. 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 아르기닌 잔기로 치환된다 (G1007R). 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 리신 잔기로 치환된다 (G1007K). 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 트립토판 잔기로 치환된다 (G1007W). 일부 실시형태에서, 1007번 위치에서 글리신 잔기는 티로신 잔기로 치환된다 (G1007Y). 추가적으로, 다른 실시형태에서, 1007번 위치에서 글리신 잔기는 류신 잔기로 치환된다 (G1007L). 다른 실시형태에서, 1007번 위치에서 글리신 잔기는 트레오닌 잔기로 치환된다 (G1007T). 다른 실시형태에서, 1007번 위치에서 글리신 잔기는 세린 잔기로 치환된다 (G1007S).

[0169] 일부 실시형태에서, 아데노신 데아미나제는 hADAR1-D 아미노산 서열의 글루탐산1008에서, 또는 상동성 ADAR 단백질 내 상응하는 위치에 돌연변이를 포함한다. 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 상대적으로 큰 측쇄를 갖는 극성 아미노산 잔기로 치환된다. 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 글루타민 잔기로 치환된다 (E1008Q). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 히스티딘 잔기로 치환된다 (E1008H). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 아르기닌 잔기로 치환된다 (E1008R). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 리신 잔기로 치환된다 (E1008K). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 비극성 또는 작은 극성의 아미노산 잔기로 치환된다. 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 페닐알라닌 잔기로 치환된다 (E1008F). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 트립토판 잔기로 치환된다 (E1008W). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 글리신 잔기로 치환된다 (E1008G). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 아이소류신 잔기로 치환된다 (E1008I). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 발린 잔기로 치환된다 (E1008V). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 프롤린 잔기로 치환된다 (E1008P). 일부 일시형태에서, 1008번 위치에서 글루탐산 잔기는 세린 잔기로 치환된다 (E1008S). 다른 실시형태에서, 1008번 위치에서 글루탐산 잔기는 아스파라긴 잔기로 치환된다 (E1008N). 다른 실시형태에서, 1008번 위치에서 글루탐산 잔기는 알라닌 잔기로 치환된다 (E1008A). 다른 실시형태에서, 1008번 위치에서 글루탐산 잔기는 메티오닌 잔기로 치환된다 (E1008M). 일부 실시형태에서, 1008번 위치에서 글루탐산 잔기는 류신 잔기로 치환된다 (E1008L).

[0170] 일부 실시형태에서, 편집 효율을 개선시키기 위해서, 아데노신 데아미나제는 hADAR1-D의 아미노산 서열 위치를 기반으로, 돌연변이: E1007S, E1007A, E1007V, E1008Q, E1008R, E1008H, E1008M, E1008N, E1008K, 및 상기에 상응하는 상동성 ADAR 단백질 내 돌연변이 중 하나 이상을 포함할 수 있다.

[0171] 일부 실시형태에서, 편집 효율을 감소시키기 위해서, 아데노신 데아미나제는 hADAR1-D의 아미노산 서열 위치를 기반으로, 돌연변이: E1007R, E1007K, E1007Y, E1007L, E1007T, E1008G, E1008I, E1008P, E1008V, E1008F, E1008W, E1008S, E1008N, E1008K, 및 상기에 상응하는 상동성 ADAR 위치 내 돌연변이 중 하나 이상을 포함할 수 있다.

[0172] 일부 실시형태에서, 아데노신 데아미나제의 기질 편집 선호도, 효율 및/또는 선택성은 효소의 활성 중심부 또는 그 근처의 아미노산 잔기에 의해 영향받는다. 일부 실시형태에서, 아데노신 데아미나제는 hADAR1-D 서열 내 글루탐산 1008, 또는 상동성 ADAR 단백질 내 해당 위치에 돌연변이를 포함한다. 일부 실시형태에서, 돌연변이 E1008R이거나, 또는 상동성 ADAR 단백질 내 해당 돌연변이이다. 일부 실시형태에서, E1008R 돌연변이체는 반대쪽 가닥에 미스매치된 G 잔기를 갖는 아데노신 잔기의 표적화에 대해 증가된 편집 효율을 갖는다.

[0173] 일부 실시형태에서, 아데노신 데아미나제 단백질은 이중 가닥 핵산 기질을 인식하고 결합하기 위해서 하나 이상의 이중 가닥 RNA (dsRNA) 결합 모티프 (dsRBM) 또는 도메인 (dsRBD)를 더 포함하거나 또는 그에 연결된다. 일부 실시형태에서, 아데노신 데아미나제 및 이중 가닥 기질 간 상호작용은 CRISPR/CAS 단백질 인자를 포함하여, 하나 이상의 추가 단백질 인자(들)에 의해 매개된다. 일부 실시형태에서, 아데노신 데아미나제 및 이중 가닥 기질 간 상호작용은 가이드 RNA를 포함하여, 하나 이상의 핵산 성분(들)에 의해 더욱 매개된다.

[0174] 일정 예의 실시형태에서, 유도 진화를 사용하여 아데닌의 하이포잔틴으로의 탈아미노화 이외에도 추가의 반응을 촉매할 수 있는 변형된 ADAR 단백질을 디자인할 수 있다.

[0175] 본 발명에 따라서, 아데노신 데아미나제의 기질은 이후에 CRISPR-Cas 효소와 CRISPR-Cas 복합체를 형성하게 되는 그의 DNA 표적과 가이드 분자의 결합 시 형성되는 RNA/DNAn RNA 듀플렉스이다. RNA/DNA 또는 DNA/RNAn RNA 듀플렉스는 또한 본 명세서에서 "RNA/DNA 하이브리드", "DNA/RNA 하이브리드" 또는 "이중 가닥 기질"이라고도 한다. 가이드 분자 및 CRISPR-Cas 효소의 특정한 특성은 하기에 상술된다.

[0176] 본 명세서에서 사용되는 용어 "편집 선택성"은 아데노신 데아미나제에 의해 편집되는 이중 가닥 기질 상의 모든 부위의 분율을 의미한다. 이론에 국한하지 않지만, 아데노신 데아미나제의 편집 선택성은 이중 가닥 기질의 길이 및 2차 구조, 예컨대 미스매치된 염기, 벌지 및/또는 내부 루프의 존재에 의해 영향받는다는 것을 고려한다.

[0177] 일부 실시형태에서, 기질이 50 bp보다 긴 완벽하게 염기-쌍형성된 듀플렉스인 경우에, 아데노신 데아미나제는 듀플렉스 내 다수 아데노신 잔기를 탈아미노화시킬 수도 있다 (예를 들어, 모든 아데노신 잔기의 50%). 일부 실시형태에서, 기질이 50 bp 보다 짧은 경우에, 아데노신 데아미나제의 편집 선택성은 표적 아데노신 부위에 미스매치의 존재에 의해 영향받는다. 특히, 일부 실시형태에서, 반대쪽 가닥 상에 미스매치된 시티딘 (C) 잔기를 갖는 아데노신 (A) 잔기는 고효율로 탈아미노화된다. 일부 실특형태에서, 반대쪽 가닥 상에 미스매치된 구아노신 (G) 잔기를 갖는 아데노신 (A) 잔기는 편집없이 스키핑된다.

[0178] **CRISPR-Cas 단백질 및 가이드**

[0179] 본 발명의 방법 및 시스템에서, CRISPR-Cas 단백질 및 상응하는 가이드 분자가 사용된다. 보다 특히, CRISPR-Cas 단백질은 클래스 2 CRISPR-Cas 단백질이다. 일정 실시형태에서, 상기 CRISPR-Cas 단백질은 Cas13이다. CRISPR-Cas 시스템은 특별한 서열을 표적화하기 위해 맞춤 단백질의 생성을 요구하지 않지만 대신 단일 Cas 단백질은 특별한 핵산 표적을 인식하기 위해서 가이드 분자에 의해 프로그램될 수 있는데, 달리 말해서, Cas 효소 단백질은 상기 가이드 분자를 사용하여 관심의 특별한 핵산 표적 유전자좌로 동원될 수 있다.

[0180] *가이드 분자*

[0181] 클래스 2 V형 CRISPR-Cas 단백질의 가이드 분자 또는 가이드 RNA는 tracr-mate 서열 (내생성 CRISPR 시스템의 경우 "직접 반복부" 포괄) 및 가이드 서열 (내생성 CRISPR 시스템의 경우 "스페이서"라고도 함)을 포함한다. 실제로, II형 CRISPR-Cas 단백질과 대조적으로, Cas13은 tracr 서열의 존재에 의존하지 않는다. 일부 실시형태에서, 본 명세서에 기술된 바와 같은 CRISPR-Cas 시스템 또는 복합체는 (예를 들어, Cas 단백질 Cas13이라면) tracr 서열의 존재에 의존하지 않고/않거나 그를 포함하지 않는다. 일정 실시형태에서, 가이드 분자는 가이드 서열 또는 스페이서 서열에 융합 또는 연결된 직접 반복부 서열을 포함할 수 있거나, 그로 본질적으로 이루어질 수 있거나, 또는 그로 이루어질 수 있다.

[0182] 일반적으로, CRISPR 시스템은 표적 서열 부위에서 CRISPR 복합체의 형성을 촉진하는 구성요소를 특징으로 한다. CRISPR 복합체의 형성 경우에, "표적 서열"은 가이드 서열이 상보성을 갖도록 디자인된 서열을 의미하고, 여

기서 표적 DNA 서열과 가이드 서열 간 하이브리드화가 CRISPR 복합체의 형성을 촉진한다.

[0183] 용어 "가이드 분자" 및 "가이드 RNA"는 CRISPR-Cas 단백질과 복합체를 형성할 수 있고 표적 핵산 서열과 충분한 상보성을 갖는 가이드 서열을 포함하여 표적 핵산 서열과 하이브리드화하고 표적 핵산 서열과 복합체의 서열-특이적 결합을 유도하는 RNA-기반 분자를 의미하고자 본 명세서에서 상호교환적으로 사용된다. 가이드 분자 또는 가이드 RNA는 특별히 본 명세서에 기술된 바와 같이, (예를 들어, 2개의 리보뉴클레오티드를 화학적으로 연결하거나 또는 하나 이상의 리보뉴클레오티드를 하나 이상의 데옥시리보뉴클레오티드로 치환시켜서) 하나 이상의 화학적 변형을 갖는 RNA-기반 분자를 포괄한다.

[0184] 본 명세서에서 사용시, CRISPR-Cas 시스템에서 용어 "가이드 서열"은 표적 핵산 서열과 하이브리드하고 표적 핵산 서열과 핵산-표적화 복합체의 서열-특이적 결합을 유도하도록 표적 핵산 서열과 충분한 상보성을 갖는 임의의 폴리뉴클레오티드 서열을 포함한다. 본 발명에서, 표적 핵산 서열 또는 표적 서열은 본 명세서에서 "표적 아데노신"이라고도 하는 탈아미노화되는 표적 아데노신을 포함하는 서열이다. 일부 실시형태에서, 의도하는 dA-C 미스매치를 제외하고, 적합한 정렬 알고리즘을 사용해 최적으로 정렬했을 때, 상보성 정도는 약 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, 또는 그 이상이다. 최적 정렬은 서열을 정렬하기 위한 임의의 적합한 알고리즘을 사용하여 결정할 수 있고, 이의 비제한적인 예는 스미스-워터만 (Smith-Waterman) 알고리즘, 니들만-분취 (Needleman-Wunsch) 알고리즘, 버로우스-윌러스 (Burrows-Wheeler) 전환 기반 알고리즘 (예를 들어, Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; www.novocraft.com에서 입수가능), ELAND (Illumina, San Diego, CA), SOAP (soap.genomics.org.cn에서 입수가능), 및 Maq (maq.sourceforge.net에서 입수가능)을 포함한다. 표적 핵산 서열과 핵산-표적화 복합체의 서열-특이적 결합을 유도하는 (핵산-표적화 가이드 RNA 내) 가이드 서열의 능력은 임의의 적합한 어세이를 통해 평가될 수 있다. 예를 들어, 시험하려는 가이드 서열을 포함하여, 핵산-표적화 복합체를 형성하기 위해 충분한 핵산-표적화 CRISPR 시스템의 성분은 예컨대 핵산-표적화 복합체의 성분을 코딩하는 벡터에 의한 형질감염 이후, 예컨대 본 명세서에 기술된 바와 같은 Surveyor 어세이를 통하여, 표적 핵산 서열 내 선호 표적화 (예를 들어, 절단)를 평가하여, 상응하는 표적 핵산 서열을 갖는 숙주 세포에 제공될 수 있다. 유사하게, 표적 핵산 서열 (또는 이의 부근 서열)의 절단은 표적 핵산 서열, 시험하려는 가이드 서열을 포함하여, 핵산-표적화 복합체의 성분 및 시험 가이드 서열과 상이한 대조군 가이드 서열을 제공하고, 시험 및 대조군 가이드 서열 반응 간에 표적 서열 또는 그 부근에서의 결합 또는 절단율을 비교하여 시험관에서 평가될 수 있다. 다른 어세이도 가능하고, 당업자에게 자명할 것이다. 가이드 서열, 및 따라서 핵산-표적화 가이드 RNA는 임의의 표적 핵산 서열을 표적화하도록 선택될 수 있다.

[0185] 일부 실시형태에서, 가이드 분자는 표적 서열과 적어도 하나의 미스매치를 갖도록 디자인된 가이드 서열을 포함하여서, 가이드 서열과 표적 서열 간에 형성된 RNA 듀플렉스는 표적 서열 상에서 탈아미노화를 위한 표적 A의 반대쪽 가이드 서열 에 쌍-비형성된 C를 포함한다. 일부 실시형태에서, 이러한 A-C 미스매치이외에도, 적합한 정렬 알고리즘을 사용해 최적으로 정렬했을때, 상보성 정도는 약 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, 또는 그 이상이다.

[0186] 일정 실시형태에서, 가이드 분자의 가이드 서열 또는 스페이서 길이는 15 내지 50 nt이다. 일정 실시형태에서, 가이드 RNA의 스페이서 길이는 적어도 15 뉴클레오티드이다. 일정 실시형태에서, 스페이서 길이는 15 내지 17 nt, 예를 들어, 15, 16, 또는 17 nt, 17 내지 20 nt, 예를 들어 17, 18, 19, 또는 20 nt, 20 내지 24 nt, 예를 들어, 20, 21, 22, 23, 또는 24 nt, 23 내지 25 nt, 예를 들어, 23, 24, 또는 25 nt, 24 내지 27 nt, 예를 들어, 24, 25, 26, 또는 27 nt, 27-30 nt, 예를 들어, 27, 28, 29, 또는 30 nt, 30-35 nt, 예를 들어, 30, 31, 32, 33, 34, 또는 35 nt, 또는 35 nt 또는 그 이상의 길이이다. 일정 예의 실시형태에서, 가이드 서열은 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, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100 nt이다.

[0187] 일부 실시형태에서, 가이드 서열은 10 내지 50 nt 길이, 보다 특히 약 20-30 nt, 유리하게 약 20 nt, 23-25 nt 또는 24 nt의 RNA 서열이다. 가이드 서열은 탈아미노화되는 아데노신을 포함하는 표적 서열과의 하이브리드화를 보장하도록 선택된다. 이것은 하기에 보다 상술된다. 선택은 탈아미노화의 효율 및 특이성을 증가시키는 단계를 더 포괄한다.

[0188] 일부 실시형태에서, 가이드 서열은 약 20 nt 내지 약 30 nt 길이이고 표적 DNA 가닥과 하이브리드화하여 표적

아데노신 부위에서 dA-C 미스매치를 갖는 것을 제외하고는, 거의 완벽하게 일치되는 듀플렉스를 형성한다. 특히, 일부 실시형태에서, dA-C 미스매치는 표적 서열의 중심 (따라서 표적 서열과 가이드 서열의 하이브리드화 시 듀플렉스의 중심에 가깝게 위치되어서, 아데노신 데아미나제를 협소한 편집창 (예를 들어, 약 4 bp 너비)으로 제한시킨다. 일부 실시형태에서, 표적 서열은 탈아미노화하려는 하나 초과의 표적 아데노신을 포함할 수 있다. 추가 실시형태에서 표적 서열은 표적 아데노신 부위의 3'에 하나 이상의 dA-C 미스매치를 더 포함할 수 있다. 일부 실시형태에서, 표적 서열 내 의도하지 않은 아데닌 부위에서 오프-표적 편집을 피하기 위해서, 가이드 서열은 일정 아데노신 데아미나제 예컨대 ADAR1 및 ADAR2에 대해 촉매적으로 비호의적인, dA-G 미스매치를 도입시키기 위해 상기 의도하지 않은 아데닌에 상응하는 위치에 비-쌍형성 구아닌을 포함하도록 디자인될 수 있다. 본 명세서에 그 전문이 참조로 편입되는 [Wong et al., RNA 7:846-858 (2001)]을 참조한다.

[0189] 일부 실시형태에서, 정규 길이 (예를 들어, AsCas13의 경우 약 24 nt)를 갖는 Cas13 가이드 서열은 표적 DNA와 RNA 듀플렉스를 형성시키는데 사용된다. 일부 실시형태에서, 정규 길이보다 긴 (예를 들어, AsCas13의 경우 >24 nt) Cas13 가이드 분자는 Cas13-가이드 RNA-표적 DNA 복합체의 외부를 포함하는 표적 DNA와 RNA 듀플렉스를 형성시키는데 사용된다. 이것은 소정 스트레치의 뉴클레오티드 내 하나 초과의 아데닌의 탈아미노화에 관심 있을 때 특히 흥미로울 수 있다. 대안적인 실시형태에서, 정규 가이드 서열 길이의 제한성을 유지시키는 것이 흥미롭다. 일부 실시형태에서, 가이드 서열은 Cas13 가이드의 정규 길이 외부에 dA-C 미스매치를 도입시키도록 디자인되어서, 이것은 아데노신 데아미나제 및 dA-C 미스매치 간 접촉 빈도를 증가시키고 Cas13에 의한 입체 방해를 감소시킬 수 있다.

[0190] 일부 실시형태에서, 가이드 분자 (직접 반복부 및/또는 스페이서)의 서열은 가이드 분자 내 2차 구조 정도를 감소시키도록 선택된다. 일부 실시형태에서, 핵산-표적화 가이드 RNA의 뉴클레오티드의 약 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1% 이하가 최적으로 폴딩되었을 때 자기-상보성 염기쌍에 참여한다. 최적 폴딩은 임의의 적합한 폴리뉴클레오티드 폴딩 알고리즘을 통해 결정될 수 있다. 일부 프로그램은 최소 깁스 (Gibbs) 자유 에너지의 계산을 기반으로 한다. 이러한 알고리즘의 일례는 [Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148)]에 기술된 바와 같은, mFold이다. 다른 예의 폴딩 알고리즘은 중심 구조 예측 알고리즘을 사용하여, 유니버시티 오브 비엔나 (University of Vienna)의 이론 화학 연구소 (Institute for Theoretical Chemistry)에서 개발한, 온라인 웹서버 RNAfold이다 (예를 들어, [A.R. Gruber et al., 2008, Cell 106(1): 23-24]; 및 [PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62] 참조).

[0191] 일부 실시형태에서, RNA 절단, 예컨대 Cas13에 의한 절단에 대한 가이드 분자의 감수성을 감소시키는 것에 관심을 갖는다. 따라서, 특정 실시형태에서, 가이드 분자는 Cas13 또는 다른 RNA-절단 효소에 의한 절단을 피하도록 조정된다.

[0192] 일정 실시형태에서, 가이드 분자는 비천연 발생 핵산 및/또는 비천연 발생 뉴클레오티드 및/또는 뉴클레오티드 유사체, 및/또는 화학적 변형을 포함한다. 바람직하게, 이들 비천연 발생 핵산 및 비천연 발생 뉴클레오티드는 가이드 서열 외부에 위치된다. 비천연 발생 핵산은 예를 들어 천연 및 비천연 발생 뉴클레오티드의 혼합물을 포함한다. 비천연 발생 뉴클레오티드 및/또는 뉴클레오티드 유사체는 리보스, 포스페이트, 및/또는 염기 모이어티에서 변형될 수 있다. 본 발명의 일 실시형태에서, 가이드 핵산은 리보뉴크레오티드 및 비-리보뉴클레오티드를 포함한다. 이러한 일 실시형태에서, 가이드는 하나 이상의 리보뉴클레오티드 및 하나 이상의 데옥시리보뉴클레오티드를 포함한다. 본 발명의 일 실시형태에서, 가이드는 하나 이상의 비천연 발생 뉴클레오티드 또는 뉴클레오티드 유사체 예컨대 포스포로티오에이트 연결부를 갖는 뉴클레오티드, 리보스 고리의 2' 및 4' 탄소 사이에 메틸렌 가교를 포함하는 잠김 핵산 (LNA) 뉴클레오티드, 또는 가교 핵산 (BNA)을 포함한다. 변형된 뉴클레오티드의 다른 예는 2'-O-메틸 유사체, 2'-데옥시 유사체, 또는 2'-플루오로 유사체를 포함한다. 변형된 염기의 추가 예는 제한없이 2-아미노푸린, 5-브로모-우리딘, 슈도우리딘, 이노신, 7-메틸구아노신을 포함한다. 가이드 RNA의 화학적 변형의 예는 제한없이, 하나 이상의 말단 뉴클레오티드에 2'-O-메틸 (M), 2'-O-메틸 3' 포스포로티오에이트 (MS), S-구속형 에틸 (cEt), 또는 2'-O-메틸 3'티오PACE (MSP)의 도입을 포함한다. 이러한 화학적으로 변형된 가이드는 온-표적 대 오프-표적 특이성이 예측불가하더라도, 비변형된 가이드와 비교하여 증가된 안정성 및 증가된 활성을 포함할 수 있다 (다음의 문헌들을 참조한다: Hendel, 2015, Nat Biotechnol. 33(9):985-9, doi: 10.1038/nbt.3290, 2015년 6월 29일 온라인 공개; Ragdarm et al., 0215, PNAS, E7110-E7111; Allerson et al., J. Med. Chem. 2005, 48:901-904; Bramsen et al., Front. Genet., 2012, 3:154; Deng et al., PNAS, 2015, 112:11870-11875; Sharma et al., MedChemComm., 2014, 5:1454-1471; Hendel et al., Nat. Biotechnol. (2015) 33(9): 985-989; Li et al., Nature Biomedical Engineering, 2017, 1, 0066 DOI:10.1038/s41551-017-0066). 일부 실시형태에서, 가이드 RNA의 5' 및/또는

3' 말단은 형광성 염료, 폴리에틸렌 글리콜, 콜레스테롤, 단백질, 또는 검출 태그를 포함한 다양한 기능성 모이어티에 의해 변형된다 ([Kelly et al., 2016, J. Biotech. 233:74-83] 참조). 일정 실시형태에서, 가이드는 표적 DNA에 결합하는 영역 내 리보뉴클레오티드 및 Cas13에 결합하는 영역 내 하나 이상의 데옥시리보뉴클레오티드 및/또는 뉴클레오티드를 포함한다. 본 발명의 일 실시형태에서, 데옥시리보뉴클레오티드 및/또는 뉴클레오티드 유사체는 조작된 가이드 구조, 예컨대 제한없이, 스템-루프 영역, 및 씨드 영역에 도입된다. Cas13 가이드의 경우에, 일정 실시형태에서, 변형은 스템-루프 영역의 5'-핸들에 존재하지 않는다. 가이드의 스템-루프 영역의 5'-핸들 내 화학적 변형은 이의 기능을 폐기시킬 수 있다 ([Li, et al., Nature Biomedical Engineering, 2017, 1:0066] 참조). 일정 실시형태에서, 가이드의 적어도 1, 2, 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, 35, 40, 45, 50, 또는 75개 뉴클레오티드는 화학적으로 변형된다. 일부 실시형태에서, 가이드의 3' 또는 5' 말단에서 3-5개 뉴클레오티드는 화학적으로 변형된다. 일부 실시형태에서, 오직 소수의 변형, 예컨대 2'-F 변형이 씨드 영역에 도입된다. 일부 실시형태에서, 2'-F 변형이 가이드의 3' 말단에 도입된다. 일정 실시형태에서, 가이드의 5' 및/또는 3' 말단에서 3개 내지 5개 뉴클레오티드는 2'-O-메틸 (M), 2'-O-메틸 3' 포스포로티오에이트 (MS), S-구속형 에틸 (cEt), 또는 2'-O-메틸 3' 티오PACE (MSP)로 화학적으로 변형된다. 이러한 변형은 게놈 편집 효율을 증강시킬 수 있다 ([Hendel et al., Nat. Biotechnol. (2015) 33(9): 985-989] 참조). 일정 실시형태에서, 가이드의 모든 포스포디에스테르 결합은 유전자 파괴 수준을 증강시키기 위해 포스포로티오에이트 (PS)로 치환된다. 일정 실시형태에서, 가이드의 5' 및/또는 3' 말단에서 5개 초과의 뉴클레오티드가 2'-O-Me, 2'-F 또는 S-구속형 에틸(cEt)로 화학적으로 치환된다. 이러한 화학적으로 변형된 가이드는 유전자 파괴의 증강된 수준을 매개할 수 있다 ([Ragdarm et al., 0215, PNAS, E7110-E7111] 참조). 본 발명의 일 실시형태에서, 가이드는 이의 3' 및/또는 5' 말단에서 화학적 모이어티를 포함하도록 변형된다.

이러한 모이어티는 제한없이, 아민, 아지드, 알킨, 티오, 디벤조시클로옥틴 (DBCO), 또는 로다민을 포함한다. 일정 실시형태에서, 화학적 모이어티는 링커, 예컨대 알킬 사슬을 통해서 가이드에 접합된다. 일정 실시형태에서, 변형된 가이드의 화학적 모이어티는 다른 분자, 예컨대 DNA, RNA, 단백질, 또는 나노입자에 가이드를 부착시키는데 사용될 수 있다. 이러한 화학적으로 변형된 가이드는 CRISPR 시스템에 의해 유전적으로 편집된 세포를 확인하거나 또는 농축시키는데 사용될 수 있다 ([Lee et al., eLife, 2017, 6:e25312, DOI:10.7554] 참조).

[0193] 일부 실시형태에서, 가이드는 3'-말단 및 씨드 영역을 더 포함하는 가이드 절편 및 5'-핸들을 가지는, 변형된 Cas13 crRNA를 포함한다. 일부 실시형태에서, 변형된 가이드는 아시다미노코커스 (*Acidaminococcus*) sp. BV3L6 Cas13 (AsCas13); 프란시셀라 툴라렌시스 (*Francisella tularensis*) subsp. 노비시다 (Novicida) U112 Cas13 (FnCas13); 엘. 박테리움 (*L. bacterium*) MC2017 Cas13 (Lb3Cas13); 부티리비브리오 프로테오클라스티커스 (*Butyrivibrio proteoclasticus*) Cas13 (BpCas13); 파르쿠박테리아 박테리움 (*Parcubacteria bacterium*) GWC2011_GWC2_44_17 Cas13 (PbCas13); 페레그리니박테리아 박테리움 (*Peregrinibacteria bacterium*) GW2011_GWA_33_10 Cas13 (PeCas13); 렙토스피라 이나다이 (*Leptospira inadai*) Cas13 (LiCas13); 스미텔라 (*Smithella*) sp. SC_K08D17 Cas13 (SsCas13); 엘. 박테리움 MA2020 Cas13 (Lb2Cas13); 포르피로모나스 크레비오리카니스 (*Porphyromonas crevioricanis*) Cas13 (PcCas13); 포르피로모나스 마카카 (*Porphyromonas macacae*) Cas13 (PmCas13); 칸디다터스 메타노플라스마 터미툼 (*Candidatus Methanoplasma termitum*) Cas13 (CMtCas13); 유박테리움 엘리겐스 (*Eubacterium eligens*) Cas13 (EeCas13); 모라셀라 보보쿨리 (*Moraxella bovoculi*) 237 Cas13 (MbCas13); 프레보텔라 디시엔스 (*Prevotella disiens*) Cas13 (PdCas13); 또는 엘. 박테리움 ND2006 Cas13 (LbCas13) 중 어느 하나의 Cas13과 함께 사용될 수 있다.

[0194] 일부 실시형태에서, 가이드에 대한 변형은 화학적 변형, 삽입, 결실 또는 분할이다. 일부 실시형태에서, 화학적 변형은 제한없이 2'-O-메틸 (M) 유사체, 2'-데옥시 유사체, 2-티오우리딘 유사체, N6-메틸아데노신 유사체, 2'-플루오로 유사체, 2-아미노푸린, 5-브로모-우리딘, 슈도우리딘 (Ψ), N1-메틸슈도우리딘 (me1Ψ), 5-메톡시우리딘 (5moU), 이노신, 7-메틸구아노신, 2'-O-메틸 3'포스포로티오에이트 (MS), S-구속형 에틸(cEt), 포스포로티오에이트 (PS), 또는 2'-O-메틸 3'티오PACE (MSP)의 도입을 포함한다. 일부에서 가이드는 하나 이상의 포스포로티오에이트 변형을 포함한다. 일정 실시형태에서, 가이드의 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 25개 뉴클레오티드는 화학적으로 변형된다. 일정 실시형태에서, 씨드 영역 내 하나 이상의 뉴클레오티드는 화학적으로 변형된다. 일정 실시형태에서, 3'-말단 내 하나 이상의 뉴클레오티드는 화학적으로 변형된다. 일정 실시형태에서, 5'-핸들 내 뉴클레오티드는 화학적으로 변형되지 않는다. 일부에서 씨드 영역 내 화학적 변형은 소수의 변형, 예컨대 2'-플루오로 유사체의 도입이다. 특별한 실시형태에서, 씨드 영역의 하나의 뉴클레오티드는 2'-플루오로 유사체로 치환된다.

일부에서 3'-말단에서 5 내지 10개 뉴클레오티드가 화학적으로 변형된다. Cas13 CrRNA의 3'-말단에서 이러한 화학 변형은 Cas13 활성을 개선시킬 수 있다 ([Li, et al., Nature Biomedical Engineering, 2017, 1:0066] 참조). 특별한 실시형태에서, 3'-말단에서 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 뉴클레오티드가 2'-플루오로 유사체로 치환된다. 특별한 실시형태에서, 3'-말단에서 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 뉴클레오티드가 2'-O-메틸 (M) 유사체로 치환된다.

[0195] 일부 실시형태에서, 가이드의 5'-핸들의 루프가 변형된다. 일부 실시형태에서, 가이드의 5'-핸들의 루프는 결실, 삽입, 분할, 또는 화학적 변형을 갖도록 변형된다. 일정 실시형태에서, 변형된 루프는 3, 4, 또는 5개 뉴클레오티드를 포함한다. 일정 실시형태에서, 루프는 UCUU, UUUU, UAUU, 또는 UGUU의 서열을 포함한다.

[0196] 일부 실시형태에서, 가이드 분자는 DNA 또는 RNA일 수 있는 별개의 비공유적으로 연결된 서열과 스템 루프를 형성한다. 특정 실시형태에서, 가이드를 형성하는 서열은 표준 포스포르아미다이트 합성 프로토콜을 사용해 먼저 합성된다 (Herdewijn, P., ed., Methods in Molecular Biology Col 288, Oligonucleotide Synthesis: Methods and Applications, Humana Press, New Jersey (2012)). 일부 실시형태에서, 이들 서열은 당업자에게 공지된 표준 프로토콜을 사용하여 결찰을 위한 적절한 작용기를 함유하도록 작용화될 수 있다 (Hermanson, G. T., Bioconjugate Techniques, Academic Press (2013)). 작용기의 예는 제한없이 히드록실, 아민, 카르복실산, 카르복실산 할라이드, 카르복실산 활성 에스테르, 알데히드, 카르보닐, 클로로카르보닐, 이미다졸릴카르보닐, 히드로지드, 세미카르바지드, 티오 세미카르바지드, 티올, 말레이미드, 할로알킬, 술포닐, 알릴, 프로파르길, 디엔, 및 아지드를 포함한다. 이러한 서열이 작용화되면, 공유적 화학 결합 또는 연결부가 이러한 서열과 직접 반복부 서열 사이에 형성될 수 있다. 화학적 결합의 예는 제한없이 카바메이트, 에테르, 에스테르, 아미드, 이민, 아미딘, 아미노트리진, 히드로존, 디술피드, 티오에테르, 티오에스테르, 포스포로티오에이트, 포스포로디티오에이트, 술폰아미드, 술포네이트, 풀폰, 술폭시드, 우레아, 티오우레아, 히드라지드, 옥심, 트리아졸, 광불안정 연결부, C-C 결합 형성기 예컨대 딜스-알더 (Diels-Alder) 고리-부가 쌍 또는 고리-폐쇄 복분해 쌍, 및 마이클 (Michael) 반응 쌍을 기반으로 하는 것을 포함한다.

[0197] 일부 실시형태에서, 이들 스템-루프 형성 서열이 화학적으로 합성될 수 있다. 일부 실시형태에서, 화학 합성은 2'-아세톡시에틸 오르쏘에스테르 (2'-ACE) (Scaringe et al., J. Am. Chem. Soc. (1998) 120: 11820-11821; Scaringe, Methods Enzymol. (2000) 317: 3-18) 또는 2'-티오노카바메이트 (2'-TC) 화학 (Dellinger et al., J. Am. Chem. Soc. (2011) 133: 11540-11546; Hendel et al., Nat. Biotechnol. (2015) 33:985-989)에 의한 자동화 고체상 올리고뉴클레오티드 합성 기계를 사용한다.

[0198] 일정 실시형태에서, (표적 유전자좌로 Cas13을 가이드할 수 있는) 가이드 분자는 (1) 표적 유전자좌와 하이브리드화할 수 있는 가이드 서열 및 (2) 직접 반복부 서열이 가이드 서열의 상류 (즉, 5')에 위치되는 것인 직접 반복부 서열 또는 tracr mate를 포함한다. 특정한 실시형태에서, Cas13 가이드 서열의 씨드 서열 (즉, 표적 유전자좌를 인식하고/하거나 하이브리드화하는데 결정적인 필수 서열)은 대략 가이드 서열의 처음 10개 뉴클레오티드 내에 존재한다. 특정 실시형태에서, Cas13은 FnCas13이고 씨드 서열은 대략 가이드 서열의 5' 말단상의 처음 5 nt 내에 존재한다.

[0199] 특정한 실시형태에서, 가이드 분자는 직접 반복부 서열에 연결된 가이드 서열을 포함하고, 여기서 직접 반복부 서열은 하나 이상의 스템 루프 또는 최적화된 2차 구조를 포함한다. 특정 실시형태에서, 직접 반복부는 16 nt의 최소 길이 및 단일 스템 루프를 갖는다. 추가 실시형태에서, 직접 반복부는 16 nt 초과, 바람직하게 17 nt 초과의 길이를 가지고, 하나 초과의 스템 루프 또는 최적화된 2차 구조를 갖는다. 특정 실시형태에서 가이드 분자는 천연 직접 반복부 서열의 전체 또는 일부에 연결된 가이드 서열을 포함하거나 또는 그로 이루어진다. 전형적인 V형 Cas13 가이드 분자는 (3'에서 5' 방향으로): 가이드 서열 제1 상보성 스트레치 ("반복부"), (전형적으로 4 또는 5개 뉴클레오티드 길이인) 루프, 제2 상보성 스트레치 ("반복부에 상보성인 "안티-반복부"), 및 폴리 A (RNA에서 종종 폴리 U) 꼬리부 (종결자)를 포함한다. 일정 실시형태에서, 직접 반복부 서열은 이의 천연 아키텍처를 유지하고 단일 스템 루프를 형성한다. 특정 실시형태에서, 가이드 아키텍처의 일정 양상은 예를 들어 특성의 첨가, 차감, 또는 치환을 통해 변형될 수 있는 반면, 가이드 아키텍처의 일정한 다른 양상은 유지된다. 제한없이 삽입, 결실, 및 치환을 포함하는, 조작된 가이드 분자 변형을 위한 바람직한 위치는 Cas13 단백질 및/또는 표적과 복합체를 형성했을 때, 예를 들어 직접 반복부 서열의 스템루프를 노출시키는 가이드 분자의 영역 및 가이드 말단을 포함한다.

[0200] 특정 실시형태에서, 스템은 상보성 X 및 Y 서열을 포함하는 적어도 약 4 bp를 포함하지만, 더 많거나, 예를 들어, 5, 6, 7, 8, 9, 10, 11 또는 12개, 또는 더 적은, 예를 들어, 3, 2개 염기쌍의 스템이 또한 고려된다.

따라서, 예를 들어 X2-10 및 Y2-10 (여기서 X 및 Y는 뉴클레오티드의 임의의 상보성 세트를 나타냄)이 고려될 수 있다. 일 양상에서, 루프와 함께, X 및 Y 뉴클레오티드로 만들어진 스템은 전체 2차 구조에서 완전한 헤어핀을 형성하게 될 것이고, 이것은 유리할 수 있으며, 염기쌍의 양은 완전한 헤어핀을 형성하는 임의의 양일 수 있다. 일 양상에서, 임의의 상보성 X:Y 염기쌍 서열 (예를 들어, 길이에 대해)은 전체 가이드 분자의 2차 구조가 보존되는 한, 허용된다. 일 양상에서, X:Y 염기쌍으로 만들어진 스템을 연결시키는 루프는 가이드 분자의 전체 2차 구조를 파괴하지 않는 동일 길이 (예를 들어, 4 또는 5개 뉴클레오티드)이거나 또는 더 긴 임의 서열일 수 있다. 일 양상에서, 스템루프는 예를 들어, MS2 압타머를 더 포함할 수 있다. 일 양상에서, 스템은 상보성 X 및 Y 서열을 포함하는 약 5-7 bp를 포함하지만, 그 이상 또는 그 이하의 염기쌍의 스템도 고려된다. 일 양상에서, 달리 이러한 쌍형성이 일반적으로 그 위치에서 스템루프의 아키텍처를 보존하는 경우에, 비-왓슨 크릭 염기쌍이 고려된다.

[0201] 특정 실시형태에서 가이드 분자의 천연 헤어핀 또는 스템루프 구조는 연장되거나 또는 연장된 스템루프로 대체된다. 스템의 연장은 CRISPR-Cas 단백질과 가이드 분자의 조립을 증강시킬 수 있다는 것이 입증되었다 (Chen et al. Cell. (2013); 155(7): 1479-1491). 특정 실시형태에서 스템루프의 스템은 적어도 1, 2, 3, 4, 5 또는 그 이상의 상보성 염기쌍 (즉, 가이드 분자 내 2,4, 6, 8, 10개 또는 그 이상의 뉴클레오티드의 첨가에 상응)만큼 연장된다. 특정 실시형태에서 이들은 스템루프의 루프에 인접하는, 스템의 말단에 위치된다.

[0202] 특정 실시형태에서, RNAse 또는 감소된 발현에 대한 가이드 분자의 감수성은 이의 기능에 영향을 미치지 않는 가이드 분자의 서열의 약간의 변형에 의해 감소될 수 있다. 예를 들어, 특정 실시형태에서, 전사물의 미성숙 종결, 예컨대 U6 Pol-III의 미성숙 전사는 가이드 분자 서열 내 추정의 Pol-III 종결자 (4개의 연속 U)를 변형시켜 제거될 수 있다. 이러한 서열 변형이 가이드 분자의 스템루프에서 필요한 경우에, 바람직하게는 염기쌍 플립에 의해 보장된다.

[0203] 바람직한 실시형태에서, 직접 반복부는 하나 이상의 단백질-결합 RNA 압타머를 포함하도록 변형될 수 있다. 특정한 실시형태에서, 하나 이상의 압타머는 예컨대 최적 2차 구조의 일부분을 포함할 수 있다. 이러한 압타머는 본 명세서에서 더욱 상술되는 바와 같은 박테리오파지 외피 단백질에 결합할 수도 있다.

[0204] 일부 실시형태에서, 가이드 분자는 편집하려는 적어도 하나의 표적 아데노신 잔기를 포함하는 표적 DNA 가닥과 듀플렉스를 형성한다. 표적 DNA 가닥과 가이드 RNA 분자의 하이브리드화 시에, 아데노신 데아미나제는 듀플렉스에 결합하고 DNA-RNA 듀플렉스 내에 포함된 하나 이상의 표적 아데노신 잔기의 탈아미노화를 촉매한다.

[0205] 가이드 서열, 및 따라서 핵산-표적화 가이드 RNA는 임의의 표적 핵산 서열을 표적화하도록 선택될 수 있다. 표적 서열은 DNA일 수 있다. 표적 서열은 게놈 DNA일 수 있다. 표적 서열은 미토콘드리아 DNA일 수 있다.

[0206] 일정 실시형태에서, 표적 서열은 PAM (protospacer adjacent motif) 또는 PFS (protospacer flanking sequence 또는 site), 다시 말해서, CRISPR 복합체에 의해 인식되는 짧은 서열과 회합되어야 한다. CRISPR-Cas 단백질의 성질에 따라서, 표적 서열은 DNA 듀플렉스 내 이의 상보성 서열 (본 명세서에서 비-표적 서열이라고도 함)이 PAM의 상류 또는 하류에 존재하도록 선택되어야 한다. CRISPR-Cas 단백질이 Cas13 단백질인 본 발명의 실시형태에서, a 내 표적 서열의 상보성 서열은 PAM의 하류 또는 3'에 존재한다. PAM을 위한 정확한 서열 및 길이 요건은 사용되는 Cas13 단백질에 따라 상이하지만, PAM은 전형적으로 프로토 스페이서 (즉, 표적 서열)에 인접한 2-5 염기쌍 서열이다. 상이한 Cas13 오솔로그에 대한 천연 PAM 서열의 예는 하기 본 명세서에서 제공하고 당업자는 소정 Cas13 단백질과 사용을 위한 추가 PAM 서열을 확인할 수 있을 것이다.

[0207] 추가로, PAM 상호작용 (PI) 도메인의 조작은 PAM 특이성의 프로그래밍을 가능하게 할 수 있고, 표적 부위 인식 신뢰성을 개선시킬 수 있으며, 예를 들어 [Kleinstiver BP et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015 Jul 23;523(7561):481-5. doi: 10.1038/nature14592]의 Cas9에 대해 기술된 바와 같이, CRISPR-Cas 단백질의 다재다능성을 증가시킨다. 본 명세서에서 더욱 상술하는 바와 같이, 당업자는 Cas13 단백질을 유사하게 변형시킬 수 있다는 것을 이해하게 될 것이다.

[0208] 특정 실시형태에서, 가이드 서열은 탈아미노화시키려는 아데닌에 대한 데아미나제의 최적 효율을 보장하기 위해 선택된다. Cas13 닉카제의 절단 부위에 대해서 표적 가닥 내 아데닌의 위치를 고려할 수 있다. 특정 실시형태에서 닉카제는 비-표적 가닥 상에서, 탈아미노화시키려는 아데닌 부근에서 작용할 것임을 보장하는 것에 관심있다. 예를 들어, 특정 실시형태에서, Cas13 닉카제는 PAM의 하류에 17개 뉴클레오티드 (예를 들어, AsCas13, LbCas13) 또는 PAM의 하류에 18개 뉴클레오티드 (FnCas13)의 비표적화 가닥을 절단하고, 탈아미노화시

키려는 아데닌에 상응하는 시토신이 상응하는 비표적 가닥의 서열 내 닉카제 절단 부위의 10 bp 상류 또는 하류 내에서 가이드 서열에 위치되는 가이드를 디자인하는 것이 흥미로울 수 있다.

[0209] 특정 실시형태에서, 가이드는 에스코트 가이드이다. "에스코트"란 Cas13 CRISPR-Cas 시스템 또는 복합체 또는 가이드가 세포 내에서 선택된 시간 또는 위치로 전달되어서, Cas13 CRISPR-Cas 시스템 또는 복합체 또는 가이드의 활성이 공간적으로 또는 시간적으로 제어되는 것을 의미한다. 예를 들어, Cas13 CRISPR-Cas 시스템 또는 복합체 또는 가이드의 활성 및 목적은 압타머 리간드, 예컨대 세포 표면 단백질 또는 다른 국재화된 세포 성분에 대해 결합 친화성을 갖는 에스코트 RNA 압타머 서열에 의해 제어될 수 있다. 대안적으로, 에스코트 압타머는 예를 들어 세포 상 또는 세포 내 압타머 이펙터, 예컨대 일시적 이펙터, 예컨대 특정한 시기에 세포에 적용되는 외부 에너지원에 반응성일 수 있다.

[0210] 에스코트된 Cas13 CRISPR-Cas 시스템 또는 복합체는 가이드 분자 구조, 아키텍처, 안정성, 유전자 발현, 또는 이의 임의 조합을 개선시키도록 디자인된 기능성 구조를 갖는 가이드 분자를 갖는다. 이러한 구조는 압타머를 포함할 수 있다.

[0211] 압타머는 예를 들어 기하급수적 농축을 통한 리간드의 체계적 진화 (systematic evolution of ligands by exponential enrichment)라고 불리는 기술을 사용하여, 다른 리간드에 단단하게 결합되도록 디자인되거나 또는 선택된 생물분자이다 (SELEX; Tuerk C, Gold L: "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." Science 1990, 249:505-510). 핵산 압타머는 예를 들어 생물의학적으로 관련된 광범위한 표적에 대해 높은 결합 친화성 및 특이성을 갖는, 무작위-서열 올리고뉴클레오티드의 풀로부터 선택할 수 있어서, 압타머에 대한 광범위한 치료적 활용성을 시사한다 (Keefe, Anthony D., Supriya Pai, and Andrew Ellington. "Aptamers as therapeutics." Nature Reviews Drug Discovery 9.7 (2010): 537-550). 이들 특징은 또한 약물 전달 비히클로서 압타머에 대한 광범위한 용도를 시사한다 (Levy-Nissenbaum, Etgar, et al. "Nanotechnology and aptamers: applications in drug delivery." Trends in biotechnology 26.8 (2008): 442-449; and, Hicke BJ, Stephens AW. "Escort aptamers: a delivery service for diagnosis and therapy." J Clin Invest 2000, 106:923-928.). 압타머는 또한 녹색 형광 단백질의 활성을 모방하도록 형광단에 결합하는 RNA 압타머와 같이, 속성을 변화시켜 큐 (que)에 반응하는, 분자 스위치로서 기능하게 구축될 수 있다 (Paige, Jeremy S., Karen Y. Wu, and Samie R. Jaffrey. "RNA mimics of green fluorescent protein." Science 333.6042 (2011): 642-646). 압타머는 예를 들어 세포 표면 단백질을 표적화하는, 표적화된 siRNA 치료제 전달 시스템의 성분으로서 사용될 수 있다는 것이 또한 제안되었다 (Zhou, Jiehua, and John J. Rossi. "Aptamer-targeted cell-specific RNA interference." Silence 1.1 (2010): 4).

[0212] 따라서, 특정 실시형태에서, 가이드 분자는 예를 들어 세포막을 통해서, 세포내 구획, 또는 핵으로의 전달을 포함하여, 가이드 분자 전달을 개선시키도록 디자인된 하나 이상의 압타머(들)를 통해 변형된다. 이러한 구조는 가이드 분자를 선택된 이펙터에 대해 전달가능하거나, 유도가능하거나 또는 반응성이도록 만들기 위해서, 하나 이상의 압타머(들)의 첨가 또는 이러한 하나 이상의 압타머(들), 모이어티(들)없이 포함될 수 있다. 따라서 본 발명은 제한없이 pH, 저산소, O_2 농도, 온도, 단백질 농도, 효소 농도, 지질 구조, 광노출, 기계적 파괴 (예를 들어, 초음파), 자기장, 전기장 또는 전자기 방사선을 포함한 정상 또는 병적 생리학적 조건에 반응하는 가이드 분자를 이해한다.

[0213] 유도성 시스템의 광 반응성은 크립토크롬-2 및 CIB1의 활성화 및 결합을 통해 획득될 수 있다. 파란색 빛 자극은 크립토크롬-2의 활성화 입체형태 변화를 유도하여서, 이의 결합 파트너 CIB1의 동원을 야기시킨다. 이러한 결합은 빠르고 가역적이어서, 펄스된 자극 후에 <15초의 포화를 획득하고 자극 종료 후 <15분에 기준치로 복귀된다. 이들 신속한 결합 동역학은 유도제의 흡수 및 청소보다는, 전사/번역 및 전사물/단백질 분해의 속도에 의해서만 시간적으로 제한되는 시스템을 야기시킨다. 크립토크롬-2 활성화는 또한 고도로 민감하여서, 낮은 및 강도 자극의 사용을 허용하고 광독성 위험성을 완화시킨다. 또한, 예컨대 온전한 포유동물 뇌의 경우에, 다양한 빛 광도를 사용하여 자극 영역의 크기를 제어하여서, 벡터 전달 단독으로 제공될 수 있는 것보다 더 큰 정밀도를 허용한다.

[0214] 본 발명은 가이드를 유도시키기 위해서 에너지원, 예컨대 전자기 방사선, 소리 에너지 또는 열 에너지를 고려한다. 유리하게, 전자기 방사선은 가시광선의 성분이다. 바람직한 실시형태에서, 빛은 약 450 내지 약 495 nm 파장의 파란색 빛이다. 특히 바람직한 실시형태에서, 파장은 약 488 nm이다. 다른 바람직한 실시형태에서, 광 자극은 펄스를 통한다. 빛 동력은 약 0-9 mW/cm2 범위일 수 있다. 바람직한 실시형태에

서, 15초마다 0.25초 만큼 낮은 자극 파라다임이 최대 활성화를 야기시켜야 한다.

[0215] 화학적 또는 에너지 민감성 가이드는 가이드로서 작용하고 Cas13 CRISPR-Cas 시스템 또는 복합체 기능을 갖도록 허용하는 에너지 또는 화학적 소스에 의한 결합을 통한 유도 시에 입체형태적 변화를 겪을 수 있다. 본 발명은 가이드 기능 및 Cas13 CRISPR-Cas 시스템 또는 복합체 기능을 갖도록 화학적 소스 또는 에너지를 인가하는 단계; 및 임의로는 게놈 유전자좌의 발현이 변경된다는 것을 결정하는 추가 단계를 포함할 수 있다.

[0216] 이러한 화학적 유도성 시스템의 몇몇 상이한 디자인이 존재한다: 1. 압시스산 (ABA)에 의해 유도가능한 ABI-PYL 기반 시스템 (예를 들어, http://stke.sciencemag.org/cgi/content/abstract/sigtrans;4/164/rs2 참조), 2. 라파마이신 (또는 라파마이신 기반의 관련 화학물)에 의해 유도가능한 FKBP-FRB 기반 시스템 (예를 들어, http://www.nature.com/nmeth/journal/v2/n6/full/nmeth763.html 참조), 3. 지베렐린 (GA)에 의해 유도가능한 GID1-GAI 기반 시스템 (예를 들어, http://www.nature.com/nchembio/journal/v8/n5/full/nchembio.922.html 참조).

[0217] 화학적 유도성 시스템은 4-히드록시타목시펜 (4OHT) (예를 들어, http://www.pnas.org/content/104/3/1027.abstract 참조)에 의해 유도가능한 에스트로겐 수용체 (ER) 기반 시스템일 수 있다. ERT2라고 하는 에스트로겐 수용체의 돌연변이된 리간드-결합 도메인은 4-히드록시타목시펜의 결합 시 세포의 핵으로 전위된다. 본 발명의 추가 실시형태에서 임의의 핵 수용체, 갑상선 호르몬 수용체, 레티노산 수용체, 에스트로겐 수용체, 에스트로겐-관련 수용체, 글루코코르티코이드 수용체, 프로게스테론 수용체, 안드로겐 수용체의 임의의 천연 발생 또는 조작된 유도체가 ER 기반 유도성 시스템과 유사한 유도성 시스템에서 사용될 수 있다.

[0218] 다른 유도성 시스템은 에너지, 열, 또는 전파에 의해 유도가능한 일시적 수용체 전위 (TRP) 이온 채널 기반 시스템을 사용하는 디자인을 기반으로 한다 (예를 들어, http://www.sciencemag.org/content/336/6081/604 참조). 이들 TRP 패밀리 단백질은 빛과 열을 포함한, 상이한 자극에 반응한다. 이러한 단백질은 빛 또는 열에 의해 활성화될 때, 이온 채널이 열리게 되어 칼슘 이온과 같은 이온의 형질막으로의 진입을 허용한다. 이러한 이온 유입은 가이드 및 Cas13 CRISPR-Cas 복합체 또는 시스템의 다른 성분을 포함한 폴리펩티드에 연결된 세포내 이온 상호작용 파트너에 결합하게 될 것이고 이러한 결합은 폴리펩티드의 세포하 국재화의 변화를 유도하여, 전체 폴리펩티드가 세포의 핵으로 진입하게 한다. 핵 내부에 존재하게 되면, 가이드 단백질 및 Cas13 CRISPR-Cas 복합체의 다른 성분이 활성화되어 세포에서 표적 유전자 발현을 조절하게 될 것이다.

[0219] 빛 활성화가 유리한 실시형태일 수 있지만, 때때로 빛이 피부 또는 다른 장기를 침투할 수 없는 생체내 적용에서는 특히 불리할 수 있다. 이러한 예에서, 에너지 활성화의 다른 방법, 특히 유사한 효과를 갖는 전기장 에너지 및/또는 초음파가 고려된다.

[0220] 전기장 에너지는 바람직하게는 생체내 조건 하에서 약 1 Volt/cm 내지 약 10 kVolts/cm의 하나 이상의 전기 펄스를 사용하여, 당분야에 기술된 바와 같이, 실질적으로 투여된다. 펄스 대신에 또는 그이외에도, 전기장은 연속적인 방식으로 전달될 수 있다. 전기 펄스는 1 μs 내지 500 밀리초, 바람직하게 1 μs 내지 100 밀리초 동안 인가될 수 있다. 전기장은 약 5분 동안 펄스식 방식으로 또는 연속적으로 인가될 수 있다.

[0221] 본 명세서에서 사용되는 "전기장 에너지"는 세포가 노출되는 전기 에너지이다. 바람직하게, 전기장은 생체내 조건 하에서 약 1 Volt/cm 내지 약 10 kVolts/cm 이상의 강도를 갖는다 (WO97/49450 참조).

[0222] 본 명세서에서 사용되는 용어 "전기장"은 가변 용량 및 전압에서 하나 이상의 펄스를 포함하고, 지수파 및/또는 사각파 및/또는 변조파 및/또는 변조 사각파 형태를 포함한다. 전기장 및 전기장은 세포 환경에서 전기적 전위차의 존재에 대한 참조를 포함하도록 언급되어야 한다. 이러한 환경은 당업계에 공지된 바와 같은 정전기, 교류 (AC), 직류 (DC)에 의해 셋업될 수 있다. 전기장은 균일, 비균일 등일 수 있으며, 시간 의존적 방식으로 강도 및/또는 방향이 다를 수 있다.

[0223] 전기장의 단회 또는 다회 인가뿐만 아니라 초음파의 단회 또는 다회 인가는 임의의 순서로 그리고 임의의 조합으로 가능하다. 초음파 및/또는 전기장은 단회 또는 다회 연속 인가로서, 또는 펄스 (박동성 전달)로서 전달될 수 있다.

[0224] 외래 물질을 살아있는 세포 내로 도입하기 위해 시험관내 절차와 생체내 절차 둘 모두에서 전기천공법이 사용되어 왔다. 시험관내 적용에 의해, 생존 세포의 샘플은 먼저 관심의 작용제와 혼합되고, 전극, 예컨대, 평행판 사이에 위치된다. 이어서, 전극은 전기장을 세포/이식 혼합물에 인가된다. 시험관내 전기천공법을 수행하는 시스템의 예는 Electro Cell Manipulator ECM600 제품, 및 Electro Square Porator T820를 포함하며,

이들은 Genetronics, Inc의 BTX Division에서 제조되었다 (미국 특허 제5,869,326호 참조).

[0225] 공지된 전기천공법 기법은 (시험관내 및 생체내) 처리 영역 주변에 위치된 전극에 짧은 고전압 펄스를 인가함으로써 작용한다. 전극 사이에서 발생된 전기장은 세포막을 일시적으로 다공성이 되게 하여서, 그때에 관심 작용제의 분자가 세포로 진입한다. 공지된 전기천공법 적용에서, 전기장은 약 100 μs 지속기간의, 1000V/cm 규모로 단일 사각파 펄스를 포함한다. 이러한 펄스는, 예를 들어, Electro Square Porator T820의 공지된 어플리케이션에서 발생시킬 수 있다.

[0226] 바람직하게는, 전기장은 시험관내 조건 하에서 약 1 V/cm 내지 약 10 kV/cm의 강도를 갖는다. 따라서, 전기장은 시험관내 조건 하에서 1 V/cm, 2 V/cm, 3 V/cm, 4 V/cm, 5 V/cm, 6 V/cm, 7 V/cm, 8 V/cm, 9 V/cm, 10 V/cm, 20 V/cm, 50 V/cm, 100 V/cm, 200 V/cm, 300 V/cm, 400 V/cm, 500 V/cm, 600 V/cm, 700 V/cm, 800 V/cm, 900 V/cm, 1 kV/cm, 2 kV/cm, 5 kV/cm, 10 kV/cm, 20 kV/cm, 50 kV/cm 이상, 보다 바람직하게는 약 0.5 kV/cm 내지 약 4.0 kV/cm의 강도를 갖는다. 바람직하게는 전기장은 생체내 조건 하에서 약 1 V/cm 내지 약 10 kV/cm의 강도를 갖는다. 그러나, 표적 부위에 전달된 펄스 수가 증가되는 경우에 전기장 강도는 낮아질 수 있다. 따라서, 더 낮은 전기장 강도에서 전기장의 박동성 전달이 계획된다.

[0227] 바람직하게는 전기장의 인가는 동일한 강도 및 용량의 이중 펄스 또는 다양한 강도 및/또는 용량의 순차적 펄스와 같은 다중 펄스 형태이다. 본 명세서에서 사용되는 용어 "펄스"는 가변 용량 및 전압에서 하나 이상의 전기적 펄스를 포함하고, 지수 및/또는 사각파 및/또는 변조파/사각파 형태를 포함한다.

[0228] 바람직하게는 전기적 펄스는 지수 파형, 사각 파형, 변조 파형으로부터 선택된 파형으로서 전달된다.

[0229] 바람직한 실시형태는 저전압에서 직류를 사용한다. 따라서, 출원인은 1 V/cm 내지 20 V/cm의 전기장 강도에서, 100 밀리초 이상, 바람직하게는 15분 이상의 기간 동안 세포, 조직 또는 조직 덩어리에 인가되는 전기장의 사용을 개시한다.

[0230] 초음파는 유리하게는 약 0.05 W/㎠ 내지 약 100 W/㎠의 전력 수준으로 투여된다. 진단 또는 치료 초음파, 또는 이들의 조합이 사용될 수 있다.

[0231] 본 명세서에서 사용되는 용어 "초음파"는 기계적 진동으로 이루어지고, 이의 주파수는 너무 높아서 인간 청력 범위 이상인 에너지의 형태를 의미한다. 초음파 스펙트럼의 하한 주파수는 일반적으로 약 20 kHz로서 취해질 수 있다. 초음파의 대부분의 진단적 적용은 1 내지 15 ㎒' 범위의 주파수를 사용한다 (Ultrasonics in Clinical Diagnosis, P. N. T. Wells, ed., 2nd. Edition, Publ. Churchill Livingstone [Edinburgh, London & NY, 1977]).

[0232] 초음파는 진단 및 치료 적용분야 둘 모두에서 사용되었다. 진단 도구로서 사용될 때 ("진단적 초음파"), 최대 750 mW/㎠의 에너지 밀도가 사용되었지만, 초음파는 전형적으로 최대 약 100 mW/㎠ (FDA 권장사항)의 에너지 밀도 범위에서 사용된다. 물리치료에서, 초음파는 전형적으로 약 3 내지 4 W/㎠ (WHO 권장사항) 범위에서 에너지 공급원으로서 사용된다. 다른 치료적 적용분야에서, 더 높은 강도의 초음파는 단기간 동안, 예를 들어, 100 W/cm 내지 1 kW/㎠ (또는 훨씬 높은)에서의 HIFU가 적용될 수 있다. 본 명세서에서 사용되는 용어 "초음파"는 진단, 치료 및 집속 초음파를 포함하는 것으로 의도된다.

[0233] 집속 초음파 (Focused ultrasound: FUS)는 비침습 프로브 없이 열 에너지를 전달할 수 있게 한다 ([Morocz et al 1998 Journal of Magnetic Resonance Imaging Vol.8, No. 1, pp.136-142] 참조). 집속 초음파의 다른 형태는 [Moussatov et al in Ultrasonics (1998) Vol.36, No.8, pp.893-900] 및 [TranHuuHue et al in Acustica (1997) Vol.83, No.6, pp.1103-1106]에서 고찰된 고강도 집속 초음파 (high intensity focused ultrasound: HIFU)이다.

[0234] 바람직하게는, 진단 초음파와 치료 초음파의 조합이 사용된다. 그러나, 이 조합은 제한하려는 의도가 아니고, 당업자는 초음파의 임의의 다양한 조합이 사용될 수 있다는 것을 인식할 것이다. 추가적으로, 에너지 밀도, 초음파 주파수 및 노출 시간은 변할 수 있다.

[0235] 바람직하게는 초음파 에너지원에 대한 노출은 약 0.05 내지 약 100 Wcm^{-2}의 전력 밀도에서이다. 훨씬 더 바람직하게는, 초음파 에너지원에 대한 노출은 약 1 내지 약 15 Wcm^{-2}의 전력 밀도에서이다.

[0236] 바람직하게는 초음파 에너지원에 대한 노출은 약 0.015 내지 약 10.0 ㎒의 주파수에서이다. 더 바람직하게는 초음파 에너지원에 대한 노출은 약 0.02 내지 약 5.0 ㎒ 또는 약 6.0 ㎒의 주파수에서이다. 가장 바람직

하게는, 초음파는 3 ㎒의 주파수에서 인가된다.

[0237] 바람직하게는 노출은 약 10 밀리초 내지 약 60분의 기간 동안이다. 바람직하게는 노출은 약 1초 내지 약 5분의 기간 동안이다. 더 바람직하게는, 초음파는 약 2분 동안 인가된다. 그러나, 붕괴하려는 특정 표적 세포에 따라서, 노출은 더 긴 지속기간, 예를 들어, 15분 동안일 수 있다.

[0238] 유리하게, 표적 조직이 약 0.0 5Wcm^{-2} 내지 약 10 Wcm^{-2}의 음향 출력 밀도에서 약 0.015 내지 약 10㎒ 범위의 주파수로 초음파 에너지원에 노출된다 (WO 98/52609 참조). 그러나, 대안, 예를 들어, 100 Wcm^{-2} 초과의 음향 출력 밀도에서, 그러나 감소된 시간 기간 동안, 예를 들어, 밀리초 범위 이하의 기간 동안 1000 Wcm^{-2}에서 초음파 에너지에 대한 노출이 또한 가능하다.

[0239] 바람직하게는 초음파의 인가는 다중 펄스 형태이고, 따라서, 지속파와 펄스파 (초음파 맥동성 전달)는 둘 모두 임의의 조합으로 사용될 수 있다. 예를 들어, 지속파 초음파, 다음에 펄스파 초음파 또는 그 반대로 적용될 수 있다. 이는 임의의 횟수, 임의의 순서 및 조합으로 반복될 수 있다. 펄스파 초음파는 지속파 초음파의 배경에 대해 적용될 수 있고, 임의의 펄스 수가 임의의 그룹 수에서 사용될 수 있다.

[0240] 바람직하게는, 초음파는 펄스파 초음파를 포함할 수 있다. 고도로 바람직한 실시형태에서, 초음파는 지속파로서 0.7 Wcm-2 또는 1.25 Wcm-2의 출력 밀도로 적용된다. 펄스파 초음파가 사용된다면, 더 높은 출력 밀도가 사용될 수 있다.

[0241] 초음파의 사용은 표적 상에서 정확하게 집속시킬 수 있기 때문에 빛만큼 유리하다. 게다가, 초음파는 빛과 달리 조직에 더 깊게 집속시킬 수 있기 때문에 유리하다. 따라서 전체 조직 침투 (예컨대, 이에 제한없이, 간엽) 또는 전체 장기 (예컨대, 제한없이, 전체 간 또는 전체 근육, 예컨대, 심장) 요법에 더 적합하게 된다.
다른 중요한 이점은 초음파가 매우 다양한 진단적 및 치료적 적용분야에서 사용되는 비침습성 자극이라는 것이다. 예로서, 초음파는 의학적 영상화 기법에서, 추가적으로 정형외과적 요법에서 잘 공지되어 있다. 더 나아가, 대상 척추동물에 대한 초음파의 적용에 적합한 기기가 널리 이용 가능하며, 그들의 용도는 당업계에 충분히 공지되어 있다.

[0242] 특정 실시형태에서, 가이드 분자는 CRISPR-Cas 시스템의 특이성을 증가시키기 위해 2차 구조에 의해 변형되고, 그 2차 구조는 엑소뉴클레아제 활성에 대해 보호될 수 있고, 본 명세서에서 또한 보호된 가이드 분자라고도 언급되는 가이드 서열에 대한 5' 첨가를 허용한다.

[0243] 일 양상에서, 본 발명은 가이드 분자의 서열과 "보호자 RNA"를 하이브리드화하기 위해 제공되고, 여기서 "보호자 RNA"는 가이드 분자의 3' 말단에 상보성이어서 부분적으로 이중-가닥인 가이드 RNA를 생성시키는 RNA 가닥이다. 본 발명의 일 실시형태에서, 완벽하게 상보성인 보호자 서열로 미스매치된 염기 (즉, 가이드 서열의 일부를 형성하지 않는 가이드 분자의 염기)를 보호하는 것은 3' 말단에서 미스매치된 염기쌍과 표적 DNA 결합 가능성을 감소시킨다. 본 발명의 특정 실시형태에서, 연장된 길이를 포함하는 추가 서열은 가이드가 가이드 분자 내에 보호자 서열을 포함하도록 가이드 분자 내에 존재할 수 있다. 이러한 "보호자 서열"은 가이드 분자가 "노출된 서열" (표적 서열과 하이브리드화되는 가이드 서열의 일부 포함) 이외에도 "보호된 서열"을 포함하는 것을 보장한다. 특정 실시형태에서, 가이드 분자는 2차 구조 예컨대 헤어핀을 포함하도록 보호자 가이드의 존재에 의해 변형된다. 유리하게, 보호된 서열, 가이드 서열 또는 둘 모두에 대해 상보성을 갖는 3 또는 4 내지 30 이상, 예를 들어 약 10 이상의 인접한 염기쌍이 존재한다. 보호된 부분은 이의 표적과 상호작용하는 CRISPR-Cas 시스템의 열역학을 방해하지 않는 것이 유리하다. 부분적으로 이중 가닥인 가이드 분자를 포함한 이러한 연장부를 제공하여, 가이드 분자가 보호되는 것을 고려하고 그 결과로 비활성을 유지하면서, CRISPR-Cas 복합체의 개선된 비결합을 야기시킨다.

[0244] 특정 실시형태에서, 절단된 가이드 (tru-가이드), 즉 정규 가이드 서열 길이에 비해서 길이가 절단된 가이드 서열을 포함하는 가이드 분자가 이용된다. Nowak 등 (Nucliec Acids Res (2016) 44 (20): 9555-9564)이 기술한 바와 같이, 이러한 가이드는 표적 DNA를 절단하지 않고 촉매적 활성 CRISPR-Cas 효소가 이의 표적에 결합할 수 있게 한다. 특정 실시형태에서, 절단된 가이드는 CRISPR-Cas 효소의 닉카제 활성만을 보유하면서 표적의 결합을 허용하는 것이 사용된다.

[0245] *Crispr-Cas 효소*

[0246] 이의 비변형된 형태에서, CRISPR-Cas 단백질은 촉매적 활성 단백질이다. 이것은 핵산-표적화 복합체 (표적 서열과 하이브리드화되는 가이드 RNA 포함)의 형성 시에 표적 서열 내 또는 그 근처 (예를 들어, 그로부터 1,

2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 염기쌍 이상 내) DNA 가닥 중 하나 또는 둘 모두가 변형 (예를 들어, 절단)된다는 것을 의미한다. 본 명세서에서 사용되는 용어 "관심 표적 유전자좌와 연관된 서열(들)"은 표적 서열 부근 근처 (예를 들어, 표적 서열로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50개 이상의 염기쌍 내이고, 표적 서열은 관심 표적 유전자좌 내에 포함됨)의 서열을 의미한다. 비변형된 촉매적 활성 Cas13 단백질은 스태거드 절단부를 생성시키고, 그 절단 부위는 전형적으로 표적 서열 내에 있다. 보다 특히, 스태거드 절단부는 전형적으로 PAM에 대해 13-23개 뉴클레오티드 멀리있다. 특정 실시형태에서, 비표적 가닥 상의 절단부는 PAM 하류 17개 뉴클레오티드 (즉, PAM 하류의 뉴클레오티드 17 내지 18)에 있는 한편, 표적 가닥 (즉, 가이드 서열과 하이브리드화하는 가닥) 상의 절단은 PAM에 상보성인 서열로부터 멀리 4개 뉴클레오티드만큼 멀리에서 발생된다 (이것은 3' 가닥 상의 PAM의 상보체의 상류 21개 뉴클레오티드 또는 PAM의 상보체의 상류 뉴클레오티드 21개 내지 22개이다).

[0247] 본 발명에 따른 방법에서, CRISPR-Cas 단백질은 바람직하게는 상응하는 야생형 효소에 대해 돌연변이되어서 돌연변이된 CRISPR-Cas 단백질이 표적 서열을 함유하는 표적 유전자좌의 한쪽 또는 양쪽 DNA 가닥을 절단하는 능력이 결여된다. 특정 실시형태에서, Cas13 단백질의 하나 이상의 촉매 도메인은 표적 서열의 오직 하나의 DNA 가닥을 절단하는 돌연변이된 Cas 단백질을 생성시키도록 돌연변이된다.

[0248] 특정 실시형태에서, CRISPR-Cas 단백질은 상응하는 야생형 효소에 대해 돌연변이될 수 있어서 돌연변이된 CRISPR-Cas 단백질은 실질적으로 모든 DNA 절단 활성이 결여된다. 일부 실시형태에서, CRISPR-Cas 단백질은 돌연변이된 효소의 절단 활성이 효소의 비돌연변이된 형태의 핵산 절단 활성의 약 25%, 10%, 5%, 1%, 0.1%, 0.01% 이하일 때, 모든 DNA 및/또는 RNA 절단 활성이 실질적으로 결여된 것으로 간주될 수 있고, 예로는 돌연변이된 형태의 핵산 절단 활성이 비돌연변이된 형태와 비교하여 무시할만하거나 또는 전무할 때일 수 있다.

[0249] 본 명세서에서 제공되는 방법의 일정 실시형태에서, CRISPR-Cas 단백질은 오직 하나의 DNA 가닥을 절단하는 돌연변이된 CRISPR-Cas 단백질, 즉 닉카제이다. 보다 특히, 본 발명에서, 닉카제는 비-표적 서열, 즉 표적 서열의 반대쪽 DNA 가닥 상에 있고 PAM 서열의 3'에 있는 서열 내에서 절단을 보장한다. 추가 지침으로서, 제한없이, 아시다미노코커스 sp. 유래 Cas13의 Nuc 도메인에서 아르기닌-아르기닌 (arginine-to-alanine) 치환 (R1226A)은 Cas13을 양쪽 가닥을 절단하는 뉴클레아제에서 (단일 가닥을 절단하는) 닉카제로 전환시킨다. 효소가 AsCas13이 아닌 경우에, 돌연변이는 상응하는 위치 내 잔기에서 만들어질 수 있다는 것을 당업자는 이해하게 될 것이다. 특정 실시형태에서, Cas13은 FnCas13이고 돌연변이는 위치 R1218의 아르기닌에서 존재한다. 특정 실시형태에서, Cas13은 LbCas13이고 돌연변이는 위치 R1138의 아르기닌에서 존재한다. 특정 실시형태에서, Cas13은 MbCas13이고 돌연변이는 위치 R1293의 아르기닌에서 존재한다.

[0250] 본 명세서에서 제공되는 방법의 일정 실시형태에서, CRISPR-Cas 단백질은 촉매 활성이 감소되었거나 또는 전무하다. CRISPR-Cas 단백질이 Cas13 단백질인 경우에, 돌연변이는 제한없이 AsCas13의 위치에 대해서 촉매적 RuvC-유사 도메인 내 하나 이상의 돌연변이, 예컨대 D908A 또는 E993A를 포함할 수 있다.

[0251] 일부 실시형태에서, CRISPR-Cas 단백질은 돌연변이된 효소의 DNA 절단 활성이 효소의 비돌연변이된 DNA 절단 활성의 약 25%, 10%, 5%, 1%, 0.1%, 0.01% 이하일 때 모든 DNA 절단 활성이 실질적으로 결여된 것으로 간주되고, 예로는 돌연변이된 형태의 DNA 절단 활성이 비돌연변이된 형태와 비교하여 무시할만하거나 또는 전무한 경우일 수 있다. 이들 실시형태에서, CRISPR-Cas 단백질은 포괄적인 DNA 결합 단백질로서 사용된다. 돌연변이는 인공적으로 도입된 돌연변이일 수 있거나 또는 기능 획득 또는 기능 상실 돌연변이일 수 있다.

[0252] 상기 기술된 돌연변이이외에도, CRISPR-Cas 단백질은 추가적으로 변형될 수 있다. 본 명세서에서 사용되는, CRISPR-Cas 단백질에 대한 용어 "변형된"은 일반적으로 그것이 유래된 야생형 Cas 단백질과 비교하여 하나 이상의 변형 또는 돌연변이 (점 돌연변이, 절단, 삽입, 결실, 키메라, 융합 단백질 등 포함)를 갖는 CRISPR-Cas 단백질을 의미한다. 유래된 이란 유래된 효소가 대체로 높은 서열 상동성 정도를 갖는다는 의미에서, 야생형 효소를 기반으로 하지만, 당분야에 공지되거나 또는 본 명세서에 기술된 바와 같이 일부 방식으로 돌연변이 (변형)되었다는 것을 의미한다.

[0253] 일부 실시형태에서, Cas13b 이펙터 및 하나 이상의 기능성 도메인의 융합 단백질의 크기를 축소시키기 위해서, Cas13b 이펙터의 C-말단은 이의 RNA 결합 기능은 여전히 유지하면서 절단될 수 있다. 예를 들어, 적어도 20개 아미노산, 적어도 50개 아미노산, 적어도 80개 아미노산, 또는 적어도 100개 아미노산, 또는 적어도 150개 아미노산, 또는 적어도 200개 아미노산, 또는 적어도 250개 아미노산, 또는 적어도 300개 아미노산, 또는 적어도 350개 아미노산, 또는 최대 120개 아미노산, 또는 최대 140개 아미노산, 또는 최대 160개 아미노산, 또는 최대 180개 아미노산, 또는 최대 200개 아미노산, 또는 최대 250개 아미노산, 또는 최대 300개 아미노산, 또는 최

대 350개 아미노산, 또는 최대 400개 아미노산이 Cas13b 이펙터의 C-말단에서 절단될 수 있다. Cas13b 절단의 특별한 예는 C-말단 Δ984-1090, C-말단 Δ1026-1090, 및 C-말단 Δ1053-1090, C-말단 Δ934-1090, C-말단 Δ884-1090, C-말단 Δ834-1090, C-말단 Δ784-1090, 및 C-말단 Δ734-1090을 포함하고, 여기서 아미노산 위치는 프레보텔라 sp. P5-125 Cas13b 단백질의 아미노산 위치에 해당된다. 도 67을 또한 참조한다.

[0254] CRISPR-Cas 단백질의 추가 변형은 변경된 기능성을 야기할 수도 있거나 또는 그렇지 않을 수도 있다. 예로서, 특히 CRISPR-Cas 단백질을 참조하여, 변경된 기능성을 야기하지 않는 변형은 예를 들어 특정 숙주에서 발현을 위한 코돈 최적화, 또는 뉴클레아제에 특정 마커 (예를 들어, 가시화용)를 제공하는 것을 포함한다. 변경된 기능성을 야기시킬 수 있는 변형은 또한 점 돌연변이, 삽입, 결실, 절단 (분할 뉴클레아제 포함) 등을 포함하는, 돌연변이를 포함할 수 있다. 융합 단백질은 제한없이 예를 들어 이종성 도메인 또는 기능성 도메인 (예를 들어, 국재화 신호, 촉매 도메인 등)과의 융합을 포함할 수 있다. 일정 실시형태에서, 다양한 상이한 변형은 돌연변이, 결실, 삽입, 치환, 결찰, 분해, 파괴 또는 재조합을 조합할 수 있다 (예를 들어, 예컨대 제한없이 (예를 들어, 상이한 뉴클레아제 (도메인)에 의한) 파괴를 포함하여, 예컨대 예를 들어 DNA 메틸화 또는 다른 핵산 변형을 유도하도록, 기능성 도메인과 더욱 융합되고 촉매적 불활성인 돌연변이된 뉴클레아제). 본 명세서에서 사용되는, "변경된 기능성"은 제한없이 변경된 특이성 (예를 들어, 변경된 표적 인식, 증가 (예를 들어, "증강된" Cas 단백질) 또는 감소된 특이성, 또는 변경된 PAM 인식), 변경된 활성 (예를 들어, 촉매적 불활성 뉴클레아제 또는 닉카제를 포함하는 증가 또는 감소된 촉매 활성), 및/또는 변경된 안정성 (예를 들어, 탈안정화 도메인과의 융합)을 포함한다. 적합한 이종성 도메인은 제한없이 뉴클레아제, 리가제, 복구 단백질, 메틸트랜스퍼라제, (바이러스) 인테그라제, 리콤비나제, 트랜스포사제, 아르고노트, 시티딘 데아미나제, 레트론, 그룹 II 인트론, 포스파타제, 포스포릴라제, 술프퓨릴라제, 키나제, 폴리머라제, 엑소뉴클레아제 등을 포함한다. 모든 이들 변형의 예는 당업자에게 공지되어 있다. 본 명세서에서 언급되는 "변형된" 뉴클레아제, 특히 "변형된" Cas 또는 "변형된" CRISPR-Cas 시스템 또는 복합체는 바람직하게는 (예를 들어, 가이드 분자와 복합체인) 폴리핵산과 상호작용하거나 또는 그와 결합하는 능력을 여전히 갖는다는 것을 이해하게 될 것이다. 이러한 변형된 Cas 단백질은 본 명세서에 기술된 바와 같은 데아미나제 단백질 또는 이의 활성 도메인과 조합될 수 있다.

[0255] 일정 실시형태에서, CRISPR-Cas 단백질은 예컨대 표적화 또는 비표적화 가닥을 안정화시키는 돌연변이된 잔기를 포함하여, 증강된 활성 및/또는 특이성을 야기시키는 하나 이상의 변형을 포함할 수 있다 (예를 들어, 그 전문이 참조로 본 명세서에 편입되는, [eCas9; "Rationally engineered Cas9 nucleases with improved specificity", Slaymaker et al. (2016), Science, 351(6268):84-88] 참조). 일정 실시형태에서, 조작된 CRISPR 단백질의 변경 또는 변형 활성은 증가된 표적화 효율 또는 감소된 오프-표적 결합을 포함한다. 일정 실시형태에서, 조작된 CRISPR 단백질의 변경된 활성은 변형된 절단 활성을 포함한다. 일정 실시형태에서, 변경된 활성은 표적 폴리뉴클레오티드 유전자좌에 대해 증가된 변형된 절단 활성을 포함한다. 일정 실시형태에서, 변경된 활성은 표적 폴리뉴클레오티드 유전자좌에 대해 감소된 변형된 절단 활성을 포함한다. 일정 실시형태에서, 변경된 활성은 오프-표적 폴리뉴클레오티드 유전자좌에 대해 감소된 변형된 절단 활성을 포함한다. 일정 실시형태에서, 변형된 뉴클레아제의 변경 또는 변형된 활성은 변경된 헬리카제 동역학을 포함한다. 일정 실시형태에서, 변형된 뉴클레아제는 RNA (Cas 단백질의 경우)를 포함하는 핵산 분자, 또는 표적 폴리뉴클레오티드 유전자좌의 가닥, 또는 오프-표적 폴리뉴클레오티드 유전자좌의 가닥과 단백질의 회합을 변경시키는 변형을 포함한다. 본 발명의 일 양상에서, 조작된 CRISPR 단백질은 CRISPR 복합체의 형성을 변경시키는 변형을 포함한다. 일정 실시형태에서, 변경된 활성은 오프-표적 폴리뉴클레오티드 유전자좌에 대해 증가된 변형된 절단 활성을 포함한다. 따라서, 일정 실시형태에서, 오프-표적 폴리뉴클레오티드 유전자좌와 비교하여 표적 폴리뉴클레오티드 유전자좌에 대한 증가된 특이성이 존재한다. 다른 실시형태에서, 오프-표적 폴리뉴클레오티드 유전자좌와 비교하여 표적 폴리뉴클레오티드 유전자좌에 대해 감소된 특이성이 존재한다. 일정 실시형태에서, 돌연변이는 예컨대 Cas 단백질이 예를 들어 표적과 가이드 RNA 간 미스매치에 대해 낮아진 내성이 야기된 경우에, 감소된 오프-표적 효과 (예를 들어, 절단 또는 결합 속성, 활성, 또는 동역학)를 야기한다. 다른 돌연변이는 증가된 오프-표적 효과 (예를 들어, 절단 또는 결합 성질, 활성 또는 동역학)를 야기시킬 수 있다. 다른 돌연변이는 증가되거나 또는 감소된 온-표적 효과 (예를 들어, 절단 또는 결합 성질, 활성 또는 동역학)를 야기시킬 수 있다. 일정 실시형태에서, 돌연변이는 기능성 뉴클레아제 복합체 (예를 들어, CRISPR-Cas 복합체)의 변경 (예를 들어, 증가 또는 감소)된 헬리카제 활성, 회합 또는 형성을 야기시킨다. 일정 실시형태에서, 상기 기술된 바와 같이, 돌연변이는 변경된 PAM 인식을 야기시키며, 즉 상이한 PAM이, 비변형된 Cas 단백질과 비교하여, (추가로 또는 대안적으로) 인식될 수 있다. 특히 바람직한 돌연변이는 특이성을 증강시키기 위해서 양으로 하전된 잔기 및/또는 (진화적으로) 보존된 잔기, 예컨대 보존된

양으로 하전된 잔기를 포함한다. 일정 실시형태에서, 이러한 잔기는 비하전된 잔기, 예컨대 알라닌으로 돌연변이될 수 있다.

[0256] **염기 절제 복구 억제제**

[0257] 일부 실시형태에서, AD-기능화된 CRISPR 시스템은 염기 절제 복구 (base excision repair: BER) 억제제를 추가로 포함한다. 임의의 특정 이론에 국한하고 싶지 않지만, I:T 쌍형성의 존재에 대한 세포의 DNA-복구 반응은 세포 내 핵염기 편집 효율의 감소를 초래할 수 있다. 알킬아데닌 DNA 글리코실라제 (또한 DNA-3-메틸아데닌 글리코실라제, 3-알킬아데닌 DNA 글리코실라제, 또는 N-메틸퓨린 DNA 글리코실라제로서 알려짐)는 세포에서 DNA로부터 하이포잔틴의 제거를 촉매하는데, 이는 결과로서 I:T 쌍의 A:T 쌍으로의 반전과 함께 염기 절제 복구를 개시한다.

[0258] 일부 실시형태에서, BER 억제제는 알킬아데닌 DNA 글리코실라제의 억제제이다. 일부 실시형태에서, BER 억제제는 인간 알킬아데닌 DNA 글리코실라제의 억제제이다. 일부 실시형태에서, BER 억제제는 폴리펩티드 억제제이다. 일부 실시형태에서, BER 억제제는 하이포잔틴에 결합하는 단백질이다. 일부 실시형태에서, BER 억제제는 DNA에서 하이포잔틴에 결합하는 단백질이다. 일부 실시형태에서, BER 억제제는 촉매적 불활성인 알킬아데닌 DNA 글리코실라제 단백질 또는 이의 결합 도메인이다. 일부 실시형태에서, BER 억제제는 DNA로부터 하이포잔틴을 절단하지 않는 촉매적 불활성인 알킬아데닌 DNA 글리코실라제 단백질 또는 이의 결합 도메인이다. 알킬아데닌 DNA 글리코실라제 염기-절단 복구 효소를 억제할 수 있는 (예를 들어, 입체적으로 차단하는) 다른 단백질은 본 개시내용의 범주 내이다. 추가적으로, 염기-절단 복구를 차단 또는 억제하는 임의의 단백질은 본 개시내용의 범주 내이다.

[0259] 임의의 특정 이론으로 구속되는 일 없이, 염기 절제 복구은 편집된 가닥에 결합하고/하거나, 편집 염기를 차단하고/하거나, 알킬아데닌 DNA 글리코실라제를 억제하고/하거나, 염기 절제 복구를 억제하고/하거나, 편집 염기를 보호하고/하거나 비편집 가닥의 고정을 촉진시키는 분자에 의해 억제될 수 있다. 본 명세서에 기재된 BER 억제제의 사용은 A의 I로의 변화를 촉매할 수 있는 아데노신 데아미나제의 편집 효율을 증가시킬 수 있는 것으로 여겨진다.

[0260] 따라서, 상기 논의한 AD-기능화된 CRISPR 시스템의 제1 디자인에서, CRISPR-Cas 단백질 또는 아데노신 데아미나제는 BER 억제제(예를 들어, 알킬아데닌 DNA 글리코실라제의 억제제)에 융합되거나 또는 연결될 수 있다. 일부 실시형태에서, BER 억제제는 다음의 구조 중 하나를 포함할 수 있다 (nCas13=Cas13 닉카제; dCas13=데드 Cas13): [AD]-[선택적 링커]-[nCas13/dCas13]-[선택적 링커]-[BER 억제제]; [AD]-[선택적 링커]-[BER 억제제]-[선택적 링커]-[nCas13/dCas13]; [BER 억제제]-[선택적 링커]-[AD]-[선택적 링커]-[nCas13/dCas13]; [BER 억제제]-[선택적 링커]-[nCas13/dCas13]-[선택적 링커]-[AD]; [nCas13/dCas13]-[선택적 링커]-[AD]-[선택적 링커]-[BER 억제제]; [nCas13/dCas13]-[선택적 링커]-[BER 억제제]-[선택적 링커]-[AD].

[0261] 유사하게는, 상기 논의한 AD-기능화된 CRISPR 시스템의 제2 디자인에서, CRISPR-Cas 단백질, 아데노신 데아미나제 또는 어댑터 단백질이 BER 억제제 (예를 들어, 알킬아데닌 DNA 글리코실라제의 억제제)에 융합되거나 또는 연결될 수 있다. 일부 실시형태에서, BER 억제제는 다음의 구조 중 하나를 포함할 수 있다 (nCas13=Cas13 닉카제; dCas13=데드 Cas13): [nCas13/dCas13]-[선택적 링커]-[BER 억제제]; [BER 억제제]-[선택적 링커]-[nCas13/dCas13]; [AD]-[선택적 링커]-[어댑터]-[선택적 링커]-[BER 억제제]; [AD]-[선택적 링커]-[BER 억제제]-[선택적 링커]-[어댑터]; [BER 억제제]-[선택적 링커]-[AD]-[선택적 링커]-[어댑터]; [BER 억제제]-[선택적 링커]-[어댑터]-[선택적 링커]-[AD]; [어댑터]-[선택적 링커]-[AD]-[선택적 링커]-[BER 억제제]; [어댑터]-[선택적 링커]-[BER 억제제]-[선택적 링커]-[AD].

[0262] 상기 논의한 AD-기능화된 CRISPR 시스템의 제3 디자인에서, BER 억제제는 CRISPR-Cas 단백질의 내부 루프 또는 비구조화된 영역에 삽입될 수 있다.

[0263] **핵으로의 표적화**

[0264] 일부 실시형태에서, 본 발명의 방법은 표적 유전자좌가 세포 내에 있어서, 관심 표적 유전자좌에서 아데닌을 변형시키는 것에 관한 것이다. 본 발명의 방법에서 사용되는 CRISPR-Cas 단백질 및/또는 아데노신 데아미나제 단백질 또는 이의 촉매적 도메인의 핵에 대한 표적화를 개선시키기 위해, 이들 성분 중 하나 또는 둘 모두에 하나 이상의 핵 국재화 서열(NLS)을 제공하는 것이 유리할 수 있다.

[0265] 바람직한 실시형태에서, 본 발명과 관련하여 사용되는 NLS는 단백질에 대해 이종성이다. NLS의 비제한적 예는 하기로부터 유래된 NLS 서열을 포함한다: 아미노산 서열 PKKKRKV (SEQ ID No. 17) 또는 PKKKRKVEAS (SEQ ID

No. 18)을 갖는, SV40 바이러스 거대 T-항원의 NLS; 뉴클레오플라스민으로부터의 NLS(예를 들어, 서열 KRPAATKKAGQAKKKK (SEQ ID No. 19)을 갖는 뉴클레오플라스민 이분형 NLS); 아미노산 서열 PAAKRVKLD (SEQ ID No. 20) 또는 RQRRNELKRSP (SEQ ID No. 21)을 갖는 c-myc NLS; 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID No. 22)을 갖는 hRNPA1 M9 NLS; 임포틴-알파 유래 IBB 도메인의 서열 RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV(SEQ ID No. 23); 근종 T 단백질의 서열 VSRKRPRP (SEQ ID No. 24) 및 PPKKARED (SEQ ID No. 25); 인간 p53의 서열 PQPKKKPL (SEQ ID No. 26); 마우스 c-abl IV의 서열 SALIKKKKKMAP (SEQ ID No. 27); 인플루엔자 바이러스 NS1의 서열 DRLRR (SEQ ID No. 28) 및 PKQKKRK (SEQ ID No. 29); 간염 바이러스 델타 항원의 서열 RKLKKKIKKL (SEQ ID No. 30); 마우스 Mx1 단백질의 서열 REKKKFLKRR (SEQ ID No. 31); 인간 폴리(ADP-리보스) 중합효소의 서열 KRKGDEVDGVDEVAKKKSKK (SEQ ID No. 32); 및 스테로이드 호르몬 수용체(인간) 글루코코티코이드의 서열 RKCLQAGMNLEARKTKK (SEQ ID No. 33). 일반적으로, 하나 이상의 NLS는 진핵생물 세포의 핵에서 검출 가능한 양으로 DNA-표적화 Cas 단백질 축적을 유도하기에 충분한 강도이다. 일반적으로, 핵 국재화 활성의 강도는 CRISPR-Cas 단백질 내 NLS의 수, 사용되는 특정 NLS(들) 또는 이들 인자의 조합물로부터 유도될 수 있다. 핵에서의 축적의 검출은 임의의 적합한 기법에 의해 수행될 수 있다. 예를 들어, 검출 가능한 마커는, 세포 내의 위치가 시각화될 수 있도록, 예컨대, 핵 위치를 검출하기 위한 수단 (예를 들어, DAPI와 같은 핵에 특이적인 염색)과 조합하여, 핵산-표적화 단백질에 융합될 수 있다. 세포핵은 또한 세포로부터 단리될 수 있으며, 이어서, 이의 함량은 단백질을 검출하기 위한 임의의 적합한 공정, 예컨대, 면역조직화학, 웨스턴 블롯, 또는 효소 활성 분석에 의해 분석될 수 있다. 핵에서의 축적은 또한 CRISPR-Cas 단백질 및 데아미나제 단백질에 노출되지 않은, 또는 하나 이상의 NLS를 결여하는 CRISPR-Cas 및/또는 데아미나제 단백질에 노출된 대조군에 비교하여, 간접적으로, 예컨대, 표적 서열에서 핵산-표적화 복합체 형성 효과에 대한 분석 (예를 들어, 데아미나제 활성에 대한 분석), 또는 DNA-표적화 복합체 형성 및/또는 DNA-표적화에 의해 영향받는 변경된 유전자 발현 활성에 대한 분석에 의해 결정될 수 있다.

[0266] CRISPR-Cas 및/또는 아데노신 데아미나제 단백질은 1개 이상, 예컨대, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 이종성 NLS가 제공될 수 있다. 일부 실시형태에서, 단백질은 아미노-말단에서 또는 근처에서 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 NLS, 카복시-말단에서 또는 근처에서 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 NLS, 또는 이들의 조합 (예를 들어, 아미노-말단에서 0 또는 적어도 하나 이상의 NLS 또는 카복시 말단에서 0 또는 하나 이상의 NLS)을 포함한다. 하나 초과의 NLS가 존재할 때, 각각은 서로 독립적으로 선택될 수 있어서, 단일 NLS가 하나 초과의 카피수로 존재하고/하거나 하나 초과의 카피수로 존재하는 하나 초과의 다른 NLS와 조합하여 존재할 수 있다. 일부 실시형태에서, NLS의 가장 가까운 아미노산이 N-말단 또는 C-말단으로부터 폴리펩티드를 따라서 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50개 이상의 아미노산 내에 있을 때, NLS는 N-말단 또는 C-말단 근처에서 고려된다. CRISPR-Cas 단백질의 바람직한 실시형태에서, NLS는 단백질의 C-말단에 부착된다.

[0267] 본 명세서에 제공된 방법의 일정 실시형태에서, CRISPR-Cas 단백질 및 데아미나제 단백질이 세포에 전달되거나 또는 세포 내에서 별개의 단백질로서 발현된다. 이들 실시형태에서, CRISPR-Cas 및 데아미나제 단백질의 각각은 본 명세서에 기재된 바와 같은 하나 이상의 NLS가 제공될 수 있다. 일정 실시형태에서, CRISPR-Cas 및 데아미나제 단백질이 세포에 전달되거나 또는 세포 내에서 융합 단백질로서 발현된다. 이들 실시형태에서, CRISPR-Cas 및 데아미나제 단백질 중 하나 또는 둘 모두는 하나 이상의 NLS가 제공된다. 아데노신 데아미나제가 상기 기재한 바와 같은 어댑터 단백질 (예컨대, MS2)에 융합되는 경우에, 하나 이상의 NLS는 어댑터 단백질 상에서 제공될 수 있으며, 단 이는 압타머 결합을 방해하지 않는 것을 조건으로 한다. 특정 실시형태에서, 하나 이상의 NLS 서열은 또한 아데노신 데아미나제와 CRISPR-Cas 단백질 사이에 링커 서열로서 작용할 수 있다.

[0268] 일정 실시형태에서, 본 발명의 가이드는 아데노신 데아미나제 또는 이의 촉매적 도메인에 연결되거나 또는 융합될 수 있는 어댑터 단백질에 대해 특정 결합 부위 (예를 들어, 압타머)를 포함한다. 이러한 가이드가 CRISPR 복합체 (즉, 가이드 및 표적에 대한 CRISPR-Cas 단백질 결합)를 형성할 때, 어댑터 단백질이 결합하며, 어댑터 단백질과 회합된 아데노신 데아미나제 또는 이의 촉매적 도메인은 부여받은 작용이 효과적이게 되는 유리한 공간적 배향으로 위치된다.

[0269] 당업자는 어댑터 + 아데노신 데아미나제의 결합을 가능하게 하지만, 어댑터 + 아데노신 데아미나제의 적절하지 않은 배치 (예를 들어, CRISPR 복합체의 3차원 구조 내에서의 입체 장애에 기인)를 허용하는 가이드에 대한 변형이 의도하지 않은 변형이라는 것을 이해할 것이다. 하나 이상의 변형된 가이드는 본 명세서에 기재된 바와 같은 테트라 루프, 스템 루프 1, 스템 루프 2, 또는 스템 루프 3에서, 바람직하게는 테트라 루프 또는 스템

루프 2에서, 가장 바람직하게는 테트라 루프와 스템 루프 2 둘 모두에서 변형될 수 있다.

[0270] **직교성 촉매적 불활성 CRISPR-Cas 단백질의 용도**

[0271] 특정 실시형태에서, Cas13 닉카제는 상기 Cas13 닉카제의 효율을 증가시키기 위해 직교성 촉매적 불활성 CRISPR-Cas 단백질과 조합하여 사용된다 ([Chen et al. 2017, Nature Communications 8:14958; doi:10.1038/ncomms14958]에 기술된 바와 같음). 보다 특히, 직교성 촉매적 불활성 CRISPR-Cas 단백질은 AD-기능화된 CRISPR 시스템에서 사용되는 Cas13 닉카제와 상이한 PAM 인식 부위를 특징으로 하고 상응하는 가이드 서열은 AD-기능화된 CRISPR 시스템의 Cas13 닉카제의 것에 근접한 표적 서열에 결합하도록 선택된다. 본 발명의 상황에서 사용되는 직교성 촉매적 불활성 CRISPR-Cas 단백질은 AD-기능화된 CRISPR 시스템의 일부분을 형성하지 않지만 단지 상기 Cas13 닉카제의 효율을 증가시키는 기능을 하고 상기 CRISPR-Cas 단백질에 대해 당분야에 기술된 바와 같은 표준 가이드 분자와 조합하여 사용된다. 특정 실시형태에서, 상기 직교성 촉매적 불활성 CRISPR-Cas 단백질은 데드 CRISPR-Cas 단백질로서, 다시 말해서, 상기 CRISPR-Cas 단백질의 뉴클레아제 활성을 폐기시키는 하나 이상의 돌연변이를 포함한다. 특정 실시형태에서, 촉매적 불활성 직교성 CRISPR-Cas 단백질은 Cas13 닉카제의 표적 서열에 근접해 있는 표적 서열과 하이브리드화할 수 있는 둘 이상의 가이드 분자가 제공된다. 특정 실시형태에서, 적어도 2종의 가이드 분자가 상기 촉매적 불활성 CRISPR-Cas 단백질을 표적화하는데 사용되고, 이 중에서 적어도 하나의 가이드 분자는 Cas13 닉카제의 표적 서열의 5'의 표적 서열과 하이브리드화할 수 있고, 적어도 하나의 가이드 분자는 AD-기능화된 CRISPR 시스템의 Cas13 닉카제의 표적 서열의 3'의 표적 서열과 하이브리드할 수 있고, 그리하여 상기 하나 이상의 표적 서열은 Cas13 닉카제의 표적 서열과 동일하거나 또는 반대쪽에 있는 DNA 가닥 상에 존재할 수 있다. 특정 실시형태에서, 직교성 촉매적 불활성 CRISPR-Cas 단백질의 하나 이상의 가이드 분자에 대한 가이드 서열은 표적 서열이 AD-기능화된 CRISPR의 표적화를 위해서, 다시 말해, Cas13 닉카제의 표적화를 위해서 가이드 분자의 것과 근접하여 존재하도록 선택된다. 특정 실시형태에서, 직교성 촉매적 불활성 CRISPR-Cas 효소의 하나 이상의 표적 서열은 5 염기쌍 초과 450 염기쌍 미만으로 Cas13 닉카제의 표적 서열로부터 각각 분리된다. 직교성 촉매적 불활성 CRISPR-Cas 단백질과 사용을 위한 가이드의 표적 서열 및 AD-기능화된 CRISPR 시스템의 표적 서열 간 최적 거리는 당업자가 결정할 수 있다. 특정 실시형태에서, 직교성 CRISPR-Cas 단백질은 클래스 II, II형 CRISPR 단백질이다. 특정 실시형태에서, 직교성 CRISPR-Cas 단백질은 클래스 II, V형 CRISPR 단백질이다. 특정 실시형태에서, 촉매적 불활성 직교성 CRISPR-Cas 단백질의 특정 실시형태에서, 촉매적 불활성 직교성 CRISPR-Cas 단백질은 본 명세서의 다른 곳에 기술된 바와 같이 이의 PAM 특이성을 변경시키기 위해 변형되었다. 특정 실시형태에서, Cas13 단백질 닉카제는 그 자체로 인간 세포에서 제한된 활성을 가지지만, 불활성 직교성 CRISPR-Cas 단백질 및 하나 이상의 상응하는 근접한 가이드와 조합하여서, 필요한 닉카제 활성을 보장하는 닉카제이다.

[0272] **CRISPR 개발 및 용도**

[0273] 본 발명은 하기 논문들에 기재된 바와 같고, 특히 CRISPR 단백질 복합체의 전달 및 세포 및 유기체에서 RNA 가이드된 엔도뉴클레아제의 용도와 관련된 CRISPR-Cas 개발 및 용도의 측면들을 기반으로 더욱 예시되고 확장될 수 있다:

- Multiplex genome engineering using CRISPR-Cas systems. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., & Zhang, F. Science Feb 15;339(6121):819-23 (2013);
- RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Jiang W., Bikard D., Cox D., Zhang F, Marraffini LA. Nat Biotechnol Mar;31(3):233-9 (2013);
- One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR-Cas-Mediated Genome Engineering. Wang H., Yang H., Shivalila CS., Dawlaty MM., Cheng AW., Zhang F., Jaenisch R. Cell May 9;153(4):910-8 (2013);
- Optical control of mammalian endogenous transcription and epigenetic states. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F. Nature. Aug 22;500(7463):472-6. doi: 10.1038/Nature12466. Epub 2013 Aug 23 (2013);
- Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Ran, FA., Hsu, PD., Lin, CY., Gootenberg, JS., Konermann, S., Trevino, AE., Scott, DA., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. Cell Aug 28. pii: S0092-8674(13)01015-5 (2013-A);
- DNA targeting specificity of RNA-guided Cas9 nucleases. Hsu, P., Scott, D., Weinstein, J., Ran, FA., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, TJ., Marraffini, LA., Bao, G., & Zhang, F. Nat Biotechnol doi:10.1038/nbt.2647 (2013);

[0274]

- Genome engineering using the CRISPR-Cas9 system. Ran, FA., Hsu, PD., Wright, J., Agarwala, V., Scott, DA., Zhang, F. Nature Protocols Nov;8(11):2281-308 (2013-B);
- Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Shalem, O., Sanjana, NE., Hartenian, E., Shi, X., Scott, DA., Mikkelson, T., Heckl, D., Ebert, BL., Root, DE., Doench, JG., Zhang, F. Science Dec 12. (2013);
- Crystal structure of cas9 in complex with guide RNA and target DNA. Nishimasu, H., Ran, FA., Hsu, PD., Konermann, S., Shehata, SI., Dohmae, N., Ishitani, R., Zhang, F., Nureki, O. Cell Feb 27, 156(5):935-49 (2014);
- Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Wu X., Scott DA., Kriz AJ., Chiu AC., Hsu PD., Dadon DB., Cheng AW., Trevino AE., Konermann S., Chen S., Jaenisch R., Zhang F., Sharp PA. Nat Biotechnol. Apr 20. doi: 10.1038/nbt.2889 (2014);
- CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, Graham DB, Jhunjhunwala S, Heidenreich M, Xavier RJ, Langer R, Anderson DG, Hacohen N, Regev A, Feng G, Sharp PA, Zhang F. Cell 159(2): 440-455 DOI: 10.1016/j.cell.2014.09.014(2014);
- Development and Applications of CRISPR-Cas9 for Genome Engineering, Hsu PD, Lander ES, Zhang F., Cell. Jun 5;157(6):1262-78 (2014).
- Genetic screens in human cells using the CRISPR-Cas9 system, Wang T, Wei JJ, Sabatini DM, Lander ES., Science. January 3; 343(6166): 80–84. doi:10.1126/science.1246981 (2014);
- Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation, Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE., (published online 3 September 2014) Nat Biotechnol. Dec;32(12):1262-7 (2014);
- *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9, Swiech L, Heidenreich M, Banerjee A, Habib N, Li Y, Trombetta J, Sur M, Zhang F., (published online 19 October 2014) Nat Biotechnol. Jan;33(1):102-6 (2015);
- Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex, Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F., Nature. Jan 29;517(7536):583-8 (2015).

[0275]

- A split-Cas9 architecture for inducible genome editing and transcription modulation, Zetsche B, Volz SE, Zhang F., (published online 02 February 2015) Nat Biotechnol. Feb;33(2):139-42 (2015);
- Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis, Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R, Lee H, Zhang F, Sharp PA. Cell 160, 1246–1260, March 12, 2015 (multiplex screen in mouse), and
- In *vivo* genome editing using *Staphylococcus aureus* Cas9, Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F., (published online 01 April 2015), Nature. Apr 9;520(7546):186-91 (2015).
- Shalem et al., "High-throughput functional genomics using CRISPR-Cas9," Nature Reviews Genetics 16, 299-311 (May 2015).
- Xu et al., "Sequence determinants of improved CRISPR sgRNA design," Genome Research 25, 1147-1157 (August 2015).
- Parnas et al., "A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks," Cell 162, 675-686 (July 30, 2015).
- Ramanan et al., CRISPR-Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus," Scientific Reports 5:10833. doi: 10.1038/srep10833 (June 2, 2015)
- Nishimasu et al., Crystal Structure of Staphylococcus aureus Cas9," Cell 162, 1113-1126 (Aug. 27, 2015)
- BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis, Canver et al., Nature 527(7577):192-7 (Nov. 12, 2015) doi: 10.1038/nature15521. Epub 2015 Sep 16.
- *Cas13 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System*, Zetsche et al., Cell 163, 759-71 (Sep 25, 2015).
- *Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems*, Shmakov et al., Molecular Cell, 60(3), 385–397 doi: 10.1016/j.molcel.2015.10.008 Epub October 22, 2015.
- *Rationally engineered Cas9 nucleases with improved specificity*, Slaymaker et al., Science 2016 Jan 1 351(6268): 84-88 doi: 10.1126/science.aad5227. Epub 2015 Dec 1.

[0276]

- Gao *et al*, "Engineered Cas13 Enzymes with Altered PAM Specificities," bioRxiv 091611; doi: http://dx.doi.org/10.1101/091611 (Dec. 4, 2016).

[0277]

[0278] 참조로 본 명세서에 편입되는 이들 각각은 본 발명의 실시에서 고려될 수 있고, 하기에 간략하게 설명한다:

[0279] ▷ 문헌 [Cong *et al.*]은 스트렙토코커스 써모필러스 (*Streptococcus thermophilus*) Cas9와 또한 스트렙토코커스 피오게네스 (*Streptococcus pyogenes*) Cas9 둘 모두에 기반하여 진핵생물 세포에서 사용하기 위한 II형 CRISPR-Cas 시스템을 조작하였고, Cas9 뉴클레아제가 짧은 RNA에 의해 인간 및 마우스 세포에서 DNA의 정확한 절단을 유도하도록 지정될 수 있다는 것을 입증하였다. 그들의 연구는 추가로 닉킹 효소로 전환되는 Cas9가 최소 돌연변이유발 활성을 갖는 진핵생물 세포에서 상동성-지정 복구 (homology-directed repair)를 촉진하도록 사용될 수 있다는 것을 보여주었다. 추가로, 그들의 연구는 포유동물 게놈 내의 내생성 게놈 유전자좌 부위에서 몇몇의 자발적 편집을 가능하게 하기 위해 다중 가이드 서열이 단일 CRISPR 어레이로 코딩될 수 있다는 것을 입증하였는데, 이는 RNA-가이드된 뉴클레아제 기술의 용이한 프로그램 가능성 (programmability) 및 넓은 적

용 가능성을 입증하였다. 세포에서 서열 특이적 DNA 절단을 프로그램하기 위해 RNA를 사용하는 이런 능력은 새로운 부류의 게놈 조작 도구를 정하였다. 이들 연구는 다른 CRISPR 유전자좌가 포유동물 세포에 이식 가능하게 될 가능성이 있고, 또한 포유동물 게놈 절단을 매개할 수 있다는 것을 추가로 보여주었다. 중요하게는, CRISPR-Cas 시스템의 몇몇 양상은 그의 효율 및 융통성을 증가시키기 위해 더 개선될 수 있다는 것을 상정할 수 있다.

[0280] ▷ 문헌 [Jiang *et al.*]은 스트렙토코커스 뉴모니아 (*Streptococcus pneumoniae*) 및 에스케리치아 콜라이 (*Escherichia coli*)의 게놈에서 정확한 돌연변이를 도입하기 위해 이중-RNA와 복합체화된 주기적 간격으로 분포하는, 짧은 회문구조 반복부 (CRISPR)-연관 Cas9 엔도뉴클레아제를 사용하였다. 상기 접근법은 비돌연변이 세포를 사멸시키고 선택 가능한 마커 또는 반대-선택 시스템에 대한 필요를 피하기 위해 표적화된 게놈 부위에서 이중-RNA:Cas9-지정 절단에 의존한다. 상기 연구는 편집 주형 상에서 수행되는 단일- 및 다중뉴클레오티드 변화를 생성하기 위해 짧은 CRISPR RNA(crRNA) 서열을 변화시킴으로써 재프로그래밍 이중-RNA:Cas9 특이성을 보고하였다. 본 연구는 2개의 crRNA의 동시 사용이 다중복합 돌연변이유발을 가능하게 하였다는 것을 보여주었다. 더 나아가서, 상기 접근이 스트렙토코커스 뉴모니아에서 리컴비니어링 (recombineering)과 조합하여 사용될 때, 기술된 접근법을 사용하여 회수된 세포의 거의 100%는 목적하는 돌연변이를 함유하였고, 이. 콜라이에서 사용될 때, 회수된 세포의 65%는 돌연변이를 함유하였다.

[0281] ▷ 문헌 [Wang *et al* (2013)]은 배아 줄기 세포에서의 순차적 재조합 및/또는 단일 돌연변이를 갖는 마우스의 시간-소모적 상호교잡 (intercrossing)에 의해 다중 단계에서 전통적으로 생성된 다중 유전자에서의 돌연변이를 운반하는 마우스의 1단계 생성을 위한 CRISPR-Cas 시스템을 사용하였다. CRISPR-Cas 시스템은 기능적으로 중복되는 유전자 및 상위 유전자 상호작용의 생체내 연구를 크게 가속화시킬 것이다.

[0282] ▷ 문헌 [Konermann *et al.* (2013)]은 CRISPR Cas9 효소 및 또한 전사 활성체 유사 이펙터에 기반하여 DNA-결합 도메인의 광학적 그리고 화학적 조절을 가능하게 하는 다재다능하고 강한 기술에 대한 당업계의 요구를 해결하였다.

[0283] ▷ 문헌 [Ran *et al.* (2013-A)]은 표적화된 이중-가닥 파괴를 도입하기 위해 짝지어진 가이드 RNA와 Cas9 닉카제 돌연변이체를 조합한 접근을 기재하였다. 이는 가이드 서열에 의해 특정 게놈 유전자좌로 표적화되는 미생물 CRISPR-Cas 시스템으로부터의 Cas9 뉴클레아제의 문제를 해결하는데, 이는 DNA 표적에 대한 소정의 미스매치를 용인하고 이에 의해 원치않는 오프-표적 돌연변이유발을 촉진시킬 수 있다. 게놈에서 개개의 틈은 높은 충실도로 복구되기 때문에, 적절하게 오프셋 가이드 RNA를 통한 동시 닉킹은 이중-가닥 파괴에 필요하고, 표적 절단을 위해 특이적으로 인식되는 염기의 수를 연장시킨다. 저자는 짝지어진 닉킹의 사용이 세포주에서 50배 내지 1,500배 만큼 오프-표적 활성을 감소시킬 수 있고 온-표적의 절단 효율을 희생시키지 않고 마우스 접합체에서 유전자 녹아웃을 용이하게 한다는 것을 입증하였다. 이 다재다능한 전략은 높은 특이성을 필요로 하는 매우 다양한 게놈 편집 적용을 가능하게 한다.

[0284] ▷ 문헌 [Hsu *et al.* (2013)]은 표적 부위의 선택을 알아내기 위해 그리고 오프-표적 효과를 피하기 위해 인간 세포에서 SpCas9 표적화 특이성을 특징규명하였다. 연구는 293T 및 293FT 세포에서의 100개 초과의 예측된 게놈 오프-표적 유전자좌에서 700개 초과의 가이드 RNA 변이체 및 SpCas9-유도된 삽입결실 돌연변이 수준을 평가하였다. 저자들은 SpCas9가 미스매치의 수, 위치 및 분포에 민감한 서열-의존적 방식으로 상이한 위치에서 가이드 RNA와 표적 DNA 사이의 미스매치를 용인한다는 것을 보여주었다. 저자들은 추가로 SpCas9-매개 절단이 DNA 메틸화에 의해 영향받지 않는다는 것과, SpCas9 및 가이드 RNA의 투여량이 오프-표적 변형을 최소화시키도록 적정될 수 있다는 것을 보여주었다. 추가적으로, 포유동물 게놈 조작 적용을 용이하게 하기 위해, 저자들은 표적 서열뿐만 아니라 오프-표적 분석의 선택 및 확인을 가이드하기 위한 웹-기반 소프트웨어 툴의 제공을 보고하였다.

[0285] ▷ 문헌 [Ran *et al.* (2013-B)]은 포유동물 세포에서 비-상동성 말단 결합 (NHEJ) 또는 상동성-지정 복구 (HDR)뿐만 아니라 하류의 기능성 연구를 위해 변형된 세포주의 생성을 통한 Cas9-매개 게놈 편집을 위한 도구의 세트를 기재하였다. 오프-표적 절단을 최소화하기 위해, 저자들은 짝지워진 가이드 RNA와 함께 Cas9 닉카제 돌연변이체를 이용하는 이중-닉킹 전략을 추가로 기재하였다. 저자에 의해 제공되는 프로토콜은 표적 부위의 선택, 절단 효율의 평가 및 오프-표적 활성의 분석을 위한 가이드라인을 실험적으로 유도하였다. 상기 연구들은 표적 디자인에서 시작하여, 유전자 변형이 1 내지 2주 내에 달성될 수 있고, 변형된 클론 세포주가 2 내지 3주 내에 유도될 수 있다는 것을 보여주었다.

[0286] ▷ 문헌 [Shalem *et al.*]은 게놈-와이드 규모에 대한 유전자 기능의 정보를 얻기 위한 새로운 방법을 기재하였

다. 그들의 연구는 64,751개의 독특한 가이드 서열에 의해 18,080개의 유전자를 표적화한 게놈-규모 CRISPR-Cas9 녹아웃 (GeCKO) 라이브러리의 전달이 인간 세포에서 음성 선별 및 양성 선별 스크리닝 둘 모두를 가능하게 하였다는 것을 보여주었다. 처음에, 저자는 암 및 다능성 줄기 세포에서 세포 생존도에 필수적인 유전자를 식별하기 위한 GeCKO 라이브러리의 사용을 보여주었다. 다음에, 흑색종 모델에서, 저자는 유전자의 상실이 돌연변이체 단백질 키나제 BRAF를 억제하는 치료제인 베무라페닙에 대한 내성에 연루되는 유전자를 스크리닝하였다. 그들의 연구는 가장 높은-순위의 후보가 이전에 입증한 유전자 NF1 및 MED12 뿐만 아니라 신규한 히트 NF2, CUL3, TADA2B 및 TADA1을 포함하였다는 것을 보여주었다. 저자는 동일한 유전자를 표적화하는 독립적 가이드 RNA와 높은 비율의 히트 확인 간 높은 수준의 일관성을 관찰하였고, 따라서, Cas9에 의한 게놈-규모 스크리닝의 유망함을 입증하였다.

[0287] ▷ 문헌 [Nishimasu *et al.*]은 2.5 A° 분해능으로 sgRNA 및 그의 표적 DNA와의 복합체에서 스트렙토코커스 피오게네스 Cas9의 결정 구조를 보고하였다. 상기 구조는 표적 인식 및 뉴클레아제 로브 (lobe)로 구성된 이중로브 (bilobed) 구조를 나타내었는데, 이는 그들의 계면에서 양으로 하전된 그루브에 sgRNA:DNAn RNA 듀플렉스를 수용한다. 인식 로브는 sgRNA 및 DNA의 결합에 필수적인 반면, 뉴클레아제 로브는 HNH 및 RuvC 뉴클레아제 도메인을 함유하는데, 이는 각각 표적 DNA의 상보성 및 비상보성 가닥의 절단을 위해 적절하게 위치된다.
뉴클레아제 로브는 또한 프로토스페이서 인접 모티프 (PAM)와의 상호작용을 초래하는 카복실-말단의 도메인을 함유한다. 이런 고-분해능 구조 및 수반하는 기능성 분석은 Cas9에 의한 RNA-가이드된 DNA 표적화의 분자 메커니즘을 나타내었고, 따라서 새롭고, 다재다능한 게놈-편집 기법의 합리적 디자인을 위한 방법을 조성한다.

[0288] ▷ 문헌 [Wu *et al.*]은 마우스 배아 줄기 세포 (mESC)에서 단일 가이드 RNA(sgRNA)와 함께 로딩된 스트렙토코커스 피오게네스로부터의 촉매적 불활성인 Cas9(dCas9)의 맵핑된 게놈-와이드 결합 부위이다. 저자들은 시험한 4개의 shRNA의 각각이 dCas9를 sgRNA 및 NGG 프로토스페이서 인접 모티프 (PAM)에서 5-뉴클레오티드 씨드 영역을 빈번하게 특징으로 하는, 수십 내지 수천개의 게놈 부위로 표적화한다는 것을 나타내었다. 염색질 비접근성은 매칭 씨드 서열을 갖는 다른 부위에 대한 dCas9 결합을 감소시키고; 따라서 오프-표적 부위의 70%는 유전자와 관련된다. 저자들은 촉매적으로 활성인 Cas9로 형질감염시킨 mESC에서 295 dCas9 결합 부위의 표적화된 시퀀싱이 배경 수준 초과로 돌연변이된 단지 하나의 부위를 식별하였다는 것을 보여주었다. 저자들은 Cas9 결합 및 절단에 대한 2-상태 모델을 제안하였는데, 여기서 씨드 매치는 결합을 촉발시키지만 표적 DNA와의 광범위 쌍형성이 절단에 필요하다.

[0289] ▷ 문헌 [Platt *et al.*]은 Cre-의존적 Cas9 녹인 (knock in) 마우스를 확립하였다. 저자들은 뉴런, 면역세포 및 내피 세포에서 가이드 RNA의 아데노-연합 바이러스(AAV)-, 렌티바이러스-, 또는 입자-매개 전달을 이용하는 생체내 및 생체외 게놈 편집을 입증하였다.

[0290] ▷ 문헌 [Hsu *et al.* (2014)]은 세포의 유전자 스크리닝을 포함하여, 요거트에서부터 게놈 편집까지 CRISPR-Cas9 역사를 일반적으로 논의하는 리뷰 논문이다.

[0291] ▷ 문헌 [Wang *et al.* (2014)]은 게놈-규모 렌티바이러스 단일 가이드 RNA (sgRNA) 라이브러리를 사용하는 양성 선택 및 음성 선별 둘 모두에 적합한 풀링된, 기능상실 유전자 스크리닝 접근법에 관한 것이다.

[0292] ▷ 문헌 [Doench *et al.*]은 6개의 내생성 마우스 및 3개의 내생성 인간 유전자 패널의 모든 가능한 표적 부위에 걸쳐 타일링된 sgRNA의 풀을 생성하였고, 항체 염색 및 유세포분석에 의해 그들의 표적 유전자의 비대립유전자를 생성하는 그들의 능력을 정량적으로 평가하였다. 저자들은 PAM의 최적화가 활성을 개선시키고 또한 sgRNA를 디자인하기 위한 온라인 툴을 제공하였다는 것을 보여주었다.

[0293] ▷ 문헌 [Swiech *et al.*]은 AAV-매개 SpCas9 게놈 편집이 뇌에서 유전자 기능의 역 유전자 연구를 가능하게 할 수 있다는 것을 입증한다.

[0294] ▷ 문헌 [Konermann *et al.* (2015)]은 링커 존재 및 부재로 스템 또는 테트라루프와 같은 가이드 상의 적절한 위치에서 다중 이펙터 도메인, 예를 들어, 전사 활성체, 기능성 및 후생적 조절자에 부착하는 능력을 논의한다.

[0295] ▷ 문헌 [Zetsche *et al.*]은 Cas9 효소가 둘로 분할될 수 있고, 따라서 활성화를 위한 Cas9의 조립체가 제어될 수 있다는 것을 입증한다.

[0296] ▷ 문헌 [Chen *et al.*]은 마우스에서의 게놈-와이드 생체내 CRISPR-Cas9 스크린이 폐 전이를 조절하는 유전자를 밝혀준다는 것을 입증한 다중복합 스크리닝에 관한 것이다.

[0297] ▷ 문헌 [Ran *et al.* (2015)]은 SaCas9 및 게놈을 편집하는 그의 능력에 관한 것이며, 생화학적 분석으로부터 추론할 수 없다는 것을 입증한다.

[0298] ▷ 문헌 [Shalem *et al.* (2015)]은 촉매적 불활성인 Cas9 (dCas9) 융합체를 합성적으로 발현을 억제 (CRISPRi) 또는 활성화 (CRISPRa)시키는데 사용하여, 전사 활성을 조절하는 전략 및 게놈 유전자좌를 불활성화시키는 정렬되고 풀링된 스크린, 녹아웃 접근법을 포함하여, 게놈-규모 스크린에 Cas9를 사용하는 장점을 보여주는 방식을 기술한다.

[0299] ▷ 문헌 [Xu *et al.* (2015)]은 CRISPR-기반 스크리닝에서 단일 가이드 RNA (sgRNA) 효율에 기여하는 DNA 서열 특징을 평가하였다. 저자들은 CRISPR-Cas9 녹아웃의 효율 및 절단 부위에서 뉴클레오티드 선호도를 연구하였다. 저자들은 또한 CRISPRi/a에 대한 서열 선호도가 CRISPR-Cas9 녹아웃과 실질적으로 상이하다는 것을 발견하였다.

[0300] ▷ 문헌 [Parnas *et al.* (2015)]은 박테리아성 지질다당체 (LPS)에 의한 종양 괴사 인자 (Tnf)의 유도를 제어하는 유전자를 확인하기 위해 게놈-와이드 풀링된 CRISPR-Cas9 라이브러리를 수지상 세포 (DC) 내로 도입하였다. Tlr4 신호전달의 알려진 조절자 및 이전에 알려지지 않은 후보를 동정하였고, LPS에 대한 정규 반응에 대해 별개의 효과를 갖는 3개의 기능성 모듈로 분류하였다.

[0301] ▷ 문헌 [Ramanan *et al* (2015)]은 감염 세포에서 바이러스 에피솜 DNA (cccDNA)의 절단을 입증하였다. HBV 게놈은 HBV 생활사에서 중요한 성분인 공유적으로 폐쇄된 원형 DNA (cccDNA)로 불리는 3.2kb의 이중-가닥 에피솜 DNA로서, 감염된 간세포의 핵에 존재하며, 이의 복제는 현재의 요법에 의해 억제되지 않는다. 저자들은 HBV의 고도로 보존된 영역을 특이적으로 표적화하는 sgRNA가 바이러스 복제 및 감손된 cccDNA를 억제한다는 것을 보여주었다.

[0302] ▷ 문헌 [Nishimasu *et al.* (2015)]은 5'-TTGAAT-3' PAM 및 5'-TTGGGT-3' PAM을 함유하는 단일 가이드 RNA (sgRNA) 및 그의 이중-가닥 DNA와의 복합체에서 SaCas9의 결정 구조를 보고하였다. SaCas9와 SpCas9의 구조적 비교는 그들의 별개의 PAM 특이성 및 오솔로그 sgRNA 인식을 설명하는, 구조적 보존과 분기를 둘 모두 강조하였다.

[0303] ▷ 문헌 [Canver *et al.* (2015)]은 넌코딩 게놈 구성요소의 CRISPR-Cas9-기반 기능성 연구를 입증하였다. 저자들은 인핸서의 중요한 특징을 나타낸 인간과 마우스 BCL11A 인핸서의 인시추 포화 돌연변이유발을 수행하기 위해 풀링된 CRISPR-Cas9 가이드 RNA 라이브러리를 개발하였다.

[0304] ▷ 문헌 [Zetsche et al. (2015)]은 Cas9와 별개의 특징을 갖는 프란시셀라 노비시다 U112로부터의 클래스 2 CRISPR 뉴클레아제인 Cas13의 특징규명을 보고하였다. Cas13은 tracrRNA를 결여하는 단일 RNA-가이드 엔도뉴클레아제이고, T-풍부 프로토스페이서-인접 모티프를 이용하고, 스태거드 DNA 이중-가닥 파괴를 통해 DNA를 절단한다.

[0305] ▷ 문헌 [Shmakov et al. (2015)]은 3종의 별개 클래스 2 CRISPR-Cas 시스템을 보고하였다. 2종의 시스템 CRISPR 효소 (C2c1 및 C2c3)는 Cas13과 멀리 관련된 RuvC-유사 엔도뉴클레아제 도메인을 함유한다. Cas13과 달리, C2c1은 DNA 절단을 위해 crRNA와 tracrRNA 둘 모두에 의존한다. 제3 효소 (C2c2)는 2종의 예측된 HEPN RNase 도메인을 함유하고, tracrRNA 독립적이다.

[0306] ▷ 문헌 [Slaymaker et al (2016)]은 스트렙토코커스 피오게네스 Cas9 (SpCas9)의 특이성을 개선시키기 위한 구조-가이드 조작의 사용을 보고하였다. 저자들은 오프-표적 효과가 감소된 강한 온-표적의 절단이 유지된 "증강된 특이성" SpCas9 (eSpCas9) 변이체를 개발하였다.

[0307] 본 명세서에 제공된 방법 및 도구는 tracrRNA를 이용하지 않는 II 형 뉴클레아제, Cas13을 예시한다. Cas13의 오솔로그는 본 명세서에 기재된 바와 같은 상이한 박테리아 종에서 동정되었다. 유사한 특성을 갖는 추가적인 II형 뉴클레아제는 당업계에 기재된 방법을 이용하여 동정될 수 있다 (Shmakov et al. 2015, 60:385-397; Abudayeh et al. 2016, Science, 5;353(6299)). 특정 실시형태에서, 신규한 CRISPR 이펙터 단백질을 동정하기 위한 이러한 방법은 CRISPR Cas 유전자좌의 존재를 동정하는 씨드를 코딩하는 데이터베이스로부터 서열을 선택하는 단계, 선택된 서열에서 오픈 리딩 프레임 (ORF)을 포함하는 10kb의 씨드 내에 위치된 유전자좌를 동정하는 단계, 이로부터, 단일 ORF만이 700개 초과의 아미노산 및 공지된 CRISPR 이펙터에 대해 90% 이하의 상동성을 갖는 신규한 CRISPR 이펙터를 코딩하는 ORF를 포함하는 유전자좌를 선택하는 단계를 포함할 수 있다. 특정 실시형태에서, 씨드는 CRISPR-Cas 시스템, 예컨대, Cas1에 대해 통상적인 단백질이다. 추가 실시형

태에서, CRISPR 어레이는 새로운 이펙터 단백질을 동정하기 위한 씨드로서 사용된다.

[0308] 본 발명의 유효성은 입증되었다. Cas13 및 crRNA를 포함하는 사전 조립된 재조합 CRISPR-Cas13 복합체는, 예를 들어 전기천공으로 형질감염되어, 높은 돌연변이율 및 검출 가능한 오프-표적 돌연변이의 부재를 초래한다. Hur, J.K. et al, targeted mutagenesis in mice by electroporation of Cas13 ribonucleoproteins, Nat Biotechnol. 2016 Jun 6. doi: 10.1038/nbt.3596. 게놈-와이드 분석은 Cas13이 고도로 특이적이라는 것을 보여준다. 한 측정에 의해, 인간 HEK293T 세포에서 Cas13에 대해 결정된 시험관내 절단 부위는 SpCas9에 대한 것보다 상당히 더 낮았다. Kim, D. et al., Genome-wide analysis reveals specificities of Cas13 endonucleases in human cells, Nat Biotechnol. 2016 Jun 6. doi: 10.1038/nbt.3609. Cas13을 사용하는 효율적인 다중복합 시스템은 본 발명의 tRNA를 함유하는 어레이로부터 가공된 gRNA를 사용하여 초파리에서 입증되었다. Port, F. et al, Expansion of the CRISPR toolbox in an animal with tRNA-flanked Cas9 and Cas13 gRNAs. doi: http://dx.doi.org/10.1101/046417.

[0309] 또한, ["Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome edit", Shengdar Q. Tsai, Nicolas Wyvekens, Cyd Khayter, Jennifer A. Foden, Vishal Thapar, Deepak Reyon, Mathew J. Goodwin, Martin J. Aryee, J. Keith Joung Nature Biotechnology 32(6): 569-77 (2014)]은 연장된 서열을 인식하고 인간 세포에서 고효율로 내생성 유전자를 편집할 수 있는 이량체 RNA-가이드 FokI 뉴클레아제에 관한 것이다.

[0310] CRISPR-Cas-발현 진핵생물 세포, CRISPR-Cas 발현 진핵생물, 예컨대 마우스를 비롯하여, 양 및 제제에 관한 것을 포함해, 방법, 재료, 전달 비히클, 벡터, 입자 및 이의 제조 및 사용 방법을 포함하여, CRISPR/Cas 시스템, 이의 성분 및 이러한 성분의 전달에 관한 일반 정보에 대해서, 다음을 참조한다: 미국 특허 제8,999,641호, 제8,993,233호, 제8,697,359호, 제8,771,945호, 제8,795,965호, 제8,865,406호, 제8,871,445호, 제8,889,356호, 제8,889,418호, 제8,895,308호, 제8,906,616호, 제8,932,814호 및 제8,945,839호; 미국 특허 공개 제2014-0310830호 (미국 특허 출원 제14/105,031호), 미국 특허 제2014-0287938 A1호 (미국 특허 출원 제14/213,991호), 미국 특허 제2014-0273234 A1호 (미국 특허 출원 제14/293,674호), 미국 특허 제2014-0273232 A1호 (미국 특허 출원 제14/290,575호), 미국 특허 제2014-0273231 (미국 특허 출원 제14/259,420호), 미국 특허 제2014-0256046 A1호 (미국 특허 출원 제14/226,274호), 미국 특허 제2014-0248702 A1호 (미국 특허 출원 제14/258,458호), 미국 특허 제2014-0242700 A1호 (미국 특허 출원 제14/222,930호), 미국 특허 제2014-0242699 A1호 (미국 특허 출원 제14/183,512호), 미국 특허 제2014-0242664 A1호 (미국 특허 출원 제14/104,990호), 미국 특허 제2014-0234972 A1호 (미국 특허 출원 제14/183,471호), 미국 특허 제2014-0227787 A1호 (미국 특허 출원 제14/256,912호), 미국 특허 제2014-0189896 A1호 (미국 특허 출원 제14/105,035호), 미국 특허 제2014-0186958 (미국 특허 출원 제14/105,017호), 미국 특허 제2014-0186919 A1호 (미국 특허 출원 제14/104,977호), 미국 특허 제2014-0186843 A1호 (미국 특허 출원 제14/104,900호), 미국 특허 제2014-0179770 A1호 (미국 특허 출원 제14/104,837호) 및 미국 특허 제2014-0179006 A1호 (미국 특허 출원 제14/183,486호), 미국 특허 제2014-0170753호 (미국 특허 출원 제14/183,429호); 미국 특허 제2015-0184139호 (미국 특허 출원 제14/324,960호); 제14/054,414호 유럽 특허 출원 제2 771 468호(유럽 특허 제13818570.7호), 유럽 특허 제2 764 103호(유럽 특허 제13824232.6호) 및 유럽 특허 제2 784 162호(유럽 특허 제14170383.5호); 및 국제 특허 출원 공개 번호WO2014/093661 (PCT/US2013/074743), WO2014/093694 (PCT/US2013/074790), WO2014/093595 (PCT/US2013/074611), WO2014/093718 (PCT/US2013/074825), WO2014/093709 (PCT/US2013/074812), WO2014/093622 (PCT/US2013/074667), WO2014/093635 (PCT/US2013/074691), WO2014/093655 (PCT/US2013/074736), WO2014/093712 (PCT/US2013/074819), WO2014/093701 (PCT/US2013/074800), WO2014/018423 (PCT/US2013/051418), WO2014/204723 (PCT/US2014/041790), WO2014/204724 (PCT/US2014/041800), WO2014/204725 (PCT/US2014/041803), WO2014/204726 (PCT/US2014/041804), WO2014/204727 (PCT/US2014/041806), WO2014/204728 (PCT/US2014/041808), WO2014/204729 (PCT/US2014/041809), WO2015/089351 (PCT/US2014/069897), WO2015/089354 (PCT/US2014/069902), WO2015/089364 (PCT/US2014/069925), WO2015/089427 (PCT/US2014/070068), WO2015/089462 (PCT/US2014/070127), WO2015/089419 (PCT/US2014/070057), WO2015/089465 (PCT/US2014/070135), WO2015/089486 (PCT/US2014/070175), WO2015/058052 (PCT/US2014/061077), WO2015/070083 (PCT/US2014/064663), WO2015/089354 (PCT/US2014/069902), WO2015/089351 (PCT/US2014/069897), WO2015/089364 (PCT/US2014/069925), WO2015/089427 (PCT/US2014/070068), WO2015/089473 (PCT/US2014/070152), WO2015/089486 (PCT/US2014/070175), WO2016/049258 (PCT/US2015/051830), WO2016/094867 (PCT/US2015/065385), WO2016/094872 (PCT/US2015/065393),

WO2016/094874 (PCT/US2015/065396), WO2016/106244 (PCT/US2015/067177).

[0311] 또한 다음을 언급할 수 있다: 2015년 6월 17일자로 출원된 미국 특허 출원 제62/180,709호 (발명의 명칭: PROTECTED GUIDE RNAS(PGRNAS)); 2014년 12월 12일자로 출원된 미국 특허 출원 제62/091,455호 (발명의 명칭: PROTECTED GUIDE RNAS (PGRNAS); 2014년 12월 24일자로 출원된 미국 특허 출원 제62/096,708호 (발명의 명칭: PROTECTED GUIDE RNAS (PGRNAS)); 2014년 12월 12일자로 출원된 미국 특허 출원 제62/091,462호, 2014년 12월 23일자로 출원된 제62/096,324호, 2015년 6월 17일자로 출원된 제62/180,681호 및 2015년 10월 5일자로 출원된 제62/237,496호 (발명의 명칭: DEAD GUIDES FOR CRISPR TRANSCRIPTION FACTORS); 2014년 12월 12일자로 출원된 미국 특허 출원 제62/091,456호 및 2015년 6월 17일자로 출원된 제62/180,692호 (발명의 명칭: ESCORTED AND FUNCTIONALIZED GUIDES FOR CRISPR-CAS SYSTEMS); 2014년 12월 12일자로 출원된 미국 특허 출원 제62/091,461호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR GENOME EDITING AS TO HEMATOPOETIC STEM CELLS (HSCs); 2014년 12월 19일자로 출원된 미국 특허 출원 제62/094,903호 (발명의 명칭: UNBIASED IDENTIFICATION OF DOUBLE-STRAND BREAKS AND GENOMIC REARRANGEMENT BY GENOME-WISE INSERT CAPTURE SEQUENCING); 2014년 12월 24일자로 출원된 미국 특허 출원 제62/096,761호 (발명의 명칭: ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED ENZYME AND GUIDE SCAFFOLDS FOR SEQUENCE MANIPULATION); 2014년 12월 30일자로 출원된 미국 특허 출원 제62/098,059호, 2015년 6월 18일자로 출원된 제62/181,641호 및 2015년 6월 18일자로 출원된 제62/181,667호 (발명의 명칭: RNA-TARGETING SYSTEM); 2014년 12월 24일자로 출원된 미국 특허 출원 제62/096,656호 및 2015년 6월 17일자로 출원된 제62/181,151호 (발명의 명칭: CRISPR HAVING 또는 ASSOCIATED WITH DESTABILIZATION DOMAINS); 2014년 12월 24일자로 출원된 미국 특허 출원 제62/096,697호 (발명의 명칭: CRISPR HAVING 또는 ASSOCIATED WITH AAV); 2014년 12월 30일자로 출원된 미국 특허 출원 제62/098,158호 (발명의 명칭: ENGINEERED CRISPR COMPLEX INSERTIONAL TARGETING SYSTEMS); 2015년 4월 22일자로 출원된 미국 특허 출원 제62/151,052호 (발명의 명칭: CELLULAR TARGETING FOR EXTRACELLULAR EXOSOMAL REPORTING); 2014년 9월 24일자로 출원된 미국 특허 출원 제62/054,490호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING PARTICLE DELIVERY COMPONENTS); 2014년 2월 12일자로 출원된 미국 특허 출원 제 61/939,154호 (발명의 명칭: SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS); 2014년 9월 25일자로 출원된 미국 특허 출원 제62/055,484호 (발명의 명칭: SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS); 2014년 12월 4일자로 출원된 미국 특허 출원 제62/087,537호 (발명의 명칭: SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS); 2014년 9월 24일자로 출원된 미국 특허 출원 제62/054,651호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO); 2014년 10월 23일자로 출원된 미국 특허 출원 제62/067,886호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO); 2014년 9월 24일자로 출원된 미국 특허 출원 제62/054,675호, 및 2015년 6월 17일자로 출원된 제62/181,002호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN NEURONAL CELLS/TISSUES); 2014년 9월 24일자로 출원된 미국 특허 출원 제62/054,528호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN IMMUNE DISEASES 또는 DISORDERS); 2014년 9월 25일자로 출원된 미국 특허 출원 제62/055,454호 (발명의 명칭: DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING CELL PENETRATION PEPTIDES (CPP); 2014년 9월 25일자로 출원된 미국 특허 출원 제62/055,460호 (발명의 명칭: MULTIFUNCTIONAL-CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES; 2014년 12월 4일자로 출원된 미국 특허 출원 제62/087,475호 및 2015년 6월 18일자로 출원된 제62/181,690호 (발명의 명칭: FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS); 2014년 9월 25일자로 출원된 미국 특허 출원 제62/055,487호 (발명의 명칭: FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS); 2014년 12월 4일자로 출원된 미국 특허 출원 제62/087,546호 및 2015년 6월 18일자로 출원된 제62/181,687호 (발명의 명칭: MULTIFUNCTIONAL CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES); 및 2014년 12월 30일자로 출원된 미국 특허 출원 제62/098,285호 (발명의 명칭: CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND METASTASIS).

[0312] 2015년 6월 18일자로 출원된 미국 특허 출원 제62/181,659호 및 2015년 8월 19일자로 출원된 제62/207,318호

(발명의 명칭: ENGINEERING AND OPTIMIZATION OF SYSTEMS, METHODS, ENZYME AND GUIDE SCAFFOLDS OF CAS9 ORTHOLOGS AND VARIANTS FOR SEQUENCE MANIPULATION)가 언급된다. 2015년 6월 18일자로 출원된 미국 특허 출원 제62/181,663 및 2015년 10월 22일자로 출원된 제62/245,264호 (발명의 명칭: NOVEL CRISPR ENZYMES AND SYSTEMS), 2015년 6월 18일자로 출원된 미국 특허 출원 제62/181,675호, 2015년 10월 22일자로 출원된 제62/285,349호, 2016년 2월 17일자로 출원된 제62/296,522호 및 2016년 4월 8일자로 출원된 제62/320,231호 (발명의 명칭: NOVEL CRISPR ENZYMES AND SYSTEMS), 2015년 9월 24일자로 출원된 미국 특허 출원 제62/232,067호, 2015년 12월 18일자로 출원된 미국 특허 출원 제14/975,085호, 유럽 특허 출원 제16150428.7호, 2015년 8월 16일자로 출원된 미국 특허 출원 제62/205,733호, 2015년 8월 5일자로 출원된 미국 특허 출원 제62/201,542호, 2015년 7월 16일자로 출원된 미국 특허 출원 제62/193,507호 및 2015년 6월 18일자로 출원된 미국 특허 출원 제62/181,739호 (발명의 명칭: NOVEL CRISPR ENZYMES AND SYSTEMS) 및 2015년 10월 22일자로 출원된 미국 특허 출원 제62/245,270호 (발명의 명칭: NOVEL CRISPR ENZYMES AND SYSTEMS)가 언급된다. 2014년 2월 12일자로 출원된 미국 특허 출원 제61/939,256호 및 2014년 12월 12일자로 출원된 WO 2015/089473 (PCT/US2014/070152) (발명의 명칭: ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED GUIDE COMPOSITIONS WITH NEW ARCHITECTURES FOR SEQUENCE MANIPULATION)이 언급된다. 또한 2015년 8월 15일자로 출원된 PCT/US2015/045504, 2015년 6월 17일자로 출원된 미국 특허 출원 제62/180,699호 및 2014년 8월 17일자로 출원된 미국 특허 출원 제62/038,358호 (발명의 명칭: GENOME EDIT USING CAS9 NIKCASES)가 언급된다.

[0313] 각각의 이들 특허, 특허 공보 및 출원, 및 그들 심사 동안 또는 그에 인용된 모든 문서 ("출원 인용 문헌") 및 인출원 인용 문헌에서 인용 또는 참조된 모든 문서는 거기에서 또는 거기에 인용된 임의 문서에서 언급되는 임의 제품에 대한 임의의 지시서, 설명서, 제품 사양서, 및 제품 시트가 참조로 본 명세서에 편입되고, 본 발명의 실시에서 적용될 수 있다. 모든 문서 (예를 들어, 이들 특허, 특허 공보 및 출원 및 출원 인용 문서)는 각각의 개별 문서가 특별히 개별적으로 참조로 편입된다고 표시된 바와 동일한 정도로 참조로 본 명세서에 편입된다.

[0314] **V형 CRISPR-Cas 단백질**

[0315] 본 출원은 V형 CRISPR-Cas 단백질을 사용하는 방법을 기술한다. 본 명세서에서는 Cas13을 예시하고, 그리하여 다수의 오솔로그 또는 상동체가 확인되었다. 추가의 오솔로그 또는 상동체를 확인할 수 있고 본 명세서에 기술된 임의의 기능성은 다수 오솔로그로부터의 단편을 포함하는 키메라 효소를 포함하여, 다른 오솔로그로 조작될 수 있다는 것은 당업자에게 자명할 것이다.

[0316] 신규한 CRISPR-Cas 유전자좌를 확인하기 위한 컴퓨터 방법은 EP3009511 또는 US2016208243에 기술되어 있고, 하기 단계들을 포함할 수 있다: Cas1 단백질을 코딩하는 모든 contig를 검출하는 단계; cas1 유전자의 20 kB 내의 모든 예측되는 단백질 코딩 유전자를 확인하는 단계; 확인된 유전자와 Cas 단백질-특이적 프로파일을 비교하고 CRISPR 어레이를 예측하는 단계; 500개 아미노산 초과 (>500 aa)의 단백질을 함유하는 미분류된 후보 CRISPR-Cas 유전자좌를 선택하는 단계; PSI-BLAST 및 HHPred와 같은 방법을 사용하여 선택된 후보를 분석하여서 기지의 단백질 도메인을 스크리닝하고, 그리하여 신규한 클래스 2 CRISPR-Cas 유전자좌를 확인하는 단계 (또한, [Schmakov et al. 2015, Mol Cell. 60(3):385-97] 참조). 상기 언급된 단계들 이외에도, 후보에 대한 추가 분석은 추가 상동체에 대한 메타게노믹스 데이터베이스를 검색하여 수행될 수 있다. 추가적으로 또는 대안적으로, 비자율적 CRISPR-Cas 시스템으로 검색을 확대하기 위해서, 씨드로서 사용되는 CRISPR 어레이를 사용하여 동일한 절차를 수행할 수 있다.

[0317] 일 양상에서, Cas1 단백질을 코딩하는 모든 contig를 검출하는 단계는 본 명세서에 참조로 편입되는, ["GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions." John Besemer, Alexandre Lomsadze and Mark Borodovsky, Nucleic Acids Research (2001) 29, pp 2607-2618]에 더욱 기술된 바와 같은 유전자 예측 프로그램인 GenemarkS를 통해 수행된다.

[0318] 일 양상에서 모든 예측되는 단백질 코딩 유전자를 확인하는 단계는 확인된 유전자를 Cas 단백질-특이적 프로파일과 비교하고 조상 도메인 및 전체 길이 단백질에 대해 주석이 잘 달려진 다중 서열 정렬 모델의 컬렉션으로 이루어진 단백질 주석 리소스인 NCBI 보존성 도메인 데이타베이스 (CDD)에 따라서 주석을 달아서 수행된다. 이들은 RPS-BLAST를 통해 단백질 서열 내 보존된 도메인의 신속한 확인을 위한 단백질-특이적 점수 매트릭스 (PSSM)로 이용가능하다. CDD 내용은 수많은 외부 자원 데이터베이스로부터 입력된 도메인 모델 (Pfam, SMART, COG, PRK, TIGRFAM)을 비롯하여, 서열/구조/기능 관련성으로 통찰력을 제공하고 도메인 경계를 명확하게

정의하기 위해 3D-구조 정보를 사용하는, NCBI-큐레이트된 도메인을 포함한다. 추가 양상에서, CRISPR 어레이는 본 명세서에 참조로 편입되는 ["PILER-CR: fast and accurate identification of CRISPR repeats", Edgar, R.C., BMC Bioinformatics, Jan 20;8:18(2007)]에 기술된 바와 같은 CRISPR 반복부를 찾기 위한 공공 도메인 소프트웨어인 PILER-CR 프로그램을 사용하여 예측되었다.

[0319] 추가 양상에서, 사례별 분석은 PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool)를 사용하여 수행된다. PSI-BLAST는 단백질-단백질 BLAST를 사용하여 소정 점수 한계치 이상에서 검출된 서열의 다중 서열 정렬로부터의 프로파일 또는 위치-특이적 채점 매트릭스 (Position-specific scoring matrix; PSSM)를 구동한다. 이러한 PSSM은 새로운 매치를 위해 데이터베이스를 더 검색하는데 사용되고, 이들 새롭게 검출된 서열로 후속 반복을 위해 업데이트된다. 따라서, PSI-BLAST는 단백질 간 장거리 상관도를 검출하는 수단을 제공한다.

[0320] 다른 양상에서, 사례별 분석은 BLAST 또는 PSI-BLAST 만큼 사용이 용이하고 동시에 멀리 있는 상동체 조사에서 훨씬 더 민감한 서열 데이터베이스 검색 및 구조 예측을 위한 방법인, HHpred를 사용하여 수행된다. 사실, HHpred의 감도는 현재 이용가능한 구조 예측을 위한 가장 강력한 서버와 경쟁력이 있다. HHpred는 프로파일 HMM (hidden Markov model)의 쌍별 비교를 기반으로 하는 최초 서버이다. 대부분의 통상의 서열 검색 방법은 UniProt 또는 NR과 같은 서열 데이터베이스를 검색하는데 반해, HHpred는 Pfam 또는 SMART와 같은 정렬 데이터베이스를 검색한다. 이것은 단일 서열의 클러터 (clutter) 대신에 다수의 서열 패밀리에 대한 히트 목록을 상당히 단순화시킨다. 모든 주요한 공공으로 입수가능한 프로파일 및 정렬 데이터베이스는 HHpred를 통해 이용가능하다. HHpred는 입력치로서 단일 문의 서열 또는 다수 정렬을 수용한다. 단지 수분 이내에, PSI-BLAST와 유사한 판독이 용이한 포맷으로 검색 결과를 돌려준다. 검색 옵션은 국소 또는 전체 정렬 및 2차 구조 유사성 채점을 포함한다. HHpred는 쌍별 문의-주형 서열 정렬, 병합 문의-주형 다중 정렬 (예를 들어, 추이적 검색)을 비롯하여, HHpred 정렬로부터 MODELLER 소프트웨어를 통해 계산된 3D 구조 모델을 생성시킬 수 있다.

[0321] **Cas13의 오솔로그**

[0322] 용어 "오솔로그" ("orthologue" 또는 "ortholog") 및 "상동체" ("homologue" 또는 "homolog")는 당분야에 충분히 공지되어 있다. 추가 지침으로, 본 명세서에서 사용되는 단백질의 "상동체"는 상동체인 단백질과 동일하거나 또는 유사한 기능을 수행하는 동일 종의 단백질이다. 그러나 상동성 단백질은 구조적으로 관련될 필요가 없거나, 또는 오직 부분적으로 구조적으로 관련된다. 본 명세서에서 사용되는 단백질의 "오솔로그"는 오솔로그인 단백질과 동일하거나 또는 상이한 기능을 수행하는 상이한 종의 단백질이다. 그러나 오솔로그 단백질은 구조적으로 관련될 필요가 없거나, 또는 오직 부분적으로 구조적으로 관련된다. 상동체 및 오솔로그는 상동성 모델링 (예를 들어, [Greer, Science vol. 228 (1985) 1055], 및 [Blundell et al. Eur J Biochem vol 172 (1988), 513] 참조) 또는 "구조적 BLAST" (Dey F, Cliff Zhang Q, Petrey D, Honig B. Toward a "structural BLAST": using structural relationships to infer function. Protein Sci. 2013 Apr;22(4):359-66. doi: 10.1002/pro.2225.)를 통해 확인될 수 있다. 또한, CRISPR-Cas 유전자좌 분야에서 적용을 위한 [Shmakov et al. (2015)]를 참조한다. 그러나 상동성 단백질은 구조적으로 관련될 필요가 없거나, 또는 오직 부분적으로만 구조적으로 관련된다.

[0323] Cas13 유전자는 몇몇 다양한 박테리아 게놈, 전형적으로 cas1, cas2, 및 cas4 유전자 및 CRISPR 카세트 (예를 들어, 프란시셀라 cf. 노비시다 (*Francisella cf. novicida*) Fx1의 FNFX1_1431-FNFX1_1428)와 동일한 유전자좌에서 발견된다. 따라서, 이러한 추정의 신규 CRISPR-Cas 시스템의 레이아웃은 II-B형의 것과 유사한 것으로 나타난다. 뿐만 아니라, Cas9와 유사하게, Cas13 단백질은 트랜스포존 ORF-B와 상동성이고 활성 RuvC-유사 뉴클레아제, 아르기닌-풍부 영역, 및 Zn 핑거 (Cas9에는 부재)를 포함하는 쉽게 확인가능한 C-말단 영역을 함유한다. 그러나, Cas9와 달리, Cas13은 또한 CRISPR-Cas 환경없이 몇몇 게놈에 존재하고 ORF-B와 이의 매우 높은 유사성은 이것이 트랜스포존 성분일 수 있음을 시사한다. 이것이 진짜 CRISPR-Cas 시스템이고 Cas13이 Cas9의 기능적 유사체라면 이것은 신규한 CRISPR-Cas 유형, 즉 V형이라는 것을 시사하였다 (Annotation and Classification of CRISPR-Cas Systems. Makarova KS, Koonin EV. Methods Mol Biol. 2015;1311:47-75). 그러나, 본 명세서에 기술된 바와 같이, Cas13은 동일한 도메인 구조를 갖지 않아서 V-B 아형이라고 표시되는 C2c1p와 구별하기 위해 V-A형이라고 한다.

[0324] 본 발명은 V-A 아형으로 표시되는 Cas13 유전자좌로부터 유래된, Cas13 이펙터 단백질의 용도를 포괄한다. 본 명세서에서 이러한 이펙터 단백질은 또한 "Cas13p", 예를 들어, Cas13 단백질 (및 이러한 이펙터 단백질 또

는 Cas13 단백질 또는 Cas13 유전자좌로부터 유래된 단백질은 "CRISPR-Cas 단백질이라고도 함)이라고도 한다.

[0325] 특정 실시형태에서, 이펙터 단백질은 스트렙토코커스 (*Streptococcus*), 캄필로박터 (*Campylobacter*), 니트라티프락토르 (*Nitratifractor*), 스타필로코커스 (*Staphylococcus*), 파르비바큘럼 (*Parvibaculum*), 로세부리아 (*Roseburia*), 네이세리아 (*Neisseria*), 글루콘아세토박터 (*Gluconacetobacter*), 아조스피릴럼 (*Azospirillum*), 스파에로차에타 (*Sphaerochaeta*), 락토바실러스 (*Lactobacillus*), 유박테리움 (*Eubacterium*), 코리네박터 (*Corynebacter*), 카르노박테리움 (*Carnobacterium*), 로도박터 (*Rhodobacter*), 리스테리아 (*Listeria*), 팔루디박터 (*Paludibacter*), 클로스트리듐 (*Clostridium*), 라크노스피라과 (*Lachnospiraceae*), 클로스트리디아리듐 (*Clostridiaridium*), 렙토트리키아 (*Leptotrichia*), 프란시셀라 (*Francisella*), 레지오넬라 (*Legionella*), 알리시클로바실러스 (*Alicyclobacillus*), 메타노메티오필러스 (*Methanomethyophilus*), 포르피로모나스 (*Porphyromonas*), 프레보텔라 (*Prevotella*), 박테로이데테스 (*Bacteroidetes*), 헬코코커스 (*Helcococcus*), 렙토스피라 (*Leptospira*), 데설포비브리오 (*Desulfovibrio*), 데설포나트로넘 (*Desulfonatronum*), 오피투타세아 (*Opitutaceae*), 투베리바실러스 (*Tuberibacillus*), 바실러스 (*Bacillus*), 브레비바실러스 (*Brevibacilus*), 메틸로박테리움 (*Methylobacterium*), 부티비브리오 (*Butyvibrio*), 페리그리니박테리움 (*Perigrinibacterium*), 파레우박테리움 (*Pareubacterium*), 모라셀라 (*Moraxella*), 티오미크로스피라 (*Thiomicrospira*) 또는 아시다미노코커스 (*Acidaminococcus*)를 포함하는 속 유래의 유기체 유래 Cas13 이펙터 단백질이다. 특정 실시형태에서, Cas13 이펙터 단백질은 유박테리움, 라크노스피라과, 렙토트리키아, 프란시셀라, 메타노메티오필러스, 포르피로모나스, 프레보텔라, 렙토스피라, 부티비브리오, 페리그리니박테리움, 파레우박테리움, 모라셀라, 티오미크로스피라 또는 아시다미노코커스로부터 선택된 속 유래 유기체로부터 선택된다.

[0326] 추가의 특정 실시형태에서, Cas13 이펙터 단백질은 에스. 뮤탄스 (*S. mutans*), 에스. 아갈락티아 (*S. agalactiae*), 에스. 에퀴시밀리스 (*S. equisimilis*), 에스. 산귀니스 (*S. sanguinis*), 에스. 뉴모니아 (*S. pneumonia*); 씨. 제주니 (*C. jejuni*), 씨. 콜라이 (*C. coli*); 엔. 살수기니스 (*N. salsuginis*), 엔. 테르가쿠스 (*N. tergarcus*); 에스. 아우리쿨라리스 (*S. auricularis*), 에스. 카르노서스 (*S. carnosus*); 엔. 메닌지티데스 (*N. meningitides*), 엔. 고노레아 (*N. gonorrhoeae*); 엘. 모노사이토게네스 (*L. monocytogenes*), 엘. 이바노비 (*L. ivanovii*); 씨. 보툴리눔 (*C. botulinum*), 씨. 디피실 (*C. difficile*), 씨. 테타니 (*C. tetani*), 씨. 소르델리 (*C. sordellii*), 엘. 이나다이 (*L. inadai*), 에프. 툴라렌시스 (*F. tularensis*) 1, 피. 알벤시스 (*P. albensis*), 엘. 박테리움 (*L. bacterium*), 비. 프로테오클라스티쿠스 (*B. proteoclasticus*), 피. 박테리움 (*P. bacterium*), 피. 크레비오리카니스 (*P. crevioricanis*), 피. 디시엔스 (*P. disiens*) 및 피. 마카카 (*P. macacae*)로부터 선택된 유기체로부터 유래된다.

[0327] 이펙터 단백질은 제1 이펙터 단백질 (예를 들어, Cas13) 오솔로그로부터의 제1 단편 및 제2 이펙터 (예를 들어, Cas13) 단백질 오솔로그로부터의 제2 단편을 포함하는 키메라 이펙터 단백질을 포함할 수 있고, 여기서 제1 이펙터 단백질 오솔로그와 제2 이펙터 단백질 오솔로그는 상이하다. 제1 및 제2 이펙터 단백질 (예를 들어, Cas13) 오솔로그 중 적어도 하나는 스트렙토코커스, 캄필로박터, 니트라티프락토르, 스타필로코커스, 파르비바큘럼, 로세부리아, 네이세리아, 글루콘아세토박터, 아조스피릴럼, 스파에로차에타, 락토바실러스, 유박테리움, 코리네박터, 카르노박테리움, 로도박터, 리스테리아, 팔루디박터, 클로스트리듐, 라크노스피라과, 클로스트리디아리듐, 렙토트리키아, 프란시셀라, 레지오넬라, 알리시클로바실러스, 메타노메티오필러스, 포르피로모나스, 프레보텔라, 박테로이데테스, 헬코코커스, 렙토스피라, 데설포비브리오, 데설포나트로넘, 오피투타세아, 투베리바실러스, 바실러스, 브레비바실러스, 메틸로박테리움, 부티비브리오, 페리그리니박테리움, 파레우박테리움, 모라셀라, 티오미크로스피라 또는 아시다미노코커스를 포함하는 유기체 유래 이펙터 단백질 (예를 들어, Cas13)을 포함할 수 있고; 예를 들어, 제1 단편 및 제2 단편을 포함하는 키메라 이펙터 단백질을 포함할 수 있는데, 여기서 제1 단편 및 제2 단편의 각각은 스트렙토코커스, 캄필로박터, 니트라티프락토르, 스타필로코커스, 파르비바큘럼, 로세부리아, 네이세리아, 글루콘아세토박터, 아조스피릴럼, 스파에로차에타, 락토바실러스, 유박테리움, 코리네박터, 카르노박테리움, 로도박터, 리스테리아, 팔루디박터, 클로스트리듐, 라크노스피라과, 클로스트리디아리듐, 렙토트리키아, 프란시셀라, 레지오넬라, 알리시클로바실러스, 메타노메티오필러스, 포르피로모나스, 프레보텔라, 박테로이데테스, 헬코코커스, 렙토스피라, 데설포비브리오, 데설포나트로넘, 오피투타세아, 투베리바실러스, 바실러스, 브레비바실러스, 메틸로박테리움, 부티비브리오, 페리그리니박테리움, 파레우박테리움, 모라셀라, 티오미크로스피라 또는 아시다미노코커스를 포함하는 유기체의 Cas13으로부터 선택되고; 제1 단편과 제2 단편은 동일한 박테리아 유래가 아니고; 예를 들어, 제1 단편 및 제2 단편을 포함하는 키메라 이펙터 단백질로서, 제1 단편 및 제2 단편의 각각은 에스. 뮤탄스, 에스. 아갈락티아, 에스. 에퀴시밀리스, 에스. 산귀니스, 에스. 뉴모니아; 씨. 제주니, 씨. 콜라이; 엔. 살수기니스, 엔. 테르가쿠스; 에스. 아우리쿨라리스, 에스. 카르노서스; 엔. 메닌기티데스, 엔. 고노레아; 엘. 모노사이토게네스, 엘. 이바노비; 씨. 보툴리눔, 씨. 디피실, 씨.

테타니, 씨. 소르델리; 프란시셀라 툴라렌시스 1, 프레보텔라 알벤시스 (*Prevotella albensis*), 라크노스피라과 박테리움 MC2017 1, 부티리비브리오 프로테오클라스티쿠스 (*Butyrivibrio proteoclasticus*), 페레그리니박테리아 박테리움 (*Peregrinibacteria bacterium*) GW2011_GWA2_33_10, 파르쿠박테리아 박테리움 (*Parcubacteria bacterium*) GW2011_GWC2_44_17, 스미텔라 (*Smithella*) sp. SCADC, 아시다미노코커스 sp. BV3L6, 라크노스피라과 박테리아 MA2020, 칸디다투스 메타노플라스마 터미툼 (*Candidatus Methanoplasma termitum*), 유박테리움 엘리겐스 (*Eubacterium eligens*), 모라셀라 보보쿨리 (*Moraxella bovoculi*) 237, 렙토스피라 이나다이 (*Leptospira inadai*), 라크노스피라과 박테리아 ND2006, 포르피로모나스 크레비오리카니스 (*Porphyromonas crevioricanis*) 3, 프레보텔라 디시엔스 (*Prevotella disiens*) 및 포르피로모나스 마카카 (*Porphyromonas macacae*)의 Cas13으로부터 선택되고, 제1 단편 및 제2 단편은 동일한 박테리아로부터 유래되지 않는다.

[0328] 보다 바람직한 실시형태에서, Cas13p는 프란시셀라 툴라렌시스 1, 프레보텔라 알벤시스, 라크노스피라과 박테리아 MC2017 1, 부티리비브리오 프로테오클라스티쿠스, 페레그리니박테리아 박테리움 GW2011_GWA2_33_10, 파르쿠박테리아 박테리움 GW2011_GWC2_44_17, 스미텔라 sp. SCADC, 아시다미노코커스 sp. BV3L6, 라크노스피라과 박테리아 MA2020, 칸디다투스 메타노플라스마 터미툼, 유박테리움 엘리겐스, 모라셀라 보보쿨리 237, 모라셀라 보보쿨리 AAX08_00205, 모라셀라 보보쿨리 AAX11_00205, 부티리비브리오 sp. NC3005, 티오미크로스피라 sp. XS5, 렙토스피라 이나다이, 라크노스피라과 박테리아 ND2006, 포르피로모나스 크레비오리카니스 3, 프레보텔라 디시엔스 및 포르피로모나스 마카카로부터 선택된 박테리아 종으로부터 유래된다. 일정 실시형태에서, Cas13p는 아시다미노코커스 sp. BV3L6, 라크노스피라과 박테리아 MA2020으로부터 선택된 박테리아 종으로부터 유래된다.

일정 실시형태에서, 이펙터 단백질은 제한없이 프란시셀라 툴라렌시스 아종 노비시다를 포함하는, 프란시셀라 툴라렌시스 1의 아종으로부터 유래된다. 바람직한 일정 실시형태에서, Cas13p는 아시다미노코커스 sp. BV3L6, 라크노스피라과 박테리아 ND2006, 라크노스피라과 박테리아 MA2020, 모라셀라 보보쿨리 AAX08_00205, 모라셀라 보보쿨리 AAX11_00205, 부티리비브리오 sp. NC3005 또는 티오미크로스피라 sp. XS5로부터 선택된 박테리아 종으로부터 유래된다.

[0329] 특정 실시형태에서, 본 명세서에서 언급되는 Cas13의 상동체 또는 오솔로그는 본 명세서에 개시된 예로서의 Cas13 단백질과 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예컨대, 예를 들어, 적어도 95%의 서열 상동성 또는 동일성을 갖는다. 추가 실시형태에서, 본 명세서에서 언급되는 Cas13의 상동체 또는 오솔로그는 야생형 Cas13과 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예컨대, 예를 들어, 적어도 95%의 서열 동일성을 갖는다. Cas13이 하나 이상의 돌연변이를 갖는 경우 (돌연변이됨)에, 본 명세서에서 언급되는 상기 Ca13의 상동체 또는 오솔로그는 돌연변이된 Cas13과 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예컨대, 예를 들어, 적어도 95%의 서열 동일성을 갖는다.

[0330] 일 실시형태에서, Cas13 단백질은 제한없이 아시다미노코커스 sp, 라크노스피라과 박테리아 또는 모라셀라 보보쿨리를 포함하는 속의 유기체의 오솔로그일 수 있고; 특정 실시형태에서, V형 Cas 단백질은 제한없이 아시다미노코커스 sp. BV3L6; 라크노스피라과 박테리아 ND2006 (LbCas13) 또는 모라셀라 보보쿨리 237을 포함하는 종의 유기체의 오솔로그일 수 있다. 특정 실시형태에서, 본 명세서에서 언급되는 Cas13의 상동체 또는 오솔로그는 본 명세서에 개시된 Cas13 서열 중 하나 이상과 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예를 들어, 적어도 95%의 서열 상동성 또는 동일성을 갖는다. 추가 실시형태에서, 본 명세서에 언급된 바와 같은 Cas13의 상동체 또는 오솔로그는 야생형 FnCas13, AsCas13 또는 LbCas13과 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예를 들어, 적어도 95%의 서열 동일성을 갖는다.

[0331] 특정 실시형태에서, 본 발명의 Cas13 단백질은 FnCas13, AsCas13 또는 LbCas13과 적어도 60%, 더 특별하게는 적어도 70%, 예컨대 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예를 들어, 적어도 95%의 서열 상동성 또는 동일성을 갖는다. 추가 실시형태에서, 본 명세서에서 언급되는 바와 같은 Cas13 단백질은 야생형 AsCas13 또는 LbCas13과 적어도 60%, 예컨대 적어도 70%, 더 특별하게는 적어도 80%, 더 바람직하게는 적어도 85%, 훨씬 더 바람직하게는 적어도 90%, 예를 들어, 적어도 95%의 서열 동일성을 갖는다. 특정 실시형태에서, 본 발명의 Cas13 단백질은 FnCas13과 60% 미만의 서열 동일성을 갖는다. 당업자는 이것이 Cas13 단백질의 절단된 형태를 포함하고, 그에 따라 서열 동일성은 절단된 형태의 길이 상에서 결정된다는 것을 이해할 것이다. 특정 실시형태에서, Cas13 효소는 FnCas13이 아니다.

[0332] **변형된 Cas13 효소**

[0333] 특정 실시형태에서, 본 명세서에 정의한 바와 같은 조작된 Cas13 단백질, 예컨대, Cas13을 사용하게 하는 데 관심이 있고, 여기서 단백질은 RNA를 포함하는 핵산 분자와 복합체를 형성하여 CRISPR 복합체를 형성하고, CRISPR 복합체에서, 핵산 분자가 하나 이상의 표적 폴리뉴클레오티드 유전자좌를 표적화할 때, 단백질은 비변형 Cas13 단백질에 비해 적어도 하나의 변형을 포함하고, 변형된 단백질을 포함하는 CRISPR 복합체는 비변형 Cas13 단백질을 포함하는 복합체에 비해 변경된 활성을 갖는다. 본 명세서에서 CRISPR "단백질"에 대해 언급할 때, Cas13 단백질은 바람직하게는, 예컨대, 제한없이 Cas13을 포함하여, (예를 들어, 증가 또는 감소 (또는 부재)된 효소 활성을 갖는) 변형된 CRISPR-Cas 단백질이라는 것을 이해해야 한다. 용어 "CRISPR 단백질"은 야생형 CRISPR 단백질에 비해서 CRISPR 단백질이 변경, 예컨대, 증가되거나 또는 감소 (또는 부재)된 활성을 갖는지 여부와 무관하게, "CRISPR-Cas 단백질"과 상호교환 가능하게 사용될 수 있다.

[0334] Cas13 뉴클레아제의 1차 구조의 컴퓨터 분석은 3개의 별개 영역을 밝혀준다. 첫번째, C-말단 RuvC 유사 도메인은, 유일한 기능적으로 특징규명된 도메인이다. 두번째, N-말단의 알파-나선 영역, 및 세번째로 RuvC 유사 도메인과 알파-나선 영역 사이에 위치된 혼합 알파 및 베타 영역이다.

[0335] 비구조화된 영역의 몇몇 소형 스트레치는 Cas13 1차 구조 내에서 예측된다. 용매에 노출되고 상이한 Cas13 오솔로그 내에서 보존되지 않은 비구조화된 영역은 작은 단백질 서열의 분할 및 삽입에 바람직한 측면이다. 추가로, 이들 측면은 Cas13 오솔로그 간에 키메라 단백질을 생성하는 데 사용될 수 있다.

[0336] 상기 정보에 기반하여, 효소의 불활성화를 야기하거나 또는 닉카제 활성으로 이중 가닥 뉴클레아제를 변형시키는 돌연변이체가 생성될 수 있다. 대안의 실시형태에서, 이러한 정보는 오프-표적 효과가 감소된 효고를 개발하는데 사용된다 (본 명세서의 다른 곳에 기재됨).

[0337] 일정한 상기 기술된 Cas13 효소에서, 상기 효소는 제한없이, AsCas13 (아시다미노코커스 sp. BV3L6)의 아미노산 위치 번호매김에 대해 위치 R909, R912, R930, R947, K949, R951, R955, K965, K968, K1000, K1002, R1003, K1009, K1017, K1022, K1029, K1035, K1054, K1072, K1086, R1094, K1095, K1109, K1118, K1142, K1150, K1158, K1159, R1220, R1226, R1242, 및/또는 R1252를 포함하는 (RuvC 도메인 내) 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되고, 더 바람직하게는 표적에 대한 특이성이 증가된다.

[0338] 일정한 상기-기술된 비-천연 발생 CRISPR-Cas 단백질에서, 상기 효소는 제한없이 AsCas13 (아시다미노코커스 sp. BV3L6)의 아미노산 위치 번호매김에 대해 위치 K324, K335, K337, R331, K369, K370, R386, R392, R393, K400, K404, K406, K408, K414, K429, K436, K438, K459, K460, K464, R670, K675, R681, K686, K689, R699, K705, R725, K729, K739, K748, 및/또는 K752를 포함하는 (RAD50) 도메인 내 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되고, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0339] 일정한 Cas13 효소에서, 효소는 제한없이 AsCas13 (아시다미노코커스 sp. BV3L6)의 아미노산 위치 번호매김에 대해 위치 R912, T923, R947, K949, R951, R955, K965, K968, K1000, R1003, K1009, K1017, K1022, K1029, K1072, K1086, F1103, R1226, 및/또는 R1252를 포함하는, 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되고, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0340] 일정 실시형태에서, Cas13 효소는 제한없이, LbCas13 (라크노스피라과 박테리아 ND2006)의 아미노산 위치 번호매김에 대해 위치 R833, R836, K847, K879, K881, R883, R887, K897, K900, K932, R935, K940, K948, K953, K960, K984, K1003, K1017, R1033, R1138, R1165, 및/또는 R1252를 포함하는 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되며, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0341] 일정 실시형태에서, Cas13 효소는 제한없이, AsCas13 (아시다미노코커스 sp. BV3L6)의 아미노산 위치 번호매김에 대해 위치 K15, R18, K26, Q34, R43, K48, K51, R56, R84, K85, K87, N93, R103, N104, T118, K123, K134, R176, K177, R192, K200, K226, K273, K275, T291, R301, K307, K369, S404, V409, K414, K436, K438, K468, D482, K516, R518, K524, K530, K532, K548, K559, K570, R574, K592, D596, K603, K607, K613, C647, R681, K686, H720, K739, K748, K757, T766, K780, R790, P791, K796, K809, K815, T816, K860, R862, R863, K868, K897, R909, R912, T923, R947, K949, R951, R955, K965, K968, K1000, R1003, K1009, K1017, K1022, K1029, A1053, K1072, K1086, F1103, S1209, R1226, R1252, K1273, K1282, 및/또는 K1288을 포함하는 하나 이상의 잔

기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되며, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0342] 일정 실시형태에서, 효소는 제한없이 FnCas13 (프란시셀라 노비시다 U112)의 아미노산 위치 번호매김에 대해 위치 K15, R18, K26, R34, R43, K48, K51, K56, K87, K88, D90, K96, K106, K107, K120, Q125, K143, R186, K187, R202, K210, K235, K296, K298, K314, K320, K326, K397, K444, K449, E454, A483, E491, K527, K541, K581, R583, K589, K595, K597, K613, K624, K635, K639, K656, K660, K667, K671, K677, K719, K725, K730, K763, K782, K791, R800, K809, K823, R833, K834, K839, K852, K858, K859, K869, K871, R872, K877, K905, R918, R921, K932, I960, K962, R964, R968, K978, K981, K1013, R1016, K1021, K1029, K1034, K1041, K1065, K1084, 및/또는 K1098을 포함하는 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되며, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0343] 일정 실시형태에서, 효소는 제한없이, LbCas13 (라크노스피라과 박테리아 ND2006)의 아미노산 위치 번호매김에 대해 위치 K15, R18, K26, K34, R43, K48, K51, R56, K83, K84, R86, K92, R102, K103, K116, K121, R158, E159, R174, R182, K206, K251, K253, K269, K271, K278, P342, K380, R385, K390, K415, K421, K457, K471, A506, R508, K514, K520, K522, K538, Y548, K560, K564, K580, K584, K591, K595, K601, K634, K640, R645, K679, K689, K707, T716, K725, R737, R747, R748, K753, K768, K774, K775, K785, K787, R788, Q793, K821, R833, R836, K847, K879, K881, R883, R887, K897, K900, K932, R935, K940, K948, K953, K960, K984, K1003, K1017, R1033, K1121, R1138, R1165, K1190, K1199 및/또는 K1208을 포함하는 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되며, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0344] 일정 실시형태에서, 상기 효소는 제한없이, MbCas13 (모라셀라 보보쿨리 237)의 아미노산 위치 번호매김에 대해 위치 K14, R17, R25, K33, M42, Q47, K50, D55, K85, N86, K88, K94, R104, K105, K118, K123, K131, R174, K175, R190, R198, I221, K267, Q269, K285, K291, K297, K357, K403, K409, K414, K448, K460, K501, K515, K550, R552, K558, K564, K566, K582, K593, K604, K608, K623, K627, K633, K637, E643, K780, Y787, K792, K830, Q846, K858, K867, K876, K890, R900, K901, M906, K921, K927, K928, K937, K939, R940, K945, Q975, R987, R990, K1001, R1034, I1036, R1038, R1042, K1052, K1055, K1087, R1090, K1095, N1103, K1108, K1115, K1139, K1158, R1172, K1188, K1276, R1293, A1319, K1340, K1349, 및/또는 K1356을 포함하는 하나 이상의 잔기의 돌연변이에 의해 변형된다. 일정 실시형태에서, 상기 하나 이상의 돌연변이를 포함하는 Cas13 효소는 변형되며, 더 바람직하게는 표적에 대한 특이성이 증가되었다.

[0345] 일 실시형태에서, Cas13 단백질은 AsCas13의 아미노산 위치 번호매김에 대해 S1228에서 돌연변이 (예를 들어, S1228A)를 갖게 변형된다. 그 전문이 참조로 본 명세서에 편입되는 문헌 [Yamano *et al., Cell* 165:949-962 (2016)]을 참조한다.

[0346] 일정 실시형태에서, Cas13 단백질은 비천연 PAM을 인식하고, 예컨대 서열 YCN, YCV, AYV, TYV, RYN, RCN, TGYV, NTTN, TTN, TRTN, TYTV, TYCT, TYCN, TRTN, NTTN, TACT, TYCC, TRTC, TATV, NTTV, TTV, TSTG, TVTS, TYYS, TCYS, TBYS, TCYS, TNYS, TYYS, TNTN, TSTG, TTCC, TCCC, TATC, TGTG, TCTG, TYCV 또는 TCTC을 갖거나 또는 이들 서열을 포함하는 PAM을 인식하도록 변형되었다. 특정 실시형태에서, 상기 돌연변이된 Cas13은 AsCas13의 위치 11, 12, 13, 14, 15, 16, 17, 34, 36, 39, 40, 43, 46, 47, 50, 54, 57, 58, 111, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 642, 643, 644, 645, 646, 647, 648, 649, 651, 652, 653, 654, 655, 656, 676, 679, 680, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 707, 711, 714, 715, 716, 717, 718, 719, 720, 721, 722, 739, 765, 768, 769, 773, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884 또는 1048 또는 Cas13 오솔로그에서 이에 상응하는 위치에서 하나 이상의 돌연변이된 아미노산 잔기; 바람직하게는 위치 130, 131, 132, 133, 134, 135, 136, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 570, 571, 572, 573, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610,

611, 612, 613, 614, 615, 630, 631, 632, 646, 647, 648, 649, 650, 651, 652, 653, 683, 684, 685, 686, 687, 688, 689 또는 690에서 하나 이상의 돌연변이된 아미노산 잔기를 포함한다.

[0347] 일정 실시형태에서, Cas13 단백질은 증가된 활성, 즉 더 넓은 PAM 특이성을 갖도록 변형된다. 특정한 실시형태에서, Cas13 단백질은 제한없이 AsCas13의 위치 539, 542, 547, 548, 550, 551, 552, 167, 604, 및/또는 607, 또는 AsCas13 오솔로그, 상동체, 또는 변이체의 상응하는 위치를 포함한 하나 이상의 잔기의 돌연변이, 바람직하게는 위치 542 또는 542 및 607에서 돌연변이된 아미노산 잔기로서, 상기 돌연변이는 바람직하게 542R 및 607R, 예컨대 S542R 및 K607R인 돌연변이된 아미노산 잔기; 또는 바람직하게 위치 542 및 548 (및 임의로는 552)에서 돌연변이된 아미노산 잔기로서, 상기 돌연변이는 바람직하게 542R 및 548V (및 임의로는 552R), 예컨대 S542R 및 K548V (및 임의로는 N552R)인 돌연변이된 아미노산 잔기; 또는 LbCas13의 위치 532, 538, 542, 및/또는 595, 또는 AsCas13 오솔로그, 상동체, 또는 변이체의 상응하는 위치를 포함하는 아미노산 잔기에서의 돌연변이, 바람직하게 위치 532 또는 532 및 595에서 돌연변이된 아미노산 잔기로서, 상기 돌연변이는 바람직하게 532R 및 595R, 예컨대 G532R 및 K595R인 돌연변이된 아미노산 잔기; 또는 바람직하게 위치 532 및 538 (및 임의로는 542)에서 돌연변이된 아미노산 잔기로서, 상기 돌연변이는 바람직하게 532R 및 538V (및 임의로는 542R), 예컨대 G532R 및 K538V (및 임의로는 Y542R)인 돌연변이된 아미노산 잔기에 의해 변형되고, 가장 바람직하게 상기 돌연변이는 AsCas13의 S542R 및 K607R, S542R 및 K548V, 또는 S542R, K548V 및 N552R이다.

[0348] **탈활성화/불활성화 Cas13 단백질**

[0349] Cas13 단백질이 뉴클레아제 활성을 갖는 경우에, Cas13 단백질은 야생형 효소에 비해 감소된 뉴클레아제 활성, 예를 들어, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95%, 적어도 97% 또는 100%의 뉴클레아제 불활성화를 갖도록 변형될 수 있거나; 또는 달리 말해서, Cas13 효소는 유리하게는 비돌연변이 또는 야생형 Cas13 효소 또는 CRISPR-Cas 단백질의 뉴클레아제 활성의 약 0%, 또는 비돌연변이 또는 야생형 Cas13 효소의 뉴클레아제 활성, 예를 들어, 비돌연변이 또는 야생형 프란시셀라 노비시다 U112 (FnCas13), 아시다미노코커스 sp. BV3L6 (AsCas13), 라크노스피라과 박테리아 ND2006 (LbCas13) 또는 모라셀라 보보쿨리 237 (MbCas13 Cas13 효소 또는 CRISPR-Cas 단백질의 뉴클레아제 활성의 약 3% 또는 약 5% 또는 약 10% 이하를 갖는다. 이것은 돌연변이를 Cas13 및 이의 오솔로그의 뉴클레아제 도메인 내로 도입시켜 가능하다.

[0350] 본 발명의 바람직한 실시형태에서, Cas13 닉카제인 적어도 하나의 Cas13 단백질이 사용된다. 보다 특히, Cas13 닉카제는 표적 가닥을 절단하지 않지만, 표적 가닥에 상보성인 가닥, 즉, 가이드 서열에 상보성이 아닌 가닥으로서 본 명세서에서 또한 언급되는 비표적 DNA 가닥만을 절단할 수 있는 것이 사용된다. 보다 특히, Cas13 닉카제는 아시다미노코커스 sp. 유래 Cas13의 Nuc 도메인 내 위치 1226A의 아르기닌, 또는 Cas13 오솔로그 내 상응하는 위치에 돌연변이를 포함하는 Cas13 단백질이다. 추가의 특정 실시형태에서, 효소는 아르기닌-알라닌 (arginine-to-alanine) 치환 또는 R1226A 돌연변이를 포함한다. 효소가 AsCas13이 아닌 경우에, 돌연변이는 상응하는 위치의 잔기에서 만들어질 수 있다는 것을 당업자는 이해하게 될 것이다. 특정한 실시형태에서, Cas13은 FnCas13이고, 돌연변이는 위치 R1218의 아르기닌에서 존재한다. 특정 실시형태에서, Cas13은 LbCas13이고, 돌연변이는 위치 R1138의 아르기닌에서 존재한다. 특정 실시형태에서, Cas13은 MbCas13이고, 돌연변이는 위치 R1293의 아르기닌에서 존재한다.

[0351] 일정 실시형태에서, 추가로 또는 대안적으로 뉴클레아제 활성이 감소 또는 제거된 하나 이상의 돌연변이를 포함할 수 있고, 조작된 CRISPR-Cas 단백질이 이용된다. FnCas13p RuvC 도메인 내 아미노산 위치는 제한없이 D917A, E1006A, E1028A, D1227A, D1255A, N1257A, D917A, E1006A, E1028A, D1227A, D1255A 및 N1257A를 포함한다. 출원인은 또한 PD-(D/E)XK 뉴클레아제 수퍼패밀리와 가장 유사하고 HincII 엔도뉴클레아제와 비슷한 추정의 제2 뉴클레아제 도메인을 확인하였다. 뉴클레아제 활성을 실질적으로 감소시키기 위해 이러한 추정의 뉴클레아제 도메인에 생성시키려는 점 돌연변이는 제한없이, N580A, N584A, T587A, W609A, D610A, K613A, E614A, D616A, K624A, D625A, K627A 및 Y629A를 포함한다. 바람직한 실시형태에서, FnCas13p RuvC 도메인 내 돌연변이는 D917A 또는 E1006A이고, 여기서 D917A 또는 E1006A 돌연변이는 FnCas13 이펙터 단백질의 DNA 절단 활성을 완전하게 불활성화시킨다. 다른 실시형태에서, FnCas13p RuvC 도메인 내 돌연변이는 D1255A이고, 여기서 돌연변이된 FnCas13 이펙터 단백질은 핵산용해 활성을 상당히 감소시켰다.

[0352] 보다 특히, 불활성화된 Cas13 효소는 AsCas13의 아미노산 위치 As908, As993, As1263 또는 Cas13 오솔로그 내 상응하는 위치에서 돌연변이된 효소를 포함한다. 추가로, 불활성화된 Cas13 효소는 LbCas13의 아미노산 위치 Lb832, 925, 947 또는 1180 또는 Cas13 오솔로그 내 상응하는 위치에서 돌연변이된 효소를 포함한다. 보다 특히, 불활성화된 Cas13은 AsCas13의 돌연변이 AsD908A, AsE993A, AsD1263A 또는 Cas13 오솔로그 내 상응하

는 돌연변이 중 하나 이상을 포함하는 효소를 포함한다. 추가적으로, 불활성화된 Cas13 효소는 LbCas13의 돌연변이 LbD832A, E925A, D947A 또는 D1180A 또는 Cas14 오솔로그 내 상응하는 돌연변이 중 하나 이상을 포함하는 효소를 포함한다.

[0353] 돌연변이는 또한 이웃 잔기, 예를 들어 뉴클레아제 활성에 관여하는 상기 표시된 것 근처 아미노산에서 만들어질 수 있다. 일부 실시형태에서, 오직 RuvC 도메인만이 불활성화되고, 다른 실시형태에서는, 다른 추정의 뉴클레아제 도메인이 불활성화되며, 여기서 이펙터 단백질 복합체는 닉카제로 기능하고 오직 하나의 DNA 가닥만을 절단한다. 바람직한 실시형태에서, 다른 추정의 뉴클레아제 도메인은 HincII-유사 엔도뉴클레아제 도메인이다.

[0354] 불활성화된 Cas13 또는 Cas13 닉카제는 예를 들어 아데노신 데아미나제 또는 이의 촉매 도메인을 포함하여, (예를 들어, 융합 단백질을 통해서) 회합된 하나 이상의 기능성 도메인을 가질 수 있다. 일부 경우에서, 추가로 적어도 하나의 이종성 NLS가 제공되는 것이 유리하다. 일부 예에서, N 말단에 NLS를 위치시키는 것이 유리하다. 일반적으로, 불활성화된 Cas13 또는 Cas13 닉카제 상에 하나 이상의 기능성 도메인의 위치잡이는 기여되는 기능적 효과로 표적에 영향을 미칠 수 있도록 기능적 도메인에 대한 공간적 배향을 교정할 수 있게 하는 것이다. 예를 들어, 기능성 도메인이 아데노신 데아미나제의 촉매 도메인일 때, 아데노신 데아미나제 촉매 도메인은 표적 아데닌과 접촉하여 디아민노화를 허용하는 공간 배향으로 위치된다. 이것은 Cas13의 N-말단/C-말단 이외의 위치를 포함할 수 있다. 일부 실시형태에서, 아데노신 데아미나제 단백질 또는 이의 촉매 도메인은 Cas13의 내부 루프에 삽입된다.

[0355] **PAM의 결정**

[0356] PAM의 결정은 다음과 같이 보장될 수 있다. 이 실험은 StCas9의 이종성 발현을 위해 이. 콜라이에서 유사한 작업을 긴밀히 병행한다 (Sapranauskas, R. et al. Nucleic Acids Res 39, 9275-9282 (2011)). 출원인은 PAM과 내성 유전자 둘 모두를 함유하는 플라스미드를 이종성 이. 콜라이 내로 도입하고, 이어서, 상응하는 항생제 상에서 플레이팅한다. 플라스미드의 DNA 절단이 있다면, 출원인은 생존 가능한 콜로니를 관찰하지 못한다.

[0357] 추가 상세한 설명에서, 어세이는 DNA 표적에 대해 다음과 같다. 2종의 이. 콜라이 균주가 본 어세이에서 사용된다. 하나는 박테리아 균주로부터 내생성 이펙터 단백질 유전자좌를 코딩하는 플라스미드를 운반한다. 다른 균주는 빈 플라스미드 (예를 들어, pACYC184, 대조군 균주)를 운반한다. 모든 가능한 7 또는 8 bp PAM 서열은 항생체 내성 플라스미드 (암피실린 내성 유전자를 갖는 pUC19) 상에 존재한다. PAM은 프로토-스페이서 1 (내생성 이펙터 단백질 유전자좌에서 제1 스페이서에 대한 DNA 표적)의 서열 다음에 위치된다. 2개의 PAM 라이브러리를 클로닝하였다. 하나는 프로토-스페이서의 5'에 8개의 무작위 bp를 갖는다 (예를 들어, 총 65536개의 상이한 PAM 서열 = 복잡성). 다른 라이브러리는 프로토-스페이서의 3'에 7개의 무작위 bp를 갖는다(예를 들어, 총 복잡성은 16384 상이한 PAM임). 라이브러리는 둘 모두 가능한 PAM 당 평균 500개의 플라스미드를 갖도록 클로닝되었다. 시험 균주 및 대조군 균주는 별개의 형질전환에서 5'PAM 및 3'PAM 라이브러리로 형질전환되었고, 형질전환된 세포는 암피실린 플레이트 상에서 별개로 플레이팅되었다. 플라스미드에 의한 인식 및 후속적 절단/간섭은 암피실린에 취약한 세포를 제공하며, 성장을 방지한다. 형질전환의 대략 12시간 후에, 채취한 시험 및 대조군 균주에 의해 형성되는 모든 콜로니 및 플라스미드 DNA가 단리되었다. 플라스미드 DNA는 PCR 증폭 및 후속적 심층 시퀀싱을 위한 주형으로서 사용되었다. 비형질전환 라이브러리에서 모든 PAM의 제시는 형질전환 세포에서 PAM의 예상되는 표현을 나타내었다. 대조군 균주에서 발견되는 모든 PAM의 표현은 실제 표현을 보였다. 시험 균주에서 모든 PAM의 표현은 PAM이 효소에 의해 인식되지 않으며 대조군 균주에 대한 비교가 감손된 PAM 서열의 추출을 허용한다는 것을 나타낸다.

[0358] 다음의 PAM은 소정의 야생형 Cas13 오솔로그에 대해 동정되었다: 아시다미노코커스 sp. BV3L6 Cas13(AsCas13), 라크노스피라과 박테리아 ND2006 Cas13(LbCas13) 및 프레보텔라 알벤시스(PaCas13)는 TTTV PAM 앞의 표적 부위를 절단할 수 있고 (여기서, V는 A/C 또는 G임), FnCas13p는 TTN 앞의 부위를 절단할 수 있다 (여기서, N은 A/C/G 또는 T임). 모라셀라 보보쿨리 AAX08_00205, 모라셀라 보보쿨리 AAX11_00205, 부티리비브리오 종 NC3005, 티오미크로스피라 종 XS5, 또는 라크노스피라과 박테리아 MA2020 PAM은 5' TTN이며, 여기서 N은 A/C/G 또는 T이다. 천연 PAM 서열은 TTTV 또는 BTTV이고, 여기서, B는 T/C 또는 G이고, V는 A/C 또는 G이며, 이펙터 단백질은 모라셀라 라쿠나타 Cas13이다.

[0359] ***코돈 최적화된 핵산 서열***

[0360] 이펙터 단백질이 핵산으로서 투여되는 경우에, 본 출원은 코돈-최적화된 CRISPR-Cas V형 단백질, 보다 특히 Cas13-코딩 핵산 서열 (및 임의로는 단백질 서열)의 사용을 예상한다. 코돈-최적화된 서열의 예로서 예를 들어 진핵생물, 예를 들어 인간 (즉, 인간에서 발현을 위해 최적화됨), 또는 다른 진핵생물의 경우에, 본 명세서에 논의된 바와 같은 동물 또는 포유동물에서 에서의 발현을 위해 최적화된 서열이고; 예를 들어 WO 2014/093622 (PCT/US2013/074667)의 SaCas9 인간 코돈-최적화된 서열을 코돈 최적화된 서열의 예로서 참조한다 (당업계 및 본 개시의 지식으로부터, 특히 이펙터 단백질 (예를 들어, Cas13)에 대한, 코돈 최적화 코딩 핵산 분자(들)는 당업자의 영역 내임). 이것이 바람직하지만, 다른 예가 가능하며, 인간 이외의 숙주 종에 대한 코돈 최적화 또는 특정 장기에 대한 코돈 최적화가 공지되어 있다는 것을 이해할 것이다. 일부 실시형태에서, DNA/RNA-표적화 Cas 단백질을 코딩하는 효소 코딩 서열은 특정 세포, 예컨대, 진핵생물 세포에서의 발현을 위해 코돈 최적화된다. 진핵생물 세포는 특정 유기체, 예컨대, 제한없이 인간, 또는 비-인간 진핵생물 또는 본 명세서에서 논의되는 동물 또는 포유동물, 예를 들어, 마우스, 래트, 토끼, 개, 가축 또는 인간이외의 포유동물 또는 영장류를 포함하는, 포유동물 또는 식물의 것일 수 있거나 또는 그로부터 유래될 수 있다. 일부 실시형태에서, 인간의 생식 계통 유전적 동일성을 변형시키기 위한 과정 및/또는 사람 또는 동물, 및 이러한 과정으로부터 초래된 동물에 임의의 실질적인 의학적 이점 없이 그들에게 고통을 야기할 가능성이 있는 동물의 유전적 동일성을 변형시키는 과정은 배제될 수 있다. 일반적으로, 코돈 최적화는 천연 서열의 적어도 하나의 코돈 (예를 들어, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 50개 이상의 코돈)을 해당 숙주 세포의 유전자에서 더 빈번하게 또는 가장 빈번하게 사용되는 코돈으로 대체함으로써 관심 숙주 세포에서 증강된 발현을 위한 핵산 서열을 변형시키는 한편, 천연 아미노산 서열을 유지시키는 과정을 의미한다. 다양한 종은 특정 아미노산의 소정의 코돈에 대한 특정 편향성을 나타낸다. 코돈 편향성 (유기체 사이의 코돈 용법 차이)은 종종 메신저 RNA (mRNA)의 번역 효율과 상관관계가 있는데, 이는 결국 특히 번역 중인 코돈의 특성 및 특정 전달 RNA(tRNA) 분자의 이용 가능성에 의존하는 것으로 여겨진다. 세포에서 선택된 tRNA의 우세성은 일반적으로 펩티드 합성에서 가장 빈번하게 사용되는 코돈의 반영이다. 따라서, 유전자는 코돈 최적화에 기반하여 주어진 유기체에서 최적의 유전자 발현을 위해 맞춤될 수 있다. 코돈 용법 표는, 예를 들어, www.kazusa.orjp/codon/에서 이용 가능한 "코돈 용법 데이터베이스"에서 용이하게 이용 가능하며, 이들 표는 다수의 방법에서 적합하게 될 수 있다. 문헌 [Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000)] 참조한다. 특정 숙주 세포에서 발현을 위해 특정 서열을 최적화하는 코돈에 대한 컴퓨터 알고리즘을 또한 이용 가능하며, 예컨대, Gene Forge (Aptagen; Jacobus, PA)가 또한 이용 가능하다. 일부 실시형태에서, DNA/RNA-표적화 Cas 단백질을 코딩하는 서열에서 하나 이상의 코돈(예를 들어, 1, 2, 3, 4, 5, 10, 15, 20, 25, 50개 이상, 또는 모든 코돈)은 특정 아미노산에 대해 가장 빈번하게 사용되는 코돈에 대응한다. 효모에서의 코돈 용법에 관해서, http://www.yeastgenome.org/community/codon_usage.shtml에서 입수가능한 온라인 Yeast Genome 데이타베이스 또는 [Codon selection in yeast, Bennetzen and Hall, J Biol Chem. 1982 Mar 25;257(6):3026-31]를 참조한다. 조류를 포함한 식물에서 코돈 요법에 관해서, [Codon usage in higher plants, green algae, and cyanobacteria, Campbell and Gowri, Plant Physiol. 1990 Jan; 92(1): 1-11.]; 및 [Codon usage in plant genes, Murray et al, Nucleic Acids Res. 1989 Jan 25;17(2):477-98]; 또는 [Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages, Morton BR, J Mol Evol. 1998 Apr;46(4):449-59]을 참조한다.

[0361] 일정 예의 실시형태에서, CRISPR Cas 단백질은 하기 **표 1**로부터 선택된다.

[0362] 표 1

C2c2 오솔로그	코드	다글자
렙토트리키아 샤히	C2-2	Lsh
엘. 웨이데이 F0279 (Lw2)	C2-3	Lw2
리스테리아 실리게리	C2-4	Lse
라크노스피라과 박테리아 MA2020	C2-5	LbM
라크노스피라과 박테리아 NK4A179	C2-6	LbNK179
[클로스트리듐] 아미노필럼 DSM 10710	C2-7	Ca
카르노박테리움 갈리나럼 DSM 4847	C2-8	Cg
카르노박테리움 갈리나럼 DSM 4847	C2-9	Cg2
팔루디박터 프로피오니시게네스 WB4	C2-10	Pp
리스테리아 웨이헨스테파넨시스 FSL R9-0317	C2-11	Lwei
리스테리아과 박테리아 FSL M6-0635	C2-12	LbFSL
렙토트리키아 웨이데이 F0279	C2-13	Lw
로도박터 캡슐라터스 SB 1003	C2-14	Rc
로도박터 캡슐라터스 R121	C2-15	Rc
로도박터 캡슐라터스 DE442	C2-16	Rc

[0363]

[0364] 일정 예의 실시형태에서, CRISPR 이펙터 단백질은 하기 표 2로부터 선택되는 Cas13a 단백질이다.

[0365] 표 2

c2c2-5	1	라크노스피라과 박테리아 MA2020 (SEQ ID No. 34)	MQISKVNHKHVAVGQKDRERITGFIYNDPVGDEKSLEDVVA KRANDTKVLFNVFNTKDLYDSQESDKSEKDKEIISKGAKFV AKSFNSAITILKKQNKIYSTLTSQQVIKELKDKFGGARIYDDD IEEALTETLKKSFRKENVRNSIKVLIENAAGIRSSLSKDEEELI QEYFVKQLVEEYTKTKLQKNVVKSIKNQNMVIQPDSDSQVL SLSESRREKQSSAVSSDTLVNCKEKDVLKAFLTDYAVLDEDE RNSLLWKLRNLVNLYFYGSESIRDYSYTKEKSVWKEHDEQK ANKTLFIDEICHITKIGKNGKEQKVLDYEENRSRCRKQNINY YRSALNYAKNNTSGIFENEDSNHFWIHLIENEVERLYNGIEN GEEFKFETGYISEKVWKAVINHLSIKYIALGKAVYNYAMKEL SSPGDIEPGKIDDSYINGITSFDYEIIKAEESLQRDISMNVVFAT NYLACATVDTDKDFLLFSKEDIRSCTKKDGNLCKNIMQFWG GYSTWKNFCEEYLKDDKDALELLYSLKSMLYSMRNSSFHFS TENVDNGSWDTELIGKLFEEDCNRAARIEKEKFYNNNLHMF YSSSLLEKVLERLYSSHHERASQVPSFNRVFVRKNFPSSLSEQ RITPKFTDSKDEQIWQSAVYYLCKEIYYNDFLQSKEAYKLFR EGVKNLDKNDINNQKAADSFKQAVVYYGKAIGNATLSQVC QAIMTEYNRQNNDGLKKKSAYAEKQNSNKYKHYPLFLKQV LQSAFWEYLDENKEIYGFISAQIHKSNVEIKAEDFIANYSSQQ YKKLVDKVKKTPELQKWYTLGRLINPRQANQFLGSIRNYVQ FVKDIQRRAKENGNPIRNYYEVLESDSIIKILEMCTKLNGTTS NDIHDYFRDEDEYAEYISQFVNFGDVHSGAALNAFCNSESEG KKNGIYYDGINPIVNRNWVLCKLYGSPDLISKIISRVNENMIH DFHKQEDLIREYQIKGICSNKKEQQDLRTFQVLKNRVELRDI VEYSEIINELYGQLIKWCYLRERDLMYFQLGFHYLCLNNASS KEADYIKINVDDRNISGAILYQIAAMYINGLPVYYKKDDMY VALKSGKKASDELNSNEQTSKKINYFLKYGNNILGDKKDQL YLAGLELFENVAEHENIIIFRNEIDHFHYFYDRDRSMLDLYSE VFDRFFTYDMKLRKNVVNMLYNILLDHNIVSSFVFETGEKK VGRGDSEVIKPSAKIRLRANNGVSSDVFTYKVGSKDELKIAT LPAKNEEFLLNVARLIYYPDMEAVSENMVREGVVKVEKSND KKGKISRGSNTRSSNQSKYNNKSKNRMNYSMGSIFEKMDLK

[0366]

			FD
c2c2-6	2	라크노스피라과 박테리아 NK4A179 (SEQ ID No. 35)	MKISKVREENRGAKLTVNAKTAVVSENRSQEGILYNDPSRY GKSRKNDEDRDRYIESRLKSSGKLYRIFNEDKNKRETDELQ WFLSEIVKKINRRNGLVLSDMLSVDDRAFEKAFEKYAELSYT NRRNKVSGSPAFETCGVDAATAERLKGIISETNFINRIKNNID NKVSEDIIDRIIAKYLKKSLCRERVKRGLKKLLMNAFDLPYS DPDIDVQRDFIDYVLEDFYHVRAKSQVSRSIKNMNMPVQPE GDGKFAITVSKGGTESGNKRSAEKEAFKKFLSDYASLDERV RDDMLRRMRRLVVLYFYGSDDSKLSDVNEKFDVWEDHAA RRVDNREFIKLPLENKLANGKTDKDAERIRKNTVKELYRNQ NIGCYRQAVKAVEEDNNGRYFDDKMLNMFFIHRIEYGVEKI YANLKQVTEFKARTGYLSEKIWKDLINYISIKYIAMGKAVYN YAMDELNASDKKEIELGKISEEYLSGISSFDYELIKAEEMLQR ETAVYVAFAARHLSSQTVELDSENSDFLLLKPKGTMDKNDK NKLASNNILNFLKDKETLRDTILQYFGGHSLWTDFPFDKYLA GGKDDVDFLTDLKDVIYSMRNDSFHYATENHNNGKWNKEL ISAMFEHETERMTVVMKDKFYSNNLPMFYKNDDLKKLLIDL YKDNVERASQVPSFNKVFVRKNFPALVRDKDNLGIELDLKA DADKGENELKFYNALYYMFKEIYYNAFLNDKNVRERFITKA TKVADNYDRNKERNLKDRIKSAGSDEKKKLREQLQNYIAEN DFGQRIKNIVQVNPDYTLAQICQLIMTEYNQQNNGCMQKKS AARKDINKDSYQHYKMLLLVNLRKAFLEFIKENYAFVLKPY KHDLCDKADFVPDFAKYVKPYAGLISRVAGSSELQKWYIVS RFLSPAQANHMLGFLHSYKQYVWDIYRRASETGTEINHSIAE DKIAGVDITDVDAVIDLSVKLCGTISSEISDYFKDDEVYAEYI SSYLDFEYDGGNYKDSLNRFCNSDAVNDQKVALYYDGEHP KLNRNIILSKLYGERRFLEKITDRVSRSDIVEYYKLKKETSQY QTKGIFDSEDEQKNIKKFQEMKNIVEFRDLMDYSEIADELQG QLINWIYLRERDLMNFQLGYHYACLNNDSNKQATYVTLDY QGKKNRKINGAILYQICAMYINGLPLYYVDKDSSEWTVSDG KESTGAKIGEFYRYAKSFENTSDCYASGLEIFENISEHDNITEL RNYIEHFRYYSSFDRSFLGIYSEVFDRFFTYDLKYRKNVPTIL

[0367]

			YNILLQHFVNVRFEFVSGKKMIGIDKKDRKIAKEKECARITIR EKNGVYSEQFTYKLKNGTVYVDARDKRYLQSIIRLLFYPEK VNMDEMIEVKEKKKPSDNNTGKGYSKRDRQQDRKEYDKY KEKKKKEGNFLSGMGGNINWDEINAQLKN
c2c2-7	3	[클로스트리듐] 아미노필럼 DSM 10710 SEQ ID No. 36)	MKFSKVDHTRSAVGIQKATDSVHGMLYTDPKKQEVNDLDK RFDQLNVKAKRLYNVFNQSKAEEDDDEKRFGKVVKKLNRE LKDLLFHREVSRYNSIGNAKYNYYGIKSNPEEIVSNLGMVES LKGERDPQKVISKLLLYYLRKGLKPGTDGLRMILEASCGLRK LSGDEKELKVFLQTLDEDFEKKTFKKNLIRSIENQNMAVQPS NEGDPIIGITQGRFNSQKNEEKSAIERMMSMYADLNEDHRED VLRKLRRLNVLYFNVDTEKTEEPTLPGEVDTNPVFEVWHDH EKGKENDRQFATFAKILTEDRETRKKEKLAVKEALNDLKSAI RDHNIMAYRCSIKVTEQDKDGLFFEDQRINRFWIHHIESAVE RILASINPEKLYKLRIGYLGEKVWKDLLNYLSIKYIAVGKAV FHFAMEDLGKTGQDIELGKLSNSVSGGLTSFDYEQIRADETL QRQLSVEVAFAANNLFRAVVGQTGKKIEQSKSEENEEDFLL WKAEKIAESIKKEGEGNTLKSILQFFGGASSWDLNHFCAAYG NESSALGYETKFADDLRKAIYSLRNETFHFTTLNKGSFDWNA KLIGDMFSHEAATGIAVERTRFYSNNLPMFYRESDLKRIMDH LYNTYHPRASQVPSFNSVFVRKNFRLFLSNTLNTNTSFDTEV YQKWESGVYYLFKEIYYNSFLPSGDAHHLFFEGLRRIRKEAD NLPIVGKEAKKRNAVQDFGRRCDELKNLSLSAICQMIMTEY NEQNNGNRKVKSTREDKRKPDIFQHYKMLLLRTLQEAFAIYI RREEFKFIFDLPKTLYVMKPVEEFLPNWKSGMFDSLVERVK QSPDLQRWYVLCKFLNGRLLNQLSGVIRSYIQFAGDIQRRAK ANHNRLYMDNTQRVEYYSNVLEVVDFCIKGTSRFSNVFSDY FRDEDAYADYLDNYLQFKDEKIAEVSSFAALKTFCNEEEVK AGIYMDGENPVMQRNIVMAKLFGPDEVLKNVVPKVTREEIE EYYQLEKQIAPYRQNGYCKSEEDQKKLLRFQRIKNRVEFQTI TEFSEIINELLGQLISWSFLRERDLLYFQLGFHYLCLHNDTEK PAEYKEISREDGTVIRNAILHQVAAMYVGGLPVYTLADKKL AAFEKGEADCKLSISKDTAGAGKKIKDFFRYSKYVLIKDRML

[0368]

			TDQNQKYTIYLAGLELFENTDEHDNITDVRKYVDHFKYYAT SDENAMSILDLYSEIHDRFFTYDMKYQKNVANMLENILLRH FVLIRPEFFTGSKKVGEGKKITCKARAQIEIAENGMRSEDFTY KLSDGKKNISTCMIAARDQKYLNTVARLLYYPHEAKKSIVD TREKKNNKKTNRGDGTFNKQKGTARKEKDNGPREFNDTGF SNTPFAGFDPFRNS
c2c2-8	5	카르노박테리움 갈리나럼 DSM 4847 (SEQ ID No. 37)	MRITKVKIKLDNKLYQVTMQKEEKYGTLKLNEESRKSTAEIL RLKKASFNKSFHSKTINSQKENKNATIKKNGDYISQIFEKLVG VDTNKNIRKPKMSLTDLKDLPKKDLALFIKRKFKNDDIVEIK NLDLISLFYNALQKVPGEHFTDESWADFCQEMMPYREYKNK FIERKIILLANSIEQNKGFSINPETFSKRKRVLHQWAIEVQERG DFSILDEKLSKLAEIYNFKKMCKRVQDELNDLEKSMKKGKN PEKEKEAYKKQKNFKIKTIWKDYPYKTHIGLIEKIKENEELN QFNIEIGKYFEHYFPIKKERCTEDEPYYLNSETIATTVNYQLK NALISYLMQIGKYKQFGLENQVLDSKKLQEIGIYEGFQTKFM DACVFATSSLKNIIEPMRSGDILGKREFKEAIATSSFVNYHHF FPYFPFELKGMKDRESELIPFGEQTEAKQMQNIWALRGSVQQ IRNEIFHSFDKNQKFNLPQLDKSNFEFDASENSTGKSQSYIET DYKFLFEAEKNQLEQFFIERIKSSGALEYYPLKSLEKLFAKKE MKFSLGSQVVAFAPSYKKLVKKGHSYQTATEGTANYLGLS YYNRYELKEESFQAQYYLLKLIYQYVFLPNFSQGNSPAFRET VKAILRINKDEARKKMKKNKKFLRKYAFEQVREMEFKETPD QYMSYLQSEMREEKVRKAEKNDKGFEKNITMNFEKLLMQIF VKGFDVFLTTFAGKELLLSSEEKVIKETEISLSKKINEREKTLK ASIQVEHQLVATNSAISYWLFCKLLDSRHLNELRNEMIKFKQ SRIKFNHTQHAELIQNLLPIVELTILSNDYDEKNDSQNVDVSA YFEDKSLYETAPYVQTDDRTRVSFRPILKLEKYHTKSLIEALL KDNPQFRVAATDIQEWMHKREEIGELVEKRKNLHTEWAEG QQTLGAEKREEYRDYCKKIDRFNWKANKVTLTYLSQLHYLI TDLLGRMVGFSALFERDLVYFSRSFSELGGETYHISDYKNLS GVLRLNAEVKPIKIKNIKVIDNEENPYKGNEPEVKPFLDRLH AYLENVIGIKAVHGKIRNQTAHLSVLQLELSMIESMNNLRDL

[0369]

			MAYDRKLKNAVTKSMIKILDKHGMILKLKIDENHKNFEIESL IPKEIIHLKDKAIKTNQVSEEYCQLVLALLTTNPGNQLN
c2c2-9	6	카르노박테리움 갈리나럼 DSM 4847 (SEQ ID No. 38)	MRMTKVKINGSPVSMNRSKLNGHLVWNGTTNTVNILTKKE QSFAASFLNKTLVKADQVKGYKVLAENIFIIFEQLEKSNSEKP SVYLNNIRRLKEAGLKRFFKSKYHEEIKYTSEKNQSVPTKLN LIPLFFNAVDRIQEDKFDEKNWSYFCKEMSPYLDYKKSYLNR KKEILANSIQQNRGFSMPTAEEPNLLSKRKQLFQQWAMKFQ ESPLIQQNNFAVEQFNKEFANKINELAAVYNVDELCTAITEK LMNFDKDKSNKTRNFEIKKLWKQHPHNKDKALIKLFNQEG NEALNQFNIELGKYFEHYFPKTGKKESAESYYLNPQTIIKTVG YQLRNAFVQYLLQVGKLHQYNKGVLDSQTLQEIGMYEGFQ TKFMDACVFASSSLRNIIQATTNEDILTREKFKKELEKNVELK HDLFFKTEIVEERDENPAKKIAMTPNELDLWAIRGAVQRVR NQIFHQQINKRHEPNQLKVGSFENGDLGNVSYQKTIYQKLFD AEIKDIEIYFAEKIKSSGALEQYSMKDLEKLFSNKELTLSLGG QVVAFAPSYKKLYKQGYFYQNEKTIELEQFTDYDFSNDVFK ANYYLIKLIYHYVFLPQFSQANNKLFKDTVHYVIQQNKELNT TEKDKKNNKKIRKYAFEQVKLMKNESPEKYMQYLQREMQE ERTIKEAKKTNEEKPNYNFEKLLIQIFIKGFDTFLRNFDLNLNP AEELVGTVKEKAEGLRKRKERIAKILNVDEQIKTGDEEIAFW IFAKLLDARHLSELRNEMIKFKQSSVKKGLIKNGDLIEQMQPI LELCILSNDSESMEKESFDKIEVFLEKVELAKNEPYMQEDKL TPVKFRFMKQLEKYQTRNFIENLVIENPEFKVSEKIVLNWHE EKEKIADLVDKRTKLHEEWASKAREIEEYNEKIKKNKSKKL DKPAEFAKFAEYKIICEAIENFNRLDHKVRLTYLKNLHYLMI DLMGRMVGFSVLFERDFVYMGRSYSALKKQSIYLNDYDTF ANIRDWEVNENKHLFGTSSSDLTFQETAEFKNLKKPMENQL KALLGVTNHSFEIRNNIAHLHVLRNDGKGEGVSLLSCMNDL RKLMSYDRKLKNAVTKAIIKILDKHGMILKLTNNDHTKPFEI ESLKPKKIIHLEKSNHSFPMDQVSQEYCDLVKKMLVFTN
c2c2-	7	팔루디박터	MRVSKVKVKDGGKDKMVLVHRKTTGAQLVYSGQPVSNET

[0370]

10		프로피오니시게네스 WB4 (SEQ ID No. 39)	SNILPEKKRQSFDLSTLNKTIIKFDTAKKQKLNVDQYKIVEKI FKYPKQELPKQIKAEEILPFLNHKFQEPVKYWKNGKEESFNL TLLIVEAVQAQDKRKLQPYYDWKTWYIQTKSDLLKKSIENN RIDLTENLSKRKKALLAWETEFTASGSIDLTHYHKVYMTDV LCKMLQDVKPLTDDKGKINTNAYHRGLKKALQNHQPAIFGT REVPNEANRADNQLSIYHLEVVKYLEHYFPIKTSKRRNTADD IAHYLKAQTLKTTIEKQLVNAIRANIIQQGKTNHHELKADTT SNDLIRIKTNEAFVLNLTGTCAFAANNIRNMVDNEQTNDILG KGDFIKSLLKDNTNSQLYSFFFGEGLSTNKAEKETQLWGIRG AVQQIRNNVNHYKKDALKTVFNISNFENPTITDPKQQTNYA DTIYKARFINELEKIPEAFAQQLKTGGAVSYYTIENLKSLLTT FQFSLCRSTIPFAPGFKKVFNGGINYQNAKQDESFYELMLEQ YLRKENFAEESYNARYFMLKLIYNNLFLPGFTTDRKAFADSV GFVQMQNKKQAEKVNPRKKEAYAFEAVRPMTAADSIADY MAYVQSELMQEQNKKEEKVAEETRINFEKFVLQVFIKGFDS FLRAKEFDFVQMPQPQLTATASNQQKADKLNQLEASITADC KLTPQYAKADDATHIAFYVFCKLLDAAHLSNLRNELIKFRES VNEFKFHHLLEIIEICLLSADVVPTDYRDLYSSEADCLARLRP FIEQGADITNWSDLFVQSDKHSPVIHANIELSVKYGTTKLLEQ IINKDTQFKTTEANFTAWNTAQKSIEQLIKQREDHHEQWVK AKNADDKEKQERKREKSNFAQKFIEKHGDDYLDICDYINTY NWLDNKMHFVHLNRLHGLTIELLGRMAGFVALFDRDFQFF DEQQIADEFKLHGFVNLHSIDKKLNEVPTKKIKEIYDIRNKIIQ INGNKINESVRANLIQFISSKRNYYNNAFLHVSNDEIKEKQM YDIRNHIAHFNYLTKDAADFSLIDLINELRELLHYDRKLKNA VSKAFIDLFDKHGMILKLKLNADHKLKVESLEPKKIYHLGSS AKDKPEYQYCTNQVMMAYCNMCRSLLEMKK
c2c2-11	9	리스테리아 웨이헨스테파넨시스 FSL R9-0317 (SEQ	MLALLHQEVPSQKLHNLKSLNTESLTKLFKPKFQNMISYPPS KGAEHVQFCLTDIAVPAIRDLDEIKPDWGIFFEKLKPYTDWA ESYIHYKQTTIQKSIEQNKIQSPDSPRKLVLQKYVTAFLNGEP LGLDLVAKKYKLADLAESFKVVDLNEDKSANYKIKACLQQ HQRNILDELKEDPELNQYGIEVKKYIQRYFPIKRAPNRSKHA

[0371]

		ID No. 40)	RADFLKKELIESTVEQQFKNAVYHYVLEQGKMEAYELTDPK TKDLQDIRSGEAFSFKFINACAFASNNLKMILNPECEKDILGK GDFKKNLPNSTTQSDVVKKMIPFFSDEIQNVNFDEAIWAIRG SIQQIRNEVYHCKKHSWKSILKIKGFEFEPNNMKYTDSDMQK LMDKDIAKIPDFIEEKLKSSGIIRFYSHDKLQSIWEMKQGFSL LTTNAPFVPSFKRVYAKGHDYQTSKNRYYDLGLTTFDILEY GEEDFRARYFLTKLVYYQQFMPWFTADNNAFRDAANFVLR LNKNRQQDAKAFINIREVEEGEMPRDYMGYVQGQIAIHEDS TEDTPNHFEKFISQVFIKGFDSHMRSADLKFIKNPRNQGLEQS EIEEMSFDIKVEPSFLKNKDDYIAFWTFCKMLDARHLSELRN EMIKYDGHLTGEQEIIGLALLGVDSRENDWKQFFSSEREYEK IMKGYVGEELYQREPYRQSDGKTPILFRGVEQARKYGTETVI QRLFDASPEFKVSKCNITEWERQKETIEETIERRKELHNEWE KNPKKPQNNAFFKEYKECCDAIDAYNWHKNKTTLVYVNEL HHLLIEILGRYVGYVAIADRDFQCMANQYFKHSGITERVEY WGDNRLKSIKKLDTFLKKEGLFVSEKNARNHIAHLNYLSLK SECTLLYLSERLREIFKYDRKLKNAVSKSLIDILDRHGMSVVF ANLKENKHRLVIKSLEPKKLRHLGEKKIDNGYIETNQVSEEY CGIVKRLLEI
c2c2-12	10	리스테리아과 박테리아 FSL M6-0635 = 리스테리아 뉴요켄시스 FSL M6-0635 (SEQ ID No. 41)	MKITKMRVDGRTIVMERTSKEGQLGYEGIDGNKTTEIIFDKK KESFYKSILNKTVRKPDEKEKNRRKQAINKAINKEITELMLA VLHQEVPSQKLHNLKSLNTESLTKLFKPKFQNMISYPPSKGA EHVQFCLTDIAVPAIRDLDEIKPDWGIFFEKLKPYTDWAESYI HYKQTTIQKSIEQNKIQSPDSPRKLVLQKYVTAFLNGEPLGL DLVAKKYKLADLAESFKLVDLNEDKSANYKIKACLQQHQR NILDELKEDPELNQYGIEVKKYIQRYFPIKRAPNRSKHARADF LKKELIESTVEQQFKNAVYHYVLEQGKMEAYELTDPKTKDL QDIRSGEAFSFKFINACAFASNNLKMILNPECEKDILGKGNFK KNLPNSTTRSDVVKKMIPFFSDELQNVNFDEAIWAIRGSIQQI RNEVYHCKKHSWKSILKIKGFEFEPNNMKYADSDMQKLMD KDIAKIPEFIEEKLKSSGVVRFYRHDELQSIWEMKQGFSLLTT NAPFVPSFKRVYAKGHDYQTSKNRYYNLDLTTFDILEYGEE

[0372]

			DFRARYFLTKLVYYQQFMPWFTADNNAFRDAANFVLRLNK NRQQDAKAFINIREVEEGEMPRDYMGYVQGQIAIHEDSIEDT PNHFEKFISQVFIKGFDRHMRSANLKFIKNPRNQGLEQSEIEE MSFDIKVEPSFLKNKDDYIAFWIFCKMLDARHLSELRNEMIK YDGHLTGEQEIIGLALLGVDSRENDWKQFFSSEREYEKIMKG YVVEELYQREPYRQSDGKTPILFRGVEQARKYGTETVIQRLF DANPEFKVSKCNLAEWERQKETIEETIKRRKELHNEWAKNP KKPQNNAFFKEYKECCDAIDAYNWHKNKTTLAYVNELHHL LIEILGRYVGYVAIADRDFQCMANQYFKHSGITERVEYWGD NRLKSIKKLDTFLKKEGLFVSEKNARNHIAHLNYLSLKSECT LLYLSERLREIFKYDRKLKNAVSKSLIDILDRHGMSVVFANL KENKHRLVIKSLEPKKLRHLGGKKIDGGYIETNQVSEEYCGI VKRLLEM
c2c2-13	12	렙토트리키아 웨이데이 F0279 (SEQ ID No. 42)	MKVTKVDGISHKKYIEEGKLVKSTSEENRTSERLSELLSIRLD IYIKNPDNASEEENRIRRENLKKFFSNKVLHLKDSVLYLKNR KEKNAVQDKNYSEEDISEYDLKNKNSFSVLKKILLNEDVNSE ELEIFRKDVEAKLNKINSLKYSFEENKANYQKINENNVEKVG GKSKRNIIYDYYRESAKRNDYINNVQEAFDKLYKKEDIEKLF FLIENSKKHEKYKIREYYHKIIGRKNDKENFAKIIYEEIQNVN NIKELIEKIPDMSELKKSQVFYKYYLDKEELNDKNIKYAFCH FVEIEMSQLLKNYVYKRLSNISNDKIKRIFEYQNLKKLIENKL LNKLDTYVRNCGKYNYYLQVGEIATSDFIARNRQNEAFLRNI IGVSSVAYFSLRNILETENENDITGRMRGKTVKNNKGEEKYV SGEVDKIYNENKQNEVKENLKMFYSYDFNMDNKNEIEDFFA NIDEAISSIRHGIVHFNLELEGKDIFAFKNIAPSEISKKMFQNEI NEKKLKLKIFKQLNSANVFNYYEKDVIIKYLKNTKFNFVNK NIPFVPSFTKLYNKIEDLRNTLKFFWSVPKDKEEKDAQIYLLK NIYYGEFLNKFVKNSKVFFKITNEVIKINKQRNQKTGHYKYQ KFENIEKTVPVEYLAIIQSREMINNQDKEEKNTYIDFIQQIFLK GFIDYLNKNNLKYIESNNNNDNNDIFSKIKIKKDNKEKYDKIL KNYEKHNRNKEIPHEINEFVREIKLGKILKYTENLNMFYLILK LLNHKELTNLKGSLEKYQSANKEETFSDELELINLLNLDNNR

[0373]

			VTEDFELEANEIGKFLDFNENKIKDRKELKKFDTNKIYFDGE NIIKHRAFYNIKKYGMLNLLEKIADKAKYKISLKELKEYSNK KNEIEKNYTMQQNLHRKYARPKKDEKFNDEDYKEYEKAIG NIQKYTHLKNKVEFNELNLLQGLLLKILHRLVGYTSIWERDL RFRLKGEFPENHYIEEIFNFDNSKNVKYKSGQIVEKYINFYKE LYKDNVEKRSIYSDKKVKKLKQEKKDLYIRNYIAHFNYIPHA EISLLEVLENLRKLLSYDRKLKNAIMKSIVDILKEYGFVATFK IGADKKIEIQTLESEKIVHLKNLKKKKLMTDRNSEELCELVK VMFEYKALE
c2c2-14	15	로도박터 캡슐라터스 SB 1003 (SEQ ID No. 43)	MQIGKVQGRTISEFGDPAGGLKRKISTDGKNRKELPAHLSSD PKALIGQWISGIDKIYRKPDSRKSDGKAIHSPTPSKMQFDARD DLGEAFWKLVSEAGLAQDSDYDQFKRRLHPYGDKFQPADS GAKLKFEADPPEPQAFHGRWYGAMSKRGNDAKELAAALYE HLHVDEKRIDGQPKRNPKTDKFAPGLVVARALGIESSVLPRG MARLARNWGEEEIQTYFVVDVAASVKEVAKAAVSAAQAFD PPRQVSGRSLSPKVGFALAEHLERVTGSKRCSFDPAAGPSVL ALHDEVKKTYKRLCARGKNAARAFPADKTELLALMRHTHE NRVRNQMVRMGRVSEYRGQQAGDLAQSHYWTSAGQTEIK ESEIFVRLWVGAFALAGRSMKAWIDPMGKIVNTEKNDRDLT AAVNIRQVISNKEMVAEAMARRGIYFGETPELDRLGAEGNE GFVFALLRYLRGCRNQTFHLGARAGFLKEIRKELEKTRWGK AKEAEHVVLTDKTVAAIRAIIDNDAKALGARLLADLSGAFV AHYASKEHFSTLYSEIVKAVKDAPEVSSGLPRLKLLLKRADG VRGYVHGLRDTRKHAFATKLPPPPAPRELDDPATKARYIALL RLYDGPFRAYASGITGTALAGPAARAKEAATALAQSVNVTK AYSDVMEGRTSRLRPPNDGETLREYLSALTGETATEFRVQIG YESDSENARKQAEFIENYRRDMLAFMFEDYIRAKGFDWILKI EPGATAMTRAPVLPEPIDTRGQYEHWQAALYLVMHFVPASD VSNLLHQLRKWEALQGKYELVQDGDATDQADARREALDL VKRFRDVLVLFLKTGEARFEGRAAPFDLKPFRALFANPATFD RLFMATPTTARPAEDDPEGDGASEPELRVARTLRGLRQIARY NHMAVLSDLFAKHKVRDEEVARLAEIEDETQEKSQIVAAQE

[0374]

			LRTDLHDKVMKCHPKTISPEERQSYAAAIKTIEEHRFLVGRV YLGDHLRLHRLMMDVIGRLIDYAGAYERDTGTFLINASKQL GAGADWAVTIAGAANTDARTQTRKDLAHFNVLDRADGTPD LTALVNRAREMMAYDRKRKNAVPRSILDMLARLGLTLKWQ MKDHLLQDATITQAAIKHLDKVRLTVGGPAAVTEARFSQDY LQMVAAVFNGSVQNPKPRRRDDGDAWHKPPKPATAQSQPD QKPPNKAPSAGSRLPPPQVGEVYEGVVVKVIDTGSLGFLAVE GVAGNIGLHISRLRRIREDAIIVGRRYRFRVEIYVPPKSNTSKL NAADLVRID
c2c2-15	16	로도박터 캡슐라터스 R121 (SEQ ID No. 44)	MQIGKVQGRTISEFGDPAGGLKRKISTDGKNRKELPAHLSSD PKALIGQWISGIDKIYRKPDSRKSDGKAIHSPTPSKMQFDARD DLGEAFWKLVSEAGLAQDSDYDQFKRRLHPYGDKFQPADS GAKLKFEADPPEPQAFHGRWYGAMSKRGNDAKELAAALYE HLHVDEKRIDGQPKRNPKTDKFAPGLVVARALGIESSVLPRG MARLARNWGEEEIQTYFVVDVAASVKEVAKAAVSAAQAFD PPRQVSGRSLSPKVGFALAEHLERVTGSKRCSFDPAAGPSVL ALHDEVKKTYKRLCARGKNAARAFPADKTELLALMRHTHE NRVRNQMVRMGRVSEYRGQQAGDLAQSHYWTSAGQTEIK ESEIFVRLWVGAFALAGRSMKAWIDPMGKIVNTEKNDRDLT AAVNIRQVISNKEMVAEAMARRGIYFGETPELDRLGAEGNE GFVFALLRYLRGCRNQTFHLGARAGFLKEIRKELEKTRWGK AKEAEHVVLTDKTVAAIRAIIDNDAKALGARLLADLSGAFV AHYASKEHFSTLYSEIVKAVKDAPEVSSGLPRLKLLLKRADG VRGYVHGLRDTRKHAFATKLPPPPAPRELDDPATKARYIALL RLYDGPFRAYASGITGTALAGPAARAKEAATALAQSVNVTK AYSDVMEGRSSRLRPPNDGETLREYLSALTGETATEFRVQIG YESDSENARKQAEFIENYRRDMLAFMFEDYIRAKGFDWILKI EPGATAMTRAPVLPEPIDTRGQYEHWQAALYLVMHFVPASD VSNLLHQLRKWEALQGKYELVQDGDATDQADARREALDL VKRFRDVLVLFLKTGEARFEGRAAPFDLKPFRALFANPATFD RLFMATPTTARPAEDDPEGDGASEPELRVARTLRGLRQIARY NHMAVLSDLFAKHKVRDEEVARLAEIEDETQEKSQIVAAQE

[0375]

			LRTDLHDKVMKCHPKTISPEERQSYAAAIKTIEEHRFLVGRV YLGDHLRLHRLMMDVIGRLIDYAGAYERDTGTFLINASKQL GAGADWAVTIAGAANTDARTQTRKDLAHFNVLDRADGTPD LTALVNRAREMMAYDRKRKNAVPRSILDMLARLGLTLKWQ MKDHLLQDATITQAAIKHLDKVRLTVGGPAAVTEARFSQDY LQMVAAVFNGSVQNPKPRRRDDGDAWHKPPKPATAQSQPD QKPPNKAPSAGSRLPPPQVGEVYEGVVVKVIDTGSLGFLAVE GVAGNIGLHISRLRRIREDAIIVGRRYRFRVEIYVPPKSNTSKL NAADLVRID
c2c2-16	17	로도박터 캡슐라터스 DE442 (SEQ ID No. 45)	MQIGKVQGRTISEFGDPAGGLKRKISTDGKNRKELPAHLSSD PKALIGQWISGIDKIYRKPDSRKSDGKAIHSPTPSKMQFDARD DLGEAFWKLVSEAGLAQDSDYDQFKRRLHPYGDKFQPADS GAKLKFEADPPEPQAFHGRWYGAMSKRGNDAKELAAALYE HLHVDEKRIDGQPKRNPKTDKFAPGLVVARALGIESSVLPRG MARLARNWGEEEIQTYFVVDVAASVKEVAKAAVSAAQAFD PPRQVSGRSLSPKVGFALAEHLERVTGSKRCSFDPAAGPSVL ALHDEVKKTYKRLCARGKNAARAFPADKTELLALMRHTHE NRVRNQMVRMGRVSEYRGQQAGDLAQSHYWTSAGQTEIK ESEIFVRLWVGAFALAGRSMKAWIDPMGKIVNTEKNDRDLT AAVNIRQVISNKEMVAEAMARRGIYFGETPELDRLGAEGNE GFVFALLRYLRGCRNQTFHLGARAGFLKEIRKELEKTRWGK AKEAEHVVLTDKTVAAIRAIIDNDAKALGARLLADLSGAFV AHYASKEHFSTLYSEIVKAVKDAPEVSSGLPRLKLLLKRADG VRGYVHGLRDTRKHAFATKLPPPPAPRELDDPATKARYIALL RLYDGPFRAYASGITGTALAGPAARAKEAATALAQSVNVTK AYSDVMEGRSSRLRPPNDGETLREYLSALTGETATEFRVQIG YESDSENARKQAEFIENYRRDMLAFMFEDYIRAKGFDWILKI EPGATAMTRAPVLPEPIDTRGQYEHWQAALYLVMHFVPASD VSNLLHQLRKWEALQGKYELVQDGDATDQADARREALDL VKRFRDVLVLFLKTGEARFEGRAAPFDLKPFRALFANPATFD RLFMATPTTARPAEDDPEGDGASEPELRVARTLRGLRQIARY NHMAVLSDLFAKHKVRDEEVARLAEIEDETQEKSQIVAAQE

[0376]

			LRTDLHDKVMKCHPKTISPEERQSYAAAIKTIEEHRFLVGRV YLGDHLRLHRLMMDVIGRLIDYAGAYERDTGTFLINASKQL GAGADWAVTIAGAANTDARTQTRKDLAHFNVLDRADGTPD LTALVNRAREMMAYDRKRKNAVPRSILDMLARLGLTLKWQ MKDHLLQDATITQAAIKHLDKVRLTVGGPAAVTEARFSQDY LQMVAAVFNGSVQNPKPRRRDDGDAWHKPPKPATAQSQPD QKPPNKAPSAGSRLPPPQVGEVYEGVVVKVIDTGSLGFLAVE GVAGNIGLHISRLRRIREDAIIVGRRYRFRVEIYVPPKSNTSKL NAADLVRID
c2c2-2		(SEQ ID No. 46)	MGNLFGHKRWYEVRDKKDFKIKRKVKVKRNYDGNKYILNI NENNNKEKIDNNKFIRKYINYKKNDNILKEFTRKFHAGNILF KLKGKEGIIRIENNDDFLETEEVVLYIEAYGKSEKLKALGITK KKIIDEAIRQGITKDDKKIEIKRQENEEEIEIDIRDEYTNKTLND CSIILRIIENDELETKKSIYEIFKNINMSLYKIIEKIIENETEKVFE NRYYEEHLREKLLKDDKIDVILTNFMEIREKIKSNLEILGFVK FYLNVGGDKKKSKNKKMLVEKILNINVDLTVEDIADFVIKEL EFWNITKRIEKVKKVNNEFLEKRRNRTYIKSYVLLDKHEKFK IERENKKDKIVKFFVENIKNNSIKEKIEKILAEFKIDELIKKLEK ELKKGNCDTEIFGIFKKHYKVNFDSKKFSKKSDEEKELYKIIY RYLKGRIEKILVNEQKVRLKKMEKIEIEKILNESILSEKILKRV KQYTLEHIMYLGKLRHNDIDMTTVNTDDFSRLHAKEELDLE LITFFASTNMELNKIFSRENINNDENIDFFGGDREKNYVLDKK ILNSKIKIIRDLDFIDNKNNITNNFIRKFTKIGTNERNRILHAISK ERDLQGTQDDYNKVINIIQNLKISDEEVSKALNLDVVFKDKK NIITKINDIKISEENNNDIKYLPSFSKVLPEILNLYRNNPKNEPF DTIETEKIVLNALIYVNKELYKKLILEDDLEENESKNIFLQELK KTLGNIDEIDENIIENYYKNAQISASKGNNKAIKKYQKKVIEC YIGYLRKNYEELFDFSDFKMNIQEIKKQIKDINDNKTYERITV KTSDKTIVINDDFEYIISIFALLNSNAVINKIRNRFFATSVWLN TSEYQNIIDILDEIMQLNTLRNECITENWNLNLEEFIQKMKEIE KDFDDFKIQTKKEIFNNYYEDIKNNILTEFKDDINGCDVLEKK LEKIVIFDDETKFEIDKKSNILQDEQRKLSNINKKDLKKKVDQ

[0377]

			YIKDKDQEIKSKILCRIIFNSDFLKKYKKEIDNLIEDMESENEN KFQEIYYPKERKNELYIYKKNLFLNIGNPNFDKIYGLISNDIK MADAKFLFNIDGKNIRKNKISEIDAILKNLNDKLNGYSKEYK EKYIKKLKENDDFFAKNIQNKNYKSFEKDYNRVSEYKKIRD LVEFNYLNKIESYLIDINWKLAIQMARFERDMHYIVNGLREL GIIKLSGYNTGISRAYPKRNGSDGFYTTTAYYKFFDEESYKKF EKICYGFGIDLSENSEINKPENESIRNYISHFYIVRNPFADYSIA EQIDRVSNLLSYSTRYNNSTYASVFEVFKKDVNLDYDELKK KFKLIGNNDILERLMKPKKVSVLELESYNSDYIKNLIIELLTKI ENTNDTL
c2c2-3		엘. 웨이데이 (Lw2) (SEQ ID No. 47)	MKVTKVDGISHKKYIEEGKLVKSTSEENRTSERLSELLSIRLD IYIKNPDNASEEENRIRRENLKKFFSNKVLHLKDSVLYLKNR KEKNAVQDKNYSEEDISEYDLKNKNSFSVLKKILLNEDVNSE ELEIFRKDVEAKLNKINSLKYSFEENKANYQKINENNVEKVG GKSKRNIIYDYYRESAKRNDYINNVQEAFDKLYKKEDIEKLF FLIENSKKHEKYKIREYYHKIIGRKNDKENFAKIIYEEIQNVN NIKELIEKIPDMSELKKSQVFYKYYLDKEELNDKNIKYAFCH FVEIEMSQLLKNYVYKRLSNISNDKIKRIFEYQNLKKLIENKL LNKLDTYVRNCGKYNYYLQVGEIATSDFIARNRQNEAFLRNI IGVSSVAYFSLRNILETENENDITGRMRGKTVKNNKGEEKYV SGEVDKIYNENKQNEVKENLKMFYSYDFNMDNKNEIEDFFA NIDEAISSIRHGIVHFNLELEGKDIFAFKNIAPSEISKKMFQNEI NEKKLKLKIFKQLNSANVFNYYEKDVIIKYLKNTKFNFVNK NIPFVPSFTKLYNKIEDLRNTLKFFWSVPKDKEEKDAQIYLLK NIYYGEFLNKFVKNSKVFFKITNEVIKINKQRNQKTGHYKYQ KFENIEKTVPVEYLAIIQSREMINNQDKEEKNTYIDFIQQIFLK GFIDYLNKNNLKYIESNNNNDNNDIFSKIKIKKDNKEKYDKIL KNYEKHNRNKEIPHEINEFVREIKLGKILKYTENLNMFYLILK LLNHKELTNLKGSLEKYQSANKEETFSDELELINLLNLDNNR VTEDFELEANEIGKFLDFNENKIKDRKELKKFDTNKIYFDGE NIIKHRAFYNIKKYGMLNLLEKIADKAKYKISLKELKEYSNK KNEIEKNYTMQQNLHRKYARPKKDEKFNDEDYKEYEKAIG

[0378]

			NIQKYTHLKNKVEFNELNLLQGLLLKILHRLVGYTSIWERDL RFRLKGEFPENHYIEEIFNFDNSKNVKYKSGQIVEKYINFYKE LYKDNVEKRSIYSDKKVKKLKQEKKDLYIRNYIAHFNYIPHA EISLLEVLENLRKLLSYDRKLKNAIMKSIVDILKEYGFVATFK IGADKKIEIQTLESEKIVHLKNLKKKKLMTDRNSEELCELVK VMFEYKALEKRPAATKKAGQAKKKKGSYPYDVPDYAYPY DVPDYAYPYDVPDYA*
c2c2-4		리스테리아 실리게리 (SEQ ID No. 48)	MWISIKTLIHHLGVLFFCDYMYNRREKKIIEVKTMRITKVEV DRKKVLISRDKNGGKLVYENEMQDNTEQIMHHKKSSFYKS VVNKTICRPEQKQMKKLVHGLLQENSQEKIKVSDVTKLNIS NFLNHRFKKSLYYFPENSPDKSEEYRIEINLSQLLEDSLKKQQ GTFICWESFSKDMELYINWAENYISSKTKLIKKSIRNNRIQST ESRSGQLMDRYMKDILNKNKPFDIQSVSEKYQLEKLTSALK ATFKEAKKNDKEINYKLKSTLQNHERQIIEELKENSELNQFNI EIRKHLETYFPIKKTNRKVGDIRNLEIGEIQKIVNHRLKNKIVQ RILQEGKLASYEIESTVNSNSLQKIKIEEAFALKFINACLFASN NLRNMVYPVCKKDILMIGEFKNSFKEIKHKKFIRQWSQFFSQ EITVDDIELASWGLRGAIAPIRNEIIHLKKHSWKKFFNNPTFK VKKSKIINGKTKDVTSEFLYKETLFKDYFYSELDSVPELIINK MESSKILDYYSSDQLNQVFTIPNFELSLLTSAVPFAPSFKRVY LKGFDYQNQDEAQPDYNLKLNIYNEKAFNSEAFQAQYSLFK MVYYQVFLPQFTTNNDLFKSSVDFILTLNKERKGYAKAFQDI RKMNKDEKPSEYMSYIQSQLMLYQKKQEEKEKINHFEKFIN QVFIKGFNSFIEKNRLTYICHPTKNTVPENDNIEIPFHTDMDDS NIAFWLMCKLLDAKQLSELRNEMIKFSCSLQSTEEISTFTKAR EVIGLALLNGEKGCNDWKELFDDKEAWKKNMSLYVSEELL QSLPYTQEDGQTPVINRSIDLVKKYGTETILEKLFSSSDDYKV SAKDIAKLHEYDVTEKIAQQESLHKQWIEKPGLARDSAWTK KYQNVINDISNYQWAKTKVELTQVRHLHQLTIDLLSRLAGY MSIADRDFQFSSNYILERENSEYRVTSWILLSENKNKNKYND YELYNLKNASIKVSSKNDPQLKVDLKQLRLTLEYLELFDNRL KEKRNNISHFNYLNGQLGNSILELFDDARDVLSYDRKLKNA VSKSLKEILSSHGMEVTFKPLYQTNHHLKIDKLQPKKIHHLG EKSTVSSNQVSNEYCQLVRTLLTMK

[0379]

C2-17		렙토트리키아 부칼리스 C-1013-b (SEQ ID No. 49)	MKVTKVGGISHKKYTSEGRLVKSESEENRTDERLSALLNMR LDMYIKNPSSTETKENQKRIGKLKKFFSNKMVYLKDNTLSL KNGKKENIDREYSETDILESDVRDKKNFAVLKKIYLNENVNS EELEVFRNDIKKKLNKINSLKYSFEKNKANYQKINENNIEKV EGKSKRNIIYDYYRESAKRDAYVSNVKEAFDKLYKEEDIAK LVLEIENLTKLEKYKIREFYHEIIGRKNDKENFAKIIYEEIQNV NNMKELIEKVPDMSELKKSQVFYKYYLDKEELNDKNIKYAF CHFVEIEMSQLLKNYVYKRLSNISNDKIKRIFEYQNLKKLIEN KLLNKLDTYVRNCGKYNYYLQDGEIATSDFIARNRQNEAFL RNIIGVSSVAYFSLRNILETENENDITGRMRGKTVKNNKGEE KYVSGEVDKIYNENKKNEVKENLKMFYSYDFNMDNKNEIE DFFANIDEAISSIRHGIVHFNLELEGKDIFAFKNIAPSEISKKMF QNEINEKKLKLKIFRQLNSANVFRYLEKYKILNYLKRTRFEF VNKNIPFVPSFTKLYSRIDDLKNSLGIYWKTPKTNDDNKTKEI IDAQIYLLKNIYYGEFLNYFMSNNGNFFEISKEIIELNKNDKR NLKTGFYKLQKFEDIQEKIPKEYLANIQSLYMINAGNQDEEE KDTYIDFIQKIFLKGFMTYLANNGRLSLIYIGSDEETNTSLAE KKQEFDKFLKKYEQNNNIKIPYEINEFLREIKLGNILKYTERL NMFYLILKLLNHKELTNLKGSLEKYQSANKEEAFSDQLELIN LLNLDNNRVTEDFELEADEIGKFLDFNGNKVKDNKELKKFD TNKIYFDGENIIKHRAFYNIKKYGMLNLLEKIADKAGYKISIE ELKKYSNKKNEIEKNHKMQENLHRKYARPRKDEKFTDEDY ESYKQAIENIEEYTHLKNKVEFNELNLLQGLLLRILHRLVGY TSIWERDLRFRLKGEFPENQYIEEIFNFENKKNVKYKGGQIVE KYIKFYKELHQNDEVKINKYSSANIKVLKQEKKDLYIRNYIA HFNYIPHAEISLLEVLENLRKLLSYDRKLKNAVMKSVVDILK EYGFVATFKIGADKKIGIQTLESEKIVHLKNLKKKKLMTDRN SEELCKLVKIMFEYKMEEKKSEN

[0380]

C2-18		허비닉스 헤미셀룰로실리티카 (SEQ ID No. 50)	MKLTRRRISGNSVDQKITAAFYRDMSQGLLYYDSEDNDCTD KVIESMDFERSWRGRILKNGEDDKNPFYMFVKGLVGSNDKI VCEPIDVDSDPDNLDILINKNLTGFGRNLKAPDSNDTLENLIR KIQAGIPEEEVLPELKKIKEMIQKDIVNRKEQLLKSIKNNRIPF SLEGSKLVPSTKKMKWLFKLIDVPNKTFNEKMLEKYWEIYD YDKLKANITNRLDKTDKKARSISRAVSEELREYHKNLRTNY NRFVSGDRPAAGLDNGGSAKYNPDKEEFLLFLKEVEQYFKK YFPVKSKHSNKSKDKSLVDKYKNYCSYKVVKKEVNRSIINQ LVAGLIQQGKLLYYFYYNDTWQEDFLNSYGLSYIQVEEAFK KSVMTSLSWGINRLTSFFIDDSNTVKFDDITTKKAKEAIESNY FNKLRTCSRMQDHFKEKLAFFYPVYVKDKKDRPDDDIENLI VLVKNAIESVSYLRNRTFHFKESSLLELLKELDDKNSGQNKI DYSVAAEFIKRDIENLYDVFREQIRSLGIAEYYKADMISDCFK TCGLEFALYSPKNSLMPAFKNVYKRGANLNKAYIRDKGPKE TGDQGQNSYKALEEYRELTWYIEVKNNDQSYNAYKNLLQLI YYHAFLPEVRENEALITDFINRTKEWNRKETEERLNTKNNKK HKNFDENDDITVNTYRYESIPDYQGESLDDYLKVLQRKQMA RAKEVNEKEEGNNNYIQFIRDVVVWAFGAYLENKLKNYKN ELQPPLSKENIGLNDTLKELFPEEKVKSPFNIKCRFSISTFIDNK GKSTDNTSAEAVKTDGKEDEKDKKNIKRKDLLCFYLFLRLL DENEICKLQHQFIKYRCSLKERRFPGNRTKLEKETELLAELEE LMELVRFTMPSIPEISAKAESGYDTMIKKYFKDFIEKKVFKNP KTSNLYYHSDSKTPVTRKYMALLMRSAPLHLYKDIFKGYYL ITKKECLEYIKLSNIIKDYQNSLNELHEQLERIKLKSEKQNGK DSLYLDKKDFYKVKEYVENLEQVARYKHLQHKINFESLYRI FRIHVDIAARMVGYTQDWERDMHFLFKALVYNGVLEERRF EAIFNNNDDNNDGRIVKKIQNNLNNKNRELVSMLCWNKKL NKNEFGAIIWKRNPIAHLNHFTQTEQNSKSSLESLINSLRILLA YDRKRQNAVTKTINDLLLNDYHIRIKWEGRVDEGQIYFNIKE KEDIENEPIIHLKHLHKKDCYIYKNSYMFDKQKEWICNGIKE EVYDKSILKCIGNLFKFDYEDKNKSSANPKHT
C2-19		[유박테리	MLRRDKEVKKLYNVFNQIQVGTKPKKWNNDEKLSPEENER

[0381]

		움] 렉탈레 (SEQ ID No. 51)	RAQQKNIKMKNYKWREACSKYVESSQRIINDVIFYSYRKAK NKLRYMRKNEDILKKMQEAEKLSKFSGGKLEDFVAYTLRKS LVVSKYDTQEFDSLAAMVVFLECIGKNNISDHEREIVCKLLE LIRKDFSKLDPNVKGSQGANIVRSVRNQNMIVQPQGDRFLFP QVYAKENETVTNKNVEKEGLNEFLLNYANLDDEKRAESLR KLRRILDVYFSAPNHYEKDMDITLSDNIEKEKFNVWEKHEC GKKETGLFVDIPDVLMEAEAENIKLDAVVEKRERKVLNDRV RKQNIICYRYTRAVVEKYNSNEPLFFENNAINQYWIHHIENA VERILKNCKAGKLFKLRKGYLAEKVWKDAINLISIKYIALGK AVYNFALDDIWKDKKNKELGIVDERIRNGITSFDYEMIKAHE NLQRELAVDIAFSVNNLARAVCDMSNLGNKESDFLLWKRN DIADKLKNKDDMASVSAVLQFFGGKSSWDINIFKDAYKGKK KYNYEVRFIDDLRKAIYCARNENFHFKTALVNDEKWNTELF GKIFERETEFCLNVEKDRFYSNNLYMFYQVSELRNMLDHLY SRSVSRAAQVPSYNSVIVRTAFPEYITNVLGYQKPSYDADTL GKWYSACYYLLKEIYYNSFLQSDRALQLFEKSVKTLSWDDK KQQRAVDNFKDHFSDIKSACTSLAQVCQIYMTEYNQQNNQI KKVRSSNDSIFDQPVYQHYKVLLKKAIANAFADYLKNNKDL FGFIGKPFKANEIREIDKEQFLPDWTSRKYEALCIEVSGSQEL QKWYIVGKFLNARSLNLMVGSMRSYIQYVTDIKRRAASIGN ELHVSVHDVEKVEKWVQVIEVCSLLASRTSNQFEDYFNDKD DYARYLKSYVDFSNVDMPSEYSALVDFSNEEQSDLYVDPKN PKVNRNIVHSKLFAADHILRDIVEPVSKDNIEEFYSQKAEIAY CKIKGKEITAEEQKAVLKYQKLKNRVELRDIVEYGEIINELLG QLINWSFMRERDLLYFQLGFHYDCLRNDSKKPEGYKNIKVD ENSIKDAILYQIIGMYVNGVTVYAPEKDGDKLKEQCVKGGV GVKVSAFHRYSKYLGLNEKTLYNAGLEIFEVVAEHEDIINLR NGIDHFKYYLGDYRSMLSIYSEVFDRFFTYDIKYQKNVLNLL QNILLRHNVIVEPILESGFKTIGEQTKPGAKLSIRSIKSDTFQY KVKGGTLITDAKDERYLETIRKILYYAENEEDNLKKSVVVTN ADKYEKNKESDDQNKQKEKKNKDNKGKKNEETKSDAEKN NNERLSYNPFANLNFKLSN

[0382]

C2-20		유박테리아과 박테리아 CHKCI004 (SEQ ID No. 52)	MKISKESHKRTAVAVMEDRVGGVVYVPGGSGIDLSNNLKK RSMDTKSLYNVFNQIQAGTAPSEYEWKDYLSEAENKKREAQ KMIQKANYELRRECEDYAKKANLAVSRIIFSKKPKKIFSDDDI ISHMKKQRLSKFKGRMEDFVLIALRKSLVVSTYNQEVFDSR KAATVFLKNIGKKNISADDERQIKQLMALIREDYDKWNPDK DSSDKKESSGTKVIRSIEHQNMVIQPEKNKLSLSKISNVGKKT KTKQKEKAGLDAFLKEYAQIDENSRMEYLKKLRRLLDTYFA APSSYIKGAAVSLPENINFSSELNVWERHEAAKKVNINFVEIP ESLLNAEQNNNKINKVEQEHSLEQLRTDIRRRNITCYHFANA LAADERYHTLFFENMAMNQFWIHHMENAVERILKKCNVGT LFKLRIGYLSEKVWKDMLNLLSIKYIALGKAVYHFALDDIW KADIWKDASDKNSGKINDLTLKGISSFDYEMVKAQEDLQRE MAVGVAFSTNNLARVTCKMDDLSDAESDFLLWNKEAIRRH VKYTEKGEILSAILQFFGGRSLWDESLFEKAYSDSNYELKFL DDLKRAIYAARNETFHFKTAAIDGGSWNTRLFGSLFEKEAGL CLNVEKNKFYSNNLVLFYKQEDLRVFLDKLYGKECSRAAQI PSYNTILPRKSFSDFMKQLLGLKEPVYGSAILDQWYSACYYL FKEVYYNLFLQDSSAKALFEKAVKALKGADKKQEKAVESFR KRYWEISKNASLAEICQSYITEYNQQNNKERKVRSANDGMF NEPIYQHYKMLLKEALKMAFASYIKNDKELKFVYKPTEKLF EVSQDNFLPNWNSEKYNTLISEVKNSPDLQKWYIVGKFMNA RMLNLLLGSMRSYLQYVSDIQKRAAGLGENQLHLSAENVG QVKKWIQVLEVCLLLSVRISDKFTDYFKDEEEYASYLKEYV DFEDSAMPSDYSALLAFSNEGKIDLYVDASNPKVNRNIIQAK LYAPDMVLKKVVKKISQDECKEFNEKKEQIMQFKNKGDEVS WEEQQKILEYQKLKNRVELRDLSEYGELINELLGQLINWSYL RERDLLYFQLGFHYSCLMNESKKPDAYKTIRRGTVSIENAVL YQIIAMYINGFPVYAPEKGELKPQCKTGSAGQKIRAFCQWAS MVEKKKYELYNAGLELFEVVKEHDNIIDLRNKIDHFKYYQG NDSILALYGEIFDRFFTYDMKYRNNVLNHLQNILLRHNVIIKP IISKDKKEVGRGKMKDRAAFLLEEVSSDRFTYKVKEGERKID AKNRLYLETVRDILYFPNRAVNDKGEDVIICSKKAQDLNEK

[0383]

			KADRDKNHDKSKDTNQKKEGKNQEEKSENKEPYSDRMTW KPFAGIKLE
C2-21		플라우티아 sp. 마르세이유-P2398 (SEQ ID No. 53)	MKISKVDHVKSGIDQKLSSQRGMLYKQPQKKYEGKQLEEH VRNLSRKAKALYQVFPVSGNSKMEKELQIINSFIKNILLRLDS GKTSEEIVGYINTYSVASQISGDHIQELVDQHLKESLRKYTCV GDKRIYVPDIIVALLKSKFNSETLQYDNSELKILIDFIREDYLK EKQIKQIVHSIENNSTPLRIAEINGQKRLIPANVDNPKKSYIFE FLKEYAQSDPKGQESLLQHMRYLILLYLYGPDKITDDYCEEI EAWNFGSIVMDNEQLFSEEASMLIQDRIYVNQQIEEGRQSKD TAKVKKNKSKYRMLGDKIEHSINESVVKHYQEACKAVEEK DIPWIKYISDHVMSVYSSKNRVDLDKLSLPYLAKNTWNTWI SFIAMKYVDMGKGVYHFAMSDVDKVGKQDNLIIGQIDPKFS DGISSFDYERIKAEDDLHRSMSGYIAFAVNNFARAICSDEFRK KNRKEDVLTVGLDEIPLYDNVKRKLLQYFGGASNWDDSIIDI IDDKDLVACIKENLYVARNVNFHFAGSEKVQKKQDDILEEIV RKETRDIGKHYRKVFYSNNVAVFYCDEDIIKLMNHLYQREK PYQAQIPSYNKVISKTYLPDLIFMLLKGKNRTKISDPSIMNMF RGTFYFLLKEIYYNDFLQASNLKEMFCEGLKNNVKNKKSEK PYQNFMRRFEELENMGMDFGEICQQIMTDYEQQNKQKKKT ATAVMSEKDKKIRTLDNDTQKYKHFRTLLYIGLREAFIIYLK DEKNKEWYEFLREPVKREQPEEKEFVNKWKLNQYSDCSELI LKDSLAAAAWYVVAHFINQAQLNHLIGDIKNYIQFISDIDRRA KSTGNPVSESTEIQIERYRKILRVLEFAKFFCGQITNVLTDYY QDENDFSTHVGHYVKFEKKNMEPAHALQAFSNSLYACGKE KKKAGFYYDGMNPIVNRNITLASMYGNKKLLENAMNPVTE QDIRKYYSLMAELDSVLKNGAVCKSEDEQKNLRHFQNLKN RIELVDVLTLSELVNDLVAQLIGWVYIRERDMMYLQLGLHY IKLYFTDSVAEDSYLRTLDLEEGSIADGAVLYQIASLYSFNLP MYVKPNKSSVYCKKHVNSVATKFDIFEKEYCNGDETVIENG LRLFENINLHKDMVKFRDYLAHFKYFAKLDESILELYSKAYD FFFSYNIKLKKSVSYVLTNVLLSYFINAKLSFSTYKSSGNKTV QHRTTKISVVAQTDYFTYKLRSIVKNKNGVESIENDDRRCEV

[0384]

			VNIAARDKEFVDEVCNVINYNSDK
C2-22		렙토트리키아 sp. 경구분류군 879 str. F0557 (SEQ ID No. 54)	MGNLFGHKRWYEVRDKKDFKIKRKVKVKRNYDGNKYILNI NENNNKEKIDNNKFIGEFVNYKKNNNVLKEFKRKFHAGNIL FKLKGKEEIIRIENNDDFLETEEVVLYIEVYGKSEKLKALEITK KKIIDEAIRQGITKDDKKIEIKRQENEEEIEIDIRDEYTNKTLND CSIILRIIENDELETKKSIYEIFKNINMSLYKIIEKIIENETEKVFE NRYYEEHLREKLLKDNKIDVILTNFMEIREKIKSNLEIMGFVK FYLNVSGDKKKSENKKMFVEKILNTNVDLTVEDIVDFIVKEL KFWNITKRIEKVKKFNNEFLENRRNRTYIKSYVLLDKHEKFK IERENKKDKIVKFFVENIKNNSIKEKIEKILAEFKINELIKKLEK ELKKGNCDTEIFGIFKKHYKVNFDSKKFSNKSDEEKELYKIIY RYLKGRIEKILVNEQKVRLKKMEKIEIEKILNESILSEKILKRV KQYTLEHIMYLGKLRHNDIVKMTVNTDDFSRLHAKEELDLE LITFFASTNMELNKIFNGKEKVTDFFGFNLNGQKITLKEKVPS FKLNILKKLNFINNENNIDEKLSHFYSFQKEGYLLRNKILHNS YGNIQETKNLKGEYENVEKLIKELKVSDEEISKSLSLDVIFEG KVDIINKINSLKIGEYKDKKYLPSFSKIVLEITRKFREINKDKL FDIESEKIILNAVKYVNKILYEKITSNEENEFLKTLPDKLVKKS NNKKENKNLLSIEEYYKNAQVSSSKGDKKAIKKYQNKVTNA YLEYLENTFTEIIDFSKFNLNYDEIKTKIEERKDNKSKIIIDSIST NINITNDIEYIISIFALLNSNTYINKIRNRFFATSVWLEKQNGTK EYDYENIISILDEVLLINLLRENNITDILDLKNAIIDAKIVENDE TYIKNYIFESNEEKLKKRLFCEELVDKEDIRKIFEDENFKFKSF IKKNEIGNFKINFGILSNLECNSEVEAKKIIGKNSKKLESFIQNI IDEYKSNIRTLFSSEFLEKYKEEIDNLVEDTESENKNKFEKIYY PKEHKNELYIYKKNLFLNIGNPNFDKIYGLISKDIKNVDTKIL FDDDIKKNKISEIDAILKNLNDKLNGYSNDYKAKYVNKLKE NDDFFAKNIQNENYSSFGEFEKDYNKVSEYKKIRDLVEFNYL NKIESYLIDINWKLAIQMARFERDMHYIVNGLRELGIIKLSGY NTGISRAYPKRNGSDGFYTTTAYYKFFDEESYKKFEKICYGF GIDLSENSEINKPENESIRNYISHFYIVRNPFADYSIAEQIDRVS NLLSYSTRYNNSTYASVFEVFKKDVNLDYDELKKKFRLIGN

[0385]

			NDILERLMKPKKVSVLELESYNSDYIKNLIIELLTKIENTNDTL
C2-23		라크노스피라과박테리아 NK4A144 (SEQ ID No. 55)	MKISKVDHTRMAVAKGNQHRRDEISGILYKDPTKTGSIDFDE RFKKLNCSAKILYHVFNGIAEGSNKYKNIVDKVNNNLDRVL FTGKSYDRKSIIDIDTVLRNVEKINAFDRISTEEREQIIDDLLEI QLRKGLRKGKAGLREVLLIGAGVIVRTDKKQEIADFLEILDE DFNKTNQAKNIKLSIENQGLVVSPVSRGEERIFDVSGAQKGK SSKKAQEKEALSAFLLDYADLDKNVRFEYLRKIRRLINLYFY VKNDDVMSLTEIPAEVNLEKDFDIWRDHEQRKEENGDFVGC PDILLADRDVKKSNSKQVKIAERQLRESIREKNIKRYRFSIKTI EKDDGTYFFANKQISVFWIHRIENAVERILGSINDKKLYRLRL GYLGEKVWKDILNFLSIKYIAVGKAVFNFAMDDLQEKDRDI EPGKISENAVNGLTSFDYEQIKADEMLQREVAVNVAFAANN LARVTVDIPQNGEKEDILLWNKSDIKKYKKNSKKGILKSILQ FFGGASTWNMKMFEIAYHDQPGDYEENYLYDIIQIIYSLRNK SFHFKTYDHGDKNWNRELIGKMIEHDAERVISVEREKFHSN NLPMFYKDADLKKILDLLYSDYAGRASQVPAFNTVLVRKNF PEFLRKDMGYKVHFNNPEVENQWHSAVYYLYKEIYYNLFL RDKEVKNLFYTSLKNIRSEVSDKKQKLASDDFASRCEEIEDR SLPEICQIIMTEYNAQNFGNRKVKSQRVIEKNKDIFRHYKML LIKTLAGAFSLYLKQERFAFIGKATPIPYETTDVKNFLPEWKS GMYASFVEEIKNNLDLQEWYIVGRFLNGRMLNQLAGSLRSY IQYAEDIERRAAENRNKLFSKPDEKIEACKKAVRVLDLCIKIS TRISAEFTDYFDSEDDYADYLEKYLKYQDDAIKELSGSSYAA LDHFCNKDDLKFDIYVNAGQKPILQRNIVMAKLFGPDNILSE VMEKVTESAIREYYDYLKKVSGYRVRGKCSTEKEQEDLLKF QRLKNAVEFRDVTEYAEVINELLGQLISWSYLRERDLLYFQL GFHYMCLKNKSFKPAEYVDIRRNNGTIIHNAILYQIVSMYIN GLDFYSCDKEGKTLKPIETGKGVGSKIGQFIKYSQYLYNDPS YKLEIYNAGLEVFENIDEHDNITDLRKYVDHFKYYAYGNKM SLLDLYSEFFDRFFTYDMKYQKNVVNVLENILLRHFVIFYPK FGSGKKDVGIRDCKKERAQIEISEQSLTSEDFMFKLDDKAGE EAKKFPARDERYLQTIAKLLYYPNEIEDMNRFMKKGETINK

[0386]

			KVQFNRKKKITRKQKNNSSNEVLSSTMGYLFKNIKL
C2-24		클로로플렉서스 아그레간스 (SEQ ID No. 56)	MTDQVRREEVAAGELADTPLAAAQTPAADAAVAATPAPAE AVAPTPEQAVDQPATTGESEAPVTTAQAAAHEAEPAEATGA SFTPVSEQQPQKPRRLKDLQPGMELEGKVTSIALYGIFVDVG VGRDGLVHISEMSDRRIDTPSELVQIGDTVKVWVKSVDLDA RRISLTMLNPSRGEKPRRSRQSQPAQPQPRRQEVDREKLASL KVGEIVEGVITGFAPFGAFADIGVGKDGLIHISELSEGRVEKP EDAVKVGERYQFKVLEIDGEGTRISLSLRRAQRTQRMQQLEP GQIIEGTVSGIATFGAFVDIGVGRDGLVHISALAPHRVAKVED VVKVGDKVKVKVLGVDPQSKRISLTMRLEEEQPATTAGDEA AEPAEEVTPTRRGNLERFAAAAQTARERSERGERSERGERRE RRERRPAQSSPDTYIVGEDDDESFEGNATIEDLLTKFGGSSSR RDRDRRRRHEDDDDEEMERPSNRRQREAIRRTLQQIGYDE
C2-25		데메퀴나 아우란티아카 (SEQ ID No. 57)	MDLTWHALLILFIVALLAGFLDTLAGGGGLLTVPALLLTGIP PLQALGTNKLQSSFGTGMATYQVIRKKRVHWRDVRWPMV WAFLGSAAGAVAVQFIDTDALLIIIPVVLALVAAYFLFVPKS HLPPPEPRMSDPAYEATLVPIIGAYDGAFGPGTGSLYALSGV ALRAKTLVQSTAIAKTLNFATNFAALLVFAFAGHMLWTVGA VMIAGQLIGAYAGSHMLFRVNPLVLRVLIVVMSLGMLIRVL LD
C2-26		탈라소스피라 sp. TSL5-1 (SEQ ID No. 58)	MRIIKPYGRSHVEGVATQEPRRKLRLNSSPDISRDIPGFAQSH DALIIAQWISAIDKIATKPKPDKKPTQAQINLRTTLGDAAWQ HVMAENLLPAATDPAIREKLHLIWQSKIAPWGTARPQAEKD GKPTPKGGWYERFCGVLSPEAITQNVARQIAKDIYDHLHVA AKRKGREPAKQGESSNKPGKFKPDRKRGLIEERAESIAKNAL RPGSHAPCPWGPDDQATYEQAGDVAGQIYAAARDCLEEKK RRSGNRNTSSVQYLPRDLAAKILYAQYGRVFGPDTTIKAALD EQPSLFALHKAIKDCYHRLINDARKRDILRILPRNMAALFRL VRAQYDNRDINALIRLGKVIHYHASEQGKSEHHGIRDYWPS QQDIQNSRFWGSDGQADIKRHEAFSRIWRHIIALASRTLHDW ADPHSQKFSGENDDILLLAKDAIEDDVFKAGHYERKCDVLF

[0387]

			GAQASLFCGAEDFEKAILKQAITGTGNLRNATFHFKGKVRFE KELQELTKDVPVEVQSAIAALWQKDAEGRTRQIAETLQAVL AGHFLTEEQNRHIFAALTAAMAQPGDVPLPRLRRVLARHDSI CQRGRILPLSPCPDRAKLEESPALTCQYTVLKMLYDGPFRAW LAQQNSTILNHYIDSTIARTDKAARDMNGRKLAQAEKDLITS RAADLPRLSVDEKMGDFLARLTAATATEMRVQRGYQSDGE NAQKQAAFIGQFECDVIGRAFADFLNQSGFDFVLKLKADTP QPDAAQCDVTALIAPDDISVSPPQAWQQVLYFILHLVPVDDA SHLLHQIRKWQVLEGKEKPAQIAHDVQSVLMLYLDMHDAK FTGGAALHGIEKFAEFFAHAADFRAVFPPQSLQDQDRSIPRR GLREIVRFGHLPLLQHMSGTVQITHDNVVAWQAARTAGAT GMSPIARRQKQREELHALAVERTARFRNADLQNYMHALVD VIKHRQLSAQVTLSDQVRLHRLMMGVLGRLVDYAGLWERD LYFVVLALLYHHGATPDDVFKGQGKKNLADGQVVAALKPK NRKAAAPVGVFDDLDHYGIYQDDRQSIRNGLSHFNMLRGG KAPDLSHWVNQTRSLVAHDRKLKNAVAKSVIEMLAREGFD LDWGIQTDRGQHILSHGKIRTRQAQHFQKSRLHIVKKSAKPD KNDTVKIRENLHGDAMVERVVQLFAAQVQKRYDITVEKRL DHLFLKPQDQKGKNGIHTHNGWSKTEKKRRPSRENRKGNH EN
C2-27		SAMN044 87830_139 20 [슈도부티 리비브리오 sp. OR37] (SEQ ID No. 59)	MKFSKESHRKTAVGVTESNGIIGLLYKDPLNEKEKIEDVVNQ RANSTKRLFNLFGTEATSKDISRASKDLAKVVNKAIGNLKGN KKFNKKEQITKGLNTKIIVEELKNVLKDEKKLIVNKDIIDEAC SRLLKTSFRTAKTKQAVKMILTAVLIENTNLSKEDEAFVHEY FVKKLVNEYNKTSVKKQIPVALSNQNMVIQPNSVNGTLEISE TKKSKETKTTEKDAFRAFLRDYATLDENRRHKMRLCLRNLV NLYFYGETSVSKDDFDEWRDHEDKKQNDELFVKKIVSIKTD RKGNVKEVLDVDATIDAIRTNNIACYRRALAYANENPDVFF SDTMLNKFWIHHVENEVERIYGHINNNTGDYKYQLGYLSEK VWKGIINYLSIKYIAEGKAVYNYAMNALAKDNNSNAFGKLD EKFVNGITSFEYERIKAEETLQRECAVNIAFAANHLANATVD LNEKDSDFLLLKHEDNKDTLGAVARPNILRNILQFFGGKSRW

[0388]

			NDFDFSGIDEIQLLDDLRKMIYSLRNSSFHFKTENIDNDSWNT KLIGDMFAYDFNMAGNVQKDKMYSNNVPMFYSTSDIEKML DRLYAEVHERASQVPSFNSVFVRKNFPDYLKNDLKITSAFGV DDALKWQSAVYYVCKEIYYNDFLQNPETFTMLKDYVQCLPI DIDKSMDQKLKSERNAHKNFKEAFATYCKECDSLSAICQMI MTEYNNQNKGNRKVISARTKDGDKLIYKHYKMILFEALKN VFTIYLEKNINTYGFLKKPKLINNVPAIEEFLPNYNGRQYETL VNRITEETELQKWYIVGRLLNPKQVNQLIGNFRSYVQYVND VARRAKQTGNNLSNDNIAWDVKNIIQIFDVCTKLNGVTSNIL EDYFDDGDDYARYLKNFVDYTNKNNDHSATLLGDFCAKEI DGIKIGIYHDGTNPIVNRNIIQCKLYGATGIISDLTKDGSILSV DYEIIKKYMQMQKEIKVYQQKGICKTKEEQQNLKKYQELKN IVELRNIIDYSEILDELQGQLINWGYLRERDLMYFQLGFHYLC LHNESKKPVGYNNAGDISGAVLYQIVAMYTNGLSLIDANGK SKKNAKASAGAKVGSFCSYSKEIRGVDKDTKEDDDPIYLAG VELFENINEHQQCINLRNYIEHFHYYAKHDRSMLDLYSEVFD RFFTYDMKYTKNVPNMMYNILLQHLVVPAFEFGSSEKRLDD NDEQTKPRAMFTLREKNGLSSEQFTYRLGDGNSTVKLSARG DDYLRAVASLLYYPDRAPEGLIRDAEAEDKFAKINHSNPKSD NRNNRGNFKNPKVQWYNNKTKRK
C2-28		SAMN02910398_00008 [부티리비브리오 sp. YAB3001] (SEQ ID No. 60)	MKISKVDHRKTAVKITDNKGAEGFIYQDPTRDSSTMEQIISN RARSSKVLFNIFGDTKKSKDLNKYTESLIIYVNKAIKSLKGDK RNNKYEEITESLKTERVLNALIQAGNEFTCSENNIEDALNKY LKKSFRVGNTKSALKKLLMAAYCGYKLSIEEKEEIQNYFVD KLVKEYNKDTVLKYTAKSLKHQNMVVQPDTDNHVFLPSRI AGATQNKMSEKEALTEFLKAYAVLDEEKRHNLRIILRKLVN LYFYESPDFIYPENNEWKEHDDRKNKTETFVSPVKVNEEKN GKTFVKIDVPATKDLIRLKNIECYRRSVAETAGNPITYFTDHN ISKFWIHHIENEVEKIFALLKSNWKDYQFSVGYISEKVWKEII NYLSIKYIAIGKAVYNYALEDIKKNDGTLNFGVIDPSFYDGIN SFEYEKIKAEETFQREVAVYVSFAVNHLSSATVKLSEAQSDM LVLNKNDIEKIAYGNTKRNILQFFGGQSKWKEFDFDRYINPV

[0389]

			NYTDIDFLFDIKKMVYSLRNESFHFTTTDTESDWNKNLISAM FEYECRRISTVQKNKFFSNNLPLFYGENSLERVLHKLYDDYV DRMSQVPSFGNVFVRKKFPDYMKEIGIKHNLSSEDNLKLQG ALYFLYKEIYYNAFISSEKAMKIFVDLVNKLDTNARDDKGRI THEAMAHKNFKDAISHYMTHDCSLADICQKIMTEYNQQNT GHRKKQTTYSSEKNPEIFRHYKMILFMLLQKAMTEYISSEEIF DFIMKPNSPKTDIKEEEFLPQYKSCAYDNLIKLIADNVELQK WYITARLLSPREVNQLIGSFRSYKQFVSDIERRAKETNNSLSK SGMTVDVENITKVLDLCTKLNGRFSNELTDYFDSKDDYAVY VSKFLDFGFKIDEKFPAALLGEFCNKEENGKKIGIYHNGTEPI LNSNIIKSKLYGITDVVSRAVKPVSEKLIREYLQQEVKIKPYL ENGVCKNKEEQAALRKYQELKNRIEFRDIVEYSEIINELMGQ LINFSYLRERDLMYFQLGFHYLCLNNYGAKPEGYYSIVNDK RTIKGAILYQIVAMYTYGLPIYHYVDGTISDRRKNKKTVLDT LNSSETVGAKIKYFIYYSDELFNDSLILYNAGLELFENINEHE NIVNLRKYIDHFKYYVSQDRSLLDIYSEVFDRYFTYDRKYKK NVMNLFSNIMLKHFIITDFEFSTGEKTIGEKNTAKKECAKVRI KRGGLSSDKFTYKFKDAKPIELSAKNTEFLDGVARILYYPEN VVLTDLVRNSEVEDEKRIEKYDRNHNSSPTRKDKTYKQDVK KNYNKKTSKAFDSSKLDTKSVGNNLSDNPVLKQFLSESKKK R
C2-29		플라우티아 sp. 마르세이유-P2398 (SEQ ID No. 61)	MKISKVDHVKSGIDQKLSSQRGMLYKQPQKKYEGKQLEEH VRNLSRKAKALYQVFPVSGNSKMEKELQIINSFIKNILLRLDS GKTSEEIVGYINTYSVASQISGDHIQELVDQHLKESLRKYTCV GDKRIYVPDIIVALLKSKFNSETLQYDNSELKILIDFIREDYLK EKQIKQIVHSIENNSTPLRIAEINGQKRLIPANVDNPKKSYIFE FLKEYAQSDPKGQESLLQHMRYLILLYLYGPDKITDDYCEEI EAWNFGSIVMDNEQLFSEEASMLIQDRIYVNQQIEEGRQSKD TAKVKKNKSKYRMLGDKIEHSINESVVKHYQEACKAVEEK DIPWIKYISDHVMSVYSSKNRVDLDKLSLPYLAKNTWNTWI SFIAMKYVDMGKGVYHFAMSDVDKVGKQDNLIIGQIDPKFS DGISSFDYERIKAEDDLHRSMSGYIAFAVNNFARAICSDEFRK

[0390]

			KNRKEDVLTVGLDEIPLYDNVKRKLLQYFGGASNWDDSIIDI IDDKDLVACIKENLYVARNVNFHFAGSEKVQKKQDDILEEIV RKETRDIGKHYRKVFYSNNVAVFYCDEDIIKLMNHLYQREK PYQAQIPSYNKVISKTYLPDLIFMLLKGKNRTKISDPSIMNMF RGTFYFLLKEIYYNDFLQASNLKEMFCEGLKNNVKNKKSEK PYQNFMRRFEELENMGMDFGEICQQIMTDYEQQNKQKKKT ATAVMSEKDKKIRTLDNDTQKYKHFRTLLYIGLREAFIIYLK DEKNKEWYEFLREPVKREQPEEKEFVNKWKLNQYSDCSELI LKDSLAAAWYVVAHFINQAQLNHLIGDIKNYIQFISDIDRRA KSTGNPVSESTEIQIERYRKILRVLEFAKFFCGQITNVLTDYY QDENDFSTHVGHYVKFEKKNMEPAHALQAFSNSLYACGKE KKKAGFYYDGMNPIVNRNITLASMYGNKKLLENAMNPVTE QDIRKYYSLMAELDSVLKNGAVCKSEDEQKNLRHFQNLKN RIELVDVLTLSELVNDLVAQLIGWVYIRERDMMYLQLGLHY IKLYFTDSVAEDSYLRTLDLEEGSIADGAVLYQIASLYSFNLP MYVKPNKSSVYCKKHVNSVATKFDIFEKEYCNGDETVIENG LRLFENINLHKDMVKFRDYLAHFKYFAKLDESILELYSKAYD FFFSYNIKLKKSVSYVLTNVLLSYFINAKLSFSTYKSSGNKTV QHRTTKISVVAQTDYFTYKLRSIVKNKNGVESIENDDRRCEV VNIAARDKEFVDEVCNVINYNSDK
C2-30		렙토트리키아 sp. 마르세이유-P3007 (SEQ ID No. 62)	MKITKIDGISHKKYIKEGKLVKSTSEENKTDERLSELLTIRLD TYIKNPDNASEEENRIRRENLKEFFSNKVLYLKDGILYLKDR REKNQLQNKNYSEEDISEYDLKNKNNFLVLKKILLNEDINSE ELEIFRNDFEKKLDKINSLKYSLEENKANYQKINENNIKKVE GKSKRNIFYNYYKDSAKRNDYINNIQEAFDKLYKKEDIENLF FLIENSKKHEKYKIRECYHKIIGRKNDKENFATIIYEEIQNVNN MKELIEKVPNVSELKKSQVFYKYYLNKEKLNDENIKYVFCH FVEIEMSKLLKNYVYKKPSNISNDKVKRIFEYQSLKKLIENKL LNKLDTYVRNCGKYSFYLQDGEIATSDFIVGNRQNEAFLRNI IGVSSTAYFSLRNILETENENDITGRMRGKTVKNNKGEEKYIS GEIDKLYDNNKQNEVKKNLKMFYSYDFNMNSKKEIEDFFSN IDEAISSIRHGIVHFNLELEGKDIFTFKNIVPSQISKKMFHDEIN

[0391]

			EKKLKLKIFKQLNSANVFRYLEKYKILNYLNRTRFEFVNKNI PFVPSFTKLYSRIDDLKNSLGIYWKTPKTNDDNKTKEITDAQI YLLKNIYYGEFLNYFMSNNGNFFEITKEIIELNKNDKRNLKT GFYKLQKFENLQEKTPKEYLANIQSLYMINAGNQDEEEKDT YIDFIQKIFLKGFMTYLANNGRLSLIYIGSDEETNTSLAEKKQ EFDKFLKKYEQNNNIEIPYEINEFVREIKLGKILKYTERLNMF YLILKLLNHKELTNLKGSLEKYQSANKEEAFSDQLELINLLN LDNNRVTEDFELEADEIGKFLDFNGNKVKDNKELKKFDTNK IYFDGENIIKHRAFYNIKKYGMLNLLEKISDEAKYKISIEELKN YSKKKNEIEENHTTQENLHRKYARPRKDEKFTDEDYKKYEK AIRNIQQYTHLKNKVEFNELNLLQSLLLRILHRLVGYTSIWER DLRFRLKGEFPENQYIEEIFNFDNSKNVKYKNGQIVEKYINFY KELYKDDTEKISIYSDKKVKELKKEKKDLYIRNYIAHFNYIPN AEISLLEMLENLRKLLSYDRKLKNAIMKSIVDILKEYGFVVTF KIEKDKKIRIESLKSEEVVHLKKLKLKDNDKKKEPIKTYRNS KELCKLVKVMFEYKMKEKKSEN
C2-31		박테로이데스 이후아에(SEQ ID No. 63)	MRITKVKVKESSDQKDKMVLIHRKVGEGTLVLDENLADLTA PIIDKYKDKSFELSLLKQTLVSEKEMNIPKCDKCTAKERCLSC KQREKRLKEVRGAIEKTIGAVIAGRDIIPRLNIFNEDEICWLIK PKLRNEFTFKDVNKQVVKLNLPKVLVEYSKKNDPTLFLAYQ QWIAAYLKNKKGHIKKSILNNRVVIDYSDESKLSKRKQALEL WGEEYETNQRIALESYHTSYNIGELVTLLPNPEEYVSDKGEIR PAFHYKLKNVLQMHQSTVFGTNEILCINPIFNENRANIQLSAY NLEVVKYFEHYFPIKKKKKNLSLNQAIYYLKVETLKERLSLQ LENALRMNLLQKGKIKKHEFDKNTCSNTLSQIKRDEFFVLNL VEMCAFAANNIRNIVDKEQVNEILSKKDLCNSLSKNTIDKEL CTKFYGADFSQIPVAIWAMRGSVQQIRNEIVHYKAEAIDKIF ALKTFEYDDMEKDYSDTPFKQYLELSIEKIDSFFIEQLSSNDV LNYYCTEDVNKLLNKCKLSLRRTSIPFAPGFKTIYELGCHLQ DSSNTYRIGHYLMLIGGRVANSTVTKASKAYPAYRFMLKLI YNHLFLNKFLDNHNKRFFMKAVAFVLKDNRENARNKFQYA FKEIRMMNNDESIASYMSYIHSLSVQEQEKKGDKNDKVRYN

[0392]

			TEKFIEKVFVKGFDDFLSWLGVEFILSPNQEERDKTVTREEYE NLMIKDRVEHSINSNQESHIAFFTFCKLLDANHLSDLRNEWI KFRSSGDKEGFSYNFAIDIIELCLLTVDRVEQRRDGYKEQTEL KEYLSFFIKGNESENTVWKGFYFQQDNYTPVLYSPIELIRKY GTLELLKLIIVDEDKITQGEFEEWQTLKKVVEDKVTRRNELH QEWEDMKNKSSFSQEKCSIYQKLCRDIDRYNWLDNKLHLV HLRKLHNLVIQILSRMARFIALWDRDFVLLDASRANDDYKL LSFFNFRDFINAKKTKTDDELLAEFGSKIEKKNAPFIKAEDVP LMVECIEAKRSFYQKVFFRNNLQVLADRNFIAHYNYISKTAK CSLFEMIIKLRTLMYYDRKLRNAVVKSIANVFDQNGMVLQL SLDDSHELKVDKVISKRIVHLKNNNIMTDQVPEEYYKICRRL LEMKK
C2-32		SAMN05216357_1045 [포르피로모나다과 박테리아 KH3CP3RA] (SEQ ID No. 64)	MEFRDSIFKSLLQKEIEKAPLCFAEKLISGGVFSYYPSERLKEF VGNHPFSLFRKTMPFSPGFKRVMKSGGNYQNANRDGRFYD LDIGVYLPKDGFGDEEWNARYFLMKLIYNQLFLPYFADAEN HLFRECVDFVKRVNRDYNCKNNNSEEQAFIDIRSMREDESIA DYLAFIQSNIIIEENKKKETNKEGQINFNKFLLQVFVKGFDSFL KDRTELNFLQLPELQGDGTRGDDLESLDKLGAVVAVDLKLD ATGIDADLNENISFYTFCKLLDSNHLSRLRNEIIKYQSANSDF SHNEDFDYDRIISIIELCMLSADHVSTNDNESIFPNNDKDFSGI RPYLSTDAKVETFEDLYVHSDAKTPITNATMVLNWKYGTDK LFERLMISDQDFLVTEKDYFVWKELKKDIEEKIKLREELHSL WVNTPKGKKGAKKKNGRETTGEFSEENKKEYLEVCREIDRY VNLDNKLHFVHLKRMHSLLIELLGRFVGFTYLFERDYQYYH LEIRSRRNKDAGVVDKLEYNKIKDQNKYDKDDFFACTFLYE KANKVRNFIAHFNYLTMWNSPQEEEHNSNLSGAKNSSGRQN LKCSLTELINELREVMSYDRKLKNAVTKAVIDLFDKHGMVI KFRIVNNNNNDNKNKHHLELDDIVPKKIMHLRGIKLKRQDG KPIPIQTDSVDPLYCRMWKKLLDLKPTPF
C2-33		리스테리아 리파리아	MHDAWAENPKKPQSDAFLKEYKACCEAIDTYNWHKNKAT LVYVNELHHLLIDILGRLVGYVAIADRDFQCMANQYLKSSG

[0393]

		(SEQ ID No. 65)	HTERVDSWINTIRKNRPDYIEKLDIFMNKAGLFVSEKNGRNY IAHLNYLSPKHKYSLLYLFEKLREMLKYDRKLKNAVTKSLID LLDKHGMCVVFANLKNNKHRLVIASLKPKKIETFKWKKIK
C2-34		인솔리티 스피릴럼 페레그리넘 (SEQ ID No. 66)	MRIIRPYGSSTVASPSPQDAQPLRSLQRQNGTFDVAEFSRRHP ELVLAQWVAMLDKIIRKPAPGKNSTALPRPTAEQRRLRQQV GAALWAEMQRHTPVPPELKAVWDSKVHPYSKDNAPATAKT PSHRGRWYDRFGDPETSAATVAEGVRRHLLDSAQPFRANGG QPKGKGVIEHRALTIQNGTLLHHHQSEKAGPLPEDWSTYRA DELVSTIGKDARWIKVAASLYQHYGRIFGPTTPISEAQTRPEF VLHTAVKAYYRRLFKERKLPAERLERLLPRTGEALRHAVTV QHGNRSLADAVRIGKILHYGWLQNGEPDPWPDDAALYSSR YWGSDGQTDIKHSEAVSRVWRRALTAAQRTLTSWLYPAGT DAGDILLIGQKPDSIDRNRLPLLYGDSTRHWTRSPGDVWLFL KQTLENLRNSSFHFKTLSAFTSHLDGTCESEPAEQQAAQALW QDDRQQDHQQVFLSLRALDATTYLPTGPLHRIVNAVQSTDA TLPLPRFRRVVTRAANTRLKGFPVEPVNRRTMEDDPLLRCR YGVLKLLYERGFRAWLETRPSIASCLDQSLKRSTKAAQTING KNSPQGVEILSRATKLLQAEGGGGHGIHDLFDRLYAATARE MRVQVGYHHDAEAARQQAEFIEDLKCEVVARAFCAYLKTL GIQGDTFRRQPEPLPTWPDLPDLPSSTIGTAQAALYSVLHLMP VEDVGSLLHQLRRWLVALQARGGEDGTAITATIPLLELYLN RHDAKFSGGGAGTGLRWDDWQVFFDCQATFDRVFPPGPAL DSHRLPLRGLREVLRFGRVNDLAALIGQDKITAAEVDRWHT AEQTIAAQQQRREALHEQLSRKKGTDAEVDEYRALVTAIAD HRHLTAHVTLSNVVRLHRLMTTVLGRLVDYGGLWERDLTF VTLYEAHRLGGLRNLLSESRVNKFLDGQTPAALSKKNNAEE NGMISKVLGDKARRQIRNDFAHFNMLQQGKKTINLTDEINN ARKLMAHDRKLKNAITRSVTTLLQQDGLDIVWTMDASHRL TDAKIDSRNAIHLHKTHNRANIREPLHGKSYCRWVAALFGA TSTPSATKKSDKIR

[0394]

[0395] 일정 예의 실시형태에서, CRISPR 이펙터 단백질은 하기 표 3으로부터 선택되는 Cas13b 단백질이다.

[0396] 표 3

베르게이 엘라 주헬컴 (SEQ ID No. 67)	1	MENKTSLGNNIYYNPFKPQDKSYFAGYFNAAMENTDSVFRELG KRLKGKEYTSENFFDAIFKENISLVEYERYVKLLSDYFPMARLL DKKEVPIKERKENFKKNFKGIIKAVRDLRNFYTHKEHGEVEITD EIFGVLDEMLKSTVLTVKKKKVKTDKTKEILKKSIEKQLDILCQ KKLEYLRDTARKIEEKRRNQRERGEKELVAPFKYSDKRDDLIA AIYNDAFDVYIDKKKDSLKESSKAKYNTKSDPQQEEGDLKIPIS KNGVVFLLSLFLTKQEIHAFKSKIAGFKATVIDEATVSEATVSHG KNSICFMATHEIFSHLAYKKLKRKVRTAEINYGEAENAEQLSVY AKETLMMQMLDELSKVPDVVYQNLSEDVQKTFIEDWNEYLKE NNGDVGTMEEEQVIHPVIRKRYEDKFNYFAIRFLDEFAQFPTLR FQVHLGNYLHDSRPKENLISDRRIKEKITVFGRLSELEHKKALFI KNTETNEDREHYWEIFPNPNYDFPKENISVNDKDFPIAGSILDRE KQPVAGKIGIKVKLLNQQYVSEVDKAVKAHQLKQRKASKPSIQ NIIEEIVPINESNPKEAIVFGGQPTAYLSMNDIHSILYEFFDKWEK KKEKLEKKGEKELRKEIGKELEKKIVGKIQAQIQQIIDKDTNAKI LKPYQDGNSTAIDKEKLIKDLKQEQNILQKLKDEQTVREKEYN DFIAYQDKNREINKVRDRNHKQYLKDNLKRKYPEAPARKEVL YYREKGKVAVWLANDIKRFMPTDFKNEWKGEQHSLLQKSLAY YEQCKEELKNLLPEKVFQHLPFKLGGYFQQKYLYQFYTCYLDK RLEYISGLVQQAENFKSENKVFKKVENECFKFLKKQNYTHKEL DARVQSILGYPIFLERGFMDEKPTIIKGKTFKGNEALFADWFRY YKEYQNFQTFYDTENYPLVELEKKQADRKRKTKIYQQKKNDV FTLLMAKHIFKSVFKQDSIDQFSLEDLYQSREERLGNQERARQT GERNTNYIWNKTVDLKLCDGKITVENVKLKNVGDFIKYEYDQR VQAFLKYEENIEWQAFLIKESKEEENYPYVVEREIEQYEKVRRE ELLKEVHLIEEYILEKVKDKEILKKGDNQNFKYYILNGLLKQLK NEDVESYKVFNLNTEPEDVNINQLKQEATDLEQKAFVLTYIRN KFAHNQLPKKEFWDYCQEKYGKIEKEKTYAEYFAEVFKKEKE ALIK
프레보텔라	2	MEDDKKTTDSIRYELKDKHFWAAFLNLARHNVYITVNHINKIL

[0397]

인터메디아 (SEQ ID No. 68)		EEGEINRDGYETTLKNTWNEIKDINKKDRLSKLIIKHFPFLEAAT YRLNPTDTTKQKEEKQAEAQSLESLRKSFFVFIYKLRDLRNHYS HYKHSKSLERPKFEEGLLEKMYNIFNASIRLVKEDYQYNKDINP DEDFKHLDRTEEEFNYYFTKDNEGNITESGLLFFVSLFLEKKDAI WMQQKLRGFKDNRENKKKMTNEVFCRSRMLLPKLRLQSTQTQ DWILLDMLNELIRCPKSLYERLREEDREKFRVPIEIADEDYDAEQ EPFKNTLVRHQDRFPYFALRYFDYNEIFTNLRFQIDLGTYHFSIY KKQIGDYKESHHLTHKLYGFERIQEFTKQNRPDEWRKFVKTFN SFETSKEPYIPETTPHYHLENQKIGIRFRNDNDKIWPSLKTNSEK NEKSKYKLDKSFQAEAFLSVHELLPMMFYYLLLKTENTDNDNE IETKKKENKNDKQEKHKIEEIIENKITEIYALYDTFANGEIKSIDE LEEYCKGKDIEIGHLPKQMIAILKDEHKVMATEAERKQEEMLV DVQKSLESLDNQINEEIENVERKNSSLKSGKIASWLVNDMMRF QPVQKDNEGKPLNNSKANSTEYQLLQRTLAFFGSEHERLAPYF KQTKLIESSNPHPFLKDTEWEKCNNILSFYRSYLEAKKNFLESLK PEDWEKNQYFLKLKEPKTKPKTLVQGWKNGFNLPRGIFTEPIRK WFMKHRENITVAELKRVGLVAKVIPLFFSEEYKDSVQPFYNYH FNVGNINKPDEKNFLNCEERRELLRKKKDEFKKMTDKEKEENP SYLEFKSWNKFERELRLVRNQDIVTWLLCMELFNKKKIKELNV EKIYLKNINTNTTKKEKNTEEKNGEEKNIKEKNNILNRIMPMRL PIKVYGRENFSKNKKKKIRRNTFFTVYIEEKGTKLLKQGNFKAL ERDRRLGGLFSFVKTPSKAESKSNTISKLRVEYELGEYQKARIEII KDMLALEKTLIDKYNSLDTDNFNKMLTDWLELKGEPDKASFQ NDVDLLIAVRNAFSHNQYPMRNRIAFANINPFSLSSANTSEEKG LGIANQLKDKTHKTIEKIIEIEKPIETKE
프레보텔라 부카에 (SEQ ID No. 69)	3	MQKQDKLFVDRKKNAIFAFPKYITIMENKEKPEPIYYELTDKHF WAAFLNLARHNVYTTINHINRRLEIAELKDDGYMMGIKGSWNE QAKKLDKKVRLRDLIMKHFPFLEAAAYEMTNSKSPNNKEQRE KEQSEALSLNNLKNVLFIFLEKLQVLRNYYSHYKYSEESPKPIFE TSLLKNMYKVFDANVRLVKRDYMHHENIDMQRDFTHLNRKK QVGRTKNIIDSPNFHYHFADKEGNMTIAGLLFFVSLFLDKKDAI WMQKKLKGFKDGRNLREQMTNEVFCRSRISLPKLKLENVQTK DWMQLDMLNELVRCPKSLYERLREKDRESFKVPFDIFSDDYNA

[0398]

		EEEPFKNTLVRHQDRFPYFVLRYFDLNEIFEQLRFQIDLGTYHFS IYNKRIGDEDEVRHLTHHLYGFARIQDFAPQNQPEEWRKLVKD LDHFETSQEPYISKTAPHYHLENEKIGIKFCSAHNNLFPSLQTDK TCNGRSKFNLGTQFTAEAFLSVHELLPMMFYYLLLTKDYSRKE SADKVEGIIRKEISNIYAIYDAFANNEINSIADLTRRLQNTNILQG HLPKQMISILKGRQKDMGKEAERKIGEMIDDTQRRLDLLCKQT NQKIRIGKRNAGLLKSGKIADWLVNDMMRFQPVQKDQNNIPIN NSKANSTEYRMLQRALALFGSENFRLKAYFNQMNLVGNDNPH PFLAETQWEHQTNILSFYRNYLEARKKYLKGLKPQNWKQYQH FLILKVQKTNRNTLVTGWKNSFNLPRGIFTQPIREWFEKHNNSK RIYDQILSFDRVGFVAKAIPLYFAEEYKDNVQPFYDYPFNIGNRL KPKKRQFLDKKERVELWQKNKELFKNYPSEKKKTDLAYLDFLS WKKFERELRLIKNQDIVTWLMFKELFNMATVEGLKIGEIHLRDI DTNTANEESNNILNRIMPMKLPVKTYETDNKGNILKERPLATFY IEETETKVLKQGNFKALVKDRRLNGLFSFAETTDLNLEEHPISKL SVDLELIKYQTTRISIFEMTLGLEKKLIDKYSTLPTDSFRNMLER WLQCKANRPELKNYVNSLIAVRNAFSHNQYPMYDATLFAEVK KFTLFPSVDTKKIELNIAPQLLEIVGKAIKEIEKSENKN
포르피로모나스 진지발리스 (SEQ ID No. 70)	4	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFAVFFKPDDFVLA KNRKEQLISVADGKECLTVSGFAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLDEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMDQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV

[0399]

		AELRLLDPSSGHPFLSATMETAHRYTEGFYKCYLEKKREWLAK IFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKVMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVRDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
박테로이데스 피오제네스 (SEQ ID No. 71)	5	MESIKNSQKSTGKTLQKDPPYFGLYLNMALLNVRKVENHIRKW LGDVALLPEKSGFHSLLTTDNLSSAKWTRFYYKSRKFLPFLEMF DSDKKSYENRRETAECLDTIDRQKISSLLKEVYGKLQDIRNAFS HYHIDDQSVKHTALIISSEMHRFIENAYSFALQKTRARFTGVFVE TDFLQAEEKGDNKKFFAIGGNEGIKLKDNALIFLICLFLDREEAF KFLSRATGFKSTKEKGFLAVRETFCALCCRQPHERLLSVNPREA LLMDMLNELNRCPDILFEMLDEKDQKSFLPLLGEEEQAHILENS LNDELCEAIDDPFEMIASLSKRVRYKNRFPYLMLRYIEEKNLLPF IRFRIDLGCLELASYPKKMGEENNYERSVTDHAMAFGRLTDFH NEDAVLQQITKGITDEVRFSLYAPRYAIYNNKIGFVRTSGSDKIS FPTLKKKGGEGHCVAYTLQNTKSFGFISIYDLRKILLLSFLDKDK AKNIVSGLLEQCEKHWKDLSENLFDAIRTELQKEFPVPLIRYTLP RSKGGKLVSSKLADKQEKYESEFERRKEKLTEILSEKDFDLSQIP RRMIDEWLNVLPTSREKKLKGYVETLKLDCRERLRVFEKREKG EHPLPPRIGEMATDLAKDIIRMVIDQGVKQRITSAYYSEIQRCLA QYAGDDNRRHLDSIIRELRLKDTKNGHPFLGKVLRPGLGHTEK LYQRYFEEKKEWLEATFYPAASPKRVPRFVNPPTGKQKELPLII RNLMKERPEWRDWKQRKNSHPIDLPSQLFENEICRLLKDKIGKE PSGKLKWNEMFKLYWDKEFPNGMQRFYRCKRRVEVFDKVVE YEYSEEGGNYKKYYEALIDEVVRQKISSSKEKSKLQVEDLTLSV RRVFKRAINEKEYQLRLLCEDDRLLFMAVRDLYDWKEAQLDL DKIDNMLGEPVSVSQVIQLEGGQPDAVIKAECKLKDVSKLMRY

[0400]

		CYDGRVKGLMPYFANHEATQEQVEMELRHYEDHRRRVFNWV FALEKSVLKNEKLRRFYEESQGGCEHRRCIDALRKASLVSEEEY EFLVHIRNKSAHNQFPDLEIGKLPPNVTSGFCECIWSKYKAIICRI IPFIDPERRFFGKLLEQK
알리스티페스 sp. ZOR0009 (SEQ ID No. 72)	6	MSNEIGAFREHQFAYAPGNEKQEEATFATYFNLALSNVEGMMF GEVESNPDKIEKSLDTLPPAILRQIASFIWLSKEDHPDKAYSTEE VKVIVTDLVRRLCFYRNYFSHCFYLDTQYFYSDELVDTTAIGEK LPYNFHHFITNRLFRYSLPEITLFRWNEGERKYEILRDGLIFFCCL FLKRGQAERFLNELRFFKRTDEEGRIKRTIFTKYCTRESHKHIGIE EQDFLIFQDIIGDLNRVPKVCDGVVDLSKENERYIKNRETSNESD ENKARYRLLIREKDKFPYYLMRYIVDFGVLPCITFKQNDYSTKE GRGQFHYQDAAVAQEERCYNFVVRNGNVYYSYMPQAQNVVR ISELQGTISVEELRNMVYASINGKDVNKSVEQYLYHLHLLYEKI LTISGQTIKEGRVDVEDYRPLLDKLLLRPASNGEELRRELRKLLP KRVCDLLSNRFDCSEGVSAVEKRLKAILLRHEQLLLSQNPALHI DKIKSVIDYLYLFFSDDEKFRQQPTEKAHRGLKDEEFQMYHYL VGDYDSHPLALWKELEASGRLKPEMRKLTSATSLHGLYMLCL KGTVEWCRKQLMSIGKGTAKVEAIADRVGLKLYDKLKEYTPE QLEREVKLVVMHGYAAAATPKPKAQAAIPSKLTELRFYSFLGK REMSFAAFIRQDKKAQKLWLRNFYTVENIKTLQKRQAAADAA CKKLYNLVGEVERVHTNDKVLVLVAQRYRERLLNVGSKCAVT LDNPERQQKLADVYEVQNAWLSIRFDDLDFTLTHVNLSNLRKA YNLIPRKHILAFKEYLDNRVKQKLCEECRNVRRKEDLCTCCSPR YSNLTSWLKENHSESSIEREAATMMLLDVERKLLSFLLDERRKA IIEYGKFIPFSALVKECRLADAGLCGIRNDVLHDNVISYADAIGK LSAYFPKEASEAVEYIRRTKEVREQRREELMANSSQ
프레보텔라 sp. MA2016 (SEQ ID No. 73)	7a	MSKECKKQRQEKKRRLQKANFSISLTGKHVFGAYFNMARTNF VKTINYILPIAGVRGNYSENQINKMLHALFLIQAGRNEELTTEQK QWEKKLRLNPEQQTKFQKLLFKHFPVLGPMMADVADHKAYL NKKKSTVQTEDETFAMLKGVSLADCLDIICLMADTLTECRNFY THKDPYNKPSQLADQYLHQEMIAKKLDKVVVASRRILKDREGL SVNEVEFLTGIDHLHQEVLKDEFGNAKVKDGKVMKTFVEYDD FYFKISGKRLVNGYTVTTKDDKPVNVNTMLPALSDFGLLYFCV

[0401]

		LFLSKPYAKLFIDEVRLFEYSPFDDKENMIMSEMLSIYRIRTPRL HKIDSHDSKATLAMDIFGELRRCPMELYNLLDKNAGQPFFHDE VKHPNSHTPDVSKRLRYDDRFPTLALRYIDETELFKRIRFQLQL GSFRYKFYDKENCIDGRVRVRRIQKEINGYGRMQEVADKRMD KWGDLIQKREERSVKLEHEELYINLDQFLEDTADSTPYVTDRRP AYNIHANRIGLYWEDSQNPKQYKVFDENGMYIPELVVTEDKKA PIKMPAPRCALSVYDLPAMLFYEYLREQQDNEFPSAEQVIIEYE DDYRKFFKAVAEGKLKPFKRPKEFRDFLKKEYPKLRMADIPKK LQLFLCSHGLCYNNKPETVYERLDRLTLQHLEERELHIQNRLEH YQKDRDMIGNKDNQYGKKSFSDVRHGALARYLAQSMMEWQP TKLKDKEKGHDKLTGLNYNVLTAYLATYGHPQVPEEGFTPRTL EQVLINAHLIGGSNPHPFINKVLALGNRNIEELYLHYLEEELKHI RSRIQSLSSNPSDKALSALPFIHHDRMRYHERTSEEMMALAARY TTIQLPDGLFTPYILEILQKHYTENSDLQNALSQDVPVKLNPTCN AAYLITLFYQTVLKDNAQPFYLSDKTYTRNKDGEKAESFSFKR AYELFSVLNNNKKDTFPFEMIPLFLTSDEIQERLSAKLLDGDGNP VPEVGEKGKPATDSQGNTIWKRRIYSEVDDYAEKLTDRDMKIS FKGEWEKLPRWKQDKIIKRRDETRRQMRDELLQRMPRYIRDIK DNERTLRRYKTQDMVLFLLAEKMFTNIISEQSSEFNWKQMRLS KVCNEAFLRQTLTFRVPVTVGETTIYVEQENMSLKNYGEFYRFL TDDRLMSLLNNIVETLKPNENGDLVIRHTDLMSELAAYDQYRS TIFMLIQSIENLIITNNAVLDDPDADGFWVREDLPKRNNFASLLE LINQLNNVELTDDERKLLVAIRNAFSHNSYNIDFSLIKDVKHLPE VAKGILQHLQSMLGVEITK
프레보텔라 sp. MA2016 (SEQ ID No. 74)	7b	MSKECKKQRQEKKRRLQKANFSISLTGKHVFGAYFNMARTNF VKTINYILPIAGVRGNYSENQINKMLHALFLIQAGRNEELTTEQK QWEKKLRLNPEQQTKFQKLLFKHFPVLGPMMADVADHKAYL NKKKSTVQTEDETFAMLKGVSLADCLDIICLMADTLTECRNFY THKDPYNKPSQLADQYLHQEMIAKKLDKVVVASRRILKDREGL SVNEVEFLTGIDHLHQEVLKDEFGNAKVKDGKVMKTFVEYDD FYFKISGKRLVNGYTVTTKDDKPVNVNTMLPALSDFGLLYFCV LFLSKPYAKLFIDEVRLFEYSPFDDKENMIMSEMLSIYRIRTPRL HKIDSHDSKATLAMDIFGELRRCPMELYNLLDKNAGQPFFHDE

[0402]

		VKHPNSHTPDVSKRLRYDDRFPTLALRYIDETELFKRIRFQLQL GSFRYKFYDKENCIDGRVRVRRIQKEINGYGRMQEVADKRMD KWGDLIQKREERSVKLEHEELYINLDQFLEDTADSTPYVTDRRP AYNIHANRIGLYWEDSQNPKQYKVFDENGMYIPELVVTEDKKA PIKMPAPRCALSVYDLPAMLFYEYLREQQDNEFPSAEQVIIEYE DDYRKFFKAVAEGKLKPFKRPKEFRDFLKKEYPKLRMADIPKK LQLFLCSHGLCYNNKPETVYERLDRLTLQHLEERELHIQNRLEH YQKDRDMIGNKDNQYGKKSFSDVRHGALARYLAQSMMEWQP TKLKDKEKGHDKLTGLNYNVLTAYLATYGHPQVPEEGFTPRTL EQVLINAHLIGGSNPHPFINKVLALGNRNIEELYLHYLEEELKHI RSRIQSLSSNPSDKALSALPFIHHDRMRYHERTSEEMMALAARY TTIQLPDGLFTPYILEILQKHYTENSDLQNALSQDVPVKLNPTCN AAYLITLFYQTVLKDNAQPFYLSDKTYTRNKDGEKAESFSFKR AYELFSVLNNNKKDTFPFEMIPLFLTSDEIQERLSAKLLDGDGNP VPEVGEKGKPATDSQGNTIWKRRIYSEVDDYAEKLTDRDMKIS FKGEWEKLPRWKQDKIIKRRDETRRQMRDELLQRMPRYIRDIK DNERTLRRYKTQDMVLFLLAEKMFTNIISEQSSEFNWKQMRLS KVCNEAFLRQTLTFRVPVTVGETTIYVEQENMSLKNYGEFYRFL TDDRLMSLLNNIVETLKPNENGDLVIRHTDLMSELAAYDQYRS TIFMLIQSIENLIITNNAVLDDPDADGFWVREDLPKRNNFASLLE LINQLNNVELTDDERKLLVAIRNAFSHNSYNIDFSLIKDVKHLPE VAKGILQHLQSMLGVEITK
리에메렐라 아나티페 스티퍼 (SEQ ID No. 75)	8	MEKPLLPNVYTLKHKFFWGAFLNIARHNAFITICHINEQLGLKT PSNDDKIVDVVCETWNNILNNDHDLLKKSQLTELILKHFPFLTA MCYHPPKKEGKKKGHQKEQQKEKESEAQSQAEALNPSKLIEAL EILVNQLHSLRNYYSHYKHKKPDAEKDIFKHLYKAFDASLRMV KEDYKAHFTVNLTRDFAHLNRKGKNKQDNPDFNRYRFEKDGF FTESGLLFFTNLFLDKRDAYWMLKKVSGFKASHKQREKMTTE VFCRSRILLPKLRLESRYDHNQMLLDMLSELSRCPKLLYEKLSE ENKKHFQVEADGFLDEIEEEQNPFKDTLIRHQDRFPYFALRYLD LNESFKSIRFQVDLGTYHYCIYDKKIGDEQEKRHLTRTLLSFGRL QDFTEINRPQEWKALTKDLDYKETSNQPFISKTTPHYHITDNKIG FRLGTSKELYPSLEIKDGANRIAKYPYNSGFVAHAFISVHELLPL

[0403]

		MFYQHLTGKSEDLLKETVRHIQRIYKDFEEERINTIEDLEKANQ GRLPLGAFPKQMLGLLQNKQPDLSEKAKIKIEKLIAETKLLSHR LNTKLKSSPKLGKRREKLIKTGVLADWLVKDFMRFQPVAYDA QNQPIKSSKANSTEFWFIRRALALYGGEKNRLEGYFKQTNLIGN TNPHPFLNKFNWKACRNLVDFYQQYLEQREKFLEAIKNQPWEP YQYCLLLKIPKENRKNLVKGWEQGGISLPRGLFTEAIRETLSED LMLSKPIRKEIKKHGRVGFISRAITLYFKEKYQDKHQSFYNLSY KLEAKAPLLKREEHYEYWQQNKPQSPTESQRLELHTSDRWKD YLLYKRWQHLEKKLRLYRNQDVMLWLMTLELTKNHFKELNL NYHQLKLENLAVNVQEADAKLNPLNQTLPMVLPVKVYPATAF GEVQYHKTPIRTVYIREEHTKALKMGNFKALVKDRRLNGLFSFI KEENDTQKHPISQLRLRRELEIYQSLRVDAFKETLSLEEKLLNKH TSLSSLENEFRALLEEWKKEYAASSMVTDEHIAFIASVRNAFCH NQYPFYKEALHAPIPLFTVAQPTTEEKDGLGIAEALLKVLREYC EIVKSQI
프레보텔라 아우란티아카 (SEQ ID No. 76)	9	MEDDKKTTGSISYELKDKHFWAAFLNLARHNVYITINHINKLLE IREIDNDEKVLDIKTLWQKGNKDLNQKARLRELMTKHFPFLET AIYTKNKEDKKEVKQEKQAEAQSLESLKDCLFLFLDKLQEARN YYSHYKYSEFSKEPEFEEGLLEKMYNIFGNNIQLVINDYQHNKD INPDEDFKHLDRKGQFKYSFADNEGNITESGLLFFVSLFLEKKD AIWMQQKLNGFKDNLENKKKMTHEVFCRSRILMPKLRLESTQT QDWILLDMLNELIRCPKSLYERLQGDDREKFKVPFDPADEDYN AEQEPFKNTLIRHQDRFPYFVLRYFDYNEIFKNLRFQIDLGTYHF SIYKKLIGGQKEDRHLTHKLYGFERIQEFAKQNRPDEWKAIVKD LDTYETSNKRYISETTPHYHLENQKIGIRFRNGNKEIWPSLKTND ENNEKSKYKLDKQYQAEAFLSVHELLPMMFYYLLLKKEKPNN DEINASIVEGFIKREIRNIFKLYDAFANGEINNIDDLEKYCADKGI PKRHLPKQMVAILYDEHKDMVKEAKRKQKEMVKDTKKLLAT LEKQTQKEKEDDGRNVKLLKSGEIARWLVNDMMRFQPVQKD NEGKPLNNSKANSTEYQMLQRSLALYNNEEKPTRYFRQVNLIE SNNPHPFLKWTKWEECNNILTFYYSYLTKKIEFLNKLKPEDWK KNQYFLKLKEPKTNRETLVQGWKNGFNLPRGIFTEPIREWFKR HQNNSKEYEKVEALDRVGLVTKVIPLFFKEEYFKDKEENFKED

[0404]

		TQKEINDCVQPFYNFPYNVGNIHKPKEKDFLHREERIELWDKKK DKFKGYKEKIKSKKLTEKDKEEFRSYLEFQSWNKFERELRLVR NQDIVTWLLCKELIDKLKIDELNIEELKKLRLNNIDTDTAKKEK NNILNRVMPMELPVTVYEIDDSHKIVKDKPLHTIYIKEAETKLL KQGNFKALVKDRRLNGLFSFVKTNSEAESKRNPISKLRVEYELG EYQEARIEIIQDMLALEEKLINKYKDLPTNKFSEMLNSWLEGKD EADKARFQNDVDFLIAVRNAFSHNQYPMHNKIEFANIKPFSLYT ANNSEEKGLGIANQLKDKTKETTDKIKKIEKPIETKE
프레보텔라 사카롤리티카 (SEQ ID No. 77)	10	MEDKPFWAAFFNLARHNVYLTVNHINKLLDLEKLYDEGKHKEI FEREDIFNISDDVMNDANSNGKKRKLDIKKIWDDLDTDLTRKY QLRELILKHFPFIQPAIIGAQTKERTTIDKDKRSTSTSNDSLKQTG EGDINDLLSLSNVKSMFFRLLQILEQLRNYYSHVKHSKSATMPN FDEDLLNWMRYIFIDSVNKVKEDYSSNSVIDPNTSFSHLIYKDE QGKIKPCRYPFTSKDGSINAFGLLFFVSLFLEKQDSIWMQKKIPG FKKASENYMKMTNEVFCRNHILLPKIRLETVYDKDWMLLDML NEVVRCPLSLYKRLTPAAQNKFKVPEKSSDNANRQEDDNPFSRI LVRHQNRFPYFVLRFFDLNEVFTTLRFQINLGCYHFAICKKQIGD KKEVHHLIRTLYGFSRLQNFTQNTRPEEWNTLVKTTEPSSGNDG KTVQGVPLPYISYTIPHYQIENEKIGIKIFDGDTAVDTDIWPSVST EKQLNKPDKYTLTPGFKADVFLSVHELLPMMFYYQLLLCEGML KTDAGNAVEKVLIDTRNAIFNLYDAFVQEKINTITDLENYLQDK PILIGHLPKQMIDLLKGHQRDMLKAVEQKKAMLIKDTERRLKL LDKQLKQETDVAAKNTGTLLKNGQIADWLVNDMMRFQPVKR DKEGNPINCSKANSTEYQMLQRAFAFYATDSCRLSRYFTQLHLI HSDNSHLFLSRFEYDKQPNLIAFYAAYLKAKLEFLNELQPQNW ASDNYFLLLRAPKNDRQKLAEGWKNGFNLPRGLFTEKIKTWFN EHKTIVDISDCDIFKNRVGQVARLIPVFFDKKFKDHSQPFYRYDF NVGNVSKPTEANYLSKGKREELFKSYQNKFKNNIPAEKTKEYR EYKNFSLWKKFERELRLIKNQDILIWLMCKNLFDEKIKPKKDIL EPRIAVSYIKLDSLQTNTSTAGSLNALAKVVPMTLAIHIDSPKPK GKAGNNEKENKEFTVYIKEEGTKLLKWGNFKTLLADRRIKGLF SYIEHDDIDLKQHPLTKRRVDLELDLYQTCRIDIFQQTLGLEAQL LDKYSDLNTDNFYQMLIGWRKKEGIPRNIKEDTDFLKDVRNAF

[0405]

		SHNQYPDSKKIAFRRIRKFNPKELILEEEEGLGIATQMYKEVEKV VNRIKRIELFD
HMPREF9 712_03108 [미로이데스 오도라티미 머스 CCUG 10230] (SEQ ID No. 78)	11	MKDILTTDTTEKQNRFYSHKIADKYFFGGYFNLASNNIYEVFEE VNKRNTFGKLAKRDNGNLKNYIIHVFKDELSISDFEKRVAIFAS YFPILETVDKKSIKERNRTIDLTLSQRIRQFREMLISLVTAVDQLR NFYTHYHHSDIVIENKVLDFLNSSFVSTALHVKDKYLKTDKTKE FLKETIAAELDILIEAYKKKQIEKKNTRFKANKREDILNAIYNEA FWSFINDKDKDKDKETVVAKGADAYFEKNHHKSNDPDFALNIS EKGIVYLLSFFLTNKEMDSLKANLTGFKGKVDRESGNSIKYMA TQRIYSFHTYRGLKQKIRTSEEGVKETLLMQMIDELSKVPNVVY QHLSTTQQNSFIEDWNEYYKDYEDDVETDDLSRVIHPVIRKRYE DRFNYFAIRFLDEFFDFPTLRFQVHLGDYVHDRRTKQLGKVESD RIIKEKVTVFARLKDINSAKASYFHSLEEQDKEELDNKWTLFPN PSYDFPKEHTLQHQGEQKNAGKIGIYVKLRDTQYKEKAALEEA RKSLNPKERSATKASKYDIITQIIEANDNVKSEKPLVFTGQPIAY LSMNDIHSMLFSLLTDNAELKKTPEEVEAKLIDQIGKQINEILSK DTDTKILKKYKDNDLKETDTDKITRDLARDKEEIEKLILEQKQR ADDYNYTSSTKFNIDKSRKRKHLLFNAEKGKIGVWLANDIKRF MFKESKSKWKGYQHTELQKLFAYFDTSKSDLELILSNMVMVK DYPIELIDLVKKSRTLVDFLNKYLEARLEYIENVITRVKNSIGTP QFKTVRKECFTFLKKSNYTVVSLDKQVERILSMPLFIERGFMDD KPTMLEGKSYKQHKEKFADWFVHYKENSNYQNFYDTEVYEIT TEDKREKAKVTKKIKQQQKNDVFTLMMVNYMLEEVLKLSSND RLSLNELYQTKEERIVNKQVAKDTQERNKNYIWNKVVDLQLC DGLVHIDNVKLKDIGNFRKYENDSRVKEFLTYQSDIVWSAYLS NEVDSNKLYVIERQLDNYESIRSKELLKEVQEIECSVYNQVANK ESLKQSGNENFKQYVLQGLLPIGMDVREMLILSTDVKFKKEEII QLGQAGEVEQDLYSLIYIRNKFAHNQLPIKEFFDFCENNYRSISD NEYYAEYYMEIFRSIKEKYAN
프레보텔라 인터메디아 (SEQ ID No. 79)	12	MEDDKKTTDSIRYELKDKHFWAAFLNLARHNVYITVNHINKIL EEDEINRDGYENTLENSWNEIKDINKKDRLSKLIIKHFPFLEATT YRQNPTDTTKQKEEKQAEAQSLESLKKSFFVFIYKLRDLRNHYS HYKHSKSLERPKFEEDLQNKMYNIFDVSIQFVKEDYKHNTDINP

[0406]

		KKDFKHLDRKRKGKFHYSFADNEGNITESGLLFFVSLFLEKKDA IWVQKKLEGFKCSNKSYQKMTNEVFCRSRMLLPKLRLESTQTQ DWILLDMLNELIRCPKSLYERLQGVNRKKFYVSFDPADEDYDA EQEPFKNTLVRHQDRFPYFALRYFDYNEVFANLRFQIDLGTYHF SIYKKLIGGQKEDRHLTHKLYGFERIQEFDKQNRPDEWKAIVKD SDTFKKKEEKEEEKPYISETTPHYHLENKKIGIAFKNHNIWPSTQ TELTNNKRKKYNLGTSIKAEAFLSVHELLPMMFYYLLLKTENT KNDNKVGGKKETKKQGKHKIEAIIESKIKDIYALYDAFANGEIN SEDELKEYLKGKDIKIVHLPKQMIAILKNEHKDMAEKAEAKQE KMKLATENRLKTLDKQLKGKIQNGKRYNSAPKSGEIASWLVN DMMRFQPVQKDENGESLNNSKANSTEYQLLQRTLAFFGSEHER LAPYFKQTKLIESSNPHPFLNDTEWEKCSNILSFYRSYLKARKNF LESLKPEDWEKNQYFLMLKEPKTNRETLVQGWKNGFNLPRGFF TEPIRKWFMEHWKSIKVDDLKRVGLVAKVTPLFFSEKYKDSVQ PFYNYPFNVGDVNKPKEEDFLHREERIELWDKKKDKFKGYKA KKKFKEMTDKEKEEHRSYLEFQSWNKFERELRLVRNQDIVTWL LCTELIDKLKIDELNIKELKKLRLKDINTDTAKKEKNNILNRVMP MELPVTVYKVNKGGYIIKNKPLHTIYIKEAETKLLKQGNFKALV KDRRLNGLFSFVKTPSEAESESNPISKLRVEYELGKYQNARLDII EDMLALEKKLIDKYNSLDTDNFHNMLTGWLELKGEAKKARFQ NDVKLLTAVRNAFSHNQYPMYDENLFGNIERFSLSSSNIIESKGL DIAAKLKEEVSKAAKKIQNEEDNKKEKET
카프노사이토파가카니몰서스 (SEQ ID No. 80)	13	MKNIQRLGKGNEFSPFKKEDKFYFGGFLNLANNNIEDFFKEIITR FGIVITDENKKPKETFGEKILNEIFKKDISIVDYEKWVNIFADYFP FTKYLSLYLEEMQFKNRVICFRDVMKELLKTVEALRNFYTHYD HEPIKIEDRVFYFLDKVLLDVSLTVKNKYLKTDKTKEFLNQHIG EELKELCKQRKDYLVGKGKRIDKESEIINGIYNNAFKDFICKREK QDDKENHNSVEKILCNKEPQNKKQKSSATVWELCSKSSSKYTE KSFPNRENDKHCLEVPISQKGIVFLLSFFLNKGEIYALTSNIKGFK AKITKEEPVTYDKNSIRYMATHRMFSFLAYKGLKRKIRTSEINY NEDGQASSTYEKETLMLQMLDELNKVPDVVYQNLSEDVQKTFI EDWNEYLKENNGDVGTMEEEQVIHPVIRKRYEDKFNYFAIRFL DEFAQFPTLRFQVHLGNYLCDKRTKQICDTTTEREVKKKITVFG

[0407]

		RLSELENKKAIFLNEREEIKGWEVFPNPSYDFPKENISVNYKDFP IVGSILDREKQPVSNKIGIRVKIADELQREIDKAIKEKKLRNPKNR KANQDEKQKERLVNEIVSTNSNEQGEPVVFIGQPTAYLSMNDIH SVLYEFLINKISGEALETKIVEKIETQIKQIIGKDATTKILKPYTNA NSNSINREKLLRDLEQEQQILKTLLEEQQQREKDKKDKKSKRK HELYPSEKGKVAVWLANDIKRFMPKAFKEQWRGYHHSLLQKY LAYYEQSKEELKNLLPKEVFKHFPFKLKGYFQQQYLNQFYTDY LKRRLSYVNELLLNIQNFKNDKDALKATEKECFKFFRKQNYIIN PINIQIQSILVYPIFLKRGFLDEKPTMIDREKFKENKDTELADWF MHYKNYKEDNYQKFYAYPLEKVEEKEKFKRNKQINKQKKND VYTLMMVEYIIQKIFGDKFVEENPLVLKGIFQSKAERQQNNTHA ATTQERNLNGILNQPKDIKIQGKITVKGVKLKDIGNFRKYEIDQR VNTFLDYEPRKEWMAYLPNDWKEKEKQGQLPPNNVIDRQISK YETVRSKILLKDVQELEKIISDEIKEEHRHDLKQGKYYNFKYYIL NGLLRQLKNENVENYKVFKLNTNPEKVNITQLKQEATDLEQKA FVLTYIRNKFAHNQLPKKEFWDYCQEKYGKIEKEKTYAEYFAE VFKREKEALIK
포르피로모나스 굴라에 (SEQ ID No. 81)	14	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNFDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSGSSELPLFDGNMLQRLYNV FDVSVQRVKIDHEHNDEVDPHYHFNHLVRKGKKDRYGHNDNP SFKHHFVDGEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTETYQQMTNEVFCRSRISLPKLKLESLRMDDWMLLDMLNEL VRCPKPLYDRLREDDRACFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTSPHYHIEKGKIGLRFMPEGQHLWPSPEVGTTRTGRS KYAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAERV QGRIKRVIEDVYAVYDAFARDEINTRDELDACLADKGIRRGHLP RQMIAILSQEHKDMEEKIRKKLQEMMADTDHRLDMLDRQTDR KIRIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDASGKPLNNS KANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFL

[0408]

		HETRWESHTNILSFYRSYLRARKAFLERIGRSDRVENRPFLLLKE PKTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGHDEVASYK EVGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRFL SKEERAEEWERGKERFRDLEAWSYSAARRIEDAFAGIEYASPG NKKKIEQLLRDLSLWEAFESKLKVRADRINLAKLKKEILEAQEH PYHDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLD TGTLYLKDIRPNVQEQGSLNVLNRVKPMRLPVVVYRADSRGH VHKEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDT GGLAMEQYPISKLRVEYELAKYQTARVCVFELTLRLEESLLTRY PHLPDESFREMLESWSDPLLAKWPELHGKVRLLIAVRNAFSHN QYPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAK ETVERIIQA
프레보텔라 sp. P5-125 (SEQ ID No. 82)	15	MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQ NENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIERLQSYFPF LKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMY RDLTNHYKTYEEKLNDGCEFLTSTEQPLSGMINNYYTVALRNM NERYGYKTEDLAFIQDKRFKFVKDAYGKKKSQVNTGFFLSLQD YNGDTQKKLHLSGVGIALLICLFLDKQYINIFLSRLPIFSSYNAQS EERRIIIRSFGINSIKLPKDRIHSEKSNKSVAMDMLNEVKRCPDEL FTTLSAEKQSRFRIISDDHNEVLMKRSSDRFVPLLLQYIDYGKLF DHIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFGRLE EAETMRKQENGTFGNSGIRIRDFENMKRDDANPANYPYIVDTY THYILENNKVEMFINDKEDSAPLLPVIEDDRYVVKTIPSCRMSTL EIPAMAFHMFLFGSKKTEKLIVDVHNRYKRLFQAMQKEEVTAE NIASFGIAESDLPQKILDLISGNAHGKDVDAFIRLTVDDMLTDTE RRIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLF QPSVNDGENKITGLNYRIMQSAIAVYDSGDDYEAKQQFKLMFE KARLIGKGTTEPHPFLYKVFARSIPANAVEFYERYLIERKFYLTG LSNEIKKGNRVDVPFIRRDQNKWKTPAMKTLGRIYSEDLPVELP RQMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEYMKRVLDDD FQTFYQWNRNYRYMDMLKGEYDRKGSLQHCFTSVEEREGLW KERASRTERYRKQASNKIRSNRQMRNASSEEIETILDKRLSNSR NEYQKSEKVIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEI

[0409]

		MPDAEKGILSEIMPMSFTFEKGGKKYTITSEGMKLKNYGDFFVL ASDKRIGNLLELVGSDIVSKEDIMEEFNKYDQCRPEISSIVFNLE KWAFDTYPELSARVDREEKVDFKSILKILLNNKNINKEQSDILR KIRNAFDHNNYPDKGVVEIKALPEIAMSIKKAFGEYAIMK
플라보박테리움 브란키오필럼 (SEQ ID No. 83)	16	MENLNKILDKENEICISKIFNTKGIAAPITEKALDNIKSKQKNDL NKEARLHYFSIGHSFKQIDTKKVFDYVLIEELKDEKPLKFITLQK DFFTKEFSIKLQKLINSIRNINNHYVHNFNDINLNKIDSNVFHFLK ESFELAIIEKYYKVNKKYPLDNEIVLFLKELFIKDENTALLNYFT NLSKDEAIEYILTFTITENKIWNINNEHNILNIEKGKYLTFEAMLF LITIFLYKNEANHLLPKLYDFKNNKSKQELFTFFSKKFTSQDIDA EEGHLIKFRDMIQYLNHYPTAWNNDLKLESENKNKIMTTKLIDS IIEFELNSNYPSFATDIQFKKEAKAFLFASNKKRNQTSFSNKSYN EEIRHNPHIKQYRDEIASALTPISFNVKEDKFKIFVKKHVLEEYFP NSIGYEKFLEYNDFTEKEKEDFGLKLYSNPKTNKLIERIDNHKL VKSHGRNQDRFMDFSMRFLAENNYFGKDAFFKCYKFYDTQEQ DEFLQSNENNDDVKFHKGKVTTYIKYEEHLKNYSYWDCPFVEE NNSMSVKISIGSEEKILKIQRNLMIYFLENALYNENVENQGYKL VNNYYRELKKDVEESIASLDLIKSNPDFKSKYKKILPKRLLHNY APAKQDKAPENAFETLLKKADFREEQYKKLLKKAEHEKNKED FVKRNKGKQFKLHFIRKACQMMYFKEKYNTLKEGNAAFEKKD PVIEKRKNKEHEFGHHKNLNITREEFNDYCKWMFAFNGNDSYK KYLRDLFSEKHFFDNQEYKNLFESSVNLEAFYAKTKELFKKWIE TNKPTNNENRYTLENYKNLILQKQVFINVYHFSKYLIDKNLLNS ENNVIQYKSLENVEYLISDFYFQSKLSIDQYKTCGKLFNKLKSN KLEDCLLYEIAYNYIDKKNVHKIDIQKILTSKIILTINDANTPYKIS VPFNKLERYTEMIAIKNQNNLKARFLIDLPLYLSKNKIKKGKDS AGYEIIIKNDLEIEDINTINNKIINDSVKFTEVLMELEKYFILKDKC ILSKNYIDNSEIPSLKQFSKVWIKENENEIINYRNIACHFHLPLLET FDNLLLNVEQKFIKEELQNVSTINDLSKPQEYLILLFIKFKHNNF YLNLFNKNESKTIKNDKEVKKNRVLQKFINQVILKKK
미로이데스 오도라티미머스 (SEQ ID	17	MKDILTTDTTEKQNRFYSHKIADKYFFGGYFNLASNNIYEVFEE VNKRNTFGKLAKRDNGNLKNYIIHVFKDELSISDFEKRVAIFAS YFPILETVDKKSIKERNRTIDLTLSQRIRQFREMLISLVTAVDQLR

[0410]

No. 84)		NFYTHYHHSDIVIENKVLDFLNSSFVSTALHVKDKYLKTDKTKE FLKETIAAELDILIEAYKKKQIEKKNTRFKANKREDILNAIYNEA FWSFINDKDKDKDKETVVAKGADAYFEKNHHKSNDPDFALNIS EKGIVYLLSFFLTNKEMDSLKANLTGFKGKVDRESGNSIKYMA TQRIYSFHTYRGLKQKIRTSEEGVKETLLMQMIDELSKVPNVVY QHLSTTQQNSFIEDWNEYYKDYEDDVETDDLSRVTHPVIRKRY EDRFNYFAIRFLDEFFDFPTLRFQVHLGDYVHDRRTKQLGKVES DRIIKEKVTVFARLKDINSAKASYFHSLEEQDKEELDNKWTLFP NPSYDFPKEHTLQHQGEQKNAGKIGIYVKLRDTQYKEKAALEE ARKSLNPKERSATKASKYDIITQIIEANDNVKSEKPLVFTGQPIA YLSMNDIHSMLFSLLTDNAELKKTPEEVEAKLIDQIGKQINEILS KDTDTKILKKYKDNDLKETDTDKITRDLARDKEEIEKLILEQKQ RADDYNYTSSTKFNIDKSRKRKHLLFNAEKGKIGVWLANDIKR FMFKESKSKWKGYQHIELQKLFAYFDTSKSDLELILSNMVMVK DYPIELIDLVKKSRTLVDFLNKYLEARLEYIENVITRVKNSIGTP QFKTVRKECFTFLKKSNYTVVSLDKQVERILSMPLFIERGFMDD KPTMLEGKSYKQHKEKFADWFVHYKENSNYQNFYDTEVYEIT TEDKREKAKVTKKIKQQQKNDVFTLMMVNYMLEEVLKLSSND RLSLNELYQTKEERIVNKQVAKDTQERNKNYIWNKVVDLQLC DGLVHIDNVKLKDIGNFRKYENDSRVKEFLTYQSDIVWSAYLS NEVDSNKLYVIERQLDNYESIRSKELLKEVQEIECSVYNQVANK ESLKQSGNENFKQYVLQGLLPIGMDVREMLILSTDVKFKKEEII QLGQAGEVEQDLYSLIYIRNKFAHNQLPIKEFFDFCENNYRSISD NEYYAEYYMEIFRSIKEKYAN
플라보박테리움 콜럼나레 (SEQ ID No. 85)	18	MSSKNESYNKQKTFNHYKQEDKYFFGGFLNNADDNLRQVGKE FKTRINFNHNNNELASVFKDYFNKEKSVAKREHALNLLSNYFP VLERIQKHTNHNFEQTREIFELLLDTIKKLRDYYTHHYHKPITIN PKIYDFLDDTLLDVLITIKKKKVKNDTSRELLKEKLRPELTQLKN QKREELIKKGKKLLEENLENAVFNHCLIPFLEENKTDDKQNKTV SLRKYRKSKPNEETSITLTQSGLVFLMSFFLHRKEFQVFTSGLER FKAKVNTIKEEEISLNKNNIVYMITHWSYSYYNFKGLKHRIKTD QGVSTLEQNNTTHSLTNTNTKEALLTQIVDYLSKVPNEIYETLSE KQQKEFEEDINEYMRENPENEDSTFSSIVSHKVIRKRYENKFNY

[0411]

		FAMRFLDEYAELPTLRFMVNFGDYIKDRQKKILESIQFDSERIIK KEIHLFEKLSLVTEYKKNVYLKETSNIDLSRFPLFPNPSYVMAN NNIPFYIDSRSNNLDEYLNQKKKAQSQNKKRNLTFEKYNKEQS KDAIIAMLQKEIGVKDLQQRSTIGLLSCNELPSMLYEVIVKDIKG AELENKIAQKIREQYQSIRDFTLDSPQKDNIPTTLIKTINTDSSVT FENQPIDIPRLKNALQKELTLTQEKLLNVKEHEIEVDNYNRNKN TYKFKNQPKNKVDDKKLQRKYVFYRNEIRQEANWLASDLIHF MKNKSLWKGYMHNELQSFLAFFEDKKNDCIALLETVFNLKED CILTKGLKNLFLKHGNFIDFYKEYLKLKEDFLSTESTFLENGFIG LPPKILKKELSKRLKYIFIVFQKRQFIIKELEEKKNNLYADAINLS RGIFDEKPTMIPFKKPNPDEFASWFVASYQYNNYQSFYELTPDI VERDKKKKYKNLRAINKVKIQDYYLKLMVDTLYQDLFNQPLD KSLSDFYVSKAEREKIKADAKAYQKLNDSSLWNKVIHLSLQNN RITANPKLKDIGKYKRALQDEKIATLLTYDARTWTYALQKPEK ENENDYKELHYTALNMELQEYEKVRSKELLKQVQELEKKILDK FYDFSNNASHPEDLEIEDKKGKRHPNFKLYITKALLKNESEIINL ENIDIEILLKYYDYNTEELKEKIKNMDEDEKAKIINTKENYNKIT NVLIKKALVLIIIRNKMAHNQYPPKFIYDLANRFVPKKEEEYFAT YFNRVFETITKELWENKEKKDKTQV
포르피로모나스 진지발리스 (SEQ ID No. 86)	19	MTEQNEKPYNGTYYTLEDKHFWAAFLNLARHNAYITLAHIDR QLAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFS FLEGAAYGKKLFESQSSGNKSSKKKELSKKEKEELQANALSLD NLKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYN VFDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDKYGNND NPFFKHHFVDREGTVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTEAYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKSLYDRLREEDRARFRVPVDILSDEDDTDGTEEDPFKNTL VRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGE QPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGD KPYITQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGATRTGRSK YAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQ GRIKRVIEDVYAVYDAFARDEINTRDELDACLADKGIRRGHLPR QMIAILSQEHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKI

[0412]

		RIGRKNAGLPKSGVVADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLEARKAFLQSIGRSDRVENHRFLLLKEP KTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGYDEVGSYKE VGFMAKAVPLYFERASKDRVQPFYDYPFNVGNSLKPKKGRFLS KEKRAEEWESGKERFRLAKLKKEILEAKEHPYHDFKSWQKFER ELRLVKNQDIITWMMCRDLMEENKVEGLDTGTLYLKDIRTDV QEQGSLNVLNRVKPMRLPVVVYRADSRGHVHKEQAPLATVYI EERDTKLLKQGNFKSFVKDRRLNGLFSFVDTGALAMEQYPISK LRVEYELAKYQTARVCAFEQTLELEESLLTRYPHLPDKNFRKM LESWSDPLLDKWPDLHGNVRLLIAVRNAFSHNQYPMYDETLFS SIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKEMVERIIQA
포르피로모나스 sp. COT-052 OH4946 (SEQ ID No. 87)	20	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNFDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGHNDN PSFKHHFVDSEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREDDRACFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGTTRTGRSK YAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQ GRIKRVIEDVYAIYDAFARDEINTLKELDACLADKGIRRGHLPK QMIGILSQERKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKI RIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLRARKAFLERIGRSDRVENCPFLLLKEP KTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGYDEVGSYRE VGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRFLS KEDRAEEWERGKERFRDLEAWSHSAARRIKDAFAGIEYASPGN KKKIEQLLRDLSLWEAFESKLKVRADKINLAKLKKEILEAQEHP

[0413]

		YHDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDT GTLYLKDIRPNVQEQGSLNVLNRVKPMRLPVVVYRADSRGHV HKEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDTG GLAMEQYPISKLRVEYELAKYQTARVCVFELTLRLEESLLSRYP HLPDESFREMLESWSDPLLAKWPELHGKVRLLIAVRNAFSHNQ YPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKE TVERIIQA
프레보텔라 인터메디아 (SEQ ID No. 88)	21	MEDDKKTKESTNMLDNKHFWAAFLNLARHNVYITVNHINKVL ELKNKKDQDIIIDNDQDILAIKTHWEKVNGDLNKTERLRELMTK HFPFLETAIYTKNKEDKEEVKQEKQAKAQSFDSLKHCLFLFLEK LQEARNYYSHYKYSESTKEPMLEKELLKKMYNIFDDNIQLVIK DYQHNKDINPDEDFKHLDRTEEEFNYYFTTNKKGNITASGLLFF VSLFLEKKDAIWMQQKLRGFKDNRESKKKMTHEVFCRSRMLL PKLRLESTQTQDWILLDMLNELIRCPKSLYERLQGEYRKKFNVP FDSADEDYDAEQEPFKNTLVRHQDRFPYFALRYFDYNEIFTNLR FQIDLGTYHFSIYKKLIGGQKEDRHLTHKLYGFERIQEFAKQNR TDEWKAIVKDFDTYETSEEPYISETAPHYHLENQKIGIRFRNDN DEIWPSLKTNGENNEKRKYKLDKQYQAEAFLSVHELLPMMFY YLLLKKEEPNNDKKNASIVEGFIKREIRDIYKLYDAFANGEINNI DDLEKYCEDKGIPKRHLPKQMVAILYDEHKDMAEEAKRKQKE MVKDTKKLLATLEKQTQGEIEDGGRNIRLLKSGEIARWLVNDM MRFQPVQKDNEGNPLNNSKANSTEYQMLQRSLALYNKEEKPT RYFRQVNLINSSNPHPFLKWTKWEECNNILSFYRSYLTKKIEFLN KLKPEDWEKNQYFLKLKEPKTNRETLVQGWKNGFNLPRGIFTE PIREWFKRHQNDSEEYEKVETLDRVGLVTKVIPLFFKKEDSKDK EEYLKKDAQKEINNCVQPFYGFPYNVGNIHKPDEKDFLPSEERK KLWGDKKYKFKGYKAKVKSKKLTDKEKEEYRSYLEFQSWNK FERELRLVRNQDIVTWLLCTELIDKLKVEGLNVEELKKLRLKDI DTDTAKQEKNNILNRVMPMQLPVTVYEIDDSHNIVKDRPLHTV YIEETKTKLLKQGNFKALVKDRRLNGLFSFVDTSSETELKSNPIS KSLVEYELGEYQNARIETIKDMLLLEETLIEKYKTLPTDNFSDM LNGWLEGKDEADKARFQNDVKLLVAVRNAFSHNQYPMRNRIA FANINPFSLSSADTSEEKKLDIANQLKDKTHKIIKRIIEIEKPIETK

[0414]

		E
PIN17_0200 [프레보텔라 인터메디아 17] (SEQ ID No. 89)	AFJ07523	MKMEDDKKTKESTNMLDNKHFWAAFLNLARHNVYITVNHIN KVLELKNKKDQDIIIDNDQDILAIKTHWEKVNGDLNKTERLREL MTKHFPFLETAIYTKNKEDKEEVKQEKQAKAQSFDSLKHCLFL FLEKLQEARNYYSHYKYSESTKEPMLEKELLKKMYNIFDDNIQ LVIKDYQHNKDINPDEDFKHLDRTEEEFNYYFTTNKKGNITASG LLFFVSLFLEKKDAIWMQQKLRGFKDNRESKKKMTHEVFCRSR MLLPKLRLESTQTQDWILLDMLNELIRCPKSLYERLQGEYRKKF NVPFDSADEDYDAEQEPFKNTLVRHQDRFPYFALRYFDYNEIFT NLRFQIDLGTYHFSIYKKLIGGQKEDRHLTHKLYGFERIQEFAK QNRTDEWKAIVKDFDTYETSEEPYISETAPHYHLENQKIGIRFRN DNDEIWPSLKTNGENNEKRKYKLDKQYQAEAFLSVHELLPMM FYYLLLKKEEPNNDKKNASIVEGFIKREIRDIYKLYDAFANGEIN NIDDLEKYCEDKGIPKRHLPKQMVAILYDEHKDMAEEAKRKQ KEMVKDTKKLLATLEKQTQGEIEDGGRNIRLLKSGEIARWLVN DMMRFQPVQKDNEGNPLNNSKANSTEYQMLQRSLALYNKEEK PTRYFRQVNLINSSNPHPFLKWTKWEECNNILSFYRSYLTKKIEF LNKLKPEDWEKNQYFLKLKEPKTNRETLVQGWKNGFNLPRGIF TEPIREWFKRHQNDSEEYEKVETLDRVGLVTKVIPLFFKKEDSK DKEEYLKKDAQKEINNCVQPFYGFPYNVGNIHKPDEKDFLPSEE RKKLWGDKKYKFKGYKAKVKSKKLTDKEKEEYRSYLEFQSW NKFERELRLVRNQDIVTWLLCTELIDKLKVEGLNVEELKKLRLK DIDTDTAKQEKNNILNRVMPMQLPVTVYEIDDSHNIVKDRPLHT VYIEETKTKLLKQGNFKALVKDRRLNGLFSFVDTSSETELKSNPI SKSLVEYELGEYQNARIETIKDMLLLEETLIEKYKTLPTDNFSDM LNGWLEGKDEADKARFQNDVKLLVAVRNAFSHNQYPMRNRIA FANINPFSLSSADTSEEKKLDIANQLKDKTHKIIKRIIEIEKPIETK E
프레보텔라 인터메디아 (SEQ ID No. 90)	BAU18623	MEDDKKTTDSISYELKDKHFWAAFLNLARHNVYITVNHINKVL ELKNKKDQDIIIDNDQDILAIKTHWEKVNGDLNKTERLRELMTK HFPFLETAIYSKNKEDKEEVKQEKQAKAQSFDSLKHCLFLFLEK LQETRNYYSHYKYSESTKEPMLEKELLKKMYNIFDDNIQLVIKD YQHNKDINPDEDFKHLDRTEEDFNYYFTRNKKGNITESGLLFFV

[0415]

		SLFLEKKDAIWMQQKLRGFKDNRESKKKMTHEVFCRSRMLLP KLRLESTQTQDWILLDMLNELIRCPKSLYERLQGEDREKFKVPF DPADEDYDAEQEPFKNTLVRHQDRFPYFALRYFDYNEIFTNLRF QIDLGTFHFSIYKKLIGGQKEDRHLTHKLYGFERIQEFAKQNRPD EWKAIVKDLDTYETSNERYISETTPHYHLENQKIGIRFRNDNDEI WPSLKTNGENNEKSKYKLDKQYQAEAFLSVHELLPMMFYYLL LKKEEPNNDKKNASIVEGFIKREIRDMYKLYDAFANGEINNIDD LEKYCEDKGIPKRHLPKQMVAILYDEHKDMVKEAKRKQRKMV KDTEKLLAALEKQTQEKTEDGGRNIRLLKSGEIARWLVNDMM RFQPVQKDNEGNPLNNSKANSTEYQMLQRSLALYNKEEKPTRY FRQVNLINSSNPHPFLKWTKWEECNNILSFYRSYLTKKIEFLNKL KPEDWEKNQYFLKLKEPKTNRETLVQGWKNGFNLPRGIFTEPIR EWFKRHQNDSKEYEKVEALDRVGLVTKVIPLFFKKEDSKDKEE DLKKDAQKEINNCVQPFYSFPYNVGNIHKPDEKDFLHREERIEL WDKKKDKFKGYKAKVKSKKLTDKEKEEYRSYLEFQSWNKFER ELRLVRNQDIVTWLLCTELIDKLKVEGLNVEELKKLRLKDIDTD TAKQEKNNILNRVMPMQLPVTVYEIDDSHNIVKDRPLHTVYIEE TKTKLLKQGNFKALVKDRRLNGLFSFVDTSSEAELKSNPISKSL VEYELGEYQNARIETIKDMLLLEETLIEKYKNLPTDNFSDMLNG WLEGKDEADKARFQNDVKLLVAVRNAFSHNQYPMRNRIAFAN INPFSLSSADTSEEKKLDIANQLKDKTHKIIKRIIEIEKPIETKE
HMPREF6485_0083 [프레보텔라 부카에 ATCC 33574] (SEQ ID No. 91)	EFU31981	MQKQDKLFVDRKKNAIFAFPKYITIMENKEKPEPIYYELTDKHF WAAFLNLARHNVYTTINHINRRLEIAELKDDGYMMGIKGSWNE QAKKLDKKVRLRDLIMKHFPFLEAAAYEMTNSKSPNNKEQRE KEQSEALSLNNLKNVLFIFLEKLQVLRNYYSHYKYSEESPKPIFE TSLLKNMYKVFDANVRLVKRDYMHHENIDMQRDFTHLNRKK QVGRTKNIIDSPNFHYHFADKEGNMTIAGLLFFVSLFLDKKDAI WMQKKLKGFKDGRNLREQMTNEVFCRSRISLPKLKLENVQTK DWMQLDMLNELVRCPKSLYERLREKDRESFKVPFDIFSDDYNA EEEPFKNTLVRHQDRFPYFVLRYFDLNEIFEQLRFQIDLGTYHFS IYNKRIGDEDEVRHLTHHLYGFARIQDFAPQNQPEEWRKLVKD LDHFETSQEPYISKTAPHYHLENEKIGIKFCSAHNNLFPSLQTDK TCNGRSKFNLGTQFTAEAFLSVHELLPMMFYYLLLTKDYSRKE

[0416]

		SADKVEGIIRKEISNIYAIYDAFANNEINSIADLTRRLQNTNILQG HLPKQMISILKGRQKDMGKEAERKIGEMIDDTQRRLDLLCKQT NQKIRIGKRNAGLLKSGKIADWLVNDMMRFQPVQKDQNNIPIN NSKANSTEYRMLQRALALFGSENFRLKAYFNQMNLVGNDNPH PFLAETQWEHQTNILSFYRNYLEARKKYLKGLKPQNWKQYQH FLILKVQKTNRNTLVTGWKNSFNLPRGIFTQPIREWFEKHNNSK RIYDQILSFDRVGFVAKAIPLYFAEEYKDNVQPFYDYPFNIGNRL KPKKRQFLDKKERVELWQKNKELFKNYPSEKKKTDLAYLDFLS WKKFERELRLIKNQDIVTWLMFKELFNMATVEGLKIGEIHLRDI DTNTANEESNNILNRIMPMKLPVKTYETDNKGNILKERPLATFY IEETETKVLKQGNFKALVKDRRLNGLFSFAETTDLNLEEHPISKL SVDLELIKYQTTRISIFEMTLGLEKKLIDKYSTLPTDSFRNMLER WLQCKANRPELKNYVNSLIAVRNAFSHNQYPMYDATLFAEVK KFTLFPSVDTKKIELNIAPQLLEIVGKAIKEIEKSENKN
HMPREF9144_1146 [프레보텔라 팔렌스 ATCC 700821] (SEQ ID No. 92)	EGQ18444	MKEEEKGKTPVVSTYNKDDKHFWAAFLNLARHNVYITVNHIN KILGEGEINRDGYENTLEKSWNEIKDINKKDRLSKLIIKHFPFLE VTTYQRNSADTTKQKEEKQAEAQSLESLKKSFFVFIYKLRDLRN HYSHYKHSKSLERPKFEEDLQEKMYNIFDASIQLVKEDYKHNT DIKTEEDFKHLDRKGQFKYSFADNEGNITESGLLFFVSLFLEKK DAIWVQKKLEGFKCSNESYQKMTNEVFCRSRMLLPKLRLQSTQ TQDWILLDMLNELIRCPKSLYERLREEDRKKFRVPIEIADEDYD AEQEPFKNALVRHQDRFPYFALRYFDYNEIFTNLRFQIDLGTYH FSIYKKQIGDYKESHHLTHKLYGFERIQEFTKQNRPDEWRKFVK TFNSFETSKEPYIPETTPHYHLENQKIGIRFRNDNDKIWPSLKTNS EKNEKSKYKLDKSFQAEAFLSVHELLPMMFYYLLLKTENTDND NEIETKKKENKNDKQEKHKIEEIIENKITEIYALYDAFANGKINSI DKLEEYCKGKDIEIGHLPKQMIAILKSEHKDMATEAKRKQEEM LADVQKSLESLDNQINEEIENVERKNSSLKSGEIASWLVNDMM RFQPVQKDNEGNPLNNSKANSTEYQMLQRSLALYNKEEKPTRY FRQVNLIESSNPHPFLNNTEWEKCNNILSFYRSYLEAKKNFLESL KPEDWEKNQYFLMLKEPKTNCETLVQGWKNGFNLPRGIFTEPI RKWFMEHRKNITVAELKRVGLVAKVIPLFFSEEYKDSVQPFYN YLFNVGNINKPDEKNFLNCEERRELLRKKKDEFKKMTDKEKEE

[0417]

		NPSYLEFQSWNKFERELRLVRNQDIVTWLLCMELFNKKKIKEL NVEKIYLKNINTNTTKKEKNTEEKNGEEKIIKEKNNILNRIMPMR LPIKVYGRENFSKNKKKKIRRNTFFTVYIEEKGTKLLKQGNFKA LERDRRLGGLFSFVKTHSKAESKSNTISKSRVEYELGEYQKARIE IIKDMLALEETLIDKYNSLDTDNFHNMLTGWLKLKDEPDKASF QNDVDLLIAVRNAFSHNQYPMRNRIAFANINPFSLSSANTSEEK GLGIANQLKDKTHKTIEKIIEIEKPIETKE
HMPREF9 714_02132 [미로이데스 오도라티 미머스 CCUG 12901] (SEQ ID No. 93)	EHO08 761	MKDILTTDTTEKQNRFYSHKIADKYFFGGYFNLASNNIYEVFEE VNKRNTFGKLAKRDNGNLKNYIIHVFKDELSISDFEKRVAIFAS YFPILETVDKKSIKERNRTIDLTLSQRIRQFREMLISLVTAVDQLR NFYTHYHHSEIVIENKVLDFLNSSLVSTALHVKDKYLKTDKTKE FLKETIAAELDILIEAYKKKQIEKKNTRFKANKREDILNAIYNEA FWSFINDKDKDKETVVAKGADAYFEKNHHKSNDPDFALNISEK GIVYLLSFFLTNKEMDSLKANLTGFKGKVDRESGNSIKYMATQ RIYSFHTYRGLKQKIRTSEEGVKETLLMQMIDELSKVPNVVYQH LSTTQQNSFIEDWNEYYKDYEDDVETDDLSRVIHPVIRKRYEDR FNYFAIRFLDEFFDFPTLRFQVHLGDYVHDRRTKQLGKVESDRII KEKVTVFARLKDINSAKANYFHSLEEQDKEELDNKWTLFPNPS YDFPKEHTLQHQGEQKNAGKIGIYVKLRDTQYKEKAALEEARK SLNPKERSATKASKYDIITQIIEANDNVKSEKPLVFTGQPIAYLS MNDIHSMLFSLLTDNAELKKTPEEVEAKLIDQIGKQINEILSKDT DTKILKKYKDNDLKETDTDKITRDLARDKEEIEKLILEQKQRAD DYNYTSSTKFNIDKSRKRKHLLFNAEKGKIGVWLANDIKRFMT EEFKSKWKGYQHTELQKLFAYYDTSKSDLDLILSDMVMVKDY PIELIALVKKSRTLVDFLNKYLEARLGYMENVITRVKNSIGTPQF KTVRKECFTFLKKSNYTVVSLDKQVERILSMPLFIERGFMDDKP TMLEGKSYQQHKEKFADWFVHYKENSNYQNFYDTEVYEITTE DKREKAKVTKKIKQQQKNDVFTLMMVNYMLEEVLKLSSNDRL SLNELYQTKEERIVNKQVAKDTQERNKNYIWNKVVDLQLCEG LVRIDKVKLKDIGNFRKYENDSRVKEFLTYQSDIVWSAYLSNEV DSNKLYVIERQLDNYESIRSKELLKEVQEIECSVYNQVANKESL KQSGNENFKQYVLQGLVPIGMDVREMLILSTDVKFIKEEIIQLG QAGEVEQDLYSLIYIRNKFAHNQLPIKEFFDFCENNYRSISDNEY

[0418]

		YAEYYMEIFRSIKEKYTS
HMPREF9711_00870 [미로이데스 오도라티 미머스 CCUG 3837] (SEQ ID No. 94)	EKB06014	MKDILTTDTTEKQNRFYSHKIADKYFFGGYFNLASNNIYEVFEEVNKRNTFGKLAKRDNGNLKNYIIHVFKDELSISDFEKRVAIFASYFPILETVDKKSIKERNRTIDLTLSQRIRQFREMLISLVTAVDQLRNFYTHYHHSEIVIENKVLDFLNSSLVSTALHVKDKYLKTDKTKEFLKETIAAELDILIEAYKKKQIEKKNTRFKANKREDILNAIYNEAFWSFINDKDKDKETVVAKGADAYFEKNHHKSNDPDFALNISEKGIVYLLSFFLTNKEMDSLKANLTGFKGKVDRESGNSIKYMATQRIYSFHTYRGLKQKIRTSEEGVKETLLMQMIDELSKVPNVVYQHLSTTQQNSFIEDWNEYYKDYEDDVETDDLSRVIHPVIRKRYEDRFNYFAIRFLDEFFDFPTLRFQVHLGDYVHDRRTKQLGKVESDRIIKEKVTVFARLKDINSAKASYFHSLEEQDKEELDNKWTLFPNPSYDFPKEHTLQHQGEQKNAGKIGIYVKLRDTQYKEKAALEEARKSLNPKERSATKASKYDIITQIIEANDNVKSEKPLVFTGQPIAYLSMNDIHSMLFSLLTDNAELKKTPEEVEAKLIDQIGKQINEILSKDTDTKILKKYKDNDLKETDTDKITRDLARDKEEIEKLILEQKQRADDYNYTSSTKFNIDKSRKRKHLLFNAEKGKIGVWLANDIKRFMFKESKSKWKGYQHTELQKLFAYFDTSKSDLELILSDMVMVKDYPIELIDLVRKSRTLVDFLNKYLEARLGYIENVITRVKNSIGTPQFKTVRKECFAFLKESNYTVASLDKQIERILSMPLFIERGFMDSKPTMLEGKSYQQHKEDFADWFVHYKENSNYQNFYDTEVYEIITEDKREQAKVTKKIKQQQKNDVFTLMMVNYMLEEVLKLPSNDRLSLNELYQTKEERIVNKQVAKDTQERNKNYIWNKVVDLQLCEGLVRIDKVKLKDIGNFRKYENDSRVKEFLTYQSDIVWSGYLSNEVDSNKLYVIERQLDNYESIRSKELLKEVQEIECIVYNQVANKESLKQSGNENFKQYVLQGLLPRGTDVREMLILSTDVKFKKEEIMQLGQVREVEQDLYSLIYIRNKFAHNQLPIKEFFDFCENNYRPISDNEYYAEYYMEIFRSIKEKYAS
HMPREF9699_02005 [베르게이 엘라 주헬컴 ATCC	EKB54193	MENKTSLGNNIYYNPFKPQDKSYFAGYFNAAMENTDSVFRELGKRLKGKEYTSENFFDAIFKENISLVEYERYVKLLSDYFPMARLLDKKEVPIKERKENFKKNFKGIIKAVRDLRNFYTHKEHGEVEITDEIFGVLDEMLKSTVLTVKKKKVKTDKTKEILKKSIEKQLDILCQKKLEYLRDTARKIEEKRRNQRERGEKELVAPFKYSDKRDDLIA

[0419]

43767] (SEQ ID No. 95)		AIYNDAFDVYIDKKKDSLKESSKAKYNTKSDPQQEEGDLKIPIS KNGVVFLLSLFLTKQEIHAFKSKIAGFKATVIDEATVSEATVSHG KNSICFMATHEIFSHLAYKKLKRKVRTAEINYGEAENAEQLSVY AKETLMMQMLDELSKVPDVVYQNLSEDVQKTFIEDWNEYLKE NNGDVGTMEEEQVIHPVIRKRYEDKFNYFAIRFLDEFAQFPTLR FQVHLGNYLHDSRPKENLISDRRIKEKITVFGRLSELEHKKALFI KNTETNEDREHYWEIFPNPNYDFPKENISVNDKDFPIAGSILDRE KQPVAGKIGIKVKLLNQQYVSEVDKAVKAHQLKQRKASKPSIQ NIIEEIVPINESNPKEAIVFGGQPTAYLSMNDIHSILYEFFDKWEK KKEKLEKKGEKELRKEIGKELEKKIVGKIQAQIQQIIDKDTNAKI LKPYQDGNSTAIDKEKLIKDLKQEQNILQKLKDEQTVREKEYN DFIAYQDKNREINKVRDRNHKQYLKDNLKRKYPEAPARKEVL YYREKGKVAVWLANDIKRFMPTDFKNEWKGEQHSLLQKSLAY YEQCKEELKNLLPEKVFQHLPFKLGGYFQQKYLYQFYTCYLDK RLEYISGLVQQAENFKSENKVFKKVENECFKFLKKQNYTHKEL DARVQSILGYPIFLERGFMDEKPTIIKGKTFKGNEALFADWFRY YKEYQNFQTFYDTENYPLVELEKKQADRKRKTKIYQQKKNDV FTLLMAKHIFKSVFKQDSIDQFSLEDLYQSREERLGNQERARQT GERNTNYIWNKTVDLKLCDGKITVENVKLKNVGDFIKYEYDQR VQAFLKYEENIEWQAFLIKESKEEENYPYVVEREIEQYEKVRRE ELLKEVHLIEEYILEKVKDKEILKKGDNQNFKYYILNGLLKQLK NEDVESYKVFNLNTEPEDVNINQLKQEATDLEQKAFVLTYIRN KFAHNQLPKKEFWDYCQEKYGKIEKEKTYAEYFAEVFKKEKE ALIK
HMPREF9151_01387 [프레보텔라 사카롤리티카 F0055] (SEQ ID No. 96)	EKY00089	MMEKENVQGSHIYYEPTDKCFWAAFYNLARHNAYLTIAHINSF VNSKKGINNDDKVLDIIDDWSKFDNDLLMGARLNKLILKHFPFL KAPLYQLAKRKTRKQQGKEQQDYEKKGDEDPEVIQEAIANAFK MANVRKTLHAFLKQLEDLRNHFSHYNYNSPAKKMEVKFDDGF CNKLYYVFDAALQMVKDDNRMNPEINMQTDFEHLVRLGRNR KIPNTFKYNFTNSDGTINNNGLLFFVSLFLEKRDAIWMQKKIKG FKGGTENYMRMTNEVFCRNRMVIPKLRLETDYDNHQLMFDML NELVRCPLSLYKRLKQEDQDKFRVPIEFLDEDNEADNPYQENA NSDENPTEETDPLKNTLVRHQHRFPYFVLRYFDLNEVFKQLRFQ

[0420]

		INLGCYHFSIYDKTIGERTEKRHLTRTLFGFDRLQNFSVKLQPEH WKNMVKHLDTEESSDKPYLSDAMPHYQIENEKIGIHFLKTDTE KKETVWPSLEVEEVSSNRNKYKSEKNLTADAFLSTHELLPMMF YYQLLSSEEKTRAAAGDKVQGVLQSYRKKIFDIYDDFANGTINS MQKLDERLAKDNLLRGNMPQQMLAILEHQEPDMEQKAKEKL DRLITETKKRIGKLEDQFKQKVRIGKRRADLPKVGSIADWLVND MMRFQPAKRNADNTGVPDSKANSTEYRLLQEALAFYSAYKDR LEPYFRQVNLIGGTNPHPFLHRVDWKKCNHLLSFYHDYLEAKE QYLSHLSPADWQKHQHFLLLKVRKDIQNEKKDWKKSLVAGW KNGFNLPRGLFTESIKTWFSTDADKVQITDTKLFENRVGLIAKLI PLYYDKVYNDKPQPFYQYPFNINDRYKPEDTRKRFTAASSKLW NEKKMLYKNAQPDSSDKIEYPQYLDFLSWKKLERELRMLRNQ DMMVWLMCKDLFAQCTVEGVEFADLKLSQLEVDVNVQDNLN VLNNVSSMILPLSVYPSDAQGNVLRNSKPLHTVYVQENNTKLL KQGNFKSLLKDRRLNGLFSFIAAEGEDLQQHPLTKNRLEYELSI YQTMRISVFEQTLQLEKAILTRNKTLCGNNFNNLLNSWSEHRTD KKTLQPDIDFLIAVRNAFSHNQYPMSTNTVMQGIEKFNIQTPKL EEKDGLGIASQLAKKTKDAASRLQNIINGGTN
A343_1752 [포르피로모나스 진지발리스 JCVI SC001] (SEQ ID No. 97)	EOA10535	MTEQNEKPYNGTYYTLEDKHFWAAFFNLARHNAYITLTHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRCGNNDN PFFKHHFVDREEKVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNELV RCPKSLYDRLREEDRARFRVPVDILSDEDDTDGTEEDPFKNTLV RHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGEQ PEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGDK PYITQTTPHYHIEKGKIGLRFVPEGQLLWPSPEVGATRTGRSKY AQDKRFTAEAFLSVHELMPMMFYYFLLREKYSEEASAERVQGR IKRVIEDVYAVYDAFARGEIDTLDRLDACLADKGIRRGHLPRQ MIAILSQEHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKIR IGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSKA

[0421]

		NSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLHE TRWESHTNILSFYRSYLKARKAFLQSIGRSDRVENHRFLLLKEP KTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGLDEVGSYKE VGFMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGRFLS KEKRAEEWESGKERFRDLEAWSHSAARRIEDAFAGIENASREN KKKIEQLLQDLSLWETFESKLKVKADKINIAKLKKEILEAKEHP YLDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDT GTLYLKDIRTDVHEQGSLNVLNRVKPMRLPVVVYRADSRGHV HKEQAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDT GALAMEQYPISKLRVEYELAKYQTARVCAFEQTLELEESLLTRY PHLPDKNFRKMLESWSDPLLDKWPDLHGNVRLLIAVRNAFSHN QYPMYDETLFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAK EMVERIIQA
HMPREF1981_03090 [박테로이데스 피오제네스 F0041] (SEQ ID No. 98)	ERI81700	MESIKNSQKSTGKTLQKDPPYFGLYLNMALLNVRKVENHIRKW LGDVALLPEKSGFHSLLTTDNLSSAKWTRFYYKSRKFLPFLEMF DSDKKSYENRRETTECLDTIDRQKISSLLKEVYGKLQDIRNAFS HYHIDDQSVKHTALIISSEMHRFIENAYSFALQKTRARFTGVFVE TDFLQAEEKGDNKKFFAIGGNEGIKLKDNALIFLICLFLDREEAF KFLSRATGFKSTKEKGFLAVRETFCALCCRQPHERLLSVNPREA LLMDMLNELNRCPDILFEMLDEKDQKSFLPLLGEEEQAHILENS LNDELCEAIDDPFEMIASLSKRVRYKNRFPYLMLRYIEEKNLLPF IRFRIDLGCLELASYPKKMGEENNYERSVTDHAMAFGRLTDFH NEDAVLQQITKGITDEVRFSLYAPRYAIYNNKIGFVRTGGSDKIS FPTLKKKGGEGHCVAYTLQNTKSFGFISIYDLRKILLLSFLDKDK AKNIVSGLLEQCEKHWKDLSENLFDAIRTELQKEFPVPLIRYTLP RSKGGKLVSSKLADKQEKYESEFERRKEKLTEILSEKDFDLSQIP RRMIDEWLNVLPTSREKKLKGYVETLKLDCRERLRVFEKREKG EHPVPPRIGEMATDLAKDIIRMVIDQGVKQRITSAYYSEIQRCLA QYAGDDNRRHLDSIIRELRLKDTKNGHPFLGKVLRPGLGHTEK LYQRYFEEKKEWLEATFYPAASPKRVPRFVNPPTGKQKELPLII RNLMKERPEWRDWKQRKNSHPIDLPSQLFENEICRLLKDKIGKE PSGKLKWNEMFKLYWDKEFPNGMQRFYRCKRRVEVFDKVVE YEYSEEGGNYKKYYEALIDEVVRQKISSSKEKSKLQVEDLTLSV

[0422]

		RRVFKRAINEKEYQLRLLCEDDRLLFMAVRDLYDWKEAQLDL DKIDNMLGEPVSVSQVIQLEGGQPDAVIKAECKLKDVSKLMRY CYDGRVKGLMPYFANHEATQEQVEMELRHYEDHRRRVFNWV FALEKSVLKNEKLRRFYEESQGGCEHRRCIDALRKASLVSEEEY EFLVHIRNKSAHNQFPDLEIGKLPPNVTSGFCECIWSKYKAIICRI IPFIDPERRFFGKLLEQK
HMPREF1553_02065 [포르피로모나스 진지발리스 F0568] (SEQ ID No. 99)	ERJ65637	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPRSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMDQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLQKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRHQFRAIV AELRLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
HMPREF1	ERJ819	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK

[0423]

988_01768 [포르피로모나스 진지발리스 F0185] (SEQ ID No. 100)	87	FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHDLRKLLLMELLCEGSFSRMQSGFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDHENRFFGKLLNNMSQPINDL
HMPREF1990_01800 [포르피로모나스 진지발리스 W4087] (SEQ ID No. 101)	ERJ87335	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD

[0424]

		LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPRSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMDQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLQKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRHQFRAIV AELRLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKVMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVRDKKRELRTAGKPVPPDLAAYIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKIMTDREEDILPGLKNIDSILDKENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEIPLIYRDVSAKVGSIEGSSAKDLPEG SSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
M573_117042 [프레보텔라 인터메디아 ZT] (SEQ ID No. 102)	KJJ86756	MKMEDDKKTTESTNMLDNKHFWAAFLNLARHNVYITVNHINK VLELKNKKDQDIIIDNDQDILAIKTHWEKVNGDLNKTERLRELM TKHFPFLETAIYTKNKEDKEEVKQEKQAEAQSLESLKDCLFLFL EKLQEARNYYSHYKYSESTKEPMLEEGLLEKMYNIFDDNIQLVI KDYQHNKDINPDEDFKHLDRKGQFKYSFADNEGNITESGLLFF VSLFLEKKDAIWMQQKLTGFKDNRESKKKMTHEVFCRRRMLL PKLRLESTQTQDWILLDMLNELIRCPKSLYERLQGEYRKKFNVP FDSADEDYDAEQEPFKNTLVRHQDRFPYFALRYFDYNEIFTNLR FQIDLGTYHFSIYKKLIGGQKEDRHLTHKLYGFERIQEFAKQNRP DEWKALVKDLDTYETSNERYISETTPHYHLENQKIGIRFRNGNK EIWPSLKTNGENNEKSKYKLDKPYQAEAFLSVHELLPMMFYYL LLKKEEPNNDKKNASIVEGFIKREIRDMYKLYDAFANGEINNIG DLEKYCEDKGIPKRHLPKQMVAILYDEPKDMVKEAKRKQKEM VKDTKKLLATLEKQTQEEIEDGGRNIRLLKSGEIARWLVNDMM RFQPVQKDNEGNPLNNSKANSTEYQMLQRSLALYNKEEKPTRY

[0425]

		FRQVNLINSSNPHPFLKWTKWEECNNILSFYRNYLTKKIEFLNK LKPEDWEKNQYFLKLKEPKTNRETLVQGWKNGFNLPRGIFTEPI REWFKRHQNDSKEYEKVEALKRVGLVTKVIPLFFKEEYFKEDA QKEINNCVQPFYSFPYNVGNIHKPDEKDFLPSEERKKLWGDKK DKFKGYKAKVKSKKLTDKEKEEYRSYLEFQSWNKFERELRLV RNQDIVTWLLCTELIDKMKVEGLNVEELQKLRLKDIDTDTAKQ EKNNILNRIMPMQLPVTVYEIDDSHNIVKDRPLHTVYIEETKTKL LKQGNFKALVKDRRLNGLFSFVDTSSKAELKDKPISKSVVEYEL GEYQNARIETIKDMLLLEKTLIKKYEKLPTDNFSDMLNGWLEG KDESDKARFQNDVKLLVAVRNAFSHNQYPMRNRIAFANINPFS LSSADISEEKKLDIANQLKDKTHKIIKKIIEIEKPIETKE
A2033_10205 [박테로이데테스 박테리아 GWA2_31_9] (SEQ ID No. 103)	OFX18020.1	MENQTQKGKGIYYYYTKNEDKHYFGSFLNLANNNIEQIIEEFRI RLSLKDEKNIKEIINNYFTDKKSYTDWERGINILKEYLPVIDYLD LAITDKEFEKIDLKQKETAKRKYFRTNFSLLIDTIIDLRNFYTHYF HKPISINPDVAKFLDKNLLNVCLDIKKQKMKTDKTKQALKDGL DKELKKLIELKKAELKEKKIKTWNITENVEGAVYNDAFNHMVY KNNAGVTILKDYHKSILPDDKIDSELKLNFSISGLVFLLSMFLSK KEIEQFKSNLEGFKGKVIGENGEYEISKFNNSLKYMATHWIFSY LTFKGLKQRVKNTFDKETLLMQMIDELNKVPHEVYQTLSKEQQ NEFLEDINEYVQDNEENKKSMENSIVVHPVIRKRYDDKFNYFAI RFLDEFANFPTLKFFVTAGNFVHDKREKQIQGSMLTSDRMIKEK INVFGKLTEIAKYKSDYFSNENTLETSEWELFPNPSYLLIQNNIPV HIDLIHNTEEAKQCQIAIDRIKCTTNPAKKRNTRKSKEEIIKIIYQ KNKNIKYGDPTALLSSNELPALIYELLVNKKSGKELENIIVEKIV NQYKTIAGFEKGQNLSNSLITKKLKKSEPNEDKINAEKIILAINRE LEITENKLNIIKNNRAEFRTGAKRKHIFYSKELGQEATWIAYDLK RFMPEASRKEWKGFHHSELQKFLAFYDRNKNDAKALLNMFW NFDNDQLIGNDLNSAFREFHFDKFYEKYLIKRDEILEGFKSFISN FKDEPKLLKKGIKDIYRVFDKRYYIIKSTNAQKEQLLSKPICLPR GIFDNKPTYIEGVKVESNSALFADWYQYTYSDKHEFQSFYDMP RDYKEQFEKFELNNIKSIQNKKNLNKSDKFIYFRYKQDLKIKQIK SQDLFIKLMVDELFNVVFKNNIELNLKKLYQTSDERFKNQLIAD VQKNREKGDTSDNKMNENFIWNMTIPLSLCNGQIEEPKVKLKD

[0426]

		IGKFRKLETDDKVIQLLEYDKSKVWKKLEIEDELENMPNSYERI RREKLLKGIQEFEHFLLEKEKFDGINHPKHFEQDLNPNFKTYVIN GVLRKNSKLNYTEIDKLLDLEHISIKDIETSAKEIHLAYFLIHVRN KFGHNQLPKLEAFELMKKYYKKNNEETYAEYFHKVSSQIVNEF KNSLEKHS
SAMN054 21542_066 6 [크리세오 박테리움 제주엔스] (SEQ ID No. 104)	SDI272 89.1	MEKTQTGLGIYYDHTKLQDKYFFGGFFNLAQNNIDNVIKAFIIK FFPERKDKDINIAQFLDICFKDNDADSDFQKKNKFLRIHFPVIGF LTSDNDKAGFKKKFALLLKTISELRNFYTHYYHKSIEFPSELFEL LDDIFVKTTSEIKKLKKKDDKTQQLLNKNLSEEYDIRYQQQIER LKELKAQGKRVSLTDETAIRNGVFNAAFNHLIYRDGENVKPSR LYQSSYSEPDPAENGISLSQNSILFLLSMFLERKETEDLKSRVKG FKAKIIKQGEEQISGLKFMATHWVFSYLCFKGIKQKLSTEFHEET LLIQIIDELSKVPDEVYSAFDSKTKEKFLEDINEYMKEGNADLSL EDSKVIHPVIRKRYENKFNYFAIRFLDEYLSSTSLKFQVHVGNY VHDRRVKHINGTGFQTERIVKDRIKVFGRLSNISNLKADYIKEQ LELPNDSNGWEIFPNPSYIFIDNNVPIHVLADEATKKGIELFKDK RRKEQPEELQKRKGKISKYNIVSMIYKEAKGKDKLRIDEPLALL SLNEIPALLYQILEKGATPKDIELIIKNKLTERFEKIKNYDPETPAP ASQISKRLRNNTTAKGQEALNAEKLSLLIEREIENTETKLSSIEEK RLKAKKEQRRNTPQRSIFSNSDLGRIAAWLADDIKRFMPAEQRK NWKGYQHSQLQQSLAYFEKRPQEAFLLLKEGWDTSDGSSYWN NWVMNSFLENNHFEKFYKNYLMKRVKYFSELAGNIKQHTHNT KFLRKFIKQQMPADLFPKRHYILKDLETEKNKVLSKPLVFSRGL FDNNPTFIKGVKVTENPELFAEWYSYGYKTEHVFQHFYGWERD YNELLDSELQKGNSFAKNSIYYNRESQLDLIKLKQDLKIKKIKIQ DLFLKRIAEKLFENVFNYPTTLSLDEFYLTQEERAEKERIALAQS LREEGDNSPNIIKDDFIWSKTIAFRSKQIYEPAIKLKDIGKFNRFV LDDEESKASKLLSYDKNKIWNKEQLERELSIGENSYEVIRREKL FKEIQNLELQILSNWSWDGINHPREFEMEDQKNTRHPNFKMYL VNGILRKNINLYKEDEDFWLESLKENDFKTLPSEVLETKSEMVQ LLFLVILIRNQFAHNQLPEIQFYNFIRKNYPEIQNNTVAELYLNLI KLAVQKLKDNS
SAMN054	SHM52	MNTRVTGMGVSYDHTKKEDKHFFGGFLNLAQDNITAVIKAFCI

[0427]

44360_11366 [크리세오박테리움 카르니풀로럼] (SEQ ID No. 105)	812.1	KFDKNPMSSVQFAESCFTDKDSDTDFQNKVRYVRTHLPVIGYL NYGGDRNTFRQKLSTLLKAVDSLRNFYTHYYHSPLALSTELFEL LDTVFASVAVEVKQHKMKDDKTRQLLSKSLAEELDIRYKQQLE RLKELKEQGKNIDLRDEAGIRNGVLNAAFNHLIYKEGEIAKPTL SYSSFYYGADSAENGITISQSGLLFLLSMFLGKKEIEDLKSRIRGF KAKIVRDGEENISGLKFMATHWIFSYLSFKGMKQRLSTDFHEET LLIQIIDELSKVPDEVYHDFDTATREKFVEDINEYIREGNEDFSLG DSTIIHPVIRKRYENKFNYFAVRFLDEFIKFPSLRFQVHLGNFVH DRRIKDIHGTGFQTERVVKDRIKVFGKLSEISSLKTEYIEKELDL DSDTGWEIFPNPSYVFIDNNIPIYISTNKTFKNGSSEFIKLRRKEKP EEMKMRGEDKKEKRDIASMIGNAGSLNSKTPLAMLSLNEMPAL LYEILVKKTTPEEIELIIKEKLDSHFENIKNYDPEKPLPASQISKRL RNNTTDKGKKVINPEKLIHLINKEIDATEAKFALLAKNRKELKE KFRGKPLRQTIFSNMELGREATWLADDIKRFMPDILRKNWKGY QHNQLQQSLAFFNSRPKEAFTILQDGWDFADGSSFWNGWIINSF VKNRSFEYFYEAYFEGRKEYFSSLAENIKQHTSNHRNLRRFIDQ QMPKGLFENRHYLLENLETEKNKILSKPLVFPRGLFDTKPTFIKG IKVDEQPELFAEWYQYGYSTEHVFQNFYGWERDYNDLLESELE KDNDFSKNSIHYSRTSQLELIKLKQDLKIKKIKIQDLFLKLIAGHI FENIFKYPASFSLDELYLTQEERLNKEQEALIQSQRKEGDHSDNII KDNFIGSKTVTYESKQISEPNVKLKDIGKFNRFLLDDKVKTLLS YNEDKVWNKNDLDLELSIGENSYEVIRREKLFKKIQNFELQTLT DWPWNGTDHPEEFGTTDNKGVNHPNFKMYVVNGILRKHTDW FKEGEDNWLENLNETHFKNLSFQELETKSKSIQTAFLIIMIRNQF AHNQLPAVQFFEFIQKKYPEIQGSTTSELYLNFINLAVVELLELL EK
SAMN05421786_1011119 [크리세오박테리움 우레일티컴] (SEQ ID	SIS70481.1	METQILGNGISYDHTKTEDKHFFGGFLNTAQNNIDLLIKAYISKF ESSPRKLNSVQFPDVCFKKNDSDADFQHKLQFIRKHLPVIQYLK YGGNREVLKEKFRLLLQAVDSLRNFYTHFYHKPIQLPNELLTLL DTIFGEIGNEVRQNKMKDDKTRHLLKKNLSEELDFRYQEQLER LRKLKSEGKKVDLRDTEAIRNGVLNAAFNHLIFKDAEDFKPTVS YSSYYYDSDTAENGISISQSGLLFLLSMFLGRREMEDLKSRVRG FKARIIKHEEQHVSGLKFMATHWVFSEFCFKGIKTRLNADYHEE

[0428]

[0429]

No. 106)		TLLIQLIDELSKVPDELYRSFDVATRERFIEDINEYIRDGKEDKSL IESKIVHPVIRKRYESKFNYFAIRFLDEFVNFPTLRFQVHAGNYV HDRRIKSIEGTGFKTERLVKDRIKVFGKLSTISSLKAEYLAKAVN ITDDTGWELLPHPSYVFIDNNIPIHLTVDPSFKNGVKEYQEKRKL QKPEEMKNRQGGDKMHKPAISSKIGKSKDINPESPVALLSMNEI PALLYEILVKKASPEEVEAKIRQKLTAVFERIRDYDPKVPLPASQ VSKRLRNNTDTLSYNKEKLVELANKEVEQTERKLALITKNRRE CREKVKGKFKRQKVFKNAELGTEATWLANDIKRFMPEEQKKN WKGYQHSQLQQSLAFFESRPGEARSLLQAGWDFSDGSSFWNG WVMNSFARDNTFDGFYESYLNGRMKYFLRLADNIAQQSSTNK LISNFIKQQMPKGLFDRRLYMLEDLATEKNKILSKPLIFPRGIFD DKPTFKKGVQVSEEPEAFADWYSYGYDVKHKFQEFYAWDRD YEELLREELEKDTAFTKNSIHYSRESQIELLAKKQDLKVKKVRI QDLYLKLMAEFLFENVFGHELALPLDQFYLTQEERLKQEQEAIV QSQRPKGDDSPNIVKENFIWSKTIPFKSGRVFEPNVKLKDIGKFR NLLTDEKVDILLSYNNTEIGKQVIENELIIGAGSYEFIRREQLFKEI QQMKRLSLRSVRGMGVPIRLNLK
프레보텔라 부카에 (SEQ ID No. 107)	WP_00434358 1	MQKQDKLFVDRKKNAIFAFPKYITIMENQEKPEPIYYELTDKHF WAAFLNLARHNVYTTINHINRRLEIAELKDDGYMMDIKGSWNE QAKKLDKKVRLRDLIMKHFPFLEAAAYEITNSKSPNNKEQREK EQSEALSLNNLKNVLFIFLEKLQVLRNYYSHYKYSEESPKPIFET SLLKNMYKVFDANVRLVKRDYMHHENIDMQRDFTHLNRKKQ VGRTKNIIDSPNFHYHFADKEGNMTIAGLLFFVSLFLDKKDAIW MQKKLKGFKDGRNLREQMTNEVFCRSRISLPKLKLENVQTKD WMQLDMLNELVRCPKSLYERLREKDRESFKVPFDIFSDDYDAE EEPFKNTLVRHQDRFPYFVLRYFDLNEIFEQLRFQIDLGTYHFSI YNKRIGDEDEVRHLTHHLYGFARIQDFAQQNQPEVWRKLVKD LDYFEASQEPYIPKTAPHYHLENEKIGIKFCSTHNNLFPSLKTEK TCNGRSKFNLGTQFTAEAFLSVHELLPMMFYYLLLTKDYSRKE SADKVEGIIRKEISNIYAIYDAFANGEINSIADLTCRLQKTNILQG HLPKQMISILEGRQKDMEKEAERKIGEMIDDTQRRLDLLCKQTN QKIRIGKRNAGLLKSGKIADWLVNDMMRFQPVQKDQNNIPINN SKANSTEYRMLQRALALFGSENFRLKAYFNQMNLVGNDNPHP

		FLAETQWEHQTNILSFYRNYLEARKKYLKGLKPQNWKQYQHF LILKVQKTNRNTLVTGWKNSFNLPRGIFTQPIREWFEKHNNSKR IYDQILSFDRVGFVAKAIPLYFAEEYKDNVQPFYDYPFNIGNKL KPQKGQFLDKKERVELWQKNKELFKNYPSEKKKTDLAYLDFL SWKKFERELRLIKNQDIVTWLMFKELFNMATVEGLKIGEIHLRD IDTNTANEESNNILNRIMPMKLPVKTYETDNKGNILKERPLATF YIEETETKVLKQGNFKVLAKDRRLNGLLSFAETTDIDLEKNPITK LSVDHELIKYQTTRISIFEMTLGLEKKLINKYPTLPTDSFRNMLE RWLQCKANRPELKNYVNSLIAVRNAFSHNQYPMYDATLFAEV KKFTLFPSVDTKKIELNIAPQLLEIVGKAIKEIEKSENKN
포르피로모나스 진지발리스 (SEQ ID No. 108)	WP_00587351 1	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHNLRKLLLMELLCEGSFSRMQSDFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI

[0430]

		MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
포르피로모나스 진지발리스 (SEQ ID No. 109)	WP_00587419 5	MTEQNEKPYNGTYYTLEDKHFWAAFFNLARHNAYITLAHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDKYGNNDN PFFKHHFVDREEKVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTEAYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNELV RCPKSLYDRLREEDRARFRVPVDILSDEDDTDGTEEDPFKNTLV RHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGEQ PEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGDK PYITQTTPHYHIEKGKIGLRFVPEGQLLWPSPEVGATRTGRSKY AQDKRFTAEAFLSVHELMPMMFYYFLLREKYSEEASAEKVQG RIKRVIEDVYAVYDAFARDEINTRDELDACLADKGIRRGHLPRQ MIAILSQEHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKIR IGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSKA NSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLHE TRWESHTNILSFYRSYLKARKAFLQSIGRSDREENHRFLLLKEPK TDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYKEV GFMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGRFLSK EKRAEEWESGKERFRDLEAWSHSAARRIEDAFVGIEYASWENK KKIEQLLQDLSLWETFESKLKVKADKINIAKLKKEILEAKEHPY HDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDTG TLYLKDIRTDVQEQGSLNVLNHVKPMRLPVVVYRADSRGHVH KEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDTGA LAMEQYPISKLRVEYELAKYQTARVCAFEQTLELEESLLTRYPH LPDESFREMLESWSDPLLDKWPDLQREVRLLIAVRNAFSHNQY PMYDETIFSSIRKYDPSSLDAIEERMGLNIAHRLSEEVKLAKEMV ERIIQA
프레보텔라 팔렌스	WP_00604483	MKEEEKGKTPVVSTYNKDDKHFWAAFLNLARHNVYITVNHIN KILGEGEINRDGYENTLEKSWNEIKDINKKDRLSKLIIKHFPFLE

[0431]

(SEQ ID No. 110)	3	VTTYQRNSADTTKQKEEKQAEAQSLESLKKSFFVFIYKLRDLRN HYSHYKHSKSLERPKFEEDLQEKMYNIFDASIQLVKEDYKHNT DIKTEEDFKHLDRKGQFKYSFADNEGNITESGLLFFVSLFLEKK DAIWVQKKLEGFKCSNESYQKMTNEVFCRSRMLLPKLRLQSTQ TQDWILLDMLNELIRCPKSLYERLREEDRKKFRVPIEIADEDYD AEQEPFKNALVRHQDRFPYFALRYFDYNEIFTNLRFQIDLGTYH FSIYKKQIGDYKESHHLTHKLYGFERIQEFTKQNRPDEWRKFVK TFNSFETSKEPYIPETTPHYHLENQKIGIRFRNDNDKIWPSLKTNS EKNEKSKYKLDKSFQAEAFLSVHELLPMMFYYLLLKTENTDND NEIETKKKENKNDKQEKHKIEEIIENKITEIYALYDAFANGKINSI DKLEEYCKGKDIEIGHLPKQMIAILKSEHKDMATEAKRKQEEM LADVQKSLESLDNQINEEIENVERKNSSLKSGEIASWLVNDMM RFQPVQKDNEGNPLNNSKANSTEYQMLQRSLALYNKEEKPTRY FRQVNLIESSNPHPFLNNTEWEKCNNILSFYRSYLEAKKNFLESL KPEDWEKNQYFLMLKEPKTNCETLVQGWKNGFNLPRGIFTEPI RKWFMEHRKNITVAELKRVGLVAKVIPLFFSEEYKDSVQPFYN YLFNVGNINKPDEKNFLNCEERRELLRKKKDEFKKMTDKEKEE NPSYLEFQSWNKFERELRLVRNQDIVTWLLCMELFNKKKIKEL NVEKIYLKNINTNTTKKEKNTEEKNGEEKIIKEKNNILNRIMPMR LPIKVYGRENFSKNKKKIRRNTFFTVYIEEKGTKLLKQGNFKA LERDRRLGGLFSFVKTHSKAESKSNTISKSRVEYELGEYQKARIE IIKDMLALEETLIDKYNSLDTDNFHNMLTGWLKLKDEPDKASF QNDVDLLIAVRNAFSHNQYPMRNRIAFANINPFSLSSANTSEEK GLGIANQLKDKTHKTIEKIIEIEKPIETKE
미로이데스 오도라티미머스(SEQ ID No. 111)	WP_00626141 4	MKDILTTDTTEKQNRFYSHKIADKYFFGGYFNLASNNIYEVFEE VNKRNTFGKLAKRDNGNLKNYIIHVFKDELSISDFEKRVAIFAS YFPILETVDKKSIKERNRTIDLTLSQRIRQFREMLISLVTAVDQLR NFYTHYHHSEIVIENKVLDFLNSSLVSTALHVKDKYLKTDKTKE FLKETIAAELDILIEAYKKKQIEKKNTRFKANKREDILNAIYNEA FWSFINDKDKDKETVVAKGADAYFEKNHHKSNDPDFALNISEK GIVYLLSFFLTNKEMDSLKANLTGFKGKVDRESGNSIKYMATQ RIYSFHTYRGLKQKIRTSEEGVKETLLMQMIDELSKVPNVVYQH LSTTQQNSFIEDWNEYYKDYEDDVETDDLSRVIHPVIRKRYEDR

[0432]

		FNYFAIRFLDEFFDFPTLRFQVHLGDYVHDRRTKQLGKVESDRII KEKVTVFARLKDINSAKANYFHSLEEQDKEELDNKWTLFPNPS YDFPKEHTLQHQGEQKNAGKIGIYVKLRDTQYKEKAALEEARK SLNPKERSATKASKYDIITQIIEANDNVKSEKPLVFTGQPIAYLS MNDIHSMLFSLLTDNAELKKTPEEVEAKLIDQIGKQINEILSKDT DTKILKKYKDNDLKETDTDKITRDLARDKEEIEKLILEQKQRAD DYNYTSSTKFNIDKSRKRKHLLFNAEKGKIGVWLANDIKRFMT EEFKSKWKGYQHTELQKLFAYYDTSKSDLDLILSDMVMVKDY PIELIALVKKSRTLVDFLNKYLEARLGYMENVITRVKNSIGTPQF KTVRKECFTFLKKSNYTVVSLDKQVERILSMPLFIERGFMDDKP TMLEGKSYQQHKEKFADWFVHYKENSNYQNFYDTEVYEITTE DKREKAKVTKKIKQQQKNDVFTLMMVNYMLEEVLKLSSNDRL SLNELYQTKEERIVNKQVAKDTQERNKNYIWNKVVDLQLCEG LVRIDKVKLKDIGNFRKYENDSRVKEFLTYQSDIVWSAYLSNEV DSNKLYVIERQLDNYESIRSKELLKEVQEIECSVYNQVANKESL KQSGNENFKQYVLQGLVPIGMDVREMLILSTDVKFIKEEIIQLG QAGEVEQDLYSLIYIRNKFAHNQLPIKEFFDFCENNYRSISDNEY YAEYYMEIFRSIKEKYTS
미로이데스 오도라티미머스(SEQ ID No. 112)	WP_006265509	MKDILTTDTTEKQNRFYSHKIADKYFFGGYFNLASNNIYEVFEE VNKRNTFGKLAKRDNGNLKNYIIHVFKDELSISDFEKRVAIFAS YFPILETVDKKSIKERNRTIDLTLSQRIRQFREMLISLVTAVDQLR NFYTHYHHSEIVIENKVLDFLNSSLVSTALHVKDKYLKTDKTKE FLKETIAAELDILIEAYKKKQIEKKNTRFKANKREDILNAIYNEA FWSFINDKDKDKETVVAKGADAYFEKNHHKSNDPDFALNISEK GIVYLLSFFLTNKEMDSLKANLTGFKGKVDRESGNSIKYMATQ RIYSFHTYRGLKQKIRTSEEGVKETLLMQMIDELSKVPNVVYQH LSTTQQNSFIEDWNEYYKDYEDDVETDDLSRVIHPVIRKRYEDR FNYFAIRFLDEFFDFPTLRFQVHLGDYVHDRRTKQLGKVESDRII KEKVTVFARLKDINSAKASYFHSLEEQDKEELDNKWTLFPNPS YDFPKEHTLQHQGEQKNAGKIGIYVKLRDTQYKEKAALEEARK SLNPKERSATKASKYDIITQIIEANDNVKSEKPLVFTGQPIAYLS MNDIHSMLFSLLTDNAELKKTPEEVEAKLIDQIGKQINEILSKDT DTKILKKYKDNDLKETDTDKITRDLARDKEEIEKLILEQKQRAD

[0433]

		DYNYTSSTKFNIDKSRKRKHLLFNAEKGKIGVWLANDIKRFMF KESKSKWKGYQHTELQKLFAYFDTSKSDLELILSDMVMVKDYP IELIDLVRKSRTLVDFLNKYLEARLGYIENVITRVKNSIGTPQFKT VRKECFAFLKESNYTVASLDKQIERILSMPLFIERGFMDSKPTML EGKSYQQHKEDFADWFVHYKENSNYQNFYDTEVYEIITEDKRE QAKVTKKIKQQQKNDVFTLMMVNYMLEEVLKLPSNDRLSLNE LYQTKEERIVNKQVAKDTQERNKNYIWNKVVDLQLCEGLVRID KVKLKDIGNFRKYENDSRVKEFLTYQSDIVWSGYLSNEVDSNK LYVIERQLDNYESIRSKELLKEVQEIECIVYNQVANKESLKQSGN ENFKQYVLQGLLPRGTDVREMLILSTDVKFKKEEIMQLGQVRE VEQDLYSLIYIRNKFAHNQLPIKEFFDFCENNYRPISDNEYYAEY YMEIFRSIKEKYAS
프레보텔라 sp. MSX73 (SEQ ID No. 113)	WP_00741216 3	MQKQDKLFVDRKKNAIFAFPKYITIMENQEKPEPIYYELTDKHF WAAFLNLARHNVYTTINHINRRLEIAELKDDGYMMGIKGSWNE QAKKLDKKVRLRDLIMKHFPFLEAAAYEITNSKSPNNKEQREK EQSEALSLNNLKNVLFIFLEKLQVLRNYYSHYKYSEESPKPIFET SLLKNMYKVFDANVRLVKRDYMHHENIDMQRDFTHLNRKKQ VGRTKNIIDSPNFHYHFADKEGNMTIAGLLFFVSLFLDKKDAIW MQKKLKGFKDGRNLREQMTNEVFCRSRISLPKLKLENVQTKD WMQLDMLNELVRCPKSLYERLREKDRESFKVPFDIFSDDYDAE EEPFKNTLVRHQDRFPYFVLRYFDLNEIFEQLRFQIDLGTYHFSI YNKRIGDEDEVRHLTHHLYGFARIQDFAPQNQPEEWRKLVKDL DHFETSQEPYISKTAPHYHLENEKIGIKFCSTHNNLFPSLKREKT CNGRSKFNLGTQFTAEAFLSVHELLPMMFYYLLLTKDYSRKES ADKVEGIIRKEISNIYAIYDAFANNEINSIADLTCRLQKTNILQGH LPKQMISILEGRQKDMEKEAERKIGEMIDDTQRRLDLLCKQTNQ KIRIGKRNAGLLKSGKIADWLVSDMMRFQPVQKDTNNAPINNS KANSTEYRMLQHALALFGSESSRLKAYFRQMNLVGNANPHPFL AETQWEHQTNILSFYRNYLEARKKYLKGLKPQNWKQYQHFLIL KVQKTNRNTLVTGWKNSFNLPRGIFTQPIREWFEKHNNSKRIYD QILSFDRVGFVAKAIPLYFAEEYKDNVQPFYDYPFNIGNKLKPQ KGQFLDKKERVELWQKNKELFKNYPSEKNKTDLAYLDFLSWK KFERELRLIKNQDIVTWLMFKELFKTTTVEGLKIGEIHLRDIDTN

[0434]

		TANEESNNILNRIMPMKLPVKTYETDNKGNILKERPLATFYIEET ETKVLKQGNFKVLAKDRRLNGLLSFAETTDIDLEKNPITKLSVD YELIKYQTTRISIFEMTLGLEKKLIDKYSTLPTDSFRNMLERWLQ CKANRPELKNYVNSLIAVRNAFSHNQYPMYDATLFAEVKKFTL FPSVDTKKIELNIAPQLLEIVGKAIKEIEKSENKN
포르피로모나스 진지발리스 (SEQ ID No. 114)	WP_012458414	MTEQNERPYNGTYYTLEDKHFWAAFFNLARHNAYITLAHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGNNDN PFFKHHFVDREEKVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNELV RCPKSLYDRLREEDRARFRVPVDILSDEDDTDGTEEDPFKNTLV RHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGEQ PEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGDK PYITQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGATRTGRSKY AQDKRLTAEAFLSVHELMPMMFYYFLLREKYSDEASAERVQG RIKRVIEDVYAVYDAFARGEINTRDELDACLADKGIRRGHLPRQ MIGILSQEHKDMEEKVRKKLQEMIVDTDHRLDMLDRQTDRKIR IGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSKA NSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLHE TRWESHTNILSFYRSYLKARKAFLQSIGRSDRVENHRFLLLKEP KTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGLDEVGSYKE VGFMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGRFLS KEKRAEEWESGKERFRLAKLKKEILEAKEHPYLDFKSWQKFER ELRLVKNQDIITWMICRDLMEENKVEGLDTGTLYLKDIRTDVQ EQGNLNVLNRVKPMRLPVVVYRADSRGHVHKEQAPLATVYIE ERDTKLLKQGNFKSFVKDRRLNGLFSFVDTGALAMEQYPISKL RVEYELAKYQTARVCAFEQTLELEESLLTRYPHLPDKNFRKML ESWSDPLLDKWPDLHGNVRLLIAVRNAFSHNQYPMYDEAVFSS IRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKEMAERIIQA
팔루디박터	WP_013446102	MKTSANNIYFNGINSFKKIFDSKGAIAPIAEKSCRNFDIKAQNDV NKEQRIHYFAVGHTFKQLDTENLFEYVLDENLRAKRPTRFISLQ

[0435]

프로피오니시게네스(SEQ ID No. 115)	7	QFDKEFIENIKRLISDIRNINSHYIHRFDPLKIDAVPTNIIDFLKESF ELAVIQIYLKEKGINYLQFSENPHADQKLVAFLHDKFLPLDEKK TSMLQNETPQLKEYKEYRKYFKTLSKQAAIDQLLFAEKETDYI WNLFDSHPVLTISAGKYLSFYSCLFLLSMFLYKSEANQLISKIKG FKKNTTEEEKSKREIFTFFSKRFNSMDIDSEENQLVKFRDLILYL NHYPVAWNKDLELDSSNPAMTDKLKSKIIELEINRSFPLYEGNE RFATFAKYQIWGKKHLGKSIEKEYINASFTDEEITAYTYETDTCP ELKDAHKKLADLKAAKGLFGKRKEKNESDIKKTETSIRELQHEP NPIKDKLIQRIEKNLLTVSYGRNQDRFMDFSARFLAEINYFGQD ASFKMYHFYATDEQNSELEKYELPKDKKKYDSLKFHQGKLVH FISYKEHLKRYESWDDAFVIENNAIQLKLSFDGVENTVTIQRAL LIYLLEDALRNIQNNTAENAGKQLLQEYYSHNKADLSAFKQILT QQDSIEPQQKTEFKKLLPRRLLNNYSPAINHLQTPHSSLPLILEK ALLAEKRYCSLVVKAKAEGNYDDFIKRNKGKQFKLQFIRKAW NLMYFRNSYLQNVQAAGHHKSFHIERDEFNDFSRYMFAFEELS QYKYYLNEMFEKKGFFENNEFKILFQSGTSLENLYEKTKQKFEI WLASNTAKTNKPDNYHLNNYEQQFSNQLFFINLSHFINYLKSTG KLQTDANGQIIYEALNNVQYLIPEYYYTDKPERSESKSGNKLYN KLKATKLEDALLYEMAMCYLKADKQIADKAKHPITKLLTSDVE FNITNKEGIQLYHLLVPFKKIDAFIGLKMHKEQQDKKHPTSFLA NIVNYLELVKNDKDIRKTYEAFSTNPVKRTLTYDDLAKIDGHLI SKSIKFTNVTLELERYFIFKESLIVKKGNNIDFKYIKGLRNYYNN EKKKNEGIRNKAFHFGIPDSKSYDQLIRDAEVMFIANEVKPTHA TKYTDLNKQLHTVCDKLMETVHNDYFSKEGDGKKKREAAGQ KYFENIISAK
포르피로모나스 진지발리스(SEQ ID No. 116)	WP_013816155	MTEQNEKPYNGTYYTLEDKHFWAAFFNLARHNAYITLAHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGNNDN PFFKHHFVDREGTVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNELV RCPKSLYDRLREEDRARFRVPVDILSDEEDTDGAEEDPFKNTLV

[0436]

		RHQDRFPYFALRYFDLKKVFTSLRFQIDLGTYHFAIYKKNIGEQ PEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGDK PYITQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGATRTGRSKY AQDKRFTAEAFLSAHELMPMMFYYFLLREKYSEEASAERVQGR IKRVIEDVYAVYDAFARDEINTRDELDACLADKGIRRGHLPRQ MIGILSQEHKDMEEKIRKKLQEMMADTDHRLDMLDRQTDRKIR IGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSKA NSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLHE TRWESHTNILSFYRSYLKARKAFLQSIGRSDRVENHRFLLLKEP KTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGLDEVGSYKE VGFMAKAVPLYFERACKDWVQPFYNYPFNVGNSLKPKKGRFL SKEKRAEEWESGKERFRLAKLKKEILEAKEHPYLDFKSWQKFE RELRLVKNQDIITWMICGDLMEENKVEGLDTGTLYLKDIRTDV QEQGSLNVLNRVKPMRLPVVVYRADSRGHVHKEQAPLATVYI EERDTKLLKQGNFKSFVKDRRLNGLFSFVDTGALAMEQYPISK LRVEYELAKYQTARVCAFEQTLELEESLLTRCPHLPDKNFRKM LESWSDPLLDKWPDLHRKVRLLIAVRNAFSHNQYPMYDEAVFS SIRKYDPSFPDAIEERMGLNIAHRLSEEVKQAKETVERIIQA
플라보박테리움 콜럼나레 (SEQ ID No. 117)	WP_01416554 1	MSSKNESYNKQKTFNHYKQEDKYFFGGFLNNADDNLRQVGKE FKTRINFNHNNNELASVFKDYFNKEKSVAKREHALNLLSNYFP VLERIQKHTNHNFEQTREIFELLLDTIKKLRDYYTHHYHKPITIN PKIYDFLDDTLLDVLITIKKKKVKNDTSRELLKEKLRPELTQLKN QKREELIKKGKKLLEENLENAVFNHCLRPFLEENKTDDKQNKT VSLRKYRKSKPNEETSITLTQSGLVFLMSFFLHRKEFQVFTSGLE GFKAKVNTIKEEEISLNKNNIVYMITHWSYSYYNFKGLKHRIKT DQGVSTLEQNNTTHSLTNTNTKEALLTQIVDYLSKVPNEIYETL SEKQQKEFEEDINEYMRENPENEDSTFSSIVSHKVIRKRYENKFN YFAMRFLDEYAELPTLRFMVNFGDYIKDRQKKILESIQFDSERII KKEIHLFEKLSLVTEYKKNVYLKETSNIDLSRFPLFPNPSYVMA NNNIPFYIDSRSNNLDEYLNQKKKAQSQNKKRNLTFEKYNKEQ SKDAIIAMLQKEIGVKDLQQRSTIGLLSCNELPSMLYEVIVKDIK GAELENKIAQKIREQYQSIRDFTLDSPQKDNIPTTLIKTINTDSSV TFENQPIDIPRLKNAIQKELTLTQEKLLNVKEHEIEVDNYNRNKN

[0437]

		TYKFKNQPKNKVDDKKLQRKYVFYRNEIRQEANWLASDLIHF MKNKSLWKGYMHNELQSFLAFFEDKKNDCIALLETVFNLKED CILTKGLKNLFLKHGNFIDFYKEYLKLKEDFLNTESTFLENGLIG LPPKILKKELSKRFKYIFIVFQKRQFIIKELEEKKNNLYADAINLS RGIFDEKPTMIPFKKPNPDEFASWFVASYQYNNYQSFYELTPDI VERDKKKKYKNLRAINKVKIQDYYLKLMVDTLYQDLFNQPLD KSLSDFYVSKAEREKIKADAKAYQKRNDSSLWNKVIHLSLQNN RITANPKLKDIGKYKRALQDEKIATLLTYDDRTWTYALQKPEK ENENDYKELHYTALNMELQEYEKVRSKELLKQVQELEKQILEE YTDFLSTQIHPADFEREGNPNFKKYLAHSILENEDDLDKLPEKV EAMRELDETITNPIIKKAIVLIIIRNKMAHNQYPPKFIYDLANRFV PKKEEEYFATYFNRVFETITKELWENKEKKDKTQV
사이크로플렉서스 토르쿠스 (SEQ ID No. 118)	WP_01502476 5	MESIIGLGLSFNPYKTADKHYFGSFLNLVENNLNAVFAEFKERIS YKAKDENISSLIEKHFIDNMSIVDYEKKISILNGYLPIIDFLDDELE NNLNTRVKNFKKNFIILAEAIEKLRDYYTHFYHDPITFEDNKEPL LELLDEVLLKTILDVKKKYLKTDKTKEILKDSLREEMDLLVIRK TDELREKKKTNPKIQHTDSSQIKNSIFNDAFQGLLYEDKGNNKK TQVSHRAKTRLNPKDIHKQEERDFEIPLSTSGLVFLMSLFLSKKE IEDFKSNIKGFKGKVVKDENHNSLKYMATHRVYSILAFKGLKY RIKTDTFSKETLMMQMIDELSKVPDCVYQNLSETKQKDFIEDW NEYFKDNEENTENLENSRVVHPVIRKRYEDKFNYFAIRFLDEFA NFKTLKFQVFMGYYIHDQRTKTIGTTNITTERTVKEKINVFGKL SKMDNLKKHFFSQLSDDENTDWEFFPNPSYNFLTQADNSPANN IPIYLELKNQQIIKEKDAIKAEVNQTQNRNPNKPSKRDLLNKILK TYEDFHQGDPTAILSLNEIPALLHLFLVKPNNKTGQQIENIIRIKIE KQFKAINHPSKNNKGIPKSLFADTNVRVNAIKLKKDLEAELDM LNKKHIAFKENQKASSNYDKLLKEHQFTPKNKRPELRKYVFYK SEKGEEATWLANDIKRFMPKDFKTKWKGCQHSELQRKLAFYD RHTKQDIKELLSGCEFDHSLLDINAYFQKDNFEDFFSKYLENRIE TLEGVLKKLHDFKNEPTPLKGVFKNCFKFLKRQNYVTESPEIIK KRILAKPTFLPRGVFDERPTMKKGKNPLKDKNEFAEWFVEYLE NKDYQKFYNAEEYRMRDADFKKNAVIKKQKLKDFYTLQMVN YLLKEVFGKDEMNLQLSELFQTRQERLKLQGIAKKQMNKETG

[0438]

		DSSENTRNQTYIWNKDVPVSFFNGKVTIDKVKLKNIGKYKRYE RDERVKTFIGYEVDEKWMMYLPHNWKDRYSVKPINVIDLQIQE YEEIRSHELLKEIQNLEQYIYDHTTDKNILLQDGNPNFKMYVLN GLLIGIKQVNIPDFIVLKQNTNFDKIDFTGIASCSELEKKTIILIAIR NKFAHNQLPNKMIYDLANEFLKIEKNETYANYYLKVLKKMISD LA
리에메렐라 아나티페스티퍼 (SEQ ID No. 119)	WP_01534562 0	MFFSFHNAQRVIFKHLYKAFDASLRMVKEDYKAHFTVNLTRDF AHLNRKGKNKQDNPDFNRYRFEKDGFFTESGLLFFTNLFLDKR DAYWMLKKVSGFKASHKQREKMTTEVFCRSRILLPKLRLESRY DHNQMLLDMLSELSRCPKLLYEKLSEENKKHFQVEADGFLDEI EEEQNPFKDTLIRHQDRFPYFALRYLDLNESFKSIRFQVDLGTYH YCIYDKKIGDEQEKRHLTRTLLSFGRLQDFTEINRPQEWKALTK DLDYKETSNQPFISKTTPHYHITDNKIGFRLGTSKELYPSLEIKDG ANRIAKYPYNSGFVAHAFISVHELLPLMFYQHLTGKSEDLLKET VRHIQRIYKDFEEERINTIEDLEKANQGRLPLGAFPKQMLGLLQ NKQPDLSEKAKIKIEKLIAETKLLSHRLNTKLKSSPKLGKRREKL IKTGVLADWLVKDFMRFQPVAYDAQNQPIKSSKANSTEFWFIR RALALYGGEKNRLEGYFKQTNLIGNTNPHPFLNKFNWKACRNL VDFYQQYLEQREKFLEAIKHQPWEPYQYCLLLKVPKENRKNLV KGWEQGGISLPRGLFTEAIRETLSKDLTLSKPIRKEIKKHGRVGFI SRAITLYFKEKYQDKHQSFYNLSYKLEAKAPLLKKEEHYEYWQ QNKPQSPTESQRLELHTSDRWKDYLLYKRWQHLEKKLRLYRN QDIMLWLMTLELTKNHFKELNLNYHQLKLENLAVNVQEADAK LNPLNQTLPMVLPVKVYPTTAFGEVQYHETPIRTVYIREEQTKA LKMGNFKALVKDRRLNGLFSFIKEENDTQKHPISQLRLRRELEI YQSLRVDAFKETLSLEEKLLNKHASLSSLENEFRTLLEEWKKKY AASSMVTDKHIAFIASVRNAFCHNQYPFYKETLHAPILLFTVAQ PTTEEKDGLGIAEALLKVLREYCEIVKSQI
프레보텔라 플레우리티디스 (SEQ ID No. 120)	WP_02158463 5	MENDKRLEESACYTLNDKHFWAAFLNLARHNVYITVNHINKTL ELKNKKNQEIIIDNDQDILAIKTHWAKVNGDLNKTDRLRELMIK HFPFLEAAIYSNNKEDKEEVKEEKQAKAQSFKSLKDCLFLFLEK LQEARNYYSHYKYSESSKEPEFEEGLLEKMYNTFDASIRLVKED YQYNKDIDPEKDFKHLERKEDFNYLFTDKDNKGKITKNGLLFF

[0439]

		VSLFLEKKDAIWMQQKFRGFKDNRGNKEKMTHEVFCRSRMLL PKIRLESTQTQDWILLDMLNELIRCPKSLYERLQGAYREKFKVP FDSIDEDYDAEQEPFRNTLVRHQDRFPYFALRYFDYNEIFKNLR FQIDLGTYHFSIYKKLIGGKKEDRHLTHKLYGFERIQEFTKQNRP DKWQAIIKDLDTYETSNERYISETTPHYHLENQKIGIRFRNDNN DIWPSLKTNGEKNEKSKYNLDKPYQAEAFLSVHELLPMMFYYL LLKMENTDNDKEDNEVGTKKKGNKNNKQEKHKIEEIIENKIKDI YALYDAFTNGEINSIDELAEQREGKDIEIGHLPKQLIVILKNKSK DMAEKANRKQKEMIKDTKKRLATLDKQVKGEIEDGGRNIRLL KSGEIARWLVNDMMRFQPVQKDNEGKPLNNSKANSTEYQMLQ RSLALYNKEEKPTRYFRQVNLIKSSNPHPFLEDTKWEECYNILSF YRNYLKAKIKFLNKLKPEDWKKNQYFLMLKEPKTNRKTLVQG WKNGFNLPRGIFTEPIKEWFKRHQNDSEEYKKVEALDRVGLVA KVIPLFFKEEYFKEDAQKEINNCVQPFYSFPYNVGNIHKPEEKNF LHCEERRKLWDKKKDKFKGYKAKEKSKKMTDKEKEEHRSYLE FQSWNKFERELRLVRNQDILTWLLCTKLIDKLKIDELNIEELQKL RLKDIDTDTAKKEKNNILNRVMPMRLPVTVYEIDKSFNIVKDKP LHTVYIEETGTKLLKQGNFKALVKDRRLNGLFSFVKTSSEAESK SKPISKLRVEYELGAYQKARIDIIKDMLALEKTLIDNDENLPTNK FSDMLKSWLKGKGEANKARLQNDVGLLVAVRNAFSHNQYPM YNSEVFKGMKLLSLSSDIPEKEGLGIAKQLKDKIKETIERIIEIEKE IRN
포르피로모나스 진지발리스 (SEQ ID No. 121)	WP_021663197	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPRSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANR

[0440]

		ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMDQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLQKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRHQFRAIV AELRLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
포르피로모나스 진지발리스 (SEQ ID No. 122)	WP_02166547 5	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTNENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHDLRKLLLMELLCEGSFSRMQSGFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW

[0441]

		WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDKENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDHENRFFGKLLNNMSQPINDL
포르피로모나스 진지발리스 (SEQ ID No. 123)	WP_02167765 7	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHDLRKLLLMELLCEGSFSRMQSGFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE

[0442]

		GSSLVDSLWKKYEMIIRKILPILDHENRFFGKLLNNMSQPINDL
포르피로모나스 진지발리스 (SEQ ID No. 124)	WP_021680012	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIKFGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYFDPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRLDGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLAKNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGFKRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNELNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLLWDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVDLGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSRMISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEKRALSNPRSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANRILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKYKQEIKGRKDKLNSQLLSAFDMDQRQLPSRLLDEWMNIRPASHSVKLRTYVKQLNEDCRLRLQKFRKDGDGKARAIPLVGEMATFLSQDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRHQFRAIVAELRLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAKTFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQDWIRNKQAHPIDLPSHLFDSKVMELLKVKDGKKKWNEAFKDWWSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHLMEKTVRDKKRELRTAGKPVPPDLAAYIKRSFHRAVNEREFMLRLVQEDDRLMLMAINKIMTDREEDILPGLKNIDSILDKENQFSLAVHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDRRVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAIMSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQYLILIRNKAAHNQFPCAAEIPLIYRDVSAKVGSIEGSSAKDLPEGSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
포르피로모나스 진지발리스 (SEQ ID No. 125)	WP_021680012	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIKFGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYFDPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRLDGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLAKNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGFKRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE

[0443]

		LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPRSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
프레보텔라 팔세니 (SEQ ID No. 126)	WP_036884929	MKNDNNSTKSTDYTLGDKHFWAAFLNLARHNVYITVNHINKV LELKNKKDQEIIIDNDQDILAIKTLWGKVDTDINKKDRLRELIM KHFPFLEAATYQQSSTNNTKQKEEEQAKAQSFESLKDCLFLFLE KLREARNYYSHYKHSKSLEEPKLEEKLLENMYNIFDTNVQLVIK DYEHNKDINPEEDFKHLGRAEGEFNYYFTRNKKGNITESGLLFF VSLFLEKKDAIWAQTKIKGFKDNRENKQKMTHEVFCRSRMLLP KLRLESTQTQDWILLDMLNELIRCPKSLYKRLQGEKREKFRVPF DPADEDYDAEQEPFKNTLVRHQDRFPYFALRYFDYNEIFTNLRF QIDLGTYHFSIYKKQIGDKKEDRHLTHKLYGFERIQEFAKENRP DEWKALVKDLDTFEESNEPYISETTPHYHLENQKIGIRNKNKKK KKTIWPSLETKTTVNERSKYNLGKSFKAEAFLSVHELLPMMFY YLLLNKEEPNNGKINASKVEGIIEKKIRDIYKLYGAFANEEINNE EELKEYCEGKDIAIRHLPKQMIAILKNEYKDMAKKAEDKQKKM

[0444]

		IKDTKKRLAALDKQVKGEVEDGGRNIKPLKSGRIASWLVNDM MRFQPVQRDRDGYPLNNSKANSTEYQLLQRTLALFGSERERLA PYFRQMNLIGKDNPHPFLKDTKWKEHNNILSFYRSYLEAKKNF LGSLKPEDWKKNQYFLKLKEPKTNRETLVQGWKNGFNLPRGIF TEPIREWFIRHQNESEEYKKVKDFDRIGLVAKVIPLFFKEDYQKE IEDYVQPFYGYPFNVGNIHNSQEGTFLNKKEREELWKGNKTKF KDYKTKEKNKEKTNKDKFKKKTDEEKEEFRSYLDFQSWKKFE RELRLVRNQDIVTWLLCMELIDKLKIDELNIEELQKLRLKDIDTD TAKKEKNNILNRIMPMELPVTVYETDDSNNIIKDKPLHTIYIKEA ETKLLKQGNFKALVKDRRLNGLFSFVETSSEAELKSKPISKSLVE YELGEYQRARVEIIKDMLRLEETLIGNDEKLPTNKFRQMLDKW LEHKKETDDTDLKNDVKLLTEVRNAFSHNQYPMRDRIAFANIK PFSLSSANTSNEEGLGIAKKLKDKTKETIDRIIEIEEQTATKR
프레보텔라 플레우리티디스 (SEQ ID No. 127)	WP_03693148 5	MENDKRLEESTCYTLNDKHFWAAFLNLARHNVYITINHINKLL EIRQIDNDEKVLDIKALWQKVDKDINQKARLRELMIKHFPFLEA AIYSNNKEDKEEVKEEKQAKAQSFKSLKDCLFLFLEKLQEARN YYSHYKSSESSKEPEFEEGLLEKMYNTFGVSIRLVKEDYQYNKD IDPEKDFKHLERKEDFNYLFTDKDNKGKITKNGLLFFVSLFLEK KDAIWMQQKLRGFKDNRGNKEKMTHEVFCRSRMLLPKIRLES TQTQDWILLDMLNELIRCPKSLYERLQGAYREKFKVPFDSIDED YDAEQEPFRNTLVRHQDRFPYFALRYFDYNEIFKNLRFQIDLGT YHFSIYKKLIGDNKEDRHLTHKLYGFERIQEFAKQKRPNEWQA LVKDLDIYETSNEQYISETTPHYHLENQKIGIRFKNKKDKIWPSL ETNGKENEKSKYNLDKSFQAEAFLSIHELLPMMFYDLLLKKEEP NNDEKNASIVEGFIKKEIKRMYAIYDAFANEEINSKEGLEEYCK NKGFQERHLPKQMIAILTNKSKNMAEKAKRKQKEMIKDTKKR LATLDKQVKGEIEDGGRNIRLLKSGEIARWLVNDMMRFQSVQK DKEGKPLNNSKANSTEYQMLQRSLALYNKEQKPTPYFIQVNLI KSSNPHPFLEETKWEECNNILSFYRSYLEAKKNFLESLKPEDWK KNQYFLMLKEPKTNRKTLVQGWKNGFNLPRGIFTEPIKEWFKR HQNDSEEYKKVEALDRVGLVAKVIPLFFKEEYFKEDAQKEINN CVQPFYSFPYNVGNIHKPEEKNFLHCEERRKLWDKKKDKFKGY KAKEKSKKMTDKEKEEHRSYLEFQSWNKFERELRLVRNQDIVT

[0445]

		WLLCTELIDKLKIDELNIEELQKLRLKDIDTDTAKKEKNNILNRI MPMQLPVTVYEIDKSFNIVKDKPLHTIYIEETGTKLLKQGNFKA LVKDRRLNGLFSFVKTSSEAESKSKPISKLRVEYELGAYQKARI DIIKDMLALEKTLIDNDENLPTNKFSDMLKSWLKGKGEANKAR LQNDVDLLVAIRNAFSHNQYPMYNSEVFKGMKLLSLSSDIPEKE GLGIAKQLKDKIKETIERIIEIEKEIRN
[포르피로모나스 진지발리스 (SEQ ID No. 128)	WP_03941739 0	MTEQNERPYNGTYYTLEDKHFWAAFFNLARHNAYITLAHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGNNDN PFFKHHFVDREGTVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTEAYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNELV RCPKSLYDRLREEDRARFRVPIDILSDEDDTDGTEEDPFKNTLVR HQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGEQPE DRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGDKPY ITQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGATRTGRSKYAQ DKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQGRIK RVIEDVYAVYDAFARGEIDTLDRLDACLADKGIRRGHLPRQMI AILSQEHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKIRIG RKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSKANS TEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLHETR WESHTNILSFYRSYLKARKAFLQSIGRSDREENHRFLLLKEPKT DRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYKEVG FMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGRFLSKE KRAEEWESGKERFRLAKLKKEILEAKEHPYLDFKSWQKFEREL RLVKNQDIITWMMCRDLMEENKVEGLDTGTLYLKDIRTDVHE QGSLNVLNRVKPMRLPVVVYRADSRGHVHKEQAPLATVYIEE RDTKLLKQGNFKSFVKDRRLNGLFSFVDTGALAMEQYPISKLR VEYELAKYQTARVCAFEQTLELEESLLTRYPHLPDKNFRKMLE SWSDPLLDKWPDLHRKVRLLIAVRNAFSHNQYPMYDEAVFSSI RKYDPSSPDAIEERMGLNIAHRLSEEVKQAKEMAERIIQV
포르피로	WP_03	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ

[0446]

모나스 굴라에 (SEQ ID No. 129)	9418912	LAYSKADITNDQDVLSFKALWKNLDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSGSSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGHNDN PSFKHHFVDSEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTETYQQMTNEVFCRSRISLPKLKLESLRMDDWMLLDMLNE LVRCPKPLYDRLREDDRACFRVPVDILPDEDDTDGGGEDPFKN TLVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMI GEQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFET GDKPYISQTSPHYHIEKGKIGLRFMPEGQHLWPSPEVGTTRTGR SKYAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKV QGRIKRVIEDVYAIYDAFARDEINTLKELDACLADKGIRRGHLP KQMIAILSQEHKNMEEKVRKKLQEMIADTDHRLDMLDRQTDR KIRIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDASGKPLNNS KANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFL HDTRWESHTNILSFYRSYLRARKAFLERIGRSDRMENRPFLLLK EPKTDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSY REVGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRF LSKEERAEEWERGKERFRDLEAWSHSAARRIEDAFAGIEYASPG NKKKIEQLLRDLSLWEAFESKLKVRADKINLAKLKKEILEAQEH PYHDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLD TGTLYLKDIRTNVQEQGSLNVLNHVKPMRLPVVVYRADSRGH VHKEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDT GGLAMEQYPISKLRVEYELAKYQTARVCAFEQTLELEESLLTRY PHLPDKNFRKMLESWSDPLLAKWPELHGKVRLLIAVRNAFSHN QYPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAK ETVERIIQA
포르피로 모나스 굴라에 (SEQ ID No. 130)	WP_03941979 2	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNLDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSGSSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGHNDN PSFKHHFVDGEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK

[0447]

		GGTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREKDRARFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKVIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGTTRTGRSK YAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQ GRIKRVIEDVYAIYDAFARDEINTRDELDACLADKGIRRGHLPK QMIGILSQEHKNMEEKVRKKLQEMIADTDHRLDMLDRQTDRKI RIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLD ETRWESHTNILSFYRSYLRARKAFLERIGRSDRVENRPFLLLKEP KTDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYKE VGFMAKAVPLYFERACKDRVQPFYDSPFNVGNSLKPKKGRFLS KEKRAEEWESGKERFRLAKLKKEILEAQEHPYHDFKSWQKFER ELRLVKNQDIITWMMCRDLMEENKVEGLDTGTLYLKDIRPNVQ EQGSLNVLNRVKPMRLPVVVYRADSRGHVHKEEAPLATVYIEE RDTKLLKQGNFKSFVKDRRLNGLFSFVDTGGLAMEQYPISKLR VEYELAKYQTARVCVFELTLRLEESLLSRYPHLPDESFREMLES WSDPLLAKWPELHGKVRLLIAVRNAFSHNQYPMYDEAVFSSIR KYDPSSPDAIEERMGLNIAHRLSEEVKQAKETVERIIQA
포르피로모나스 굴라에 (SEQ ID No. 131)	WP_03942617 6	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNFDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSGSSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHYHFNHLVRKGKKDRYGHNDN PSFKHHFVDSEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTGPYEQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREKDRACFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTTPHYHIEKGKIGLRFMPEGQHLWPSPEVGTTRTGRS KYAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKV QGRIKRVIKDVYAIYDAFARDEINTLKELDACSADKGIRRGHLP

[0448]

		KQMIGILSQEHKNMEEKVRKKLQEMIADTDHRLDMLDRQTDR KIRIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNS KANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFL DETRWESHTNILSFYRSYLRARKAFLERIGRSDRVENRPFLLLKE PKNDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYK EVGFMAKAVPLYFERACKDRVQPFYDSPFNVGNSLKPKKGRFL SKEKRAEEWESGKERFRLAKLKKEILEAKEHPYHDFKSWQKFE RELRLVKNQDIITWMMCRDLMEENKVEGLDTGTLYLKDIRTDV HEQGSLNVLNRVKPMRLPVVVYRADSRGHVHKEQAPLATVYI EERDTKLLKQGNFKSFVKDRRLNGLFSFVDTGGLAMEQYPISK LRVEYELAKYQTARVCAFEQTLELEESLLTRYPHLPDENFREML ESWSDPLLGKWPDLHGKVRLLIAVRNAFSHNQYPMYDEAVFSS IRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKETVERIIQA
포르피로모나스 굴라에 (SEQ ID No. 132)	WP_03943177 8	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNFDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGHNDN PSFKHHFVDGEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREDDRACFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTSPHYHIEKGKIGLRFMPEGQHLWPSPEVGTTRTGRS KYAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKV QGRIKRVIEDVYAIYDAFARDEINTLKELDACLADKGIRRGHLP KQMIAILSQEHKDMEEKIRKKLQEMIADTDHRLDMLDRQTDRK IRIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKKRLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLRARKAFLERIGRSDRMENRPFLLLKEP KTDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYRE VGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRFLS KEERAEEWERGKERFRDLEAWSHSAARRIEDAFAGIEYASPGN

[0449]

		KKKIEQLLRDLSLWEAFESKLKVRADKINLAKLKKEILEAQEHP YHDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDT GTLYLKDIRPNVQEQGSLNVLNRVKPMRLPVVVYRADSRGHV HKEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDTG GLAMEQYPISKLRVEYELAKYQTARVCVFELTLRLEESLLTRYP HLPDESFRKMLESWSDPLLAKWPELHGKVRLLIAVRNAFSHNQ YPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKE TVERIIQV
포르피로모나스 콜라에 (SEQ ID No. 133)	WP_03943719 9	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKSRLRSLILKHFSF LEGAAYGKKFFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSGSSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDEVDPHYHFNHLVRKGKKDRYGHNDN PSFKHHFVDGEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTEPYEQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREKDRACFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGTTRTGRSK YAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQ GRIKRVIEDVYAIYDAFARDEINTLKELDACLADKGIRRGHLPK QMIGILSQERKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKI RIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLRARKAFLERIGRSDRVENCPFLLLKEP KTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGYDEVGSYRE VGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRFLS KEKRAEEWESGKERFRLAKLKKEILEAQEHPYHDFKSWQKFER ELRLVKNQDIITWMMCRDLMEENKVEGLDTGTLYLKDIRPNVQ EQGSLNVLNRVKPMRLPVVVYRADSRGHVHKEEAPLATVYIEE RDTKLLKQGNFKSFVKDRRLNGLFSFVDTGALAMEQYPISKLR VEYELAKYQTARVCAFEQTLELEESLLTRYPHLPDESFREMLES WSDPLLTKWPELHGKVRLLIAVRNAFSHNQYPMYDEAVFSSIW

[0450]

		KYDPSSPDAIEERMGLNIAHRLSEEVKQAKETIERIIQA
포르피로모나스 굴라에 (SEQ ID No. 134)	WP_03944217 1	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNLDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSGSSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHYHFNHLVRKGKKDRYGHNDN PSFKHHFVDSEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTGPYEQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREKDRACFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYLETG DKPYISQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGTTRTGRSK CAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQ GRIKRVIEDVYAIYDAFARDEINTLKELDTCLADKGIRRGHLPK QMITILSQERKDMKEKIRKKLQEMIADTDHRLDMLDRQTDRKI RIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDASGKPLNNSK ANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLRARKAFLERIGRSDRVENCPFLLLKEP KTDRQTLVAGWKDEFHLPRGIFTEAVRDCLIEMGYDEVGSYRE VGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRFLS KEDRAEEWERGMERFRDLEAWSHSAARRIKDAFAGIEYASPGN KKKIEQLLRDLSLWEAFESKLKVRADKINLAKLKKEILEAQEHP YHDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDT GTLYLKDIRPNVQEQGSLNVLNRVKPMRLPVVVYRADSRGHV HKEAPLATVYIEERNTKLLKQGNFKSFVKDRRLNGLFSFVDTG GLAMEQYPISKLRVEYELAKYQTARVCVFELTLRLEESLLSRYP HLPDESFREMLESWSDPLLAKWPELHGKVRLLIAVRNAFSHNQ YPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKE TVERIIQA
포르피로모나스 굴라에 (SEQ ID No. 135)	WP_03944505 5	MNTVPATENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRI KFGKKKLNEESLKQSLLCDHLLSIDRWTKVYGHSRRYLPFLHCF DPDSGIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLKVSPDISSFITGAYTFACERAQSRFADFFKPDDFLLA

[0451]

		KNRKEQLISVADGKECLTVSGFAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHDLRKLLLMELLCEGSFSRMQSDFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVRDKKRELRTAGKPVPPDLAAYIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDEENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDHENRFFGKLLNNMSQPINDL
카프노사이토파가 시노데그미 (SEQ ID No. 136)	WP_041989581	MENKTSLGNNIYYNPFKPQDKSYFAGYLNAAMENIDSVFRELG KRLKGKEYTSENFFDAIFKENISLVEYERYVKLLSDYFPMARLL DKKEVPIKERKENFKKNFRGIIKAVRDLRNFYTHKEHGEVEITD EIFGVLDEMLKSTVLTVKKKKIKTDKTKEILKKSIEKQLDILCQK KLEYLKDTARKIEEKRRNQRERGEKKLVPRFEYSDRRDDLIAAI YNDAFDVYIDKKKDSLKESSKTKYNTESYPQQEEGDLKIPISKN GVVFLLSLFLSKQEVHAFKSKIAGFKATVIDEATVSHRKNSICF MATHEIFSHLAYKKLKRKVRTAEINYSEAENAEQLSIYAKETLM MQMLDELSKVPDVVYQNLSEDVQKTFIEDWNEYLKENNGDVG TMEEEQVIHPVIRKRYEDKFNYFAIRFLDEFAQFPTLRFQVHLG NYLHDSRPKEHLISDRRIKEKITVFGRLSELEHKKALFIKNTETN

[0452]

		EDRKHYWEVFPNPNYDFPKENISVNDKDFPIAGSILDREKQPTA GKIGIKVNLLNQKYISEVDKAVKAHQLKQRNNKPSIQNIIEEIVPI NGSNPKEIIVFGGQPTAYLSMNDIHSILYEFFDKWEKKKEKLEK KGEKELRKEIGKELEEKIVGKIQTQIQQIIDKDINAKILKPYQDDD STAIDKEKLIKDLKQEQKILQKLKNEQTAREKEYQECIAYQEES RKIKRSDKSRQKYLRNQLKRKYPEVPTRKEILYYQEKGKVAVW LANDIKRFMPTDFKNEWKGEQHSLLQKSLAYYEQCKEELKNLL PQQKVFKHLPFELGGHFQQKYLYQFYTRYLDKRLEHISGLVQQ AENFKNENKVFKKVENECFKFLKKQNYTHKGLDAQAQSVLGY PIFLERGFMDEKPTIIKGKTFKGNESLFTDWFRYYKEYQNFQTF YDTENYPLVELEKKQADRKRETKIYQQKKNDVFTLLMAKHIFK SVFKQDSIDRFSLEDLYQSREERLENQEKAKQTGERNTNYIWNK TVDLNLCDGKVTVENVKLKNVGNFIKYEYDQRVQTFLKYEENI KWQAFLIKESKEEENYPYIVEREIEQYEKVRREELLKEVHLIEEY ILEKVKDKEILKKGDNQNFKYYILNGLLKQLKNEDVESYKVFN LNTKPEDVNINQLKQEATDLEQKAFVLTYIRNKFAHNQLPKKEF WDYCQEKYGKIEKEKTYAEYFAEVFKREKEALMK
프레보텔라 sp. P5-119 (SEQ ID No. 137)	WP_042518169	MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQ NENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIERLQSYFPF LKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMY RDLTNHYKTYEEKLIDGCEFLTSTEQPLSGMISKYYTVALRNTK ERYGYKTEDLAFIQDNIKKITKDAYGKRKSQVNTGFFLSLQDYN GDTQKKLHLSGVGIALLICLFLDKQYINIFLSRLPIFSSYNAQSEE RRIIIRSFGINSIKLPKDRIHSEKSNKSVAMDMLNEVKRCPDELFT TLSAEKQSRFRIISDDHNEVLMKRSTDRFVPLLLQYIDYGKLFD HIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFGRLEE AETMRKQENGTFGNSGIRIRDFENVKRDDANPANYPYIVDTYT HYILENNKVEMFISDKGSSAPLLPLIEDDRYVVKTIPSCRMSTLEI PAMAFHMFLFGSKKTEKLIVDVHNRYKRLFQAMQKEEVTAENI ASFGIAESDLPQKILDLISGNAHGKDVDAFIRLTVDDMLTDTER RIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLFQ PSVNDGENKITGLNYRIMQSAIAVYDSGDDYEAKQQFKLMFEK ARLIGKGTTEPHPFLYKVFARSIPANAVDFYERYLIERKFYLTGL

[0453]

		CNEIKRGNRVDVPFIRRDQNKWKTPAMKTLGRIYSEDLPVELPR QMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEYMKRVLNDDF QTFYQWKRNYHYMDMLKGEYDRKGSLQHCFTSVEEREGLWK ERASRTERYRKLASNKIRSNRQMRNASSEEIETILDKRLSNCRNE YQKSEKVIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEIMP DAEKGILSEIMPMSFTFEKGGKKYTITSEGMKLKNYGDFFVLAS DKRIGNLLELVGSDIVSKEDIMEEFNKYDQCRPEISSIVFNLEKW AFDTYPELSARVDREEKVDFKSILKILLNNKNINKEQSDILRKIR NAFDHNNYPDKGIVEIKALPEIAMSIKKAFGEYAIMK
프레보텔라 sp. P4-76 (SEQ ID No. 138)	WP_04 407214 7	MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQ NENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIERLQSYFPF LKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMY RDQASHYKTYDEKLIDGCEFLTSTEQPLSGMINNYYTVALRNM NERYGYKTEDLAFIQDKRFKFVKDAYGKKKSQVNTGFFLSLQD YNGDTQKKLHLSGVGIALLICLFLDKQYINIFLSRLPIFSSYNAQS EERRIIIRSFGINSIKQPKDRIHSEKSNKSVAMDMLNEIKRCPNEL FETLSAEKQSRFRIISNDHNEVLMKRSSDRFVPLLLQYIDYGKLF DHIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFGRLE EVETMRKQENGTFGNSGIRIRDFENMKRDDANPANYPYIVDTY THYILENNKVEMFISDEETPAPLLPVIEDDRYVVKTIPSCRMSTL EIPAMAFHMFLFGSKKTEKLIVDVHNRYKRLFKAMQKEEVTAE NIASFGIAESDLPQKIIDLISGNAHGKDVDAFIRLTVDDMLADTE RRIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLF QPSVNDGENKITGLNYRIMQSAIAVYNSGDDYEAKQQFKLMFE KARLIGKGTTEPHPFLYKVFVRSIPANAVDFYERYLIERKFYLIG LSNEIKKGNRVDVPFIRRDQNKWKTPAMKTLGRIYDEDLPVELP RQMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEYMKRVLNDD FQTFYQWKRNYRYMDMLRGEYDRKGSLQSCFTSVEEREGLWK ERASRTERYRKLASNKIRSNRQMRNASSEEIETILDKRLSNSRNE YQKSEKVIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEIMP DAEKGILSEIMPMSFTFEKGGKKYTITSEGMKLKNYGDFFVLAS DKRIGNLLELVGSDTVSKEDIMEEFKKYDQCRPEISSIVFNLEKW AFDTYPELSARVDREEKVDFKSILKILLNNKNINKEQSDILRKIR

[0454]

		NAFDHNNYPDKGVVEIRALPEIAMSIKKAFGEYAIMK
프레보텔라 sp. P5-60 (SEQ ID No. 139)	WP_04 407478 0	MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQ NENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIERLQSYFPF LKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMY RDLTNHYKTYEEKLIDGCEFLTSTEQPFSGMISKYYTVALRNTK ERYGYKAEDLAFIQDNRYKFTKDAYGKRKSQVNTGSFLSLQDY NGDTTKKLHLSGVGIALLICLFLDKQYINLFLSRLPIFSSYNAQSE ERRIIIRSFGINSIKQPKDRIHSEKSNKSVAMDMLNEVKRCPDELF TTLSAEKQSRFRIISDDHNEVLMKRSSDRFVPLLLQYIDYGKLFD HIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFGRLEE VETMRKQENGTFGNSGIRIRDFENMKRDDANPANYPYIVETYT HYILENNKVEMFISDEENPTPLLPVIEDDRYVVKTIPSCRMSTLEI PAMAFHMFLFGSEKTEKLIIDVHDRYKRLFQAMQKEEVTAENI ASFGIAESDLPQKIMDLISGNAHGKDVDAFIRLTVDDMLTDTER RIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLFQ PSVNDGENKITGLNYRIMQSAIAVYDSGDDYEAKQQFKLMFEK ARLIGKGTTEPHPFLYKVFVRSIPANAVDFYERYLIERKFYLIGL SNEIKKGNRVDVPFIRRDQNKWKTPAMKTLGRIYSEDLPVELPR QMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEYMKRVLNDDF QTFYQWKRNYRYMDMLRGEYDRKGSLQHCFTSIEEREGLWKE RASRTERYRKLASNKIRSNRQMRNASSEEIETILDKRLSNCRNE YQKSEKIIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEIMP DAEKGILSEIMPMSFTFEKGGKIYTITSGGMKLKNYGDFFVLAS DKRIGNLLELVGSNTVSKEDIMEEFKKYDQCRPEISSIVFNLEKW AFDTYPELPARVDRKEKVDFWSILDVLSNNKDINNEQSYILRKI RNAFDHNNYPDKGIVEIKALPEIAMSIKKAFGEYAIMK
파에오닥틸리박터 시아메넨시스 (SEQ ID No. 140)	WP_04 421823 9	MTNTPKRRTLHRHPSYFGAFLNIARHNAFMIMEHLSTKYDMED KNTLDEAQLPNAKLFGCLKKRYGKPDVTEGVSRDLRRYFPFLN YPLFLHLEKQQNAEQAATYDINPEDIEFTLKGFFRLLNQMRNNY SHYISNTDYGKFDKLPVQDIYEAAIFRLLDRGKHTKRFDVFESK HTRHLESNNSEYRPRSLANSPDHENTVAFVTCLFLERKYAFPFL SRLDCFRSTNDAAEGDPLIRKASHECYTMFCCRLPQPKLESSDIL LDMVNELGRCPSALYNLLSEEDQARFHIKREEITGFEEDPDEELE

[0455]

		QEIVLKRHSDRFPYFALRYFDDTEAFQTLRFDVYLGRWRTKPV YKKRIYGQERDRVLTQSIRTFTRLSRLLPIYENVKHDAVRQNEE DGKLVNPDVTSQFHKSWIQIESDDRAFLSDRIEHFSPHYNFGDQ VIGLKFINPDRYAAIQNVFPKLPGEEKKDKDAKLVNETADAIIST HEIRSLFLYHYLSKKPISAGDERRFIQVDTETFIKQYIDTIKLFFED IKSGELQPIADPPNYQKNEPLPYVRGDKEKTQEERAQYRERQKE IKERRKELNTLLQNRYGLSIQYIPSRLREYLLGYKKVPYEKLAL QKLRAQRKEVKKRIKDIEKMRTPRVGEQATWLAEDIVFLTPPK MHTPERKTTKHPQKLNNDQFRIMQSSLAYFSVNKKAIKKFFQK ETGIGLSNRETSHPFLYRIDVGRCRGILDFYTGYLKYKMDWLDD AIKKVDNRKHGKKEAKKYEKYLPSSIQHKTPLELDYTRLPVYLP RGLFKKAIVKALAAHADFQVEPEEDNVIFCLDQLLDGDTQDFY NWQRYYRSALTEKETDNQLVLAHPYAEQILGTIKTLEGKQKNN KLGNKAKQKIKDELIDLKRAKRRLLDREQYLRAVQAEDRALW LMIQERQKQKAEHEEIAFDQLDLKNITKILTESIDARLRIPDTKV DITDKLPLRRYGDLRRVAKDRRLVNLASYYHVAGLSEIPYDLV KKELEEYDRRRVAFFEHVYQFEKEVYDRYAAELRNENPKGEST YFSHWEYVAVAVKHSADTHFNELFKEKVMQLRNKFHHNEFPY FDWLLPEVEKASAALYADRVFDVAEGYYQKMRKLMRQ
플라보박테리움 sp. 316 (SEQ ID No. 141)	WP_04596837 7	MDNNITVEKTELGLGITYNHDKVEDKHYFGGFFNLAQNNIDLV AQEFKKRLLIQGKDSINIFANYFSDQCSITNLERGIKILAEYFPVV SYIDLDEKNKSKSIREHLILLLETINNLRNYYTHYYHKKIIIDGSL FPLLDTILLKVVLEIKKKKLKEDKTKQLLKKGLEKEMTILFNLM KAEQKEKKIKGWNIDENIKGAVLNRAFSHLLYNDELSDYRKSK YNTEDETLKDTLTESGILFLLSFFLNKKEQEQLKANIKGYKGKIA SIPDEEITLKNNSLRNMATHWTYSHLTYKGLKHRIKTDHEKETL LVNMVDYLSKVPHEIYQNLSEQNKSLFLEDINEYMRDNEENHD SSEASRVIHPVIRKRYENKFAYFAIRFLDEFAEFPTLRFMVNVGN YIHDNRKKDIGGTSLITNRTIKQQINVFGNLTEIHKKKNDYFEKE ENKEKTLEWELFPNPSYHFQKENIPIFIDLEKSKETNDLAKEYAK EKKKIFGSSRKKQQNTAKKNRETIINLVFDKYKTSDRKTVTFEQ PTALLSFNELNSFLYAFLVENKTGKELEKIIIEKIANQYQILKNCS STVDKTNDNIPKSIKKIVNTTTDSFYFEGKKIDIEKLEKDITIEIEK

[0456]

		TNEKLETIKENEESAQNYKRNERNTQKRKLYRKYVFFTNEIGIE ATWITNDILRFLDNKENWKGYQHSELQKFISQYDNYKKEALGL LESEWNLESDAFFGQNLKRMFQSNSTFETFYKKYLDNRKNTLE TYLSAIENLKTMTDVRPKVLKKKWTELFRFFDKKIYLLSTIETKI NELITKPINLSRGIFEEKPTFINGKNPNKENNQHLFANWFIYAKK QTILQDFYNLPLEQPKAITNLKKHKYKLERSINNLKIEDIYIKQM VDFLYQKLFEQSFIGSLQDLYTSKEKREIEKGKAKNEQTPDESFI WKKQVEINTHNGRIIAKTKIKDIGKFKNLLTDNKIAHLISYDDRI WDFSLNNDGDITKKLYSINTELESYETIRREKLLKQIQQFEQFLL EQETEYSAERKHPEKFEKDCNPNFKKYIIEGVLNKIIPNHEIEEIEI LKSKEDVFKINFSDILILNNDNIKKGYLLIMIRNKFAHNQLIDKN LFNFSLQLYSKNENENFSEYLNKVCQNIIQEFKEKLK
포르피로모나스 굴라에 (SEQ ID No. 142)	WP_046201018	MTEQSERPYNGTYYTLEDKHFWAAFLNLARHNAYITLTHIDRQ LAYSKADITNDQDVLSFKALWKNFDNDLERKSRLRSLILKHFSF LEGAAYGKKLFESKSSGNKSSKNKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHSESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRYGHNDN PSFKHHFVDSEGMVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKPLYDRLREKDRARFRVPVDILPDEDDTDGGGEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKMIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETG DKPYISQTTPHYHIEKGKIGLRFMPEGQHLWPSPEVGTTRTGRS KYAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKV QGRIKRVIEDVYAIYDAFARDEINTLKELDACLADKGIRRGHLP KQMIAILSQEHKDMEEKIRKKLQEMIADTDHRLDMLDRQTDRK IRIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKKRLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLRARKAFLERIGRSDRMENRPFLLLKEP KTDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYRE VGFMAKAVPLYFERACEDRVQPFYDSPFNVGNSLKPKKGRFLS KEERAEEWERGKERFRDLEAWSHSAARRIEDAFAGIEYASPGN KKKIEQLLRDLSLWEAFESKLKVRADKINLAKLKKEILEAQEHP

[0457]

		YHDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDT GTLYLKDIRPNVQEQGSLNVLNRVKPMRLPVVVYRADSRGHV HKEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDTG GLAMEQYPISKLRVEYELAKYQTARVCVFELTLRLEESLLTRYP HLPDESFRKMLESWSDPLLAKWPELHGKVRLLIAVRNAFSHNQ YPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQAKE TVERIIQV
WP_04743 1796 (SEQ ID No. 143)	크리세오 박테리움 sp. YR477	METQTIGHGIAYDHSKIQDKHFFGGFLNLAENNIKAVLKAFSEK FNVGNVDVKQFADVSLKDNLPDNDFQKRVSFLKMYFPVVDFIN IPNNRAKFRSDLTTLFKSVDQLRNFYTHYYHKPLDFDASLFILLD DIFARTAKEVRDQKMKDDKTRQLLSKSLSEELQKGYELQLERL KELNRLGKKVNIHDQLGIKNGVLNNAFNHLIYKDGESFKTKLT YSSALTSFESAENGIEISQSGLLFLLSMFLKRKEIEDLKNRNKGF KAKVVIDEDGKVNGLKFMATHWVFSYLCFKGLKSKLSTEFHEE TLLIQIIDELSKVPDELYCAFDKETRDKFIEDINEYVKEGHQDFSL EDAKVIHPVIRKRYENKFNYFAIRFLDEFVKFPSLRFQVHVGNY VHDRRIKNIDGTTFETERVVKDRIKVFGRLSEISSYKAQYLSSVS DKHDETGWEIFPNPSYVFINNNIPIHISVDTSFKKEIADFKKLRRA QVPDELKIRGAEKKRKFEITQMIGSKSVLNQEEPIALLSLNEIPAL LYEILINGKEPAEIERIIKDKLNERQDVIKNYNPENWLPASQISRR LRSNKGERIINTDKLLQLVTKELLVTEQKLKIISDNREALKQKKE GKYIRKFIFTNSELGREAIWLADDIKRFMPADVRKEWKGYQHS QLQQSLAFYNSRPKEALAILESSWNLKDEKIIWNEWILKSFTQN KFFDAFYNEYLKGRKKYFAFLSEHIVQYTSNAKNLQKFIKQQM PKDLFEKRHYIIEDLQTEKNKILSKPFIFPRGIFDKKPTFIKGVKV EDSPESFANWYQYGYQKDHQFQKFYDWKRDYSDVFLEHLGKP FINNGDRRTLGMEELKERIIIKQDLKIKKIKIQDLFLRLIAENLFQ KVFKYSAKLPLSDFYLTQEERMEKENMAALQNVREEGDKSPNI IKDNFIWSKMIPYKKGQIIENAVKLKDIGKLNVLSLDDKVQTLL SYDDAKPWSKIALENEFSIGENSYEVIRREKLFKEIQQFESEILFR SGWDGINHPAQLEDNRNPKFKMYIVNGILRKSAGLYSQGEDIW FEYNADFNNLDADVLETKSELVQLAFLVTAIRNKFAHNQLPAK EFYFYIRAKYGFADEPSVALVYLNFTKYAINEFKKVMI

[0458]

리에메렐라 아나티페스티퍼 (SEQ ID No. 144)	WP_049354263	MFFSFHNAQRVIFKHLYKAFDASLRMVKEDYKAHFTVNLTRDF AHLNRKGKNKQDNPDFNRYRFEKDGFFTESGLLFFTNLFLDKR DAYWMLKKVSGFKASHKQREKMTTEVFCRSRILLPKLRLESRY DHNQMLLDMLSELSRCPKLLYEKLSEENKKHFQVEADGFLDEI EEEQNPFKDTLIRHQDRFPYFALRYLDLNESFKSIRFQVDLGTYH YCIYDKKIGDEQEKRHLTRTLLSFGRLQDFTEINRPQEWKALTK DLDYKETSNQPFISKTTPHYHITDNKIGFRLGTSKELYPSLEIKDG ANRIAKYPYNSGFVAHAFISVHELLPLMFYQHLTGKSEDLLKET VRHIQRIYKDFEEERINTIEDLEKANQGRLPLGAFPKQMLGLLQ NKQPDLSEKAKIKIEKLIAETKLLSHRLNTKLKSSPKLGKRREKL IKTGVLADWLVKDFMRFQPVAYDAQNQPIKSSKANSTEFWFIR RALALYGGEKNRLEGYFKQTNLIGNTNPHPFLNKFNWKACRNL VDFYQQYLEQREKFLEAIKNQPWEPYQYCLLLKIPKENRKNLV KGWEQGGISLPRGLFTEAIRETLSEDLMLSKPIRKEIKKHGRVGF ISRAITLYFKEKYQDKHQSFYNLSYKLEAKAPLLKREEHYEYW QQNKPQSPTESQRLELHTSDRWKDYLLYKRWQHLEKKLRLYR NQDVMLWLMTLELTKNHFKELNLNYHQLKLENLAVNVQEAD AKLNPLNQTLPMVLPVKVYPATAFGEVQYHKTPIRTVYIREEHT KALKMGNFKALVKDRRLNGLFSFIKEENDTQKHPISQLRLRREL EIYQSLRVDAFKETLSLEEKLLNKHTSLSSLENEFRALLEEWKK EYAASSMVTDEHIAFIASVRNAFCHNQYPFYKEALHAPIPLFTV AQPTTEEKDGLGIAEALLKVLREYCEIVKSQI
포르피로모나스 진지발리스 (SEQ ID No. 145)	WP_052912312	MTEQNEKPYNGTYYTLEDKHFWAAFFNLARHNAYITLAHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDKYGNNDN PFFKHHFVDREEKVTEAGLLFFVSLFLEKRDAIWMQKKIRGFKG GTEAYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNELV RCPKLLYDRLREEDRARFRVPVDILSDEDDTDGTEEDPFKNTLV RHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGEQ PEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGDK PYITQTTPHYHIEKGKIGLRFVPEGQLLWPSPEVGATRTGRSKY

[0459]

		AQDKRFTAEAFLSVHELMPMMFYYFLLREKYSEEASAEKVQG RIKRVIEDVYAVYDAFARDEINTRDELDACLADKGIRRGHLPRQ MIAILSQEHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKIR IGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSKA NSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLHE TRWESHTNILSFYRSYLKARKAFLQSIGRSDREENHRFLLLKEPK TDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYKEV GFMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGRFLSK EKRAEEWESGKERFRDLEAWSHSAARRIEDAFVGIEYASWENK KKIEQLLQDLSLWETFESKLKVKADKINIAKLKKEILEAKEHPY HDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGLDTG TLYLKDIRTDVQEQGSLNVLNHVKPMRLPVVVYRADSRGHVH KEEAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFVDTGA LAMEQYPISKLRVEYELAKYQTARVCAFEQTLELEESLLTRYPH LPDESFREMLESWSDPLLDKWPDLQREVRLLIAVRNAFSHNQY PMYDETIFSSIRKYDPSSLDAIEERMGLNIAHRLSEEVKLAKEMV ERIIQA
포르피로모나스 진지발리스 (SEQ ID No. 146)	WP_058019250	MTEQNEKPYNGTYYTLKDKHFWAAFFNLARHNAYITLTHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFESQSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPMFDGNMLQRLYN VFDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRCGNND NPFFKHHFVDREGKVTEAGLLFFVSLFLEKRDAIWMQKKIRGF KGGTETYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNE LVRCPKSLYDRLREEDRACFRVPVDILSDEDDTDGAEEDPFKNT LVRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIG EQPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDCFETG DKPYITQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGATRTGRS KYAQDKRFTAEAFLSVHELMPMMFYYFLLREKYSEEVSAERV QGRIKRVIEDVYAVYDAFARDEINTRDELDACLADKGIRRGHLP RQMIAILSQKHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDR KIRIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNS KANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFL

[0460]

		HETRWESHTNILSFYRSYLKARKAFLQSIGRSDRVENHRFLLLK EPKTDRQTLVAGWKGEFHLPRGIFTEAVRDCLIEMGLDEVGSY KEVGFMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGR FLSKEKRAEEWESGKERFRDLEAWSHSAARRIEDAFAGIENASR ENKKKIEQLLQDLSLWETFESKLKVKADKINIAKLKKEILEAKE HPYLDFKSWQKFERELRLVKNQDIITWMMCRDLMEENKVEGL DTGTLYLKDIRTDVQEQGSLNVLNHVKPMRLPVVVYRADSRG HVHKEQAPLATVYIEERDTKLLKQGNFKSFVKDRRLNGLFSFV DTGALAMEQYPISKLRVEYELAKYQTARVCAFEQTLELEESLLT RYPHLPDENFRKMLESWSDPLLDKWPDLHRKVRLLIAVRNAFS HNQYPMYDEAVFSSIRKYDPSSPDAIEERMGLNIAHRLSEEVKQ AKEMAERIIQA
플라보박테리움 콜럼나레 (SEQ ID No. 147)	WP_060381855	MSSKNESYNKQKTFNHYKQEDKYFFGGFLNNADDNLRQVGKE FKTRINFNHNNNELASVFKDYFNKEKSVAKREHALNLLSNYFP VLERIQKHTNHNFEQTREIFELLLDTIKKLRDYYTHHYHKPITIN PKVYDFLDDTLLDVLITIKKKKVKNDTSRELLKEKFRPELTQLK NQKREELIKKGKKLLEENLENAVFNHCLRPFLEENKTDDKQNK TVSLRKYRKSKPNEETSITLTQSGLVFLISFFLHRKEFQVFTSGLE GFKAKVNTIKEEEISLNKNNIVYMITHWSYSYYNFKGLKHRIKT DQGVSTLEQNNTTHSLTNTNTKEALLTQIVDYLSKVPNEIYETL SEKQQKEFEEDINEYMRENPENEDSTFSSIVSHKVIRKRYENKFN YFAMRFLDEYAELPTLRFMVNFGDYIKDRQKKILESIQFDSERII KKEIHLFEKLGLVTEYKKNVYLKETSNIDLSRFPLFPSPSYVMA NNNIPFYIDSRSNNLDEYLNQKKKAQSQNRKRNLTFEKYNKEQ SKDAIIAMLQKEIGVKDLQQRSTIGLLSCNELPSMLYEVIVKDIK GAELENKIAQKIREQYQSIRDFTLDSPQKDNIPTTLTKTISTDTSV TFENQPIDIPRLKNALQKELTLTQEKLLNVKQHEIEVDNYNRNK NTYKFKNQPKDKVDDNKLQRKYVFYRNEIGQEANWLASDLIH FMKNKSLWKGYMHNELQSFLAFFEDKKNDCIALLETVFNLKED CILTKDLKNLFLKHGNFIDFYKEYLKLKEDFLNTESTFLENGFIG LPPKILKKELSKRLNYIFIVFQKRQFIIKELEEKKNNLYADAINLS RGIFDEKPTMIPFKKPNPDEFASWFVASYQYNNYQSFYELTPDK IENDKKKKYKNLRAINKVKIQDYYLKLMVDTLYQDLFNQPLDK

[0461]

		SLSDFYVSKTDREKIKADAKAYQKRNDSFLWNKVIHLSLQNNR ITANPKLKDIGKYKRALQDEKIATLLTYDDRTWTYALQKPEKE NENDYKELHYTALNMELQEYEKVRSKKLLKQVQELEKQILDKF YDFSNNATHPEDLEIEDKKGKRHPNFKLYITKALLKNESEIINLE NIDIEILIKYYDYNTEKLKEKIKNMDEDEKAKIVNTKENYNKITN VLIKKALVLIIIRNKMAHNQYPPKFIYDLATRFVPKKEEEYFACY FNRVFETITTELWENKKKAKEIV
포르피로모나스 진지발리스 (SEQ ID No. 148)	WP_06115647 0	MTEQNERPYNGTYYTLEDKHFWAAFFNLARHNAYITLTHIDRQ LAYSKADITNDEDILFFKGQWKNLDNDLERKARLRSLILKHFSF LEGAAYGKKLFENKSSGNKSSKKKELTKKEKEELQANALSLDN LKSILFDFLQKLKDFRNYYSHYRHPESSELPLFDGNMLQRLYNV FDVSVQRVKRDHEHNDKVDPHRHFNHLVRKGKKDRCGNNDN PFFKHHFVDREGKVTEAGLLFFVSLFLEKRDAIWMQKKIRGFK GGTEAYQQMTNEVFCRSRISLPKLKLESLRTDDWMLLDMLNEL VRCPKSLYDRLREEDRARFRVPVDILSDEDDTDGTEEDPFKNTL VRHQDRFPYFALRYFDLKKVFTSLRFHIDLGTYHFAIYKKNIGE QPEDRHLTRNLYGFGRIQDFAEEHRPEEWKRLVRDLDYFETGD KPYITQTTPHYHIEKGKIGLRFVPEGQHLWPSPEVGATRTGRSK YAQDKRLTAEAFLSVHELMPMMFYYFLLREKYSEEVSAEKVQ GRIKRVIEDVYAVYDAFARGEIDTLDRLDACLADKGIRRGHLPR QMIAILSQEHKDMEEKVRKKLQEMIADTDHRLDMLDRQTDRKI RIGRKNAGLPKSGVIADWLVRDMMRFQPVAKDTSGKPLNNSK ANSTEYRMLQRALALFGGEKERLTPYFRQMNLTGGNNPHPFLH ETRWESHTNILSFYRSYLKARKAFLQSIGRSDREENHRFLLLKEP KTDRQTLVAGWKSEFHLPRGIFTEAVRDCLIEMGYDEVGSYKE VGFMAKAVPLYFERACKDRVQPFYDYPFNVGNSLKPKKGRFLS KEKRAEEWESGKERFRLAKLKKEILEAKEHPYLDFKSWQKFER ELRLVKNQDIITWMMCRDLMEENKVEGLDTGTLYLKDIRTEVQ EQGSLNVLNRVKPMRLPVVVYRADSRGHVHKEQAPLATVYIEE RDTKLLKQGNFKSFVKDRRLNGLFSFVDTGGLAMEQYPISKLR VEYELAKYQTARVCAFEQTLELEESLLTRCPHLPDKNFRKMLES WSDPLLDKWPDLQREVWLLIAVRNAFSHNQYPMYDEAVFSSIR KYDPSSPDAIEERMGLNIAHRLSEEVKQAKEMAERIIQA

[0462]

포르피로모나스 진지발리스 (SEQ ID No. 149)	WP_061156637	MNTVPASENKGQSRTVEDDPQYFGLYLNLARENLIEVESHVRIK FGKKKLNEESLKQSLLCDHLLSVDRWTKVYGHSRRYLPFLHYF DPDSQIEKDHDSKTGVDPDSAQRLIRELYSLLDFLRNDFSHNRL DGTTFEHLEVSPDISSFITGTYSLACGRAQSRFADFFKPDDFVLA KNRKEQLISVADGKECLTVSGLAFFICLFLDREQASGMLSRIRGF KRTDENWARAVHETFCDLCIRHPHDRLESSNTKEALLLDMLNE LNRCPRILYDMLPEEERAQFLPALDENSMNNLSENSLNEESRLL WDGSSDWAEALTKRIRHQDRFPYLMLRFIEEMDLLKGIRFRVD LGEIELDSYSKKVGRNGEYDRTITDHALAFGKLSDFQNEEEVSR MISGEASYPVRFSLFAPRYAIYDNKIGYCHTSDPVYPKSKTGEK RALSNPQSMGFISVHDLRKLLLMELLCEGSFSRMQSGFLRKANR ILDETAEGKLQFSALFPEMRHRFIPPQNPKSKDRREKAETTLEKY KQEIKGRKDKLNSQLLSAFDMNQRQLPSRLLDEWMNIRPASHS VKLRTYVKQLNEDCRLRLRKFRKDGDGKARAIPLVGEMATFLS QDIVRMIISEETKKLITSAYYNEMQRSLAQYAGEENRRQFRAIV AELHLLDPSSGHPFLSATMETAHRYTEDFYKCYLEKKREWLAK TFYRPEQDENTKRRISVFFVPDGEARKLLPLLIRRRMKEQNDLQ DWIRNKQAHPIDLPSHLFDSKIMELLKVKDGKKKWNEAFKDW WSTKYPDGMQPFYGLRRELNIHGKSVSYIPSDGKKFADCYTHL MEKTVQDKKRELRTAGKPVPPDLAADIKRSFHRAVNEREFMLR LVQEDDRLMLMAINKMMTDREEDILPGLKNIDSILDKENQFSLA VHAKVLEKEGEGGDNSLSLVPATIEIKSKRKDWSKYIRYRYDR RVPGLMSHFPEHKATLDEVKTLLGEYDRCRIKIFDWAFALEGAI MSDRDLKPYLHESSSREGKSGEHSTLVKMLVEKKGCLTPDESQ YLILIRNKAAHNQFPCAAEMPLIYRDVSAKVGSIEGSSAKDLPE GSSLVDSLWKKYEMIIRKILPILDPENRFFGKLLNNMSQPINDL
리에메렐라 아나티페스티퍼 (SEQ ID No. 150)	WP_061710138	MFFSFHNAQRVIFKHLYKAFDASLRMVKEDYKAHFTVNLTRDF AHLNRKGKNKQDNPDFNRYRFEKDGFFTESGLLFFTNLFLDKR DAYWMLKKVSGFKASHKQSEKMTTEVFCRSRILLPKLRLESRY DHNQMLLDMLSELSRCPKLLYEKLSEKDKKCFQVEADGFLDEI EEEQNPFKDTLIRHQDRFPYFALRYLDLNESFKSIRFQVDLGTYH YCIYDKKIGYEQEKRHLTRTLLNFGRLQDFTEINRPQEWKALTK DLDYNETSNQPFISKTTPHYHITDNKIGFRLRTSKELYPSLEVKD

[0463]

		GANRIAKYPYNSDFVAHAFISISVHELLPLMFYQHLTGKSEDLL KETVRHIQRIYKDFEEERINTIEDLEKANQGRLPLGAFPKQMLGL LQNKQPDLSEKAKIKIEKLIAETKLLSHRLNTKLKSSPKLGKRRE KLIKTGVLADWLVKDFMRFQPVVYDAQNQPIKSSKANSTESRLI RRALALYGGEKNRLEGYFKQTNLIGNTNPHPFLNKFNWKACRN LVDFYQQYLEQREKFLEAIKHQPWEPYQYCLLLKVPKENRKNL VKGWEQGGISLPRGLFTEAIRETLSKDLTLSKPIRKEIKKHGRVG FISRAITLYFKEKYQDKHQSFYNLSYKLEAKAPLLKKEEHYEYW QQNKPQSPTESQRLELHTSDRWKDYLLYKRWQHLEKKLRLYR NQDIMLWLMTLELTKNHFKELNLNYHQLKLENLAVNVQEADA KLNPLNQTLPMVLPVKVYPTTAFGEVQYHETPIRTVYIREEQTK ALKMGNFKALVKDRHLNGLFSFIKEENDTQKHPISQLRLRRELE IYQSLRVDAFKETLSLEEKLLNKHASLSSLENEFRTLLEEWKKK YAASSMVTDKHIAFIASVRNAFCHNQYPFYKETLHAPILLFTVA QPTTEEKDGLGIAEALLRVLREYCEIVKSQI
플라보박테리움 콜럼나레 (SEQ ID No. 151)	WP_063744070	MSSKNESYNKQKTFNHYKQEDKYFFGGFLNNADDNLRQVGKE FKTRINFNHNNNELASVFKDYFNKEKSVAKREHALNLLSNYFP VLERIQKHTNHNFEQTREIFELLLDTIKKLRDYYTHHYHKPITIN PKIYDFLDDTLLDVLITIKKKKVKNDTSRELLKEKLRPELTQLKN QKREELIKKGKKLLEENLENAVFNHCLRPFLEENKTDDKQNKT VSLRKYRKSKPNEETSITLTQSGLVFLMSFFLHRKEFQVFTSGLE GFKAKVNTIKEEKISLNKNNIVYMITHWSYSYYNFKGLKHRIKT DQGVSTLEQNNTTHSLTNTNTKEALLTQIVDYLSKVPNEIYETL SEKQQKEFEEDINEYMRENPENEDSTFSSIVSHKVIRKRYENKFN YFAMRFLDEYAELPTLRFMVNFGDYIKDRQKKILESIQFDSERII KKEIHLFEKLGLVTEYKKNVYLKETSNIDLSRFPLFPSPSYVMA NNNIPFYIDSRSNNLDEYLNQKKKAQSQNRKRNLTFEKYNKEQ SKDAIIAMLQKEIGVKDLQQRSTIGLLSCNELPSMLYEVIVKDIK GAELENKIAQKIREQYQSIRDFTLNSPQKDNIPTTLIKTISTDTSV TFENQPIDIPRLKNAIQKELALTQEKLLNVKQHEIEVNNYNRNK NTYKFKNQPKDKVDDNKLQRKYVFYRNEIGQEANWLASDLIH FMKNKSLWKGYMHNELQSFLAFFEDKKNDCIALLETVFNLKED CILTKDLKNLFLKHGNFIDFYKEYLKLKEDFLNTESTFLENGFIG

[0464]

		LPPKILKKELSKRLNYIFIVFQKRQFIIKELEEKKNNLYADAINLS RGIFDEKPTMIPFKKPNPDEFASWFVASYQYNNYQSFYELTPDK IENDKKKKYKNLRAINKVKIQDYYLKLMVDTLYQDLFNQPLDK SLSDFYVSKTDREKIKADAKAYQKRNDSFLWNKVIHLSLQNNR ITANPKLKDIGKYKRALQDEKIATLLTYDDRTWTYALQKPEKE NENDYKELHYTALNMELQEYEKVRSKKLLKQVQELEKQILDKF YDFSNNATHPEDLEIEDKKGKRHPNFKLYITKALLKNESEIINLE NIDIEILIKYYDYNTEKLKEKIKNMDEDEKAKIVNTKENYNKITN VLIKKALVLIIIRNKMAHNQYPPKFIYDLATRFVPKKEEEYFACY FNRVFETITTELWENKKKAKEIV
리에메렐라 아나티페 스티퍼 (SEQ ID No. 152)	WP_06 497088 7	MEKPLPPNVYTLKHKFFWGAFLNIARHNAFITICHINEQLGLTTP PNDDKIADVVCGTWNNILNNDHDLLKKSQLTELILKHFPFLAA MCYHPPKKEGKKKGSQKEQQKEKENEAQSQAEALNPSELIKVL KTLVKQLRTLRNYYSHHSHKKPDAEKDIFKHLYKAFDASLRMV KEDYKAHFTVNLTQDFAHLNRKGKNKQDNPDFDRYRFEKDGF FTESGLLFFTNLFLDKRDAYWMLKKVSGFKASHKQSEKMTTEV FCRSRILLPKLRLESRYDHNQMLLDMLSELSRYPKLLYEKLSEE DKKRFQVEADGFLDEIEEEQNPFKDTLIRHQDRFPYFALRYLDL NESFKSIRFQVDLGTYHYCIYDKKIGDEQEKRHLTRTLLSFGRL QDFTEINRPQEWKALTKDLDYKETSKQPFISKTTPHYHITDNKIG FRLGTSKELYPSLEVKDGANRIAQYPYNSDFVAHAFISVHELLP LMFYQHLTGKSEDLLKETVRHIQRIYKDFEEERINTIEDLEKANQ GRLPLGAFPKQMLGLLQNKQPDLSEKAKIKIEKLIAETKLLSHR LNTKLKSSPKLGKRREKLIKTGVLADWLVKDFMRFQPVAYDA QNQPIESSKANSTEFQLIQRALALYGGEKNRLEGYFKQTNLIGN TNPHPFLNKFNWKACRNLVDFYQQYLEQREKFLEAIKNQPWEP YQYCLLLKIPKENRKNLVKGWEQGGISLPRGLFTEAIRETLSKD LTLSKPIRKEIKKHGRVGFISRAITLYFREKYQDDHQSFYDLPYK LEAKASPLPKKEHYEYWQQNKPQSPTELQRLELHTSDRWKDYL LYKRWQHLEKKLRLYRNQDVMLWLMTLELTKNHFKELNLNY HQLKLENLAVNVQEADAKLNPLNQTLPMVLPVKVYPATAFGE VQYQETPIRTVYIREEQTKALKMGNFKALVKDRRLNGLFSFIKE ENDTQKHPISQLRLRRELEIYQSLRVDAFKETLNLEEKLLKKHTS

[0465]

		LSSVENKFRILLEEWKKEYAASSMVTDEHIAFIASVRNAFCHNQ YPFYEEALHAPIPLFTVAQQTTEEKDGLGIAEALLRVLREYCEIV KSQI
시노미크로비움 오세아니 (SEQ ID No. 153)	WP_07231947 6.1	MESTTTLGLHLKYQHDLFEDKHYFGGGVNLAVQNIESIFQAFA ERYGIQNPLRKNGVPAINNIFHDNISISNYKEYLKFLKQYLPVVG FLEKSNEINIFEFREDFEILINAIYKLRHFYTHYYHSPIKLEDRFYT CLNELFVAVAIQVKKHKMKSDKTRQLLNKNLHQLLQQLIEQKR EKLKDKKAEGEKVSLDTKSIENAVLNDAFVHLLDKDENIRLNY SSRLSEDIITKNGITLSISGLLFLLSLFLQRKEAEDLRSRIEGFKGK GNELRFMATHWVFSYLNVKRIKHRLNTDFQKETLLIQIADELSK VPDEVYKTLDHENRSKFLEDINEYIREGNEDASLNESTVVHGVI RKRYENKFHYLVLRYLDEFVDFPSLRFQVHLGNYIHDRRDKVI DGTNFITNRVIKEPIKVFGKLSHVSKLKSDYMESLSREHKNGWD VFPNPSYNFVGHNIPIFINLRSASSKGKELYRDLMKIKSEKKKKS REEGIPMERRDGKPTKIEISNQIDRNIKDNNFKDIYPGEPLAMLS LNELPALLFELLRRPSITPQDIEDRMVEKLYERFQIIRDYKPGDG LSTSKISKKLRKADNSTRLDGKKLLRAIQTETRNAREKLHTLEE NKALQKNRKRRTVYTTREQGREASWLAQDLKRFMPIASRKEW RGYHHSQLQQILAFYDQNPKQPLELLEQFWDLKEDTYVWNSWI HKSLSQHNGFVPMYEGYLKGRLGYYKKLESDIIGFLEEHKVLK RYYTQQHLNVIFRERLYFIKTETKQKLELLARPLVFPRGIFDDKP TFVQDKKVVDHPELFADWYVYSYKDDHSFQEFYHYKRDYNEI FETELSWDIDFKDNKRQLNPSEQMDLFRMKWDLKIKKIKIQDIF LKIVAEDIYLKIFGHKIPLSLSDFYISRQERLTLDEQAVAQSMRLP GDTSENQIKESNLWQTTVPYEKEQIREPKIKLKDIGKFKYFLQQ QKVLNLLKYDPQHVWTKAELEEELYIGKHSYEVVRREMLLQK CHQLEKHILEQFRFDGSNHPRELEQGNHPNFKMYIVNGILTKRG ELEIEAENWWLELGNSKNSLDKVEVELLTMKTIPEQKAFLLILIR NKFAHNQLPADNYFHYASNLMNLKKSDTYSLFWFTVADTIVQ EFMSL
레이켄바키엘라 아가리퍼포란스	WP_07312444 1.1	MKTNPLIASSGEKPNYKKFNTESDKSFKKIFQNKGSIAPIAEKAC KNFEIKSKSPVNRDGRLHYFSVGHAFKNIDSKNVFRYELDESQM DMKPTQFLALQKEFFDFQGALNGLLKHIRNVNSHYVHTFEKLEI

[0466]

ans (SEQ ID No. 154)		QSINQKLITFLIEAFELAVIHSYLNEEELSYEAYKDDPQSGQKLV QFLCDKFYPNKEHEVEERKTILAKNKRQALEHLLFIEVTSDIDW KLFEKHKVFTISNGKYLSFHACLFLLSLFLYKSEANQLISKIKGF KRNDDNQYRSKRQIFTFFSKKFTSQDVNSEEQHLVKFRDVIQYL NHYPSAWNKHLELKSGYPQMTDKLMRYIVEAEIYRSFPDQTDN HRFLLFAIREFFGQSCLDTWTGNTPINFSNQEQKGFSYEINTSAEI KDIETKLKALVLKGPLNFKEKKEQNRLEKDLRREKKEQPTNRV KEKLLTRIQHNMLYVSYGRNQDRFMDFAARFLAETDYFGKDA KFKMYQFYTSDEQRDHLKEQKKELPKKEFEKLKYHQSKLVDY FTYAEQQARYPDWDTPFVVENNAIQIKVTLFNGAKKIVSVQRN LMLYLLEDALYSEKRENAGKGLISGYFVHHQKELKDQLDILEK ETEISREQKREFKKLLPKRLLHRYSPAQINDTTEWNPMEVILEEA KAQEQRYQLLLEKAILHQTEEDFLKRNKGKQFKLRFVRKAWH LMYLKELYMNKVAEHGHHKSFHITKEEFNDFCRWMFAFDEVP KYKEYLCDYFSQKGFFNNAEFKDLIESSTSLNDLYEKTKQRFEG WSKDLTKQSDENKYLLANYESMLKDDMLYVNISHFISYLESKG KINRNAHGHIAYKALNNVPHLIEEYYYKDRLAPEEYKSHGKLY NKLKTVKLEDALLYEMAMHYLSLEPALVPKVKTKVKDILSSNI AFDIKDAAGHHLYHLLIPFHKIDSFVALINHQSQQEKDPDKTSFL AKIQPYLEKVKNSKDLKAVYHYYKDTPHTLRYEDLNMIHSHIV SQSVQFTKVALKLEEYFIAKKSITLQIARQISYSEIADLSNYFTDE VRNTAFHFDVPETAYSMILQGIESEFLDREIKPQKPKSLSELSTQ QVSVCTAFLETLHNNLFDRKDDKKERLSKARERYFEQIN

[0467]

[0468] 일정 예의 실시형태에서, RNA-표적화 이펙터 단백질은 2017년 6월 26일 출원된 미국 가출원 제62/525,165호, 및 2017년 8월 16일 출원된 PCT 출원 번호 US 2017/047193에 개시된 바와 같은 Cas13c 이펙터 단백질이다. Cas13c의 예시적인 야생형 오솔로그 서열은 하기 표 4B에 제공된다. 일정 예의 실시형태에서, CRISPR 이펙터 단백질은 표 4a 또는 4b의 Cas13c 단백질이다.

[0469] **표 4a**

푸소박테리움 네크로포럼 subsp.		MEKFRRQNRNSIIKIIISNYDTKGIKELKVRYRKQAQLDTFIIKTEI VNNDIFIKSIIEKAREKYRYSFLFDGEEKYHFKNKSSVEIVKKDIF SQTPDNMIRNYKITLKISEKNPRVVEAEIEDLMNSTILKDGRRSA

[0470]

푼둘리포르메 ATCC 51357 contig00003 (SEQ ID No. 155)		RREKSMTERKLIEEKVAKNYSLLANCPMEEVDSIKIYKIKRFLT YRSNMLLYFASINSFLCEGIKGKDNETEEIWHLKDNDVRKEKV RENFKNKLIQSTENYNSSLKNQIEEKEKLLRKEFKKGAFYRTIIK KLQQERIKELSEKSLTEDCEKIIKLYSKLRHSLMHYDYQYFENLF ENKKNDDLMKDLNLDLFKSLPLIRKMKLNNKVNYLEDGDTLF VLQKTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVF KQIINEKFQSEMEFLEKRISESEKKNEKLKKKLDSMKAHFRNINS EDTKEAYFWDIHSSRNYKTKYNERKNLVNEYTELLGSSKEKKL LREEITKINRQLLKLKQEMEEITKKNSLFRLEYKMKIAFGFLFCE FDGNISKFKDEFDASNQEKIIQYHKNGEKYLTSFLKEEEKEKFNL EKMQKIIQKTEEEDWLLPETKNNLFKFYLLTYLLLPYELKGDFL GFVKKHYYDIKNVDFIDENQNNIQVSQTVEKQEDYFYHKIRLFE KNTKKYEIVKYSIVPNEKLKQYFEDLGIDIKYLTVEQKSEVSEEK NKKVSLKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFFLRE KSGKPLEIFRKELESKMKDGYLNFGQLLYVVYEVLVKNKDLDK ILSKKIDYRKDKSFSPEIAYLRNFLSHLNYSKFLDNFMKINTNKS DENKEVLIPSIKIQKMIQFIEKCNLQNQIDFDFNFVNDFYMRKEK MFFIQLKQIFPDINSTEKQKMNEKEEILRNRYHLTDKKNEQIKDE HEAQSQLYEKILSLQKIYSSDKNNFYGRLKEEKLLFLEKQGKKK LSMEEIKDKIAGDISDLLGILKKEITRDIKDKLTEKFRYCEEKLLN LSFYNHQDKKKEESIRVFLIRDKNSDNFKFESILDDGSNKIFISKN GKEITIQCCDKVLETLIIEKNTLKISSNGKIISLIPHYSYSIDVKY
푸소박테리움 네크로포럼 DJ-2 contig0065, 전체 게놈 샷건 서열 (SEQ ID No. 156)		MEKFRRQNRSSIIKIIISNYDTKGIKELKVRYRKQAQLDTFIIKTEI VNNDIFIKSIIEKAREKYRYSFLFDGEEKYHFKNKSSVEIVKKDIF SQTPDNMIRNYKITLKISEKNPRVVEAEIEDLMNSTILKDGRRSA RREKSMTERKLIEEKVAENYSLLANCPMEEVDSIKIYKIKRFLTY RSNMLLYFASINSFLCEGIKGKDNETEEIWHLKDNDVRKEKVKE NFKNKLIQSTENYNSSLKNQIEEKEKLLRKESKKGAFYRTIIKKL QQERIKELSEKSLTEDCEKIIKLYSELRHPLMHYDYQYFENLFEN KENSELTKNLNLDIFKSLPLVRKMKLNNKVNYLEDNDTLFVLQ KTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVFKQII NEKFQSEIEFLEKRISESEKKNEKLKKKLDSMKAHFRNINSEDTK EAYFWDIHSSRNYKTKYNERKNLVNEYTELLGSSKEKKLLREEI

[0471]

		TKINRQLLKLKQEMEEITKKNSLFRLEYKMKMAFGFLFCEFDG NISRFKDEFDASNQEKIIQYHKNGEKYLTYFLKEEEKEKFNLKK LQETIQKTGEENWLLPQNKNNLFKFYLLTYLLLPYELKGDFLGF VKKHYYDIKNVDFMDENQSSKIIESKEDDFYHKIRLFEKNTKKY EIVKYSIVPDKKLKQYFKDLGIDTKYLILDQKSEVSGEKNKKVS LKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFFLREKSGKPF EVFLKELKDKMIGKQLNFGQLLYVVYEVLVKNKDLSEILSERID YRKDMCFSAEIADLRNFLSHLNYSKFLDNFMKINTNKSDENKE VLIPSIKIQKMIKFIEECNLQSQIDFDFNFVNDFYMRKEKMFFIQL KQIFPDINSTEKQKMNEKEEILRNRYHLTDKKNEQIKDEHEAQS QLYEKILSLQKIYSSDKNNFYGRLKEEKLLFLEKQEKKKLSMEEI KDKIAGDISDLLGILKKEITRDIKDKLTEKFRYCEEKLLNLSFYN HQDKKKEESIRVFLIRDKNSDNFKFESILDDGSNKIFISKNGKEIT IQCCDKVLETLIIEKNTLKISSNGKIISLIPHYSYSIDVKY
푸소박테리움 네크로포럼 BFTR-1 contig0068 (SEQ ID No. 157)		MKVRYRKQAQLDTFIIKTEIVNNDIFIKSIIEKAREKYRYSFLFDG EEKYHFKNKSSVEIVKNDIFSQTPDNMIRNYKITLKISEKNPRVV EAEIEDLMNSTILKDGRRSARREKSMTERKLIEEKVAENYSLLA NCPIEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGKDNE TEEIWHLKDNDVRKEKVKENFKNKLIQSTENYNSSLKNQIEEKE KLSSKEFKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKLYSE LRHPLMHYDYQYFENLFENKENSELTKNLNLDIFKSLPLVRKM KLNNKVNYLEDNDTLFVLQKTKKAKTLYQIYDALCEQKNGFN KFINDFFVSDGEENTVFKQIINEKFQSEMEFLEKRISESEKKNEKL KKKLDSMKAHFRNINSEDTKEAYFWDIHSSRNYKTKYNERKNL VNEYTKLLGSSKEKKLLREEITKINRQLLKLKQEMEEITKKNSLF RLEYKMKIAFGFLFCEFDGNISKFKDEFDASNQEKIIQYHKNGE KYLTSFLKEEEKEKFNLEKMQKIIQKTEEEDWLLPETKNNLFKF YLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFMDENQNNIQVS QTVEKQEDYFYHKIRLFEKNTKKYEIVKYSIVPNEKLKQYFEDL GIDIKYLTGSVESGEKWLGENLGIDIKYLTVEQKSEVSEEKNKK VSLKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFFLREKSEK PFEVFLEELKDKMIGKQLNFGQLLYVVYEVLVKNKDLDKILSK KIDYRKDKSFSPEIAYLRNFLSHLNYSKFLDNFMKINTNKSDEN

[0472]

		KEVLIPSIKIQKMIQFIEKCNLQNQIDFDFNFVNDFYMRKEKMFF IQLKQIFPDINSTEKQKKSEKEEILRKRYHLINKKNEQIKDEHEA QSQLYEKILSLQKIFSCDKNNFYRRLKEEKLLFLEKQGKKKISM KEIKDKIASDISDLLGILKKEITRDIKDKLTEKFRYCEEKLLNISFY NHQDKKKEEGIRVFLIRDKNSDNFKFESILDDGSNKIFISKNGKEI TIQCCDKVLETLMIEKNTLKISSNGKIISLIPHYSYSIDVKY
푸소박테리움 네크로포럼 subsp. 푼둘리포르메 1_1_36S cont1.14 (SEQ ID No. 158)		MTEKKSIIFKNKSSVEIVKKDIFSQTPDNMIRNYKITLKISEKNPR VVEAEIEDLMNSTILKDGRRSARREKSMTERKLIEEKVAENYSL LANCPMEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGK DNETEEIWHLKDNDVRKEKVKENFKNKLIQSTENYNSSLKNQIE EKEKLLRKESKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKL YSELRHPLMHYDYQYFENLFENKENSELTKNLNLDIFKSLPLVR KMKLNNKVNYLEDNDTLFVLQKTKKAKTLYQIYDALCEQKNG FNKFINDFFVSDGEENTVFKQIINEKFQSEMEFLEKRISESEKKNE KLKKKFDSMKAHFHNINSEDTKEAYFWDIHSSSNYKTKYNERK NLVNEYTELLGSSKEKKLLREEITQINRKLLKLKQEMEEITKKNS LFRLEYKMKIAFGFLFCEFDGNISKFKDEFDASNQEKIIQYHKNG EKYLTYFLKEEEKEKFNLEKMQKIIQKTEEEDWLLPETKNNLFK FYLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFMDENQNNIQV SQTVEKQEDYFYHKIRLFEKNTKKYEIVKYSIVPNEKLKQYFED LGIDIKYLTGSVESGEKWLGENLGIDIKYLTVEQKSEVSEEKIKK FL
푸소박테리움 퍼포에텐스 ATCC 29250 T364DRAFT_scaffold00009.9_C (SEQ ID No. 159)		MGKPNRSSIIKIIISNYDNKGIKEVKVRYNKQAQLDTFLIKSELK DGKFILYSIVDKAREKYRYSFEIDKTNINKNEILIIKKDIYSNKED KVIRKYILSFEVSEKNDRTIVTKIKDCLETQKKEKFERENTRRLIS ETERKLLSEETQKTYSKIACCSPEDIDSVKIYKIKRYLAYRSNML LFFSLINDIFVKGVVKDNGEEVGEIWRIIDSKEIDEKKTYDLLVE NFKKRMSQEFINYKQSIENKIEKNTNKIKEIEQKLKKEKYKKEIN RLKKQLIELNRENDLLEKDKIELSDEEIREDIEKILKIYSDLRHKL MHYNYQYFENLFENKKISKEKNEDVNLTELLDLNLFRYLPLVR QLKLENKTNYLEKEDKITVLGVSDSAIKYYSYYNFLCEQKNGF NNFINSFFSNDGEENKSFKEKINLSLEKEIEIMEKETNEKIKEINK NELQLMKEQKELGTAYVLDIHSLNDYKISHNERNKNVKLQNDI

[0473]

		MNGNRDKNALDKINKKLVELKIKMDKITKRNSILRLKYKLQVA YGFLMEEYKGNIKKFKDEFDISKEKIKSYKSKGEKYLEVKSEKK YITKILNSIEDIHNITWLKNQEENNLFKFYVLTYILLPFEFRGDFL GFVKKHYYDIKNVEFLDENNDRLTPEQLEKMKNDSFFNKIRLFE KNSKKYDILKESILTSERIGKYFSLLNTGAKYFEYGGEENRGIFN KNIIIPIFKYYQIVLKLYNDVELAMLLTLSESDEKDINKIKELVTL KEKVSPKKIDYEKKYKFSVLLDCFNRIINLGKKDFLASEEVKEV AKTFTNLAYLRNKICHLNYSKFIDDLLTIDTNKSTTDSEGKLLIN DRIRKLIKFIRENNQKMNISIDYNYINDYYMKKEKFIFGQRKQA KTIIDSGKKANKRNKAEELLKMYRVKKENINLIYELSKKLNELT KSELFLLDKKLLKDIDFTDVKIKNKSFFELKNDVKEVANIKQAL QKHSSELIGIYKKEVIMAIKRSIVSKLIYDEEKVLSIIIYDKTNKK YEDFLLEIRRERDINKFQFLIDEKKEKLGYEKIIETKEKKKVVVKI QNNSELVSEPRIIKNKDKKKAKTPEEISKLGILDLTNHYCFNLKI TL
푸소박테리움 울세란스 ATCC 49185 cont2.38 (SEQ ID No. 160)		MENKGNNKKIDFDENYNILVAQIKEYFTKEIENYNNRIDNIIDKK ELLKYSEKKEESEKNKKLEELNKLKSQKLKILTDEEIKADVIKII KIFSDLRHSLMHYEYKYFENLFENKKNEELAELLNLNLFKNLTL LRQMKIENKTNYLEGREEFNIIGKNIKAKEVLGHYNLLAEQKNG FNNFINSFFVQDGTENLEFKKLIDEHFVNAKKRLERNIKKSKKLE KELEKMEQHYQRLNCAYVWDIHTSTTYKKLYNKRKSLIEEYN KQINEIKDKEVITAINVELLRIKKEMEEITKSNSLFRLKYKMQIA YAFLEIEFGGNIAKFKDEFDCSKMEEVQKYLKKGVKYLKYYKD KEAQKNYEFPFEEIFENKDTHNEEWLENTSENNLFKFYILTYLLL PMEFKGDFLGVVKKHYYDIKNVDFTDESEKELSQVQLDKMIGD SFFHKIRLFEKNTKRYEIIKYSILTSDEIKRYFRLLELDVPYFEYE KGTDEIGIFNKNIILTIFKYYQIIFRLYNDLEIHGLFNISSDLDKILR DLKSYGNKNINFREFLYVIKQNNNSSTEEEYRKIWENLEAKYLR LHLLTPEKEEIKTKTKEELEKLNEISNLRNGICHLNYKEIIEEILKT EISEKNKEATLNEKIRKVINFIKENELDKVELGFNFINDFFMKKE QFMFGQIKQVKEGNSDSITTERERKEKNNKKLKETYELNCDNL SEFYETSNNLRERANSSSLLEDSAFLKKIGLYKVKNNKVNSKVK DEEKRIENIKRKLLKDSSDIMGMYKAEVVKKLKEKLILIFKHDE

[0474]

		EKRIYVTVYDTSKAVPENISKEILVKRNNSKEEYFFEDNNKKYV TEYYTLEITETNELKVIPAKKLEGKEFKTEKNKENKLMLNNHYC FNVKIIY
아나에로살리박터 sp. ND1 게놈 어셈블리 아나에로살리박터 마실리엔시스 ND1 (SEQ ID No. 161)		MKSGRREKAKSNKSSIVRVIISNFDDKQVKEIKVLYTKQGGIDVI KFKSTEKDEKGRMKFNFDCAYNRLEEEEFNSFGGKGKQSFFVT TNEDLTELHVTKRHKTTGEIIKDYTIQGKYTPIKQDRTKVTVSIT DNKDHFDSNDLGDKIRLSRSLTQYTNRILLDADVMKNYREIVCS DSEKVDETINIDSQEIYKINRFLSYRSNMIIYYQMINNFLLHYDG EEDKGGNDSINLINEIWKYENKKNDEKEKIIERSYKSIEKSINQYI LNHNTEVESGDKEKKIDISEERIKEDLKKTFILFSRLRHYMVHYN YKFYENLYSGKNFIIYNKDKSKSRRFSELLDLNIFKELSKIKLVK NRAVSNYLDKKTTIHVLNKNINAIKLLDIYRDICETKNGFNNFIN NMMTISGEEDKEYKEMVTKHFNENMNKLSIYLENFKKHSDFKT NNKKKETYNLLKQELDEQKKLRLWFNAPYVYDIHSSKKYKEL YVERKKYVDIHSKLIEAGINNDNKKKLNEINVKLCELNTEMKE MTKLNSKYRLQYKLQLAFGFILEEFNLDIDKFVSAFDKDNNLTI SKFMEKRETYLSKSLDRRDNRFKKLIKDYKFRDTEDIFCSDREN NLVKLYILMYILLPVEIRGDFLGFVKKNYYDLKHVDFIDKRNND NKDTFFHDLRLFEKNVKRLEVTSYSLSDGFLGKKSREKFGKELE KFIYKNVSIALPTNIDIKEFNKSLVLPMMKNYQIIFKLLNDIEISA LFLIAKKEGNEGSITFKKVIDKVRKEDMNGNINFSQVMKMALN EKVNCQIRNSIAHINMKQLYIEPLNIYINNNQNKKTISEQMEEIID ICITKGLTGKELNKNIINDYYMKKEKLVFNLKLRKRNNLVSIDA QQKNMKEKSILNKYDLNYKDENLNIKEIILKVNDLNNKQKLLK ETTEGESNYKNALSKDILLLNGIIRKNINFKIKEMILGIIQQNEYR YVNINIYDKIRKEDHNIDLKINNKYIEISCYENKSNESTDERINFK IKYMDLKVKNELLVPSCYEDIYIKKKIDLEIRYIENCKVVYIDIY YKKYNINLEFDGKTLFVKFNKDVKKNNQKVNLESNYIQNIKFIV S

[0475]

[0476] 표 4b

명칭	서열
EHO19081	MTEKKSIIFKNKSSVEIVKKDIFSQTPDNMIRNYKITLKISEKNPRVVEAEIEDLMNSTILKDGRRSARREKSMTERKLIEEKVAENYSLLANCPMEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGKDNETEEIWHLKDNDVRKEKVKENFKNKLIQSTENYNSSLKNQIEEKEKLLRKESKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKLYSELRHPLMHYDYQYFENLFENKENSELTKNLNLDIFKSLPLVRKMKLNNKVNYLEDNDTLFVLQKTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVFKQIINEKFQSEMEFLEKRISESEKKNEKLKKKFDSMKAHFHNINSEDTKEAYFWDIHSSSNYKTKYNERKNLVNEYTELLGSSKEKKLLREEITQINRKLLKLKQEMEEITKKNSLFRLEYKMKIAFGFLFCEFDGNISKFKDEFDASNQEKIIQYHKNGEKYLTYFLKEEEKEKFNLEKMQKIIQKTEEEDWLLPETKNNLFKFYLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFMDENQNNIQVSQTVEKQEDYFYHKIRLFEKNTKKYEIVKYSIVPNEKLKQYFEDLGIDIKYLTGSVESGEKWLGENLGIDIKYLTVEQKSEVSEEKIKKFL
WP_094899336	MEKDKKGEKIDISQEMIEEDLRKILILFSRLRHSMVHYDYEFYQALYSGKDFVISDKNNLENRMISQLLDLNIFKELSKVKLIKDKAISNYLDKNTTIHVLGQDIKAIRLLDIYRDICGSKNGFNKFINTMITISGEEDREYKEKVIEHFNKKMENLSTYLEKLEKQDNAKRNNKRVYNLLKQKLIEQQKLKEWFGGPYVYDIHSSKRYKELYIERKKLVDRHSKLFEEGLDEKNKKELTKINDELSKLNSEMKEMTKLNSKYRLQYKLQLAFGFILEEFDLNIDTFINNFDKDKDLIISNFMKKRDIYLNRVLDRGDNRLKNIIKEYKFRDTEDIFCNDRDNNLVKLYILMYILLPVEIRGDFLGFVKKNYYDMKHVDFIDKKDKEDKDTFFHDLRLFEKNIRKLEITDYSLSSGFLSKEHKVDIEKKINDFINRNGAMKLPEDITIEEFNKSLILPIMKNYQINFKLLNDIEISALFKIAKDRSITFKQAIDEIKNEDIKKNSKKNDKNNHKDKNINFTQLMKRALHEKIPYKAGMYQIRNNISHIDMEQLYIDPLNSYMNSNKNNITISEQIEKIIDVCVTGGVTGKELNNNIINDYYMKKEKLVFNLKLRKQNDIVSIESQEKNKREEFVFKKYGLDYKDGEINIIEVIQKVNSLQEELRNIKETSKEKLKNKETLFRDISLINGTIRKNINFKIKEMVLDIVRMDEIRHINIHIYYKGENYTRSNIIKFKYAIDGENKKYYLKQHEINDINLELKDKFVTLICNMDKHPNKNKQTINLESNYIQNVKFIIP
WP_040490876	MENKGNNKKIDFDENYNILVAQIKEYFTKEIENYNNRIDNIIDKKELLKYSEKKEESEKNKKLEELNKLKSQKLKILTDEEIKADVIKIIKIFSDLRHSLMHYEYKYFENLFENKKNEELAELLNLNLFKNLTLLRQMKIENKTNYLEGREEFNIIGKNIKAKEVLGHYNLLAEQKNGFNNFINSFFVQDGTENLEFKKLIDEHFVNAKKRLERNIKKSKKLEKELEKMEQHYQRLNCAYVWDIHTSTTYKKLYNKRKSLIEEYNKQINEIKDKEVITAINVELLRIKKEMEEITKSNSLFRLKYKMQIAYAFLEIEFGGNIAKFKDEFDCSKMEEVQKYLKKGVKYLKYYKDKEAQKNYEFPFEEIFENKDTHNEEWLENTSENNLFKFYILTYLLLPMEFKGDFLGVVKKHYYDIKNVDFTDESEKELSQVQLDKMIGDSFFHKIRLFEKNTKRYEIIKYSILTSDEIKRYFRLLELDVPYFEYEKGTDEIGIFNKNIILTIFKYYQIIFRLYNDLEIHGLFNISSDLDKILRDLKSYGNKNINFREFLYVIKQNNNSSTEEEYRKIWENLEAKYLRLHLLTPEKEEIKTKTKEELEKLNEISNLRNGICHLNYKEIIEEILKTEISEKNKEATLNEKIRKVINFIKENELDKVELGFNFINDFFMKKEQFMFGQIKQVKEGNSDSITTERERKEKNNKKLKETYELNCDNLSEFYETSNNLRERANSSSLLEDSAFLKKIGLYKVKNNKVNSKVKDEEKRIENIKRKLLKDSSDIMGMYKAEVVKKLKEKLILIFKHDEEKRIYVTVYDTSKAVPENISKEILVKRNNSKEEYFFEDNNKKYVTEYYTLEITETNELKVIPAKKLEGKEFKTEKNKENKLMLNNHYCFNVKIIY
WP_0	MEEIKHKKNKSSIIRVIVSNYDMTGIKEIKVLYQKQGGVDTFNLKTIINLE

[0477]

47396607	SGNLEIISCKPKEREKYRYEFNCKTEINTISITKKDKVLKKEIRKYSLELYF KNEKKDTVVAKVTDLLKAPDKIEGERNHLRKLSSSTERKLLSKTLCKNY SEISKTPIEEIDSIKIYKIKRFLNYRSNFLIYFALINDFLCAGVKEDDINEVW LIQDKEHTAFLENRIEKITDYIFDKLSKDIENKKNQFEKRIKKYKTSLEEL KTETLEKNKTFYIDSIKTKITNLENKITELSLYNSKESLKEDLIKIISIFTNLR HSLMHYDYKSFENLFENIENEELKNLLDLNLFKSIRMSDEFKTKNRTNYL DGTESFTIVKKHQNLKKLYTYYNNLCDKKNGFNTFINSFFVTDGIENTDF KNLIILHFEKEMEEYKKSIEYYKIKISNEKNKSKKEKLKEKIDLLQSELIN MREHKNLLKQIYFFDIHNSIKYKELYSERKNLIEQYNLQINGVKDVTAIN HINTKLLSLKNKMDKITKQNSLYRLKYKLKIAYSFLMIEFDGDVSKFKN NFDPTNLEKRVEYLDKKEEYLNYTAPKNKFNFAKLEEELQKIQSTSEMG ADYLNVSPENNLFKFYILTYIMLPVEFKGDFLGFVKNHYYNIKNVDFMD ESLLDENEVDSNKLNEKIENLKDSSFFNKIRLFEKNIKKYEIVKYSVSTQE NMKEYFKQLNLDIPYLDYKSTDEIGIFNKNMILPIFKYYQNVFKLCNDIEI HALLALANKKQQNLEYAIYCCSKKNSLNYNELLKTFNRKTYQNLSFIRN KIAHLNYKELFSDLFNNELDLNTKVRCLIEFSQNNKFDQIDLGMNFINDY YMKKTRFIFNQRRLRDLNVPSKEKIIDGKRKQQNDSNNELLKKYGLSRT NIKDIFNKAWY
WP_035935671	MKVRYRKQAQLDTFIIKTEIVNNDIFIKSIIEKAREKYRYSFLFDGEEKYH FKNKSSVEIVKNDIFSQTPDNMIRNYKITLKISEKNPRVVEAEIEDLMNSTI LKDGRRSARREKSMTERKLIEEKVAENYSLLANCPIEEVDSIKIYKIKRFL TYRSNMLLYFASINSFLCEGIKGKDNETEEIWHLKDNDVRKEKVKENFK NKLIQSTENYNSSLKNQIEEKEKLSSKEFKKGAFYRTIIKKLQQERIKELS EKSLTEDCEKIIKLYSELRHPLMHYDYQYFENLFENKENSELTKNLNLDI FKSLPLVRKMKLNNKVNYLEDNDTLFVLQKTKKAKTLYQIYDALCEQK NGFNKFINDFFVSDGEENTVFKQIINEKFQSEMEFLEKRISESEKKNEKLK KKLDSMKAHFRNINSEDTKEAYFWDIHSSRNYKTKYNERKNLVNEYTK LLGSSKEKKLLREEITKINRQLLKLKQEMEEITKKNSLFRLEYKMKIAFGF LFCEFDGNISKFKDEFDASNQEKIIQYHKNGEKYLTSFLKEEEKEKFNLE KMQKIIQKTEEEDWLLPETKNNLFKFYLLTYLLLPYELKGDFLGFVKKH YYDIKNVDFMDENQNNIQVSQTVEKQEDYFYHKIRLFEKNTKKYEIVKY SIVPNEKLKQYFEDLGIDIKYLTGSVESGEKWLGENLGIDIKYLTVEQKSE VSEEKNKKVSLKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFFLREK SEKPFEVFLEELKDKMIGKQLNFGQLLYVVYEVLVKNKDLDKILSKKID YRKDKSFSPEIAYLRNFLSHLNYSKFLDNFMKINTNKSDENKEVLIPSIKI QKMIQFIEKCNLQNQIDFDFNFVNDFYMRKEKMFFIQLKQIFPDINSTEK QKKSEKEEILRKRYHLINKKNEQIKDEHEAQSQLYEKILSLQKIFSCDKN NFYRRLKEEKLLFLEKQGKKKISMKEIKDKIASDISDLLGILKKEITRDIK DKLTEKFRYCEEKLLNISFYNHQDKKKEEGIRVFLIRDKNSDNFKFESILD DGSNKIFISKNGKEITIQCCDKVLETLMIEKNTLKISSNGKIISLIPHYSYSI DVKY
WP_035906563	MEKFRRQNRSSIIKIIISNYDTKGIKELKVRYRKQAQLDTFIIKTEIVNNDIF IKSIIEKAREKYRYSFLFDGEEKYHFKNKSSVEIVKKDIFSQTPDNMIRNY KITLKISEKNPRVVEAEIEDLMNSTILKDGRRSARREKSMTERKLIEEKVA ENYSLLANCPMEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGKD NETEEIWHLKDNDVRKEKVKENFKNKLIQSTENYNSSLKNQIEEKEKLL RKESKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKLYSELRHPLMHY DYQYFENLFENKENSELTKNLNLDIFKSLPLVRKMKLNNKVNYLEDNDT LFVLQKTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVFKQIIN EKFQSEIEFLEKRISESEKKNEKLKKKLDSMKAHFRNINSEDTKEAYFWD

[0478]

	IHSSRNYKTKYNERKNLVNEYTELLGSSKEKKLLREEITKINRQLLKLKQ EMEEITKKNSLFRLEYKMKMAFGFLFCEFDGNISRFKDEFDASNQEKIIQ YHKNGEKYLTYFLKEEEKEKFNLKKLQETIQKTGEENWLLPQNKNNLF KFYLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFMDENQSSKIIESKEDD FYHKIRLFEKNTKKYEIVKYSIVPDKKLKQYFKDLGIDTKYLILDQKSEV SGEKNKKVSLKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFFLREKS GKPFEVFLKELKDKMIGKQLNFGQLLYVVYEVLVKNKDLSEILSERIDY RKDMCFSAEIADLRNFLSHNYSKFLDNFMKINTNKSDENKEVLIPSIKIQ KMIKFIEECNLQSQIDFDFNFVNDFYMRKEKMFFIQLKQIFPDINSTEKQK MNEKEEILRNRYHLTDKKNEQIKDEHEAQSQLYEKILSLQKIYSSDKNNF YGRLKEEKLLFLEKQEKKKLSMEEIKDKIAGDISDLLGILKKEITRDIKDK LTEKFRYCEEKLLNLSFYNHQDKKKEESIRVFLIRDKNSDNFKFESILDD GSNKIFISKNGKEITIQCCDKVLETLIIEKNTLKISSNGKIISLIPHYSYSIDV KY
WP_042678931	MKSGRREKAKSNKSSIVRVIISNFDDKQVKEIKVLYTKQGGIDVIKFKST EKDEKGRMKFNFDCAYNRLEEEEFNSFGGKGKQSFFVTTNEDLTELHVT KRHKTTGEIIKDYTIQGKYTPIKQDRTKVTVSITDNKDHFDSNDLGDKIR LSRSLTQYTNRILLDADVMKNYREIVCSDSEKVDETINIDSQEIYKINRFL SYRSNMIIYYQMINNFLLHYDGEEDKGGNDSINLINEIWKYENKKNDEK EKIIERSYKSIEKSINQYILNHNTEVESGDKEKKIDISEERIKEDLKKTFILF SRLRHYMVHYNYKFYENLYSGKNFIIYNKDKSKSRRFSELLDLNIFKELS KIKLVKNRAVSNYLDKKTTIHVLNKNINAIKLLDIYRDICETKNGFNNFIN NMMTISGEEDKEYKEMVTKHFNENMNKLSIYLENFKKHSDFKTNNKKK ETYNLLKQELDEQKKLRLWFNAPYVYDIHSSKKYKELYVERKKYVDIH SKLIEAGINNDNKKKLNEINVKLCELNTEMKEMTKLNSKYRLQYKLQLA FGFILEEFNLDIDKFVSAFDKDNNLTISKFMEKRETYLSKSLDRRDNRFK KLIKDYKFRDTEDIFCSDRENNLVKLYILMYILLPVEIRGDFLGFVKKNY YDLKHVDFIDKRNNDNKDTFFHDLRLFEKNVKRLEVTSYSLSDGFLGKK SREKFGKELEKFIYKNVSIALPTNIDIKEFNKSLVLPMMKNYQIIFKLLNDI EISALFLIAKKEGNEGSITFKKVIDKVRKEDMNGNINFSQVMKMALNEK VNCQIRNSIAHINMKQLYIEPLNIYINNNQNKKTISEQMEEIIDICITKGLT GKELNKNIINDYYMKKEKLVFNLKLRKRNNLVSIDAQQKNMKEKSILN KYDLNYKDENLNIKEIILKVNDLNNKQKLLKETTEGESNYKNALSKDILL LNGIIRKNINFKIKEMILGIIQQNEYRYVNINIYDKIRKEDHNIDLKINNKYI EISCYENKSNESTDERINFKIKYMDLKVKNELLVPSCYEDIYIKKKIDLEIR YIENCKVVYIDIYYKKYNINLEFDGKTLFVKFNKDVKKNNQKVNLESNY IQNIKFIVS
WP_062627846	MEKFRRQNRNSIIKIIISNYDTKGIKELKVRYRKQAQLDTFIIKTEIVNNDI FIKSIIEKAREKYRYSFLFDGEEKYHFKNKSSVEIVKKDIFSQTPDNMIRN YKITLKISEKNPRVVEAEIEDLMNSTILKDGRRSARREKSVTERKLIEEKV AENYSLLANCPMEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGK ENETEEIWHLKDNDVRKEKVKENFKNKLIQSTENYNSSLKNQIEEKEKL LRKESKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKLYSKLRHSLMH YDYQYFENLFENKETPELKDKLDLHLFKSLPLIRKMKLNNKVNYLEDGD TLFVLQKTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVFKQII NEKFQSEMEFLGKRISESEEKNPKLKKKFDSMKAHFHNINSEDTKEAYF WDIHSSSNYKTKYNERKNLVNEYTELLGSSKEKKLLREEITQINRKLLKL KQEMEEITKKNSLFRLEYKMKMAFGFLFCEFDGNISRFKDEFDASNQEKI IQYHKNGEKYLTYFLKEEEKEKFNLKKLQETIQKTGKENWLLPQNKNNL FKFYLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFMDENQSSKIIESKED

[0479]

	DFYHKIRLFEKNTKKYEIVKYSIVPDEKLKQYFKDLGIDTKYLILEQKSE VSGEKNKKVSLKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFFLREK SGKPFEVFLKELKDKMIGKQLNFGQLLYVIYEVLVKNKDLSEILSERIDY RKDMCFSAEIADLRNFLSHLNYSKFLDNFMKINTNKSDENKEVLIPSIKIQ KMIKFIEECNLQSQIDFDFNFVNDFYMRKEKMFFIQLKQIFPDINSTEKQK MNEKEEILRNRYHLTDKKNEQIKDEHEAQSQLYEKILSLQKIYSSDKNNF YGRLKEEKLLFLGKQGKKKLSMEEIKDKIAGDISDLLGILKKEITRDIKD KLTEKFRYCEEKLLNLSFYNHQDKKKEESIRVFLIRDKNSDNFKFESILD DGSNKIFISKNGKEITIQCCDKVLETLMIEKNTLKISSNGKIISLVPHYSYSI DVKY
WP_0 05959 231	MEKFRRQNRNSIIKIIISNYDTKGIKELKVRYRKQAQLDTFIIKTEIVNNDI FIKSIIEKAREKYRYSFLFDGEEKYHFKNKSSVEIVKKDIFSQTPDNMIRN YKITLKISEKNPRVVEAEIEDLMNSTILKDGRRSARREKSMTERKLIEEKV AKNYSLLANCPMEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGK DNETEEIWHLKDNDVRKEKVRENFKNKLIQSTENYNSSLKNQIEEKEKL LRKEFKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKLYSKLRHSLMH YDYQYFENLFENKKNDDLMKDLNLDLFKSLPLIRKMKLNNKVNYLEDG DTLFVLQKTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVFKQ IINEKFQSEMEFLEKRISESEKKNEKLKKKLDSMKAHFRNINSEDTKEAY FWDIHSSRNYKTKYNERKNLVNEYTELLGSSKEKKLLREEITKINRQLLK LKQEMEEITKKNSLFRLEYKMKIAFGFLFCEFDGNISKFKDEFDASNQEK IIQYHKNGEKYLTSFLKEEEKEKFNLEKMQKIIQKTEEEDWLLPETKNNL FKFYLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFIDENQNNIQVSQTVE KQEDYFYHKIRLFEKNTKKYEIVKYSIVPNEKLKQYFEDLGIDIKYLTVE QKSEVSEEKNKKVSLKNNGMFNKTILLFVFKYYQIAFKLFNDIELYSLFF LREKSGKPLEIFRKELESKMKDGYLNFGQLLYVVYEVLVKNKDLDKILS KKIDYRKDKSFSPEIAYLRNFLSHLNYSKFLDNFMKINTNKSDENKEVLI PSIKIQKMIQFIEKCNLQNQIDFDFNFVNDFYMRKEKMFFIQLKQIFPDINS TEKQKMNEKEEILRNRYHLTDKKNEQIKDEHEAQSQLYEKILSLQKIYSS DKNNFYGRLKEEKLLFLEKQGKKKLSMEEIKDKIAGDISDLLGILKKEIT RDIKDKLTEKFRYCEEKLLNLSFYNHQDKKKEESIRVFLIRDKNSDNFKF ESILDDGSNKIFISKNGKEITIQCCDKVLETLIIEKNTLKISSNGKIISLIPHY SYSIDVKY
WP_0 27128 616	MGKPNRSSIIKIIISNYDNKGIKEVKVRYNKQAQLDTFLIKSELKDGKFIL YSIVDKAREKYRYSFEIDKTNINKNEILIIKKDIYSNKEDKVIRKYILSFEV SEKNDRTIVTKIKDCLETQKKEKFERENTRRLISETERKLLSEETQKTYSK IACCSPEDIDSVKIYKIKRYLAYRSNMLLFFSLINDIFVKGVVKDNGEEVG EIWRIIDSKEIDEKKTYDLLVENFKKRMSQEFINYKQSIENKIEKNTNKIK EIEQKLKKEKYKKEINRLKKQLIELNRENDLLEKDKIELSDEEIREDIEKIL KIYSDLRHKLMHYNYQYFENLFENKKISKEKNEDVNLTELLDLNLFRYL PLVRQLKLENKTNYLEKEDKITVLGVSDSAIKYYSYYNFLCEQKNGFNN FINSFFSNDGEENKSFKEKINLSLEKEIEIMEKETNEKIKEINKNELQLMKE QKELGTAYVLDIHSLNDYKISHNERNKNVKLQNDIMNGNRDKNALDKI NKKLVELKIKMDKITKRNSILRLKYKLQVAYGFLMEEYKGNIKKFKDEF DISKEKIKSYKSKGEKYLEVKSEKKYITKILNSIEDIHNITWLKNQEENNL FKFYVLTYILLPFEFRGDFLGFVKKHYYDIKNVEFLDENNDRLTPEQLEK MKNDSFFNKIRLFEKNSKKYDILKESILTSERIGKYFSLLNTGAKYFEYGG EENRGIFNKNIIIPIFKYYQIVLKLYNDVELAMLLTLSESDEKDINKIKELV TLKEKVSPKKIDYEKKYKFSVLLDCFNRIINLGKKDFLASEEVKEVAKTF TNLAYLRNKICHLNYSKFIDDLLTIDTNKSTTDSEGKLLINDRIRKLIKFIR

[0480]

	ENNQKMNISIDYNYINDYYMKKEKFIFGQRKQAKTIIDSGKKANKRNKA EELLKMYRVKKENINLIYELSKKLNELTKSELFLLDKKLLKDIDFTDVKI KNKSFFELKNDVKEVANIKQALQKHSSELIGIYKKEVIMAIKRSIVSKLIY DEEKVLSIIIYDKTNKKYEDFLLEIRRERDINKFQFLIDEKKEKLGYEKIIE TKEKKKVVVKIQNNSELVSEPRIIKNKDKKKAKTPEEISKLGILDLTNHY CFNLKITL
WP_0 62624 740	MEKFRRQNRNSIIKIIISNYDTKGIKELKVRYRKQAQLDTFIIKTEIVNNDI FIKSIIEKAREKYRYSFLFDGEEKYHFKNKSSVEIVKKDIFSQTPDNMIRN YKITLKISEKNPRVVEAEIEDLMNSTILKDGRRSARREKSMTERKLIEEKV AKNYSLLANCPMEEVDSIKIYKIKRFLTYRSNMLLYFASINSFLCEGIKGK DNETEEIWHLKDNDVRKEKVRENFKNKLIQSTENYNSSLKNQIEEKEKL LRKEFKKGAFYRTIIKKLQQERIKELSEKSLTEDCEKIIKLYSKLRHSLMH YDYQYFENLFENKKNDDLMKDLNLDLFKSLPLIRKMKLNNKVNYLEDG DTLFVLQKTKKAKTLYQIYDALCEQKNGFNKFINDFFVSDGEENTVFKQ IINEKFQSEMEFLEKRISESEKKNEKLKKKLDSMKAHFRNINSEDTKEAY FWDIHSSRNYKTKYNERKNLVNEYTELLGSSKEKKLLREEITKINRQLLK LKQEMEEITKKNSLFRLEYKMKIAFGFLFCEFDGNISKFKDEFDASNQEK IIQYHKNGEKYLTSFLKEEEKEKFNLEKMQKIIQKTEEEDWLLPETKNNL FKFYLLTYLLLPYELKGDFLGFVKKHYYDIKNVDFIDENQNNIQVSQTVE KQEDYFYHKIRLFEKNTKKYEIVKYSIVPNEKLKQYFEDLGIDIKYLTGS VESGEKWLGENLGIDIKYLTVEQKSEVSEEKNKKVSLKNNGMFNKTILL FVFKYYQIAFKLFNDIELYSLFFLREKSGKPLEIFRKELESKMKDGYLNFG QLLYVVYEVLVKNKDLDKILSKKIDYRKDKSFSPEIAYLRNFLSHLNYSK FLDNFMKINTNKSDENKEVLIPSIKIQKMIQFIEKCNLQNQIDFDFNFVND FYMRKEKMFFIQLKQIFPDINSTEKQKMNEKEEILRNRYHLTDKKNEQIK DEHEAQSQLYEKILSLQKIYSSDKNNFYGRLKEEKLLFLEKQGKKKLSM EEIKDKIAGDISDLLGILKKEITRDIKDKLTEKFRYCEEKLLNLSFYNHQD KKKEESIRVFLIRDKNSDNFKFESILDDGSNKIFISKNGKEITIQCCDKVLE TLIIEKNTLKISSNGKIISLIPHYSYSIDVKY
WP_0 96402 050	MENKNKPNRGSIVRIIISNYDMKGIKELKVRYRKQAQLDTFILQTTLDKS NNSILINDFRVKAREKYRYSFTYDGKEKFSVPSNSIIVTKIDNAAPEKSKEI RKYKITLGIDEKCKTGSMITAAIEDLLEDDRVREGIRNPRRKASKTERKLI TESICHNYAQITQCPVEEIDAVKIYKVKRFLSYRSNMLLFFALINDFLCKN LKNEKGEKINEIWEMENKGNNKKIDFDENYNILVAQIKEYFTKEIENYN NRIDNIIDKKELLKYSEEKEESEKNKKLEELNKLESQKLKILTDEEIKADV IKIIKIFSDLRHSLMHYEYKYFENLFENKKNEELAELLNLNLFKNLTLLRQ MKIENKTNYLEGDEKFNILGKDVRAKNALGHYDLLVEQKNGFNNFINSF FVQDGTENLEFKKFIDENFIKAQKELEEDIKNCKESVKKLEKKLKENPKK SEDLEKKLEKKQKKLKELKKELEKMKQHYKRLNCAYVWDIHSSTVYK KLYNERKNLIEKYNKQLNGLQDKNAITGINAQLLRIKKEMEEITKSNSLF RLKYKMQIAYAFLEMEYEGNIAKFKNEFDCSKTEKIQEWLEKSEEYLNY CMEKEEDGKNYKFHFKEISEIKDTHNEEWLENTSENNLFKFYILTYLLLP MEFKGDFLGVVKKHYYDIKNVDFTDESEKELSQEQIDKMIGDSFFHKIR LFEKNTKRYEIIKYSILTSDEIKKYFELLELKVPYLEYKGIDEIGIFNKNIIL PIFKYYQIIFRLYNDLEIHGLFNVSFDINKILSDLKSYGNENINFREFLYVIK QNNNSSTEEEYQKIWEKLESKYLKEPLLTPEKKEINKKTEKELKKLDGIS FLRNKISHLEYEKIIEGVLKTAVNGENKKTSETNADKVFLNEKIKKIINFI KENELDKIELGFNFINDFFMKKEQFMFGQIKQVKEGNSDSITTERKRKEE NNKRLKITYGLNYNNLSKIYEFSNTLREIVNSPLFLKDSTLLKKVDLSKV MLKEKPICSLQYENNTKLEDDIKRILLKDSSDIMGIYKAEVVKKLKEKLV

[0481]

	LIFKYDEEKKIYVTVYDTSKAVPENISKEILVKRNNSKEEYFFEDNKKKY TTQYYTLEITKENELKVIPAKKLEGKEFKTEKKEENKLMLNNHYCFNVK IIY

[0482]

[0483] 일부 실시형태에서, AD-기능화된 CRISPR-Cas 시스템의 성분은 다양한 형태, 예컨대 DNA/RNA 또는 RNA/RNA 또는 단백질 RNA의 조합으로 전달될 수 있다. 예를 들어, Cas13 단백질은 DNA-코딩 폴리뉴클레오티드 또는 RNA-코딩 폴리뉴클레오티드로서 또는 단백질로서 전달될 수 있다. 가이드는 DNA-코딩 폴리뉴클레오티드 또는 RNA로서 전달될 수 있다. 혼합된 전달 형태를 포함하여, 모든 가능한 조합이 고려된다.

[0484] 일부 양상에서, 본 발명은 하나 이상의 폴리뉴클레오티드, 예컨대 또는 본 명세서에 기술된 바와 같은 하나 이상의 벡터, 이의 하나 이상의 전사물, 및/또는 그로부터 전사되는 하나 이상의 단백질을 숙주 세포에 전달하는 단계를 포함하는 방법을 제공한다.

[0485] **벡터**

[0486] 일반적으로, 용어 "벡터"는 그것이 연결된 다른 핵산을 수송할 수 있는 핵산 분자를 의미한다. 이는 삽입된 절편의 복제를 초래하기 위해 다른 DNA 세그먼트가 삽입될 수 있는 레플리콘, 예컨대, 플라스미드, 파지 또는 코스미드이다. 일반적으로, 벡터는 적절한 제어 구성요소와 회합될 때 복제될 수 있다. 벡터는 단일-가닥, 이중-가닥, 또는 부분적으로 이중-가닥인 핵산 분자; 하나 이상의 유리 단부를 포함하거나, 유리 단부 부재(예를 들어, 원형)의 핵산 분자; DNA, RNA 또는 둘 모두를 포함하는 핵산 분자; 및 당업계에 공지된 폴리뉴클레오티드의 다른 변이체를 포함하지만, 이들로 제한되지 않는다. 한 유형의 벡터는 추가적인 DNA 절편이, 예컨대, 표준 분자 클로닝 기법에 의해 삽입될 수 있는 원형 이중 가닥 DNA 루프를 지칭하는 "플라스미드"이다.

다른 유형의 벡터는 바이러스 벡터이되, 바이러스 유래된 DNA 또는 RNA 서열은 바이러스 (예를 들어, 레트로바이러스, 복제 결함 레트로바이러스, 아데노바이러스, 복제 결함 아데노바이러스 및 아데노-연합 바이러스) 내로 패키징을 위해 벡터에 존재한다. 바이러스 벡터는 또한 숙주 세포 내로 형질감염을 위해 바이러스에 의해 수행되는 폴리뉴클레오티드를 포함한다. 소정의 벡터는 그들이 도입되는 숙주 세포에서 자가 복제할 수 있다 (예를 들어, 바이러스 복제 기원을 갖는 박테리아 벡터 및 에피솜 포유동물 벡터). 다른 벡터 (예를 들어, 비-에피솜 포유동물 벡터)는 숙주 세포 내로 도입 시 숙주 세포의 게놈에 통합되고, 이에 의해 숙주 게놈과 함께 복제된다. 게다가, 소정의 벡터는 그들이 작동적으로 연결된 유전자의 발현을 유도할 수 있다. 이러한 벡터는 본 명세서에서 "발현 벡터"라고 한다. 진핵생물 세포에서 발현을 위한 그리고 발현을 초래하는 벡터는 본 명세서에서 "진핵생물 발현 벡터"라고 할 수 있다. 재조합 DNA 기법에서 유용성이 있는 통상적인 발현 벡터는 종종 플라스미드 형태이다.

[0487] 재조합 발현 벡터는 숙주 세포에서 핵산의 발현에 적합한 형태로 본 발명의 핵산을 포함할 수 있는데, 이는 재조합 발현 벡터가 발현을 위해 사용될 숙주 세포에 기반하여 선택될 수 있는, 즉, 발현될 핵산 서열에 작동가능하게 연결되는 하나 이상의 조절 구성요소를 포함한다는 것을 의미한다. 재조합 발현 벡터 내에서, "작동적으로 연결된"은 관심 뉴클레오티드 서열이 뉴클레오티드 서열의 발현을 허용하는 방식으로 (예를 들어, 시험관내 전사/번역 시스템에서 또는 벡터가 숙주 세포에 도입될 때 숙주 세포에서) 조절 구성요소(들)에 연결된다는 것을 의미한다. 유리한 벡터는 렌티바이러스 및 아데노-연합 바이러스를 포함하고, 이러한 벡터의 유형은 또한 특정 세포 유형을 표적화하도록 선택된다.

[0488] 재조합 및 클로닝 방법에 관해, 미국 특허 제2004-0171156 A1호로서 2004년 9월 2일자로 공개된 미국 특허 출원 제10/815,730호가 언급되며, 이의 내용은 본 명세서에 그들의 전문이 참고로 편입된다.

[0489] 용어 "조절 구성요소"는 프로모터, 인핸서, 내부 리보솜 진입 부위 (IRES) 및 다른 발현 제어 구성요소 (예를 들어, 전사 종결 신호, 예컨대, 폴리아데닐화 신호 및 폴리-U 서열)를 포함하도록 의도된다. 이러한 조절 구성요소는, 예를 들어, 문헌 [Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)]에 기재되어 있다. 조절 구성요소는 다수 유형의 숙주 세포에서 뉴클레오티드의 구성적 발현을 지시하는 것 및 소정의 숙주 세포에서만 뉴클레오티드 서열(예를 들어, 조직-특이적 조절 서열)의 발현을 유도시키는 것을 포함한다. 조직-특이적 프로모터는 목적하는 관심의 조직, 예컨대, 근육, 뉴런, 뼈, 피부, 혈액, 특정 기관(예를 들어, 간, 췌장) 또는 특정 세포 유형(예를 들어, 림프구)에서 주로 발현을 유도할 수 있다. 조절 구성요소는 또한 조직 또는 세포 유형 특이적일 수도 있고 그렇지 않을 수도 있는, 일시적-의존적 방식으로, 예컨대, 세포-주기 의존적 또는 발생 단계-의존적 방식으로 발현을 지시할 수 있다. 일부 실시형태에서, 벡터는 하나 이상의 pol III 프로모터 (예를 들어, 1, 2, 3, 4, 5개 이상의 pol III 프로모터), 하나 이상의 pol II 프로모터 (예를 들어, 1, 2, 3, 4, 5개 이상의 pol II 프로모터), 하나 이상의 pol I 프로모터 (예를 들어, 1, 2, 3, 4, 5개 이상의 pol I 프로모터), 또는 이들의 조합을 포함한다. pol III 프로모터의 예는 U6 및 H1 프로모터를 포함하지만, 이들로 제한되지 않는다. pol II 프로모터의 예는 레트로바이러스 라우스 육종 바이러스 (RSV) LTR 프로모터(선택적으로 RSV 인핸서를 가짐), 거대세포바이러스 (CMV) 프로모터 (선택적으로 CMV 인핸서를 가짐)[예를 들어, 문헌 [Boshart et al, Cell, 41:521-530 (1985)] 참조], SV40 프로모터, 디히드로폴레이트 리덕타제 프로모터, β-액틴 프로모터, 포스포글리세롤 키나제 (PGK) 프로모터 및 EF1α 프로모터를 포함하지만, 이들로 제한되지 않는다. 또한 용어 "조절 구성요소"에 의해 인핸서 구성요소, 예컨대, WPRE; CMV 인핸서; HTLV-I의 LTR에서 R-U5' 절편 (Mol. Cell. Biol., Vol. 8(1), p. 466-472, 1988); SV40 인핸서; 및 토끼 β-글로빈의 엑손 2와 3 사이의 인트론 서열 (Proc. Natl. Acad. Sci. USA., Vol. 78(3), p. 1527-31, 1981)이 포함된다. 발현 벡터의 디자인은 형질전환될 숙주 세포의 선택, 목적하는 발현 수준 등과 같은 인자에 의존할 수 있다는 것이 당업자에 의해 인식될 것이다. 벡터는 숙주 세포 내로 도입될 수 있으며, 이에 의해 본 명세서에 기재된 바와 같은 핵산에 의해 코딩되는 융합 단백질 또는 펩티드를 비롯한 전사체, 단백질 또는 펩티드 (예를 들어, 주기적으로 분포하는 짧은 회문구조 반

복부 (CRISPR) 전사물, 단백질, 효소, 이들의 돌연변이체 형태, 이들의 융합 단백질 등)를 생성한다. 조절 서열에 관해, 미국 특허 출원 제10/491,026호가 언급되며, 이의 내용은 본 명세서에 그의 전문이 참고로 편입된다. 프로모터에 관해, 국제 특허 출원 WO 2011/028929 및 미국 특허 출원 제12/511,940호가 언급되며, 이의 내용은 본 명세서에 그들의 전문이 참고로 편입된다.

[0490] 유리한 벡터는 렌티바이러스 및 아데노-연합 바이러스를 포함하며, 이러한 벡터의 유형은 또한 특정 유형의 세포를 표적화하기 위해 선택될 수 있다.

[0491] 특정 실시형태에서, 가이드 RNA 및 아데노신 데아미나제에 융합된 (선택적으로 변형되거나 또는 돌연변이된) CRISPR-Cas 단백질에 대한 바이시스트론 벡터가 사용된다. 가이드 RNA 및 아데노신 데아미나제에 융합된 (선택적으로 변형 또는 돌연변이된) CRISPR-Cas 단백질에 대한 바이시스트론 발현 벡터가 바람직하다. 일반적으로 그리고 특히 본 실시형태에서, 아데노신 데아미나제에 융합된 (선택적으로 변형 또는 돌연변이된) CRISPR-Cas 단백질은 바람직하게는 CBh 프로모터에 의해 유도된다. RNA는 바람직하게는 Pol III 프로모터, 예컨대, U6 프로모터에 의해 유도될 수 있다. 이상적으로는 2개가 조합된다.

[0492] 벡터는 원핵 또는 진핵생물 세포에서 CRISPR 전사물 (예를 들어, 핵산 전사물, 단백질 또는 효소)의 발현을 위해 디자인될 수 있다. 예를 들어, CRISPR 전사물은 박테리아 세포, 예컨대, 에스케리치아 콜라이, 곤충 세포 (배큘로바이러스 발현 벡터를 이용), 효모 세포 또는 포유동물 세포에서 발현될 수 있다. 적합한 숙주 세포는 문헌 [Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)]에서 추가로 논의된다. 대안적으로, 재조합 발현 벡터는, 예를 들어 T7 프로모터 조절 서열 및 T7 중합효소를 이용하여 시험관내에서 전사 및 번역될 수 있다.

[0493] 벡터는 원핵생물 또는 원핵 세포에 도입되고 증식될 수 있다. 일부 실시형태에서, 원핵생물은 진핵생물 세포 내로 도입될 벡터의 복제물을 증식시키기 위해 또는 진핵생물 세포 내로 도입될 벡터 생성에서의 중간 벡터로서 사용된다 (예를 들어, 바이러스 벡터 패키징 시스템의 일부분으로서 플라스미드를 증폭시킴). 일부 실시형태에서, 원핵생물은 벡터의 복제물을 증폭시키기 위해 그리고, 예컨대, 숙주 세포 또는 숙주 유기체에 전달을 위한 하나 이상의 단백질 공급원을 제공하기 위해 하나 이상의 핵산을 발현시키기 위해 사용된다. 원핵생물 내 단백질의 발현은 융합 또는 비융합 단백질 중 하나의 발현을 지시하는 구성적 또는 유도성 프로모터를 함유하는 벡터로 에스케리치아 콜라이에서 가장 흔히 수행된다. 융합 벡터는 그 안에서 코딩되는 단백질에, 예컨대, 재조합 단백질의 아미노 말단에 다수의 아미노산을 부가한다. 이러한 융합 벡터는 하나 이상의 목적, 예컨대: (i) 재조합 단백질 발현을 증가시키고; (ii) 재조합 단백질의 용해도를 증가시키고; (iii) 친화도 정제에서 리간드로서 작용함으로써 재조합 단백질의 정제에 도움을 주는 작용을 할 수 있다. 종종, 융합 발현 벡터에서, 단백질 절단 부위는 융합 모이어티 및 재조합 단백질의 접합부에 도입되어 융합 단백질의 후속 정제를 위해 융합 모이어티로부터 재조합 단백질의 분리를 가능하게 한다. 이러한 효소, 및 그들의 동족 인식 서열은 인자 Xa, 트롬빈 및 엔테로키나제를 포함한다. 예시적 융합 발현 벡터는 각각 글루타티온 S-트랜스퍼라제(GST), 말토스 E 결합 단백질 또는 단백질 A를 표적 재조합 단백질에 융합시키는, pGEX(Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) 및 pRIT5 (Pharmacia, Piscataway, N.J.)를 포함한다. 적합한 유도성 비융합 이. 콜라이 발현 벡터의 예는 pTrc (Amrann et al., (1988) Gene 69:301-315) 및 pET 11d (Studier et al., GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89)를 포함한다. 일부 실시형태에서, 벡터는 효모 발현 벡터이다. 효모 사카로마이세스 세레비시에에서 발현을 위한 벡터의 예는 pYepSec1 (Baldari, et al., 1987. EMBO J. 6: 229-234), pMFa (Kuijan and Herskowitz, 1982. Cell 30: 933-943), pJRY88 (Schultz et al., 1987. Gene 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.) 및 picZ (InVitrogen Corp, San Diego, Calif.)를 포함한다. 일부 실시형태에서, 벡터는 배큘로바이러스 발현 벡터를 이용하여 곤충 세포에서 단백질 발현을 유도한다. 배양된 곤충 세포 (예를 들어, SF9 세포)에서 단백질 발현에 이용 가능한 배큘로바이러스 벡터는 pAc 시리즈 (Smith, et al., 1983. Mol. Cell. Biol. 3: 2156-2165) 및 pVL 시리즈 (Lucklow and Summers, 1989. Virology 170: 31-39)를 포함한다.

[0494] 일부 실시형태에서, 벡터는 포유동물 발현 벡터를 이용하여 포유동물 세포에서 하나 이상의 서열의 발현을 유도할 수 있다. 포유동물 발현 벡터의 예는 pCDM8 (Seed, 1987. Nature 329: 840) 및 pMT2PC (Kaufman, et al., 1987. EMBO J. 6: 187-195)를 포함한다. 포유동물 세포에서 사용될 때, 발현 벡터의 제어 기능은 전형적으로 하나 이상의 조절 구성요소에 의해 제공된다. 예를 들어, 통상적으로 사용되는 프로모터는 폴리오마바이러스, 아데노바이러스 2, 거대세포바이러스, 유인원 바이러스 40, 및 본 명세서에 개시되고 당업계에 공지된 다른 것으로부터 유래된다. 원핵 세포와 진핵생물 세포 둘 모두에 대한 다른 적합한 발현 시스템에 대

해, 예를 들어, 문헌 [Chapters 16 and 17 of Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989]을 참조한다.

[0495] 일부 실시형태에서, 재조합 포유동물 발현 벡터는 특정 세포 유형에서 핵산의 발현을 우선적으로 유도할 수 있다 (예를 들어, 조직-특이적 조절 구성요소는 핵산을 발현시키기 위해 사용된다). 조직-특이적 조절 구성요소는 당업계에 공지되어 있다. 적합한 조직-특이적 프로모터의 비제한적 예는 알부민 프로모터(간-특이적; Pinkert, et al., 1987. *Genes Dev.* 1: 268-277), 림구-특이적 프로모터 (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), 특히 T 세포 수용체의 프로모터 (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) 및 면역글로불린 (Baneiji, et al., 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), 뉴런-특이적 프로모터 (예를 들어, 신경세사 프로모터; 문헌 [Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477]), 췌장-특이적 프로모터 (Edlund, et al., 1985. *Science* 230: 912-916) 및 유선-특이적 프로모터 (예를 들어, 유청 프로모터; 미국 특허 제4,873,316호 및 유럽 특허 출원 공개 제264,166호)를 포함한다. 발생 조절된 프로모터, 예를 들어, 쥐과 hox 프로모터 (Kessel and Gruss, 1990. *Science* 249: 374-379) 및 α-태아단백질 프로모터 (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546)가 또한 포함된다. 이들 원핵 및 진핵 벡터에 대해, 미국 특허 제6,750,059호가 언급되며, 이의 내용은 본 명세서에 그들의 전문이 참고로 편입된다. 본 발명의 다른 실시형태는 바이러스 벡터의 용도에 관한 것일 수 있으며, 미국 특허 출원 제13/092,085호가 언급되고, 이의 내용은 본 명세서에 전문이 참고로 편입된다.

조직-특이적 조절 구성요소는 당업계에 공지되어 있고, 이와 관련하여, 미국 특허 제7,776,321호가 언급되며, 이의 내용은 본 명세서에서 그들의 전문이 참고로 편입된다. 일부 실시형태에서, 조절 구성요소는 CRISPR 시스템의 하나 이상의 요소의 발현을 유도하기 위해 CRISPR 시스템의 하나 이상의 요소에 작동적으로 연결된다.

[0496] 일부 실시형태에서, 핵산-표적화 시스템의 하나 이상의 구성요소의 발현을 구동하는 하나 이상의 벡터는 핵산-표적화 시스템의 구성요소의 발현이 하나 이상의 표적 부위에서 핵산-표적화 복합체의 형성을 유도하도록 숙주 세포에 도입된다. 예를 들어, 핵산-표적화 이펙터 효소 및 핵산-표적화 가이드 RNA는 별개의 벡터 상에서 별개의 조절 구성요소에 각각 작동가능하게 연결될 수 있다. 핵산-표적화 시스템의 RNA(들)는 유전자이식 핵산-표적화 이펙터 단백질 동물 또는 포유동물, 예를 들어, 핵산-표적화 이펙터 단백질을 구성적으로 또는 유도적으로 또는 조건적으로 발현시키는 동물 또는 포유동물; 또는 핵산-표적화 이펙터 단백질을 달리 발현시키거나 또는 핵산-표적화 이펙터 단백질을 함유하는 세포를 갖는 동물 또는 포유동물에, 예컨대, 생체내 핵산-표적화 이펙터 단백질을 코딩하고 이를 발현시키는 벡터 또는 벡터들의 사전 투여에 의해 전달될 수 있다. 대안적으로, 동일 또는 상이한 조절 구성요소로부터 발현된 구성요소 중 둘 이상은 단일 벡터에서, 제1 벡터에 포함되지 않는 핵산-표적화 시스템의 임의의 성분을 제공하는 하나 이상의 추가적인 벡터와 함께 조합될 수 있다.

단일 벡터에서 조합되는 핵산-표적화 시스템 구성요소는 임의의 적합한 배향으로 배열될 수 있으며, 예컨대, 하나의 구성요소는 제2 구성요소 (의 "상류")에 대해 5' 또는 ("하류")에 대해 3'에 위치된다. 한 구성요소의 코딩 서열은 제2 구성요소의 코딩 서열의 동일하거나 또는 반대쪽 가닥 상에 위치되고, 동일 또는 반대 방향으로 배향될 수 있다. 일부 실시형태에서, 단일 프로모터는 하나 이상의 인트론 서열 내에 (예를 들어, 상이한 인트론에 각각, 적어도 하나의 인트론에 2 이상, 또는 단일 인트론에 모두) 함입된 핵산-표적화 이펙터 단백질 및 핵산-표적화 가이드 RNA를 코딩하는 전사체의 발현을 유도한다. 일부 실시형태에서, 핵산-표적화 이펙터 단백질 및 핵산-표적화 가이드 RNA는 동일한 프로모터에 작동가능하게 연결되고 동일한 프로모터로부터 발현될 수 있다. 핵산-표적화 시스템의 하나 이상의 구성요소의 발현을 위한 전달 비히클, 벡터, 입자, 나노입자, 제형 및 이들의 성분은 앞서 언급한 문헌, 예컨대, WO 2014/093622 (PCT/US2013/074667)에서 사용되는 바와 같다. 일부 실시형태에서, 벡터는 하나 이상의 삽입 부위, 예컨대, 제한 엔도뉴클레아제 인식 서열 (또한 "클로닝 부위"라고도 함)을 포함한다. 일부 실시형태에서, 하나 이상의 삽입 부위 (예를 들어, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 삽입 부위)가 하나 이상의 벡터의 하나 이상의 서열 구성요소의 상류 및/또는 하류에 위치된다. 다중의 상이한 가이드 서열이 사용될 때, 세포 내에서 다중의 상이한, 상응하는 표적 서열에 대해 핵산-표적화 활성을 표적화하기 위해 단일 발현 구성체가 사용될 수 있다. 예를 들어, 단일 벡터는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20개 이상의 가이드 서열을 포함할 수 있다. 일부 실시형태에서, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 이러한 가이드-서열-함유 벡터가 제공될 수 있으며, 선택적으로 세포에 전달된다. 일부 실시형태에서, 벡터는 핵산-표적화 이펙터 단백질을 코딩하는 효소-코딩 서열에 작동적으로 연결된 조절 구성요소를 포함한다. 핵산-표적화 이펙터 단백질 또는 핵산-표적화 가이드 RNA 또는 RNA(들)는 별개로 전달될 수 있으며; 그리고 유리하게는 이들 중 적어도 하나는 입자 복합체를 통해

전달된다. 핵산-표적화 이펙터 단백질 mRNA는 핵산-표적화 이펙터 단백질이 발현될 시간을 제공하기 위해 핵산-표적화 가이드 RNA 이전에 전달될 수 있다. 핵산-표적화 이펙터 단백질 mRNA는 핵산-표적화 가이드 RNA의 투여 전에 1 내지 12시간 (바람직하게는 대략 2 내지 6시간)에 투여될 수 있다. 대안적으로, 핵산-표적화 이펙터 단백질 mRNA 및 핵산-표적화 가이드 RNA는 함께 투여될 수 있다. 유사하게는, 가이드 RNA의 제2 부스터 용량은 핵산-표적화 이펙터 단백질 mRNA + 가이드 RNA의 초기 투여 후 1 내지 12시간 (바람직하게는 대략 2 내지 6시간)에 투여될 수 있다. 핵산-표적화 이펙터 단백질 mRNA 및/또는 가이드 RNA의 추가 투여는 게놈 변형의 가장 효율적인 수준을 달성하는 데 유용할 수 있다.

[0497] 통상적인 바이러스 및 비바이러스 기반 유전자 전달 방법은 포유동물 세포 또는 표적 조직에서 핵산을 도입하는 데 사용될 수 있다. 이러한 방법은 배양물에서 또는 숙주 세포 유기체에서 세포에 핵산-표적화 시스템의 성분을 코딩하는 핵산을 투여하는 데 사용될 수 있다. 비-바이러스 벡터 전달 시스템은 DNA 플라스미드, RNA (예를 들어, 본 명세서에 기재된 벡터의 전사체), 나형 핵산, 및 전달 비히클, 예컨대, 리포솜과 복합체화된 핵산을 포함한다. 바이러스 벡터 전달 시스템은 세포에 전달 후에 에피솜 또는 통합 게놈 중 하나를 갖는 DNA 및 RNA 바이러스를 포함한다. 유전자 요법 절차의 검토를 위해, 다음의 문헌들을 참조한다: Anderson, Science 256:808-813 (1992); Nabel & Felgner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Dillon, TIBTECH 11:167-175 (1993); Miller, Nature 357:455-460 (1992); Van Brunt, Biotechnology 6(10):1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada et al., Current Topics in Microbiology and Immunology, Doerfler and Bohm (eds) (1995); 및 Yu et al., Gene Therapy 1:13-26 (1994).

[0498] 핵산의 비바이러스 전달 방법은 리포펙션, 뉴클레오펙션, 미량주사법, 유전자 총법, 바이로좀, 리포솜, 면역리포솜, 다양이온 또는 지질:핵산 접합체, 나형 DNA, 인공 비리온, 및 DNA의 제제 증강된 흡수를 포함한다. 리포펙션은, 예를 들어, 미국 특허 제5,049,386호, 제4,946,787호; 및 4,897,355호에 기재되어 있으며, 리포펙션 시약은 상업적으로 시판되고 있다 (예를 들어, Transfectam™ 및 Lipofectin™). 폴리뉴클레오티드의 효율적인 수용체-인식 리포펙션에 적합한 양이온성 및 중성 지질은 Felgner의 WO 91/17424; WO 91/16024의 것을 포함한다. 전달은 세포 (예를 들어, 시험관내 또는 생체외 투여) 또는 표적 조직 (예를 들어, 생체내 투여)에 대한 것일 수 있다.

[0499] 플라스미드 전달은 CRISPR-Cas 단백질 발현 플라스미드 내로 가이드 RNA의 클로닝 및 세포 배양물에서의 DNA 형질감염에 의해 수행된다. 플라스미드 골격은 상업적으로 입수 가능하며, 특정 장비가 필요하지 않다. 그들은 상이한 크기의 CRISPR-Cas 코딩 서열 (더 큰 크기의 단백질을 코딩하는 것을 포함)뿐만 아니라 선택 마커를 운반할 수 있는 모듈의 이점을 갖는다. 플라스미드의 이점은 둘 모두 그들이 일시적이지만, 지속적 발현을 보장할 수 있다는 것이다. 그러나, 플라스미드의 전달은 간단하지 않으므로, 생체내 효율이 종종 낮아진다. 지속 발현은 또한 오프-표적 편집을 증가시킬 수 있다는 점에서 불리할 수 있다. 추가로, CRISPR-Cas 단백질의 과량의 증강은 세포에 대해 독성이 될 수 있다. 마지막으로, 플라스미드는 보다 특히 이중-가닥 파괴가 생성된다는 관점에서 (온 및 오프-표적), 숙주 게놈 내에서 dsDNA의 무작위 통합 위험성을 항상 보유한다.

[0500] 표적화된 리포솜을 비롯한 지질:핵산 복합체, 예컨대, 면역지질 복합체의 제조는 당업자에게 잘 공지되어 있다 (예를 들어, 문헌 [Crystal, Science 270:404-410 (1995)]; [Blaese et al., Cancer Gene Ther. 2:291-297 (1995)]; [Behr et al., Bioconjugate Chem. 5:382-389 (1994)]; [Remy et al., Bioconjugate Chem. 5:647-654 (1994)]; [Gao et al., Gene Therapy 2:710-722 (1995)]; [Ahmad et al., Cancer Res. 52:4817-4820 (1992)]; 미국 특허 제4,186,183호, 제4,217,344호, 제4,235,871호, 제4,261,975호, 제4,485,054호, 제4,501,728호, 제4,774,085호, 제4,837,028호 및 4,946,787호 참조). 이는 이하에서 더욱 상세하게 논의된다.

[0501] 핵산의 전달을 위한 RNA 또는 DNA 바이러스 기반 시스템의 사용은 체내 특정 세포에 대해 바이러스를 표적화하고 핵에 바이러스 페이로드를 수송하는 고도로 진화된 과정을 이용한다. 바이러스 벡터는 환자에게 (생체내) 직접적으로 투여될 수 있거나 또는 그들은 시험관내에서 세포를 처리하는 데 사용될 수 있고, 변형된 세포는 선택적으로 환자에게 (생체외) 투여될 수 있다. 통상적인 바이러스 기반 시스템은 유전자 전달을 위해 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 및 단순포진 바이러스 벡터를 포함할 수 있었다. 숙주 게놈 내 통합은 종종 삽입 이식유전자의 장기간 발현을 초래하는 레트로바이러스, 렌티바이러스, 및 아데노-연합 바이러스 유전자 전달 방법에 의해 가능하다. 추가적으로, 다수의 상이한 세포 유형 및 표

적 조직에서 높은 형질도입 효율이 관찰되었다.

[0502] 레트로바이러스의 향성은 외래 외피 단백질을 혼입시켜, 표적 세포의 잠재적 표적 집단을 확장시킴으로써 변경될 수 있다. 렌티바이러스 벡터는 비-분화 세포를 형질도입 또는 감염시키고 전형적으로 고 바이러스 역가를 생성하는 레트로바이러스 벡터이다. 따라서 레트로바이러스 유전자 전달 시스템의 선택은 표적 조직에 의존한다. 레트로바이러스 벡터는 최대 6 내지 10 kb의 외래 서열에 대한 패키징 능력을 갖는 시스-작용성 긴 말단 반복부를 포함한다. 최소 시스-작용성 LTR은 벡터의 복제 및 패키징에 충분하며, 이어서, 표적 세포 내로 치료 유전자를 통합시키기 위해 사용되어 영구 이식유전자 발현을 제공한다. 널리 사용되는 레트로바이러스 벡터는 쥐과 백혈병 바이러스(MuLV), 긴본 유인원 백혈병 바이러스 (GaLV), 유인원 면역결핍 바이러스 (SIV), 인간 면역결핍 바이러스 (HIV) 및 이들의 조합에 기반하는 것을 포함한다 (예를 들어, 문헌 [Buchscher et al., J. Virol. 66:2731-2739 (1992)]; [Johann et al., J. Virol. 66:1635-1640 (1992)]; [Sommnerfelt et al., Virol. 176:58-59 (1990)]; [Wilson et al., J. Virol. 63:2374-2378 (1989)]; [Miller et al., J. Virol. 65:2220-2224 (1991)]; PCT/US94/05700 참조).

[0503] 일시적 발현이 바람직한 적용에서, 아데노바이러스 기반 시스템이 사용될 수 있다. 아데노바이러스 기반 벡터는 다수의 세포 유형에서 매우 높은 형질도입 효율을 가능하게 하고, 세포 분할을 필요로 하지 않는다. 이러한 벡터에 의해, 높은 발현 역가 및 수준이 얻어졌다. 이 벡터는 상대적으로 단순한 시스템에서 다량으로 생성될 수 있다. 핵산, 예를 들어, 핵산 및 펩티드의 시험관내 생성에서, 그리고 생체내 및 생체외 유전자 요법 절차에 대해 표적 핵산을 갖는 세포를 형질도입하는데 아데노-연합 바이러스 ("AAV") 벡터가 또한 사용될 수 있다 (예를 들어, 문헌 [West et al., Virology 160:38-47 (1987)]; 미국 특허 제 4,797,368호; WO 93/24641; [Kotin, Human Gene Therapy 5:793-801(1994)]; [Muzyczka, J. Clin. Invest. 94:1351(1994)] 참조). 재조합 AAV 벡터의 구축은 미국 특허 제5,173,414호; 문헌 [Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985)]; [Tratschin, et al., Mol. Cell. Biol. 4:2072-2081(1984)]; [Hermonat & Muzyczka, PNAS 81:6466-6470 (1984)]; 및 [Samulski et al., J. Virol. 63:03822-3828 (1989)]을 포함하는 다수의 간행물에 기재되어 있다.

[0504] 본 발명은 CRISPR 시스템을 코딩하는 외생성 핵산 분자, 예를 들어 프로모터, CRISPR-연관 (Cas) 단백질 (추정 뉴클레아제 또는 헬리카제 단백질), 예를 들어 Cas13을 코딩하는 핵산 분자, 및 종결자를 포함하거나 또는 그로 본질적으로 이루어지는 제1 카세트를 포함하거나 또는 그로 이루어지는 다수의 카세트, 및 하나 이상의, 유리하게는 벡터의 최대 패키징 크기 한계까지, 예를 들어 전체 (제1 카세트 포함) 5개에, 가이드 RNA (gRNA)를 코딩하는 핵산 분자 및 종결자를 포함하거나 또는 그로 본질적으로 이루어지는 카세트 (예를 들어, 각각의 카세트는 개략적으로 프로모터-gRNA1-종결자, 프로모터-gRNA2-종결자 ... 프로모터-gRNA(N)-종결자로서 표시되고, 여기서 N은 벡터의 패키징 크기 한계의 상한치에서 삽입될 수 있는 개수임), 또는, 각각 하나 또는 하나 초과의 CRISPR 시스템을 함유하는, 둘 이상의 개별 rRNA, 예를 들어, 프로모터, Cas, 예를 들어 Cas (Cas13)을 코딩하는 핵산 분자 및 종결자를 포함하거나 또는 그로 본질적으로 이루어지는 제1 카세트를 함유하는 제1 rAAV, 및 각각 프로모터, 가이드 RNA (gRNA) 및 종결자를 포함하거나 또는 그로 본질적으로 이루어지는 하나 이상의 카세트를 함유하는 제2 rAAV (예를 들어, 각각의 카세트는 개략적으로 프로모터-gRNA1-종결자, 프로모터-gRNA2-종결자...프로모터-gRNA(N)-종결자로서 표시되고, 여기서 N은 벡터의 패키징 크기 한계의 상한치에서 삽입될 수 있는 개수임)를 함유하거나 또는 그로 본질적으로 이루어지는 AAV를 제공한다. 대안적으로, Cas13이 그 자체의 crRNA/gRNA를 프로세싱하기 때문에, 단일 crRNA/gRNA 어레이가 다중복합 유전자 편집에 사용될 수 있다. 그러한 이유로, gRNA를 전달하기 위한 다수 카세트를 포함하는 대신에, rAAV는 프로모터, 다수의 crRNA/gRNA, 및 종결자를 포함하거나 또는 그로 본질적으로 이루어지는 단일 카세트를 함유할 수도 있다 (예를 들어, 개략적으로 프로모터-gRNA1-종결자, 프로모터-gRNA2-종결자...프로모터-gRNA(N)-종결자로 표시되고, 여기서 N은 벡터의 패키징 크기 한계의 상한치에서 삽입될 수 있는 개수임). 그 전문이 참조로 본 명세서에 편입되는, [Zetsche et al Nature Biotechnology 35, 31-34 (2017)]를 참조한다. rAAV가 DNA 바이러스이므로, AAV 또는 rAAV에 관해 본 명세서에서 논의되는 핵산 분자는 유리하게는 DNA이다. 프로모터는 일부 실시형태에서, 유리하게 인간 시냅신 I 프로모터 (hSyn)이다. 세포에게 핵산의 전달을 위한 추가 방법은 당업자에게 충분히 공지되어 있다. 예를 들어, 참조로 본 명세서에 편입되는 미국 공개 특허 출원 제20030087817호를 참조한다.

[0505] 다른 실시형태에서, 코칼 베시쿨로바이러스 (Cocal vesiculovirus) 외피 위형 레트로바이러스 벡터 입자가 고려된다 (예를 들어, 미국 프레드 허친슨 암 연구 센터(Fred Hutchinson Cancer Research Center)에 양도된 미국 특허 공개 제20120164118호). 코칼 바이러스는 베시쿨로바이러스 속이며, 포유동물에서 수포성 구내염의

원인 물질이다. 코칼 바이러스는 본래 트리니다드에서 진드기로부터 단리되었고 (Jonkers et al., Am. J. Vet. Res. 25:236-242 (1964)), 트리니다드, 브라질 및 아르헨티나에서 곤충, 소 및 말로부터 감염이 확인되었다. 포유동물를 감염시키는 베시쿨로바이러스 중 다수는 자연적으로 감염된 절지동물로부터 단리되었는데, 이는 그들이 매개체 감염이라는 것을 시사한다. 베시쿨로바이러스에 대한 항체는 바이러스가 풍토성이고 실험실 획득된 시골 지역에 살고 있는 사람들 중에서는 일반적이고; 인간에서의 감염은 보통 인플루엔자-유사 증상을 초래한다. 코칼 바이러스 외피 당단백질은 아미노산 수준을 VSV-G 인디애나와 71.5% 동일성으로 공유하며, 베시쿨로바이러스의 외피 유전자의 계통 발생론적 비교는 코칼 바이러스가 베시쿨로바이러스 중에서도 VSV-G 인디애나 균주와 혈청학적으로 별개이지만, 가장 밀접하게 관련된다는 것을 나타낸다. [Jonkers et al., Am. J. Vet. Res. 25:236-242 (1964)] 및 [Travassos da Rosa et al., Am. J. Tropical Med. & Hygiene 33:999-1006 (1984)]. 코칼 베시쿨로바이러스 외피 유사형 레트로바이러스 벡터 입자는, 예를 들어, 레트로바이러스 Gag, Pol, 및/또는 하나 이상의 부속 단백질(들) 및 코칼 베시쿨로바이러스 외피 단백질을 포함할 수 있는 렌티바이러스, 알파레트로바이러스, 베타레트로바이러스, 감마레트로바이러스, 델타레트로바이러스, 및 엡실론레트로바이러스 벡터 입자를 포함할 수 있다. 이들 실시형태의 소정의 양상 내에서, Gag, Pol, 및 부속 단백질은 렌티바이러스 및/또는 감마레트로바이러스이다.

[0506] 일부 실시형태에서, 숙주 세포는 본 명세서에 기재된 하나 이상의 벡터로 일시적으로 또는 비일시적으로 형질감염된다. 일부 실시형태에서, 세포는 선택적으로 세포에 재도입될 대상체에서 자연적으로 발생함에 따라 형질감염된다. 일부 실시형태에서, 형질감염되는 세포는 대상체로부터 취해진다. 일부 실시형태에서, 세포는 대상체, 예컨대, 세포주로부터 취하는 세포로부터 유래된다. 조직 배양을 위한 매우 다양한 세포주는 당업계에 공지되어 있다. 세포주의 예는 C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panc1, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 원숭이 신장 상피, BALB/ 3T3 마우스 배아 섬유아세포, 3T3 Swiss, 3T3-L1, 132-d5 인간 태아 섬유아세포; 10.1 마우스 섬유아세포, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 세포, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr -/-, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, JY 세포, K562 세포, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK II, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN / OPCT 세포주, Peer, PNT-1A / PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 세포, Sf-9, SkBr3, T2, T-47D, T84, THP1 세포주, U373, U87, U937, VCaP, Vero 세포, WM39, WT-49, X63, YAC-1, YAR, 및 이들의 유전자이식 변이체를 포함하지만, 이들로 제한되지 않는다. 세포주는 당업자에게 공지된 다양한 공급원으로부터 입수 가능하다 (예를 들어, 미국 미생물 보존 센터 (American Type Culture Collection)(ATCC)(Manassus, Va.)).

[0507] 특정 실시형태에서, AD-기능화된 CRISPR 시스템 성분 중 하나 이상의 일시적 발현 및/또는 존재는, 예컨대, 오프-표적 효과를 감소시키는 데 관심이 있을 수 있다. 일부 실시형태에서, 본 명세서에 기재된 하나 이상의 벡터로 형질감염된 세포는 하나 이상의 벡터-유래 서열을 포함하는 새로운 세포주를 확립하는 데 사용된다. 일부 실시형태에서, 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템의 성분으로 (예컨대, 하나 이상의 벡터의 일시적 형질감염, 또는 RNA에 의한 형질감염에 의해) 일시적으로 형질감염되고 CRISPR 복합체 활성을 통해 변형된 세포는 변형을 함유하지만 임의의 다른 외생성 서열을 결여하는 세포를 포함하는 새로운 세포주를 확립하는 데 사용된다. 일부 실시형태에서, 본 명세서에 기재된 하나 이상의 벡터, 또는 이러한 세포주로부터 유래된 세포주로 일시적으로 또는 비일시적으로 형질감염된 세포는 하나 이상의 시험 화합물을 평가하는 데 사용된다.

[0508] 일부 실시형태에서, RNA 및/또는 단백질을 숙주 세포에 직접적으로 도입하는 것이 계획된다. 예를 들어, CRISPR-Cas 단백질은 시험관내 전사된 가이드 RNA와 함께 코딩 mRNA로서 전달될 수 있다. 이러한 방법은 CRISPR-Cas 단백질의 효과를 보장하기 위한 시간을 감소시키고, 추가로 CRISPR 시스템 성분의 장기간 발현을 방지할 수 있다.

[0509] 일부 실시형태에서, 본 발명의 RNA 분자는 리포솜 또는 리포펙틴 제형 등으로 전달되고, 당업자에게 잘 공지된

방법에 의해 제조될 수 있다. 이러한 방법은, 예를 들어, 미국 특허 제5,593,972호, 제5,589,466호 및 제5,580,859호에 기재되어 있으며, 이들은 본 명세서에 참고로 편입된다. siRNA의 포유동물 세포 내로의 증강되고 개선된 전달을 구체적으로 목적으로 하는 전달 시스템이 개발되었고, 본 발명에 적용될 수 있다 (예를 들어, 문헌 [Shen et al FEBS Let. 2003, 539:111-114]; [Xia et al., Nat. Biotech. 2002, 20:1006-1010]; [Reich et al., Mol. Vision. 2003, 9: 210-216]; [Sorensen et al., J. Mol. Biol. 2003, 327]: 761-766]; [Lewis et al., Nat. Gen. 2002, 32: 107-108] 및 [Simeoni et al., NAR 2003, 31, 11: 2717-2724] 참조). siRNA는 최근에 영장류에서의 유전자 발현의 억제를 위해 성공적으로 사용되었다 (예를 들어, 또한 본 발명에 적용될 수 있는 문헌 [Tolentino et al., Retina 24(4):660] 참조).

[0510] 사실, RNA 전달은 유용한 생체내 전달 방법이다. 리포솜 또는 나노입자를 이용하여 세포 내로 Cas13, 아데노신 데아미나제, 및 가이드 RNA를 전달하는 것이 가능하다. 따라서, CRISPR-Cas 단백질, 예컨대, Cas13의 전달, 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 어댑터 단백질에 융합될 수 있음)의 전달, 및/또는 본 발명의 RNA 전달은 RNA 형태로 그리고 미세소포체, 리포솜 또는 입자 또는 입자들을 통할 수 있다. 예를 들어, Cas13 mRNA, 아데노신 데아미나제 mRNA 및 가이드 RNA는 생체내 전달을 위해 리포솜 입자 내로 패키징될 수 있다. 리포솜 형질감염 시약, 예컨대, Life Technologies의 Lipofectamine 및 시판 중인 다른 시약은 RNA 분자를 간 내로 효과적으로 전달할 수 있다.

[0511] 또한 바람직한 RNA의 전달 수단은 입자 (Cho, S., Goldberg, M., Son, S., Xu, Q., Yang, F., Mei, Y., Bogatyrev, S., Langer, R. and Anderson, D., Lipid-like nano입자 for small interfering RNA delivery to endothelial cells, Advanced Functional Materials, 19: 3112-3118, 2010) 또는 엑소솜 (Schroeder, A., Levins, C., Cortez, C., Langer, R., and Anderson, D., Lipid-based nanotherapeutics for siRNA delivery, Journal of Internal Medicine, 267: 9-21, 2010, PMID: 20059641)을 통한 RNA의 전달을 포함한다. 사실, 엑소솜은 CRISPR 시스템에 일부 병행하는 시스템인 전달 siRNA에서 특히 유용한 것으로 나타났다. 예를 들어, El-Andaloussi S 등 ("Exosome-mediated delivery of siRNA in vitro and in vivo." Nat Protoc. 2012 Dec;7(12):2112-26. doi: 10.1038/nprot.2012.131. Epub 2012 Nov 15)은 엑소솜이 상이한 생물학적 장벽에 걸쳐 약물 전달을 위한 유망한 도구가 되고 시험관내 및 생체내에서 siRNA의 전달을 위해 이용될 수 있는 방법을 기재한다. 그들의 접근은 펩티드 리간드와 융합된 엑소솜 단백질을 포함하는, 발현 벡터의 형질감염을 통해 표적화된 엑소솜을 생성하는 것이다. 이어서, 엑소솜은 형질감염된 세포 상청액을 정제하고 이로부터 특징 규명되며, 이어서, RNA는 엑소솜에 로딩된다. 본 발명에 따른 전달 또는 투여는 엑소솜, 특히, 제한없이, 뇌에 의해 수행될 수 있다. 비타민 E (α-토코페롤)는 CRISPR Cas와 접합될 수 있고, 뇌에 고밀도 리포단백질 (HDL)과 함께, 예를 들어, 뇌에 짧은-간섭 RNA (siRNA)를 전달하기 위해 문헌 [Uno et al. (HUMAN GENE THERAPY 22:711-719 (June 2011))]에 의해 행해지는 것과 동일한 방식으로 전달될 수 있다. 포스페이트-완충 염수 (PBS) 또는 유리 TocsiBACE 또는 Toc-siBACE/HDL을 추전시키고 Brain Infusion 키트 3 (Alzet)와 연결된 삼투압 미니펌프 (model 1007D; Alzet, Cupertino, CA)를 통해 마우스에 주입된다. 뇌-주입 캐뉼라는 배측 제3 뇌실 내로 주입을 위해 정중선에서 브레그마 뒤쪽 약 0.5㎜에 놓인다. Uno 등은 HDL과 함께 3 nmol 정도로 적은 Toc-siRNA가 동일한 ICV 주입 방법에 의해 비슷한 정도로 표적 감소를 유도할 수 있었다는 것을 발견하였다. α-토코페롤에 접합되고, 뇌에 표적화된 HDL과 공동 투여되는 CRISPR Cas의 유사한 용량이 본 발명의 인간에 대해 고려될 수 있으며, 예를 들어, 약 3 nmol 내지 약 3 μmol의 뇌에 표적화되는 CRISPR Cas가 고려될 수 있다. Zou 등 ((HUMAN GENE THERAPY 22:465-475 (April 2011))은 래트의 척수에서 침묵화된 생체내 유전자에 대해 PKCγ를 표적화하는 짧은-헤어핀 RNA의 렌티바이러스-매개 전달 방법을 기재한다. Zou 등은 척수강내 카테터에 의해 1×10^{9}개의 형질도입 유닛 (TU)/㎖의 역가를 갖는 약 10 ㎕의 재조합 렌티바이러스를 투여하였다. 뇌에 표적화된 렌티바이러스 벡터에서 발현되는 유사한 용량의 CRISPR Cas가 본 발명에서 인간에 대해 고려될 수 있으며, 예를 들어, 1×10^{9}개의 형질도입 유닛 (TU)/㎖의 역가를 갖는 렌티바이러스 중 뇌로 표적화된 약 10 내지 50㎖의 CRISPR Cas가 고려될 수 있다.

[0512] **벡터의 용량**

[0513] 일부 실시형태에서, 벡터, 예를 들어, 플라스미드 또는 바이러스 벡터는 관심 조직에, 예를 들어, 근육내 주사에 의해 전달되는 한편, 다른 시기에 정맥내, 경피, 비강내, 경구, 점막 또는 다른 전달 방법을 통해 전달된다.

이러한 전달은 단일 용량 또는 다수 용량을 통할 수 있다. 당업자는 본 명세서에서 전달하려는 실제 용량이 다양한 인자, 예컨대, 벡터 선택, 표적 세포, 유기체, 또는 조직, 치료하려는 대상체의 일반적 병태, 추구되는 형질전환/변형 정도, 투여 경로, 투여 방식, 추구되는 형질전환/변형 유형 등에 따라 크게 다를 수 있다는

것을 이해한다.

[0514] 이러한 용량은, 예를 들어, 담체 (물, 식염수, 에탄올, 글리세롤, 락토스, 수크로스, 인산칼슘, 젤라틴, 덱스트란, 한천, 펙틴, 땅콩유, 참깨유 등), 희석제, 약학적으로 허용가능한 담체 (예를 들어, 포스페이트-완충 식염수), 약학적으로 허용가능한 부형제 및/또는 당업계에 공지된 다른 화합물을 추가로 함유할 수 있다. 용량은 1종 이상의 약학적으로 허용가능한 염, 예를 들어, 무기산염, 예컨대, 염산염, 브로민화수소산염, 포스페이트, 황산염 등; 그리고 유기산의 염, 아세트산염, 프로피온산염, 말론산염, 벤조산염 등을 추가로 함유할 수 있다. 추가적으로, 보조 물질, 예컨대, 습윤제 또는 유화제, pH 완충 물질, 겔 또는 겔화 물질, 향미제, 착색제, 미세구, 중합체, 현탁제 등이 또한 본 명세서에 제공될 수 있다. 추가로, 특히 제형이 재구성 가능한 형태라면, 1종 이상의 다른 통상적인 약제 성분, 예컨대, 보존제, 습윤제, 현탁제, 계면활성제, 항산화제, 고결방지제, 충전제, 킬레이트제, 코팅제, 화학적 안정제 등이 또한 제공될 수 있다. 적합한 예시적 성분은 미정질 셀룰로스, 카복시메틸셀룰로스 소듐, 폴리솔베이트 80, 페닐에틸 알콜, 클로로부탄올, 솔브산칼륨, 솔브산, 이산화황, 갈산프로필, 파라벤, 에틸 바닐린, 글리세린, 페놀, 파라클로로페놀, 젤라틴, 알부민 및 이들의 조합물을 포함한다. 약학적으로 허용가능한 부형제의 철저한 논의는 본 명세서에 참고로 편입된 문헌 [REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991)]에서 이용 가능하다.

[0515] 본 명세서의 실시형태에서, 용량은 본 명세서의 실시형태에서, 전달은 아데노바이러스 벡터의 적어도 1×10^{5}개의 입자 (또한 입자 단위인 pu라고 함)를 함유하는 단일 부스터 용량일 수 있는 아데노바이러스를 통한다. 바람직하게는 적어도 약 1×10^{6} 개의 입자 (예를 들어, 약 1×10^{6} 내지 1×10^{12}개의 입자), 더 바람직하게는 적어도 약 1×10^{7} 입자, 더 바람직하게는 적어도 약 1×10^{8}개의 입자 (예를 들어, 약 1×10^{8} 내지 1×10^{11} 개의 입자 또는 약 1×10^{8} 내지 1×10^{12}개의 입자), 및 가장 바람직하게는 적어도 약 1×100 개의 입자 (예를 들어, 약 1×10^{9} 내지 1×10^{10} 개의 입자 또는 약 1×10^{9} 내지 1×10^{12}개의 입자), 또는 심지어 적어도 약 1×10^{10} 개의 입자 (예를 들어, 약 1×10^{10} 내지 1×10^{12}개의 입자)의 아데노바이러스 벡터이다. 대안적으로, 상기 용량은 약 1×10^{14}개 이하의 입자, 바람직하게는 약 1×10^{13}개 이하의 입자, 훨씬 더 바람직하게는 약 1×10^{12}개 이하의 입자, 훨씬 더 바람직하게는 약 1×10^{11}개 이하의 입자, 및 가장 바람직하게는 약 1×10^{10}개 이하의 입자 (예를 들어, 약 1×10^{9}개 이하의 입자)를 포함한다. 따라서, 상기 용량은 아데노바이러스 벡터의 단일 용량을 함유하며, 예를 들어, 아데노바이러스 벡터의 약 1×10^{6}개의 입자 단위 (pu), 약 2×10^{6} pu, 약 4×10^{6} pu, 약 1×10^{7} pu, 약 2×10^{7} pu, 약 4×10^{7} pu, 약 1×10^{8} pu, 약 2×10^{8} pu, 약 4×10^{8} pu, 약 1×10^{9} pu, 약 2×10^{9} pu, 약 4×10^{9} pu, 약 1×10^{10} pu, 약 2×10^{10} pu, 약 4×10^{10} pu, 약 1×10^{11} pu, 약 2×10^{11} pu, 약 4×10^{11} pu, 약 1×10^{12} pu, 약 2×10^{12} pu, 또는 약 4×10^{12} pu이다. 예를 들어, 본 명세서에 참고로 편입된 2013년 6월 4일자로 등록된 Nabel 등의 미국 특허 제8,454,972 B2호의 아데노바이러스 벡터; 및 이의 29 칼럼, 36 내지 58 라인의 용량을 참조한다. 본 명세서의 실시형태에서, 아데노바이러스는 다수 용량을 통해 전달된다.

[0516] 본 명세서의 실시형태에서, 상기 전달은 AAV를 통한다. 인간에 대한 AAV의 생체내 전달을 위한 치료적 유효 용량은 약 1×10^{10} 내지 약 1×10^{10}개의 기능성 AAV/㎖ 용액을 함유하는 식염수 용액의 약 20 내지 약 50 ㎖ 범위인 것으로 여겨진다. 용량은 임의의 부작용에 대해 치료적 이점의 균형을 맞추도록 조절될 수 있다. 본 명세서의 실시형태에서, AAV 용량은 일반적으로 약 1×10^{5} 내지 1×10^{50}개의 게놈 AAV, 약 1×10^{8} 내지 1×10^{20} 개의 게놈 AAV, 약 1×10^{10} 내지 약 1×10^{16}개의 게놈, 또는 약 1×10^{11} 내지 약 1×10^{16}개의 게놈 AAV 농도 범위이다. 인간 용량은 약 1×10^{13}개의 게놈 AAV일 수 있다. 이러한 농도는 약 0.001 ㎖ 내지 약 100 ㎖, 약 0.05 내지 약 50 ㎖, 또는 약 10 내지 약 25 ㎖의 담체 중에 전달될 수 있다. 다른 유효 용량은 용량 반응 곡선을 확립하는 일상적 시행을 통해 당업자에 의해 용이하게 확립될 수 있다. 예를 들어, 2013년 3월 26일자로 등록된 Hajjar 등의 미국 특허 제8,404,658 B2호, 27 칼럼, 45 내지 60행 참조한다.

[0517] 본 명세서의 실시형태에서, 전달은 플라스미드를 통한다. 이러한 플라스미드 조성물에서, 용량은 반응을 유발하는 충분한 양의 플라스미드이어야 한다. 예를 들어, 플라스미드 조성물 내 적합한 양의 플라스미드 DNA는 70 kg 개체당 약 0.1 ㎎ 내지 약 2 ㎎, 또는 약 1 ㎍ 내지 약 10 ㎍일 수 있다. 본 발명의 플라스미드는 일반적으로 (i) 프로모터; (ii) 상기 프로모터에 작동적으로 연결된 CRISPR-Cas 단백질을 코딩하는 서열; (iii)

선택 가능한 마커; (iv) 복제 기원; 그리고 (v) (ii)의 하류에 작동적으로 연결된 전사 종결자를 포함할 것이다. 플라스미드는 또한 CRISPR 복합체의 RNA 성분을 코딩할 수 있지만, 이들 중 하나 이상은 대신에 상이한 벡터 상에서 코딩될 수 있다.

[0518] 본 명세서의 용량은 평균 70 kg의 개체를 기준으로 한다. 투여 빈도는 의학 또는 수의학 종사자(예를 들어, 의사, 수의사) 또는 당업계의 숙련된 과학자의 영역 내이다. 또한 실험에서 사용되는 마우스는 전형적으로 약 20g이고, 마우스 실험으로부터 70 kg까지의 개체로 확대될 수 있다는 것이 주목된다.

[0519] 본 명세서에 제공되는 조성물에 대해 사용되는 용량은 반복된 투여 또는 반복 용량을 위한 용량을 포함한다. 특정 실시형태에서, 투여는 몇 주, 몇 개월 또는 몇 년의 기간 내에 반복된다. 적합한 어세이는 최적의 용량 요법을 얻기 위해 수행될 수 있다. 반복 투여는 오프-표적 변형에 긍정적으로 영향을 미칠 수 있는 더 낮은 용량의 사용을 허용할 수 있다.

[0520] **RNA 전달**

[0521] 특정 실시형태에서, RNA 기반 전달이 사용된다. 이들 실시형태에서, CRISPR-Cas 단백질의 mRNA, 아데노신 데아미나제의 mRNA(CRISPR-Cas 단백질 또는 어댑터에 융합될 수 있음)는 시험관내 전사된 가이드 RNA와 함께 전달된다. Liang 등은 RNA 기반 전달을 이용하는 효율적인 게놈 편집을 기재한다 (Protein Cell. 2015 May; 6(5): 363-372). 일부 실시형태에서, Cas13 및/또는 아데노신 데아미나제를 코딩하는 mRNA(들)는 화학적으로 변형될 수 있는데, 이는 플라스미드-코딩된 Cas13 및/또는 아데노신 데아미나제에 비해 개선된 활성을 야기할 수 있다. 예를 들어, mRNA(들)에서 우리딘은 슈도우리딘 (Ψ), N^1-메틸슈도우리딘 ($me^1Ψ$), 5-메톡시우리딘 (5moU)으로 부분적으로 또는 완전히 치환될 수 있다. 문헌 [Li *et al.*, *Nature Biomedical Engineering* 1, 0066 DOI:10.1038/s41551-017-0066 (2017)]을 참조하며, 이의 전문은 본 명세서에 참고로 편입된다.

[0522] **RNP 전달**

[0523] 특정 실시형태에서, 사전-복합체화된 가이드 RNA, CRISPR-Cas 단백질 및 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 어댑터에 융합될 수 있음)는 리보핵단백질 (RNP)로서 전달된다. RNP는 RNA 방법보다 훨씬 더 빠른 편집 효과를 야기한다는 장점을 갖는데, 이 과정이 전사에 대한 필요를 피하기 때문이다. 중요한 잠점은 RNP 전달이 둘 모두 일시적이어서, 오프-표적 효과 및 독성 문제를 감소시킨다는 것이다. 문헌 [Kim et al. (2014, Genome Res. 24(6):1012-9)], [Paix et al. (2015, Genetics 204(1):47-54)], [Chu et al. (2016, BMC Biotechnol. 16:4)], 및 [Wang et al. (2013, Cell. 9;153(4):910-8)]에 의해 상이한 세포 유형에서의 효율적인 게놈 편집이 관찰되었다.

[0524] 특정 실시형태에서, 리보핵단백질은 WO2016161516에 기재된 바와 같은 폴리펩티드-기반 셔틀 작용제의 방법에 의해 전달된다. WO2016161516은 세포 침투성 도메인 (CPD)에, 히스티딘-풍부 도메인 및 CPD에 작동적으로 연결된 엔도솜 누출 도메인 (ELD)을 포함하는 합성 펩티드를 이용하는 폴리펩티드 카고의 효율적인 형질도입을 기재한다. 유사하게, 이들 폴리펩티드는 진핵생물 세포에서 CRISPR-이펙터 기반 RNP의 전달을 위해 사용될 수 있다.

[0525] **입자**

[0526] 일부 양상 또는 실시형태에서, 전달 입자 제형을 포함하는 조성물이 사용될 수 있다. 일부 양상 또는 실시형태에서, 제형은 CRISPR 복합체를 포함하며, 상기 복합체는 표적 서열에 대한 CRISPR 복합체의 서열 특이적 결합을 지시하는 가이드 및 CRISPR 단백질을 포함한다. 일부 실시형태에서, 전달 입자는 지질계 입자, 선택적으로 지질 나노입자, 또는 양이온성 지질 및 선택적으로 생분해성 중합체를 포함한다. 일부 실시형태에서, 양이온성 지질은 1,2-다이올레오일-3-트라이메틸암모늄-프로판 (DOTAP)을 포함한다. 일부 실시형태에서, 친수성 중합체는 에틸렌 글리콜 또는 폴리에틸렌 글리콜을 포함한다. 일부 실시형태에서, 전달 입자는 지질단백질, 바람직하게는 콜레스테롤을 추가로 포함한다. 일부 실시형태에서, 전달 입자는 직경이 500 nm 미만, 선택적으로 직경이 250 nm 미만, 선택적으로 직경이 100 nm 미만, 선택적으로 직경이 약 35 nm 내지 약 60 nm이다.

[0527] 예시적인 입자 전달 복합체는 2017년 4월 14일자로 출원된 미국 가출원 (발명의 명칭: "Nove Delivery of Large Payloads") 제62/485,625호에 추가로 개시되어 있다.

[0528] 입자 전달 시스템 및/또는 제형의 몇몇 유형은 다양한 범위의 생의학적 적용에서 유용한 것으로 알려져 있다.

일반적으로, 입자는 그의 수송 및 특성에 관해 전체 단위로서 거동하는 소형 물체로서 정의된다. 입자는 추가로 직경에 따라 분류된다. 조악한 입자는 2,500 내지 10,000 나노미터 범위를 포괄한다. 미세한 입자는 크기가 100 내지 2,500 나노미터이다. 초미세 입자, 또는 나노입자는 일반적으로 크기가 1 내지 100 나노미터이다. 100 ㎚ 한계 기준은 벌크 물질을 입자와 구별하는 신규 성질이 전형적으로 100 ㎚ 이하의 임계 길이 규모에서 발생된다는 사실이다.

[0529] 본 명세서에서 사용되는 입자 전달 시스템/제형은 본 발명에 따른 입자를 포함하는 임의의 생물학적 전달 시스템/제형으로서 정의된다. 본 발명에 따른 입자는 100 마이크론 (㎛) 미만의 최고 치수 (즉, 직경)을 갖는 임의의 독립체이다. 일부 실시형태에서, 본 발명의 입자는 10 ㎛ 미만의 최고 치수를 갖는다. 일부 실시형태에서, 본 발명의 입자는 2000 나노미터 (㎚) 미만의 최고 치수를 갖는다. 일부 실시형태에서, 본 발명의 입자는 1000 나노미터 (㎚) 미만의 최고 치수를 갖는다. 일부 실시형태에서, 본 발명의 입자는 900 ㎚, 800 ㎚, 700 ㎚, 600 ㎚, 500 ㎚, 400 ㎚, 300 ㎚, 200 ㎚, 또는 100 ㎚ 미만의 최고 치수를 갖는다. 전형적으로, 본 발명의 입자는 500 ㎚ 이하의 최고 치수 (예를 들어, 직경)를 갖는다. 일부 실시형태에서, 본 발명의 입자는 250 ㎚ 이하의 최고 치수 (예를 들어, 직경)를 갖는다. 일부 실시형태에서, 본 발명의 입자는 200 ㎚ 이하의 최고 치수 (예를 들어, 직경)를 갖는다. 일부 실시형태에서, 본 발명의 입자는 150 ㎚ 이하의 최고 치수 (예를 들어, 직경)를 갖는다. 일부 실시형태에서, 본 발명의 입자는 100 ㎚ 이하의 최고 치수 (예를 들어, 직경)를 갖는다. 예를 들어, 50 ㎚ 이하의 최고 치수를 갖는 보다 작은 입자가 본 발명의 일부 실시형태에서 사용된다. 일부 실시형태에서, 본 발명의 입자는 25 ㎚ 내지 200 ㎚의 범위의 최고 치수를 갖는다.

[0530] 본 발명에 관해, 나노입자 또는 지질 외피를 이용하여 전달되는 CRISPR 복합체, 예를 들어, CRISPR-Cas 단백질 또는 mRNA, 또는 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 어댑터에 융합될 수 있음) 또는 mRNA, 또는 가이드 RNA 중 하나 이상의 성분을 갖는 것이 바람직하다. 다른 전달 시스템 또는 벡터는 본 발명의 나노입자 양상과 함께 사용될 수 있다.

[0531] 일반적으로, "나노입자"는 1000 ㎚ 미만의 직경을 갖는 임의의 입자를 의미한다. 바람직한 일정 실시형태에서, 본 발명의 나노입자는 500 ㎚ 이하의 가장 큰 치수(예를 들어, 직경)를 갖는다. 다른 바람직한 실시형태에서, 본 발명의 나노입자는 25 ㎚ 내지 200 ㎚ 범위의 가장 큰 치수를 갖는다. 다른 바람직한 실시형태에서, 본 발명의 나노입자는 100 ㎚ 이하의 가장 큰 치수를 갖는다. 다른 바람직한 실시형태에서, 본 발명의 나노입자는 35 ㎚ 내지 60 ㎚ 범위의 가장 큰 치수를 갖는다. 입자 또는 나노입자에 대한 언급은 적절한 경우, 상호 호환 가능할 수 있다는 것이 인식될 것이다.

[0532] 입자의 크기는 로딩 전에 측정되는지 또는 로딩 후에 측정되는지에 따라 상이하다는 것이 이해될 것이다. 따라서, 특정 실시형태에서, 용어 "나노입자"는 로딩 전 입자에 대해서만 적용될 수 있다.

[0533] 본 발명에 포함되는 나노입자는 상이한 형태로, 예를 들어, 고체 나노입자 (예를 들어, 금속, 예컨대, 은, 금, 철, 티타늄), 비금속, 지질계 고체, 중합체), 나노입자의 현탁액, 또는 이들의 조합물로 제공될 수 있다. 금속, 유전체, 및 반도체 나노입자뿐만 아니라 하이브리드 구조 (예를 들어, 코어 쉘 나노입자)로서 제조될 수 있다. 반도전성 물질로 만들어진 나노입자는 또한 그들이 전자 에너지 수준의 정량화가 일어나기에 충분히 작다면 (전형적으로 10 ㎚ 미만) 표지된 양자점일 수 있다. 이러한 나노규모 입자는 생의학적 적용분야에서 약물 담체 또는 영상화제로서 사용되며, 본 발명에서 유사한 목적을 위해 적합하게 될 수 있다.

[0534] 반고체 및 연질 나노입자가 제조되었고, 본 발명의 범주 내이다. 반고체 특성의 원형 나노입자는 리포솜이다. 다양한 유형의 리포솜 나노입자는 현재 항암 약물 및 백신을 위한 전달 시스템으로서 임상적으로 사용된다. 하나의 절반은 친수성이고 다른 절반은 소수성인 나노입자는 야누스 (Janus) 입자로 칭해지며, 안정화 에멀션에 특히 효과적이다. 그들은 물/오일 계면에서 자기-조립될 수 있고, 고체 계면활성제로서 작용할 수 있다.

[0535] 입자 특징규명 (예를 들어, 형태, 치수 등의 특징규명을 포함)은 다양한 상이한 기법을 사용하여 행해진다. 통상적인 기법은 전자 현미경 (TEM, SEM), 원자력 현미경 (AFM), 동적광산란 (DLS), X-선 광전자 분광법 (XPS), 분말 X-선 회절 (XRD), 푸리에 변환 적외선 분광법 (FTIR), 매트릭스-보조 레이저 탈착/이온화 시간 비행 질량분석법 (MALDI-TOF), 자외선-가시광선 분광학, 이중 극성화 간섭계법 및 핵자기공명 (NMR)이다. 특징규명 (치수 측정)은 본 발명의 임의의 시험관내 생체외 및/또는 생체내 적용을 위해서 전달을 위한 최적 크기의 입자를 제공하기 위해, 천연 입자 (즉, 로딩 전)로서 또는 카고의 로딩 이후 (여기서 카고는 예를 들어, CRISPR-Cas 시스템 중 하나 이상의 성분, 예를 들어, CRISPR-Cas 단백질 또는 mRNA, 아데노신 데아미나제

(CRISPR-Cas 단백질 또는 어댑터에 융합될 수 있음) 또는 mRNA, 또는 가이드 RNA, 또는 이들의 임의의 조합물을 의미하고, 추가 담체 및/또는 부형제를 포함할 수 있음)에 수행될 수 있다. 바람직한 일정 실시형태에서, 입자 치수 (예를 들어, 직경) 특징규명은 동적 레이저 산란 (DLS)을 이용하는 측정에 기반한다. 입자, 및 이의 제조 및 이용 방법 및 이들의 측정에 관해 미국 특허 제8,709,843호; 미국 특허 제6,007,845호; 미국 특허 제5,855,913호; 미국 특허 제5,985,309호; 미국 특허 제5,543,158호; 및 [James E. Dahlman and Carmen Barnes et al. Nature Nanotechnology (2014)]의 간행물 (2014년 5월 11일자로 온라인 공개, doi:10.1038/nnano.2014.84)이 언급될 수 있다.

[0536] 본 발명의 범주 내의 입자 전달 시스템은 고체, 반고체, 에멀션 또는 콜로이드 입자를 포함하지만, 이들로 제한되지 않는 임의의 형태로 제공될 수 있다. 이와 같이, 예를 들어, 지질-기반 시스템, 리포솜, 마이셀, 마이크로소포체, 엑소솜, 또는 유전자 총을 포함하지만, 이들로 제한되지 않는 본 명세서에 기재된 임의의 전달 시스템은 본 발명의 범주 내에서 입자 전달 시스템으로서 제공될 수 있다.

[0537] CRISPR-Cas 단백질 mRNA, 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 어댑터에 융합될 수 있음) 또는 mRNA, 및 가이드 RNA는 입자 또는 지질 외피를 이용하여 동시에 전달될 수 있고; 예를 들어, 본 발명의 CRISPR-Cas 단백질 및 RNA는, 예를 들어, 복합체로서, Dahlman 등의 WO2015089419 A2 및 이에 인용된 문헌에서와 같은 입자, 예컨대, 7C1 (예를 들어, 문헌 [James E. Dahlman and Carmen Barnes et al. Nature Nanotechnology (2014)] 참조, 2014년 5월 11일자에 공개, doi:10.1038/nnano.2014.84)을 통해 전달될 수 있고, 예를 들어, 전달 입자는 지질 또는 리피도이드 및 친수성 중합체, 예를 들어, 양이온성 지질 및 친수성 중합체를 포함하고, 예를 들어 여기서 양이온성 지질은 1,2-이올레오일-3-트리메틸암모늄-프로판 (DOTAP) 또는 1,2-디테트라데카노일-*sn*-글리세로-3-포스포콜린 (DMPC)을 포함하고/하거나 친수성 중합체는 에틸렌 글리콜 또는 폴리에틸렌 글리콜 (PEG)을 포함하고/하거나, 입자는 콜레스테롤 (예를 들어, 제제 1의 입자 = DOTAP 100, DMPC 0, PEG 0, 콜레스테롤 0; 제제 번호 2 = DOTAP 90, DMPC 0, PEG 10, 콜레스테롤 0; 제제 번호 3 = DOTAP 90, DMPC 0, PEG 5, 콜레스테롤 5)을 더 포함하고, 입자는 효율적인, 다단계 과정을 사용해 형성되는데, 먼저, 이펙터 단백질 및 RNA는, 예를 들어, 1:1 몰비로, 예를 들어, 실온에서, 예를 들어, 30분 동안, 예를 들어, 멸균, 뉴클레아제 무함유 1X PBS에서 함께 혼합되고, 별개로, 제제에 적용가능하다면 DOTAP, DMPC, PEG, 및 콜레스테롤은 알콜, 예를 들어, 100% 에탄올에 용해되어서, 2개 용액을 함께 혼합하여 복합체를 함유하는 입자를 형성시킨다).

[0538] 핵산-표적화 이펙터 단백질 (예를 들어, V형 단백질, 예컨대, Cas13) mRNA 및 가이드 RNA는 입자 또는 지질 외피를 이용하여 동시에 전달될 수 있다. 적합한 입자의 예는 미국 특허 제9,301,923호에 기재된 것을 포함하지만, 이들로 제한되지 않는다.

[0539] 예를 들어, 문헌 [Su X, Fricke J, Kavanagh DG, Irvine DJ ("In vitro and in vivo mRNA delivery using lipid-enveloped pH-responsive polymer nanoparticles" Mol Pharm. 2011 Jun 6;8(3):774-87. doi: 10.1021/mp100390w. Epub 2011 Apr 1)]은 인지질 이중층 쉘로 밀봉된 폴리(β-아미노 에스테르)(PBAE) 코어를 갖는 생분해성 코어-쉘 구조화 나노입자를 기술한다. 이들은 생체내 mRNA 전달을 위해 개발되었다. pH-반응성 PBAE 성분은 엔도솜 하괴를 촉진하도록 선택되었고, 반면 지질 표면은 다가양이온 코어의 독성을 최소화하도록 선택되었다. 그러므로, 이것은 본 발명의 RNA 전달에 바람직하다.

[0540] 일 실시형태에서, 모두 뇌에 대한, 펩티드의 경구 전달, 펩티드의 정맥내 전달 및 펩티드의 비강 전달에 적용할 수 있는, 자기 조립형 생체부착 중합체 기반 입자/나노입자가 고려된다. 다른 실시형태, 예컨대, 경구 흡수 및 소수성 약물의 안구 전달이 또한 고려된다. 분자 외피 기법은 보호하여 질환 부 위로 전달되는 조작된 중합체 엔벨로프를 포함한다 (예를 들어, 문헌 [Mazza, M. et al. ACSNano, 2013. 7(2): 1016-1026]; [Siew, A., et al. Mol Pharm, 2012. 9(1):14-28]; [Lalatsa, A., et al. J Contr Rel, 2012. 161(2):523-36]; [Lalatsa, A., et al., Mol Pharm, 2012. 9(6):1665-80]; [Lalatsa, A., et al. Mol Pharm, 2012. 9(6):1764-74]; [Garrett, N.L., et al. J Biophotonics, 2012. 5(5-6):458-68]; [Garrett, N.L., et al. J Raman Spect, 2012. 43(5):681-688]; [Ahmad, S., et al. J Royal Soc Interface 2010. 7:S423-33]; [Uchegbu, I.F. Expert Opin Drug Deliv, 2006. 3(5):629-40]; [Qu, X.,et al. Biomacromolecules, 2006. 7(12):3452-9] 및 [Uchegbu, I.F., et al. Int J Pharm, 2001. 224:185-199] 참조). 표적 조직에 따라 단일 또는 다수 용량으로, 약 5mg/kg의 용량이 고려된다.

[0541] 일 실시형태에서, MIT의 Dan Anderson의 실험실에 의해 개발된 종양 성장을 정지시키기 위해 암 세포에 RNA를 전달할 수 있는 입자/나노입자가 본 발명의 AD-기능화된 CRISPR-Cas 시스템에 사용될 수 있고/있거나 그를 위해 적합화될 수 있다. 특히, Anderson 실험실은 신규 생물물질 및 나노제제의 합성, 정제, 특징규명 및 제제화

를 위해서, 완전 자동화, 조합 시스템을 개발하였다. 예를 들어, 문헌 [Alabi et al., Proc Natl Acad Sci U S A. 2013 Aug 6;110(32):12881-6]; [Zhang et al., Adv Mater. 2013 Sep 6;25(33):4641-5]; [Jiang et al., Nano Lett. 2013 Mar 13;13(3):1059-64]; [Karagiannis et al., ACS Nano. 2012 Oct 23;6(10):8484-7]; [Whitehead et al., ACS Nano. 2012 Aug 28;6(8):6922-9] 및 [Lee et al., Nat Nanotechnol. 2012 Jun 3;7(6):389-93]을 참조한다.

[0542] 미국 공개 특허 출원 제20110293703호는 본 발명의 AD-기능화된 CRISPR-Cas 시스템을 전달하는 데 적용될 수 있는 폴리뉴클레오티드의 투여에 특히 유용한 리피도이드 화합물에 관한 것이다. 일 양상에서, 아미노알콜 리피도이드 화합물은 세포 또는 대상체에 전달하려는 작용제와 조합되어 미세입자, 나노입자, 리포솜 또는 마이셀을 형성한다. 입자, 리포솜 또는 마이셀에 의해 전달하려는 작용제는 기체, 액체 또는 고체의 형태일 수 있고, 작용제는 폴리뉴클레오티드, 단백질, 펩티드 또는 소형 분자일 수 있다. 아미노알콜 리피도이드 화합물은 다른 아미노알콜 리피도이드 화합물, 중합체 (합성 또는 천연), 계면활성제, 콜레스테롤, 탄수화물, 단백질, 지질 등과 조합되어 입자를 형성할 수 있다. 이어서, 이들 입자는 선택적으로 약학 부형제와 조합되어 약학 조성물을 형성할 수 있다.

[0543] 미국 공개 특허 출원 제20110293703호는 또한 아미노알콜 리피도이드 화합물의 제조 방법을 제공한다. 아민의 하나 이상의 등가물은 적합한 조건 하에 에폭시드-종결 화합물의 하나 이상의 등가물과 반응되어 본 발명의 아미노알콜 리피도이드 화합물을 형성한다. 일정 실시형태에서, 아민의 모든 아미노기는 에폭시드-종결 화합물과 완전히 반응되어 3차 아민을 형성한다. 다른 실시형태에서, 아민의 모든 아미노기는 에폭시드-종결 화합물과 완전히 반응되지 않아서 3차 아민을 형성하고, 이에 의해 아미노알콜 리피도이드 화합물에서 1차 또는 2차 아민을 초래한다. 이들 1차 또는 2차 아민은 그대로 남아있거나 또는 상이한 에폭시드-종결 화합물과 같은 다른 친핵체와 반응될 수 있다. 당업자에 의해 이해되는 바와 같이, 과량 미만의 에폭시드-종결 화합물과 아민의 반응은 다양한 수의 꼬리부를 갖는 복수의 상이한 아미노알콜 리피도이드 화합물을 생성시키게 될 것이다. 일정 아민은 2개의 에폭시드-유래 화합물 꼬리부에 의해 완전히 작용화될 수 있는 반면, 다른 분자는 에폭시드-유래 화합물 꼬리부에 의해 완전히 작용화되지 않을 것이다. 예를 들어, 디아민 또는 폴리아민은 분자의 다양한 아미노 모이어티로부터 1, 2, 3 또는 4개의 에폭시드-유래 화합물 꼬리부를 포함하여, 1차, 2차 및 3차 아민을 초래할 수 있다. 일정 실시형태에서, 모든 아미노기는 완전히 작용화되지 않는다.

일정 실시형태에서, 2종의 동일한 유형의 에폭시드-종결된 화합물이 사용된다. 다른 실시형태에서, 2종 이상의 상이한 에폭시드-종결된 화합물이 사용된다. 아미노알콜 리피도이드 화합물의 합성은 용매와 함께 또는 용매 없이 수행되고, 합성은 30 내지 100℃, 바람직하게는 대략 50 내지 90℃ 범위의 고온에서 수행될 수 있다. 제조된 아미노알콜 리피도이드 화합물은 선택적으로 정제될 수 있다 예를 들어, 아미노알콜 리피도이드 화합물의 혼합물은 정제되어 특정 수의 에폭시드-유래 화합물 꼬리부를 갖는 아미노알콜 리피도이드 화합물을 수득할 수 있다. 또는 혼합물은 정제되어 특정 입체- 또는 위치이성질체를 수득할 수 있다. 아미노알콜 리피도이드 화합물은 또한 알킬 할로겐화물 (예를 들어, 메틸 아이오다이드) 또는 다른 알킬화제를 이용하여 알킬화될 수 있고/있거나 그들은 아실화될 수 있다.

[0544] 미국 특허 공개 제20110293703호는 또한 본 발명의 방법에 의아 제조된 아미노알콜 리피도이드 화합물의 라이브러리를 제공한다. 이들 아미노알콜 리피도이드 화합물은 액체 핸들러, 로봇, 미세정량판, 컴퓨터 등을 수반하는 고속대량 기법을 이용하여 제조되고/되거나 스크리닝될 수 있다. 일정 실시형태에서, 아미노알콜 리피도이드 화합물은 세포 내로 폴리뉴클레오티드 또는 다른 제제(예를 들어, 단백질, 펩티드, 소형 분자)를 형질감염시키는 그들의 능력에 대해 스크리닝된다.

[0545] 미국 공개 특허 출원 제20130302401호는 폴리(베타-아미노 알콜)(PBAA) 부류가 조합적 중합을 이용하여 제조되는 것에 관한 것이다. 본 발명의 PBAA는 생체기법 및 생의학적 적용분야에서 코팅 (예컨대, 의학적 장치 또는 이식물에 대한 필름의 코팅 또는 다중층 필름), 첨가제, 물질, 부형제, 비-바이오파울링제 (non-biofouling agent), 미세패턴화제 및 세포의 캡슐화제로서 사용될 수 있다. 표면 코팅제로서 사용될 때, 이들 PBAA는 그들의 화학 구조에 따라서 시험관내와 생체내 둘 모두에서 상이한 염증 수준을 유발하였다. 이 부류의 물질의 더 큰 화학적 다양성은 본 발명자들이 시험관내 거대세포 활성화를 억제하는 중합체 코팅을 확인할 수 있게 하였다. 더 나아가, 이들 코팅은 염증 세포의 동원을 감소시키고, 카복실화된 폴리스티렌 미세입자의 피하 이식 후에 섬유증을 감소시킨다. 이들 중합체는 세포 캡슐화를 위해 다가전해질 복합체 캡슐을 형성하는 데 사용될 수 있다. 본 발명은 또한 다수의 다른 생물학적 적용분야, 예컨대, 항미생물 코팅, DNA 또는 siRNA 전달 및 줄기 세포 조직 조작을 가질 수 있다. 미국 공개 특허 출원 제20130302401호의 교시는 본 발명의 AD-기능화된 CRISPR-Cas 시스템에 적용될 수 있다.

[0546] Cas13, (Cas13 또는 어댑터 단백질에 융합될 수 있는) 아데노신 데아미나제, 및 가이드 RNA를 포함하는 사전 조립된 재조합 CRISPR-Cas 복합체는 예를 들어 전기천공에 의해 형질감염되어, 높은 돌연변이율 및 검출 가능한 오프-표적 돌연변이의 부재를 초래한다. Hur, J.K. et al, targeted mutagenesis in mice by electroporation of Cas13 ribonucleoproteins, Nat Biotechnol. 2016 Jun 6. doi: 10.1038/nbt.3596.

[0547] 뇌로의 국소 전달과 관련하여, 이것은 다양한 방식으로 달성될 수 있다. 예를 들어, 재료는 선조체내로, 예를 들어, 주사를 통해 전달될 수 있다. 주사는 개두술을 통해 정위적으로 수행될 수 있다.

[0548] 일부 실시형태에서, 당-기반 입자, 예를 들어 GalNAc은 본 명세서에 기재된 바와 같이 사용될 수 있으며, 특히 전달에 관해 WO2014118272(본 명세서에 참고로 편입됨) 및 문헌 [Nair, JK et al., 2014, Journal of the American Chemical Society 136 (49), 16958-16961)] 및 본 명세서의 교시를 참고로 하여 달리 명확하지 않다면, 모든 입자에 적용된다. 이는 당-기반 입자인 것으로 고려될 수 있고, 다른 입자 전달 시스템 및/또는 제형에 대한 추가적인 상세한 설명이 본 명세서에 제공된다. 그러므로, GalNAc는 본 명세서에 기술된 다른 입자의 의미에서 입자라고 간주될 수 있어서, 일반적인 용도 및 다른 고려사항, 예를 들어 상기 입자의 전달이 역시 GalNAc 입자에 적용된다. 용액-상 접합 전략은 예를 들어 5'-헥실아미노 변형된 올리고뉴클레오티드 (5'-HA ASO, mol. wt. ~8000 Da; Ostergard et al., Bioconjugate Chem., 2015, 26 (8), pp 1451-1455) 상에 PFP (펜타플루오로페닐) 에스테르로서 활성화된 삼중안테나 GalNAc 클러스터 (mol. wt. ~ 2000)를 부착시키는 데 사용될 수 있다. 유사하게, 생체내 핵산 전달을 위한 폴리(아크릴레이트) 중합체가 기술된 바 있다 (참조로 본 명세서에 편입되는 WO2013158141 참조). 추가적인 대안의 실시형태에서, CRISPR 나노 입자 (또는 단백질 복합체)와 천연 발생 혈청 단백질의 사전 혼합이 전달을 개선시키기 위해 사용될 수 있다 (Akinc A et al, 2010, Molecular Therapy vol. 18 no. 7, 1357-1364).

[0549] **나노클루 (Nanoclew)**

[0550] 또한, AD-기능화된 CRISPR 시스템은 예를 들어, [Sun W et al, *Cocoon-like self-degradable DNA nanoclew for anticancer drug delivery.*, J Am Chem Soc. 2014 Oct 22;136(42):14722-5. doi: 10.1021/ja5088024. Epub 2014 Oct 13.]; 또는 [Sun W et al, *Self-Assembled DNA Nanoclews for the Efficient Delivery of CRISPR-Cas9 for Genome 편집.*, Angew Chem Int Ed Engl. 2015 Oct 5;54(41):12029-33. doi: 10.1002/anie.201506030. Epub 2015 Aug 27]에 기술된 바와 같은 나노클루를 사용하여 전달될 수도 있다.

[0551] **LNP**

[0552] 일부 실시형태에서, 전달은 지질 입자, 예컨대, LNP에 Cas13 단백질 또는 mRNA 형태의 캡슐화에 의한다. 따라서, 일부 실시형태에서, 지질 나노입자 (LNP)가 고려된다. 안티트랜스타이레틴 짧은 간섭 RNA는 지질 나노입자에 캡슐화되었고, 인간에게 전달되며 (예를 들어, 문헌 [Coelho et al., N Engl J Med 2013;369:819-29] 참조), 이러한 시스템은 본 발명의 CRISPR Cas 시스템에 적합하고, 적용될 수 있다. 정맥내로 투여되는 kg 체중당 약 0.01 내지 약 1 mg의 용량이 고려된다. 주입-관련 반응 위험을 감소시키기 위한 약물이 고려되고, 예컨대, 덱사메타손, 아세트암피노펜, 다이펜하이드라민 또는 세티리진 및 라니티딘이 고려된다. 5회 용량의 경우 4주마다 킬로그램 당 약 0.3 mg의 다수 용량이 또한 고려된다.

[0553] LNP는 간에 siRNA를 전달하는 데 고도로 효과적인 것으로 확인되었고 (예를 들어, 문헌 [Tabernero et al., Cancer Discovery, April 2013, Vol. 3, No. 4, pages 363-470 참조), 그러므로 CRISPR Cas를 코딩하는 RNA를 간에 전달하는 것이 고려된다. 2주마다 6 mg/kg의 LNP의 약 4회 용량의 용량이 고려될 수 있다. Tabernero 등은 0.7 mg/kg 용량의 LNP의 최초 2회 주기 이후에 종양 퇴행이 관찰되었고, 6회 주기의 종료까지, 환자는 림프절 전이의 완전한 퇴행 및 간 종양의 실질적 축소로 부분 반응을 달성하였다는 것을 입증하였다. 이 환자에서 40 용량 이후에 완전 반응이 수득되었고, 환자는 26개월 동안 용량을 받은 후에 관해 및 완전 치료된 채로 남았다. VEGF 경로 억제제에 의한 사전 요법 후에 진행된 신장, 폐 및 림프절을 포함하여 간외 질환 부위 및 RCC를 갖는 2명 환자는 대략 8개월 내지 12개월 동안 모든 부위에서 안정한 질환을 가졌고, PNET 및 간 전이를 갖는 환자는 안정한 질환을 가진채로 18개월 (36회 용량) 동안 계속 연장 연구되었다.

[0554] 그러나, LNP의 전하가 고려되어야만 한다. 양이온성 지질이 음으로 하전된 지질과 조합되어 세포내 전달을 용이하게 하는 비이중층 구조를 유도하기 때문이다. 하전된 LNP는 정맥내 주사 후 순환으로부터 빠르게 제거되기 때문에, pKa 값이 7 미만인 이온화 가능한 양이온성 지질이 발생되었다 (예를 들어, 문헌 [Rosin et al, Molecular Therapy, vol. 19, no. 12, pages 1286-2200, Dec. 2011] 참조). 음으로 하전된 중합체, 예컨대, RNA는 이온화가능한 지질이 양전하를 나타내는 낮은 pH 값 (예를 들어, pH 4)에서 LNP에 로딩될 수 있

다. 그러나, 생리적 pH 값에서, LNP는 더 긴 순환 시간과 상용성인 낮은 표전 전하를 나타낸다. 이온화 가능한 양이온성 지질의 4가지 종, 즉, 1,2-디리네오일-3-디메틸암모늄-프로판 (DLinDAP), 1,2-디리놀레일옥시-3-N,N-디메틸아미노프로판 (DLinDMA), 1,2-디리놀레일옥시-케토-N,N-디메틸-3-아미노프로판 (DLinKDMA) 및 1,2-디리놀레일-4-(2-디메틸아미노에틸)-[1,3]-디옥솔란 (DLinKC2-DMA)에 집중하였다. 이들 지질을 함유하는 LNP siRNA 시스템은 생체내 간세포에서 현저하게 상이한 유전자 침묵 특성을 나타내며, 역가는 인자 VII 유전자 침묵 모델을 적용하는 시리즈 DLinKC2-DMA>DLinKDMA>DLinDMA>>DLinDAP에 따라 다양하다 (예를 들어, 문헌 [Rosin et al, Molecular Therapy, vol. 19, no. 12, pages 1286-2200, Dec. 2011] 참조). 1 ㎍/㎖의 LNP 또는 LNP 내 또는 이와 회합된 CRISPR-Cas RNA의 용량이 특히 DLinKC2-DMA를 함유하는 제제의 경우에 고려될 수 있다.

[0555] LNP 및 CRISPR Cas 캡슐화의 제조는 문헌 [Rosin et al, Molecular Therapy, vol. 19, no. 12, p 1286-2200, Dec. 2011]로부터 사용되고/되거나 개조될 수 있다. 양이온성 지질 1,2-디리네오일-3-디메틸암모늄-프로판 (DLinDAP), 1,2-디리놀레일옥시-3-N,N-디메틸아미노프로판 (DLinDMA), 1,2-디리놀레일옥시케토-N,N-디메틸-3-아미노프로판 (DLinK-DMA), 1,2-디리놀레일-4-(2-디메틸아미노에틸)-[1,3]-디옥솔란 (DLinKC2-DMA), (3-o-[2"-(메톡시폴리에틸렌글리콜 2000) 숙시노일]-1,2-디미리스토일-sn-글리콜(PEG-S-DMG) 및 R-3-[(ω-메톡시-폴리(에틸렌 글리콜)2000) 카바모일]-1,2-디미리스틸옥시프로필-3-아민 (PEG-C-DOMG)은 Tekmira Pharmaceuticals (Vancouver, Canada)에 의해 제공되거나 또는 합성될 수 있다. 콜레스테롤은 Sigma (St Louis, MO)로부터 구입할 수 있다. 특정 CRISPR Cas RNA는 DLinDAP, DLinDMA, DLinK-DMA, 및 DLinKC2-DMA (40:10:40:10 몰비의 양이온성 지질:DSPC:CHOL: PEGS-DMG 또는 PEG-C-DOMG)을 함유하는 LNP에 캡슐화될 수 있다. 필요하다면, 0.2% SP-DiOC18 (Invitrogen, Burlington, Canada)은 세포 흡수, 세포내 전달 및 생체분포를 평가하기 위해 혼입될 수 있다. 캡슐화는 에탄올에 양이온성 지질:DSPC:콜레스테롤:PEG-c-DOMG (40:10:40:10 몰비)를 포함하는 지질 혼합물을 10 mmol/ℓ 의 최종 지질 농도로 용해시킴으로써 수행될 수 있다. 이러한 지질의 에탄올 용액을 50 mmol/ℓ 시트레이트, pH 4.0에 적가하여 다층판 소포체를 형성하여 30% 에탄올 vol/vol의 최종 농도를 생성할 수 있다. 대형 단층판 소포체는 압출기 (Northern Lipids, Vancouver, Canada)를 사용하여 2개의 적층된 80 nm 뉴클레포어 (Nuclepore) 폴리카보네이트 필터를 통해 다층판 소포체의 압출 후에 형성될 수 있다. 캡슐화는 30% 에탄올 vol/vol을 함유하는 50 mmol/ℓ 시트레이트, pH 4.0에 2 mg/㎖로 용해된 RNA를 압출 수행된 대형 단층판 소포체에 적가하고 0.06/1 wt/wt의 최종 RNA/지질 중량비로 일정하게 혼합하면서 31℃에서 30분 동안 인큐베이션시켜 성취될 수 있다. 에탄올의 제거 및 제제 완충액의 중화는 Spectra/Por 2 재생된 셀룰로스 투석막을 이용하여 16시간 동안 포스페이트-완충 염수 (PBS), pH 7.4에 대한 투석으로 수행하였다. NICOMP 370 입도 분석기, 소포체/강도 모드 및 가우시안 적합도 (Nicomp Particle Sizing, Santa Barbara, CA)를 이용하는 동적광산란에 의해 나노입자 크기 분포를 결정할 수 있다. 모든 3가지의 LNP 시스템에 대한 입자 크기는 ~70 nm의 직경일 수 있다. RNA 캡슐화 효율은 투석 이전 및 이후에 수집된 샘플로부터 VivaPureD MiniH 컬럼 (Sartorius Stedim Biotech)을 사용해 유리 RNA를 제거하여 결정할 수 있다. 캡슐화된 RNA는 용리된 나노입자로부터 추출될 수 있고, 260 nm에서 정량될 수 있다. RNA 대 지질비는 Wako Chemicals USA(Richmond, VA)의 콜레스테롤 E 효소 어세이를 이용하여 소포체 내 콜레스테롤 함량 측정에 의해 결정되었다. 본 명세서에서 논의된 LNP 및 PEG 지질과 함께, PEG화된 리포솜 또는 LNP가 역시 CRISPR-Cas 시스템 또는 이의 성분의 전달에 적합하다.

[0556] 지질 사전혼합 용액 (20.4 mg/㎖의 총 지질 농도)은 DLinKC2-DMA, DSPC 및 콜레스테롤을 50:10:38.5 몰비로 함유하는 에탄올 중에서 제조될 수 있다. 소듐 아세테이트는 0.75:1 (소듐 아세테이트:DLinKC2-DMA)의 몰비로 지질 사전혼합물에 첨가될 수 있다. 지질은 후속하여 격렬하게 교반시키면서 혼합물을 1.85 부피의 시트레이트 완충제 (10 mmol/ℓ, pH 3.0)와 조합하여 수하될 수 있어서, 그 결과로 35% 에탄올을 함유하는 수성 완충액 중에 자발적 리포솜 형성을 야기시킨다. 리포솜 용액은 입자 크기의 시간-의존적 증가를 가능하게 하기 위해 37℃에서 인큐베이션될 수 있다. 동적광산란 (Zetasizer Nano ZS, Malvern Instruments, Worcestershire, UK)에 의한 리포솜 크기 변화를 조사하기 위해 인큐베이션 동안 다양한 시간에 분취액을 제거할 수 있다. 목적 입자 크기가 얻어지면, 수성 PEG 지질 용액 (스톡 = 35% (vol/vol) 에탄올 중 10 mg/㎖ PEG-DMG)을 리포솜 혼합물에 첨가하여 3.5% 총 지질의 최종 PEG 몰농도가 산출될 수 있다. PEG-지질의 첨가 시, 리포솜은 그들의 크기를 조절하여 추가 성장을 효과적으로 중단시켜야 한다. 이어서, RNA는 대략 1:10 의 RNA 대 총 지질 비율 (wt:wt))로 빈 리포솜에 첨가된 후에, 30분 동안 37℃에서 인큐베이션되어 로딩된 LNP가 형성될 수 있다. 혼합물은 후속적으로 PBS에서 밤새 투석될 수 있고, 0.45 ㎛ 시린지 필터로 여과될 수 있다.

[0557] 구형 핵산 (SNA™) 구성체 및 다른 나노 입자 (특히 금 나노입자)는 또한 의도된 표적에 CRISPR-Cas 시스템을

전달하기 위한 수단으로서 고려된다. 유의한 데이터는 핵산-작용화된 금 나노입자에 기반한 AuraSense Therapeutics의 구형 핵산(SNA™) 구성체가 유용하다는 것을 보여준다.

[0558] 본 명세서의 교시와 함께 사용될 수 있는 문헌은 다음을 포함한다: Cutler et al., J. Am. Chem. Soc. 2011 133:9254-9257, Hao et al., Small. 2011 7:3158-3162, Zhang et al., ACS Nano. 2011 5:6962-6970, Cutler et al., J. Am. Chem. Soc. 2012 134:1376-1391, Young et al., Nano Lett. 2012 12:3867-71, Zheng et al., Proc. Natl. Acad. Sci. USA. 2012 109:11975-80, Mirkin, Nanomedicine 2012 7:635-638 Zhang et al., J. Am. Chem. Soc. 2012 134:16488-1691, Weintraub, Nature 2013 495:S14-S16, Choi et al., Proc. Natl. Acad. Sci. USA. 2013 110(19):7625-7630, Jensen et al., Sci. Transl. Med. 5, 209ra152 (2013) 및 Mirkin, et al., Small, 10:186-192.

[0559] RNA와 나노입자의 자가조립은 폴리에틸렌 글리콜 (PEG) 원위 단부에 부착된 Arg-Gly-Asp(RGD) 펩티드 리간드로 PEG화된 폴리에틸렌이민 (PEI)을 사용해 구축될 수 있다. 이 시스템은, 예를 들어, 종양 신생혈관 발현 인테그린을 표적화하고 siRNA 억제성 혈관 내피 성장 인자 수용체-2(VEGF R2) 발현을 전달하여, 종양 혈관신생을 달성하기 위한 수단으로서 사용되었다 (예를 들어, 문헌 [Schiffelers et al., Nucleic Acids Research, 2004, Vol. 32, No. 19] 참조). 2 내지 6 범위의 순 몰 과량의 이온화가능한 질소 (중합체) 대 포스페이트 (핵산)를 제공하도록 동일 부피의 양이온 중합체 수용액 및 핵산을 혼합하여 나노플렉스를 제조할 수 있다. 양이온성 중합체와 핵산 사이의 정전기적 상호작용은 본 명세서에서 나노플렉스라고 하는 약 100 nm의 평균 입자 크기 분포를 갖는 폴리플렉스의 형성을 야기시킨다. 약 100 내지 200 mg 용량의 CRISPR Cas가 Schiffelers 등의 자가-조립 나노입자로의 전달을 위해 계획된다.

[0560] Bartlett 등 (PNAS, September 25, 2007,vol. 104, no. 39)의 나노플렉스가 또한 본 발명에 적용될 수 있다. Bartlett 등의 나노플렉스는 2 내지 6 범위의 순 몰 과량의 이온화 가능한 질소 (중합체) 대 포스페이트 (핵산)를 제공하도록 동일 부피의 양이온성 중합체의 수용액 및 핵산을 혼합하여 제조된다. 양이온성 중합체와 핵산 간의 정전기적 상호작용은 약 100 nm의 평균 입자 크기 분포를 갖는 폴리플렉스의 형성을 야기시켰고, 그러한 이유로 본 명세서에서 나노플렉스라고 한다. Bartlett 등의 DOTA-siRNA는 다음과 같이 합성되었다: 1,4,7,10-테트라아자시클로도데칸-1,4,7,10-테트라아세트산 모노(N-히드록시숙신이미드 에스테르) (DOTA-NHS에스테르)는 Macrocyclics (Dallas, TX)에서 주문하였다. 카보네이트 완충액 (pH 9) 중 100배 몰 과량의 DOTA-NHS-에스테르로 아민 변형된 RNA 센스 가닥을 미세원심분리 튜브에 첨가하였다. 내용물은 4시간 동안 실온에서 교반시켜 반응시켰다. DOTA-RNA센스 접합체를 에탄올 침전시키고 나서, 물에서 재현탁시킨 후, 비변형 안티센스 가닥에 어닐링시켜서 DOTA-siRNA를 산출하였다. 모든 액체는 Chelex-100(Bio-Rad, Hercules, CA)로 전처리하여 미량의 금속 오염물을 제거하였다. Tf-표적화 및 비표적화 siRNA 나노입자는 시클로덱스트린-함유 다가양이온을 사용하여 형성시킬 수 있다. 전형적으로, 나노입자는 3(+/-)의 전하비 및 0.5 g/리터의 siRNA 농도로 수 중에서 형성되었다. 표적화된 나노입자 표면 상에서 아다만탄-PEG 분자의 1%가 Tf에 의해 변형되었다 (아다만탄-PEG-Tf). 나노입자는 주사용 5% (wt/vol) 글루코스 담체 용액 중에서 현탁되었다.

[0561] Davis 등(Nature, Vol 464, 15 April 2010)은 표적화된 나노입자-전달 시스템을 사용하는 RNA 임상 시험 (임상시험 등록 번호 NCT00689065)을 수행한다. 표준 치료 요법에 대해 난치성인 고형암을 갖는 환자에게 21일 주기 중 1일, 3일, 8일 및 10일에 30분 정맥내 주입으로 표적화된 나노입자의 용량을 투여한다. 나노입자는 (1) 선형, 시클로덱스트린계 중합체 (CDP), (2) 암 세포 표면 상에서 TF 수용체(TFR)와 맞물리도록 나노입자 외부에 제시되는 리간드 표적화 인간 트랜스페린 단백질 (TF), (3) 친수성 중합체 (생물학적 유체에서 나노입자 안정성을 촉진시키기 위해 사용되는 폴리에틸렌 글리콜 (PEG)), 및 (4) RRM2의 발현을 감소시키도록 디자인된 siRNA (임상에서 사용되는 서열은 이전에 표시된siR2B+5임)를 함유하는 합성 전달 시스템으로 이루어진다. TFR은 악성 세포에서 상향조절되는 것으로 오랫동안 알려져있었고, RRM2는 확립된 항암 표적이다. 이들 나노 입자 (CALAA-01로 표시되는 임상 버전)는 인간이외의 영장류에서 다회-용량 연구에서 충분히 내약성인 것으로 확인되었다. 만성 골수성 백혈병을 갖는 단일 환자에게 리포솜 전달에 의해 siRNA가 투여되었지만, Davis 등의 임상 시험은 표적화된 전달 시스템으로 siRNA를 전신에 전달하고 고형암을 갖는 환자를 치료하기 위한 초기 인간 시험이다. 표적화된 전달 시스템이 인간 종양에 기능성 siRNA의 효과적인 전달을 제공하는지의 여부를 확인하기 위해, Davis 등은 3종의 상이한 투약 코호트로부터 3명의 환자로서, 모두 전이성 흑색종을 갖고 각각 18, 24 및 30 mg m^{-2} siRNA의 CALAA-01 용량을 받은 환자 A, B 및 C로부터의 생검을 조사하였다.

본 발명의 CRISPR Cas 시스템에 대해 유사한 용량이 또한 고려될 수 있다. 본 발명의 전달은 선형, 시클로덱스트린-기반 중합체 (CDP), 암 세포의 표면 상의 TF 수용체 (TFR)와 맞물리도록 나노입자의 외면 상에 제시

되는 리간드를 표적화하는 인간 트랜스페린 단백질 (TF) 및/또는 친수성 중합체 (예를 들어, 생물학적 유체에서 나노입자 안정성을 촉진시키기 위해 사용된 폴리에틸렌 글리콜(PEG))를 함유하는 나노입자에 의해 달성될 수 있다.

[0562] 본 명세서에 참고로 편입된 미국 특허 제8,709,843호는 조직, 세포, 및 세포내 구획에 치료제-함유 입자의 표적화된 전달을 위한 약물 전달 시스템을 제공한다. 본 발명은 계면활성제, 친수성 중합체 또는 지질에 접합된 중합체를 포함하는 표적화된 입자를 제공한다. 본 명세서에 참고로 편입된 미국 특허 제6,007,845호는 다작용성 화합물을 하나 이상의 소수성 중합체 및 하나 이상의 친수성 중합체와 공유 결합시킴으로써 형성된 다중블록 공중합체의 코어를 가지며, 생물학적으로 활성인 물질을 함유하는 입자를 제공한다. 본 명세서에 참고로 편입된 미국 특허 제5,855,913호는 탭 밀도가 0.4 g/㎤ 미만이고 평균 직경이 5 ㎛ 내지 30 ㎛이며, 폐 시스템에 약물 전달을 위해 표면 상에 계면활성제를 혼입시키는 공기역학적 광 입자를 갖는 미립자 조성물을 제공한다. 본 명세서에 참고로 편입된 미국 특허 제5,985,309호는 폐 시스템에 전달을 위해 양으로 또는 음으로 하전된 치료제 또는 진단제 및 반대 전하의 하전 분자의 친수성 또는 소수성 복합체 및/또는 계면활성제를 혼입시키는 입자를 제공한다. 본 명세서에 참고로 편입된 미국 특허 제5,543,158호는 표면 상에서 생물학적 활성 물질 및 폴리(알킬렌 글리콜) 모이어티를 함유하는 생분해성 고체 코어를 갖는 생분해성 주사 가능 입자를 제공한다. 본 명세서에 참고로 편입된 WO2012135025 (또한 미국 특허 제20120251560호로서 공개됨)는 접합된 폴리에틸렌이민(PEI) 중합체 및 접합된 아자-거대고리(총괄적으로 "접합 리포머" 또는 "리포머"로서 지칭됨)를 기재한다. 일정 실시형태에서, 이러한 접합 리포머는 단백질 발현의 조율을 비롯한 유전자 발현을 변형시키는 시험관내, 생체외 및 생체내 게놈 동요를 달성하기 위해 CRISPR-Cas 시스템과 관련하여 사용될 수 있다는 것이 고려될 수 있다.

[0563] 일 실시형태에서, 나노입자는 에폭시드-변형된 지질-중합체, 유리하게는 7C1일 수 있다 (예를 들어, 2014년 5월 11일자로 온라인 상에서 공개된 [James E. Dahlman and Carmen Barnes et al. Nature Nanotechnology (2014)], doi:10.1038/nnano.2014.84] 참조). C71은 C15 에폭시드-종결된 지질을 PEI600과 14:1 몰비로 반응시킴으로써 합성되었고, C14PEG2000와 함께 제제화되어 적어도 40일 동안 PBS 용액에서 안정한 나노입자 (35 내지 60 nm의 직경)을 생성하였다.

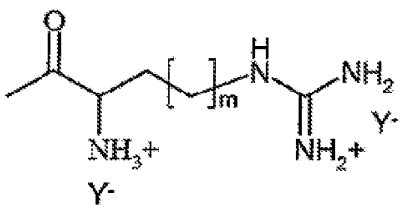
[0564] 에폭시드-변형된 지질-중합체는 폐, 심장 또는 신장 세포에 본 발명의 CRISPR-Cas 시스템을 전달하기 위해 이용될 수 있지만, 당업자는 다른 표적 기관에 전달하도록 시스템을 개조할 수 있다. 약 0.05 내지 약 0.6 mg/kg 범위의 용량이 계획된다. 수 일 또는 수 주 동안 용량이 또한 고려되고 총 용량은 약 2 mg/kg이다.

[0565] 일부 실시형태에서, RNA 분자를 전달하기 위한 LNP는 당업계에 공지된 방법, 예컨대, 본 명세서에 참고로 편입된 WO 2005/105152 (PCT/EP2005/004920), WO 2006/069782 (PCT/EP2005/014074), WO 2007/121947 (PCT/EP2007/003496), 및 WO 2015/082080 (PCT/EP2014/003274)에 기재된 것에 의해 제조된다. 특별히 포유동물 세포 내로 siRNA의 증강되고 개선된 전달을 목적으로 하는 LNP는, 예를 들어, 하기 문헌들에 기재되어 있고, 이들은 참조로 본 명세서에 편입되며, 본 기술에 적용될 수 있다: Aleku et al., Cancer Res., 68(23): 9788-98 (Dec. 1, 2008), Strumberg et al., Int. J. Clin. Pharmacol. Ther., 50(1): 76-8 (Jan. 2012), Schultheis et al., J. Clin. Oncol., 32(36): 4141-48 (Dec. 20, 2014), 및 Fehring et al., Mol. Ther., 22(4): 811-20 (Apr. 22, 2014).

[0566] 일부 실시형태에서, LNP는 WO 2005/105152 (PCT/EP2005/004920), WO 2006/069782 (PCT/EP2005/014074), WO 2007/121947 (PCT/EP2007/003496) 및 WO 2015/082080 (PCT/EP2014/003274)에 개시된 임의의 LNP를 포함한다.

[0567] 일부 실시형태에서, LNP는 화학식 I의 구조를 갖는 적어도 하나의 지질을 포함한다:

[0568] 화학식 I에서, R1 및 R2는 각각 알킬을 포함하는 군으로부터 독립적으로 선택되고, n은 1 내지 4의 임의 정수이고, R3은 리실, 오르니틸, 2,4-디아미노부티릴, 히스티딜 및 아실 모이어티 (하기 화학식 II에 따름)를 포함하는 군으로부터 선택되는 아실이다:



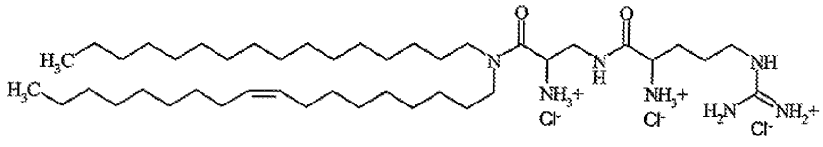
[0569] (화학식 II)

[0570] 화학식 II에서, m은 1 내지 3의 임의 정수이고, Y^- 는 약학적으로 허용가능한 음이온이다. 일부 실시형태에서, 화학식 I에 따른 지질은 적어도 2개의 비대칭 C 원자를 포함한다. 일부 실시형태에서, 화학식 I의 거울상 이성질체는 제한없이 R-R; S-S; R-S 및 S-R 거울상 이성질체를 포함한다.

[0571] 일부 실시형태에서, R1은 라우릴이고 R2는 미리스틸이다. 다른 실시형태에서, R1은 팔미틸이고 R2는 올레일이다. 일부 실시형태에서, m은 1 또는 2이다. 일부 실시형태에서, Y-는 할로게니드, 아세테이트 또는 트리플루오로아세테이트로부터 선택된다.

[0572] 일부 실시형태에서, LNP는 하기로부터 선택되는 하나 이상의 지질을 포함한다:

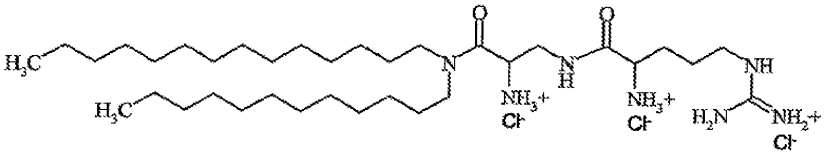
[0573] -아르기닐-2,3-디아미노 프로피온산-N-팔미틸-N-올레일-아미드 트리히드로클로라이드 (화학식 III):



(화학식 III)

[0574]

[0575] -아르기닐-2,3-디아미노 프로피온산-N-라우릴-N-미리스틸-아미드 트리히드로클로라이드 (화학식 IV):

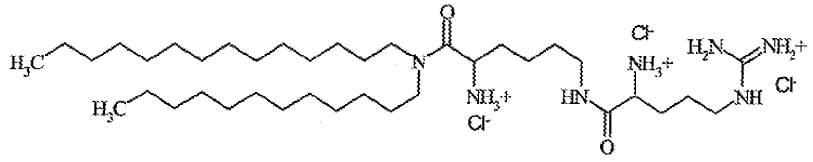


(화학식 IV)

[0576]

[0577] ; 및

[0578] -아르기닐-리신-N-라우릴-N-미리스틸-아미드 트리히드로클로라이드 (화학식 V):



(화학식 V)

[0579]

[0580] 일부 실시형태에서, LNP는 또한 구성성분을 포함한다. 예로서, 제한없이,일부 실시형태에서, 구성성분은 펩티드, 단백질, 올리고뉴클레오티드, 폴리뉴클레오티드, 핵산 또는 이들의 조합물로부터 선택된다. 일부 실시형태에서, 구성성분은 항체, 예를 들어, 단일클론 항체이다. 일부 실시형태에서, 구성성분은, 예를 들어, 리보자임, 압타머, 스피에겔머 (spiegelmer), DNA, RNA, PNA, LNA, 또는 이들의 조합으로부터 선택된 핵산이다. 일부 실시형태에서, 핵산은 가이드 RNA 및/또는 mRNA이다.

[0581] 일부 실시형태에서, LNP의 구성성분은 CRIPSR-Cas 단백질을 코딩하는 mRNA를 포함한다. 일부 실시형태에서, LNP의 구성성분은 II형 또는 V형 CRIPSR-Cas 단백질을 코딩하는 mRNA를 포함한다. 일부 실시형태에서, LNP의 구성성분은 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 어댑터 단백질에 융합될 수 있음)를 코딩하는 mRNA를 포함한다.

[0582] 일부 실시형태에서, LNP의 구성성분은 추가로 하나 이상의 가이드 RNA를 포함한다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 혈관내피에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 폐 내피에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 간에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 폐에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 심장에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 비장에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 신장에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 췌장에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 뇌에 전달하도록 구성된다. 일부 실시형태에서, LNP는 앞서 언급된 mRNA 및 가이드 RNA를 마크로파지에 전달하도록 구성된다.

[0583] 일부 실시형태에서, LNP는 또한 적어도 하나의 헬퍼 지질을 포함한다. 일부 실시형태에서, 헬퍼 지질은 인지질 및 스테로이드로부터 선택된다. 일부 실시형태에서, 인지질은 인산의 디에스테르 및/또는 모노에스테르이다. 일부 실시형태에서, 인지질은 포스포글리세라이드 및/또는 스핑고리피드이다. 일부 실시형태에서, 스테로이드는 부분적으로 수소화된 시클로펜타[a]페난트렌에 기반한 천연 발생 및/또는 합성 화합물이다. 일부 실시형태에서, 스테로이드는 21 내지 30개의 C 원자를 함유한다. 일부 실시형태에서, 스테로이드는 콜레스테롤이다. 일부 실시형태에서, 헬퍼 지질은 1,2-디피타노일-sn-글리세로-3-포스포에탄올아민(DPhyPE), 세라미드, 및 1,2-디올레일-sn-글리세로-3-포스포에탄올아민(DOPE)으로부터 선택된다.

[0584] 일부 실시형태에서, 적어도 하나의 헬퍼 지질은 PEG 모이어티, HEG 모이어티, 폴리히드록시에틸 전분 (폴리HES) 모이어티 및 폴리프로필렌 모이어티를 포함하는 군으로부터 선택된 모이어티를 포함한다. 일부 실시형태에서, 모이어티는 분자량이 약 500 내지 10,000 Da 또는 약 2,000 내지 5,000 Da이다. 일부 실시형태에서, PEG 모이어티는 1,2-디스테아로일-sn-글리세로-3 포스포에탄올아민, 1,2-디알킬-sn-글리세로-3-포스포에탄올아민, 및 세라미드-PEG로부터 선택된다. 일부 실시형태에서, PEG 모이어티는 분자량이 약 500 내지 10,000 Da 또는 약 2,000 내지 5,000 Da이다. 일부 실시형태에서, PEG 모이어티는 분자량이 2,000 Da이다.

[0585] 일부 실시형태에서, 헬퍼 지질은 조성물의 총 지질 함량의 약 20 mol% 내지 80 mol%이다. 일부 실시형태에서, 헬퍼 지질 성분은 LNP의 총 지질 함량의 약 35 mol% 내지 65 mol%이다. 일부 실시형태에서, LNP는 LNP의 총 지질 함량의 50 mol%의 지질 및 50 mol%의 헬퍼 지질을 포함한다.

[0586] 일부 실시형태에서, LNP는 DPhyPE와 조합하여 -3-아르기닐-2,3-디아미노프로피온산-N-팔미틸-N-올레일-아미드 트리히드로클로라이드, -아르기닐-2,3-디아미노프로피온산-N-라우릴-N-미리스틸-아미드 트리히드로클로라이드 또는 -아르기닐-리신-N-라우릴-N-미리스틸-아미드 트리히드로클로라이드 중 임의의 것을 포함하고, 여기서 DPhyPE의 함량은 LNP의 전체 지질 함량의 약 80 mol%, 65 mol%, 50 mol% 및 35 mol%이다. 일부 실시형태에서, LNP는 -아르기닐-2,3-디아미노 프로피온산-N-파니틸-N-올레일-아미드 트리히드로클로라이드 (지질) 및 1,2-디피타노일-sn-글리세로-3-포스포에탄올아민 (헬퍼 지질)을 포함한다. 일부 실시형태에서, LNP는 β-아르기닐-2,3-디아미노 프로피온산-N-팔미틸-N-올레일-아미드 트리히드로클로라이드(지질), 1,2-디피타노일-sn-글리세로-3-포스포에탄올아민 (제1 헬퍼 지질), 및 1,2-다이스테로일-sn-글리세로-3-포스포에탄올아민-PEG2000 (제2 헬퍼 지질)을 포함한다.

[0587] 일부 실시형태에서, 제2 헬퍼 지질은 총 지질 함량의 약 0.05 mol% 내지 4.9 mol% 또는 약 1 mol% 내지 3 mol%이다. 일부 실시형태에서, LNP는 총 지질 함량의 약 45 mol% 내지 50 mol%의 지질, 총 지질 함량의 약 45 mol% 내지 50 mol%의 제1 헬퍼 지질을 포함하며, 단 총 지질 함량 중 약 0.1 mol% 내지 5 mol%, 약 1 mol% 내지 4 mol%, 또는 약 2 mol%의 PEG화된 제2 헬퍼 지질이 존재하고, 지질, 제1 헬퍼 지질 및 제2 헬퍼 지질 함량의 총합은 총 지질 함량의 100 mol%이고, 제1 헬퍼 지질과 제2 헬퍼 지질의 합은 총 지질 함량의 50 mol%이다. 일부 실시형태에서, LNP는: (a) 50 mol%의 -아르기닐-2,3-디아미노 프로피온산-N-팔미틸-N-올레일-아미드 트리히드로클로라이드, 48 mol%의 1,2-디피타노일-sn-글리세로-3-포스포에탄올아민; 및 2 mol% 1,2-디스테아로일-sn-글리세로-3-포스포에탄올아민-PEG2000; 또는 (b) 50 mol%의 -아르기닐-2,3-디아미노 프로피온산-N-팔미틸-N-올레일-아미드 트리히드로클로라이드, 49 mol% 1,2-디피타노일-sn-글리세로-3-포스포에탄올아민; 및 1 mol% N(카보닐-메톡시폴리에틸렌글리콜-2000)-1,2-디스테아로일-sn-글리세로3-포스포에탄올아민, 또는 이의 소듐 염을 포함한다.

[0588] 일부 실시형태에서, LNP는 핵산을 함유하고, 핵산 골격 포스페이트 대 양이온성 지질 질소 원자의 전하비는 약 1: 1.5 내지 7 또는 약 1:4이다.

[0589] 일부 실시형태에서, LNP는 또한 생체내 조건 하에 지질로부터 제거 가능한 차폐 화합물을 포함한다. 일부 실시형태에서, 차폐 화합물은 생물학적 불활성 화합물이다. 일부 실시형태에서, 차폐 화합물은 그의 표면상에서 또는 이러한 분자 상에서 임의의 전하를 운반하지 않는다. 일부 실시형태에서, 차폐 화합물은 폴리에틸렌글리콜(PEG), 히드록시에틸글루코스(HEG) 기반 중합체, 폴리히드록시에틸 전분(폴리HES) 및 폴리프로필렌이다. 일부 실시형태에서, PEG, HEG, 폴리HES 및 폴리프로필렌은 약 500 내지 10,000Da 또는 약 2000 내지 5000Da의 중량이다. 일부 실시형태에서, 차폐 화합물은은 PEG2000 또는 PEG5000이다.

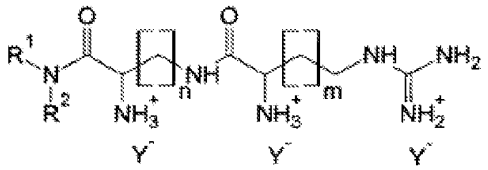
[0590] 일부 실시형태에서, LNP는 생체내 조건 하에 지질로부터 제거 가능한 적어도 하나의 지질, 제1 헬퍼 지질 및 차폐 화합물을 포함한다. 일부 실시형태에서, LNP는 또한 제2 헬퍼 지질을 포함한다. 일부 실시형태에서, 제1 헬퍼 지질은 세라미드이다. 일부 실시형태에서, 제2 헬퍼 지질은 세라미드이다. 일부 실시형태에서, 세라미드는 6 내지 10개의 탄소 원자의 적어도 하나의 짧은 탄소쇄 치환체를 포함한다. 일부 실시형태에

서, 세라미드는 8개의 탄소 원자를 포함한다. 일부 실시형태에서, 차폐 화합물은 세라미드에 부착된다. 일부 실시형태에서, 차폐 화합물은 세라미드에 부착된다. 일부 실시형태에서, 차폐 화합물은 세라미드에 공유 부착된다. 일부 실시형태에서, 차폐 화합물은 LNP에서 핵산에 부착된다. 일부 실시형태에서, 차폐 화합물은 핵산에 공유 부착된다. 일부 실시형태에서, 차폐 화합물은 링커에 의해 핵산에 부착된다. 일부 실시형태에서, 링커는 생리적 조건 하에 절단된다. 일부 실시형태에서, 링커는 ssRNA, ssDNA, dsRNA, dsDNA, 펩티드, S-S-링커 및 pH 민감성 링커로부터 선택된다. 일부 실시형태에서, 링커 모이어티는 핵산의 센스 가닥의 3' 말단에 부착된다. 일부 실시형태에서, 차폐 화합물은 pH-민감성 링커 또는 pH-민감성 모이어티를 포함한다. 일 실시형태에서, pH-민감성 링커 또는 pH-민감성 모이어티는 음이온성 링커 또는 음이온성 모이어티이다. 일부 실시형태에서, 음이온성 링커 또는 음이온성 모이어티는 산성 환경에서 덜 음이온성이거나 또는 중성이다. 일부 실시형태에서, pH-민감성 링커 또는 pH-민감성 모이어티는 올리고 (글루탐산), 올리고페놀레이트(들) 및 디에틸렌 트리아민 펜타 아세트산으로부터 선택된다.

[0591] 앞 단락의 임의의 LNP 실시형태에서, LNP는 약 50 내지 600 밀리오스몰 (mosmole)/kg, 약 250 내지 350 밀리오스몰/kg, 또는 약 280 내지 320 밀리오스몰/kg의 삼투질 농도를 가질 수 있고/있거나, 지질 및/또는 1 또는 2종의 헬퍼 지질 및 차폐 화합물에 의해 형성되는 LNP는 약 20 내지 200 nm, 약 30 내지 100 nm, 또는 약 40 내지 80 nm의 입자 크기를 갖는다.

[0592] 일부 실시형태에서, 차폐 화합물은 더 긴 생체내 순환 시간을 제공하고, 핵산 함유 LNP의 더 양호한 생체 분포를 가능하게 한다. 일부 실시형태에서, 차폐 화합물은 LNP가 투여되는 다른 체액 또는 세포질 막, 예를 들어, 혈관구조 내피벽의 세포질 막의 혈청 화합물 또는 화합물과 LNP의 즉각적인 상호작용을 방지한다. 추가적으로 또는 대안적으로, 일부 실시형태에서, 차폐 화합물은 또한 LNP와의 상호작용 직후부터 면역계의 구성요소를 방지한다. 추가적으로 또는 대안적으로, 일부 실시형태에서, 차폐 화합물은 항-옵소닌 작용 화합물로서 작용한다. 임의의 메커니즘 또는 이론에 의해 구속되는 일 없이, 일부 실시형태에서, 차폐 화합물은 그의 환경과의 상호작용을 위해 이용가능한 LNP의 표면적을 감소시키는 커버 또는 외피를 형성한다. 추가적으로 또는 대안적으로, 일부 실시형태에서, 차폐 화합물은 LNP의 전체 전하를 차폐한다.

[0593] 다른 실시형태에서, LNP는 하기 화학식 VI의 구조를 갖는 적어도 하나의 양이온성 지질을 포함한다:



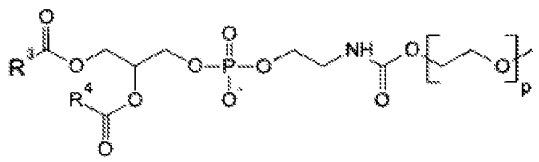
[0594] (화학식 VI)

[0595] 상기 화학식 VI에서, n은 1, 2, 3, 또는 4이고, m은 1, 2, 또는 3이고, 여기서 Y^- 는 양이온이고, 각각의 R^1 및 R^2 는 선형 C12-C18 알킬 및 선형 C12-C18 알케닐, 스테롤 화합물로 이루어진 군으로부터 개별적으로, 독립적으로 선택되고, 여기서 스테롤 화합물은 콜레스테롤 및 스티그마스테롤, 및 PEG화 지질로 이루어진 군으로부터 선택되고, 여기서 PEG화 지질은 PEG 모이어티를 포함하며, 여기서 PEG화 지질은

[0596] 하기 화학식 VII의 PEG화 포스포에탄올아민:

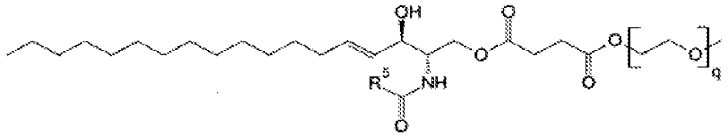
[0597] 하기 화학식 VIII의 PEG화 세라미드; 및

[0598] 하기 화학식 IX의 PEG화 디아실글리세롤로 이루어진 군으로부터 선택된다:



[0599] (화학식 VII)

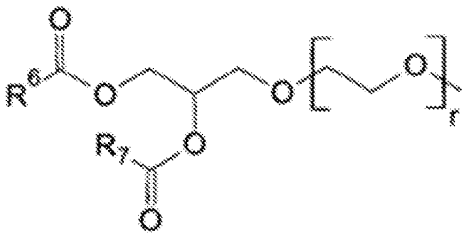
[0600] 상기 화학식 VII에서, R^3 및 R^4 는 개별적으로, 그리고 독립적으로 선형 C13-C17 알킬이고, p는 15 내지 130의 임의 정수이다.



(화학식 VIII)

[0601]

[0602] 상기 화학식 VIII에서, R^5 는 선형 C7-C15 알킬이고, q는 15 내지 130의 임의 수이다:



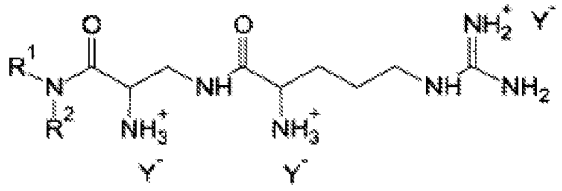
(화학식 IX)

[0603]

[0604] 상기 화학식 IX에서, 각각의 R^6 및 R^7 은 개별적으로 그리고 독립적으로 선형 C11-C17 알킬이고, r은 15 내지 130의 임의 정수이다.

[0605] 일부 실시형태에서, R^1 및 R^2 는 서로 상이하다. 일부 실시형태에서, R^1 은 팔미틸이고 R^2 는 올레일이다. 일부 실시형태에서, R^1 은 라우릴이고 R^2 는 미리스틸이다. 일부 실시형태에서, R^1 및 R^2 는 동일하다. 일부 실시형태에서, R^1 및 R^2 각각은 C12 알킬, C14 알킬, C16 알킬, C18 알킬, C12 알케닐, C14 알케닐, C16 알케닐 및 C18 알케닐로 이루어진 군으로부터 개별적으로 그리고 독립적으로 선택된다. 일부 실시형태에서, 각각의 C12 알케닐, C14 알케닐, C16 알케닐 및 C18 알케닐은 하나 또는 2개의 이중 결합을 포함한다. 일부 실시형태에서, C18 알케닐은 C9와 C10 사이에 하나의 이중 결합을 갖는 C18 알케닐이다. 일부 실시형태에서, C18 알케닐은 cis-9-옥타데실이다.

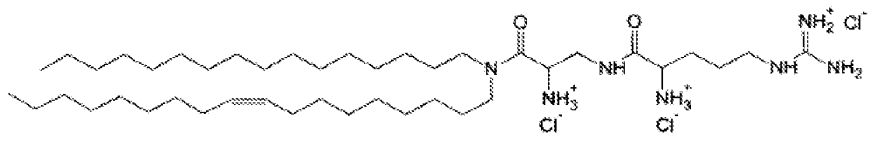
[0606] 일부 실시형태에서, 양이온성 지질은 하기 화학식 X의 화합물이다:



(화학식 X)

[0607]

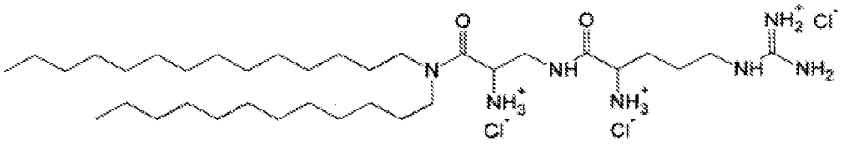
[0608] 일부 실시형태에서, Y^- 는 할로게니드, 아세테이트 및 트리플루오로아세테이트로부터 선택된다. 일부 실시형태에서, 양이온성 지질은 하기 화학식 III의 -아르기닐-2,3-디아미노 프로피온산-N-팔미틸-N-올레일-아미드 트리히드로클로라이드이다:



(화학식 III)

[0609]

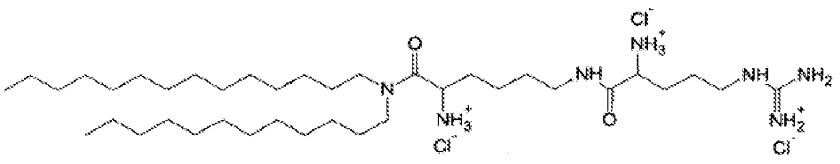
[0610] 일부 실시형태에서, 양이온성 지질은 화학식 IV의 -아르기닐-2,3-디아미노 프로피온산-N-라우릴-N-미리스틸-아미드 트리히드로클로라이드이다:



(화학식 IV)

[0611]

[0612] 일부 실시형태에서, 양이온성 지질은 화학식 V의 -아르기닐-리신-N-라우릴-N-미리스틸-아미드 트리히드로클로라이드이다:

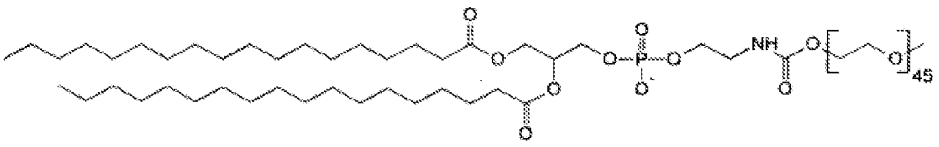


(화학식 V)

[0613]

[0614] 일부 실시형태에서, 스테롤 화합물은 콜레스테롤이다. 일부 실시형태에서, 스테롤 화합물은 스티그마스테린이다.

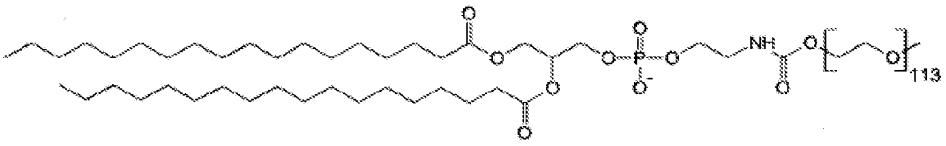
[0615] 일부 실시형태에서, PEG화 지질의 PEG 모이어티는 약 800 내지 5,000 Da의 분자량을 갖는다. 일부 실시형태에서, PEG화 지질의 PEG 모이어티의 분자량은 약 800 Da이다. 일부 실시형태에서, PEG화 지질의 PEG 모이어티의 분자량은 약 2,000 Da이다. 일부 실시형태에서, PEG화 지질의 PEG 모이어티의 분자량은 약 5,000 Da이다. 일부 실시형태에서, PEG화 지질은 화학식 VII의 PEG화 포스포에탄올아민이고, 여기서 각각의 R^3 및 R^4 는 개별적으로, 그리고 독립적으로 선형 C13-C17 알킬이고, p는 18, 19 또는 20, 또는 44, 45 또는 46 또는 113, 114 또는 115의 임의 정수이다. 일부 실시형태에서, R^3 및 R^4 는 동일하다. 일부 실시형태에서, R^3 및 R^4 는 상이하다. 일부 실시형태에서, 각각의 R^3 및 R^4 는 개별적으로 C13 알킬, C15 알킬 및 C17 알킬로 이루어진 군으로부터 독립적으로 선택된다. 일부 실시형태에서, 화학식 VII의 PEG화 포스포에탄올아민은 1,2-디스테아로일-*sn*-글리세로-3-포스포에탄올아민-N-[메톡시(폴리에틸렌 글리콜)-2000] (암모늄 염)이다:



(화학식 XI)

[0616]

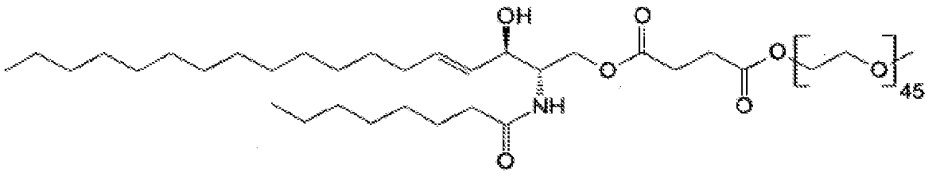
[0617] 일부 실시형태에서, 화학식 VII의 PEG화 포스포에탄올아민은 1,2-디스테아로일-*sn*-글리세로-3-포스포에탄올아민-N-[메톡시(폴리에틸렌 글리콜)-5000] (암모늄 염)이다:



(화학식 XII)

[0618]

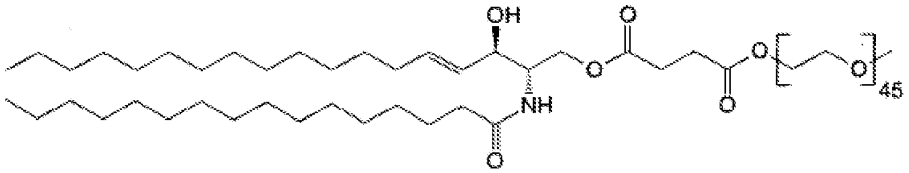
[0619] 일부 실시형태에서, PEG화 지질은 화학식 VIII의 PEG화 세라미드이고, 여기서 R^5 는 선형 C7-C15 알킬이고, q는 18, 19 또는 20, 또는 44, 45 또는 46 또는 113, 114 또는 115의 임의 정수이다. 일부 실시형태에서, R^5 는 선형 C7 알킬이다. 일부 실시형태에서, R^5 는 선형 C15 알킬이다. 일부 실시형태에서, 화학식 VIII의 PEG화 세라미드는 N-옥타노일-스핀고신-1- {숙시닐[메톡시(폴리에틸렌 글리콜)2000]}이다:



(화학식 XIII)

[0620]

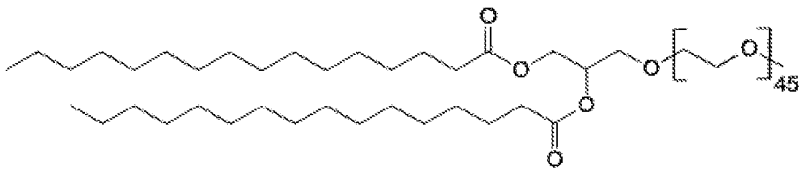
[0621] 일부 실시형태에서, 화학식 VIII의 PEG화 세라미드는 N-팔미토일-스핀고신-1- {숙시닐[메톡시(폴리에틸렌 글리콜)2000]}이다:



(화학식 XIV)

[0622]

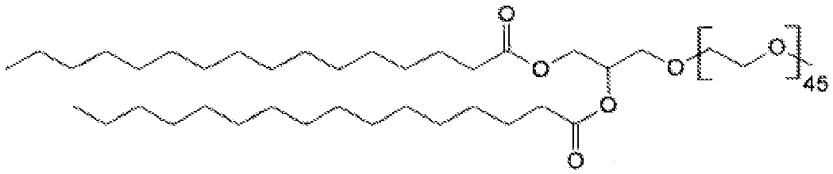
[0623] 일부 실시형태에서, PEG화 지질은 화학식 IX의 PEG화 디아실글리세롤이고, 여기서 각각의 R^6 및 R^7 은 개별적으로, 그리고 독립적으로 선형 C11-C17 알킬이고, r은 18, 19 또는 20, 또는 44, 45 또는 46 또는 113, 114 또는 115의 임의 정수이다. 일부 실시형태에서, R^6 및 R^7 은 동일하다. 일부 실시형태에서, R^6 및 R^7 은 상이하다. 일부 실시형태에서, 각각의 R^6 및 R^7 은 개별적으로, 선형 C17 알킬, 선형 C15 알킬 및 선형 C13 알킬로 이루어진 군으로부터 독립적으로 선택된다. 일부 실시형태에서, 화학식 IX의 PEG화 디아실글리세롤은 1,2-디스테아로일-sn-글리세롤 [메톡시(폴리에틸렌 글리콜)2000]이다:



(화학식 XV)

[0624]

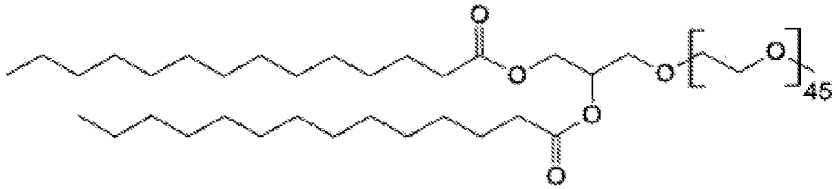
[0625] 일부 실시형태에서, 화학식 IX의 PEG화 디아실글리세롤은 1,2-디팔미토일-sn-글리세롤 [메톡시(폴리에틸렌 글리콜)2000]이다:



(화학식 XVI)

[0626]

[0627] 일부 실시형태에서, 화학식 IX의 PEG화 디아실글리세롤은 다음과 같다:



(화학식 XVII)

[0628]

[0629] 일부 실시형태에서, LNP는 화학식 III, IV, 및 V로부터 선택된 적어도 하나의 양이온성 지질, 콜레스테롤 및 스티그마스테린로부터 선택된 적어도 하나의 스테롤 화합물을 포함하고, 여기서 PEG화 지질은 화학식 XI 및 XII로

부터 선택된 적어도 하나이다. 일부 실시형태에서, LNP는 화학식 III, IV 및 V로부터 선택된 적어도 하나의 양이온성 지질, 콜레스테롤 및 스티그마스테린로부터 선택된 적어도 하나의 스테롤 화합물을 포함하고, 여기서 PEG화 지질은 화학식 XIII 및 XIV로부터 선택된 적어도 하나이다. 일부 실시형태에서, LNP는 화학식 III, IV 및 V로부터 선택된 적어도 하나의 양이온성 지질, 콜레스테롤 및 스티그마스테린으로부터 선택된 적어도 하나의 스테롤 화합물을 포함하고, 여기서 PEG화 지질은 화학식 XV 및 XVI로부터 선택된 적어도 하나이다. 일부 실시형태에서, LNP는 화학식 III의 양이온성 지질, 스테롤 화합물로서 콜레스테롤을 포함하고, 여기서 PEG화 지질은 화학식 XI이다.

[0630] 이전 단락에서 임의의 LNP 실시형태에서, 양이온성 지질 조성물의 함량은 약 65 mole% 내지 75 mole%이고, 스테롤 화합물의 함량은 약 24 mole% 내지 34 mole%이고, PEG화 지질의 함량은 약 0.5 mole% 내지 1.5 mole%이고, 지질 조성물에 대해 양이온성 지질, 스테롤 화합물 및 PEG화 지질의 함량의 총합은 100 mole%이다. 일부 실시형태에서, 양이온성 지질의 함량은 약 70 mole%이고, 스테롤 화합물의 함량은 약 29 mole%이고 PEG화 지질의 함량은 약 1 mole%이다. 일부 실시형태에서, LNP는 70 mole%의 화학식 III, 29 mole%의 콜레스테롤, 및 1 mole%의 화학식 XI이다.

[0631] **엑소솜**

[0632] 엑소솜은 RNA 및 단백질을 수송하고, 뇌 및 다른 표적 기관에 RNA를 전달할 수 있는, 내생성 나노-소포체이다. 면역원성을 감소시키기 위해, 문헌 [Alvarez-Erviti et al. (2011, Nat Biotechnol 29: 341)]은 엑소솜 생산을 위한 자가-유래 수지상 세포를 사용하였다. 뇌에 대한 표적화는 뉴런-특이적 RVG 펩티드에 융합된 엑소솜 막 단백질인 Lamp2b를 발현시키기 위해 수지상 세포를 조작함으로써 달성되었다. 정제된 엑소솜은 전기천공법에 의해 외생성 RNA와 함께 로딩되었다. 정맥내로 주사된 RVG-표적화된 엑소솜은 뇌에서 뉴런, 미세아교세포, 희돌기교세포에 GAPDH siRNA를 특이적으로 전달하여, 특이적 유전자 녹다운을 초래한다. RVG 엑소솜에 대한 사전 노출은 녹다운을 약화시키지 않았고, 다른 조직에서의 비-특이적 흡수는 관찰되지 않았다. 엑소솜-매개 siRNA 전달의 치료적 잠재력은 알츠하이머병에서의 치료적 표적인 BACE1의 강력한 mRNA (60%) 및 단백질 (62%) 녹다운에 의해 입증되었다.

[0633] 면역학적으로 불활성인 엑소솜의 풀을 얻기 위해, Alvarez-Erviti 등은 균질한 주요 조직적합 복합체 (major histocompatibility complex: MHC) 단상형을 갖는 근친교배 C57BL/6 마우스로부터 골수를 채취하였다. 미성숙 수지상 세포는 T-세포 활성체, 예컨대, MHC-II 및 CD86가 결여된 다량의 엑소솜을 생성하는데, Alvarez-Erviti 등은 7일 동안 과립구/마크로파지-콜로니 자극 인자(GM-CSF)를 갖는 수지상 세포를 선택하였다. 충분히 확립된 초원심분리 프로토콜을 이용하여 다음날에 배양 상청액으로부터 엑소솜을 정제하였다. 생성된 엑소솜은 물리적으로 균질하며, 크기 분포는 나노입자 추적 분석 (NTA) 및 전자 현미경으로 결정하여 직경이 80 nm에서 최대이다. Alvarez-Erviti 등은 10^6개의 세포당 6 내지 12 ㎍의 엑소솜을 얻었다 (단백질 농도 기반으로 측정).

[0634] 다음에, Alvarez-Erviti 등은 나노규모 적용분야에 적합한 전기천공법 프로토콜을 이용하여 외생성 카고를 변형된 엑소솜에 로딩하는 가능성을 조사하였다. 나노미터 규모에서 막 입자에 대한 전기천공법은 잘 특징규명되고 있지 않기 때문에, 전기천공법 프로토콜의 경험적 최적화를 위해 비특이적 Cy5-표지 RNA가 사용되었다. 엑소솜의 초원심분리 및 용해 후에 캡슐화된 RNA의 양이 분석되었다. 400 V 및 125 μF에서 전기천공은 RNA의 최고 체류를 초래하였고, 모든 후속 실험을 위해 사용되었다.

[0635] Alvarez-Erviti은 150 ㎍의 RVG 엑소솜에 캡슐화된 150 ㎍의 각각의 BACE1 siRNA를 정상 C57BL/6 마우스에 투여하였고, 녹다운 효율은 4종의 대조군: 비처리 마우스, RVG 엑소솜 단독 주사된 마우스, 생체내 양이온성 리포솜 시약과 복합체화된 BACE1 siRNA를 주사한 마우스 및 RVG-9R에 복합체화된 BACE1 siRNA를 주사한 마우스와 비교하였고, RVG 펩티드는 siRNA에 정전기적으로 결합되는 9 D-아르기닌에 접합된 것이다. 피질 조직 샘플은 투여 후 3일에 분석되었고, siRNA-RVG-9R-처리 마우스와 siRNARVG 엑소솜-처리 마우스 둘 모두에서 유의한 단백질 녹다운 (45%, $P < 0.05$, 대 62%, $P < 0.01$)이 관찰되었는데, 이는 BACE1 mRNA 수준의 상당한 감소를 초래하였다 (각각 66% [+ 또는 -] 15%, $P < 0.001$ 및 61% [+ 또는 -] 13%, $P < 0.01$). 게다가, 출원인들은 RVG-엑소솜-처리 동물에서 알츠하이머 병리에서 아밀로이드 플라크의 주된 성분인 총 [베타]-아밀로이드 1-42 수준의 유의한 감소 (55%, $P < 0.05$)를 입증하였다. 관찰된 감소는 BACE1 억제제의 심실내 주사 후에 정상 마우스에서 입증된 β-아밀로이드 1-40 감소보다 더 컸다. Alvarez-Erviti 등은 siRNA에 의한 RNAi-매개 녹다운 증거를 제공한 BACE1 절단 산물에 대한 cDNA 말단부의 5'-신속 증폭 (RACE)을 수행하였다.

[0636] 최종적으로, Alvarez-Erviti 등은 IL-6, IP-10, TNFα 및 IFN-α 혈청 농도를 평가함으로써 RNA-RVG 엑소솜이

생체내 면역 반응을 유도하는지의 여부를 연구하였다. 엑소솜 처리 후에, 모든 사이토카인의 유의하지 않은 변화는 IL-6 분비를 강하게 자극한 siRNA-RVG-9R와 대조적으로 siRNA-형질감염 시약 처리와 유사하게 나타나서, 엑소솜 처리의 면역학적 불활성 프로파일을 확증하였다. 엑소솜이 20%의 siRNA만을 캡슐화한다는 것을 고려하면, RVG-엑소솜에 의한 전달은 비슷한 mRNA 녹다운으로서 RVG-9R 전달보다 더 효율적인 것으로 나타나고, 상응하는 면역 자극 수준 없이 5배 더 적은 siRNA에 의해 더 큰 단백질 녹다운이 달성되었다. 이 실험은 신경퇴행성 질환과 관련된 유전자의 장기간 침묵에 잠재적으로 적합한 RVG-엑소솜 기술의 치료 가능성을 입증하였다. Alvarez-Erviti 등의 엑소솜 전달 시스템은 본 발명의 AD-기능화된 CRISPR-Cas 시스템을 치료적 표적, 특히 신경퇴행성 질환에 전달하기 위해 적용될 수 있다. 약 100 내지 1000 ㎎의 RVG 엑소솜에 캡슐화된 약 100 내지 1000㎎의 CRISPR Cas의 용량이 본 발명에서 고려될 수 있다.

[0637] El-Andaloussi 등 (Nature Protocols 7,2112-2126(2012))은 배양 세포로부터 유래된 엑소솜이 시험관내 및 생체내 RNA의 전달을 위해 이용될 수 있는 방법을 개시한다. 이 프로토콜은 먼저 펩티드 리간드에 융합된 엑소솜 단백질을 포함하는, 발현 벡터의 형질감염을 통해 표적화된 엑소솜의 생성을 기재한다. 다음에, El-Andaloussi 등은 형질감염 세포 상청액으로부터 엑소솜을 정제하고 특징규명하는 방법을 설명한다. 다음에, El-Andaloussi 등은 RNA를 엑소솜에 로딩하기 위한 중요한 단계를 상술한다. 최종적으로, El-Andaloussi 등은 마우스 뇌에서 시험관내 및 생체내에서 RNA를 효율적으로 전달하기 위해 엑소솜을 이용하는 방법을 약술한다. 엑소솜-매개 RNA 전달이 기능성 분석 및 영상화에 의해 평가된 예상 결과의 예가 또한 제공된다. 전체 프로토콜은 ~3주가 걸린다. 본 발명에 따른 전달 또는 투여는 자가-유래 수지상 세포로부터 생성된 엑소솜을 이용하여 수행될 수 있다. 본 명세서의 교시로부터, 이는 본 발명의 실행에서 사용될 수 있다.

[0638] 다른 실시형태에서, Wahlgren 등 (Nucleic Acids Research, 2012, Vol. 40, No. 17 e130)의 혈장 엑소솜이 고려된다. 엑소솜은 수지상 세포 (DC), B 세포, T 세포, 비만 세포, 상피 세포 및 종양 세포를 비롯한 다수의 세포 유형에 의해 생성된 나노-크기 소포체 (크기가 30 내지 90 ㎚)이다. 이들 소포체는 후기 엔도솜의 내향 버딩 (inward budding)에 의해 형성되고, 이어서, 혈장막과의 융합 시 세포외 환경으로 방출된다. 엑소솜은 세포 사이에서 RNA를 자연적으로 운반하기 때문에, 이 특성은 유전자 요법에서 유용할 수 있고, 본 개시내용으로부터 본 발명의 실행에서 사용될 수 있다.

[0639] 혈장을 단리시키기 위해 900g에서 20분 동안 버피코트의 원심분리 다음에 세포 상청액을 채취하고, 세포를 제거하기 위해 300g에서 10분 동안 그리고 16 500g에서 30분 동안 원심분리시킨 후 0.22 ㎜ 필터를 통해 여과시킴으로써 혈장으로부터의 엑소솜을 제조할 수 있다. 엑소솜은 120,000g에서 70분 동안 초원심분리에 의해 펠렛화된다. 엑소솜으로 siRNA의 화학적 형질감염은 RNAi 인간/마우스 Starter 키트 (Quiagen, Hilden, Germany)로 제조사의 설명서에 따라 수행된다. siRNA는 2 ㎜ol/㎖의 최종 농도로 100 ㎖ PBS에 첨가된다. HiPerFect 형질감염 시약을 첨가한 후에, 혼합물은 10분 동안 실온에서 인큐베이션된다. 과량의 마이셀을 제거하기 위해, 엑소솜은 알데히드/설페이트 라텍스 비드를 이용하여 재단리된다. 엑소솜 내로 CRISPR Cas의 화학적 형질감염은 siRNA와 유사하게 수행될 수 있다. 엑소솜은 건강한 공혈자의 말초 혈액으로부터 단리된 단핵구 및 림프구와 함께 공동배양될 수 있다. 따라서, CRISPR Cas를 함유하는 엑소솜이 인간의 단핵구 및 림프구에 도입될 수 있고, 인간에 자가적으로 재도입될 수 있다는 것이 고려될 수 있다. 따라서, 본 발명에 따른 전달 또는 투여는 혈장 엑소솜을 이용하여 수행될 수 있다.

[0640] **리포솜**

[0641] 본 발명에 따른 전달 또는 투여는 리포솜에 의해 수행될 수 있다. 리포솜은 내부 수성 구획을 둘러싸며 친유성의 인지질 이중층 밖에서 상대적으로 불침투성인 단층판 또는 다층판 지질 이중층으로 구성된 구형 소포체 구조이다. 리포솜은 약물 전달 담체로서 상당히 주목되었는데, 그들이 생체적합하고, 비독성이며, 친수성과 친유성 약물 분자를 둘 모두 전달할 수 있고, 그들의 카고를 혈장 효소에 의해 분해로부터 보호하며, 생물학적 막 및 혈액 뇌 장벽 (blood brain barrier: BBB)을 가로질러 그들의 로딩물을 수송하기 때문이다 (예를 들어, 검토를 위해 문헌 [Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[0642] 리포솜은 몇몇 상이한 유형의 지질로부터 생성될 수 있으나, 약물 담체로서 리포솜을 생성하는 데 인지질이 가장 통상적으로 사용된다. 지질막이 수용액과 혼합될 때 리포솜 형성은 자발적이지만, 이는 또한 균질기, 초음파분쇄기, 또는 압출 장치를 이용하여 진탕 형태로 힘을 적용함으로써 신속히 처리될 수 있다 (예를 들어, 고찰을 위해 [Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[0643] 몇몇 다른 첨가제가 리포솜에 첨가되어 이의 구조 및 특성을 변형시킬 수 있다. 예를 들어, 콜레스테롤 또는 스핀고미엘린 중 하나가 리포솜 혼합물에 첨가되어 리포솜 구조를 안정화시키고 리포솜 내부의 카고 누출을 방지할 수 있다. 추가로, 리포솜은 수소화된 에그 포스파티딜콜린 또는 에그 포스파티딜콜린, 콜레스테롤 및 다이세틸 포스페이트로부터 제조되며, 그들의 평균 소포체 크기는 약 50 내지 100 nm로 조절되었다. (예를 들어, 고찰을 위해 문헌 [Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[0644] 리포솜 제제는 주로 천연 인지질 및 지질, 예컨대, 1,2-디스테아로일-sn-글리세로-3-포스파티딜 콜린(DSPC), 스핀고미엘린, 에그 포스파티딜콜린 및 모노시알로강글리오시드로 구성될 수 있다. 이 제제는 인지질만으로 구성되기 때문에, 리포솜 제제는 많은 도전과 접하게 되었는데, 이 중 하나는 혈장에서의 불안정성이다. 이들 도전을 극복하기 위한 몇몇 시도가, 특히 지질막의 조작에서 만들어졌다. 이들 시도 중 하나는 콜레스테롤의 조작에 중점을 두었다. 통상적인 제제에 대한 콜레스테롤의 첨가는 혈장내로 캡슐화된 생활성 화합물의 빠른 방출을 감소시키거나 또는 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민 (DOPE)은 안정성을 증가시킨다 (예를 들어, 고찰을 위해 문헌 [Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[0645] 특히 유리한 실시형태에서, 트로이 목마 (Trojan Horse) 리포솜 (또한 분자 트로이 목마로도 공지됨)가 바람직하며, 프로토콜은 http://cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long에서 확인할 수 있다. 이들 입자는 혈관내 주사 후에 전체 뇌에 대한 이식유전자 전달을 허용한다. 이론에 의해 국한하지 않고, 표면에 접합된 특정 항체를 갖는 중성 지질 입자는 세포내이입을 통해 혈액 뇌 장벽을 가로지르는 것으로 여겨진다. 트로이 목마 리포솜은 혈관내 주사를 통해 뇌에 뉴클레아제의 CRISPR 패밀리를 전달하는 데 사용될 수 있는데, 이는 이는 배아 조작에 대한 필요 없이 전체 뇌 유전자이식 동물을 허용한다. 약 1 내지 5g의 DNA 또는 RNA는 리포솜으로 생체내 투여가 고려될 수 있다.

[0646] 다른 실시형태에서, AD-기능화된 CRISPR Cas 시스템 또는 이의 성분은 리포솜, 예컨대, 안정한 핵산-지질 입자 (SNALP)로 투여될 수 있다 (예를 들어, 문헌 [Morrissey et al., Nature Biotechnology, Vol. 23, No. 8, August 2005] 참조). SNALP에서 표적화된 약 1, 3 또는 5 mg/kg/일의 특정 CRISPR Cas의 1일 정맥내 주사가 고려된다. 1일 처리는 약 3일 동안일 수 있고, 그 다음으로 약 5주 동안 매주일 수 있다. 다른 실시형태에서, 약 1 또는 2.5mg/kg 용량으로 정맥내 주사에 의해 투여되는 특정 CRISPR Cas 캡슐화 SNALP가 또한 고려된다 (예를 들어, 문헌 [Zimmerman et al., Nature Letters, Vol. 441, 4 May 2006] 참조). SNALP 제형은 지질 3-N-[(w-메톡시 폴리(에틸렌 글리콜) 2000) 카바모일]-1,2-다이미리스틸옥시-프로필아민 (PEG-C-DMA), 1,2-디리놀레일옥시-N,N-디메틸-3-아미노프로판 (DLinDMA), 1,2-디스테아로일-sn-글리세로-3-포스포콜린(DSPC) 및 콜레스테롤을 2:40:10:48 몰 백분율 비로 함유할 수 있다 (예를 들어, 문헌 [Zimmerman et al., Nature Letters, Vol. 441, 4 May 2006] 참조).

[0647] 다른 실시형태에서, 안정한 핵산-지질 입자 (SNALP)는 고도로 혈관화된 HepG2-유래 간 종양에 대한 효과적인 전달 분자가 되지만, 불량하게 혈관화된 HCT-116 유래 간 종양에서는 그렇지 않은 것으로 증명되었다 (예를 들어, 문헌 [Li, Gene Therapy (2012) 19, 775-780] 참조). SNALP 리포솜은 콜레스테롤/D-Lin-DMA/DSPC/PEG-C-DMA의 48/40/10/2 몰비 및 25:1 지질/siRNA 비율을 사용하여 D-Lin-DMA 및 PEG-C-DMA와 디스테아로일포스파티딜콜린(DSPC), 콜레스테롤 및 siRNA를 함께 제제화하여 제조될 수 있다. 얻어진 SNALP 리포솜은 크기가 약 80 내지 100 nm이다.

[0648] 또 다른 실시형태에서, SNALP는 합성 콜레스테롤 (Sigma-Aldrich, St Louis, MO, USA), 디팔미토일포스파티딜콜린 (Avanti Polar Lipids, Alabaster, AL, USA), 3-N-[(w-메톡시 폴리(에틸렌 글리콜)2000)카바모일]-1,2-디미레스틸옥시프로필아민, 및 양이온성 1,2-디리놀레일옥시-3-N,N-디메틸아미노프로판 (예를 들어, 문헌 [Geisbert et al., Lancet 2010; 375: 1896-905] 참조)을 포함할 수 있다. 예를 들어, 볼루스 정맥내 주입으로서 투여되는 용량 당 약 2 mg/kg의 총 CRISPR Cas 용량이 고려될 수 있다.

[0649] 또 다른 실시형태에서, SNALP는 합성 콜레스테롤 (Sigma-Aldrich), 1,2-디스테아로일-sn-글리세로-3-포스포콜린 (DSPC; Avanti Polar Lipids Inc.), PEG-cDMA, 및 1,2-디리놀레일옥시-3-(N;N-디메틸)아미노프로판 (DLinDMA)을 포함할 수 있다 (예를 들어, 문헌 [Judge, J. Clin. Invest. 119:661-673 (2009)] 참조). 생체내 연구를 위해 사용되는 제제는 약 9:1의 최종 지질/RNA 질량비를 포함할 수 있다.

[0650] RNAi 나노메디신의 안전성 프로파일은 Alnylam Pharmaceuticals의 Barros 및 Gollob에 의해 검토되었다 (예를 들어, 문헌 [Advanced Drug Delivery Reviews 64 (2012) 1730-1737] 참조). 안정한 핵산 지질 입자

(SNALP)는 4가지의 상이한 지질로 구성된다 - 낮은 pH에서 양이온성인 이온화 가능한 지질 (DLinDMA), 중성 헬퍼 지질, 콜레스테롤, 및 확산성 폴리에틸렌 글리콜(PEG)-지질. 입자는 직경이 대략 80 nm이고, 생리적 pH에서 전하-중성이다. 제제화 동안에, 이온화 가능한 지질은 입자 형성 동안 음이온성 RNA와 지질을 축합시키는 역할을 한다. 점차적으로 산성인 엔도솜 조건 하에서 양으로 하전될 때, 이온화 가능한 지질은 또한 엔도솜 막과 SNALP의 융합을 매개하여 RNA이 세포질 내로 방출되는 것을 가능하게 한다. PEG-지질은 입자를 안정화시키고, 제제화 동안 응집을 감소시키며, 후속적으로 약물동태학적 특성을 개선시키는 외부의 중성 친수성을 제공한다.

[0651] 지금까지, RNA를 갖는 SNALP 제형을 이용하여 두 임상 프로그램이 개시되었다. Tekmira Pharmaceuticals는 LDL 콜레스테롤이 상승된 성인에서 SNALP-ApoB의 I상 단일-용량 연구를 최근에 완료하였다. ApoB는 간 및 공장에서 우세하게 발현되며, VLDL 및 LDL의 조립 및 분비에 필수적이다. 17명의 대상체는 SNALP-ApoB의 단일 용량을 받았다 (7회 용량 수준에 걸쳐 용량 상승). 간 독성의 증거는 없었다 (전임상 연구에 기반하여 잠재적 용량 제한 독성으로서 예상됨). 최고 용량의 (2명 중) 1명 대상체는 면역계 자극과 일관되게 감기-유사 증상을 경험하였고, 시험을 결론내리기 위해 결정되었다.

[0652] Alnylam Pharmaceuticals는 유사하게 진행된 ALN-TTR01을 갖는데, 이는 상기 기재한 SNALP 기술을 사용하며 돌연변이체와 야생형 TTR 둘 모두의 간세포 생성을 표적화하여 TTR 아밀로이드증 (ATTR)을 치료하였다. 3가지 ATTR 증후군이 기재되었다: 가족성 아밀로이드증성 말초신경병증 (FAP) 및 가족성 아밀로이드증성 심장근육병증 (FAC) - (둘 모두 TTR에서의 상염색체 우성 돌연변이에 의해 야기); 및 야생형 TTR에 의해 야기되는 노인성 전신 아밀로이드증 (SSA). ALN-TTR01의 위약-대조, 단일 용량-상승 I상 시험은 ATTR을 갖는 환자에서 최근에 완료되었다. ALN-TTR01은 0.01 내지 1.0 mg/kg (siRNA에 기반)의 용량 범위 내에서 31명의 환자 (23명은 연구 약물이고, 8명은 위약임)에게 15분 IV 주입으로서 투여되었다. 간 기능 검사에서의 유의한 증가 없이 치료는 충분히 내성이 있었다. ≥0.4 mg/kg에서 23명 환자 중 3명에서 주입-관련 반응이 확인되었으며; 모두 주입률 완화에 반응하고, 모두 연구에서 계속되었다. 2명의 환자에서 혈청 사이토카인 IL-6, IP-10 및 IL-1ra의 최소 및 일시적 상승이 1 mg/kg의 최고 용량에서 확인되었다 (전임상 및 NHP 연구로부터 예상). 혈청 TTR이 저하되면서, ALN-TTR01의 예상되는 약력학적 효과가 1mg/kg에서 관찰되었다.

[0653] 또 다른 실시형태에서, 양이온성 지질, DSPC, 콜레스테롤 및 PEG-지질을, 예를 들어, 에탄올에서, 예를 들어, 각각 40:10:40:10의 몰비로 가용화함으로써 SNALP가 생성될 수 있다 (문헌 [Semple et al., Nature Niotechnology, Volume 28 Number 2 February 2010, pp. 172-177] 참조). 지질 혼합물은 최종 에탄올과 30% (vol/vol) 및 6.1 mg/㎖ 농도의 지질에 각각 혼합하여 수성 완충제 (50 mM 시트레이트, pH 4)에 첨가되었고, 압출 전에 22℃에서 2분 동안 평형 상태가 되게 하였다. 동적광산란 분석에 의해 결정하여 70 내지 90 nm의 소포체 직경이 얻어질 때까지, 수화된 지질을 Lipex Extruder (Northern Lipids)를 이용하여 22℃에서 2 적층 80 nm 포어-크기 필터 (Nuclepore)를 통해 압출시켰다. 이는 일반적으로 1 내지 3회의 통과를 필요로 하였다. siRNA (50 mM 시트레이트에서 가용화됨, 30% 에탄올을 함유하는 pH 4 수용액)을 혼합하면서 ~5 mL/분의 속도로 사전 평형화시킨(35℃) 소포체에 첨가하였다. 0.06 (wt/wt)의 최종 표적 siRNA/지질 비에 도달된 후에, 혼합물은 추가 30분 동안 35℃에서 인큐베이션되어 siRNA의 소포체 재조직화 및 캡슐화를 허용한다. 이어서, 에탄올이 제거되고, 외부 완충제는 투석 또는 접선유동 정용여과 중 하나에 의해 PBS(155 mM NaCl, 3 mM Na_2HPO_4, 1 mM KH_2PO_4, pH 7.5)로 대체된다. 제어된 단계적 희석 방법 과정을 이용하여 siRNA는 SNALP에서 캡슐화되었다. KC2-SNALP의 지질 구성성분은 57.1:7.1:34.3:1.4의 몰비로 사용한 DLin-KC2-DMA (양이온성 지질), 다이팔미토일포스파티딜콜린 (DPPC; Avanti Polar Lipids), 합성 콜레스테롤 (Sigma) 및 PEG-C-DMA이었다. 로딩된 입자의 형성 시, SNALP는 PBS에 대해 투석되었고, 사용 전 0.2 ㎛ 필터를 통해 멸균 여과되었다. 평균 입자 크기는 75 내지 85 nm이었고, siRNA의 90 내지 95%는 지질 입자 내에서 캡슐화되었다. 생체내 시험을 위해 사용되는 제형에서의 최종 siRNA/지질 비는 ~0.15 (wt/wt)이었다. 인자 VII siRNA를 함유하는 LNP-siRNA 시스템은 사용 직전에 멸균 PBS에서 적절한 농도로 희석되었고, 제형은 10 mL/kg의 총 용적으로 옆쪽 꼬리 정맥을 통해 정맥내로 투여되었다. 이 방법 및 이들 전달 시스템은 본 발명의 AD-기능화된 CRISPR Cas 시스템에 대해 추론될 수 있다.

[0654] **기타 지질**

[0655] 다른 양이온성 지질, 예컨대, 아미노 지질 2,2-디리놀레일-4-디메틸아미노에틸-[1,3]-디옥솔란 (DLin-KC2-DMA)은, 예를 들어, SiRNA와 유사하게 CRISPR Cas 또는 이의 성분 또는 이를 코딩하는 핵산 분자(들)를 캡슐화하는 데 이용될 수 있으며 (예를 들어, 문헌 [Jayaraman, Angew. Chem. Int. Ed. 2012, 51, 8529 -8533] 참

조), 따라서 본 발명의 실행에서 사용될 수 있다. 다음의 지질 조성물이 있는 사전형성된 소포체가 고려될 수 있다: 각각 몰비 40/10/40/10로 아미노 지질, 디스테아로일포스파티딜콜린(DSPC), 콜레스테롤 및 (R)-2,3-비스(옥타데실옥시) 프로필-1-(메톡시 폴리(에틸렌 글리콜)2000)프로필카바메이트 (PEG-지질), 및 대략 0.05(w/w)의 FVII siRNA/총 지질 비율. 70 내지 90 nm 범위의 좁은 입자 크기 분포 및 0.11±0.04(n=56)의 낮은 다분산 지수를 보장하기 위해, 입자는 가이드 RNA 첨가 전에 80 nm 막을 통해 3회까지 압출될 수 있다. 고도로 강한 아미노산 지질 16을 함유하는 입자가 사용될 수 있으며, 이때 4가지 지질 성분 16, DSPC, 콜레스테롤 및 PEG-지질 (50/10/38.5/1.5)의 몰비는 생체내 활성을 증강시키기 위해 추가로 최적화될 수 있다.

[0656] Michael S D Kormann 등 ("Expression of therapeutic proteins after delivery of chemically modified mRNA in mice: Nature Biotechnology, Volume:29, Pages: 154-157 (2011))은 RNA를 전달하기 위한 지질 외피의 사용을 기재한다. 지질 외피의 사용은 또한 본 발명에서 바람직하다.

[0657] 다른 실시형태에서, 지질은 본 발명의 AD-기능화된 CRISPR Cas 시스템 또는 이의 성분(들) 또는 이를 코딩하는 핵산 분자(들)로 제제화되어 지질 나노입자(LNP)를 형성할 수 있다. 지질은 DLin-KC2-DMA4, C12-200 및 코리피드 디스테로일포스파티딜 콜린, 콜레스테롤을 포함하지만, 이들로 제한되지 않으며, PEG-DMG는 자발적 소포체 형성 절차를 이용하여 siRNA 대신에 CRISPR Cas로 제제화될 수 있다 (예를 들어, 문헌 [Novobrantseva, Molecular Therapy-Nucleic Acid (2012) 1, e4; doi:10.1038/mtna.2011.3] 참조). 성분 몰비는 약 50/10/38.5/1.5 (DLin-KC2-DMA 또는 C12-200/디스테로일포스파티딜 콜린/콜레스테롤/PEG-DMG)일 수 있다. 최종 지질:siRNA 중량비는 각각 DLin-KC2-DMA 및 C12-200 지질 나노입자 (LNP)의 경우에 대략 12:1 내지 9:1일 수 있다. 제제는 평균 입자 직경이 ~80 nm이고 포획 효율이 90% 초과일 수 있다. 3 mg/kg 용량이 고려될 수 있다.

[0658] Tekmira는 LNP 및 LNP 제제의 다양한 양상에 관해 미국 및 해외에서 대략 95개 특허 패밀리의 포트폴리오를 가지며 (예를 들어, 미국 특허 제7,982,027호; 제7,799,565호; 제8,058,069호; 제8,283,333호; 제7,901,708호; 제7,745,651호; 제7,803,397호; 제8,101,741호; 제8,188,263호; 제7,915,399호; 제8,236,943호 및 제7,838,658호 및 유럽 특허 제1766035호; 제1519714호; 제1781593호 및 제1664316호 참조), 이들 모두는 본 발명에 사용되고/되거나 적합화될 수 있다.

[0659] AD-기능화된 CRISPR Cas 시스템 또는 이의 성분 또는 이를 코딩하는 핵산 분자(들)는 단백질, 단백질 전구체를 코딩하거나, 또는 단백질 또는 단백질 전구체의 부분적으로 또는 완전히 가공된 형태를 코딩할 수 있는 변형된 핵산 분자를 포함하는 조성물의 제형의 양상에 관한 미국 특허 출원 공개 제20130252281호 및 제20130245107호 및 제20130244279호 (Moderna Therapeutics에게 양도)에 추가로 기재되는 것과 같이 PLGA 미세구에서 캡슐화되어 전달될 수 있다. 제제는 몰비 50:10:38.5:1.5 내지 3.0 (양이온성 지질:융합유도성 (fusogenic) 지질:콜레스테롤:PEG 지질)을 가질 수 있다. PEG 지질은 PEG-c-DOMG, PEG-DMG로부터 선택되지만, 이들로 제한되지 않을 수 있다. 융합유도성 지질은 DSPC일 수 있다. 또한 Schrum 등의 미국 특허 출원 공개 제20120251618호 (발명의 명칭: Delivery and Formulation of Engineered Nucleic Acid)를 참조한다.

[0660] Nanomerics의 기술은 저분자량 소수성 약물, 펩티드 및 핵산 기반 치료제 (플라스미드, siRNA, miRNA)를 비롯한 광대한 범위의 치료제에 대한 생체 이용 가능성 도전을 해결한다. 기술이 분명한 이점을 입증한 특정 투여 경로는 경구 경로, 뇌-혈관-장벽을 가로지르는 수송, 고형 종양뿐만 아니라 눈에 대한 전달을 포함한다. 예를 들어, 문헌 [Mazza et al., 2013, ACS Nano. 2013 Feb 26;7(2):1016-26], [Uchegbu and Siew, 2013, J Pharm Sci. 102(2):305-10] 및 [Lalatsa et al., 2012, J Control Release. 2012 Jul 20; 161(2):523-36]을 참조한다.

[0661] 미국 공개 특허 출원 제20050019923호는 생체 분자, 예컨대, 폴리뉴클레오티드 분자, 펩티드 및 폴리펩티드 및/또는 약학제를 포유동물 신체에 전달하기 위한 양이온성 덴드리머를 기재한다. 덴드리머는, 예를 들어, 간, 비장, 폐, 신장 또는 심장 (또는 심지어 뇌)에 대한 생활성 분자의 전달을 표적화하는 데 적합하다. 덴드리머는 단순 분지형 단량체 단위로부터 단계식-방식으로 제조되는 합성 3-차원 거대분자로서, 이의 성질 및 기능성은 쉽게 제어할 수 있으며 다양할 수 있다. 덴드리머는 다작용성 코어에 대해 (합성에 대한 분기적 접근), 또는 다작용성 코어를 향한 (합성에 대한 수렴적 접근) 빌딩 블록의 반복 첨가로부터 합성되고, 빌딩 블록의 3차원 껍질의 각각의 첨가는 더 높은 세대의 덴드리머의 형성을 야기한다. 폴리프로필렌이민 덴드리머는 디아미노부탄 코어로부터 시작되는데, 이에 대해 1차 아민에 대한 아크릴로나이트릴의 이중 마이클 첨가에 의해 2배의 아민 기 수가 첨가된 후에, 나이트릴의 수소화가 이어진다. 이는 아민 기의 배가를 초래한다. 폴리프로필렌이민 덴드리머는 100% 양성자화 가능한 질소 및 64개까지의 말단 아미노 기 (5 세대, DAB 64)

를 함유한다. 양성자화 가능한 기는 보통 중성 pH에서 양성자를 받아들일 수 있는 아민 기이다. 유전자 전달제로서 덴드리머의 사용은 대체로 유전자 전달을 위해 저급 세대 폴리프로필렌이민 덴드리머의 사용에 대한 작업이 보고되지 않은, 각각 접합 유닛으로서 아민/아미드 또는 N--P(O_2)S의 혼합물을 갖는 인 함유 화합물, 및 폴리아미도아민의 사용에 집중되었다. 폴리프로필렌이민 덴드리머는 또한 주변 아미노 기에 의해 화학적으로 변형될 때 게스트 분자의 그들 캡슐화 및 약물 절단을 위한 pH 민감성 제어 방출 시스템으로서 연구되었다.

폴리프로필렌이민 덴드리머의 DNA와의 상호작용 및 세포독성을 비롯하여 DAB 64의 형질감염 효능이 또한 연구되었다.

[0662] 미국 공개 특허 출원 제20050019923호는 보다 앞선 보고와 대조적으로, 양이온성 덴드리머, 예컨대, 폴리프로필렌이민 덴드리머가 생활성 분자, 예컨대, 유전자 물질의 표적화된 전달에서 사용하기 위한 적합한 특성, 예컨대, 특이적 표적화 및 저독성을 나타낸다는 관찰에 기반하였다. 추가로, 양이온성 덴드리머의 유도체는 또한 생활성 분자의 표적화된 전달을 위한 적합한 특성을 나타낸다. 또한 미국 공개 특허 출원 제20080267903호의 생활성 중합체를 참조하는데, 여기서는 "양이온성 폴리아민 중합체 및 덴드리머 중합체를 비롯한 다양한 중합체가 항증식 활성을 갖는 것으로 나타나며, 따라서 원치않는 세포 증식을 특징으로 하는 장애, 예컨대, 신생물 및 종양, 염증 장애(자가면역 장애를 포함), 건선 및 죽상동맥경화증의 치료에 유용할 수 있다.

중합체는 단독으로 활성제로서, 또는 다른 치료제, 예컨대, 유전자 요법을 위한 약물 분자 또는 핵산에 대한 전달 비히클로서 사용될 수 있다. 이러한 경우에, 중합체 자체의 고유한 항종양 활성은 전달될 제제의 활성을 보완할 수 있다"는 것을 개시한다. 이들 특허 공보의 개시내용은 AD-기능화된 CRISPR Cas 시스템(들) 또는 이의 성분(들) 또는 이를 코딩하는 핵산 분자(들)의 전달을 위해 본 명세서에서 함께 사용될 수 있다.

[0663] **초전하 단백질**

[0664] 초전하 단백질은 보통 높은 순 이론적 양전하 또는 음전하를 갖는 조작되거나 또는 천연 발생인 단백질의 부류이고, AD-기능화된 CRISPR Cas 시스템(들) 또는 이의 성분(들) 또는 이를 코딩하는 핵산 분자(들)의 전달에서 사용될 수 있다. 초음전하 단백질과 초양전하 단백질은 둘 모두 열적으로 또는 화학적으로 유도된 응집을 견뎌내는 현저한 능력을 나타낸다. 초양전하 단백질은 또한 포유동물 세포에 침투할 수 있다. 이들 단백질, 예컨대, 플라스미드 DNA, RNA, 또는 다른 단백질과 카고의 회합은 시험관내와 생체내 포유동물 세포 둘 모두에서 이들 거대분자의 기능적 전달을 가능하게 할 수 있다. 초전하 단백질의 생성 및 특징규명은 2007년에 보고되었다 (Lawrence et al., 2007, Journal of the American Chemical Society 129, 10110-10112).

[0665] 포유동물 세포 내로 RNA 및 플라스미드 DNA의 비바이러스 전달은 연구와 치료적 적용 둘 모두에 가치 있다 (Akinc et al., 2010, Nat. Biotech. 26, 561-569). 정제된 +36 GFP 단백질 (또는 다른 초양전하 단백질)은 적절한 무혈청 배지에서 RNA와 혼합되고, 세포에 첨가 전에 복합체를 형성되게 한다. 이 단계에서 혈청의 포함은 초하전 단백질-RNA 복합체의 형성을 억제하고, 치료 유효성을 감소시킨다. 다양한 세포주에 대해 다음의 프로토콜이 유효한 것으로 확인되었다 (McNaughton et al., 2009, Proc. Natl. Acad. Sci. USA 106, 6111-6116) (그러나, 단백질 및 RNA 용량을 달리한 파일럿 실험은 특정 세포주에 대한 절차를 최적화하기 위해 수행되어야 한다): (1) 치료 1일 전에, 48-웰 플레이트에서 웰 당 1×10^5개의 세포를 플레이팅한다. (2) 치료일에, 무혈청 배지 내 정제된 +36 GFP 단백질을 최종 농도 200 nM로 희석시킨다. 50 nM의 최종 농도로 RNA를 첨가한다. 혼합물을 교반시키고, 실온에서 10분 동안 인큐베이션시킨다. (3) 인큐베이션 동안, 세포로부터 배지를 흡입하고 나서, PBS로 1회 세척한다. (4) +36 GFP와 RNA의 인큐베이션 후에, 단백질-RNA 복합체를 세포에 첨가한다. (5) 37℃에서 4시간 동안 세포를 복합체와 함께 인큐베이션시킨다. (6) 인큐베이션 후에, 배지를 흡입하고 나서, 20 U/㎖ 헤파린 PBS로 3회 세척한다. 세포를 혈청-함유 배지와 함께 활성에 대한 어세이에 따라서 추가 48시간 이상 동안 인큐베이션시킨다. (7) 면역블롯, qPCR, 표현형 어세이, 또는 다른 적절한 방법에 의해 세포를 분석한다.

[0666] +36 GFP는 다양한 세포에서 효과적인 플라스미드 전달 시약인 것이 추가로 확인되었다. 플라스미드 DNA는 siRNA보다 더 큰 카고이기 때문에, 비례적으로 더 많은 +36 GFP 단백질이 효과적으로 복잡한 플라스미드에 필요하다. 효과적인 플라스미드 전달을 위해, 출원인들은 인플루엔자 바이러스 혈구응집소 단백질로부터 유래된 공지된 엔도솜-붕괴 펩티드인 C-말단의 HA2 펩티드 태그를 보유하는 +36 GFP의 변이체를 개발하였다. 다음의 프로토콜은 다양한 세포에서 효과적이었지만, 상기와 같이, 플라스미드 DNA 및 초전하 단백질 용량이 특정 세포주 및 전달 용도에 최적화된다는 것을 알게 되었다: (1) 처리 1일 전에, 48-웰 플레이트에서 웰 당 1×10^5를 플레이팅한다. (2) 처리일에, 무혈청 배지 내 þ36 GFP 단백질을 최종 농도 2 ㎛로 희석시킨다. 1 ㎎의 플라스미드 DNA를 첨가한다. 혼합물을 교반시키고, 실온에서 10분 동안 인큐베이션시킨다. (3) 인

큐베이션 동안, 세포로부터 배지를 흡입하고 PBS로 1회 세척한다. (4) þ36 GFP와 플라스미드 DNA의 인큐베이션 후에, 단백질-DNA 복합체를 세포에 부드럽게 첨가한다. (5) 37℃에서 4시간 동안 세포를 복합체와 인큐베이션시킨다. (6) 인큐베이션 후에, 배지를 흡입하고 나서, PBS로 세척한다. 혈청-함유 배지에서 세포를 인큐베이션시키고 나서, 추가 24 내지 48시간 동안 인큐베이션시킨다. (7) 적절하다면, 플라스미드 전달을 (예를 들어, 플라스미드-유도 유전자 발현에 의해) 분석한다.

[0667] 또한, 예를 들어, 문헌 [McNaughton et al., Proc. Natl. Acad. Sci. USA 106, 6111-6116 (2009)], [Cronican et al., ACS Chemical Biology 5, 747-752 (2010)], [Cronican et al., Chemistry & Biology 18, 833-838 (2011)], [Thompson et al., Methods in Enzymology 503, 293-319 (2012)], [Thompson, D.B., et al., Chemistry & Biology 19 (7), 831-843 (2012)]을 참조한다. 초전하 단백질 방법이 사용되고/되거나 본 발명의 AD-기능화된 CRISPR Cas 시스템의 전달에 적합화될 수 있다. 본 명세서의 교시와 함께 이들 시스템은 AD-기능화된 CRISPR Cas 시스템(들) 또는 이의 성분(들) 또는 이를 코딩하는 핵산 분자(들)의 전달에서 사용될 수 있다.

[0668] **세포 침투성 펩티드(CPP)**

[0669] 또 다른 실시형태에서, AD-기능화된 CRISPR Cas 시스템의 전달을 위한 세포 침투성 펩티드(CPP)가 고려된다. CPP는 다양한 분자 카고(나노크기 입자로부터 작은 화학적 분자 및 DNA의 거대 단편까지)의 세포 흡수를 용이하게 하는 짧은 펩티드이다. 본 명세서에서 사용되는 용어 "카고"는 치료제, 진단 프로브, 펩티드, 핵산, 안티센스 올리고뉴클레오티드, 플라스미드, 단백질, 입자(나노입자를 포함), 리포솜, 발색단, 소형 분자 및 방사성 물질로 이루어진 군을 포함하지만, 이들로 제한되지 않는다. 본 발명의 양상에서, 카고는 또한 AD-기능화된 CRISPR Cas 시스템 또는 전체 AD-기능화된 기능성 CRISPR Cas 시스템의 임의의 성분을 포함할 수 있다.

본 발명의 양상은 하기 단계들을 포함하는, 목적하는 카고를 대상체에 전달하는 방법을 추가로 제공한다: (a) 본 발명의 세포 침투성 펩티드 및 목적하는 카고를 포함하는 복합체를 준비하는 단계, 및 (b) 대상체에 복합체를 경구로, 관절내로, 복강내로, 초내로, 동맥내로, 비강내로, 실질세포내로, 피하로, 근육내로, 정맥내로, 피내로, 직장내로, 또는 국소로 투여하는 단계. 카고는 공유 결합을 통한 화학적 결합을 통해 또는 비공유 상호작용을 통해 펩티드와 회합된다.

[0670] CPP의 기능은 카고를 세포에 전달하는 것으로서, 살아있는 포유동물 세포의 엔도솜에 전달되는 카고에 의한 세포내 이입을 통해 통상적으로 일어나는 과정이다. 세포-침투성 펩티드는 상이한 크기, 아미노산 서열 및 전하를 가지지만, 모든 CPP는 한가지 구별되는 특징을 갖는데, 그것은 혈장막을 전좌하고 세포질 또는 유기체에 다양한 분자 카고의 전달을 용이하게 하는 능력이다. CPP 전좌는 3가지 주요 유입 메커니즘으로 분류될 수 있다: 막 내 직접 침투, 세포내 이입-매개 유입 및 전이 구조의 형성을 통한 전좌. CPP는 암 및 바이러스 억제제를 비롯한 상이한 질환의 치료에서의 약물 전달제뿐만 아니라 세포 표지를 위한 조영제로서 의학에서의 수많은 용도가 발견되었다. 후자의 예는 GFP에 대한 담체, MRI 조영제, 또는 양자점으로서의 작용을 포함한다. CPP는 연구 및 의학에서 사용하기 위한 시험관내 및 생체내 전달 벡터로서 큰 가능성을 보유한다.

CPP는 전형적으로 리신 또는 아르기닌과 같은 양으로 하전된 아미노산의 높은 상대적 존재비를 함유하거나 또는 극성/하전 아미노산 및 비극성의 소수성 아미노산의 교대식 패턴을 함유하는 서열을 갖는, 아미노산 조성물을 갖는다. 이들 두 유형의 구조는 각각 다가양이온 또는 양쪽성으로서 지칭된다. CPP의 제3 부류는 낮은 순전하를 갖는 비극성 잔기만을 함유하는 소수성 펩티드이거나 또는 세포 흡수에 중요한 소수성 아미노산기를 갖는다. 발견된 초기 CPP 중 하나는 배양물 내 수많은 세포 유형에 의해 주변 배지로부터 효율적으로 흡수되는 것으로 발견된 인간 면역결핍 바이러스 1 (HIV-1) 유래 트랜스-활성화 전사 활성인자 (Tat)였다. 그 이후로, 공지된 CPP의 수는 상당히 확장되며 더 효과적인 단백질 형질도입 특성을 갖는 소형 분자 합성 유사체가 생성되었다. CPP는 페네트라틴 (Penetratin), Tat (48-60), 트랜스포탄 (Transportan) 및 (R-AhX-R4)(Ahx=아미노헥사노일)을 포함하지만, 이들로 제한되지 않는다.

[0671] 미국 특허 제8,372,951호는 고도의 세포 침투성 효율 및 낮은 독성을 나타내는 호산구 양이온성 단백질 (ECP)로부터 유래된 CPP를 제공한다. CPP와 이의 카고를 척추동물 대상체에 전달하는 양태가 또한 제공된다. CPP 및 그들의 전달의 추가적인 양태는 미국 특허 제8,575,305호; 제8:614,194호 및 제8,044,019호에 기재되어 있다. CPP는 AD-기능화된 CRISPR-Cas 시스템 또는 이의 성분을 전달하는 데 사용될 수 있다. AD-기능화된 CRISPR-Cas 시스템 또는 이의 성분을 전달하는 데 사용될 수 있는 그러한 CPP는 본 명세서에 전문이 참고로 편입되는, [Suresh Ramakrishna, Abu-Bonsrah Kwaku Dad, Jagadish Beloor, et. al. Genome Res. 2014 Apr 2]에 의한 원고 "Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA"에서 제공되며, 여기서는 CPP-접합 재조합 Cas9 단백질 및 CPP-복합체화된 가이드 RNA에 의한 처리가

인간 세포주에서 내생성 유전자 붕괴를 야기한다는 것을 입증한다. 논문에서, Cas9 단백질은 티오에테르 결합을 통해 CPP에 접합된 반면, 가이드 RNA는 CPP와 복합체화되어, 축합된, 양으로 하전된 입자를 형성하였다.

변형된 Cas9 및 가이드 RNA에 의한, 배아 줄기 세포, 진피 섬유아세포, HEK293T 세포, HeLa 세포 및 배아 암종 세포를 포함하는 인간 세포의 동시 및 후속적 처리는 플라스미드 형질감염에 비해 감소된 오프-표적 돌연변이를 갖는 효율적인 유전자 붕괴를 야기하는 것으로 나타났다.

[0672] **에어로졸 전달**

[0673] 폐 질환에 대해 치료된 대상체는, 예를 들어 자발적으로 호흡하면서 폐에 대해 기관지내삽관으로 전달되는 약학적 유효량의 에어로졸화된 AAV 벡터 시스템을 받을 수 있다. 이와 같이, 일반적으로 AAV 전달에 에어로졸화된 전달이 바람직하다. 전달을 위해 아데노바이러스 또는 AAV 입자가 사용될 수 있다. 각각 하나 이상의 조절 서열에 작동적으로 연결된 적합한 유전자 구성체는 전달 벡터에 클로닝될 수 있다.

[0674] **패키징 및 프로모터**

[0675] CRISPR-Cas 단백질 및 아데노신 데아미나제 코딩 핵산 분자 발현을 구동시키기 위해 사용되는 프로모터는 AAV ITR을 포함할 수 있고, 이것이 프로모터로서 작용할 수 있다. 이는 추가 프로모터 구성요소 (벡터 내 공간을 차지할 수 있음)의 필요를 제거하는데 유리하다. 추가의 공간 개방은 추가 구성요소 (gRNA 등)의 발현을 구동하는데 사용될 수 있다. 또한, ITR 활성은 상대적으로 더 약하며, 따라서 Cas13의 과발현에 기인하여 잠재적 독성을 감소시키는 데 사용될 수 있다.

[0676] 편재성 발현을 위해, 사용될 수 있는 프로모터는 CMV, CAG, CBh, PGK, SV40, 페리틴 중쇄 또는 경쇄 등을 포함한다. 뇌 또는 다른 CNS 발현의 경우, 시냅신 I이 모든 뉴런에 대해 사용될 수 있고, CaMKII알파는 흥분 뉴런에 대해 사용될 수 있으며, GAD67 또는 GAD65 또는 VGAT는 GABAergic 뉴런에 대해 사용될 수 있다. 간 발현에 대해, 알부민 프로모터가 사용될 수 있다. 폐 발현에 대해, SP-B가 사용될 수 있다. 내피 세포에 대해, ICAM이 사용될 수 있다. 조혈 세포에 대해, IFN베타 또는 CD45가 사용될 수 있다. 골아세포에 대해, OG-2가 사용될 수 있다.

[0677] 가이드 RNA를 구동시키는데 사용되는 프로모터는 Pol III 프로모터, 예컨대, U6 또는 H1을 비롯하여, Pol II 프로모터의 사용 및 가이드 RNA의 발현을 위한 인트론 카세트를 포함할 수 있다.

[0678] **아데노 연합 바이러스 (AAV)**

[0679] CRISPR-Cas 단백질, 아데노신 데아미나제, 및 하나 이상의 가이드 RNA는 아데노 연합 바이러스 (AAV), 렌티바이러스, 아데노바이러스 또는 다른 플라스미드 또는 바이러스 벡터 유형을 이용하여, 특히, 예를 들어, 미국 특허 제8,454,972호 (제제, 아데노바이러스에 대한 용량), 제8,404,658호 (제제, AAV에 대한 용량) 및 제5,846,946호 (제제, DNA 플라스미드에 대한 용량)으로부터의 용량, 및 임상 시험 및 렌티바이러스, AAV 및 아데노바이러스를 수반하는 임상 시험에 관한 간행물로부터의 용량을 이용하여, 전달될 수 있다. 예를 들어, AAV에 대해, 투여 경로, 제형 및 용량은 미국 특허 제8,454,972호와 같고, AAV를 수반하는 임상 시험과 같을 수 있다.

아데노바이러스에 대해, 투여 경로, 제형 및 용량은 미국 특허 제8,404,658호와 같고 아데노바이러스를 수반하는 임상 시험과 같을 수 있다. 플라스미드 전달을 위해, 투여 경로, 제형 및 용량은 미국 특허 제5,846,946호에서와 같고 플라스미드를 수반하는 임상 연구에서와 같을 수 있다. 용량은 평균 70 kg 개체 (예를 들어, 남성 성인 인간)에 기반하거나 또는 이에 대해 추론될 수 있고, 상이한 체중 및 종의 환자, 대상체, 포유동물에 대해 조절될 수 있다. 투여 빈도는 환자 또는 대상체의 연령, 성별, 일반적 건강상태, 다른 병태 및 처리될 특정 병태 또는 증상을 비롯한 보통의 인자에 따라서, 의학적 또는 수의학적 실행자 (예를 들어, 의사, 수의사)의 영역 내이다. 바이러스 벡터는 관심 조직 내로 주사될 수 있다. 세포 유형 특이적 게놈 변형을 위해, Cas13 및 아데노신 데아미나제의 발현은 세포 유형 특이적 프로모터에 의해 구동될 수 있다. 예를 들어, 간-특이적 발현은 알부민 프로모터를 사용하고, 뉴런-특이적 발현은 (예를 들어, CNS 장애 표적화를 위해) 시냅신 I 프로모터를 사용할 것이다.

[0680] 생체 전달과 관련하여, AAV는 몇가지 이유때문에 다른 바이러스 벡터에 비해 유리하다: 낮은 독성 (이는 면역반응을 활성화시킬 수 있는 세포 입자의 초원심분리를 필요로 하지 않는 정제 방법에 기인할 수 있음); 및 숙주 게놈 내로 통합되지 않기 때문에 삽입 돌연변이유발을 야기할 낮은 가능성.

[0681] AAV는 4.5 또는 4.75 Kb의 패키징 한계를 갖는다. 이것은 Cas13뿐만 아니라 프로모터 및 전사 종결자가 동일한 바이러스 벡터에 모두 적합하게 될 것임 의미한다. 4.5 또는 4.75 Kb보다 더 큰 구성체는 상당히 감소된 바이러스 생성을 야기할 것이다. SpCas9는 상당히 크며, 유전자 그 자체는 4.1 Kb인데, 이는 AAV에 패키

징되는 것을 어렵게 만든다. 따라서, 본 발명의 실시형태는 더 짧은 Cas13의 상동체를 이용하는 것을 포함한다.

[0682] AAV에 대해, AAV는 AAV1, AAV2, AAV5 또는 이들의 임의의 조합일 수 있다. 표적화하려는 세포에 대한 AAV의 AAV를 선택할 수 있으며; 예를 들어, 뇌 또는 뉴런 세포를 표적화하기 위해 AAV 혈청형 1, 2, 5 또는 하이브리드 캡시드 AAV1, AAV2, AAV5 또는 이들의 임의의 조합을 선택할 수 있고; 심장 조직을 표적화하기 위해 AAV4를 선택할 수 있다. AAV8은 간에 대한 전달에 유용하다. 본 명세서의 프로모터 및 벡터는 개별적으로 바람직하다. 이들 세포에 대한 소정의 AAV 혈청형의 표 (Grimm, D. et al, J. Virol. 82: 5887-5911(2008))는 다음과 같다:

세포주	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2.5	0.0	0.1	10	0.7	0.0
HEK293	25	100	2.5	0.1	0.1	5	0.7	0.1
HeLa	3	100	2.0	0.1	6.7	1	0.2	0.1
HepG2	3	100	16.7	0.3	1.7	5	0.3	ND
Hep1A	20	100	0.2	1.0	0.1	1	0.2	0.0
911	17	100	11	0.2	0.1	17	0.1	ND
CHO	100	100	14	1.4	333	50	10	1.0
COS	33	100	33	3.3	5.0	14	2.0	0.5
MeWo	10	100	20	0.3	6.7	10	1.0	0.2
NIH3T3	10	100	2.9	2.9	0.3	10	0.3	ND
A549	14	100	20	ND	0.5	10	0.5	0.1
HT1180	20	100	10	0.1	0.3	33	0.5	0.1
단핵구	1111	100	ND	ND	125	1429	ND	ND
미성숙 DC	2500	100	ND	ND	222	2857	ND	ND
성숙 DC	2222	100	ND	ND	333	3333	ND	ND

[0683]

[0684] **렌티바이러스**

[0685] 렌티바이러스는 유사분열 세포와 유사분열 후 세포 둘 모두에서 그들의 유전자를 감염시키고 발현시키는 능력을 갖는 복잡한 레트로바이러스이다. 가장 통상적으로 알려진 렌티바이러스는 인간 면역결핍 바이러스(HIV)인데, 이는 매우 다양한 세포 유형을 표적화하기 위해 다른 바이러스의 외피 당단백질을 사용한다.

[0686] 렌티바이러스는 다음과 같이 제조될 수 있다. pCasES10 (렌티바이러스 전달 플라스미드 골격을 함유함)을 클로닝시킨 후에, 저계대수 (p=5)의 HEK293FT를10% 소 태아 혈청 존재 및 항생제 부재의 DMEM 중에서 형질감염 전날에 T-75 플라스크에서 50% 합류점까지 파종하였다. 20시간 후에, 배지를 OptiMEM (무혈청) 배지로 바꾸고 나서, 4시간 후에 형질감염을 수행하였다. 세포를 10 ㎍의 렌티바이러스 전달 플라스미드(pCasES10) 및 다음의 패키징 플라스미드: 5 ㎍의 pMD2.G (VSV-g 위형) 및 7.5 ㎍의 psPAX2(gag/pol/rev/tat)로 형질감염시켰다. 양이온성 지질 전달제 (50 ㎕의 Lipofectamine 2000 및 100 ㎕ Plus 시약) 존재의 4 ㎖ OptiMEM 중에서 형질감염을 수행하였다. 6시간 후에, 배지를 10% 소 태아 혈청이 있는 무항생제 DMEM으로 바꾸었다. 이들 방법은 세포 배양 동안 혈청을 사용하지만, 무혈청 방법이 바람직하다.

[0687] 렌티바이러스는 다음과 같이 정제될 수 있다. 48시간 후에 바이러스 상청액을 채취하였다. 상청액은 먼저 찌꺼기를 제거하고 0.45 ㎛ 저 단백질 결합 (PVDF) 필터를 통해 여과된다. 이어서, 24,000 rpm에서 2시간 동안 초원심분리로 스피닝되었다. 바이러스 펠렛은 밤새 4℃에서 50 ㎕의 DMEM 중에서 재현탁되었다. 이어서, 그들은 분취되어, 즉시 -80℃에 냉동되었다.

[0688] 다른 실시형태에서, 말 전염성 빈혈 바이러스 (EIAV)에 기반한 최소 비영장류 렌티바이러스 벡터가 또한 특히 안구 유전자 요법에 대해 고려된다 (예를 들어, 문헌 [Balagaan, J Gene Med 2006; 8: 275 - 285] 참조). 다른 실시형태에서, 노인성 황반 변성의 그물망 형태 치료를 위해 망막하 주사를 통해 전달되는 혈관억제 단백질 엔도스타틴 및 앤지오스타틴을 발현시키는 말 전염성 빈혈 바이러스-기반 렌티바이러스 유전자 요법 벡터, RetinoStat®가 또한 고려되며 (예를 들어, 문헌 [Binley et al., HUMAN GENE THERAPY 23:980-991(September 2012)] 참조), 이 벡터는 본 발명의 AD-기능화된 CRISPR-Cas 시스템에 대해 변형될 수 있다.

[0689] 다른 실시형태에서, HIV tat/rev, 핵소체-국재화 TAR 디코이, 및 항-CCR5-특이적 해머해드 리보자임에 의해 공

유되는 공통 엑손을 표적화하는 siRNA를 갖는 자가-불활성화 렌티바이러스 벡터 (예를 들어, 문헌 [DiGiusto et al. (2010) Sci Transl Med 2:36ra43] 참조)가 사용될 수 있고/있거나 본 발명의 AD-기능화된 CRISPR-Cas 시스템에 적합화 될 수 있다. 환자의 체중 킬로그램 당 최소 2.5 × 10^6개의 CD34+ 세포가 수집되고, 2 μmol/L-글루타민, 줄기 세포 인자 (100 ng/㎖), Flt-3 리간드 (Flt-3L) (100 ng/㎖), 및 트롬보포이어틴 (10 ng/㎖)(CellGenix)을 함유하는 X-VIVO 15 배지 (Lonza)에서 2 × 10^6개 세포/㎖의 밀도로 16 내지 20시간 동안 사전자극하였다. 사전 자극된 세포는 피브로넥틴 (25 ㎎/㎠)(RetroNectin, Takara Bio Inc.)으로 코팅한 75-㎠ 조직 배양 플라스크에서 16 내지 24시간 동안 5의 감염 다중도로 렌티바이러스를 형질도입시킬 수 있다.

[0690] 렌티바이러스 벡터는 파킨슨병에 대한 치료에 대해 개시되어 있고, 예를 들어, 미국 공개 특허 출원 제20120295960호 및 미국 특허 제7303910호 및 7351585호를 참조한다. 렌티바이러스 벡터는 또한 안질환의 치료에 대해 개시되었고, 예를 들어, 미국 공개 특허 출원 제20060281180호, 제20090007284호, 제20110117189; 제20090017543호; 제20070054961호, 제20100317109호를 참조한다. 렌티바이러스 벡터는 또한 뇌에 대한 전달에 대해 개시되었다. 예를 들어, 미국 공개 특허 출원 제20110293571호; 제20110293571호, 제20040013648호, 제20070025970호, 제20090111106호 및 미국 특허 제7259015호를 참조한다.

[0691] **비동물 유기체에서의 적용**

[0692] AD-기능화된 CRISPR 시스템(들)(예를 들어, 단일 또는 다중복합)이 작물 게놈에서의 최근의 진보와 함께 사용될 수 있다. 본 명세서에 기재된 시스템은 효율적이고 비용 효과적인 식물 유전자 또는 게놈 심문 (genome interrogation) 또는 편집 또는 조작을 수행하기 위해 - 예를 들어, 식물 유전자 또는 게놈의 빠른 연구 및/또는 선택 및/또는 심문 및/또는 비교 및/또는 조작 및/또는 형질전환을 위해; 예를 들어, 식물(들)에 형질(들) 또는 특징(들)을 생성하거나, 식별하거나, 발생시키거나, 최적화하거나 또는 부여하기 위해 또는 식물 게놈을 형질전환시키기 위해 사용될 수 있다. 따라서, 식물, 형질 또는 특징의 새로운 조합을 갖는 새로운 식물들 또는 증강된 형질을 갖는 새로운 식물의 개선된 생산이 있을 수 있다. AD-기능화된 CRISPR 시스템은 부위-지정 통합 (Site-Directed Integration: SDI) 또는 유전자 편집 (Gene Edit: GE) 또는 임의의 근접 리버스 육종 (Near Reverse Breeding: NRB) 또는 리버스 육종 (Reverse Breeding: RB) 기법으로 식물에 관해 사용될 수 있다. 본 명세서에 기재된 Cas13 이펙터 단백질 시스템을 이용하는 양상은 식물에서 CRISPR-Cas (예를 들어, CRISPR-Cas9) 시스템의 사용과 유사할 수 있으며, 유니버시티 오브 아리조나 웹사이트 "CRISPR-PLANT" (http://www.genome.arizona.edu/crispr/)(Penn State 및 AGI가 지원)가 언급된다. 본 발명의 실시형태는 식물에서의 게놈 편집에서 또는 RNAi 또는 유사한 게놈 편집 기법이 이전에 사용된 경우에 사용될 수 있다; 예를 들어, [Nekrasov, "Plant genome edit made easy: targeted mutagenesis in model and crop plants using the CRISPR-Cas system," Plant Methods 2013, 9:39 (doi:10.1186/1746-4811-9-39)]; [Brooks, "Efficient gene edit in tomato in the first generation using the CRISPR-Cas9 system," Plant Physiology September 2014 pp 114.247577]; [Shan, "targeted genome modification of crop plants using a CRISPR-Cas system," Nature Biotechnology 31, 686-688 (2013)]; [Feng, "Efficient genome edit in plants using a CRISPR-Cas system," Cell Research (2013) 23:1229-1232. doi:10.1038/cr.2013.114; 2013년 8월 20일자로 온라인 공개]; [Xie, "RNA-guided genome edit in plants using a CRISPR-Cas system," Mol Plant. 2013 Nov;6(6):1975-83. doi: 10.1093/mp/sst119. Epub 2013 Aug 17; [Xu, "Gene targeting using the Agrobacterium tumefaciens-mediated CRISPR-Cas system in rice," Rice 2014, 7:5 (2014), Zhou et al., "Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate: CoA ligase specificity and Redundancy," New Phytologist (2015) (Forum) 1-4 (www.newphytologist.com에서 온라인으로만 이용가능)]; [Caliando et al, "targeted DNA degradation using a CRISPR device stably carried in the host genome, NATURE COMMUNICATIONS 6:6989, DOI: 10.1038/ncomms7989, www.nature.com/naturecommunications DOI: 10.1038/ncomms7989]; 미국 특허 제6,603,061호 (발명의 명칭: Agrobacterium-Mediated Plant Transformation Method; 미국 특허 제7,868,149호 (발명의 명칭: Plant Genome sequences and Uses Thereof 및 미국 특허 제2009/0100536호 (발명의 명칭: Transgenic Plants with Enhanced Agronomic Traits)를 참조하며, 이들 각각의 모든 내용 및 개시내용은 본 명세서에 그들의 전문이 참고로 편입된다. 본 발명의 실행에서, 본 명세서의 실시형태가 식물에 대해서 사용될 수 있는 방법에 대해, 문헌 [Morrell et al "Crop genomics: advances and applications," Nat Rev Genet. 2011 Dec 29;13(2):85-96]의 내용 및 개시내용을 참조하고; 이들 각각은 본 명세서에 참고로 편입된다. 따라서, 본 명세서에서 동물 세포에 대한 언급은 또한 분명하지 않으면 준용하여 식물 세포에 적용될 수 있고, 감소된 오프-표적 효과를 갖는 본 명세서의 효소 및 이러한 효소를 적용하는 시스템은 본 명세서에 언급된 것을 포함하여, 식물 적용분야에서

사용될 수 있다.

[0693] **식물 및 효모에 AD-기능화된 CRISPR 시스템의 적용**

[0694] 일반적으로, 용어 "식물"은 세포 분할에 의해 특징적으로 성장하고, 엽록체를 함유하며 셀룰로스로 구성된 세포벽을 갖는, 식물계 (kingdom Plantae)의 임의의 다양한 광합성, 진핵, 단세포 또는 다세포 유기체에 관한 것이다. 용어 식물은 단자엽 및 쌍자엽 식물을 포함한다. 구체적으로, 식물은 제한없이 속씨식물 및 겉씨식물, 예컨대, 아카시아, 알팔파, 아마란스, 사과, 살구, 아티초크, 물푸레나무, 아스파라거스, 아보카도, 바나나, 보리, 콩, 비트, 자작나무, 너도밤나무, 블랙베리, 블루베리, 브로콜리, 양배추, 캐비지, 카놀라, 칸탈루프, 당근, 카사바, 콜리플라워, 삼나무, 곡류, 셀러리, 밤나무, 체리, 차이니즈 캐비지, 감귤류, 귤, 클로버, 커피, 옥수수, 목화, 동부, 오이, 사이프러스, 가지, 느릅나무, 엔다이브, 유칼립투스, 펜넬, 무화과, 전나무, 제라늄, 포도, 자몽, 땅콩, 파리, 솔송나무, 히코리, 케일, 키위, 콜라비, 낙엽송, 상추, 리크, 레몬, 라임, 로커스트, 소나무, 공작고사리, 메이즈, 망고, 단풍나무, 멜론, 낱알 곡물, 버섯, 겨자, 견과, 오크, 귀리, 기름야자, 오크라, 양파, 오렌지, 관상용 식물 또는 꽃 또는 나무, 파파야, 야자나무, 파슬리, 파스닙, 완두콩, 복숭아, 땅콩, 배, 피트(peat), 후추, 감나무, 피젼피, 소나무, 파인애플, 플랜틴, 자두, 석류, 감자, 호박, 라디치오, 무, 유채, 라즈베리, 벼, 호밀, 수수, 잇꽃, 갯버들, 대두, 시금치, 가문비나무, 스쿼시, 딸기, 사탕무, 사탕수수, 해바라기, 고구마, 단옥수수, 탄저린, 찻잎, 담배, 토마토, 나무, 라이밀, 잔디풀, 순무, 덩굴, 호두, 물냉이, 수박, 밀, 참마, 주목나무, 및 주키니를 포함하고자 한다. 용어 식물은 또한 뿌리, 잎 및 보다 고등 식물의 특징이 되는 다 기관의 결여에 의해 일차적으로 통합되는 광독립영양생물인 조류를 포함한다.

[0695] 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템을 이용하는 게놈 편집 방법은 본질적으로 임의의 식물 상에 목적하는 형질을 부여하는 데 사용될 수 있다. 매우 다양한 식물 및 식물 세포계는 본 개시내용의 핵산 구성체 및 상기 언급된 다양한 형질전환 방법을 이용하여 본 명세서에 기재된 목적하는 생리학적 및 작물학적 특징을 위해 조작될 수 있다. 바람직한 실시형태에서, 조작을 위한 표적 식물 및 식물 세포는 단자엽 및 쌍자엽 식물, 예컨대, 곡물 (예를 들어, 밀, 메이즈, 벼, 낱알 곡물, 보리), 과일 작물 (예를 들어, 토마토, 사과, 배, 딸기, 오렌지), 사료 작물 (예를 들어, 알팔파), 뿌리 식물 작물 (예를 들어, 당근, 감자, 사탕무, 참마), 잎줄기채소 작물 (예를 들어, 상추, 시금치); 개화 식물 (예를 들어, 페튜니아, 장미, 국화), 송백류 및 소나무 (예를 들어, 소나무 전나무, 가문비나무); 식물 환경 정화에서 사용되는 식물 (예를 들어, 중금속 축적 식물); 기름 작물(예를 들어, 해바라기, 평지씨) 및 실험 목적으로 사용되는 식물 (예를 들어, 애기장대)를 포함하는 작물을 포함하지만, 이들로 제한되지 않는다. 따라서, 상기 방법 및 시스템은 광범위한 식물, 예컨대, 마그니오랄레스 (*Magniolales*), 일리시알레스 (*Illiciales*), 라우랄레스 (*Laurales*), 피페랄레스 (*Piperales*), 아리스토치알레스 (*Aristochiales*), 님파에알레스 (*Nymphaeales*), 라눈쿠 랄레스 (*Ranunculales*), 파페베랄레스 (*Papeverales*), 사라세니아케아이 (*Sarraceniaceae*), 트로코덴드랄레스 (*Trochodendrales*), 하마멜리달레스 (*Hamamelidales*), 유코미알레스 (*Eucomiales*), 레이트네리알레스 (*Leitneriales*), 미리칼레스 (*Myricales*), 파갈레스 (*Fagales*), 카수아리날레스 (*Casuarinales*), 카리오필랄레스 (*Caryophyllales*), 바탈레스 (*Batales*), 폴리고날레스 (*Polygonales*), 플룸바지날레스 (*Plumbaginales*), 딜레니알레스 (*Dilleniales*), 테알레스 (*Theales*), 말발레스 (*Malvales*), 우르티칼레스 (*Urticales*), 레시티달레스 (*Lecythidales*), 비올랄레스 (*Violales*), 살리칼레스 (*Salicales*), 카파랄레스 (*Capparales*), 에리칼레스 (*Ericales*), 디아펜살레스 (*Diapensales*), 에베날레스 (*Ebenales*), 프리무랄레스 (*Primulales*), 로살레스 (*Rosales*), 파발레스 (Fabales), 포도스테말레스 (*Podostemales*), 할로라갈레스 (*Haloragales*), 미르탈레스 (*Myrtales*), 코르날레스 (*Cornales*), 프로테알레스 (*Proteales*), 산탈레스 (*Santales*), 라플레시알레스 (*Rafflesiales*), 셀라스트랄레스 (*Celastrales*), 유포르비알레스 (*Euphorbiales*), 람날레스 (*Rhamnales*), 사핀달레스 (*Sapindales*), 유글란달레스 (*Juglandales*), 게라니알레스 (*Geraniales*), 폴리갈랄레스 (*Polygalales*), 움벨랄레스 (*Umbellales*), 겐티아날레스 (*Gentianales*), 폴레모니알레스 (*Polemoniales*), 라미알레스 (*Lamiales*), 플란타지날레스 (*Plantaginales*), 스크로폴라리알레스 (*Scrophulariales*), 캄파누랄레스 (*Campanulales*), 루비알레스 (*Rubiales*), 딥사칼레스 (*Dipsacales*) 및 아스테랄레스 (*Asterales*) 목에 속하는 쌍자엽 식물에 대해 사용될 수 있고; 상기 방법 및 CRISPR-Cas 시스템은 알리스마탈레스 (*Alismatales*), 히드로차리탈레스 (*Hydrocharitales*), 나자달레스 (*Najadales*), 트리우리달레스 (*Triuridales*), 콤멜리날레스 (*Commelinales*), 에리오카우랄레스 (*Eriocaulales*), 레스티오날레스 (*Restionales*), 포알레스 (*Poales*), 준칼레스 (*Juncales*), 시페랄레스 (*Cyperales*), 티팔레스 (*Typhales*), 브로멜리알레스 (*Bromeliales*), 진기베랄레스 (*Zingiberales*), 아레칼레스 (*Arecales*), 시클란탈레스 (*Cyclanthales*), 판다날레스 (*Pandanales*), 아랄레스 (*Arales*), 릴리알레스 (*Lilliales*) 및 오르치달레스 (*Orchidales*) 목에 속하는 것과 같은 단자엽식물, 또는 겉씨식물문에 속하는 식물, 예를 들어 피날레스 (*Pinales*), 징코알레스 (*Ginkgoales*), 시카달레스 (*Cycadales*)

및 네탈레스 (*Gnetales*) 목에 속하는 것들에 사용될 수 있다.

[0696] 본 명세서에 기재된 AD-기능화된 CRISPR 시스템 및 사용 방법은 이하의 쌍자엽, 단자엽 또는 겉씨식물 속의 비제한적 열거에 포함되는 광범위한 식물 종에 걸쳐 사용될 수 있다: 아트로파 (*Atropa*), 알세오다프네 (*Alseodaphne*), 아나카르디움 (*Anacardium*), 아라치스 (*Arachis*), 벨리쉬 미에디아 (*Beilschmiedia*), 브라시카 (*Brassica*), 카르타무스 (*Carthamus*), 코쿠루스 (*Cocculus*), 크로톤 (*Croton*), 쿠쿠미스 (*Cucumis*), 시트러스 (*Citrus*), 시트룰루스 (*Citrullus*), 캅시쿰 (*Capsicum*), 카타란투스 (*Catharanthus*), 코코스 (*Cocos*), 코페아 (*Coffea*), 쿠쿠르비타 (*Cucurbita*), 다우쿠스 (*Daucus*), 두구에티아 (*Duguetia*), 에스크스 콜지아 (*Eschscholzia*), 피쿠스 (*Ficus*), 프라가리아 (*Fragaria*), 글라우시움 (*Glaucium*), 글리신 (글리신), 고시 피움 (*Gossypium*), 헬리안투스 (*Helianthus*), 헤베아 (*Hevea*), 히오시아무스 (*Hyoscyamus*), 락투카 (*Lactuca*), 란돌피아 (*Landolphia*), 리눔 (*Linum*), 리트세아 (*Litsea*), 리코페르시콘 (*Lycopersicon*), 루피누스 (*Lupinus*), 마니호트 (*Manihot*), 마조라나 (*Majorana*), 말루스 (*Malus*), 메디카고 (*Medicago*), 니코티아나 (*Nicotiana*), 올레아 (*Olea*), 파르테니움 (*Parthenium*), 파파베르 (*Papaver*), 페르세아 (*Persea*), 파세올루스 (*Phaseolus*), 피스타치아 (*Pistacia*), 피숨 (*Pisum*), 피루스 (*Pyrus*), 프루누스 (*Prunus*), 라파누스 (*Raphanus*), 리치누스 (*Ricinus*), 세네치오 (*Senecio*), 시노메니움 (*Sinomenium*), 스테파니아 (*Stephania*), 시나피스 (*Sinapis*), 솔라눔 (*Solanum*), 테오브로마 (*Theobroma*), 트리폴리움 (*Trifolium*), 트리고넬라 (*Trigonella*), 비키아 (*Vicia*), 빈카 (*Vinca*), 비티스 (*Vitis*) 및 비그나 (*Vigna*); 및 알리움 (*Allium*), 안드로포곤 (*Andropogon*), 아라그로스티스 (*Aragrostis*), 아스파라거스 (*Asparagus*), 아베나 (*Avena*), 시노돈 (*Cynodon*), 엘라에리스 (*Elaeis*), 페스투카 (*Festuca*), 페스투로리움 (*Festulolium*), 헤테로칼리스 (*Heterocalis*), 호르데움 (*Hordeum*), 렘나 (*Lemna*), 롤리움 (*Lolium*), 무사 (*Musa*), 오리자 (*Oryza*), 파니쿰 (*Panicum*), 판네세툼 (*Pannesetum*), 플레움 (*Phleum*), 포아 (*Poa*), 세칼레 (*Secale*), 소르굼 (*Sorghum*), 트리티쿰 (*Triticum*) 및 제아(*Zea*) 속 외떡잎식물; 또는 아비에스 (*Abies*), 쿤닝하미아 (*Cunninghamia*), 피세아 (*Picea*), 피누스 (*Pinus*) 및 프세우도추가 (*Pseudotsuga*).

[0697] AD-기능화된 CRISPR 시스템 및 사용 방법은 또한, 예를 들어, 홍조식물 (Rhodophyta) (홍조류), 녹색식물문 (Chlorophyta) (녹조류), 갈조식물문 (Phaeophyta) (갈조류), 규조문 (Bacillariophyta) (규조류), 유스티그마토피타 (Eustigmatophyta) 및 와편모류 (dinoglagellates) 를 포함하는 몇몇 진핵생물문뿐만 아니라 원핵생물문 남세균 (cyanobacteria) (청-녹조류)로부터 선택된 조류를 비롯한, 광범위한 "조류" 또는 "조류 세포"에 대해 사용될 수 있다. 용어 "조류"는, 예를 들어, 암포라 (*Amphora*), 아나바나 (*Anabaena*), 아니크스트로데스미스 (*Anikstrodesmis*), 보트리오코커스 (*Botryococcus*), 차에토세로스 (*Chaetoceros*), 클라미도모나스 (*Chlamydomonas*), 클로렐라 (*Chlorella*), 클로로코쿰 (*Chlorococcum*), 시클로텔라 (*Cyclotella*), 실린드로테카 (*Cylindrotheca*), 두날리엘라 (*Dunaliella*), 에밀리아나 (*Emiliana*), 유글레나 (*Euglena*), 헤마토코커스 (*Hematococcus*), 이소크리시스 (*Isochrysis*), 모노크리시스 (*Monochrysis*), 모노라피디움 (*Monoraphidium*), 나노클로리스 (*Nannochloris*), 나노클로롭시스 (*Nannnochloropsis*), 나비쿨라 (*Navicula*), 네프로클로리스 (*Nephrochloris*), 네프로셀미스 (*Nephroselmis*), 니츠키아 (*Nitzschia*), 노둘라리아 (*Nodularia*), 노스톡 (*Nostoc*), 오크로모나스 (*Oochromonas*), 우시스티스 (Oocystis), 오실라르토리아 (*Oscillartoria*), 파블로바 (*Pavlova*), 파에오닥틸룸 (*Phaeodactylum*), 플라이트모나스 (*Playtmonas*), 플레우로크리시스 (*Pleurochrysis*), 포르히라 (*Porhyra*), 슈도아나바나 (*Pseudoanabaena*), 피라미모나스 (*Pyramimonas*), 스티코코커스 (*Stichococcus*), 시네코코커스 (*Synechococcus*), 시네코시스티스 (*Synechocystis*), 테트라셀미스 (*Tetraselmis*), 탈라시오시라 (*Thalassiosira*) 및 트리코데스뮴 (*Trichodesmium*)으로부터 선택된 조류를 포함한다.

[0698] 식물의 일부, 즉, "식물 조직"은 개선된 식물을 생산하기 위해 본 발명의 방법에 따라 처리될 수 있다. 식물 조직은 또한 식물 세포를 포함한다. 본 명세서에서 사용되는 용어 "식물 세포"는 무손상 전체 세포에서 또는 시험관내 조직 배양으로부터 성장된 단리 형태로, 배지 또는 한천 상에서, 성장 배지 내 현탁액 또는 완충제에서 또는 더 고등의 조직 단위의 일부, 예컨대, 식물 조직, 식물 기관 또는 전체 식물로서의, 살아있는 식물의 개개 단위를 의미한다.

[0699] "원형질체"는 세포벽을 재형성, 증식 및 재생할 수 있고, 적절한 성장 조건 하에 전체 식물로 성장할 수 있는 살아있는 식물의 무손상의 생화학적 적격 단위를 생성하는 기계적 또는 효소적 수단을 이용하여, 완전히 또는 부분적으로 제거되는 보호 세포벽을 갖는 식물 세포이다.

[0700] 용어 "형질전환"은 식물 숙주가 아그로박테리움 (*Agrobacterium*)에 의해서 또는 다양한 화학적 또는 물리적 방법 중 하나에 의해서 DNA의 도입에 의해 유전자 변형되는 과정을 광범위하게 의미한다. 본 명세서에서 사

용되는 용어 "식물 숙주"는 식물의 임의의 세포, 조직, 기관 또는 자손을 포함하는 식물을 의미한다. 다수의 적합한 식물 조직 또는 식물 세포는 형질전환될 수 있고, 원형질체, 체성 배아, 화분, 잎, 묘목, 줄기, 캘러스, 주근, 미세괴경 및 순을 포함하지만, 이들로 제한되지 않는다. 식물 조직은 또한 이러한 식물, 종자, 자손, 유성 또는 무성 생식이건 간에 번식체, 및 이들 중 어느 것의 후손, 예컨대, 꺾꽂이 순 또는 종자의 임의의 클론을 의미한다.

[0701] 본 명세서에서 사용되는 용어 "형질전환된"은 외래 DNA 분자, 예컨대, 구성체가 도입되는 세포, 조직 또는 유기체를 의미한다. 도입된 DNA 분자가 후속 자손으로 전달되도록, 도입된 DNA 분자는 수용 세포, 조직, 기관 또는 유기체의 게놈 DNA에 통합될 수 있다. 이들 실시형태에서, "형질전환된" 또는 "유전자이식" 세포 또는 식물은 또한 교배에서 모체로서 이러한 형질전환된 식물을 사용하고 도입된 DNA 분자의 존재로부터 초래되는 변경된 표현형을 나타내는 육종 프로그램으로부터 생성된 자손 또는 세포 또는 식물의 자손을 포함할 수 있다. 바람직하게는, 유전자이식 식물은 생식력이 있으며, 도입된 DNA가 성적 번식을 통해 자손에게 전달될 수 있다.

[0702] 용어 "자손", 예컨대, 유전자이식 식물의 자손은 식물 또는 유전자이식 식물로부터 기인하거나, 야기되거나 또는 유래되는 것이다. 도입된 DNA 분자는 또한 도입된 DNA 분자가 후속 자손에 의해 유전되지 않고, 따라서 "유전자이식"으로 고려되지 않도록 수용 세포 내로 일시적으로 도입될 수 있다. 따라서, 본 명세서에서 사용되는, "비-유전자이식" 식물 또는 식물 세포는 게놈 내로 안정하게 통합된 외래 DNA를 함유하지 않는 식물이다.

[0703] 본 명세서에서 사용되는 용어 "식물 프로모터"는 식물 유래이건 그렇지 않건 간에 식물 세포에서 전사를 개시할 수 있는 프로모터이다. 예시적인 적합한 식물 프로모터는 식물 세포에서 발현되는 유전자를 포함하는 식물, 식물 바이러스 및 박테리아, 예컨대, 아그로박테리움 또는 리조비움 (*Rhizobium*)으로부터 얻어지는 것을 포함하지만, 이들로 제한되지 않는다.

[0704] 본 명세서에서 사용되는 "진균 세포"는 진균류 내의 임의의 유형의 진행 세포를 의미한다. 진균류 내의 문은 아스코미코타 (*Ascomycota*), 바시디오마이코타 (*Basidiomycota*), 블라스토클라디오마이코타 (Blastocladiomycota), 키트리디오마이코타 (*Chytridiomycota*), 글로메로마이코타 (*Glomeromycota*), 마이크로스포리디아 (*Microsporidia*), 및 네오칼리마스티고마이코타 (*Neocallimastigomycota*)를 포함한다. 진균 세포는 효모, 곰팡이 및 사상성진균을 포함할 수 있다. 일부 실시형태에서, 진균 세포는 효모 세포이다.

[0705] 본 명세서에서 사용되는 용어 "효모 세포"는 아스코미코타 및 바시디오마이코타 문 내의 임의의 진균 세포를 의미한다. 효모 세포는 출아 효모 세포, 분열 효모 세포 및 곰팡이 세포를 포함할 수 있다. 이들 유기체로 제한되지 않지만, 실험실 및 산업 환경에서 사용되는 다수 유형의 효모는 아스코미코타 문의 일부이다. 일부 실시형태에서, 효모 세포는 에스. 세레비지아 (*S. cerervisiae*), 클루이베로마이세스 마르시아누스 (*Kluyveromyces marxianus*), 또는 아사켄키아 오리엔탈리스 (*Issatchenkia orientalis*) 세포이다. 다른 효모 세포는 칸디다 종 (*Candida* spp.)(예를 들어, 칸디다 알비칸스 (*Candida albicans*)), 야로위아 종 (*Yarrowia* spp.) (예를 들어, 야로위아 리폴리티카 (*Yarrowia lipolytica*)), 피키아 종 (*Pichia* spp.)(예를 들어, 피키아 파스토리스 (*Phichia pastoris*), 클루이베로마이세스 종 (*Kluyveromyces* spp.) (예를 들어, 클루이베로마이세스 락티스 (*Kluyveromyces lactis*) 및 클루이베로마이세스 마르시아누스), 뉴로스포라 종 (*Neurospora* spp.) (예를 들어, 뉴로스포라 크라사 (*Neurospora crassa*)), 푸사리움 종 (*Fusarium* spp.) (예를 들어, 푸사리움 옥시스포룸 (*Fusarium oxysporum*)), 및 이사켄키아 종 (*Issatchenkia* spp.) (예를 들어, 아사켄키아 오리엔탈리스 (*Issatchenkia orientalis*) (피키아 쿠드리아베제비 (*Pichia kudriavzevii*)라고도 함) 및 칸디다 아시도써모필룸 (*Candida acidothermophilum*))을 포함할 수 있지만, 이들로 제한되지 않는다. 일부 실시형태에서, 진균 세포는 사상 진균 세포이다. 본 명세서에서 사용되는 용어 "사상 진균 세포"는 필라멘트, 즉, 균사 또는 균사체로 성장하는 임의의 유형의 진균 세포를 의미한다. 사상 진균 세포의 예는 아스퍼질러스 종 (*Aspergillus* spp.) (예를 들어, 아스퍼질러스 니거 (*Aspergillus niger*)), 트리코더마 종 (*Trichoderma* spp.) (예를 들어, 트리코더마 리세이 (*Trichoderma reesei*)), 리조푸스 종 (*Rhizopus* spp.) (예를 들어, 리조푸스 오리자에 (*Rhizopus oryzae*)), 및 모르티에렐라 종 (*Mortierella* spp.) (예를 들어, 모르티에렐라 이사벨리나 (*Mortierella isabellina*))을 포함할 수 있지만, 이들로 제한되지 않는다.

[0706] 일부 실시형태에서, 진균 세포는 산업적 균주이다. 본 명세서에서 사용되는 "산업적 균주"는 산업 공정, 예를 들어, 상업적 또는 산업적 규모의 제품 생산에서 사용되거나 또는 이로부터 단리된 진균 세포의 임의의 균주를 의미한다. 산업적 균주는 산업적 공정에서 전형적으로 사용되는 진균 종을 지칭하거나 또는 또한 비산업

적 목적 (예를 들어, 산업적 연구)을 위해 사용될 수 있는 진균 종의 단리물을 지칭할 수 있다. 산업 공정의 예는 발효 (예를 들어, 식품 또는 음료 제품의 생산에서), 증류, 바이오연료 생산, 화합물의 생산 및 폴리펩티드의 생산을 포함할 수 있다. 산업적 균주의 예는 JAY270 및 ATCC4124를 포함할 수 있지만, 이들로 제한되지 않는다.

[0707] 일부 실시형태에서, 진균 세포는 배수체 세포이다. 본 명세서에서 사용되는, "배수체" 세포는 게놈이 하나 초과의 카피수로 존재하는 임의의 세포를 지칭할 수 있다. 배수체 세포는 배수체 상태에서 자연적으로 발견되는 세포 유형을 지칭할 수 있거나, 또는 배수체 상태에 (예를 들어, 특정 조절, 변경, 불활성화, 활성화 또는 감수분열, 세포질분열 또는 DNA 복제의 변형을 통해) 존재하도록 유도된 세포를 지칭할 수 있다. 배수체 세포는 전체 게놈이 배수체인 세포를 지칭할 수 있거나, 또는 관심 특정 게놈 유전자좌에서 배수체인 세포를 지칭할 수 있다. 이론에 국한하지 않지만, 가이드 RNA의 존재비는 더 흔하게는 반수체 세포에서보다 배수체 세포의 게놈 조작에서 속도 제한적 성분이 될 수 있고, 따라서, 본 명세서에 기재된 AD-기능화된 CRISPR 시스템을 이용하는 방법은 소정의 진균 세포 유형을 이용할 수 있는 것으로 계획된다.

[0708] 일부 실시형태에서, 진균 세포는 이배체 세포이다. 본 명세서에서 사용되는, "이배체" 세포는 게놈이 2개의 카피수로 존재하는 임의의 세포를 지칭할 수 있다. 이배체 세포는 이배체 상태에서 자연적으로 발견되는 세포 유형을 지칭할 수 있거나, 또는 (예를 들어, 특정 조절, 변경, 불활성화, 활성화 또는 감수분열, 세포질분열 또는 DNA 복제의 변형을 통해) 이배체 상태로 존재하도록 유도된 세포를 지칭할 수 있다. 예를 들어, 사카로마이세스 세레비지아 균주 S228C는 반수체 또는 이배체 상태에서 유지될 수 있다. 이배체 세포는 전체 게놈이 이배체인 세포를 지칭할 수 있거나, 또는 관심 특정 게놈 유전자좌에서 이배체인 세포를 지칭할 수 있다.

일부 실시형태에서, 진균 세포는 반수체 세포이다. 본 명세서에서 사용되는, "반수체" 세포는 하나의 카피수로 게놈이 존재하는 임의의 세포를 지칭할 수 있다. 반수체 세포는 반수체 상태에서 자연적으로 발견되는 세포 유형을 지칭할 수 있거나, 또는 (예를 들어, 특정 조절, 변경, 불활성화, 활성화 또는 감수분열, 세포질분열 또는 DNA 복제의 변형을 통해) 반수체 상태로 존재하도록 유도된 세포를 지칭할 수 있다. 예를 들어, 사카로마이세스 세레비지아 균주 S228C는 반수체 또는 이배체 상태로 유지될 수 있다. 반수체 세포는 전체 게놈이 반수체인 세포를 지칭할 수 있거나, 또는 관심 특정 게놈 유전자좌에서 반수체인 세포를 지칭할 수 있다.

[0709] 본 명세서에서 사용되는, "효모 발현 벡터"는 RNA 및/또는 폴리펩티드를 코딩하는 하나 이상의 서열을 함유하는 핵산을 지칭하며, 핵산(들)의 발현을 제어하는 임의의 목적하는 구성요소뿐만 아니라 효모 세포 내부의 발현 벡터의 복제 및 유지를 가능하게 하는 임의의 구성요소를 추가로 함유할 수 있다. 다수의 적합한 효모 발현 벡터 및 이의 특징은 당업계에 공지되어 있으며; 예를 들어, 다양한 벡터 및 기법은 문헌 [Yeast Protocols, 2nd edition, Xiao, W., ed. (Humana Press, New York, 2007)] 및 [Buckholz, R.G. and Gleeson, M.A. (1991) Biotechnology (NY) 9(11): 1067-72]에 예시되어 있다. 효모 벡터는 중심립 (CEN) 서열, 자가 복제 서열 (ARS), 프로모터, 예컨대, 관심 서열 또는 유전자에 작동적으로 연결된 RNA 중합효소 III 프로모터, 종결자, 예컨대, RNA 중합효소 III 종결자, 복제 기원, 및 마커 유전자 (예를 들어, 영양요구성, 항생제, 또는 다른 선택 가능한 마커)를 함유할 수 있지만, 이들로 제한되지 않는다. 효모에서 사용하기 위한 발현 벡터의 예는 플라스미드, 효모 인공 염색체, 2μ 플라스미드, 효모 통합 플라스미드, 효모 복제성 플라스미드, 셔틀 벡터 및 에피솜 플라스미드를 포함할 수 있다.

[0710] **식물 및 식물 세포의 게놈으로 AD-기능화된 CRISPR 시스템 성분의 안정한 통합**

[0711] 특정 실시형태에서, AD-기능화된 CRISPR 시스템의 성분을 코딩하는 폴리뉴클레오티드가 식물 세포의 게놈 내로 안정한 통합을 위해 도입되는 것이 고려된다. 이들 실시형태에서, 형질전환 벡터 또는 발현 시스템의 디자인은 아데노신 데아미나제 및 Cas13의 가이드 RNA 및/또는 융합 단백질이 발현되는 시기, 위치 및 조건에 따라서 조절될 수 있다.

[0712] 특정 실시형태에서, 식물 세포의 게놈 DNA 내로 AD-기능화된 CRISPR 시스템의 성분을 안정하게 도입하는 것이 계획된다. 추가적으로 또는 대안적으로, 식물 세포기관, 예컨대, 제한없이, 색소체, 미토콘드리아 또는 엽록체의 DNA 내로 안정한 통합을 위해 AD-기능화된 CRISPR 시스템 성분을 도입하는 것이 고려된다.

[0713] 식물 세포 게놈 내로 안정한 통합을 위한 발현 시스템은 다음의 구성요소 중 하나 이상을 함유할 수 있다: 식물 세포에서 아데노신 데아미나제 및 Cas13의 RNA 및/또는 융합 단백질을 발현시키기 위해 사용될 수 있는 프로모터 구성요소; 발현을 증강시키기 위한 5' 비번역 영역; 소정의 세포, 예컨대, 단자엽 세포에서 발현을 추가로 증강시키는 인트론 구성요소; 아데노신 데아미나제의 가이드 RNA 및/또는 융합 단백질 및 Cas13 코딩 서열 및

다른 목적하는 요소를 삽입하기 위한 편리한 제한 부위를 제공하기 위한 다중 클로닝 부위; 및 발현된 전사체의 효율적인 종결을 제공하기 위한 3' 비번역 영역.

[0714] 발현 시스템의 구성요소는 원형, 예컨대, 플라스미드 또는 형질전환 벡터, 또는 비원형, 예컨대, 선형 이중 가닥 DNA 중 하나인 하나 이상의 발현 구성체일 수 있다.

[0715] 특정 실시형태에서, AD-기능화된 CRISPR 발현 시스템은 적어도: 식물에서 표적 서열에 혼성화하는 가이드 RNA (gRNA)를 코딩하는 뉴클레오티드 서열을 포함하고, 가이드 RNA는 가이드 서열 및 직접 반복 서열, 및 아데노신 데아미나제 및 Cas13의 융합 단백질을 코딩하는 뉴클레오티드 서열을 포함하며, 성분 (a) 또는 (b)는 동일하거나 또는 상이한 구성체 상에 위치되고, 이에 의해 상이한 뉴클레오티드 서열은 식물 세포에서 작동가능한 동일하거나 또는 상이한 조절 구성요소의 제어 하에 있을 수 있다.

[0716] AD-기능화된 CRISPR 시스템의 성분, 및, 적용 가능한 경우, 주형 서열을 함유하는 DNA 구성체(들)는 다양한 통상적인 기법에 의해 식물, 식물 부분 또는 식물 세포 게놈 내로 도입될 수 있다. 상기 과정은 일반적으로 적합한 숙주 세포 또는 숙주 조직을 선택하는 단계, 구성체(들)를 숙주 세포 또는 숙주 조직 내로 도입하는 단계, 및 이로부터 식물 세포 또는 식물을 재생시키는 단계를 포함한다.

[0717] 특정 실시형태에서, DNA 구성체는, 예컨대, 제한없이,전기천공법, 미량주사법, 식물 세포 원형질체의 에어로졸 빔 주사와 같은 기법을 이용하여 식물 세포 내로 도입될 수 있거나, 또는 DNA 구성체는 바이오리스틱 방법, 예컨대, DNA 유전자 총을 이용하여 식물 조직에 직접적으로 도입될 수 있다 (또한 문헌 [Fu et al., Transgenic Res. 2000 Feb;9(1):11-9] 참조). 유전자 총의 기준은 입자에 의한 원형질의 침투 및 전형적으로 게놈 내로의 안정한 통합을 초래하는, 세포에 대해 관심 유전자로 코팅된 입자의 가속도이다. (예를 들어, 문헌 [Klein et al, Nature (1987), Klein et ah, Bio/Technology (1992), Casas et ah, Proc. Natl. Acad. Sci. USA (1993).] 참조).

[0718] 특정 실시형태에서, AD-기능화된 CRISPR 시스템의 성분을 함유하는 DNA 구성체는 아그로박테리움-매개 형질전환에 의해 식물 내로 도입될 수 있다. DNA 구성체는 적합한 T-DNA 측접 영역과 합쳐지고, 통상적인 아그로박테리움 투메파시엔스 숙주 벡터 내로 도입될 수 있다. 외래 DNA는 식물을 감염시킴으로써 또는 하나 이상의 Ti (종양-유도성) 플라스미드를 함유하는 식물 원형질체를 아그로박테리움 박테리아와 함께 인큐베이션시킴으로써 게놈 내로 혼입될 수 있다. (예를 들어, 문헌 [Fraley et al., (1985), Rogers et al., (1987)] 및 미국 특허 제5,563,055호 참조).

[0719] **식물 프로모터**

[0720] 식물 세포에서 적절한 발현을 보장하기 위해, 본 명세서에 기재된 AD-기능화된 CRISPR 시스템의 성분은 전형적으로 식물 프로모터, 즉, 식물 세포에서 작동가능한 프로모터의 제어 하에 위치된다. 상이한 유형의 프로모터의 사용이 고려된다.

[0721] 항상성 식물 프로모터는 식물의 모든 또는 거의 모든 발생 단계 동안 모든 또는 거의 모든 식물 조직에서 제어되는 오픈 리딩 프레임 (ORF)을 발현 ("항상성 발현"이라함)시킬 수 있는 프로모터이다. 항상성 프로모터의 한 가지 비제한적 예는 콜리플라워 모자이크 바이러스 35S 프로모터이다. "조절된 프로모터"는 항성성이 아니지만, 시간적으로 및/또는 공간적으로 조절된 방식으로 유전자 발현을 지시하는 프로모터를 지칭하고, 조직-특이적, 조직-선호성 및 유도성 프로모터를 포함한다. 상이한 프로모터는 상이한 조직 또는 세포 유형에서 또는 상이한 발생 단계 또는 상이한 환경 조건에 반응하여 유전자 발현을 지정할 수 있다. 특정 실시형태에서, AD-기능화된 CRISPR 성분 중 하나 이상은 항상성 프로모터의 제어 하에 발현되고, 예컨대, 콜리플라워 모자이크 바이러스 35S 프로모터 문제-선호성 프로모터는 특정 식물 조직, 예를 들어, 잎 또는 뿌리에서의 맥관 세포 내에서 또는 종자의 특정 세포에서 소정의 세포 유형에서 증강된 발현을 표적화하는 데 이용될 수 있다. AD-기능화된 CRISPR 시스템에서 사용하기 위한 특정 프로모터의 예는 문헌 [Kawamata et al., (1997) Plant Cell Physiol 38:792-803]; [Yamamoto et al., (1997) Plant J 12:255-65]; [Hire et al, (1992) Plant Mol Biol 20:207-18], [Kuster et al, (1995) Plant Mol Biol 29:759-72], 및 [Capana et al., (1994) Plant Mol Biol 25:681 -91]에서 확인된다.

[0722] 유도성 프로모터는 데아미나제의 비특이적 활성을 피하기 위해 제한된 환경 하에서 AD-기능화된 CRISPR 시스템의 성분 중 하나 이상을 발현시키기 위해 흥미로울 수 있다. 특정 실시형태에서, AD-기능화된 CRISPR 시스템의 하나 이상의 요소는 유도성 프로모터의 제어 하에 발현된다. 유도성이며, 유전자 편집 또는 유전자 발현의 시공적 제어를 가능하게 하는 프로모터의 예는 에너지 형태를 사용할 수 있다. 에너지 형태는 소리 에

너지, 전자기 복사, 화학적 에너지 및/또는 열 에너지를 포함할 수 있지만, 이들로 제한되지 않는다. 유도성 시스템의 예는 테트라사이클린 유도성 프로모터 (Tet-온 또는 Tet-오프), 소형 분자 2-하이브리드 전사 활성화 시스템 (FKBP, ABA 등), 또는 광 유도성 시스템 (피토크롬, LOV 도메인 또는 크립토크롬), 예컨대, 서열-특이적 방식으로 전사 활성의 변화를 지시하는 광 유도성 전사 이펙터 (LITE)를 포함한다. 광 유도성 시스템 성분은 아데노신 데아미나제 및 Cas13, 광-반응성 시토크롬 이형이량체 (예를 들어, 아라비돕시스 탈리아나 (*Arabidopsis thaliana*)로부터 유래)의 융합 단백질을 포함할 수 있다. 유도성 DNA 결합 단백질 및 이들의 사용 방법의 추가적인 예는 본 명세서에 전문이 참고로 편입된 미국 특허 제61/736465호 및 미국 특허 제61/721,283호에서 제공된다.

[0723] 특정 실시형태에서, 일시적 또는 유도성 발현은, 예를 들어, 화학적-조절 프로모터를 이용함으로써 달성될 수 있고, 즉, 이에 의해 외생성 화학물질의 적용은 유전자 발현을 유도한다. 유전자 발현의 조절은 또한 화학물질-억제 프로모터에 의해 얻어질 수 있으며, 화학물질의 적용은 유전자 발현을 억제한다. 화학물질-유도성 프로모터는 벤젠 설폰아미드 제초제 약해 경감제에 의해 활성화되는 메이즈 ln2-2 프로모터 (De Veylder et al., (1997) Plant Cell Physiol 38:568-77), 발아전 제초제로서 사용되는 소수성 친전자성 화합물에 의해 활성화되는 메이즈 GST 프로모터 (GST-11-27, WO93/01294), 및 살리실산에 의해 활성화되는 담배 PR-1 프로모터 (Ono et al., (2004) Biosci Biotechnol Biochem 68:803-7)를 포함하지만, 이들로 제한되지 않는다. 항생제에 의해 조절되는 프로모터, 예컨대, 테트라사이클린-유도성 및 테트라사이클린-억제성 프로모터 (문헌 [Gatz et al., (1991) Mol Gen Genet 227:229-37]; 미국 특허 제5,814,618호 및 제5,789,156호 참조)가 또한 본 명세서에서 사용될 수 있다.

[0724] **특별한 식물 유기체에서 전좌 및/또는 발현**

[0725] 발현 시스템은 특별한 식물 기관으로의 전좌 및/또는 그에서의 발현을 위한 구성요소를 포함할 수 있다.

[0726] ***엽록체 표적화***

[0727] 특정 실시형태에서, AD-기능화된 CRISPR 시스템은 엽록체 유전자를 특이적으로 변형시키기 위해 또는 엽록체에서의 발현을 보장하기 위해 사용되는 것이 고려된다. 이런 목적을 위해 엽록체 형질전환 방법 또는 엽록체에 대한 AD-기능화된 CRISPR 성분의 구획화가 사용된다. 예를 들어, 색소체 게놈에서 유전자 변형의 도입은 생물학적 안전성 문제, 예컨대, 화분을 통한 유전자 유동을 감소시킬 수 있다.

[0728] 엽록체 형질전환 방법은 당업계에 공지되어 있으며, 유전자 총, PEG 처리 및 미량주사법을 포함한다. 추가적으로, 핵 게놈으로부터 색소체로의 형질전환 카세트 전위를 수반하는 방법은 WO2010061186에 기재된 바와 같이 사용될 수 있다.

[0729] 대안적으로, AD-기능화된 CRISPR 성분 중 하나 이상의 식물 엽록체에 대한 표적화가 고려된다. 이는 아데노신 데아미나제 및 Cas13의 융합 단백질을 코딩하는 서열의 5' 영역에 작동적으로 연결된 엽록체 수송 펩티드 (CTP) 또는 색소체 수송 펩티드를 코딩하는 서열을 발현 구성체에 혼입함으로써 달성된다. CTP는 엽록체 내로의 전좌 동안 가공 단계에서 제거된다. 발현된 단백질의 엽록체 표적화는 당업자에게 충분히 공지되어 있다 (예를 들어, [Protein Transport into Chloroplasts, 2010, Annual Review of Plant Biology,Vol. 61: 157-180] 참조). 이러한 실시형태에서, 또한 식물 엽록체에 대해 가이드 RNA를 표적화하는 것이 바람직하다. 엽록체 국재화 서열에 의해 엽록체에 가이드 RNA를 전좌시키기 위해 사용될 수 있는 방법 및 구성체는, 예를 들어, 본 명세서에 참고로 편입된 미국 특허 제20040142476호에 기재되어 있다. 구성체의 이러한 변형은 본 발명의 발현 시스템 내로 혼입되어 AD-기능화된 CRISPR 시스템 성분을 효율적으로 전좌시킬 수 있다.

[0730] **조류 세포로 AD-기능화된 CRISPR 시스템을 코딩하는 폴리뉴클레오티드의 도입.**

[0731] 유전자이식 조류 (또는 다른 식물, 예컨대, 유채)는 식물성 오일 또는 바이오연료, 예컨대, 알콜 (특히 메탄올 및 에탄올) 또는 다른 제품의 생산에서 특히 유용할 수 있다. 이들은 오일 또는 바이오연료 산업에서 사용하기 위한 고수준의 오일 또는 알콜을 발현 또는 과발현시키도록 조작될 수 있다.

[0732] 미국 특허 제8945839호는 Cas9를 이용하여 미세 조류 (클라미도모나스 레인하르티 (*Chlamydomonas reinhardtii*) 세포) 종을 조작하는 방법을 기술한다. 유사한 도구를 이용하여, 본 명세서에 기재된 AD-기능화된 CRISPR 시스템의 방법은 클라미도모나스 종 및 다른 조류에 적용될 수 있다. 특정 실시형태에서, CRISPR-Cas 단백질 (예를 들어, Cas13), 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 압타머-결합 어댑터 단백질에 융합될 수 있음) 및 가이드 RNA는, 항상성 프로모터, 예컨대, Hsp70A-Rbc S2 또는 Beta2-튜불린의 제어

하에 아데노신 데아미나제 및 Cas13의 아데노신 데아미나제의 융합 단백질을 발현시키는 벡터를 이용하여 발현된 조류에 도입된다. 가이드 RNA는 임의로는 T7 프로모터를 함유하는 벡터를 이용하여 전달된다. 대안적으로, Cas13 mRNA 및 시험관내 전사된 가이드 RNA는 조류 세포에 전달될 수 있다. 전기천공법 프로토콜, 예컨대, GeneArt Chlamydomonas Engineering kit로부터의 표준 권장 프로토콜은 당업자가 이용 가능하다.

[0733] **효모 세포로 AD-기능화된 CRISPR 시스템 성분의 도입**

[0734] 특정 실시형태에서, 본 발명은 효모 세포의 게놈 편집을 위한 AD-기능화된 CRISPR 시스템의 용도에 관한 것이다. AD-기능화된 CRISPR 시스템 성분을 코딩하는 폴리뉴클레오티드를 도입하기 위해 사용될 수 있는 효모 세포를 형질도입하는 방법은 문헌 [Kawai et al., 2010, Bioeng Bugs. 2010 Nov-Dec; 1(6): 395-403]에 기재되어 있다. 비제한적 예는 리튬 아세테이트 처리 (담체 DNA 및 PEG 처리를 추가로 포함할 수 있음), 유전자 총 또는 전기천공법에 의한 효모 세포의 형질전환을 포함한다.

[0735] **식물 및 식물 세포에서 AD-기능화된 CRISPR 시스템 성분의 일시적 발현**

[0736] 특정 실시형태에서, 가이드 RNA 및/또는 CRISPR-Cas 유전자가 식물 세포에서 일시적으로 발현되는 것이 고려된다. 이들 실시형태에서, AD-기능화된 CRISPR 시스템은 가이드 RNA, CRISPR-Cas 단백질 (예를 들어, Cas13) 및 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 압타머-결합 어댑터 단백질에 융합될 수 있음)가 세포에 존재하는 경우에만 표적 유전자의 변형을 보장할 수 있어서, 게놈 변형을 더욱 제어할 수 있다. CRISPR-Cas 단백질의 발현은 일시적이기 때문에, 이러한 식물 세포로부터 재생된 식물은 전형적으로 외래 DNA를 함유하지 않는다. 특정 실시형태에서, CRISPR-Cas 단백질은 식물 세포에 의해 안정하게 발현되며, 가이드 서열은 일시적으로 발현된다.

[0737] 특정 실시형태에서, AD-기능화된 CRISPR 시스템 성분은 식물 바이러스 벡터를 이용하여 식물 세포에 도입될 수 있다 (Scholthof et al. 1996, Annu Rev Phytopathol. 1996;34:299-323). 추가 특정 실시형태에서, 상기 바이러스 벡터는 DNA 바이러스로부터의 벡터이다. 예를 들어, 제미니바이러스 (예를 들어, 캐비지 잎 말림 바이러스, 콩 누른오갈병 바이러스, 밀 오갈병 바이러스, 토마토 잎 말림 바이러스, 메이즈 도말 바이러스, 담배 잎 말림 바이러스, 또는 토마토 골든 모자이크 바이러스) 또는 나노바이러스 (예를 들어, 잠두 괴사성 황화 바이러스). 다른 특정 실시형태에서, 상기 바이러스 벡터는 RNA 바이러스 유래 벡터이다. 예를 들어, 토브라바이러스 (예를 들어, 담배 얼룩 바이러스 (tobacco rattle virus), 담배 모자이크 바이러스), 포텍스바이러스 (예를 들어, 감자 바이러스 X) 또는 호르데이바이러스 (예를 들어, 보리 줄무늬병 모자이크 바이러스). 식물 바이러스의 복제 게놈은 비통합형 벡터이다.

[0738] 특정 실시형태에서, AD-기능화된 CRISPR 시스템의 일시적 발현을 위해 사용되는 벡터는, 예를 들어, 원형질체에서 아그로박테리움-매개 일시적 발현에 대해 맞춤된 pEAQ 벡터이다 (Sainsbury F. et al., Plant Biotechnol J. 2009 Sep;7(7):682-93). CRISPR 효소를 발현시키는 안정한 유전자이식 식물에서 가이드 RNA를 발현시키기 위해 변형된 양배추 잎 말림 바이러스(CaLCuV) 벡터를 이용하는 게놈 위치의 정확한 표적화가 입증되었다 (Scientific Reports 5, Article number: 14926 (2015), doi:10.1038/srep14926).

[0739] 특정 실시형태에서, 가이드 RNA 및/또는 CRISPR-Cas 유전자를 코딩하는 이중 가닥 DNA 단편은 식물 세포 내로 일시적으로 도입될 수 있다. 이러한 실시형태에서, 도입된 이중-가닥 DNA 단편은 세포를 변형시키기 위한 충분량으로 제공되지만, 통과된 시간의 고려되는 기간 후에 또는 1회 이상의 세포 분열 후에 지속되지 않는다.
식물에서 직접적 DNA 전달을 위한 방법은 당업자에게 공지되어 있다 (예를 들어, 문헌 [Davey et al. Plant Mol Biol. 1989 Sep;13(3):273-85.] 참조).

[0740] 다른 실시형태에서, CRISPR-Cas 단백질 (예를 들어, Cas13) 및/또는 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 압타머-결합 어댑터 단백질에 융합될 수 있음)을 코딩하는 RNA 폴리뉴클레오티드는 식물 세포 내로 도입되며, 그 이후에 세포를 변형시키기에 충분한 양으로 (적어도 하나의 가이드 RNA의 존재 하에) 단백질을 생성하는 숙주 세포에 의해 번역 및 처리되지만 고려되는 기간이 경과된 후 또는 1회 이상의 세포 분열 후에 지속되지 않는다. 일시적 발현을 위해 식물 원형질체에 mRNA를 도입하는 방법은 당업자에게 공지되어 있다 (예를 들어, 문헌 [Gallie, Plant Cell Reports (1993), 13:119-122] 참조).

[0741] 상기 기재된 상이한 방법의 조합이 또한 계획된다.

[0742] **식물 세포로 AD-기능화된 CRISPR 시스템 성분의 전달**

[0743] 특정 실시형태에서, AD-기능화된 CRISPR 시스템의 하나 이상의 성분을 직접적으로 식물 세포에 전달하는 데 관

심이 있다. 이는 특히, 비-유전자이식 식물의 생성에서 흥미롭다 (이하 참조). 특정 실시형태에서, AD-기능화된 CRISPR 시스템 성분 중 하나 이상은 식물 또는 식물 세포 외부에서 제조되어 세포에 전달된다. 예를 들어 특정 실시형태에서, CRISPR-Cas 단백질은 식물 세포에 대한 도입 전에 시험관내에서 제조된다. CRISPR-Cas 단백질은 당업자에게 공지된 다양한 방법에 의해 제조될 수 있으며, 재조합 생성을 포함한다. 발현 후에, CRISPR-Cas 단백질은 단리되고, 필요하다면, 재폴딩되며, 정제되고, 선택적으로 임의의 정제 태그, 예컨대, His-태그를 제거하도록 처리된다. 조질의, 부분 정제되거나 또는 더 완전하게 정제된 CRISPR-Cas 단백질이 얻어지면, 단백질은 식물 세포에 도입될 수 있다.

[0744] 특정 실시형태에서, CRISPR-Cas 단백질은 관심 대상 유전자를 표적화하는 가이드 RNA와 혼합되어 사전 조립된 리보핵단백질을 형성한다.

[0745] 개별 성분 또는 사전 조립된 리보핵단백질은 전기천공법을 통해, CRISPR-Cas-연관 유전자 산물 코팅 입자를 이용하는 유전자 총에 의해, 화학적 형질감염에 의해 또는 세포막을 가로지르는 일부 다른 수송 수단에 의해 식물 세포에 도입될 수 있다. 예를 들어, 사전 조립된 CRISPR 리보핵단백질에 의한 식물 원형질체의 형질감염은 식물 게놈의 표적화된 변형을 보장하는 것으로 입증되었다 (문헌 [Woo et al. *Nature Biotechnology*, 2015; DOI: 10.1038/nbt.3389]에 의해 기재된 바와 같음).

[0746] 특정 실시형태에서, AD-기능화된 CRISPR 시스템 성분은 나노입자를 이용하여 식물 세포 내로 도입된다. 단백질 또는 핵산으로서 또는 이들의 조합물로 성분들은 나노입자 상에 업로드되거나 또는 패키징될 수 있고, 식물에 적용될 수 있다 (예컨대, 예를 들어, WO 2008042156 및 미국 공개 특허 출원 제20130185823호에 기재). 특히, 본 발명의 실시형태는 WO2015089419에 기재된 바와 같은 CRISPR-Cas 단백질 (예를 들어, Cas13)을 코딩하는 DNA 분자(들), 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 압타머-결합 어댑터 단백질에 융합될 수 있음)를 코딩하는 DNA 분자(들), 및 가이드 RNA를 코딩하는 DNA 분자 및/또는 단리된 가이드 RNA가 업로드되거나 또는 팩킹된 나노입자를 포함한다.

[0747] 식물 세포에 AD-기능화된 CRISPR 시스템의 하나 이상의 성분을 도입하는 추가적인 수단은 세포 침투성 펩티드 (CPP)를 이용하는 것에 의한다. 따라서, 특히, 본 발명의 실시형태는 CRISPR-Cas 단백질에 연결된 세포 침투성을 포함하는 조성물을 포함한다. 본 발명의 특정 실시형태에서, CRISPR-Cas 단백질 및/또는 가이드 RNA는 하나 이상의 CPP에 결합되어 그들을 식물 원형질체 내부로 효과적으로 수송한다. Ramakrishna (인간 세포의 Cas9 경우 [Genome Res. 2014 Jun;24(6):1020-7])를 참조한다. 다른 실시형태에서, CRISPR-Cas 유전자 및/또는 가이드 RNA는 식물 원형질체 전달을 위해 하나 이상의 CPP에 결합된 하나 이상의 원형 또는 비원형 DNA 분자(들)에 의해 코딩된다. 이어서, 식물 원형질체는 식물 세포로 및 더 나아가 식물로 재생된다. CPP는 일반적으로 수용체 독립적 방식으로 세포막에 걸쳐 생체분자를 수송할 수 있는 단백질로부터 또는 키메라 서열로부터 유래된 35개보다 적은 아미노산의 짧은 펩티드로서 기재된다. CPP는 양이온성 펩티드, 소수성 서열을 갖는 펩티드, 양극성 펩티드, 프롤린-풍부 및 항-미생물 서열을 갖는 펩티드, 및 키메라 또는 이분 펩티드일 수 있다 (Pooga and Langel 2005). CPP는 생물학적 막에 침투할 수 있고, 이렇게 해서, 세포막을 가로질러서 세포질 내로 다양한 생체분자의 이동을 촉발시키고, 그들의 세포내 경로를 개선시켜서, 생체분자와 표적과의 상호작용을 용이하게 한다. CPP의 예는 특히, Tat, HIV 1형에 의한 바이러스 복제에 필요한 핵 전사 활성체 단백질, 페네트라틴, 카포시 섬유아세포 성장 인자 (FGF) 신호 펩티드 서열, 인테그린 β3 신호 펩티드 서열; 폴리아르기닌 펩티드 Arg 서열, 구아닌 풍부-분자 수송체, 스위트 애로우 펩티드를 포함한다.

[0748] **유전자 변형된 비유전자이식 식물의 제조를 위한 AD-기능화된 CRISPR 시스템의 용도**

[0749] 특정 실시형태에서, 본 명세서에 기재된 방법은 식물 게놈에서 외래 DNA의 존재를 피하기 위해 CRISPR 성분을 코딩하는 것을 포함하는, 임의의 외래 유전자의 식물 게놈 내로의 영구한 도입 없이 내생성 유전자를 변형시키거나 또는 그들의 발현을 변형시키기 위해 사용된다. 이것은 비유전자이식 식물에 대한 조절 요건이 덜 엄격하기 때문에 흥미로울 수 있다.

[0750] 특정 실시형태에서, 이는 AD-기능화된 CRISPR 시스템 성분의 일시적 발현에 의해 보장된다. 특정 실시형태에서, 성분 중 하나 이상은 본 명세서에 기재된 방법에 따라 관심 유전자의 변형을 지속적으로 꾸준하게 보장하기에 충분한 CRISPR-Cas 단백질, 아데노신 데아미나제, 및 가이드 RNA를 생성하는 하나 이상의 바이러스 벡터 상에서 발현된다.

[0751] 특정 실시형태에서, AD-기능화된 CRISPR 시스템 구성체의 일시적 발현은 식물 원형질체에서 보장되며, 따라서 게놈 내로 통합되지 않는다. 제한된 발현 창은 AD-기능화된 CRISPR 시스템이 본 명세서에 기재된 바와 같은

표적 유전자의 변형을 보장하도록 하기에 충분할 수 있다.

[0752] 특정 실시형태에서, AD-기능화된 CRISPR 시스템의 상이한 성분은 별개로 또는 혼합물로, 본 명세서에서 상기 기재한 바와 같은 나노입자 또는 CPP 분자와 같은 미립자 전달 분자의 도움에 의해, 식물 세포, 원형질체 또는 식물 조직에 도입된다.

[0753] AD-기능화된 CRISPR 시스템 성분의 발현은 아데노신 데아미나제의 데아미나제 활성에 의해 게놈의 표적화된 변형을 유도할 수 있다. 본 명세서에서 상기 기재된 상이한 전략은 식물 게놈 내로 AD-기능화된 CRISPR 시스템 성분의 도입을 필요로 하는 일 없이 CRISPR-매개 표적화된 게놈 편집을 허용한다. 식물 세포 내로 일시적으로 도입된 성분은 전형적으로 교배 시 제거된다.

[0754] **식물 배양 및 재생**

[0755] 특정 실시형태에서, 변형된 게놈을 갖고 본 명세서에 기재된 임의의 방법에 의해 생성되거나 또는 얻어지는 식물 세포는 형질전환되거나 또는 변형된 유전자형을 갖는 전체 식물 및 그에 따라 목적하는 표현형을 재생하기 위해 배양될 수 있다. 통상적인 재생 기법은 당업자에게 잘 공지되어 있다. 이러한 재생 기법의 특정 예는 조직 배양 성장 배지에서 소정의 식물 호르몬의 조작에 의존하며, 전형적으로 목적하는 뉴클레오티드 서열과 함께 도입된 항생제 및/또는 제초제 마커에 의존한다. 추가 특정 실시형태에서, 식물 재생은 배양된 원형질체, 식물 캘러스, 외식편, 기관, 화분, 배아 또는 이의 부분으로부터 얻어진다 (예를 들어, 문헌 [Evans et al. (1983), Handbook of Plant Cell Culture, Klee et al (1987) Ann. Rev. of Plant Phys.] 참조).

[0756] 특정 실시형태에서, 본 명세서에 기재된 형질전환되거나 또는 개선된 식물은 본 발명의 동형접합적 개선 식물 (DNA 변형에 대해 동형접합적)에 대한 종자를 제공하도록 자가수분되거나 또는 이형접합적 식물에 대한 종자를 제공하도록 비유전자이식 식물 또는 상이한 개선 식물과 교배될 수 있다. 재조합 DNA가 식물 세포 내로 도입되는 경우에, 이러한 교배로 얻어진 식물은 재조합 DNA 분자에 이형접합적인 식물이다. 개선된 식물로부터의 교배 그리고 유전자 변형의 포함 (재조합 DNA일 수 있음)에 의해 얻어지는 이러한 동형접합적 식물과 이형접합적 식물은 둘 모두 본 명세서에서 "자손"으로서 지칭된다. 자손 식물은 본래의 유전자이식 식물로부터의 자손이 되고 본 명세서에 제공된 방법에 의해 도입되는 게놈 변형 또는 재조합 DNA 분자를 함유하는 식물이다. 대안적으로, 유전자 변형 식물은 AD-기능화된 CRISPR 시스템을 이용하여 상기 기재한 방법 중 하나에 의해 얻어질 수 있고, 이에 의해 외래 DNA는 게놈 내로 혼입되지 않는다. 추가적인 육종에 의해 얻어지는 이러한 식물의 자손은 또한 유전자 변형을 함유할 수 있다. 육종은 상이한 작물에 대해 통상적으로 사용되는 임의의 육종 방법에 의해 수행된다 (예를 들어, 문헌 [Allard, Principles of Plant Breeding, John Wiley & Sons, NY, U. of CA, Davis, CA, 50-98 (1960)] 참조).

[0757] **증강된 작물학적 형질을 갖는 식물의 생성**

[0758] 본 명세서에 제공된 AD-기능화된 CRISPR 시스템은 표적화된 A-G 및 T-C 돌연변이를 도입하기 위해 사용될 수 있다. 단일 세포에서 다중 변형을 달성하도록 지시된 다중 표적화 RNA의 공동발현에 의해, 다중복합 게놈 변형이 보장될 수 있다. 이 기술은 증강된 영양 품질, 질환에 대해 증가된 내성 및 항생제 및 비항생제 스트레스에 대한 내성, 및 상업적으로 가치있는 식물 제품 또는 이종성 화합물의 증가된 생산을 비롯한 개선된 형질을 갖는 식물의 높은 정확도 조작에 사용될 수 있다.

[0759] 특정 실시형태에서, 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템은 표적화된 A-G 및 T-C 돌연변이를 도입하기 위해 사용된다. 이러한 돌연변이는 넌센스 돌연변이 (예를 들어, 미성숙 정지 코돈) 또는 미스센스 돌연변이 (예를 들어, 상이한 아미노산 잔기를 코딩)일 수 있다. 이는 소정의 내생성 유전자에서 A-G 및 T-C 돌연변이가 목적하는 형질을 부여하거나 또는 기여할 수 있는 경우에 관심이 있다.

[0760] 본 명세서에 기재된 방법은 일반적으로 그들이 야생형 식물에 비교되는 하나 이상의 바람직한 형질을 갖는다는 점에서 "개선된 식물"의 생성을 초래한다. 특정 실시형태에서, 얻어진 식물, 식물 세포 또는 식물 부분은 식물 세포의 모두 또는 일부에 혼입된 외생성 DNA 서열을 포함하는 유전자이식 식물이다. 특정 실시형태에서, 비-유전자이식 유전자 변형 식물, 식물 일부 또는 세포가 얻어지며, 즉, 외생성 DNA 서열은 식물의 식물 세포 중 어느 것의 게놈에 혼입된다. 이러한 실시형태에서, 개선된 식물은 비-유전자이식이다. 내생성 유전자의 변형만이 보장되고, 외래 유전자는 식물 게놈에서 도입 또는 유지되지 않는 경우에, 얻어진 유전자 변형된 작물은 외래 유전자를 함유하지 않으며, 따라서 기본적으로 비유전자이식으로 간주될 수 있다.

[0761] 특정 실시형태에서, 폴리뉴클레오티드는 DNA 바이러스 (예를 들어, 제미니바이러스) 또는 RNA 바이러스 (예를 들어, 토브라바이러스)에 의해 세포 내로 전달된다. 특정 실시형태에서, 도입하는 단계는 식물 세포에

CRISPR-Cas 단백질, 아데노신 데아미나제, 및 가이드 RNA를 코딩하는 하나 이상의 폴리뉴클레오티드 서열을 함유하는 T-DNA를 전달하는 것을 포함하며, 전달하는 것은 아그로박테리움을 통한다. AD-기능화된 CRISPR 시스템의 성분을 코딩하는 폴리뉴클레오티드 서열은 프로모터, 예컨대, 항상성 프로모터 (예를 들어, 콜리플라워 모자이크 바이러스 35S 프로모터), 또는 세포 특이적 또는 유도성 프로모터에 작동가능하게 연결될 수 있다. 특정 실시형태에서, 폴리뉴클레오티드는 미세투사 유전자 총에 의해 도입된다. 특정 실시형태에서, 상기 방법은 관심 대상 유전자의 발현이 변형되었는지의 여부를 결정하기 위한 단계들을 도입한 후에 식물 세포를 스크리닝하는 단계를 추가로 포함한다. 특정 실시형태에서, 상기 방법은 식물 세포로부터 식물을 재생시키는 단계를 포함한다. 추가 실시형태에서, 상기 방법은 유전적으로 요망되는 식물 계통을 얻기 위해 식물을 교배육종시키는 단계를 포함한다.

[0762] 상기 기재된 방법의 특정 실시형태에서, 질환 감수성 유전자 또는 식물 방어 유전자의 조절자를 코딩하는 유전자 (예를 들어, Mlo 유전자)의 표적화된 돌연변이에 의해 질환 내성 작물이 얻어진다. 특정 실시형태에서, 아세토락테이트 신타제 (ALS) 및 프로토포르피리노겐 옥시다제 (PPO)를 코딩하는 것과 같은 식물 유전자에서 특정 뉴클레오티드의 표적화된 치환에 의해 제초제-내성 작물이 생성된다. 특정 실시형태에서, 항생제 스트레스 내성의 음성 조절자를 코딩하는 유전자의 표적화된 돌연변이에 의한 가뭄 및 염 내성 작물, Waxy 유전자의 표적화된 돌연변이에 의한 저 아밀로스 곡물, 벼 또는 호분층에서의 주요 리파제 유전자의 표적화된 돌연변이에 의해 산패가 감소된 다른 곡물 등. 특정 실시형태에서, 관심 형질을 코딩하는 내생성 유전자의 더 광범위한 목록은 이하에 열거된다.

[0763] **배수체 식물을 변형하기 위한 AD-기능화된 CRISPR 시스템의 용도**

[0764] 다수의 식물은 배수체인데, 이는 그들이 게놈 복제물의 2배 (때때로 밀에서와 같이 6배만큼)를 운반한다는 것을 의미한다. AD-기능화된 CRISPR 시스템을 이용하는, 본 발명에 따른 방법은 유전자의 모든 카피수로 영향을 미치기 위해, 또는 한 번에 12개의 유전자를 표적화하기 위해 "다중복합체화"될 수 있다. 예를 들어, 특정 실시형태에서, 본 발명의 방법은 질환에 대해 방어를 억제하는 것을 초래하는 상이한 유전자에서 기능 상실 돌연변이를 동시에 보장하기 위해 사용된다. 특정 실시형태에서, 본 발명의 방법은 밀 식물이 백분병균에 내성이라는 것을 보장하기 위해 밀 식물 세포에서 TaMLO-A1, TaMLO-B1 및 TaMLO-D1 핵산 서열의 발현을 동시에 억제하고 이로부터 밀 식물을 재생하는 데 사용된다 (또한 WO2015109752 참조).

[0765] **작물학적 형질을 부여하는 예시적인 유전자**

[0766] 특정한 실시형태에서, 본 발명은 하기에 열거된 바와 같은, 내생성 유전자 및 그들 조절 구성요소 내에 표적화된 A-G 및 T-C 돌연변이를 포함하는 방법을 포괄한다:

[0767] ***1. 해충 또는 질환에 내성을 부여하는 유전자:***

[0768] 식물 질환 내성 유전자. 식물은 특정 병원성 균주에 대해 내성이 있는 식물을 조작하기 위해 클로닝된 내성 유전자로 형질전환될 수 있다. 예를 들어, 문헌 [Jones et al., Science 266:789 (1994)](클라도스포륨 풀붐(*Cladosporium fulvum*)에 대한 내성을 위한 토마토 Cf-9 유전자의 클로닝); 문헌 [Martin et al., Science 262:1432 (1993)](슈도모나스 시링가에 (*Pseudomonas syringae*) pv에 대한 내성을 위한 토마토 Pto 유전자. 토마토는 단백질 키나제를 코딩함); 문헌 [Mindrinos et al., Cell 78:1089 (1994)](아라비돕스는 슈도모나스 시링가에 대한 내성을 위한 RSP2 유전자일 수 있음). 병원균 감염 동안 상향 또는 하향 조절되는 식물 유전자는 병원균 내성을 위해 조작될 수 있다. 예를 들어, 문헌 [Thomazella et al., bioRxiv 064824; doi: https://doi.org/10.1101/064824 Epub. July 23, 2016] 참조 (병원균 감염 동안 정상적으로 상향조절된 SlDMR6-1에서 결실을 갖는 토마토 식물).

[0769] 해충, 예컨대, 대두 시스트 선충에 대한 내성을 부여하는 유전자. 예를 들어, 국제 특허 출원 WO 96/30517; 국제 특허 출원 WO 93/19181을 참조한다.

[0770] 바실러스 투링기엔시스 (*Bacillus thuringiensis*) 단백질, 예를 들어, 문헌 [Geiser et al., Gene 48:109 (1986)]을 참조한다.

[0771] 렉틴, 예를 들어, 문헌 [Van Damme et al., Plant Molec. Biol. 24:25 (1994]을 참조한다.

[0772] 비타민-결합 단백질, 예컨대, 아비딘, 국제 특허 출원 US93/06487을 참조하고, 해충에 대한 살유충제로서 아비딘 및 아비딘 상동체의 용도를 교시한다.

[0773] 효소 억제제, 예컨대, 프로테아제 또는 프로테이나제 억제제 또는 아밀라제 억제제. 예를 들어, 문헌 [Abe

et al., J. Biol. Chem. 262:16793 (1987), Huub et al., Plant Molec. Biol. 21:985 (1993)), Sumitani et al., Biosci. Biotech. Biochem. 57:1243 (1993)] 및 미국 특허 제5,494,813호를 참조한다.

[0774] 곤충-특이적 호르몬 또는 페로몬, 예컨대, 에크디스테로이드 또는 유약 호르몬, 이의 변이체, 이에 기반한 모방체 또는 이들의 길항제 또는 효현제. 예를 들어, 문헌 [Hammock et al., Nature 344:458 (1990)]을 참조한다.

[0775] 발현 시 영향받는 해충의 생리를 붕괴시키는 곤충-특이적 펩티드 또는 신경펩티드. 예를 들어, 문헌 [Regan, J. Biol. Chem. 269:9 (1994)] 및]Pratt et al., Biochem. Biophys. Res. Comm. 163:1243 (1989)]. 또한 미국 특허 제5,266,317호를 참조한다.

[0776] 뱀, 말벌 또는 임의의 다른 유기체에 의해 자연에서 생성되는 곤충-특이적 독. 예를 들어, 문헌 [Pang et al., Gene 116: 165 (1992)]를 참조한다.

[0777] 모노터펜, 세스퀴터펜, 스테로이드, 하이드록삼산, 페닐프로파노이드 유도체 또는 살충 활성을 갖는 다른 비단백질 분자의 과축적을 초래하는 효소.

[0778] 생물학적 활성 분자; 예를 들어, 천연 또는 합성이건, 해당 효소, 단백질 분해 효소, 지방분해 효소, 뉴클레아제, 사이클라제, 트랜스아미나제, 에스터라제, 가수분해효소, 포스파타제, 키나제, 포스포릴라제, 중합효소, 엘라스타제, 키티나제 및 글루카나제의 번역후 변형을 포함하는, 변형에 관여되는 효소. 국제 특허 출원 WO93/02197, 문헌 [Kramer et al., Insect Biochem. Molec. Biol. 23:691(1993)] 및 [Kawalleck et al., Plant Molec. Biol. 21 :673 (1993)]을 참조한다.

[0779] 신호 전달을 자극하는 분자. 예를 들어, 문헌 [Botella et al., Plant Molec. Biol. 24:757 (1994)], 및 [Griess et al., Plant Physiol. 104:1467 (1994)]을 참조한다.

[0780] 바이러스-침습성 단백질 또는 이로부터 유래된 복합 독소. 문헌 [Beachy et al., Ann. rev. Phytopathol. 28:451(1990)]을 참조한다.

[0781] 병원균 또는 기생충에 의해 자연에서 생성된 발생-저지 단백질. 문헌 [Lamb et al., Bio/Technology 10:1436 (1992)] 및 [Toubart et al., Plant J. 2:367 (1992)]을 참조한다.

[0782] 식물에 의해 자연계에서 생성되는 발생-저지 단백질. 예를 들어, 문헌 [Logemann et al., Bio/Technology 10:305 (1992)]을 참조한다.

[0783] 식물에서, 병원균은 종종 숙주-특이적이다. 예를 들어, 일부 푸사리움 종은 토마토 시들음을 야기할 것이지만, 토마토만을 공격하고, 다른 푸사리움 종은 밀만을 공격한다. 식물은 대부분의 병원균에 저항하는 기존의 유도된 방어성을 갖는다. 식물 세대에 걸친 돌연변이 및 재조합 사건은, 특히 병원균이 식물보다 더욱 높은 빈도로 번식함에 따라, 감수성이 증가되는 유전적 변동성을 일으킨다. 식물에서는 비숙주 저항성이 있을 수 있으며, 예를 들어, 숙주 및 병원균은 비상용성이거나, 통상적으로 많은 유전자에 의해 제어되는, 모든 종류의 병원균에 대하여 부분 저항성 및/또는 다른 종류에 대해서는 아니지만 일부 종류의 병원균에 대해 완전한 저항성일 수 있다. 이러한 내성은 전형적으로 소수의 유전자에 의해 제어된다. AD-기능화된 CRISPR 시스템의 방법 및 성분을 이용하여, 본 명세서에서 예상되는 특정 돌연변이를 유도하기 위한 새로운 도구가 이제는 존재한다. 따라서, 내성 유전자의 공급원 게놈을 분석할 수 있고, 목적하는 특징 또는 형질을 갖는 식물에서, 내성 유전자 상승을 유도하기 위해 AD-기능화된 CRISPR 시스템의 방법 및 성분을 사용할 수 있다. 본 시스템은 이전의 돌연변이유발제보다 더 정확성을 가질 수 있고, 따라서 식물 육종 프로그램을 가속화시키고 개선시킨다.

[0784] 2. 식물 질환에 관여하는 유전자, 예컨대, WO 2013046247에 열거된 것:

[0785] 벼 병해: 마그나포르테 그리세아 (*Magnaporthe grisea*), 코클리오볼루스 미야베아누스 (*Cochliobolus miyabeanus*), 리족토니아 솔라니 (*Rhizoctonia solani*) 및 지베렐라 푸지쿠로이 (*Gibberella fujikuroi*); 밀의 병해: 에리시페 그라미니스 (*Erysiphe graminis*), 푸사리움 그라미네아룸 (*Fusarium graminearum*), 푸사리움 아베나세움 (*F. avenaceum*), 푸사리움 쿨모룸 (*F. culmorum*), 마이크로도키움 니발레 (*Microdochium nivale*), 푸치니아 스트리이포르미스 (*Puccinia striiformis*), 푸치니아 그라미니스 (*P. graminis*), 푸치니아 레콘디타 (*P. recondita*), 마이크로넥트리엘라 니발레 (*Micronectriella nivale*), 티풀라 종 (*Typhula sp.*), 우스틸라고 트리티시 (*Ustilago tritici*), 틸레티아 카리에스 (*Tilletia caries*), 슈도세르코스포렐라 헤르포트리코이데스 (*Pseudocercosporella herpotrichoides*), 미코스파에렐라 그라미니콜라 (*Mycosphaerella graminicola*), 스타고

노스포라 노도룸 (*Stagonospora nodorum*), 피레노포라 트리티 시레펜티스 (*Pyrenophora tritici-repentis*); 보리의 병해: 에리시페 그라미니스 (*Erysiphe graminis*), 푸사리움 그라미네아룸 (*Fusarium graminearum*), 푸사리움 아베나세움 (*F. avenaceum*), 푸사리움 쿨모룸 (*F. culmorum*), 마이크로도키움 니발레 (*Microdochium nivale*), 푸치니아 스트리이포르미스 (*Puccinia striiformis*), 피. 그라미니스 (P. graminis), 피. 호르데이 (*P. hordei*), 우스틸라고 누다 (*Ustilago nuda*), 린코스포륨 세칼리스(*Rhynchosporium secalis*), 피레노포라 테레스(*Pyrenophora teres*), 코킬로볼루스 사비투스 (*Cochliobolus sativus*), 피레노포라 그라라미네아 (*Pyrenophora graminea*), 및 리족토니아 솔라니 (*Rhizoctonia solani*); 메이즈 병해: 우스틸라고 메이디스 (*Ustilago maydis*), 코킬로볼루스 헤테로스트로푸스 (*Cochliobolus heterostrophus*), 글로에오세르코스포라 소르기 (*Gloeocercospora sorghi*), 푸치니아 폴리소라 (*Puccinia polysora*), 세르코스포라 자에-메이디스 (*Cercospora zeae-maydis*), 리족토니아 솔라니 (*Rhizoctonia solani*);

[0786] 감귤 병해: 디아포르테 시트리 (*Diaporthe citri*), 엘시노에 파우세티 (*Elsinoe fawcetti*), 페니실리움 디지타툼 (*Penicillium digitatum*), 피. 이탈리쿰 (*P. italicum*); 피토프토라 파라시티카 (Phytophthora parasitica), 피토프토라 시트로프토라 (*Phytophthora citrophthora*); 사과 병해: 모닐리니아 말리 (*Monilinia mali*), 발사 세라토스페르마 (*Valsa ceratosperma*), 포도스파에라 류코트리차 (*Podosphaera leucotricha*), 알테르나리아 알테르나타 사과 병원형 (*Alternaria alternata apple* pathotype), 벤투리아 이나에쿠 알리스 (*Venturia inaequalis*), 콜레토트리쿰 아쿠타툼 (*Colletotrichum acutatum*) 및 피토프토라 캑토룸 (*Phytophtora cactorum*);

[0787] 배 병해: 벤투리아 나시콜라 (*Venturia nashicola*), 벤투리아 피리나 (*V. pirina*), 알테르나리아 알테르나타 일본 배 병원형 (*Alternaria alternata Japanese pear* pathotype), 짐노스포란지움 하라에아눔 (*Gymnosporangium haraeanum*) 및 피토프토라 캑토룸 (Phytophtora cactorum);

[0788] 복숭아 병해: 모닐리니아 프룩티콜라 (*Monilinia fructicola*), 클라도스포리움 카르포 필룸 (*Cladosporium carpophilum*) 및 포몹시스 종 (*Phomopsis* sp.);

[0789] 포도 병해: 엘시노에 암펠리나 (*Elsinoe ampelina*), 글로메렐라 신구라타 (*Glomerella cingulata*), 운시눌라 네카토르 (*Uncinula necator*), 파콥소라 암펠롭시디스 (*Phakopsora ampelopsidis*), 구이그나르디아 비드웰리이 (*Guignardia bidwellii*), 및 플라스모파라 비티콜라 (*Plasmopara viticola*);

[0790] 감 병해: 글로에오스포리움 카키 (*Gloeosporium kaki*) 및 세르코스포라 카키 (*Cercospora kaki*), 미코스파에렐라 나웨 (*Mycosphaerella nawae*);

[0791] 박 병해: 콜레토트리쿰 라제나리움 (*Colletotrichum lagenarium*), 스파에로테카 풀리지네아 (*Sphaerotheca fuliginea*), 미코스파에렐라 멜로니스 (*Mycosphaerella melonis*), 푸사리움 옥시스포룸 (*Fusarium oxysporum*), 슈도페로노스포라 쿠벤시스 (*Pseudoperonospora cubensis*), 피토프토라 종 (*Phytophthora* sp.) 및 피티움 종 (*Pythium* sp.);

[0792] 토마토 병해: 알테르나리아 솔라니 (*Alternaria solani*), 클라도스포리움 풀붐 (*Cladosporium fulvum*) 및 피토프토라 인페스탄스 (*Phytophthora infestans*); 슈도모나스 시링가에 pv. 토마토; 피토프토라 카프시시 (*Phytophthora capsici*); 잔토모나스 (*Xanthomonas*)

[0793] 가지 병해: 포몹시스 벡산스 (*Phomopsis vexans*) 및 에리시페 시초라세아룸 (*Erysiphe cichoracearum*); 십자화과 채소류 병해: 알테르나리아 자포니카 (*Alternaria japonica*), 세르코스포렐라 브라시카에 (*Cercosporella brassicae*), 플라스모디오포라 브라시카에 (*Plasmodiophora brassicae*) 및 페로노스포라 파라시티카 (*Peronospora parasitica*);

[0794] 웰시 양파 병해: 푸치니아 알리이 (*Puccinia allii*) 및 페로노스포라 데스트럭터 (*Peronospora destructor*);

[0795] 대두 병해: 세르코스포라 키쿠치이 (*Cercospora kikuchii*), 엘시노에 글리시네스 (*Elsinoe 글리신s*), 디아포르테 파세오로룸 변이체 소자에 (*Diaporthe phaseolorum var. sojae*), 셉토리아 글리신스 (*Septoria 글리신s*), 세르코스포라 소지나 (*Cercospora sojina*), 파콥소라 파치리지 (*Phakopsora pachyrhizi*), 피토프토라 소자에 (*Phytophthora sojae*), 리족토니아 솔라니 (*Rhizoctonia solani*), 코르니스포라 카시콜라 (*Corynespora cassiicola*), 및 스클레로니티아 스클레오티오룸 (*Sclerotinia sclerotiorum*);

[0796] 강낭콩 병해: 콜레토트리쿰 린뎀티아눔 (*Colletotrichum lindemthianum*);

[0797] 땅콩 병해: 세르코스포라 페르소나타 (*Cercospora personata*), 세르코스포라 아라키디콜라 (*Cercospora*

arachidicola) 및 스클레로티움 롤프시 (*Sclerotium rolfsii*);

[0798] 완두 병해: 에리시페 피시 (*Erysiphe pisi*);

[0799] 감자 병해: 알테르나리아 솔라니 (*Alternaria solani*) 및 피토프토라 인페스탄스 (*Phytophthora infestans*), 피토프토라 에리스로셉티카 (), 및 스폰고스포라 서브터란네안 피토프토라 종 서브터라네안 f. 종 서브터라네안 (*Spongospora subterranean f. sp. Subterranean*);

[0800] 딸기 병해: 스파에로테카 후물리 (*Sphaerotheca humuli*) 및 글로메렐라 신구라타 (*Glomerella cingulata*);

[0801] 차 병해: 엑소바시디움 레티쿨라툼 (*Exobasidium reticulatum*), 엘시노에 류코스필라 (*Elsinoe leucospila*), 페스탈로티옵시스 종 (*Pestalotiopsis sp*) 및 콜레토트리쿰 테아에시넨시스 (*Colletotrichum theaesinensis*);

[0802] 담배 병해: 알테르나리아 론지페스 (*Alternaria longipes*), 에리시페 시초라세아룸 (*Erysiphe cichoracearum*), 콜레토트리쿰 타바쿰 (*Colletotrichum tabacum*), 페로노스포라 타바시나 (*Peronospora tabacina*) 및 피토프토라 니코티아나에 (*Phytophthora nicotianae*);

[0803] 유채 병해: 스클레로티니아 스클레로티오룸 (*Sclerotinia sclerotiorum*), 및 리족토니아 솔라니 (*Rhizoctonia solani*);

[0804] 면화 병해: 리족토니아 솔라니 (*Rhizoctonia solani*);

[0805] 사탕무 병해: 세르코스포라 베티콜라 (*Cercospora beticola*), 타나테포루스 쿠쿠메리스 (*Thanatephorus cucumeris*), 타나테포루스 쿠쿠메리스 (*Thanatephorus cucumeris*) 및 아파니데르마툼 코클리오이데스 (*Aphanidermatum cochlioides*);

[0806] 장미 병해: 디플로카르폰 로사에 (*Diplocarpon rosae*), 스파에로테카 판노사 (*Sphaerotheca pannosa*) 및 페로노스포라 스파르사 (*Peronospora sparsa*);

[0807] 국화 및 국화과 병해: 브레미아 락투카에 (*Bremia lactucae*), 셉토리아 크리산테미-인디시 (*Septoria chrysanthemi-indici*) 및 푸치니아 호리아나 (*Puccinia horiana*);

[0808] 다양한 식물의 병해: 피티움 아파니데르마툼 (*Pythium aphanidermatum*), 피티움 데바리아눔 (*Pythium debarianum*), 피티움 그라미니콜라 (*Pythium graminicola*), 피티움 이레굴라레 (*Pythium irregulare*), 및 피티움 울티뭄 (*Pythium ultimum*), 보트리티스 시네레아 (*Botrytis cinerea*), 및 스클레로티니아 스클레로티오룸 (*Sclerotinia sclerotiorum*);

[0809] 왜무 병해: 알테르나리아 브라시키콜라 (*Alternaria brassicicola*);

[0810] 잔디 병해: 스클레로티니아 호메오카르파 (*Sclerotinia homeocarpa*), 리족토니아 솔라니 (*Rhizoctonia solani*);

[0811] 바나나 병해: 미코스파에렐라 피지엔시스 (*Mycosphaerella fijiensis*), 미코스파에렐라 무시콜라 (*Mycosphaerella musicola*);

[0812] 해바라기 병해: 플라스모파라 할스테디 (*Plasmopara halstedii*);

[0813] 아스퍼질러스 종, 페니실륨 종 (*Penicillium* spp.), 푸사리움 종, 지베렐라 종 (*Gibberella* spp.), 트리코더마 종, 티엘라비옵시스 종 (*Thielaviopsis* spp.), 리조푸스 종, 무코르 종 (*Mucor* spp.), 코르티시움 종 (*Corticium* spp.), 로마 종 (*Rhoma* spp.), 리족토니아 종 (*Rhizoctonia* spp.) 또는 디플로디아 종 (*Diplodia* spp.)에 의해 초래되는 다양한 식물의 초기 성장 단계에서의 병해 또는 종자 병해;

[0814] 폴릭시마 종 (*Polymixa* spp.) 또는 올피디움 종 (*Olpidium* spp.) 등에 의해 매개되는 다양한 식물의 바이러스병 등.

[0815] 3. 제초제에 내성을 부여하는 유전자의 예:

[0816] 성장점 또는 분열조직을 억제하는 제초제에 대한 내성, 예컨대, 이미다졸리논 또는 설포닐유리아, 예를 들어, 각각 문헌 [Lee et al., EMBO J. 7:1241(1988)], 및 [Miki et al., Theor. Appl. Genet. 80:449 (1990)]을 참조한다.

[0817] 글리포세이트 내성 (예를 들어, 스트렙토마이세스 하이그로스코피쿠스 (*Streptomyces hygroscopicus*) 및 스트렙

토마이세스 비리디크로모게네스 (*Streptomyces viridichromogenes*)를 비롯한 스트렙토마이세스 종으로부터의 각각 돌연변이체 5-엔올파이루빌시키메이트-3-포스페이트 신타제(EPSP) 유전자, aroA 유전자 및 글리포세이트 아세틸 트랜스퍼라제(GAT) 유전자에 의해 부여되는 내성), 또는 다른 포스포노 화합물, 예컨대, 글루포시네이트(포스피노트리신 아세틸 트랜스퍼라제(PAT) 유전자에 대한 내성), 및 ACCase 억제제-코딩 유전자에 의한 피리딘옥시 또는 페녹시 프로프리온산 및 시클로헥손에 대한 내성. 예를 들어, 미국 특허 제4,940,835호 및 미국 특허 제6,248,876호, 미국 특허 제4,769,061호, 유럽 특허 제0 333 033호 및 미국 특허 제4,975,374호 참조한다. 또한 유럽 특허 제0242246호, 문헌 [DeGreef et al., Bio/Technology 7:61(1989), Marshall et al., Theor. Appl. Genet. 83:435 (1992)], WO 2005012515 및 문헌 [Castle et. al.] 및 WO 2005107437을 참조한다.

[0818] 광합성을 억제하는 제조체에 대한 내성, 예컨대, 문헌 [Przibila et al., Plant Cell 3:169 (1991)], 미국 특허 제 4,810,648호, 및 문헌 [Hayes et al., Biochem. J. 285: 173 (1992)]의 트라이아진(psbA 및 gs+ 유전자) 또는 벤조니트릴 (니트릴라제 유전자), 및 글루타티온 S-트랜스퍼라제.

[0819] 제초제를 해독하는 효소 또는 억제에 내성이 있는 돌연변이체 글루타민 신타제를 코딩하는 유전자, 예를 들어, 미국 특허 출원 제11/760,602호. 또는 해독 효소는 포스피노트리신 아세틸트랜스퍼라제를 코딩하는 효소(예컨대, 스트렙토마이세스 종으로부터의 bar 또는 pat 단백질)이다. 포스피노트리신 아세틸트랜스퍼라제는, 예를 들어, 미국 특허 제5,561,236호; 제5,648,477호; 제5,646,024호; 제5,273,894호; 제5,637,489호; 제5,276,268호; 제5,739,082호; 제5,908,810호 및 제7,112,665호에 기재되어 있다.

[0820] 히드록시페닐피루베이트 디옥시게나제 (HPPD) 억제제, 천연 발생 HPPD 내성 유전자, 또는 WO 96/38567, WO 99/24585, 및 WO 99/24586, WO 2009/144079, WO 2002/046387, 또는 미국 특허 제6,768,044호에 기재된 바와 같은 돌연변이 또는 키메라 HPPD 효소를 코딩하는 유전자.

[0821] 4. 항생제 스트레스 내성에 관여된 유전자의 예:

[0822] WO 00/04173 또는 WO/2006/045633에 기재된 바와 같은 식물 세포 또는 식물에서의 폴리(ADP-리보스) 중합효소(PARP) 유전자의 발현 및/또는 활성을 감소시킬 수 있는 이식유전자.

[0823] WO 2004/090140에 기재된 바와 같은, 식물 또는 식물 세포의 유전자를 코딩하는 PARG의 발현 및/또는 활성을 감소시킬 수 있는 이식유전자.

[0824] 예를 들어, 유럽 특허 제04077624.7호, WO 2006/133827, PCT/EP07/002,433, 유럽 특허 제1999263호 또는 WO 2007/107326에 기재된 바와 같은 니코틴아미다제, 니코티네이트 포스포리보실트랜스퍼라제, 니코틴산 모노뉴클레오티드 아데닐 트랜스퍼라제, 니코틴아미드 아데닌 다이뉴클레오티드 신타제 또는 니코틴 아미드 포스포리보실트랜스퍼라제를 비롯한 니코틴아미드 아데닌 다이뉴클레오티드 샐비지 합성 경로의 식물-기능성 효소를 코딩하는 이식유전자.

[0825] 탄수화물 생합성에 관여되는 효소는, 예를 들어, 유럽 특허 제0571427호, WO 95/04826, 유럽 특허 제0719338호, WO 96/15248, WO 96/19581, WO 96/27674, WO 97/11188, WO 97/26362, WO 97/32985, WO 97/42328, WO 97/44472, WO 97/45545, WO 98/27212, WO 98/40503, WO99/58688, WO 99/58690, WO 99/58654, WO 00/08184, WO 00/08185, WO 00/08175, WO 00/28052, WO 00/77229, WO 01/12782, WO 01/12826, WO 02/101059, WO 03/071860, WO 2004/056999, WO 2005/030942, WO 2005/030941, WO 2005/095632, WO 2005/095617, WO 2005/095619, WO 2005/095618, WO 2005/123927, WO 2006/018319, WO 2006/103107, WO 2006/108702, WO 2007/009823, WO 00/22140, WO 2006/063862, WO 2006/072603, WO 02/034923, 유럽 특허 제06090134.5호, 유럽 특허 제06090228.5호, 유럽 특허 제06090227.7호, 유럽 특허 제07090007.1호, 유럽 특허 제07090009.7호, WO 01/14569, WO 02/79410, WO 03/33540, WO 2004/078983, WO 01/19975, WO 95/26407, WO 96/34968, WO 98/20145, WO 99/12950, WO 99/66050, WO 99/53072, 미국 특허 제6,734,341호, WO 00/11192, WO 98/22604, WO 98/32326, WO 01/98509, WO 01/98509, WO 2005/002359, 미국 특허 제5,824,790호, 미국 특허 제6,013,861호, WO 94/04693, WO 94/09144, WO 94/11520, WO 95/35026 또는 WO 97/20936에 기재된 것 또는 유럽 특허 제0663956호, WO 96/01904, WO 96/21023, WO 98/39460, 및 WO 99/24593에 개시된 바와 같은 폴리프럭토스, 특히 이눌린 및 레반형(levan-type)의 생성에 관여되는 효소, WO 95/31553, 미국 특허 제2002031826호, 미국 특허 제6,284,479호, 미국 특허 제5,712,107호, WO 97/47806, WO 97/47807, WO 97/47808 및 WO 00/14249에 개시된 것과 같은 알파-1,4-글루칸의 생성, WO 00/73422에 개시된 바와 같은 알파-1,6 분지된 알파-1,4-글루칸의 생성, 예를 들어, WO 00/47727, WO 00/73422, 유럽 특허 제06077301.7, 미국 특허 제5,908,975 및 유럽 특허 제

0728213호에 개시된 바와 같은 알테르난의 생성, 예를 들어, WO 2006/032538, WO 2007/039314, WO 2007/039315, WO 2007/039316, JP 2006304779, 및 WO 2005/012529에 개시된 바와 같은 히알루로난의 생성에 관여되는 효소를 포함한다.

[0826] 가뭄 내성을 개선시키는 유전자. 예를 들어, WO 2013122472는 기능성 유비퀴틴 단백질 리가제 단백질(UPL) 단백질, 보다 특히, UPL3의 부재 또는 감소된 수준이 물에 대한 감소된 필요 또는 상기 식물의 가뭄에 대한 개선된 내성을 야기한다는 것을 개시한다. 증가된 가뭄 내성을 갖는 유전자이식 식물의 다른 예는, 예를 들어, 미국 특허 제2009/0144850호, 미국 특허 제2007/0266453호 및 WO 2002/083911에 개시되어 있다. 미국 특허 제2009/0144850호는 DR02 핵산의 변경된 발현에 기인하여 가뭄 내성 표현형을 나타낸 식물을 기재한다. 미국 특허 제2007/0266453호는 DR03 핵산의 변경된 발현에 기인하여 가뭄 내성 표현형을 나타내는 식물을 기재하고, WO 2002/083911호는 가드 세포에서 발현되는 ABC 수송체의 감소된 활성에 기인하여 가뭄 스트레스에 대해 증가된 내성을 갖는 식물을 기재한다. 다른 예는 유전자이식 식물에서 DREB1 A를 코딩하는 cDNA의 과발현이 정상 성장 조건 하에 많은 스트레스 내성 유전자의 발현을 활성화시키고 가뭄, 염 로딩 및 결빙에 대한 개선된 내성을 초래한다는 것을 기재하는 Kasuga와 그의 공저자 (1999)에 의한 작업이다. 그러나, DREB1A의 발현은 또한 정상 성장 조건 하에 몇몇 성장 지연을 초래하였다 (Kasuga (1999) Nat Biotechnol 17(3) 287-291).

[0827] 추가 특정 실시형태에서, 작물 식물은 특정 식물 형질에 영향을 미침으로써 개선될 수 있다. 예를 들어, 살충제-내성 식물의 개발, 식물에서 질환 내성의 개선, 식물 곤충 및 선충 개성의 개선, 기생충 잡초에 대한 식물 내성의 개선, 식물 가뭄 내성의 개선, 식물 영양가의 개선, 식물 스트레스 내성, 자기-수분 회피, 식물 사료 소화흡수율 생물량, 곡물 수확률 등의 개선에 의함. 소수의 특정 비제한적 예를 본 명세서에서 이하에 제공한다.

[0828] 단일 유전자의 표적화된 돌연변이에 추가로, AD-기능화된 CRISPR 시스템은 다중 유전자의 표적화된 돌연변이, 염색체 단편의 결실, 이식유전자의 부위-특이적 통합, 생체내 부위 지정 돌연변이유발 및 정확한 유전자 대체 또는 식물에서의 대립유전자 교환을 허용하도록 디자인될 수 있다. 따라서, 본 명세서에 기재된 방법은 유전자 발견 및 확인, 돌연변이 및 동종 유전자이식(cisgenic) 육종, 및 잡종 육종에서 광범위하게 적용된다. 이들 출원은 다양한 개선된 작물학적 형질, 예컨대, 제초제 내성, 질환 내성, 항생제 스트레스 내성, 고수율 및 우수한 품질을 갖는 유전자 변형된 작물의 새로운 생성의 생산을 용이하게 한다.

[0829] **웅성 불임 식물을 생성하기 위한 AD-기능화된 CRISPR 시스템의 용도**

[0830] 잡종 식물은 전형적으로 근친교배 식물에 비해 유리한 작물학적 형질을 갖는다. 그러나, 자기-수분 식물에 대해, 잡종의 생성은 도전 중일 수 있다. 상이한 식물 유형에서, 식물 생식력, 보다 특히, 수컷 생식력에 중요한 유전자가 동정되었다. 예를 들어, 메이즈에서, 생식력에 중요한 적어도 2개의 유전자가 동정되었다 (Amitabh Mohanty International Conference on New Plant Breeding Molecular Technologies Technology Development And Regulation, Oct 9-10, 2014, Jaipur, India; Svitashev et al. Plant Physiol. 2015 Oct;169(2):931-45; Djukanovic et al. Plant J. 2013 Dec;76(5):888-99). 본 명세서에 제공된 방법 및 시스템은 잡종을 생성하기 위해 용이하게 교배될 수 있는 웅성 불임성 식물을 생성하기 위해 수컷 생식력에 필요한 유전자를 표적화하는 데 사용될 수 있다. 특정 실시형태에서, 본 명세서에 제공된 AD-기능화된 CRISPR 시스템은 시토크롬 P450-유사 유전자 (MS26) 또는 메가뉴클레아제 유전자 (MS45)의 표적화된 돌연변이유발을 위해 사용되며, 이에 의해 메이즈 식물에 대한 웅성 불임성을 부여한다. 이렇게 해서 유전자가 변경된 메이즈 식물은 잡종 육종 프로그램에서 사용될 수 있다.

[0831] **식물에서 수정 단계의 증가**

[0832] 특정 실시형태에서, 본 명세서에 제공된 방법 및 시스템은 벼 식물과 같은 식물의 수정 단계를 연장시키기 위해 사용된다. 예를 들어, 벼 생식 단계 유전자, 예컨대, Ehd3은 유전자에서 돌연변이를 생성하기 위해 표적화될 수 있고, 소식물체는 연장된 재생 식물 생식 단계를 위해 선택될 수 있다 (CN 104004782에 기재된 바와 같음).

[0833] **관심 작물에서 유전자 변이를 생성시키기 위한 AD-기능화된 CRISPR 시스템의 용도**

[0834] 작물 식물에서 야생 생식질 및 유전자 변형의 이용 가능성은 작물 개선 프로그램에서 중요하지만, 작물 식물로부터의 생식질에서 이용 가능한 다양성은 제한되어 있다. 본 발명은 관심 생식질에서 유전자 변형의 다양성을 생성하기 위한 방법을 계획한다. AD-기능화된 CRISPR 시스템의 이러한 적용에서, 식물 게놈에서 상이한 위치를 표적화하는 가이드 RNA의 라이브러리가 제공되고 CRISPR-Cas 단백질 및 아데노신 데아미나제와 함께 식

물 세포 내로 도입된다. 이 방법에서, 게놈-규모 점 돌연변이 및 유전자 녹아웃의 컬렉션이 생성될 수 있다. 특정 실시형태에서, 상기 방법은 이렇게 얻어진 세포로부터 식물 부분 또는 식물을 생성하는 단계 및 관심 형질에 대해 세포를 스크리닝하는 단계를 포함한다. 표적 유전자는 코딩 영역과 넌코딩 영역을 둘 모두 포함할 수 있다. 특정 실시형태에서, 형질은 스트레스 내성이며, 상기 방법은 스트레스-내성 작물 변종의 생성을 위한 방법이다.

[0835] **과일 숙성에 영향을 미치기 위한 AD-기능화된 CRISPR의 용도**

[0836] 숙성은 과일 및 채소의 성숙 과정에서의 정상적 단계이다. 이를 시작하고 며칠만에 과일 또는 채소는 먹을 수 없게 된다. 이 과정은 농부와 소비자 둘 모두에게 상당한 손실을 가져온다. 특정 실시형태에서, 본 발명의 방법은 에틸렌 생성을 감소시키는 데 사용된다. 이는 다음 중 하나 이상을 보장함으로써 보장된다: a. ACC 신타제 유전자 발현의 억제. ACC(1-아미노시클로프로판-1-카복실산) 신타제는 S-아데노실메티오닌(SAM)의 ACC로의 전환을 담당하는 효소이고; 에틸렌 생합성의 두번째 단계에서 마지막 단계이다.. 효소 발현은 안티센스("거울상") 또는 신타제 유전자의 절단된 복제물이 식물 게놈에 삽입될 때 입체방해된다; b. ACC 데아미나제 유전자의 삽입. 효소를 코딩하는 유전자는 공통 비병원성 토양 박테리아인 슈도모나스 클로로라피스 (*Pseudomonas chlororaphis*)로부터 얻어진다. 이것은 ACC를 상이한 화합물로 전환시킴으로써 에틸렌 생성에 이용 가능한 ACC의 양을 감소시킨다; c. SAM 가수분해효소 유전자의 삽입. 이 접근은 ACC 데아미나제와 유사한데, 여기서 에틸렌 생성은 이의 전구체 대사물질의 양이 감소될 때 방해되며; 이 경우에 SAM이 호모세린으로 전환된다. 효소를 코딩하는 유전자는 이. 콜라이 T3 박테리오파지로부터 얻어진다. 그리고 d. ACC 옥시다제 유전자 발현의 억제. ACC 옥시다제는 에틸렌 생합성 경로에서의 마지막 단계인 ACC의 에틸렌으로의 산화를 촉매하는 효소이다. 본 명세서에 기재된 방법을 이용하여, ACC 옥시다제 유전자의 하향 조절은 에틸렌 생성의 억제를 초래하고, 이에 의해 과일 숙성을 지연시킨다. 특정 실시형태에서, 상기 기재한 변형에 추가적으로 또는 대안적으로, 본 명세서에 기재된 방법은 과일에 의해 얻어지는 에틸렌 신호를 방해하기 위해 에틸렌 수용체를 변형시키는 데 사용된다. 특정 실시형태에서, 에틸렌 결합 단백질을 코딩하는 ETR1 유전자의 발현은 변형되며, 보다 특히, 억제된다. 특정 실시형태에서, 상기 기재한 변형에 추가적으로 또는 대안적으로, 본 명세서에 기재된 방법은 식물 세포벽의 완전성을 유지하는 물질인 펙틴의 분해를 초래하는 효소인 폴리갈락투로나제 (PG)를 코딩하는 유전자의 발현을 변형시키기 위해 사용된다. 펙틴 분해는 과일의 연화를 야기하는 숙성 과정의 시작 시 일어난다. 따라서, 특정 실시형태에서, 본 명세서에 기재된 방법은 생성된 PG 효소의 양을 감소시킴으로써 펙틴 분해를 지연시키기 위해 PG 유전자에 돌연변이를 도입하거나 또는 PG 유전자의 활성화를 억제하기 위해 사용된다.

[0837] 따라서 특정 실시형태에서, 상기 방법은 상기 기재한 바와 같은 식물 세포 게놈의 하나 이상의 변형을 보장하고 이로부터 식물을 재생하기 위한 AD-기능화된 CRISPR 시스템의 용도를 포함한다. 특정 실시형태에서, 식물은 토마토 식물이다.

[0838] **식물 저장 수명의 증가**

[0839] 특정 실시형태에서, 본 발명의 방법은 식물 또는 식물 부분의 저장 수명에 영향을 미치는 화합물의 생성에 관여되는 유전자를 변형시키기 위해 사용된다. 보다 특히, 변형은 감자 괴경에서 환원당 축적을 방지하는 유전자에서의 변형이다. 고온 처리 시, 이들 환원당은 유리 아미노산과 반응하여 갈색의, 쓴 맛이 나는 생성물 및 잠재적 발암 물질인 아크릴아미드의 상승된 수준을 초래한다. 특정 실시형태에서, 본 명세서에 제공되는 방법은 수크로스를 글루코스 및 프럭토스로의 분해시키는 단백질을 코딩하는 액포성 인버타제 유전자 (VInv)의 발현을 감소시키거나 또는 억제하기 위해 사용된다 (Clasen et al. DOI: 10.1111/pbi.12370).

[0840] **가치 부가 형질을 보장하기 위한 AD-기능화된 CRISPR 시스템의 용도**

[0841] 특정 실시형태에서, AD-기능화된 CRISPR 시스템은 영양적으로 개선된 농작물을 생성하기 위해 사용된다. 특정 실시형태에서, 본 명세서에 제공된 방법은 "기능성 식품", 즉, 그것이 함유하는 전통적 영양소 이상으로 건강상 이득을 제공할 수 있는 변형 식품 또는 식품 성분 및 또는 "뉴트라슈티컬 (nutraceutical)", 즉, 식품 또는 식품의 부분으로 고려되고 질환의 예방 및 치료를 비롯한 건강상 이익을 제공할 수 있는 물질을 생산하는 데 적합하다. 특정 실시형태에서, 뉴트라슈티컬은 암, 당뇨병, 심혈관 질환 및 고혈압 중 하나 이상의 예방 및/또는 치료에서 유용하다.

[0842] 영양적으로 개선된 작물의 예는 하기를 포함한다 (Newell-McGloughlin, Plant Physiology, July 2008, Vol. 147, pp. 939-953):

[0843] 변형된 단백질 품질, 함량 및/또는 아미노산 조성물은, 예컨대, 바히아그래스 (Luciani et al. 2005, Florida Genetics Conference Poster), 카놀라 (Roesler et al., 1997, Plant Physiol 113 75-81), 메이즈 (Cromwell et al, 1967, 1969 J Anim Sci 26 1325-1331, O'Quin et al. 2000 J Anim Sci 78 2144-2149, Yang et al. 2002, Transgenic Res 11 11-20, Young et al. 2004, Plant J 38 910-922), 감자 (Yu J and Ao, 1997 Acta Bot Sin 39 329-334; Chakraborty et al. 2000, Proc Natl Acad Sci USA 97 3724-3729; Li et al. 2001) Chin Sci Bull 46 482-484, 벼 (Katsube et al. 1999, Plant Physiol 120 1063-1074), 대두 (Dinkins et al. 2001, Rapp 2002, In Vitro Cell Dev Biol Plant 37 742-747), 고구마 (Egnin and Prakash 1997, In Vitro Cell Dev Biol 33 52A)에 대해 기재되어 있다.

[0844] 필수 아미노산 함량은, 예컨대, 카놀라 (Falco et al. 1995, Bio/Technology 13 577-582), 루핀 (White et al. 2001, J Sci Food Agric 81 147-154), 메이즈 (Lai and Messing, 2002, Agbios 2008 GM crop database (March 11, 2008)), 감자 (Zeh et al. 2001, Plant Physiol 127 792-802), 수수(Zhao et al. 2003, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 413-416), 대두 (Falco et al. 1995 Bio/Technology 13 577-582; Galili et al. 2002 Crit Rev Plant Sci 21 167-204)에 대해 기재되어 있다.

[0845] 오일 및 지방산, 예컨대, 카놀라 (Dehesh et al. (1996) Plant J 9 167-172 [PubMed]; Del Vecchio (1996) INFORM International News on Fats, Oils and Related Materials 7 230-243; Roesler et al. (1997) Plant Physiol 113 75-81 [PMC free article] [PubMed]; Froman and Ursin (2002, 2003) Abstracts of Papers of the American Chemical Society 223 U35; James et al. (2003) Am J Clin Nutr 77 1140-1145 [PubMed]; Agbios(2008, 상기 참조); 목화 (Chapman et al. (2001). J Am Oil Chem Soc 78 941-947; Liu et al. (2002) J Am Coll Nutr 21 205S-211S [PubMed]; O'Neill (2007) Australian Life Scientist. http://www.biotechnews.com.au/index.php/id;866694817;fp;4;fpid;2 (June 17, 2008), 아마씨 (Abbadi et al., 2004, Plant Cell 16: 2734-2748), 메이즈 (Young et al., 2004, Plant J 38 910-922), 기름야자 (Jalani et al. 1997, J Am Oil Chem Soc 74 1451-1455; Parveez, 2003, AgBiotechNet 113 1-8), 벼 (Anai et al., 2003, Plant Cell Rep 21 988-992), 대두 (Reddy and Thomas, 1996, Nat Biotechnol 14 639-642; Kinney and Kwolton, 1998, Blackie Academic and Professional, London, pp 193-213), 해바라기 (Arcadia, Biosciences 2008);

[0846] 탄수화물, 예컨대, 치커리 (Smeekens (1997) Trends Plant Sci 2 286-287, Sprenger et al. (1997) FEBS Lett 400 355-358, Sevenier et al. (1998) Nat Biotechnol 16 843-846), 메이즈 (Caimi et al. (1996) Plant Physiol 110 355-363), 감자 (Hellwege et al., 1997 Plant J 12 1057-1065), 사탕무 (Smeekens et al. 1997, 상기 참조)에 대해 기재된 프룩탄, 예컨대, 감자 (Hellewege et al. 2000, Proc Natl Acad Sci USA 97 8699-8704)에 대해 기재된 이눌린, 예컨대, 벼 (Schwall et al. (2000) Nat Biotechnol 18 551-554, Chiang et al. (2005) Mol Breed 15 125-143)에 대해 기재된 전분;

[0847] 예컨대, 카놀라 (Shintani and DellaPenna (1998) Science 282 2098-2100), 메이즈 (Rocheford et al. (2002). J Am Coll Nutr 21 191S-198S, Cahoon et al. (2003) Nat Biotechnol 21 1082-1087, Chen et al. (2003) Proc Natl Acad Sci USA 100 3525-3530), 겨자씨 (Shewmaker et al. (1999) Plant J 20 401-412), 감자 (Ducreux et al., 2005, J Exp Bot 56 81-89), 벼 (Ye et al. (2000) Science 287 303-305), 딸기 (Agius et al. (2003), Nat Biotechnol 21 177-181), 토마토 (Rosati et al. (2000) Plant J 24 413-419, Fraser et al. (2001) J Sci Food Agric 81 822-827, Mehta et al. (2002) Nat Biotechnol 20 613-618, Díaz de la Garza et al. (2004) Proc Natl Acad Sci USA 101 13720-13725, Enfissi et al. (2005) Plant Biotechnol J 3 17-27, DellaPenna (2007) Proc Natl Acad Sci USA 104 3675-3676)에 대해 기재된, 비타민 및 카로테노이드;

[0848] 예컨대, 사과 (스틸벤, Szankowski et al. (2003) Plant Cell Rep 22: 141-149), 알팔파 (레스베라트롤, Hipskind and Paiva (2000) Mol Plant Microbe Interact 13 551-562), 키위 (레스베라트롤, Kobayashi et al. (2000) Plant Cell Rep 19 904-910), 메이즈 및 대두 (플라보노이드, Yu et al. (2000) Plant Physiol 124 781-794), 감자 (안토시아닌 및 알칼로이드 글리코사이드, Lukaszewicz et al. (2004) J Agric Food Chem 52 1526-1533), 벼 (플라보노이드 및 레스베라트롤, Stark-Lorenzen et al. (1997) Plant Cell Rep 16 668-673, Shin et al. (2006) Plant Biotechnol J 4 303-315), 토마토 (+레스베라트롤, 클로로겐산, 플라보노이드, 스틸벤; Rosati et al. (2000) 상기 참조, Muir et al. (2001) Nature 19 470-474, Niggeweg et al. (2004) Nat Biotechnol 22 746-754, Giovinazzo et al. (2005) Plant Biotechnol J 3 57-69), 밀 (카페인 및 페룰산, 레스베라트롤; United Press International (2002))에 대해 기재된, 기능성 2차 대사물질; 및

[0849] 예컨대, 알팔파 (피타제, Austin-Phillips et al. (1999) http://www.molecularfarming.com/nonmedical.html), 상추 (철, Goto et al. (2000) Theor Appl Genet 100 658-664), 벼 (철, Lucca et al. (2002) J Am Coll Nutr 21 184S-190S), 메이즈, 대두 및 밀 (피타제, Drakakaki et al. (2005) Plant Mol Biol 59 869-880, Denbow et al. (1998) Poult Sci 77 878-881, Brinch-Pedersen et al. (2000) Mol Breed 6 195-206)에 대해 기재된, 미네랄 이용 가능성.

[0850] 특정 실시형태에서, 가치-부가된 형질은 식물에 존재하는 화합물의 생각되는 건강상의 이점에 관한 것이다. 예를 들어, 특정 실시형태에서, 가치-부가된 작물은 다음의 화합물 중 하나 이상의 변형을 보장하거나 또는 다음의 화합물 중 하나 이상의 합성을 유도/증가시키기 위해 본 발명의 방법을 적용함으로써 얻어진다:

[0851] ● 카로테노이드, 예컨대, 세포에 대한 손상을 야기할 수 있는 유리 라디칼을 중화시키는 당근에 존재하는 α-카로텐 또는 유리 라디칼을 중화시키는 다양한 과일 및 채소에 존재하는 β-카로텐;

[0852] ● 건강한 시력의 유지에 기여하는 녹색 채소에 존재하는 루테인;

[0853] ● 전립선 암의 위험을 감소시키는 것으로 여겨지는, 토마토 및 토마토 생성물에 존재하는 리코펜;

[0854] ● 건강한 시력의 유지에 기여하는 감귤류 및 메이즈에 존재하는 제아잔틴;

[0855] ● 유방 및/또는 결장암의 위험을 감소시킬 수 있는 밀겨에 존재하는 불용성 섬유질과 같은 식이섬유 및 귀리에 존재하는 β-글루칸, 심혈관 질환(CVD)의 위험을 감소시키는 금불초 및 전체 곡류에 존재하는 가용성 섬유질;

[0856] ● 지방산, 예컨대, CVD 위험을 감소시키고 정신 및 시각 기능을 개선시키는 ω-3 지방산, 신체 조성을 개선시킬 수 있고, 특정 암의 위험을 감소시킬 수 있는 접합된 리놀렌산 및 암 및 CVD의 염증 위험을 감소시킬 수 있고, 신체 조성을 개선시킬 수 있는 GLA;

[0857] ● 플라보노이드, 예컨대, 항산화제-유사 활성을 갖고, 퇴행성 질환의 위험을 감소시킬 수 있는, 밀에 존재하는 히드록시신나메이트, 자유 라디칼을 중화시키고 암 위험을 감소시킬 수 있는 과일 및 채소에 존재하는 플라보놀, 카테신;

[0858] ● 글루코시놀레이트, 인돌, 아이소티오시아네이트, 예컨대, 자유 라디칼을 중화시키고, 암 위험을 감소시킬 수 있는, 십자화과 채소(브로콜리, 케일), 겨자무에 존재하는 설포라판;

[0859] ● 페놀 수지류, 예컨대, 퇴행성 질환, 심장병 및 암 위험을 감소시킬 수 있고, 장기간의 효과를 가질 수 있는 포도에 존재하는 스틸벤 및 항산화제-유사 활성을 갖고, 퇴행성 질환, 심장병, 및 안질환 위험을 감소시킬 수 있는 채소 및 감귤류에 존재하는 카페인산 및 페룰산, 및 항산화제-유사 활성을 갖고, 퇴행성 질환 및 심장병의 위험을 감소시킬 수 있는 카카오에 존재하는 에피카테킨;

[0860] ● 혈액 콜레스테롤 수준을 낮춤으로써 관상동맥 심장병의 위험을 감소시키는 메이즈, 콩, 밀 및 목재 오일에 존재하는 식물 스탄올/스테롤;

[0861] ● 위장 건강을 개선시킬 수 있는 예루살렘 아티초크, 샬롯, 양파 분말에 존재하는 프럭탄, 이눌린, 프룩토-올리고당;

[0862] ● LDL 콜레스테롤을 저하시킬 수 있는 대두에 존재하는 사포닌;

[0863] ● 심장병 위험을 감소시킬 수 있는 대두에 존재하는 대두 단백질;

[0864] ● 식물성 에스트로겐, 예컨대, 폐경기 증상, 예컨대, 일과성 열감을 감소시킬 수 있고, 골다공증 및 CVD를 감소시킬 수 있는 대두에 존재하는 이소플라본, 및 심장병 및 일부 암에 대해 보호할 수 있고, LDL 콜레스테롤, 총 콜레스테롤을 낮출 수 있는 아마, 호밀 및 채소에 존재하는 리그난;

[0865] ● 설파이드 및 티올, 예컨대, 양파, 마늘, 올리브, 리크 및 스캘런에 존재하는 디알릴 설파이드 및 LDL 콜레스테롤을 저하시킬 수 있고 건강한 면역계를 유지하도록 도움을 주는 십자화과 채소에 존재하는 알릴 메틸 트리설파이드, 디티올티온; 및

[0866] ● 탄닌, 예컨대, 요로 건강상태를 개선시킬 수 있고, CVD 및 고혈압 위험을 감소시킬 수 있는 체키, 코코아에 존재하는 프로안토시아니딘.

[0867] 추가로, 본 발명의 방법은 또한 단백질/전분 작용성, 저장 수명, 맛/미관, 섬유 품질, 및 알레르겐, 항영양소(antinutrient) 및 독소 감소 형질의 변형을 계획한다.

[0868] 따라서, 본 발명은 영양적 가치가 더해진 식물을 생산하는 방법을 포함하고, 상기 방법은 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템을 이용하여 영양적 가치가 더해진 성분의 생산에 관여되는 유전자를 코딩하는 유전자를 식물 세포 내로 도입하는 단계 및 상기 식물 세포로부터 식물을 재생하는 단계를 포함하고, 상기 식물은 영양적 가치가 더해진 상기 성분의 발현 증가를 특징으로 한다. 특정 실시형태에서, AD-기능화된 CRISPR 시스템은, 예를 들어, 본 화합물의 대사를 제어하는 하나 이상의 전사 인자를 변형시킴으로써, 이들 화합물의 내생성 합성을 간접적으로 변형시키는 데 사용된다. 식물 세포 내로 관심 유전자를 도입하고/하거나 AD-기능화된 CRISPR 시스템을 이용하여 내생성 유전자를 변형시키는 방법은 본 명세서에서 상기에 기재되어 있다.

[0869] 가치-부가된 형질을 부여하도록 변형된 식물에서 변형의 일부 특정 예는 식물의 스테아르산 함량을 증가시키도록 스테아릴-ACP 불포화효소의 안티센스 유전자로 식물을 형질전환시켜서 지방산 대사작용이 변형된 식물이다. 문헌 [Knultzon et al., Proc. Natl. Acad. Sci. U.S.A. 89:2624 (1992)] 참조한다. 다른 예는 예를 들어 저수준의 피트산을 특징으로 하는 메이즈 돌연변이체의 원인일 수 있는 단일 대립유전자와 연관된 DNA를 클로닝하여 재도입시킴으로써, 피테이트 함량을 감소시키는 것을 포함한다. 문헌 [Raboy et al, Maydica 35:383 (1990)] 참조한다.

[0870] 유사하게, 강력 프로모터의 제어 하에서 메이즈 호분층 내 플라보노이드의 생산을 조절하는 메이즈 (제아 메이스 (Zea mays)) Tfs C1 및 R의 발현은 아마도 전체 경로를 활성화시킴으로써, 애기장대 (아라비돕시스 탈리아나)에서 안토시아닌의 높은 축적률을 일으켰다 (Bruce et al., 2000, Plant Cell 12:65-80). DellaPenna (Welsch et al., 2007 Annu Rev Plant Biol 57: 711-738)는 Tf RAP2.2 및 그의 상호작용 상대 SINAT2가 애기장대 잎에서 카로텐 발생을 증가시켰다는 것을 발견하였다. Tf Dof1의 발현은 탄소 골격 생성을 위한 효소를 코딩하는 유전자의 상향조절, 아미노산 함량의 현저한 증가 및 유전자이식 애기장대에서 Glc 수준의 감소 (Yanagisawa, 2004 Plant Cell Physiol 45: 386-391), 및 애기장대에서의 글루코시놀레이트 생합성 경로의 DOF Tf AtDof1.1(OBP2) 상향조절된 모든 단계를 유도하였다 (Skirycz et al., 2006 Plant J 47: 10-24).

[0871] **식물에서 알레르겐의 감소**

[0872] 특정 실시형태에서, 본 명세서에 제공되는 방법은 알레르겐 수준이 감소된 식물을 생성하여, 그들을 소비자에게 더 안전하게 만들기 위해 사용된다. 특정 실시형태에서, 상기 방법은 식물 알레르겐을 생성을 초래하는 하나 이상의 유전자 발현을 변형시키는 단계를 포함한다. 예를 들어, 특정 실시형태에서, 상기 방법은 식물 세포, 예컨대, 독보리 식물 세포에서 Lol p5 유전자의 발현을 하향 조절하는 단계 및 상기 식물의 화분의 알레르기 항원성을 감소시키기 위해 그들로부터 식물을 재생시키는 단계를 포함한다 (Bhalla et al. 1999, Proc. Natl. Acad. Sci. USA Vol. 96: 11676-11680).

[0873] 일반적으로 땅콩 알레르기 및 콩과 식물에 대한 알레르기는 실제로 심각한 건강상의 문제이다. 본 발명의 AD-기능화된 CRISPR 시스템은 이러한 콩과 식물의 알레르기성 단백질을 코딩하는 유전자를 동정하고, 돌연변이시키기 위해 사용될 수 있다. 이러한 유전자 및 단백질에 관해 제한없이, Nicolaou 등은 땅콩, 대두, 렌틸콩, 완두콩, 루핀, 그린 빈스, 및 녹두에서의 알레르기 단백질을 동정하였다. 문헌 [Nicolaou et al., Current Opinion in Allergy and Clinical Immunology 2011;11(3):222)]을 참조한다.

[0874] **관심 내생성 유전자에 대한 스크리닝 방법**

[0875] 본 명세서에 제공된 방법은 식물 종, 식물 문 및 식물계 전체에서, 영양적 부가 가치 성분의 생산에 관여되는 효소를 코딩하는 가치있는 유전자 또는 일반적으로 관심 작물적 형질에 영향을 미치는 유전자의 동정을 더욱 허용한다. 예를 들어, 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템을 이용하여 식물에서 대사 경로의 효소를 코딩하는 유전자를 선택적으로 표적화함으로써, 식물의 소정의 영양학적 양상을 초래하는 유전자가 동정될 수 있다. 유사하게는, 바람직한 작물학적 형질에 영향을 미치는 유전자를 선택적으로 표적화함으로써, 적절한 유전자가 동정될 수 있다. 따라서, 본 발명은 특정 영양학적 가치 및/또는 작물학적 형질을

갖는 화합물의 생성에 관여되는 효소를 코딩하는 유전자에 대한 스크리닝 방법을 포함한다.

[0876] **식물 및 효모에서 AD-기능화된 CRISPR 시스템의 추가 적용 분야**

[0877] **바이오연료 생산에서 AD-기능화된 CRISPR 시스템의 용도**

[0878] 본 명세서에서 사용되는 용어 "바이오연료"는 식물 및 식물-유래 자원으로부터 제조된 대안의 연료이다. 재생성 바이오연료는 에너지가 탄소 고정 방법을 통해 얻어지거나 또는 생물량의 사용 또는 전환을 통해 제조되는 유기 물질로부터 추출될 수 있다. 이러한 생물량은 바이오연료에 대해 직접적으로 사용될 수 있거나 또는 열전환, 화학적 전환 및 생화학적 전환에 의해 물질을 함유하는 편리한 에너지로 전환될 수 있다. 이 생물량 전환은 고체, 액체 또는 기체 형태로 연료를 생성할 수 있다. 2가지 유형의 바이오연료가 있다: 바이오에탄올 및 바이오디젤. 바이오에탄올은 대부분 메이즈 및 사탕수수로부터 유래된 셀룰로스 (전분)의 당 발효 공정에 의해 주로 생산된다. 반면에 바이오디젤은 주로 유채, 야자나무 및 대두와 같은 기름 작물로부터 생산된다. 바이오연료는 주로 수송을 위해 사용된다.

[0879] **바이오연료 생산을 위한 식물 성질 향상**

[0880] 특정 실시형태에서, 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템을 이용하는 방법은 발효를 위한 당의 더 효율적인 방출을 위해 중요한 가수분해제에 의한 접근을 용이하게 하기 위하여 세포벽의 특성을 변경시키는 데 사용된다. 특정 실시형태에서, 셀룰로스 및/또는 리그닌의 생합성은 변형된다. 셀룰로스는 세포벽의 주요 성분이다. 셀룰로스 및 리그닌의 생합성은 공동조절된다. 식물에서 리그닌 생산을 감소시킴으로써, 셀룰로스의 비율은 증가될 수 있다. 특정 실시형태에서, 본 명세서에 기재된 방법은 발효 가능한 탄수화물을 증가시키기 위해 식물에서 리그닌 생합성을 하향조절하는 데 사용된다. 보다 특히, 본 명세서에 기재된 방법은 WO 2008064289 A2에 개시된 바와 같이 4-쿠마레이트 3-히드록실라제 (C3H), 페닐알라닌 암모니아-리아제 (PAL), 신남에이트 4-히드록실라제 (C4H), 히드록시신남오일 트랜스퍼라제 (HCT), 카페인산 O-메틸트랜스퍼라제 (COMT), 카페오일 CoA 3-O-메틸트랜스퍼라제 (CCoAOMT), 페룰레이트 5-히드록실라제 (F5H), 신남일 알콜 디히드로게나제 (CAD), 신남오일 CoA-리덕타제 (CCR), 4-쿠마레이트-CoA 리가제 (4CL), 모노리그놀-리그닌-특이적 글리코실트랜스퍼라제, 및 알데히드 디히드로게나제 (ALDH)로 이루어진 군으로부터 선택되는 적어도 첫 번째의 리그닌 생합성 유전자를 하향조절하는 데 사용된다.

[0881] 특정 실시형태에서, 본 명세서에 기재된 방법은 발효 동안 더 낮은 아세트산 수준을 생성하는 식물 매스 (mass)를 생산하는 데 사용된다 (또한 WO 2010096488 참조). 보다 특히, 본 명세서에 개시된 방법은 다당 아세틸화를 감소시키기 위해 CasIL에 대한 상동체에서 돌연변이를 생성하는 데 사용된다.

[0882] **바이오연료 생산을 위해 변형된 효모**

[0883] 특정 실시형태에서, 본 명세서에 제공된 AD-기능화된 CRISPR 시스템은 재조합 미생물에 의한 바이오에탄올 생산을 위해 사용된다. 예를 들어, AD-기능화된 CRISPR 시스템은 미생물, 예컨대, 효모를 조작하기 위해, 발효 가능한 당으로부터 바이오연료 또는 생체중합체를 생성하기 위해 그리고 선택적으로 발효 가능한 당의 공급원으로서 농업 폐기물로부터 유래된 식물-유래 리그노셀룰로스를 분해할 수 있게 하는데 사용될 수 있다. 일부 실시형태에서, 시스템은 바이오연료 생산 경로와 경쟁하는 내생성 대사 경로를 변형하는 데 사용된다.

[0884] 따라서, 보다 특정한 실시형태에서, 본 명세서에 기술된 방법은 다음과 같이, 상기 숙주 세포 내 대사 경로의 효소를 코딩하는 적어도 하나의 핵산을 변형시키거나 또는 상기 효소의 억제제를 코딩하는 적어도 하나의 핵산을 도입시키기 위해서, 미생물을 변형시키는데 사용되며, 여기서 상기 경로는 피루베이트로부터 아세트알데히드 또는 아세트알데히드로부터 에탄올 이외의 대사산물을 생산하고, 상기 변형은 상기 대사산물의 감소된 생산을 야기시킨다.

[0885] **식물성 오일 또는 바이오연료의 생산을 위해 변형된 조류 및 식물**

[0886] 유전자이식 조류 또는 다른 식물, 예컨대, 유채는, 예를 들어, 식물성 오일 또는 바이오연료, 예컨대, 알콜(특히 메탄올 및 에탄올)의 생산에 특히 유용할 수 있다. 이들은 오일 또는 바이오연료 산업에서 사용하기 위한 고수준의 오일 또는 알콜을 발현시키거나 또는 과발현시키도록 조작될 수 있다.

[0887] 본 발명의 특정 실시형태에 따르면, AD-기능화된 CRISPR 시스템은 바이오연료 생산에서 유용한 지질-풍부 규조류를 생성하는 데 사용된다.

[0888] 특정 실시형태에서, 조류 세포에 의해 생성된 지질의 양 및/또는 지질의 질의 변형에 연루된 유전자를 특이적으

로 변형시키는 것으로 예상된다. 지방산 합성 경로에 연루된 유전자 코딩 효소의 예는, 예를 들어, 아세틸-CoA 카복실라제, 지방산 신타제, 3-케토아실_아실- 캐리어 단백질 신타제 III, 글리세롤-3-인산 디히드로게나제 (G3PDH), 엔오일-아실 캐리어 단백질 리덕타제 (엔오일-ACP-리덕타제), 글리세롤-3-포스페이트 아실트랜스퍼라제, 리소포스파티드 아실 트랜스퍼라제 또는 다이아실글리세롤 아실트랜스퍼라제, 인지질:다이아실글리세롤 아실트랜스퍼라제, 포스파티데이트 포스파타제, 지방산 티오에스터라제, 예컨대, 팔미토일 단백질 티오에스터라제, 또는 말산 효소 활성을 갖는 단백질을 코딩할 수 있다. 추가 실시형태에서, 지질 축적이 증가된 규조류를 생성시키는 것이 계획된다. 이는 지질 이화작용을 감소시키는 유전자를 표적화시킴으로써 달성될 수 있다. 특히, 트리아실글리세롤 및 유리 지방산 둘 모두의 활성화에 관여되는 유전자뿐만 아니라 지방산, 예컨대, 아실-CoA 신타제, 3-케토아실-CoA 티올라제, 아실-CoA 옥시다제 활성 및 포스포글루코뮤타제의 β-산화에 직접적으로 관여되는 유전자에 본 발명의 방법에서의 사용에 대한 관심이 있다. 본 명세서에 기재된 AD-기능화된 CRISPR 시스템 및 방법은 그들의 지질 함량을 증가시키기 위해 규조류에서 이러한 유전자를 특이적으로 활성화시키는 데 사용될 수 있다.

[0889] 유기체, 예컨대, 미세조류는 합성 생물학을 위해 널리 사용된다. Stovicek 등 (Metab. Eng. Comm., 2015; 2:13]은 산업적 생산을 위해 강한 균주를 효율적으로 생산하기 위한 산업적 효모, 예를 들어 사카로마이세스 세레비시애의 게놈 편집을 기재한다. Stovicek은 내생성 유전자의 대립유전자와 이종성 유전자에서의 녹아웃을 둘 모두를 동시에 붕괴시키는 효모에 대해 코돈-최적화된 CRISPR-Cas9 시스템을 사용하였다. Cas9 및 가이드 RNA는 게놈 또는 에피솜 2μ-기반 벡터 위치로부터 발현되었다. 저자는 또한 유전자 붕괴 효율이 Cas9 및 가이드 RNA 발현 수준의 최적화에 의해 개선될 수 있다는 것을 보여주었다. Hlavova 등 (Biotechnol. Adv. 2015)은 삽입 돌연변이유발 및 스크리닝을 위해 핵 및 엽록체를 표적화하기 위해서 CRISPR와 같은 기법을 이용하는 미세조류의 종 또는 균주 발생을 논의한다.

[0890] 미국 특허 제8945839호는 Cas9를 이용하여 미세조류 (클라미도모나스 레인하르티 세포) 종)을 조작하는 방법을 기재한다. 유사한 도구를 이용하여, 본 명세서에 기재된 AD-기능화된 CRISPR 시스템의 방법은 클라미도모나스 종 및 기타 조류에 적용될 수 있다. 특정 실시형태에서, CRISPR-Cas 단백질 (예를 들어, Cas13), 아데노신 데아미나제 (CRISPR-Cas 단백질 또는 압타머-결합 어댑터 단백질에 융합될 수 있음), 및 가이드 RNA는 항상성 프로모터, 예컨대, Hsp70A-Rbc S2 또는 베타2-튜불린의 제어 하에 CRISPR-Cas 단백질 및 선택적으로 아데노신 데아미나제를 발현시키는 벡터를 이용하여 발현된 조류에 도입된다. 가이드 RNA는 T7 프로모터를 함유하는 벡터를 이용하여 전달될 것이다. 대안적으로, mRNA 및 시험관내 전사된 가이드 RNA는 조류 세포로 전달될 수 있다. 전기천공법 프로토콜은 GeneArt 클라미도모나스 조작 키트의 표준 권장된 프로토콜에 따른다.

[0891] **지방산을 생산할 수 있는 미생물의 생성에서 AD-기능화된 CRISPR 시스템의 용도**

[0892] 특정 실시형태에서, 본 발명의 방법은 지방 에스테르, 예컨대, 지방산 메틸 에스테르 ("FAME") 및 지방산 에틸 에스테르 ("FAEE")를 생산할 수 있는 유전자 조작된 미생물의 생성을 위해 사용된다.

[0893] 전형적으로, 숙주 세포는 티오에스터라제를 코딩하는 유전자, 아실-CoA 신타제를 코딩하는 유전자, 및 에스테르 신타제를 코딩하는 유전자의 발현 또는 과발현에 의해 배지에 존재하는 탄소 공급원, 예컨대, 알콜로부터의 지방 에스터를 생산하도록 조작될 수 있다. 따라서, 본 명세서에 제공된 방법은 티오에스터라제 유전자, 아실-CoA 신타제를 코딩하는 유전자, 및 에스테르 신타제를 코딩하는 유전자를 과발현시키거나 또는 도입하기 위해 미생물을 변형하는 데 사용된다. 특정 실시형태에서, 티오에스터라제 유전자는 tesA, 'tesA, tesB, fatB, fatB2, fatB3, fatA1 또는 fatA로부터 선택된다. 특정 실시형태에서, 아실-CoA 신타제를 코딩하는 유전자는 fadDJadK, BH3103, pfl-4354, EAV15023, fadD1, fadD2, RPC_4074, fadDD35, fadDD22, faa39, 또는 동일한 특성을 갖는 효소를 코딩하는 동정된 유전자로부터 선택된다. 특정 실시형태에서, 에스테르 신타제를 코딩하는 유전자는 심몬드시아 키넨시스 (*Simmondsia chinensis*), 아시네토박터 종 (*Acinetobacter* sp.) ADP, 알카니보락스 보르쿠멘시스 (*Alcanivorax borkumensis*), 슈도모나스 애루지노사 (*Pseudomonas aeruginosa*), 펀디박터 자덴시스 (*Fundibacter jadensis*), 아라비돕시스 탈리아나 또는 알칼리게네스 유트로푸스 (*Alkaligenes eutrophus*) 유래의 신타제/아실-CoA:다이아실글리세릴 아실트랜스퍼라제 또는 이의 변이체를 코딩하는 유전자이다. 추가적으로 또는 대안적으로, 본 명세서에 제공된 방법은 상기 미생물에서 아실-CoA 디히드로게나제를 코딩하는 유전자, 외막 단백질 수용체를 코딩하는 유전자, 및 지방산 생합성의 전사 조절제를 코딩하는 유전자 중 적어도 하나의 발현을 감소시키도록 사용된다. 특정 실시형태에서, 이들 유전자 중 하나 이상은, 예컨대, 돌연변이의 도입에 의해 불활성화된다. 특정 실시형태에서, 아실-CoA 디히드로게나제를 코딩하는 유전자는 fadE이다. 특정 실시형태에서, 지방산 생합성의 전사 조절자를 코딩하는 유전자는 DNA 전

사 리프레서, 예를 들어, fabR을 코딩한다.

[0894] 추가적으로 또는 대안적으로, 상기 미생물은 피루베이트 포르메이트 리아제를 코딩하는 유전자, 락테이트 디히드로게나제를 코딩하는 유전자 중 적어도 하나, 또는 둘 모두의 발현을 감소시키도록 변형된다. 특정 실시형태에서, 피루베이트 포르메이트 리아제를 코딩하는 유전자는 pflB이다. 특정 실시형태에서, 락테이트 디히드로게나제를 코딩하는 유전자는 IdhA이다. 특정 실시형태에서, 이들 유전자 중 하나 이상은, 예컨대, 그에 대한 돌연변이의 도입에 의해 불활성화된다.

[0895] 특정 실시형태에서, 미생물은 에스케리치아 (*Escherichia*), 바실러스 (*Bacillus*), 락토바실러스 (*Lactobacillus*), 로도코커스 (*Rhodococcus*), 시네코코커스 (*Synechococcus*), 시네코이스티스 (*Synechoystis*), 슈도모나스 (*Pseudomonas*), 아스퍼질러스 (*Aspergillus*), 트리코더마 (*Trichoderma*), 뉴로스포라 (*Neurospora*), 푸사리움 (*Fusarium*), 휴미콜라 (*Humicola*), 리조무코르 (*Rhizomucor*), 클루이베로마이세스 (*Kluyveromyces*), 피키아 (*Pichia*), 무코르 (*Mucor*), 마이실리오프토라 (*Myceliophtora*), 페니실리움 (*Penicillium*), 파네로카에테 (*Phanerochaete*), 플레우로투스 (*Pleurotus*), 트라메테스 (*Trametes*), 크리소스포리움 (*Chrysosporium*), 사카로마이세스 (*Saccharomyces*), 스테노트로파모나스 (*Stenotrophamonas*), 스키조사카로마이세스 (*Schizosaccharomyces*), 야로위아 (*Yarrowia*) 또는 스트렙토마이세스 (*Streptomyces*) 속으로부터 선택된다.

[0896] **유기산을 생산할 수 있는 미생물의 생성에서 AD-기능화된 CRISPR 시스템의 용도**

[0897] 본 명세서에 제공된 방법은 유기산을, 보다 특히, 펜토스 또는 헥소스 당으로부터 생산할 수 있는 미생물을 조작하기 위해 추가로 사용된다. 특정 실시형태에서, 상기 방법은 미생물에 외생성 LDH 유전자를 도입하는 단계를 포함한다. 특정 실시형태에서, 상기 미생물에서의 유기산 생산은 관심 유기산 이외의 대사산물을 생성하는 내생성 대사 경로에 관여된 단백질을 코딩하는 내생성 유전자를 불활성화함으로써 추가적으로 또는 대안적으로 증가되고/되거나, 여기서 내생성 대사 경로는 유기산을 소비한다. 특정 실시형태에서, 변형은 관심 유기산 이외의 대사물질의 생성이 감소된다는 것을 보장한다. 특정 실시형태에 따르면, 상기 방법은 적어도 하나의 조작된 유전자 결실 및/또는 유기산이 소비되는 내생성 경로의 불활성화 또는 관심 유기산 이외의 대사물질을 생성하는 내생성 경로에 관여된 생성물을 코딩하는 유전자를 도입하기 위해 사용된다. 특정 실시형태에서, 적어도 하나의 조작된 유전자 결실 또는 불활성화는 피루베이트 데카복실라제 (pdc), 푸마레이트 리덕타제, 알콜 디히드로게나제 (adh), 아세트알데히드 디히드로게나제, 포스포엔올피루베이트 카복실라제(ppc), D-락테이트 디히드로게나제 (d-ldh), L-락테이트 디히드로게나제 (l-ldh), 락테이트 2-모노옥시게나제로 이루어진 군으로부터 선택된 효소를 코딩하는 하나 이상의 유전자에 존재한다. 추가 실시형태에서, 적어도 하나의 조작된 유전자 결실 및/또는 불활성화는 피루베이트 데카복실라제(pdc)를 코딩하는 내생성 유전자에 존재한다.

[0898] 추가 실시형태에서, 미생물은 락트산을 생성하도록 조작되고, 적어도 하나의 조작된 유전자 결실 및/또는 불활성화는 락테이트 디히드로게나제를 코딩하는 내생성 유전자에 존재한다. 추가적으로 또는 대안적으로, 미생물은 시토크롬-의존적 락테이트 디히드로게나제, 예컨대, 시토크롬 B2-의존적 L-락테이트 디히드로게나제를 코딩하는 내생성 유전자의 적어도 하나의 조작된 유전자 결실 또는 불활성화를 포함한다.

[0899] **효모 균주를 이용하여 개선된 자일로스 또는 셀로비오스의 생성에서 AD-기능화된 CRISPR 시스템의 용도**

[0900] 특정 실시형태에서, AD-기능화된 CRISPR 시스템은 효모 균주를 이용하는 개선된 자일로스 또는 셀로비오스를 선택하기 위해 적용될 수 있다. 오류 유발 PCR은 자일로스 이용 또는 셀로비오스 이용 경로에 관여되는 하나 (이상의) 유전자를 증폭시키는 데 사용될 수 있다. 자일로스 이용 경로 및 셀로비오스 이용 경로에 관여하는 유전자의 예는 문헌 [Ha, S.J., et al. (2011) Proc. Natl. Acad. Sci. USA 108(2):504-9] 및 [Galazka, J.M., et al. (2010) Science 330(6000):84-6]에 기재된 것을 포함할 수 있지만, 이들로 제한되지 않는다. 이러한 선택 유전자 내에 무작위 돌연변이를 각각 포함하는, 이중 가닥 DNA 분자의 최종 라이브러리는 효모 균주 (예를 들어, S288C)를 AD-기능화된 CRISPR 시스템의 성분과 공동 형질전환시킬 수 있고, WO2015138855에 기재된 바와 같은 증강된 자일로스 또는 셀로비오스 이용 능력을 갖는 균주가 선택될 수 있다.

[0901] **이소프레노이드 생합성에서 사용을 위한 개선된 효모 균주의 생성에서 AD-기능화된 CRISPR 시스템의 용도**

[0902] Tadas Jakociunas 등은 제빵사 효모 사카로마이세스 세레비지아에서의 한 번의 형질전환 단계에서 5개까지의 상이한 게놈 유전자좌의 게놈 조작으로 산업적으로 중요한 아이소프레노이드 생합성 경로를 위한 중요한 중간체인 높은 메발로네이트 생성을 갖는 균주를 야기하는 다중복합 CRISPR-Cas9 시스템의 성공적인 적용을 기재하였다 (Metabolic Engineering Volume 28, March 2015, Pages 213-222). 특정 실시형태에서, AD-기능화된 CRISPR

시스템은 아이소프레노이드 합성에서 사용하기 위한 추가적인 높은 생성 효모 균주를 동정하기 위해 본 명세서에 기재된 바와 같은 다중복합 게놈 조작 방법에 적용될 수 있다.

[0903] **개선된 식물 및 효모 세포**

[0904] 본 발명은 또한 본 명세서에 제공된 방법에 의해 얻을 수 있고 얻어진 식물 및 효모 세포를 제공한다. 본 명세서에 기재된 방법에 의해 얻어진 개선된 식물은, 예를 들어, 식물 해충, 제초제, 가뭄 또는 저온 또는 고온, 과량의 물 등에 대한 내성을 보장하는 유전자의 발현을 통해 식품 또는 사료 생산에서 유용할 수 있다.

[0905] 본 명세서에 기재된 방법에 의해 수득되는 개선된 식물은, 특히 작물 및 조류이고, 예를 들어, 야생형에서 정상적으로 보이는 것보다 더 높은 단백질, 탄수화물, 영양소 또는 비타민 수준의 발현을 통해 식품 또는 사료 생산에서 유용할 수 있다. 이와 관련하여, 개선된 식물, 특히 두류 및 괴경이 바람직하다.

[0906] 개선된 조류 또는 다른 식물, 예컨대, 유채는, 예를 들어, 식물성 오일 또는 바이오연료, 예컨대, 알콜 (특히 메탄올 및 에탄올)의 생산에서 특히 유용할 수 있다. 이들은 오일 또는 바이오연료 산업에서 사용하기 위한 고수준의 오일 또는 알콜을 발현시키거나 또는 과발현시키도록 조작될 수 있다

[0907] 본 발명은 또한 식물의 개선된 부분을 제공한다. 식물 부분은 잎, 줄기, 이리, 괴경, 종자, 배젖, 밑씨 및 화분을 포함하지만, 이들로 제한되지 않는다. 여기서 고려되는 식물 부분은 자생성, 비자생성, 재생성, 및/또는 비재생성일 수 있다.

[0908] 본 발명의 방법에 따라 생성된 식물 세포 및 식물을 제공하는 것을 또한 본 발명은 고려한다. 전통적인 육종 방법에 의해 생산되는, 유전자 변형을 포함하는 식물의 배우체, 종자, 배아, 접합체 또는 체세포 중 하나, 자손 또는 잡종은 본 발명의 범주 내에 포함된다. 이러한 식물은 표적 서열에 삽입되거나 또는 표적 서열 대신에 삽입되는 이종성 또는 외래 DNA 서열을 함유할 수 있다. 대안적으로, 이러한 식물은 하나 이상의 뉴클레오티드에서 오직 변경 (돌연변이, 결실, 삽입, 치환)을 함유할 수 있다. 이와 같이, 이러한 식물은 그들 조상 식물과 단지 특정 변형의 존재만이 상이할 것이다.

[0909] 따라서, 본 발명은 본 방법에 의해 생산된 식물, 동물 또는 세포, 또는 이의 자손을 제공한다. 자손은 생산된 식물 또는 동물의 클론일 수 있거나, 또는 그들 자손에 바람직한 형질을 더 유전자이입 (introgress)하기 위해 동일 종의 다른 개체와 교배시켜서 유성생식으로부터 얻어질 수 있다. 세포는 다세포 유기체, 특히 동물 또는 식물의 경우에 생체내 또는 생체외에 존재할 수 있다.

[0910] 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템을 사용하여 게놈 편집하는 방법은 본질적으로 임의의 식물, 조류, 진균, 효모 등에 목적하는 형질을 부여하기 위하여 사용될 수 있다. 매우 다양한 식물, 조류, 진균, 효모 등 및 식물 조류, 진균, 효모 세포 또는 조직 시스템이 본 개시의 핵산 구성체 및 상기 언급한 다양한 형질전환 방법을 사용하여 본 명세서에 기술된 바람직한 생리적 및 작물학적 특징을 위해 조작될 수 있다.

[0911] 특정 실시형태에서, 본 명세서에 기재된 방법은 식물 게놈 내 외래 DNA의 존재를 피하기 위해, CRISPR 성분을 코딩하는 것을 포함하는, 임의의 외래 유전자의 식물 게놈, 조류, 진균, 효모 등으로의 영구적인 도입 없이, 내생성 유전자를 변형시키거나 또는 그들의 발현을 변형시키기 위해 사용된다. 이것은 비-유전자이식 식물에 대한 규제 요건이 덜 엄격하기 때문에 흥미로울 수 있다.

[0912] 본 명세서에 기재된 방법은 일반적으로 야생형 식물에 비해 하나 이상의 바람직한 형질을 갖는다는 점에서 "개선된 식물, 조류, 진균, 효모 등"의 생성을 초래한다. 특정 실시형태에서, 비-유전자이식 유전자 변형 식물, 조류, 진균, 효모 등, 부분 또는 세포가 얻어지며, 즉, 외생성 DNA 서열 중 어떤 것도 식물의 임의 세포의 게놈에 혼입되지 않는다. 이러한 실시형태에서, 개선된 식물, 조류, 진균, 효모 등은 비-유전자이식이다. 내생성 유전자의 변형만이 보장되고, 외래 유전자 중 어떤 것도 식물, 조류, 진균, 효모 등의 게놈에 도입되거나 또는 유지되지 않는 경우에, 얻어진 유전자 변형 작물은 외래 유전자를 함유하지 않으며, 따라서 기본적으로 비유전자이식으로 간주될 수 있다. 식물, 조류, 진균, 효모 등의 게놈 편집을 위한 AD-기능화된 CRISPR 시스템의 상이한 적용은 관심 농업 형질을 부여하기 위한 내생성 유전자의 편집을 포함하지만, 이들로 제한되지 않는다. 작물학적 형질을 부여하는 예시적인 유전자는 해충 또는 병해에 내성을 부여하는 유전자; 식물 질환에 관여되는 유전자, 예컨대, WO 2013046247에 열거된 것; 제초제, 살진균제 등에 대한 내성을 부여하는 유전자; (항생제) 스트레스 내성에 관여되는 유전자를 포함하지만, 이들로 제한되지 않는다. CRISPR-Cas 시스템의 사용의 다른 양상은 (웅성) 불임 식물의 생성; 식물/조류 등에서 생식 단계의 증가; 관심 작물에서의 유전자 변형의 생성; 과일-숙성에 대한 영향; 식물/조류 등의 저장 수명 증가; 식물/조류 등에서 알레르겐의 감소; 형질(예를 들어, 영양적 개선)이 더해진 가치의 보장; 관심 내생성 유전자에 대한 스

크리닝 방법; 바이오연료, 지방산, 유기산 등의 생성을 포함하지만, 이들로 제한되지 않는다.

[0913] **AD-기능화된 CRISPR 시스템은 비-인간 유기체에서 사용될 수 있다**

[0914] 일 양상에서, 본 발명은 임의의 기재된 실시형태에 따라 진핵생물 숙주 세포를 포함하는, 비-인간 진핵생물 유기체; 바람직하게는 다세포 진핵생물 유기체를 제공한다. 다른 양상에서, 본 발명은 임의의 기재된 실시형태에 따른 진핵생물 숙주 세포를 포함하는 진핵생물 유기체; 바람직하게는 다세포 진핵생물 유기체를 제공한다.

이들 양상의 일부 실시형태에서 유기체는 동물; 예를 들어, 포유동물일 수 있다. 또한, 유기체는 절지동물, 예컨대, 곤충일 수 있다. 본 발명은 또한 다른 농업 적용분야, 예를 들어, 농장 및 생산 동물로 확대될 수 있다. 예를 들어, 돼지는 그들을 생의학적 모델, 특히 재생 의학으로서 매력있는 많은 특징을 갖는다.

특히, 중증 합병성 면역결핍 (SCID)을 갖는 돼지는 재생의학, 이종장기이식 (또한 본 명세서의 다른 곳에서 논의됨) 및 종양 발생에 유용한 모델을 제공할 수 있으며, 인간 SCID 환자를 위한 요법 개발에 도움을 줄 수 있다. Lee 등 (Proc Natl Acad Sci U S A. 2014 May 20;111(20):7260-5)은 대립유전자 둘 모두에 영향을 미치는 일부를 포함하여, 체세포에서 고효율로 재조합 활성 유전자 (RAG) 2의 표적화된 변형을 생성시키기 위해 리포터-가이드된 전사 활성체-유사 이펙터 뉴클레아제 (TALEN) 시스템을 이용하였다. AD-기능화된 CRISPR 시스템은 유사한 시스템에 적용될 수 있다.

[0915] Lee 등 (Proc Natl Acad Sci U S A. 2014 May 20;111(20):7260-5)의 방법은 다음과 유사하게 본 발명에 적용될 수 있다. 돌연변이된 돼지는 태아 섬유아세포에서 RAG2의 표적화 변형에 후속하여 SCNT 및 배아 전달을 통해 생산된다. CRISPR Cas 및 리포터를 코딩하는 구성체는 태아-유래 섬유아세포에 전기천공된다. 48시간 후에, 녹색 형광 단백질을 발현시키는 형질감염 세포는 웰당 단일 세포의 추정 희석으로 96-웰 플레이트의 개개 웰로 분류된다. RAG2의 표적화 변형은 임의의 CRISPR Cas 절단 부위가 측접된 게놈 DNA 단편을 증폭시킨 후에 PCR 산물을 시퀀싱하여 스크리닝된다. 스크리닝 및 오프-부위 (off-site) 돌연변이 확인 후, RAG2의 표적화 변형을 보유한 세포를 SCNT에 사용한다. 난모세포의 인접한 세포질의 일부분과 함께, 아마도 중기 II 플레이트를 함유하는 극체는 제거되고, 도너 세포가 위란강에 위치된다. 재구축된 배아는 이후에 전기적으로 천공되어 도너 세포를 난모세포와 융합시키고, 화학적으로 활성화된다. 활성화된 배아는 돼지 접합체 배지 3 (PZM3)에서 0.5 μM의 Scriptaid (S7817; Sigma-Aldrich)와 함께 14 내지 16시간 동안 인큐베이션된다.

이어서, 배아는 세척되어 Scriptaid가 제거되고, 그들이 대리 돼지의 수란관에 전달될 때까지 PZM3에서 배양된다.

[0916] 본 발명은 또한 다른 동물, 예컨대, 소의 SNP를 변형시키기 위해 적용가능하다. Tan 등 (Proc Natl Acad Sci U S A. 2013 Oct 8; 110(41): 16526-16531)은 플라스미드, rAAV 및 올리고뉴클레오티드 주형을 이용하여 전사 활성인자-유사(TAL) 이펙터 뉴클레아제(TALEN)- 및 주기적 간격으로 분포하는, 짧은 회문구조 반복부(CRISPR)/Cas9-자극된 상동성-지정 복구 (HDR)를 포함하도록 가축 유전자 편집 툴박스 (toolbox)를 확장시켰다.

유전자 특이적 가이드 RNA 서열은 이들 방법에 따라 Church lab 가이드 RNA 벡터 (Addgene ID: 41824)로 클로닝되었다 (Mali P, et al. (2013) RNA-Guided Human Genome Engineering via Cas9. Science 339(6121):823-826). Cas9 뉴클레아제는 RCIScript-hCas9로부터 합성된 mRNAh 또는 Cas9 플라스미드 (Addgene ID: 41815)의 공동형질감염에 의해 제공되었다. 이러한 RCIScript-hCas9는 hCas9 플라스미드 (hCas9 cDNA를 포함)로부터의 XbaI-AgeI 단편을 RCIScript 플라스미드로 서브클로닝하여 구축되었다.

[0917] Heo 등 (Stem Cells Dev. 2015 Feb 1;24(3):393-402. doi: 10.1089/scd.2014.0278. Epub 2014 Nov 3)은 소 다능성 세포 및 주기적 간격으로 분포하는, 짧은 회문구조 반복부(CRISPR)/Cas9 뉴클레아제를 이용하는 소 게놈에서의 고도로 효율적인 유전자 표적화를 보고하였다. 첫째로, Heo 등은 야마나카 인자의 이소성 발현 및 GSK3β 및 MEK 억제제(2i) 처리에 의해 소 체세포 섬유아세포로부터 유도만능줄기세포(iPSC)를 생성한다. Heo 등은 이들 소 iPSC가 기형종에서의 유전자 발현 및 발생 잠재력에 관해 미경험 다능성 줄기 세포와 고도로 유사하다는 것을 관찰하였다. 게다가, 소 NANOG 유전자좌에 특이적인 CRISPR-Cas9 뉴클레아제는 소 iPSC 및 배아에서 소 게놈의 고도로 효율적인 편집을 나타내었다.

[0918] Igenity®는 소와 같은 동물의 프로파일 분석을 제공하여, 경제적으로 중요한 경제적 특성들의 특성, 예컨대 도체 조성, 도체 품질, 모계 및 번식 특성, 및 평균 일일 체중증가를 수행 및 전송한다. 종합적 Igenity® 프로파일의 분석은 DNA 마커 (가장 흔하게는 단일 뉴클레오티드 다형성 또는 SNP)에서 시작된다. Igenity® 프로파일 배후의 모든 마커는 대학교, 연구 기관 및 USDA와 같은 정부 기관을 포함한, 연구 시설에서 독립 과학자에 의해 발굴되었다. 이어서, 마커는 검증 개체군에서 Igenity®로 분석된다. Igenity®는, 흔하게 이용가능하지 않은 표현형을 수집하기 위해 종축, 소-송아지, 가축사육장 및/또는 소고기 산업 포장 부문의 산

업 파트너와 종종 함께 작업하여, 다양한 생산 환경 및 생물학적 유형을 대표하는 다수 자원 집단을 이용한다.

소 게놈 데이터베이스는 널리 이용 가능하며, 예를 들어, NAGRP 소 게놈 협응 프로그램 (Cattle Genome Coordination Program) (http://www.animalgenome.org/cattle/maps/db.html)을 참조한다. 따라서, 본 발명은 소 SNP를 표적화하는 데 적용될 수 있다. 당업자는 SNP를 표적화하기 위해 상기 프로토콜을 이용할 수 있으며, 그들에, 예를 들어, Tan 등 또는 Heo 등이 기재한 바와 같이 소 SNP를 적용할 수 있다.

[0919] Qingjian Zou 등 (Journal of Molecular Cell Biology Advance Access published October 12, 2015)은 개 미오스타틴 (MSTN) 유전자 (골격근량의 음성 조절자)의 제1 엑손을 표적화함으로써 개에서의 증가된 근육량을 입증하였다. 먼저, sgRNA의 효율은 개 배아 섬유아세포 (CEF)로 Cas9 벡터와 MSTN 표적화 sgRNA의 공동형질감염을 이용하여, 입증되었다. 그 후, MSTN KO 개는 Cas9 mRNA 및 MSTN sgRNA의 혼합물과 함께 정상 형태학을 갖는 배아를 미세주입 및 동일한 암컷 개의 난관 내로 접합자의 자가이식에 의해 생성되었다. 녹아웃 강아지는 야생형 한배새끼 자매와 비교하여 넓적다리 상에 분명한 근육 표현형을 나타내었다. 이는 또한 본 명세서에 제공된 AD-기능화된 CRISPR 시스템을 이용하여 수행될 수 있다.

[0920] **가축 - 돼지**

[0921] 가축에서 바이러스 표적은, 일부 실시형태에서, 예를 들어 돼지 마크로파지상의 돼지 CD163을 포함할 수 있다.

CD163은 PRRSv (돼지 생식기 호흡기 증후군 바이러스 (Porcine Reproductive and Respiratory Syndrome virus, an arterivirus), 아테리바이러스)에 의한 감염과 관련된다 (바이러스 세포 진입을 통한 것으로 여겨짐). 특히 돼지 폐포 마크로파지 (폐에 존재)의 PRRSv에 의한 감염은 사육 돼지에서의 번식 장애, 체중 손실 및 높은 폐사율을 포함하여, 고통을 초래하는 사전 치유불가의 돼지 증후군 ("미확인 돼지 질환" 또는 "청색귀 질환")을 야기한다. 기회감염, 예컨대, 유행성 폐렴, 뇌수막염 및 귀 부종은 종종 마크로파지 활성 상실을 통한 면역 결핍에 기인하는 것으로 보인다. 또한 증가된 항생제 사용 및 재정 손실 (연간 $660m로 추정)로 인해 상당한 경제적 및 환경적 영향을 받는다.

[0922] Genus Plc와 협업하여 University of Missouri에서 Kristin M Whitworth 및 Dr Randall Prather (Nature Biotech 3434, 2015년 12월 7일에 온라인 공개됨)가 보고한 바와 같이, CD163은 CRISPR-Cas9를 이용하여 표적화되었고 편집된 돼지의 새끼는 PRRSv에 노출되었을 때 내성이 있었다. 모두 CD163의 엑손 7에서 돌연변이를 갖는 한 마리의 파운더 (founder) 수컷과 한 마리의 파운더 암컷을 새끼를 생산하도록 육종되었다. 파운더 수컷은 하나의 대립유전자 상의 엑손 7에서 11-bp 결실을 가졌는데, 이는 프레임 시프트 돌연변이 및 도메인 5 내 아미노산 45에서의 미스센스 번역 및 아미노산 64에서 후속적인 조기 정지 코돈을 초래한다. 다른 대립유전자는 선행 인트론에서의 엑손 7 및 377-bp 결실에서 2bp 첨가를 갖는데, 이는 도메인 5의 처음 49개 아미노산의 발현을 초래하고, 그 다음에 아미노산 85에서 조기 정지 코돈을 초래하는 것으로 예측되었다. 암퇘지는 번역되었을 때 도메인 5의 처음 48개 아미노산, 아미노산 70에서 조기 정지 코돈을 발현시키는 하나의 대립유전자에서 7 bp 첨가를 갖는다. 암퇘지의 다른 대립유전자는 증폭불가하였다. 선택된 새끼는 무효 (null) 동물(CD163-/-), 즉, CD163 녹아웃인 것으로 예측되었다.

[0923] 따라서, 일부 실시형태에서, 돼지 폐포 마크로파지는 CRISPR 단백질에 의해 표적화될 수 있다. 일부 실시형태에서, 돼지 CD163은 CRISPR 단백질에 의해 표적화될 수 있다. 일부 실시형태에서, 돼지 CD163은 DSB의 유도를 통해, 또는 삽입 또는 결실, 예를 들어, 상기 기재한 것 중 하나 이상을 포함하는, 엑손 7의 표적화 결실 또는 변형, 또는 유전자의 다른 영역, 예를 들어 엑손 5의 결실 또는 변형을 통해 녹아웃될 수 있다.

[0924] 편집된 돼지 및 그의 자손, 예를 들어 CD163 녹아웃 돼지가 또한 고려된다. 이는 가축, 육종 또는 모델링 목적 (즉, 돼지 모델)을 위한 것일 수 있다. 유전자 녹아웃을 포함하는 정액이 또한 제공된다.

[0925] CD163은 스캐빈저 수용체 시스테인-풍부 (SRCR) 슈퍼패밀리의 구성원이다. 단백질의 SRCR 도메인 5 시험관내 연구는 바이러스 게놈의 패키징 해체 (unpackagin) 및 방출을 초래하는 도메인에 기반한다. 이와 같이, SRCR 슈퍼패밀리의 다른 구성원은 또한 다른 바이러스에 대한 내성을 평가하기 위해 표적화될 수 있다. PRRSV는 또한 또한 쥐과 락테이트 디히드로게나제-상승 바이러스, 유인원 출혈열 바이러스 및 말 동맥염 바이러스를 포함하는 포유동물 아테리바이러스 그룹의 구성원이다. 아테리바이러스는 마크로파지 향성 및 중증 질환 및 지속적 감염을 유발하는 능력을 포함하는, 중요한 병원성 특성을 공유한다. 따라서, 아테리바이러스, 및 특히 쥐과 락테이트 디히드로게나제-상승 바이러스, 유인원 출혈열 바이러스 및 말 동맥염 바이러스는, 예를 들어, 돼지 CD163 또는 다른 종의 이의 상동체를 통해 표적화될 수 있고, 및 쥐과, 유인원 및 말 모델 및 녹아웃이 또한 제공된다.

[0926] 사실, 이 접근은 인간에 전염될 수 있는 다른 가축 질환을 야기하는 바이러스 또는 박테리아, 예컨대, 인플루엔자 C 및 H1N1, H1N2, H2N1, H3N1, H3N2 및 H2N3으로서 알려진 인플루엔자 A의 아형을 포함하는 돼지 인플루엔자 바이러스(SIV) 균주뿐만 아니라 상기 언급된 폐렴, 뇌수막염 및 부종에까지 확대될 수 있다.

[0927] 일부 실시형태에서, 본 명세서에 기재된 AD-기능화된 CRISPR 시스템은 돼지-대-인간 이종장기이식의 임상 적용을 용이하게 하기 위해 하나 이상의 돼지 내생성 레트로바이러스(PERV)를 불활성화시키도록 돼지 게놈을 유전자 변형시키는데 사용될 수 있다. 본 명세서에 전문이 참고로 편입된 문헌 [Yang et al., Science 350(6264):1101-1104 (2015)]을 참조한다. 일부 실시형태에서, 본 명세서에 기재된 AD-기능화된 CRISPR 시스템은 임의의 활성 돼지 내생성 레트로바이러스 (PERV) 유전자좌를 포함하지 않는 유전자 변형 돼지를 생산하는데 사용될 수 있다.

[0928] **AD-기능화된 CRISPR 시스템에 의한 치료적 표적화**

[0929] 명백한 바와 같이, AD-기능화된 CRISPR 시스템은 관심 임의의 폴리뉴클레오티드 서열을 표적화하기 위해 사용될 수 있다는 것이 예상된다. 본 발명은 비-천연 발생 또는 조작된 조성물, 또는 상기 조성물의 성분을 코딩하는 하나 이상의 폴리뉴클레오티드, 또는 생체내, 생체외 또는 시험관내 표적 세포를 변형하는 데 사용하기 위한 상기 조성물 성분을 코딩하는 하나 이상의 폴리뉴클레오티드의 벡터 또는 전달 시스템을 제공하고, 일단 변형되면, CRISPR 변형된 세포의 자손 또는 세포주는 변형된 표현형을 보유하도록 세포를 변경시키는 방식으로 수행될 수 있다. 변형된 세포 및 자손은 다중 세포 유기체, 예컨대, 목적하는 세포 유형에 대한 CRISPR 시스템의 생체외 또는 생체내 적용을 갖는 식물 또는 동물의 일부분일 수 있다. CRISPR 발명은 치료적 처치 방법일 수 있다. 치료적 처치 방법은 유전자 또는 게놈 편집, 또는 유전자 요법을 포함할 수 있다.

[0930] **양자 세포 요법**

[0931] 본 발명은 또한 양자 요법을 위해 세포를 변형시키는 본 명세서에 기재된 AD-기능화된 CRISPR 시스템의 사용을 고려한다. 따라서 본 발명의 양상은 선택된 항원, 예컨대, 종양 연관 항원에 특이적인 면역 시스템 세포, 예컨대, T 세포의 양자 전달을 수반한다 (하기의 문헌들 참조: Maus et al., 2014, Adoptive Immunotherapy for Cancer 또는 Viruses, Annual Review of Immunology, Vol. 32: 189-225; Rosenberg and Restifo, 2015, Adoptive cell transfer as personalized immunotherapy for human cancer, Science Vol. 348 no. 6230 pp. 62-68; 및, Restifo et al., 2015, Adoptive immunotherapy for cancer: harnessing the T cell response. Nat. Rev. Immunol. 12(4): 269-281; 및 Jenson and Riddell, 2014, Design and implementation of adoptive therapy with chimeric antigen receptor-modified T cells. Immunol Rev. 257(1): 127-144). 예를 들어, T 세포 수용체 (TCR)의 특이성을 변경시킴으로써, 예를 들어 선택된 펩티드 특이성을 갖는 새로운 TCR α 및 β 쇄를 도입함으로써 T 세포를 유전자 변형시키는 다양한 전략이 사용될 수 있다 (미국 특허 제8,697,854호; 국제 특허 출원 공개: WO2003020763, WO2004033685, WO2004044004, WO2005114215, WO2006000830, WO2008038002, WO2008039818, WO2004074322, WO2005113595, WO2006125962, WO2013166321, WO2013039889, WO2014018863, WO2014083173; 미국 특허 제8,088,379호 참조).

[0932] TCR 변형에 대한 대안으로서 또는 추가로, 기술되었던 매우 다양한 수용체 키메라 구성체를 사용하여, 악성 세포와 같은, 선택 표적에 특이적인 면역반응성 세포, 예컨대, T 세포를 생성하기 위해 키메라 항원 수용체 (CAR)가 사용될 수 있다 (미국 특허 제5,843,728호; 제5,851,828호; 5,912,170호; 제6,004,811호; 제6,284,240호; 제6,392,013호; 제6,410,014호; 제6,753,162호; 제8,211,422호; 및 국제 특허 출원 공개 WO9215322). 대안적인 CAR 구성체는 연속 세대에 속하는 것으로 특징규명될 수 있다. 제1-세대 CAR은 전형적으로 항원에 특이적인 항체의 단일쇄 가변 단편으로 이루어지는데, 예를 들어 가요성 링커에 의해, 예를 들어, CD8α 힌지 도메인 및 CD8α 막관통 도메인에 의해, CD3ζ 또는 FcRγ 중 하나의 막관통 및 세포내 신호전달 도메인에 연결되는 특정 항체의 VH에 연결된 VL을 포함한다 (scFv-CD3ζ 또는 scFv-FcRγ; 미국 특허 제7,741,465호; 미국 특허 제5,912,172호; 미국 특허 제5,906,936호 참조). 제2-세대 CAR은 엔도도메인 내에 하나 이상의 공자극 분자, 예컨대, CD28, OX40 (CD134), 또는 4-1BB(CD137)의 세포내 도메인을 혼입시킨다 (예를 들어 scFv-CD28/OX40/4-1BB-CD3ζ; 미국 특허 제8,911,993호; 제8,916,381호; 제8,975,071호; 제9,101,584호; 제9,102,760호; 제9,102,761호 참조). 제3-세대 CAR은 공자극 엔도도메인, 예컨대, CD3ζ-쇄, CD97, GDI la-CD18, CD2, ICOS, CD27, CD154, CDS, OX40, 4-1BB 또는 CD28 신호전달 도메인의 조합을 포함한다 (예를 들어 scFv-CD28-4-1BB-CD3ζ 또는 scFv-CD28-OX40-CD3ζ; 미국 특허 제8,906,682호; 미국 특허 제8,399,645호; 미국 특허 제5,686,281호; 국제 특허 출원 WO2014134165; 국제 특허 출원 WO2012079000 참조). 대안적으로, 공자극은 그들의 천연 αβTCR의 맞물림 후, 예를 들어 전문적 항원-제시 세포 상에서 항원에 의해, 수반하는 공

자극으로 활성화되고 확장될 수 있도록 선택되는 항원-특이적 T 세포에서 CAR을 발현시킴으로써 조직될 수 있다. 추가로, 추가의 조작된 수용체가 예를 들어 T-세포 공격의 표적화를 개선시키기 위해 그리고/또는 부작용을 최소화하기 위해 면역반응성 세포 상에 제공될 수 있다.

[0933] 표적 면역반응성 세포를 형질전환시키기 위해 대안의 기법, 예컨대, 원형질체 융합, 리포펙션, 형질감염 또는 전기천공법이 사용될 수 있다. 매우 다양한 벡터, 예컨대, 레트로바가러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노-연합 바이러스 벡터, 플라스미드 또는 트랜스포존, 예컨대, 슬리핑 뷰티 (Sleeping Beauty) 트랜스포존이 사용될 수 있고 (미국 특허 제6,489,458호; 제7,148,203호; 제7,160,682호; 제7,985,739호; 제8,227,432호 참조), 예를 들어 CD3ζ 및 CD28 또는 CD137 중 하나를 통해 2세대 항원-특이적 CAR 신호전달을 이용하여 CAR을 도입시키는데 사용될 수 있다. 바이러스 벡터는, 예를 들어, HIV, SV40, EBV, HSV 또는 BPV에 기반한 벡터를 포함할 수 있다.

[0934] 형질전환을 위해 표적화된 세포는, 예를 들어 T 세포, 자연살해 (NK) 세포, 세포독성 T 림프구(CTL), 조절 T 세포, 인간 배아 줄기 세포, 종양-침윤성 림프구(TIL) 또는 림구 세포가 분화될 수 있는 다능성 줄기 세포를 포함할 수 있다. 목적하는 CAR을 발현시키는 T 세포는, 예를 들어, 암 항원 및 공자극 분자를 공동발현시키는 γ-조사된 활성화 및 증식 세포 (AaPC)와 함께 공배양을 통해 선택될 수 있다. 조작된 CAR T-세포는 가용성 인자, 예컨대, IL-2 및 IL-21의 존재 하에 AaPC에 대한 공동 배양에 의해, 확장될 수 있다. 이러한 확장은, 예를 들어, 기억 CAR+ T 세포를 제공하기 위해 수행될 수 있다 (예를 들어, 비효소적 디지털 어레이 및/또는 다중-패널 유세포 분석기에 의해 평가될 수 있음). 이러한 방식으로, 항원-보유 종양에 대해 (선택적으로 목적하는 케모카인, 예컨대, 인터페론-γ의 생성과 함께) 특이적 세포독성 활성을 갖는 CAR T 세포가 제공될 수 있다. 이런 종류의 CAR T 세포는, 예를 들어, 동물 모델에서, 예를 들어 종양 이종이식을 위협하기 위해 사용될 수 있다.

[0935] 앞서 언급한 것과 같은 접근은 선택된 항원에 결합하는 수용체를 인식하는 항원을 포함하는 유효량의 면역반응성 세포를 투여함으로써, 질환, 예컨대, 신생물을 갖는 대상체를 치료하고/하거나 생존을 증가시키는 방법을 제공하는 데 적합할 수 있으며, 여기서 결합은 면역반응성 세포를 활성화시켜서, 질환 (예컨대, 신생물, 병원균 감염, 자가면역 장애 또는 동종이계 이식 반응)을 치료하거나 또는 예방한다. CAR T 세포 요법에서 용량은, 예를 들어, 림프구고갈 과정의 존재 또는 부재로, 예를 들어 시클로포스파미드 존재에서, 10^6 내지 10^9개의 세포/kg의 투여를 수반할 수 있다.

[0936] 일 실시형태에서, 치료는 면역억제 치료를 겪는 환자에게 투여될 수 있다. 세포 또는 세포 개체군은 이러한 면역억제제에 대한 수용체를 코딩하는 유전자의 불활성화에 기인하여 적어도 하나의 면역억제제에 대해 내성될 수 있다. 이론에 국한하지 않지만, 면역억제 치료는 환자 내에서 본 발명에 따른 면역억제 또는 T 세포의 선택 및 확장에 도움을 주어야 한다.

[0937] 본 발명에 따른 세포 또는 세포 개체군의 투여는 에어로졸 흡입, 주사, 섭취, 수혈, 이식 주입 또는 이식을 비롯한 임의의 편리한 방법으로 수행될 수 있다. 세포 또는 세포 개체군은 환자에게 피하로, 피내로, 종양내로, 낭내로, 척수내로, 근육내로, 정맥내로 또는 림프내 주사로 또는 복강내로 투여될 수 있다. 일 실시형태에서, 본 발명의 세포 조성물은 바람직하게는 정맥내 주사에 의해 투여된다.

[0938] 세포 또는 세포 개체군의 투여는 10^4 내지 10^9 세포/kg 체중, 바람직하게는 10^5 내지 10^6 세포/kg 체중 (해당 범위 내에서 세포의 모든 정수 값을 포함함)의 투여로 이루어질 수 있다. CAR T 세포 요법에서 용량은, 예를 들어 림프구제거 과정 존재 또는 부재로, 예를 들어 시클로포스파미드와 함께, 10^6 내지 10^9 세포/kg의 투여를 수반할 수 있다. 세포 또는 세포 개체군은 1회 이상의 용량으로 투여될 수 있다. 다른 실시형태에서, 유효량의 세포는 단일 용량으로서 투여된다. 다른 실시형태에서, 유효량의 세포는 일정 기간에 걸쳐 1회 초과로서 투여된다. 투여 시기는 담당 의사의 판단 내이고, 환자의 임상적 병태에 의존한다. 세포 또는 세포 개체군은 임의의 공급원, 예컨대, 혈액 은행 또는 공혈자로부터 얻을 수 있다. 각각의 요구가 다르지만, 특정 질환 또는 병태에 대해 유효량의 주어진 세포 유형의 최적 범위의 결정은 당업자의 기술 내이다. 유효량은 치료적 또는 예방적 이득을 제공하는 양을 의미한다. 투여되는 용량은 연령, 수용자의 건강상태 및 체중, 있다면 병용 치료의 종류, 치료 빈도 및 목적하는 효과의 특성에 따를 것이다.

[0939] 다른 실시형태에서, 해당 세포를 포함하는 유효량의 세포 또는 조성물은 비경구로 투여된다. 투여는 정맥내 투여일 수 있다. 투여는 종양 내에서 주사에 의해 직접적으로 행해질 수 있다.

[0940] 가능한 유해 반응에 대해 보호하기 위해, 조작된 면역반응성 세포는 특정 신호에 대한 노출에 취약한 세포를 제공하는 이식유전자의 형태로 유전자이식 안전성 스위치를 구비할 수 있다. 예를 들어, 단순 포진 바이러스 티미딘 키나제 (TK) 유전자가 이 방법에서, 줄기 세포 이식 후 도너 림프구 주입에 사용되는 동종이계 T 림프구로의 도입에 의해 사용될 수 있다 (Greco, et al., Improving the safety of cell therapy with the TK-suicide gene. Front. Pharmacol. 2015; 6: 95). 이러한 세포에서, 뉴클레오시드 프로드러그, 예컨대, 간시클로비어 또는 아시클로비어의 투여는 세포 사멸을 야기한다. 대안의 안전성 스위치 구성체는, 예를 들어 활성 효소를 형성하는 2개의 비기능성 icasp9 분자와 함께 소형 분자 이량체의 투여에 의해 촉발되는 유도성 카파제 9를 포함한다. 매우 다양한 세포 증식 제어 실행에 대한 대안의 접근법이 기술되어 있다 (미국 공개 특허 출원 제20130071414호; 국제 특허 출원 공개 번호 WO2011146862; 국제 특허 출원 공개 번호 WO2014011987; 국제 특허 출원 공개 번호 WO2013040371; 문헌 [Zhou et al. BLOOD, 2014, 123/25:3895 - 3905]; [Di Stasi et al., The New England Journal of Medicine 2011; 365:1673-1683]; [Sadelain M, The New England Journal of Medicine 2011; 365:1735-173]; [Ramos et al., Stem Cells 28(6):1107-15 (2010)] 참조).

[0941] 양자 요법의 추가적인 개선에서, 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR-Cas 시스템에 의한 게놈 편집은, 예를 들어 편집 CAR T 세포를 제공하는 대안적인 실행을 위해 면역반응성 세포를 재단하는데 사용될 수 있다 (Poirot et al., 2015, Multiplex genome edited T-cell manufacturing platform for "off-the-shelf" adoptive T-cell immunotherapies, Cancer Res 75 (18): 3853). 예를 들어, 면역반응성 세포는 HLA II형 및/또는 I형 분자 부류의 일부 또는 모두의 발현을 결실시키도록 또는 목적하는 면역 반응, 예컨대, PD1 유전자를 억제할 수 있는 선택 유전자를 녹아웃시키도록 편집될 수 있다.

[0942] 세포는 본 명세서에 기재된 바와 같은 AD-기능화된 CRISPR 시스템을 이용하여 편집될 수 있다. AD-기능화된 CRISPR 시스템은 임의의 본 명세서에 기재된 방법에 의해 면역 세포에 전달될 수 있다. 바람직한 실시형태에서, 세포는 생체외에서 편집되고, 이를 필요로 하는 대상체에게 전달된다. 면역 반응성 세포 세포, CAR-T 세포 또는 양자 세포 전달을 위해 사용되는 임의의 세포는 편집될 수 있다. 편집은 잠재적인 동종반응성 T-세포 수용체 (TCR)를 제거하기 위해, 화학치료제의 표적을 붕괴시키기 위해, 면역 체크포인트를 차단하기 위해, T 세포를 활성화시키기 위해 및/또는 기능적으로 감손되거나 또는 기능장애 CD8+ T-세포의 분화 및/또는 증식을 증가시키기 위해 수행될 수 있다 (국제 특허 출원 공개: WO2013176915, WO2014059173, WO2014172606, WO2014184744 및 WO2014191128). 편집은 유전자의 불활성화를 초래할 수 있다.

[0943] T 세포 수용체 (TCR)는 항원 제시에 반응하는 T 세포의 활성화에 참여하는 세포 표면 수용체이다. TCR은 일반적으로, 이형이량체를 형성하도록 조립되고 세포 표면 상에 제시되는 T 세포 수용체 복합체를 형성하기 위해 CD3-형질도입 서브유닛과 회합하는 2개의 쇄, 즉, α 및 β로 이루어진다. TCR의 각각의 α 및 β 쇄는 면역글로불린-유사 N-말단의 가변 (V) 및 불변 (C) 영역, 소수성 막관통 도메인 및 짧은 세포질 영역으로 이루어진다. 면역글로불린 분자에 대해서와 같이, α 및 β 쇄의 가변 영역은 V(D)J 재조합에 의해 생성되어, T 세포 개체군 내에서 매우 다양한 항원 특이성을 생성시킨다. 그러나, 온전한 항원을 인식하는 면역글로불린과 대조적으로, T 세포는 MHC 분자와 회합되는 가공된 펩티드 단편에 의해 활성화되어, MHC 제한으로서 알려진 T 세포에 의한 항원 인식에 대해 여분의 차원을 도입한다. T 세포 수용체를 통한 도너와 수용자 사이의 MHC의 인식 차이는 T 세포 증식 및 이식편대 숙주반응 (GVHD)의 잠재적 발생을 야기한다. TCR α 또는 TCR β의 불활성화는 T 세포 표면으로부터 TCR의 제거를 초래하여 동종항원의 인식을 방지하고 그에 따라 GVHD를 방지할 수 있다. 그러나, TCR 붕괴는 일반적으로 CD3 신호전달 성분의 제거를 초래하고 추가적인 T 세포 확장 수단을 변경시킨다.

[0944] 동종이계 세포는 숙주 면역계에 의해 빠르게 거부된다. 비-방사선조사 혈액 생성물에 존재하는 동종이계 백혈구는 5 내지 6일 이하 동안 지속된다는 것이 입증되었다 (Boni, Muranski et al. 2008 Blood 1;112(12):4746-54). 따라서, 동종이계 세포의 거부를 방지하기 위해, 숙주의 면역계는 일반적으로 어느 정도까지는 억제되어야 한다. 그러나, 양자 세포 전달의 경우에, 면역억제 약물의 사용이 또한 도입된 치료적 T 세포에 대해 해로운 효과를 갖는다. 그러므로, 이들 조건에서 양자 면역요법 접근을 효과적으로 사용하기 위해, 도입된 세포는 면역억제 치료에 내성일 필요가 있을 것이다. 따라서, 특정 실시형태에서, 본 발명은, 바람직하게는 면억억제제에 대한 표적을 코딩하는 적어도 하나의 유전자의 불활성화에 의해, T 세포를 면역억제제에 내성이 되도록 변형시키는 단계를 추가로 포함한다. 면역억제제는 몇몇 작용 메커니즘 중 하나에 의해 면역 기능을 억제하는 작용제이다. 면역억제제는 칼시뉴린 억제제, 라파마이신의 표적, 인터류킨-2 수용체 α-쇄 차단제, 이노신 일인산 디히드로게나제의 억제제, 다이하이드로엽산 리덕타제의 억제제, 코티코스테로이

드 또는 면역억제성 항대사물질일 수 있지만, 이들로 제한되지 않는다. 본 발명은 T 세포 내 면역억제제의 표적을 불활성화시킴으로써 면역요법을 위해 T 세포에 면역억제 내성을 부여하도록 허용한다. 비제한적 예로서, 면역억제제에 대한 표적은 면역억제제, 예컨대: CD52, 글루코코티코이드 수용체(GR), FKBP 패밀리 유전자 구성원 및 시클로필린 패밀리 유전자 구성원에 대한 수용체일 수 있다.

[0945] 면역 체크포인트는 면역반응을 늦추거나 중단시키고, 면역 세포의 제어되지 않은 활성으로부터의 과도한 조직 손상을 방지하는 억제 경로이다. 일정 실시형태에서, 표적화된 면역 체크포인트는 프로그램된 사멸-1 (PD-1 또는 CD279) 유전자 (PDCD1)이다. 다른 실시형태에서, 표적화된 면역 체크포인트는 세포독성 T-림프구-연관 항원 (CTLA-4)이다. 추가적인 실시형태에서, 표적화된 면역 체크포인트는 CD28 및 CTLA4 Ig 슈퍼패밀리의 다른 구성원, 예컨대, BTLA, LAG3, ICOS, PDL1 또는 KIR이다. 추가적인 실시형태에서, 표적화된 면역 체크포인트는 TNFR 슈퍼패밀리의 구성원, 예컨대, CD40, OX40, CD137, GITR, CD27 또는 TIM-3이다.

[0946] 추가적인 면역 체크포인트는 Src 상동성 2 도메인-함유 단백질 티로신 포스파타제 1 (SHP-1)을 포함한다 (Watson HA, et al., SHP-1: the next checkpoint target for cancer immunotherapy? Biochem Soc Trans. 2016 Apr 15;44(2):356-62). SHP-1은 널리 발현되는 억제성 단백질 티로신 포스파타제 (PTP)이다. T-세포에서, 이는 항원-의존적 활성화 및 증식의 음성 조절자이다. 이는 시토졸 단백질이고, 따라서 항체-매개 요법을 받을 수 없지만, 활성화 및 증식에서 이의 역할은 양자 전달 전략, 예컨대, 키메라 항원 수용체 (CAR) T 세포에서 유전자 조작을 위한 매력적인 표적이 되게 한다. 면역 체크포인트는 또한 Ig 및 ITIM 도메인 (TIGIT/Vstm3/WUCAM/VSIG9) 및 VISTA를 갖는 T 세포 면역수용체를 포함할 수 있다 (Le Mercier I, et al., (2015) Beyond CTLA-4 and PD-1, the generation Z of negative checkpoint regulators. Front. Immunol. 6:418).

[0947] WO2014172606은 고갈된 CD8+ T-세포의 증식 및/또는 활성을 증가시키기 위해서, 그리고 CD8+ T-세포 고갈을 감소 (예를 들어, 기능적으로 고갈되거나 또는 비반응성인 CD8+ 면역 세포를 감소)시키기 위해 MT1 및/또는 MT1 억제제의 사용에 관한 것이다. 일정 실시형태에서, 메탈로티오네인은 양자 전달된 T 세포 내 유전자 편집에 의해 표적화된다.

[0948] 일정 실시형태에서, 유전자 편집의 표적은 면역 체크포인트 단백질의 발현에 관여되는 적어도 하나의 표적화된 유전자좌일 수 있다. 이러한 표적은 CTLA4, PPP2CA, PPP2CB, PTPN6, PTPN22, PDCD1, ICOS (CD278), PDL1, KIR, LAG3, HAVCR2, BTLA, CD160, TIGIT, CD96, CRTAM, LAIR1, SIGLEC7, SIGLEC9, CD244 (2B4), TNFRSF10B, TNFRSF10A, CASP8, CASP10, CASP3, CASP6, CASP7, FADD, FAS, TGFBRII, TGFRBRI, SMAD2, SMAD3, SMAD4, SMAD10, SKI, SKIL, TGIF1, IL10RA, IL10RB, HMOX2, IL6R, IL6ST, EIF2AK4, CSK, PAG1, SIT1, FOXP3, PRDM1, BATF, VISTA, GUCY1A2, GUCY1A3, GUCY1B2, GUCY1B3, MT1, MT2, CD40, OX40, CD137, GITR, CD27, SHP-1 또는 TIM-3을 포함할 수 있지만, 이들로 제한되지 않는다. 바람직한 실시형태에서, PD-1 또는 CTLA-4 유전자의 발현에 관여되는 유전자의 좌위가 표적화된다. 다른 바람직한 실시형태에서, 유전자의 조합은, 예컨대, PD-1 및 TIGIT로 표적화되지만, 이들로 제한되지 않는다.

[0949] 다른 실시형태에서, 적어도 2개의 유전자가 편집된다. 유전자의 쌍은 PD1과 TCRα, PD1과 TCRβ, CTLA-4와 TCRα, CTLA-4와 TCRβ, LAG3과 TCRα, LAG3과 TCRβ, Tim3과 TCRα, Tim3과 TCRβ, BTLA와 TCRα, BTLA와 TCRβ, BY55와 TCRα, BY55와 TCRβ, TIGIT와 TCRα, TIGIT와 TCRβ, B7H5와 TCRα, B7H5와 TCRβ, LAIR1과 TCRα, LAIR1과 TCRβ, SIGLEC10과 TCRα, SIGLEC10과 TCRβ, 2B4와 TCRα, 2B4와 TCRβ를 포함할 수 있지만, 이들로 제한되지 않는다.

[0950] T 세포의 유전자 변형 이전이건 또는 이후이건 간에, T 세포는 일반적으로 예를 들어, 미국 특허 제6,352,694호; 제6,534,055호; 제6,905,680호; 제5,858,358호; 제6,887,466호; 제6,905,681호; 제7,144,575호; 제7,232,566호; 제7,175,843; 5,883,223호; 제6,905,874호; 제6,797,514호; 제6,867,041호; 및 제7,572,631호에 기재된 바와 같은 방법을 이용하여 활성화 및 확장될 수 있다. T 세포는 시험관내 또는 생체내에서 확장될 수 있다.

[0951] 본 발명의 실행은 당업계의 기술 내인 면역학, 생화학, 화학, 분자 생물학, 미생물학, 세포 생물학, 게놈 및 재조합 DNA 분야에 공지된 기법을 사용한다. 문헌 [MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989) (Sambrook, Fritsch and Maniatis); MOLECULAR CLONING: A LABORATORY MANUAL, 4th edition (2012) (Green and Sambrook); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (1987) (F. M. Ausubel, et al. eds.); 시리즈 METHODS IN ENZYMOLOGY (Academic Press, Inc.); PCR 2: A PRACTICAL APPROACH (1995) (M.J. MacPherson, B.D. Hames and G.R. Taylor eds.); ANTIBODIES, A LABORATORY MANUAL (1988) (Harlow and Lane, eds.);

ANTIBODIES A LABORATORY MANUAL, 2nd edition (2013) (E.A. Greenfield ed.); 및 ANIMAL CELL CULTURE (1987) (R.I. Freshney, ed.)]을 참조한다.

[0952] **질환-연관 돌연변이 및 병원성 SNP의 교정**

[0953] 일 양상에서, 본 명세서에 기재된 발명은 G>A 또는 C>T 점 돌연변이 또는 병원성 단일 뉴클레오티드 다형성(SNP)에 의해 야기되거나 또는 야기될 가능성이 있는 질환 병태를 치료 및/또는 예방할 목적으로 표적 유전자좌에서 아데노신 잔기를 변형시키는 방법을 제공한다.

[0954] **뇌 및 중추 신경계에 영향을 미치는 질환**

[0955] 뇌 및 중추 신경계에 영향을 미치는 다양한 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 ClinVar 데이터베이스에 보고되어 있고 표 A에 개시되며, 알츠하이머병, 파킨슨병, 자폐증, 근위축성 측삭 경화증(ALS), 조현병, 부신백질이영양증, 아이카디 구티에레스 증후군, 파브리병, 레쉬-니한 증후군, 및 멘케스병을 포함하지만, 이들로 제한되지 않는다. 따라서, 본 발명의 양상은 이하에 논의하는 바와 같이, 이들 질환 중 어느 것과 관련된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는 방법에 관한 것이다.

[0956] **알츠하이머병**

[0957] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템, 및 조성물은 알츠하이머병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, PSEN1, PSEN2, 및 APP로부터 선택되는 적어도 하나의 유전자에 존재한다:

[0958] NM_000021.3(PSEN1):c.796G>A (p.Gly266Ser)
NM_000484.3(APP):c.2017G>A (p.Ala673Thr)
NM_000484.3(APP):c.2149G>A (p.Val717Ile)
NM_000484.3(APP):c.2137G>A (p.Ala713Thr)
NM_000484.3(APP):c.2143G>A (p.Val715Met)
NM_000484.3(APP):c.2141C>T (p.Thr714Ile)
NM_000021.3(PSEN1):c.438G>A (p.Met146Ile)
NM_000021.3(PSEN1):c.1229G>A (p.Cys410Tyr)
NM_000021.3(PSEN1):c.487C>T (p.His163Tyr)
NM_000021.3(PSEN1):c.799C>T (p.Pro267Ser)
NM_000021.3(PSEN1):c.236C>T (p.Ala79Val)
NM_000021.3(PSEN1):c.509C>T (p.Ser170Phe)

[0959] NM_000447.2(PSEN2):c.1289C>T (p.Thr430Met)
NM_000447.2(PSEN2):c.717G>A (p.Met239Ile)
NM_000447.2(PSEN2):c.254C>T (p.Ala85Val)
NM_000021.3(PSEN1):c.806G>A (p.Arg269His)
NM_000484.3(APP):c.2018C>T (p.Ala673Val).

[0960] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 PSEN1, PSEN2, 및 APP로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 알츠하이머병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0961] **파킨슨병**

[0962] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 파킨슨병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, SNCA, PLA2G6, FBXO7, VPS35, EIF4G1, DNAJC6, PRKN, SYNJ1, CHCHD2, PINK1, PARK7, LRRK2,

ATP13A2, 및 GBA로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000345.3(SNCA):c.157G>A (p.Ala53Thr)
NM_000345.3(SNCA):c.152G>A (p.Gly51Asp)
NM_003560.3(PLA2G6):c.2222G>A (p.Arg741Gln)
NM_003560.3(PLA2G6):c.2239C>T (p.Arg747Trp)
NM_003560.3(PLA2G6):c.1904G>A (p.Arg635Gln)
NM_003560.3(PLA2G6):c.1354C>T (p.Gln452Ter)
NM_012179.3(FBXO7):c.1492C>T (p.Arg498Ter)
NM_012179.3(FBXO7):c.65C>T (p.Thr22Met)
NM_018206.5(VPS35):c.1858G>A (p.Asp620Asn)
NM_198241.2(EIF4G1):c.3614G>A (p.Arg1205His)
NM_198241.2(EIF4G1):c.1505C>T (p.Ala502Val)
NM_001256865.1(DNAJC6):c.2200C>T (p.Gln734Ter)
NM_001256865.1(DNAJC6):c.2326C>T (p.Gln776Ter)
NM_004562.2(PRKN):c.931C>T (p.Gln311Ter)
NM_004562.2(PRKN):c.1358G>A (p.Trp453Ter)
NM_004562.2(PRKN):c.635G>A (p.Cys212Tyr)

[0963]

NM_203446.2(SYNJ1):c.773G>A (p.Arg258Gln)
NM_001320327.1(CHCHD2):c.182C>T (p.Thr61Ile)
NM_001320327.1(CHCHD2):c.434G>A (p.Arg145Gln)
NM_001320327.1(CHCHD2):c.300+5G>A
NM_032409.2(PINK1):c.926G>A (p.Gly309Asp)
NM_032409.2(PINK1):c.1311G>A (p.Trp437Ter)
NM_032409.2(PINK1):c.736C>T (p.Arg246Ter)
NM_032409.2(PINK1):c.836G>A (p.Arg279His)
NM_032409.2(PINK1):c.938C>T (p.Thr313Met)
NM_032409.2(PINK1):c.1366C>T (p.Gln456Ter)
NM_007262.4(PARK7):c.78G>A (p.Met26Ile)
NM_198578.3(LRRK2):c.4321C>T (p.Arg1441Cys)
NM_198578.3(LRRK2):c.4322G>A (p.Arg1441His)
NM_198578.3(LRRK2):c.1256C>T (p.Ala419Val)
NM_198578.3(LRRK2):c.6055G>A (p.Gly2019Ser)
NM_022089.3(ATP13A2):c.1306+5G>A
NM_022089.3(ATP13A2):c.2629G>A (p.Gly877Arg)
NM_022089.3(ATP13A2):c.490C>T (p.Arg164Trp)
NM_001005741.2(GBA):c.1444G>A (p.Asp482Asn)

[0964]

[0965] m. 15950G>A.

[0966] 표 A를 참조한다. 따라서, 본 발명의 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 SNCA, PLA2G6, FBXO7, VPS35, EIF4G1, DNAJC6, PRKN, SYNJ1, CHCHD2, PINK1, PARK7, LRRK2, ATP13A2, 및 GBA로부터 선택되는 적어도 하나의 유전자 내 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된

하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 파킨슨병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0967] **자폐증**

[0968] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 자폐증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, MECP2, NLGN3, SLC9A9, EHMT1, CHD8, NLGN4X, GSPT2, 및 PTEN으로부터 선택되는 적어도 하나의 유전자에 존재한다:

[0969] NM_001110792.1(MECP2):c.916C>T (p.Arg306Ter)

NM_004992.3(MECP2):c.473C>T (p.Thr158Met)

NM_018977.3(NLGN3):c.1351C>T (p.Arg451Cys)

NM_173653.3(SLC9A9):c.1267C>T (p.Arg423Ter)

NM_024757.4(EHMT1):c.3413G>A (p.Trp1138Ter)

NM_020920.3(CHD8):c.2875C>T (p.Gln959Ter)

NM_020920.3(CHD8):c.3172C>T (p.Arg1058Ter)

NM_181332.2(NLGN4X):c.301C>T (p.Arg101Ter)

NM_018094.4(GSPT2):c.1021G>A (p.Val341Ile)

NM_000314.6(PTEN):c.392C>T (p.Thr131Ile)

[0970]

[0971] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 MECP2, NLGN3, SLC9A9, EHMT1, CHD8, NLGN4X, GSPT2, 및 PTEN으로부터 선택된 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 자폐증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0972] **근위축성 측삭 경화증 (ALS)**

[0973] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 ALS와 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, SOD1, VCP, UBQLN2, ERBB4, HNRNPA1, TUBA4A, SOD1, TARDBP, FIG4, OPTN, SETX, SPG11, FUS, VAPB, ANG, CHCHD10, SQSTM1, 및 TBK1로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000454.4(SOD1):c.289G>A (p.Asp97Asn)

NM_007126.3(VCP):c.1774G>A (p.Asp592Asn)

NM_007126.3(VCP):c.464G>A (p.Arg155His)

NM_007126.3(VCP):c.572G>A (p.Arg191Gln)

NM_013444.3(UBQLN2):c.1489C>T (p.Pro497Ser)

NM_013444.3(UBQLN2):c.1525C>T (p.Pro509Ser)

NM_013444.3(UBQLN2):c.1573C>T (p.Pro525Ser)

NM_013444.3(UBQLN2):c.1490C>T (p.Pro497Leu)

NM_005235.2(ERBB4):c.2780G>A (p.Arg927Gln)

NM_005235.2(ERBB4):c.3823C>T (p.Arg1275Trp)

NM_031157.3(HNRNPA1):c.940G>A (p.Asp314Asn)

[0974]

NM_006000.2(TUBA4A):c.643C>T (p.Arg215Cys)
NM_006000.2(TUBA4A):c.958C>T (p.Arg320Cys)
NM_006000.2(TUBA4A):c.959G>A (p.Arg320His)
NM_006000.2(TUBA4A):c.1220G>A (p.Trp407Ter)
NM_006000.2(TUBA4A):c.1147G>A (p.Ala383Thr)
NM_000454.4(SOD1):c.112G>A (p.Gly38Arg)
NM_000454.4(SOD1):c.124G>A (p.Gly42Ser)
NM_000454.4(SOD1):c.125G>A (p.Gly42Asp)
NM_000454.4(SOD1):c.14C>T (p.Ala5Val)
NM_000454.4(SOD1):c.13G>A (p.Ala5Thr)
NM_000454.4(SOD1):c.436G>A (p.Ala146Thr)
NM_000454.4(SOD1):c.64G>A (p.Glu22Lys)
NM_000454.4(SOD1):c.404G>A (p.Ser135Asn)
NM_000454.4(SOD1):c.49G>A (p.Gly17Ser)
NM_000454.4(SOD1):c.217G>A (p.Gly73Ser)
NM_007375.3(TARDBP):c.892G>A (p.Gly298Ser)
NM_007375.3(TARDBP):c.943G>A (p.Ala315Thr)
NM_007375.3(TARDBP):c.883G>A (p.Gly295Ser)
NM_007375.3(TARDBP):c.*697G>A
NM_007375.3(TARDBP):c.1144G>A (p.Ala382Thr)
NM_007375.3(TARDBP):c.859G>A (p.Gly287Ser)
NM_014845.5(FIG4):c.547C>T (p.Arg183Ter)
NM_001008211.1(OPTN):c.1192C>T (p.Gln398Ter)
NM_015046.5(SETX):c.6407G>A (p.Arg2136His)
NM_015046.5(SETX):c.8C>T (p.Thr3Ile)
NM_025137.3(SPG11):c.118C>T (p.Gln40Ter)
NM_025137.3(SPG11):c.267G>A (p.Trp89Ter)
NM_025137.3(SPG11):c.5974C>T (p.Arg1992Ter)
NM_004960.3(FUS):c.1553G>A (p.Arg518Lys)
NM_004960.3(FUS):c.1561C>T (p.Arg521Cys)
NM_004960.3(FUS):c.1562G>A (p.Arg521His)
NM_004960.3(FUS):c.1520G>A (p.Gly507Asp)
NM_004960.3(FUS):c.1483C>T (p.Arg495Ter)

[0975]

NM_004960.3(FUS):c.616G>A (p.Gly206Ser)

NM_004960.3(FUS):c.646C>T (p.Arg216Cys)

NM_004738.4(VAPB):c.166C>T (p.Pro56Ser)

NM_004738.4(VAPB):c.137C>T (p.Thr46Ile)

NM_001145.4(ANG):c.164G>A (p.Arg55Lys)

NM_001145.4(ANG):c.155G>A (p.Ser52Asn)

NM_001145.4(ANG):c.407C>T (p.Pro136Leu)

NM_001145.4(ANG):c.409G>A (p.Val137Ile)

NM_001301339.1(CHCHD10):c.239C>T (p.Pro80Leu)

NM_001301339.1(CHCHD10):c.176C>T (p.Ser59Leu)

NM_001142298.1(SQSTM1):c.-47-1924C>T

NM_003900.4(SQSTM1):c.1160C>T (p.Pro387Leu)

NM_003900.4(SQSTM1):c.1175C>T (p.Pro392Leu)

NM_013254.3(TBK1):c.1340+1G>A

NM_013254.3(TBK1):c.2086G>A (p.Glu696Lys)

[0976]

[0977] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 SOD1, VCP, UBQLN2, ERBB4, HNRNPA1, TUBA4A, SOD1, TARDBP, FIG4, OPTN, SETX, SPG11, FUS, VAPB, ANG, CHCHD10, SQSTM1, 및 TBK1로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 ALS를 치료 또는 예방하기 위한 방법에 관한 것이다.

[0978] **조현병**

[0979] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 조현병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, PRODH, SETD1A, 및 SHANK3으로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_016335.4(PRODH):c.1292G>A (p.Arg431His)

NM_016335.4(PRODH):c.1397C>T (p.Thr466Met)

NM_014712.2(SETD1A):c.2209C>T (p.Gln737Ter)

NM_033517.1(SHANK3):c.3349C>T (p.Arg1117Ter)

NM_033517.1(SHANK3):c.1606C>T (p.Arg536Trp)

[0980]

[0981] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 PRODH, SETD1A, 및 SHANK3로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 조현병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0982] **부신백질이영양증**

[0983] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 부신백질이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 ABCD1 유전자에 존재한다:

NM_000033.3(ABCD1):c.421G>A (p.Ala141Thr)

NM_000033.3(ABCD1):c.796G>A (p.Gly266Arg)

NM_000033.3(ABCD1):c.1252C>T (p.Arg418Trp)

NM_000033.3(ABCD1):c.1552C>T (p.Arg518Trp)

NM_000033.3(ABCD1):c.1850G>A (p.Arg617His)

NM_000033.3(ABCD1):c.1396C>T (p.Gln466Ter)

NM_000033.3(ABCD1):c.1553G>A (p.Arg518Gln)

NM_000033.3(ABCD1):c.1679C>T (p.Pro560Leu)

NM_000033.3(ABCD1):c.1771C>T (p.Arg591Trp)

NM_000033.3(ABCD1):c.1802G>A (p.Trp601Ter)

NM_000033.3(ABCD1):c.346G>A (p.Gly116Arg)

NM_000033.3(ABCD1):c.406C>T (p.Gln136Ter)

NM_000033.3(ABCD1):c.1661G>A (p.Arg554His)

NM_000033.3(ABCD1):c.1825G>A (p.Glu609Lys)

NM_000033.3(ABCD1):c.1288C>T (p.Gln430Ter)

NM_000033.3(ABCD1):c.1781-1G>A

NM_000033.3(ABCD1):c.529C>T (p.Gln177Ter)

NM_000033.3(ABCD1):c.1866-10G>A

[0984]

[0985] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 적어도 ABCD1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 부신백질이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0986] **아이카디 구티에레스 증후군**

[0987] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 아이카디 구티에레스 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, TREX1, RNASEH2C, ADAR, 및 IFIH1로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_016381.5(TREX1):c.794G>A (p.Trp265Ter)

NM_033629.4(TREX1):c.52G>A (p.Asp18Asn)

NM_033629.4(TREX1):c.490C>T (p.Arg164Ter)

NM_032193.3(RNASEH2C):c.205C>T (p.Arg69Trp)

NM_001111.4(ADAR):c.3019G>A (p.Gly1007Arg)

NM_022168.3(IFIH1):c.2336G>A (p.Arg779His)

NM_022168.3(IFIH1):c.2335C>T (p.Arg779Cys)

[0988]

[0989] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 TREX1, RNASEH2C, ADAR, 및 IFIH1로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 아이카디 구티에레스 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0990] **파브리병**

[0991] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 파브리병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를

포함하여, 적어도 GLA 유전자에 존재한다:

NM_000169.2(GLA):c.1024C>T (p.Arg342Ter)
NM_000169.2(GLA):c.1066C>T (p.Arg356Trp)
NM_000169.2(GLA):c.1025G>A (p.Arg342Gln)
NM_000169.2(GLA):c.281G>A (p.Cys94Tyr)
NM_000169.2(GLA):c.677G>A (p.Trp226Ter)
NM_000169.2(GLA):c.734G>A (p.Trp245Ter)
NM_000169.2(GLA):c.748C>T (p.Gln250Ter)
NM_000169.2(GLA):c.658C>T (p.Arg220Ter)
NM_000169.2(GLA):c.730G>A (p.Asp244Asn)
NM_000169.2(GLA):c.369+1G>A

[0992]

NM_000169.2(GLA):c.335G>A (p.Arg112His)
NM_000169.2(GLA):c.485G>A (p.Trp162Ter)
NM_000169.2(GLA):c.661C>T (p.Gln221Ter)
NM_000169.2(GLA):c.916C>T (p.Gln306Ter)
NM_000169.2(GLA):c.1072G>A (p.Glu358Lys)
NM_000169.2(GLA):c.1087C>T (p.Arg363Cys)
NM_000169.2(GLA):c.1088G>A (p.Arg363His)
NM_000169.2(GLA):c.605G>A (p.Cys202Tyr)
NM_000169.2(GLA):c.830G>A (p.Trp277Ter)
NM_000169.2(GLA):c.979C>T (p.Gln327Ter)
NM_000169.2(GLA):c.422C>T (p.Thr141Ile)
NM_000169.2(GLA):c.285G>A (p.Trp95Ter)
NM_000169.2(GLA):c.735G>A (p.Trp245Ter)
NM_000169.2(GLA):c.639+919G>A
NM_000169.2(GLA):c.680G>A (p.Arg227Gln)
NM_000169.2(GLA):c.679C>T (p.Arg227Ter)
NM_000169.2(GLA):c.242G>A (p.Trp81Ter)
NM_000169.2(GLA):c.901C>T (p.Arg301Ter)
NM_000169.2(GLA):c.974G>A (p.Gly325Asp)
NM_000169.2(GLA):c.847C>T (p.Gln283Ter)
NM_000169.2(GLA):c.469C>T (p.Gln157Ter)
NM_000169.2(GLA):c.1118G>A (p.Gly373Asp)

[0993]

[0994] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 적어도 GLA 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 파브리병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[0995] **레쉬-니한 증후군**

[0996] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 레쉬-니한 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 HPRT1 유전자에 존재한다:

[0997] NM_000194.2(HPRT1):c.151C>T (p.Arg51Ter)

[0998] NM_000194.2(HPRT1):c.384+1G>A

[0999] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 적어도 HPRT1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 레쉬-니한 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1000] **멘케스병**

[1001] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 멘케스병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 ATP7A 유전자에 존재한다:

NM_000052.6(ATP7A):c.601C>T (p.Arg201Ter)
NM_000052.6(ATP7A):c.2938C>T (p.Arg980Ter)
NM_000052.6(ATP7A):c.3056G>A (p.Gly1019Asp)
NM_000052.6(ATP7A):c.598C>T (p.Gln200Ter)
NM_000052.6(ATP7A):c.1225C>T (p.Arg409Ter)
NM_000052.6(ATP7A):c.1544-1G>A
NM_000052.6(ATP7A):c.1639C>T (p.Arg547Ter)
NM_000052.6(ATP7A):c.1933C>T (p.Arg645Ter)
NM_000052.6(ATP7A):c.1946+5G>A
NM_000052.6(ATP7A):c.1950G>A (p.Trp650Ter)
NM_000052.6(ATP7A):c.2179G>A (p.Gly727Arg)
NM_000052.6(ATP7A):c.2187G>A (p.Trp729Ter)
NM_000052.6(ATP7A):c.2383C>T (p.Arg795Ter)
NM_000052.6(ATP7A):c.2499-1G>A
NM_000052.6(ATP7A):c.2555C>T (p.Pro852Leu)
NM_000052.6(ATP7A):c.2956C>T (p.Arg986Ter)
NM_000052.6(ATP7A):c.3112-1G>A
NM_000052.6(ATP7A):c.3466C>T (p.Gln1156Ter)
NM_000052.6(ATP7A):c.3502C>T (p.Gln1168Ter)
NM_000052.6(ATP7A):c.3764G>A (p.Gly1255Glu)
NM_000052.6(ATP7A):c.3943G>A (p.Gly1315Arg)
NM_000052.6(ATP7A):c.4123+1G>A

[1002]

[1003] NM_000052.6(ATP7A):c.4226+5G>A

[1004] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 적어도 ATP7A 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 멘케스병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1005] **안질환**

[1006] 스타가르트병, 바르데-비들 증후군, 추체-간체 이영양증, 선천성 정지성 야맹증, 어셔 증후군, 레버 선천성 흑암시, 망막색소 변성증, 및 완전색맹을 포함하여, 다양한 안질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는

ClinVar 데이터베이스에 보고되어 있고 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기 기술된 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1007] **스타가르트병**

[1008] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 스타가르트병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 적어도 하기를 포함하여, 병원성 돌연변이/SNP는 ABCA4 유전자에 존재한다:

NM_000350.2(ABCA4):c.4429C>T (p.Gln1477Ter)
NM_000350.2(ABCA4):c.6647C>T (p.Ala2216Val)
NM_000350.2(ABCA4):c.5312+1G>A
NM_000350.2(ABCA4):c.5189G>A (p.Trp1730Ter)
NM_000350.2(ABCA4):c.4352+1G>A
NM_000350.2(ABCA4):c.4253+5G>A
NM_000350.2(ABCA4):c.3871C>T (p.Gln1291Ter)
NM_000350.2(ABCA4):c.3813G>A (p.Glu1271=)
NM_000350.2(ABCA4):c.1293G>A (p.Trp431Ter)
NM_000350.2(ABCA4):c.206G>A (p.Trp69Ter)
NM_000350.2(ABCA4):c.3322C>T (p.Arg1108Cys)
NM_000350.2(ABCA4):c.1804C>T (p.Arg602Trp)
NM_000350.2(ABCA4):c.1937+1G>A
NM_000350.2(ABCA4):c.2564G>A (p.Trp855Ter)

[1009]

NM_000350.2(ABCA4):c.4234C>T (p.Gln1412Ter)
NM_000350.2(ABCA4):c.4457C>T (p.Pro1486Leu)
NM_000350.2(ABCA4):c.4594G>A (p.Asp1532Asn)
NM_000350.2(ABCA4):c.4919G>A (p.Arg1640Gln)
NM_000350.2(ABCA4):c.5196+1G>A
NM_000350.2(ABCA4):c.6316C>T (p.Arg2106Cys)
NM_000350.2(ABCA4):c.3056C>T (p.Thr1019Met)
NM_000350.2(ABCA4):c.52C>T (p.Arg18Trp)
NM_000350.2(ABCA4):c.122G>A (p.Trp41Ter)
NM_000350.2(ABCA4):c.1903C>T (p.Gln635Ter)
NM_000350.2(ABCA4):c.194G>A (p.Gly65Glu)
NM_000350.2(ABCA4):c.3085C>T (p.Gln1029Ter)
NM_000350.2(ABCA4):c.4195G>A (p.Glu1399Lys)
NM_000350.2(ABCA4):c.454C>T (p.Arg152Ter)
NM_000350.2(ABCA4):c.45G>A (p.Trp15Ter)
NM_000350.2(ABCA4):c.4610C>T (p.Thr1537Met)
NM_000350.2(ABCA4):c.6112C>T (p.Arg2038Trp)
NM_000350.2(ABCA4):c.6118C>T (p.Arg2040Ter)
NM_000350.2(ABCA4):c.6342G>A (p.Val2114=)
NM_000350.2(ABCA4):c.6658C>T (p.Gln2220Ter)

[1010]

[1011] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 ABCA4 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 스타가르트병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1012] **바르데-비들 증후군**

[1013] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 바르데-비들 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, BBS1, BBS2, BBS7, BBS9, BBS10, BBS12, LZTFL1, 및 TRIM32로부터 선택되는 적어도 하나의 유전자에 존재한다:

[1014]
NM_024649.4(BBS1):c.416G>A (p.Trp139Ter)
NM_024649.4(BBS1):c.871C>T (p.Gln291Ter)

NM_198428.2(BBS9):c.263+1G>A
NM_001178007.1(BBS12):c.1704G>A (p.Trp568Ter)
NM_001276378.1(LZTFL1):c.271C>T (p.Arg91Ter)
NM_031885.3(BBS2):c.1864C>T (p.Arg622Ter)
NM_198428.2(BBS9):c.1759C>T (p.Arg587Ter)
NM_198428.2(BBS9):c.1789+1G>A
NM_024649.4(BBS1):c.432+1G>A
NM_176824.2(BBS7):c.632C>T (p.Thr211Ile)
NM_012210.3(TRIM32):c.388C>T (p.Pro130Ser)
NM_031885.3(BBS2):c.823C>T (p.Arg275Ter)
NM_024685.3(BBS10):c.145C>T (p.Arg49Trp)

[1015]

[1016] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 BBS1, BBS2, BBS7, BBS9, BBS10, BBS12, LZTFL1, 및 TRIM32로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 바르데-비들 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1017] **추체-간체 이영양증**

[1018] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 추체-간체 이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, RPGRIP1, DRAM2, ABCA4, ADAM9, 및 CACNA1F로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_020366.3(RPGRIP1):c.154C>T (p.Arg52Ter)

NM_178454.5(DRAM2):c.494G>A (p.Trp165Ter)

NM_178454.5(DRAM2):c.131G>A (p.Ser44Asn)

NM_000350.2(ABCA4):c.161G>A (p.Cys54Tyr)

NM_000350.2(ABCA4):c.5714+5G>A

NM_000350.2(ABCA4):c.880C>T (p.Gln294Ter)

NM_000350.2(ABCA4):c.6079C>T (p.Leu2027Phe)

NM_000350.2(ABCA4):c.3113C>T (p.Ala1038Val)

NM_000350.2(ABCA4):c.634C>T (p.Arg212Cys)

NM_003816.2(ADAM9):c.490C>T (p.Arg164Ter)

[1019]

NM_005183.3(CACNA1F):c.244C>T (p.Arg82Ter)

[1020]

[1021] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 RPGRIP1, DRAM2, ABCA4, ADAM9, 및 CACNA1F로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 추체-간체 이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1022] **선천성 정지성 야맹증**

[1023] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 선천성 정지성 야맹증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, GRM6, TRPM1, GPR179, 및 CACNA1F로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000843.3(GRM6):c.1462C>T (p.Gln488Ter)

NM_002420.5(TRPM1):c.2998C>T (p.Arg1000Ter)

NM_001004334.3(GPR179):c.673C>T (p.Gln225Ter)

NM_005183.3(CACNA1F):c.2576+1G>A

[1024]

[1025] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 GRM6, TRPM1, GPR179, 및 CACNA1F로부터 선택된 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 선천성 정지성 야맹증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1026] **어셔 증후군**

[1027] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 어셔 증후관과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, MYO7A, USH1C, CDH23, PCDH15, USH2A, ADGRV1, WHRN, 및 CLRN1로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000260.3(MYO7A):c.640G>A (p.Gly214Arg)

NM_000260.3(MYO7A):c.1200+1G>A

NM_000260.3(MYO7A):c.141G>A (p.Trp47Ter)

[1028]

NM_000260.3(MYO7A):c.1556G>A (p.Gly519Asp)
NM_000260.3(MYO7A):c.1900C>T (p.Arg634Ter)
NM_000260.3(MYO7A):c.1963C>T (p.Gln655Ter)
NM_000260.3(MYO7A):c.2094+1G>A
NM_000260.3(MYO7A):c.4293G>A (p.Trp1431Ter)
NM_000260.3(MYO7A):c.5101C>T (p.Arg1701Ter)
NM_000260.3(MYO7A):c.5617C>T (p.Arg1873Trp)
NM_000260.3(MYO7A):c.5660C>T (p.Pro1887Leu)
NM_000260.3(MYO7A):c.6070C>T (p.Arg2024Ter)
NM_000260.3(MYO7A):c.470+1G>A
NM_000260.3(MYO7A):c.5968C>T (p.Gln1990Ter)
NM_000260.3(MYO7A):c.3719G>A (p.Arg1240Gln)
NM_000260.3(MYO7A):c.494C>T (p.Thr165Met)
NM_000260.3(MYO7A):c.5392C>T (p.Gln1798Ter)
NM_000260.3(MYO7A):c.5648G>A (p.Arg1883Gln)
NM_000260.3(MYO7A):c.448C>T (p.Arg150Ter)
NM_000260.3(MYO7A):c.700C>T (p.Gln234Ter)
NM_000260.3(MYO7A):c.635G>A (p.Arg212His)
NM_000260.3(MYO7A):c.1996C>T (p.Arg666Ter)
NM_005709.3(USH1C):c.216G>A (p.Val72=)
NM_022124.5(CDH23):c.7362+5G>A
NM_022124.5(CDH23):c.3481C>T (p.Arg1161Ter)
NM_022124.5(CDH23):c.3628C>T (p.Gln1210Ter)
NM_022124.5(CDH23):c.5272C>T (p.Gln1758Ter)
NM_022124.5(CDH23):c.5712+1G>A
NM_022124.5(CDH23):c.5712G>A (p.Thr1904=)
NM_022124.5(CDH23):c.5923+1G>A
NM_022124.5(CDH23):c.6049+1G>A
NM_022124.5(CDH23):c.7776G>A (p.Trp2592Ter)
NM_022124.5(CDH23):c.9556C>T (p.Arg3186Ter)
NM_022124.5(CDH23):c.3706C>T (p.Arg1236Ter)
NM_022124.5(CDH23):c.4309C>T (p.Arg1437Ter)
NM_022124.5(CDH23):c.6050-9G>A

[1029]

NM_033056.3(PCDH15):c.3316C>T (p.Arg1106Ter)
NM_033056.3(PCDH15):c.7C>T (p.Arg3Ter)
NM_033056.3(PCDH15):c.1927C>T (p.Arg643Ter)
NM_001142772.1(PCDH15):c.400C>T (p.Arg134Ter)
NM_033056.3(PCDH15):c.3358C>T (p.Arg1120Ter)
NM_206933.2(USH2A):c.11048-1G>A
NM_206933.2(USH2A):c.1143+1G>A
NM_206933.2(USH2A):c.11954G>A (p.Trp3985Ter)
NM_206933.2(USH2A):c.12868C>T (p.Gln4290Ter)
NM_206933.2(USH2A):c.14180G>A (p.Trp4727Ter)
NM_206933.2(USH2A):c.14911C>T (p.Arg4971Ter)
NM_206933.2(USH2A):c.5788C>T (p.Arg1930Ter)
NM_206933.2(USH2A):c.5858-1G>A
NM_206933.2(USH2A):c.6224G>A (p.Trp2075Ter)
NM_206933.2(USH2A):c.820C>T (p.Arg274Ter)
NM_206933.2(USH2A):c.8981G>A (p.Trp2994Ter)
NM_206933.2(USH2A):c.9304C>T (p.Gln3102Ter)
NM_206933.2(USH2A):c.13010C>T (p.Thr4337Met)
NM_206933.2(USH2A):c.14248C>T (p.Gln4750Ter)
NM_206933.2(USH2A):c.6398G>A (p.Trp2133Ter)
NM_206933.2(USH2A):c.632G>A (p.Trp211Ter)
NM_206933.2(USH2A):c.6601C>T (p.Gln2201Ter)
NM_206933.2(USH2A):c.13316C>T (p.Thr4439Ile)
NM_206933.2(USH2A):c.4405C>T (p.Gln1469Ter)
NM_206933.2(USH2A):c.9570+1G>A
NM_206933.2(USH2A):c.8740C>T (p.Arg2914Ter)
NM_206933.2(USH2A):c.8681+1G>A
NM_206933.2(USH2A):c.1000C>T (p.Arg334Trp)
NM_206933.2(USH2A):c.14175G>A (p.Trp4725Ter)
NM_206933.2(USH2A):c.9390G>A (p.Trp3130Ter)
NM_206933.2(USH2A):c.908G>A (p.Arg303His)
NM_206933.2(USH2A):c.5776+1G>A
NM_206933.2(USH2A):c.11156G>A (p.Arg3719His)

[1030]

NM_032119.3(ADGRV1):c.2398C>T (p.Arg800Ter)
NM_032119.3(ADGRV1):c.7406G>A (p.Trp2469Ter)
NM_032119.3(ADGRV1):c.12631C>T (p.Arg4211Ter)
NM_032119.3(ADGRV1):c.7129C>T (p.Arg2377Ter)
NM_032119.3(ADGRV1):c.14885G>A (p.Trp4962Ter)
NM_015404.3(WHRN):c.1267C>T (p.Arg423Ter)
NM_174878.2(CLRN1):c.619C>T (p.Arg207Ter)

[1031]

[1032] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 MYO7A, USH1C, CDH23, PCDH15, USH2A, ADGRV1, WHRN, 및 CLRN1로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는

C>T 돌연변이/SNP를 교정하여 증강된 어셔 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1033] **레버 선천성 흑암시**

[1034] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 레버 선천성 흑암시와 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여 TULP1, RPE65, SPATA7, AIPL1, CRB1, NMNAT1, 및 PEX1로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_003322.5(TULP1):c.1495+1G>A

NM_000329.2(RPE65):c.11+5G>A

NM_018418.4(SPATA7):c.322C>T (p.Arg108Ter)

NM_014336.4(AIPL1):c.784G>A (p.Gly262Ser)

NM_201253.2(CRB1):c.1576C>T (p.Arg526Ter)

NM_201253.2(CRB1):c.3307G>A (p.Gly1103Arg)

NM_201253.2(CRB1):c.2843G>A (p.Cys948Tyr)

NM_022787.3(NMNAT1):c.769G>A (p.Glu257Lys)

NM_000466.2(PEX1):c.2528G>A (p.Gly843Asp)

[1035]

[1036] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 TULP1, RPE65, SPATA7, AIPL1, CRB1, NMNAT1, 및 PEX1로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 레버 선천성 흑암시를 치료 또는 예방하기 위한 방법에 관한 것이다.

[1037] **망막 색소변성증**

[1038] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 망막색소 변성증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, CRB1, IFT140, RP1, IMPDH1, PRPF31, RPGR, ABCA4, RPE65, EYS, NRL, FAM161A, NR2E3, USH2A, RHO, PDE6B, KLHL7, PDE6A, CNGB1, BEST1, C2orf71, PRPH2, CA4, CERKL, RPE65, PDE6B, 및 ADGRV1로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_001257965.1(CRB1):c.2711G>A (p.Cys904Tyr)
NM_014714.3(IFT140):c.3827G>A (p.Gly1276Glu)
NM_006269.1(RP1):c.2029C>T (p.Arg677Ter)
NM_000883.3(IMPDH1):c.931G>A (p.Asp311Asn)
NM_015629.3(PRPF31):c.1273C>T (p.Gln425Ter)
NM_015629.3(PRPF31):c.1073+1G>A
NM_000328.2(RPGR):c.1387C>T (p.Gln463Ter)
NM_000350.2(ABCA4):c.4577C>T (p.Thr1526Met)
NM_000350.2(ABCA4):c.6229C>T (p.Arg2077Trp)
NM_000329.2(RPE65):c.271C>T (p.Arg91Trp)
NM_001142800.1(EYS):c.2194C>T (p.Gln732Ter)
NM_001142800.1(EYS):c.490C>T (p.Arg164Ter)
NM_006177.3(NRL):c.151C>T (p.Pro51Ser)
NM_001201543.1(FAM161A):c.1567C>T (p.Arg523Ter)
NM_014249.3(NR2E3):c.166G>A (p.Gly56Arg)
NM_206933.2(USH2A):c.2209C>T (p.Arg737Ter)
NM_206933.2(USH2A):c.14803C>T (p.Arg4935Ter)
NM_206933.2(USH2A):c.10073G>A (p.Cys3358Tyr)
NM_000539.3(RHO):c.541G>A (p.Glu181Lys)
NM_000283.3(PDE6B):c.892C>T (p.Gln298Ter)
NM_001031710.2(KLHL7):c.458C>T (p.Ala153Val)
NM_000440.2(PDE6A):c.1926+1G>A
NM_001297.4(CNGB1):c.2128C>T (p.Gln710Ter)

[1039]

NM_001297.4(CNGB1):c.952C>T (p.Gln318Ter)
NM_004183.3(BEST1):c.682G>A (p.Asp228Asn)
NM_001029883.2(C2orf71):c.1828C>T (p.Gln610Ter)
NM_000322.4(PRPH2):c.647C>T (p.Pro216Leu)
NM_000717.4(CA4):c.40C>T (p.Arg14Trp)
NM_201548.4(CERKL):c.769C>T (p.Arg257Ter)
NM_000329.2(RPE65):c.118G>A (p.Gly40Ser)
NM_000322.4(PRPH2):c.499G>A (p.Gly167Ser)
NM_000539.3(RHO):c.403C>T (p.Arg135Trp)
NM_000283.3(PDE6B):c.2193+1G>A
NM_032119.3(ADGRV1):c.6901C>T (p.Gln2301Ter)

[1040]

[1041] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 CRB1, IFT140, RP1, IMPDH1, PRPF31, RPGR, ABCA4, RPE65, EYS, NRL, FAM161A, NR2E3, USH2A, RHO, PDE6B, KLHL7, PDE6A, CNGB1, BEST1, C2orf71, PRPH2, CA4, CERKL, RPE65, PDE6B, 및 ADGRV1로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 망막색소 변성증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1042] **완전색맹**

[1043] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 완전색맹과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, CNGA3, CNGB3, 및 ATF6으로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_001298.2(CNGA3):c.847C>T (p.Arg283Trp)

NM_001298.2(CNGA3):c.101+1G>A

NM_001298.2(CNGA3):c.1585G>A (p.Val529Met)

NM_019098.4(CNGB3):c.1578+1G>A

NM_019098.4(CNGB3):c.607C>T (p.Arg203Ter)

NM_019098.4(CNGB3):c.1119G>A (p.Trp373Ter)

NM_007348.3(ATF6):c.970C>T (p.Arg324Cys)

[1044]

[1045] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 CNGA3, CNGB3, 및 ATF6으로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 완전색맹을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1046] **청력에 영향을 미치는 질환**

[1047] 제한없이 난청 및 비증상성 청력 상실을 포함하여, 청력에 영향을 미치는 다양한 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 ClinVar 데이터베이스에 보고되어 있고 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기 기술된 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1048] **난청**

[1049] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 난청과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, FGF3, MYO7A, STRC, ACTG1, SLC17A8, TMC1, GJB2, MYH14, COCH, CDH23, USH1C, GJB2, MYO7A, PCDH15, MYO15A, MYO3A, WHRN, DFNB59, TMC1, LOXHD1, TMPRSS3, OTOGL, OTOF, JAG1, 및 MARVELD2로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_005247.2(FGF3):c.283C>T (p.Arg95Trp)

NM_000260.3(MYO7A):c.652G>A (p.Asp218Asn)

NM_000260.3(MYO7A):c.689C>T (p.Ala230Val)

NM_153700.2(STRC):c.4057C>T (p.Gln1353Ter)

NM_001614.3(ACTG1):c.721G>A (p.Glu241Lys)

NM_139319.2(SLC17A8):c.632C>T (p.Ala211Val)

NM_138691.2(TMC1):c.1714G>A (p.Asp572Asn)

NM_004004.5(GJB2):c.598G>A (p.Gly200Arg)

NM_004004.5(GJB2):c.71G>A (p.Trp24Ter)

NM_004004.5(GJB2):c.416G>A (p.Ser139Asn)

NM_004004.5(GJB2):c.224G>A (p.Arg75Gln)

NM_004004.5(GJB2):c.95G>A (p.Arg32His)

NM_004004.5(GJB2):c.250G>A (p.Val84Met)

NM_004004.5(GJB2):c.428G>A (p.Arg143Gln)

NM_004004.5(GJB2):c.551G>A (p.Arg184Gln)

NM_004004.5(GJB2):c.223C>T (p.Arg75Trp)

NM_024729.3(MYH14):c.359C>T (p.Ser120Leu)

[1050]

NM_004086.2(COCH):c.151C>T (p.Pro51Ser)
NM_022124.5(CDH23):c.4021G>A (p.Asp1341Asn)
NM_153700.2(STRC):c.4701+1G>A
NM_153676.3(USH1C):c.496+1G>A
NM_004004.5(GJB2):c.131G>A (p.Trp44Ter)
NM_004004.5(GJB2):c.283G>A (p.Val95Met)
NM_004004.5(GJB2):c.298C>T (p.His100Tyr)
NM_004004.5(GJB2):c.427C>T (p.Arg143Trp)
NM_004004.5(GJB2):c.109G>A (p.Val37Ile)
NM_004004.5(GJB2):c.-23+1G>A
NM_004004.5(GJB2):c.148G>A (p.Asp50Asn)
NM_004004.5(GJB2):c.134G>A (p.Gly45Glu)
NM_004004.5(GJB2):c.370C>T (p.Gln124Ter)
NM_004004.5(GJB2):c.230G>A (p.Trp77Ter)
NM_004004.5(GJB2):c.231G>A (p.Trp77Ter)
NM_000260.3(MYO7A):c.5899C>T (p.Arg1967Ter)
NM_000260.3(MYO7A):c.2005C>T (p.Arg669Ter)
NM_033056.3(PCDH15):c.733C>T (p.Arg245Ter)
NM_016239.3(MYO15A):c.3866+1G>A
NM_016239.3(MYO15A):c.6178-1G>A
NM_016239.3(MYO15A):c.8714-1G>A
NM_017433.4(MYO3A):c.2506-1G>A
NM_015404.3(WHRN):c.1417-1G>A
NM_001042702.3(DFNB59):c.499C>T (p.Arg167Ter)
NM_138691.2(TMC1):c.100C>T (p.Arg34Ter)
NM_138691.2(TMC1):c.1165C>T (p.Arg389Ter)
NM_144612.6(LOXHD1):c.2008C>T (p.Arg670Ter)
NM_144612.6(LOXHD1):c.4714C>T (p.Arg1572Ter)
NM_144612.6(LOXHD1):c.4480C>T (p.Arg1494Ter)
NM_024022.2(TMPRSS3):c.325C>T (p.Arg109Trp)
NM_173591.3(OTOGL):c.3076C>T (p.Gln1026Ter)
NM_194248.2(OTOF):c.4483C>T (p.Arg1495Ter)
NM_194248.2(OTOF):c.2122C>T (p.Arg708Ter)

[1051]

NM_194248.2(OTOF):c.2485C>T (p.Gln829Ter)
NM_001038603.2(MARVELD2):c.1498C>T (p.Arg500Ter)

[1052]

[1053] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 FGF3, MYO7A, STRC, ACTG1, SLC17A8, TMC1, GJB2, MYH14, COCH, CDH23, USH1C, GJB2, MYO7A, PCDH15, MYO15A, MYO3A, WHRN, DFNB59, TMC1, LOXHD1, TMPRSS3, OTOGL, OTOF, JAG1, 및 MARVELD2로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 난청을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1054] **비증상성 청력 상실**

[1055] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 비증상성 청력 상실과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, GJB2, POU3F4, MYO15A, TMPRSS3, LOXHD1, OTOF, MYO6, OTOA, STRC, TRIOBP, MARVELD2,

TMC1, TECTA, OTOGL, 및 GIPC3으로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_004004.5(GJB2):c.169C>T (p.Gln57Ter)
NM_000307.4(POU3F4):c.499C>T (p.Arg167Ter)
NM_016239.3(MYO15A):c.8767C>T (p.Arg2923Ter)
NM_024022.2(TMPRSS3):c.323-6G>A
NM_024022.2(TMPRSS3):c.916G>A (p.Ala306Thr)
NM_144612.6(LOXHD1):c.2497C>T (p.Arg833Ter)
NM_194248.2(OTOF):c.2153G>A (p.Trp718Ter)
NM_194248.2(OTOF):c.2818C>T (p.Gln940Ter)
NM_194248.2(OTOF):c.4799+1G>A
NM_004999.3(MYO6):c.826C>T (p.Arg276Ter)
NM_144672.3(OTOA):c.1880+1G>A
NM_153700.2(STRC):c.5188C>T (p.Arg1730Ter)
NM_153700.2(STRC):c.3670C>T (p.Arg1224Ter)
NM_153700.2(STRC):c.4402C>T (p.Arg1468Ter)
NM_024022.2(TMPRSS3):c.1192C>T (p.Gln398Ter)
NM_001039141.2(TRIOBP):c.6598C>T (p.Arg2200Ter)
NM_016239.3(MYO15A):c.7893+1G>A

[1056]

NM_016239.3(MYO15A):c.5531+1G>A
NM_016239.3(MYO15A):c.6046+1G>A
NM_144612.6(LOXHD1):c.3169C>T (p.Arg1057Ter)
NM_001038603.2(MARVELD2):c.1331+1G>A
NM_138691.2(TMC1):c.1676G>A (p.Trp559Ter)
NM_138691.2(TMC1):c.1677G>A (p.Trp559Ter)
NM_005422.2(TECTA):c.5977C>T (p.Arg1993Ter)
NM_173591.3(OTOGL):c.4987C>T (p.Arg1663Ter)
NM_153700.2(STRC):c.3493C>T (p.Gln1165Ter)
NM_153700.2(STRC):c.3217C>T (p.Arg1073Ter)
NM_016239.3(MYO15A):c.5896C>T (p.Arg1966Ter)
NM_133261.2(GIPC3):c.411+1G>A

[1057]

[1058] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 GJB2, POU3F4, MYO15A, TMPRSS3, LOXHD1, OTOF, MYO6, OTOA, STRC, TRIOBP, MARVELD2, TMC1, TECTA, OTOGL, 및 GIPC3으로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 비증상성 청력 상실을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1059] **혈액 질환**

[1060] 제한없이 베타-탈라세미아, 혈우병 A, 혈우병 B, 혈우병 C, 및 비스코트-알드리히 증후군을 포함하여, 다양한 혈액 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 ClinVar 데이터베이스에 보고되어 있고, 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기에 기술된 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1061] **베타-탈라세미아**

[1062] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 베타-탈라세미아와 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 HBB 유전자에 존재한다:

[1063]
NM_000518.4(HBB):c.-137C>T
NM_000518.4(HBB):c.-50-88C>T

NM_000518.4(HBB):c.-140C>T
NM_000518.4(HBB):c.316-197C>T
NM_000518.4(HBB):c.93-21G>A
NM_000518.4(HBB):c.114G>A (p.Trp38Ter)
NM_000518.4(HBB):c.118C>T (p.Gln40Ter)
NM_000518.4(HBB):c.92+1G>A
NM_000518.4(HBB):c.315+1G>A
NM_000518.4(HBB):c.92+5G>A
NM_000518.4(HBB):c.-50-101C>T

[1064]

[1065] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 HBB 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 베타-탈라세미아를 치료 또는 예방하기 위한 방법에 관한 것이다.

[1066] **혈우병 A**

[1067] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 혈우병 A에 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 F8 유전자에 존재한다:

[1068]
NM_000132.3(F8):c.3169G>A (p.Glu1057Lys)
NM_000132.3(F8):c.902G>A (p.Arg301His)
NM_000132.3(F8):c.1834C>T (p.Arg612Cys)

[1069] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 F8 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 혈우병 A를 치료 또는 예방하기 위한 방법에 관한 것이다.

[1070] **인자 V 레이덴**

[1071] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 인자 V 레이덴 돌연변이를 교정하는데 사용된다. 이러한 질환-유발 단일 점 돌연변이 (G1746 -> A)는 백인 집단에서 유전성 다인성 혈전성향증 중 가장 풍부한 유전적 위험 인자를 대표한다. 점 돌연변이에 기인하여, 단일 아미노산 치환 (R534[RIGHTWARDS ARROW]Q)은 혈액 응고 인자 F5의 단백질 C 의존적 단백질가수분해 절단 부위 (R533R534)를 나타낸다. 이형접합 결함은 혈전증 위험성에서 오직 소수의 증가 (ca. 8배)를 수반하지만, 동형접합 결함은 훨씬 더 두드러진 효과를 갖는다 (>80배 증가된 위험율). 19 유도된 RNA 편집은 RNA 수준에서 이의 복구에 의해 이 유전자 결함을 보상할 잠재력을 갖는다.

[1072] **혈우병 B**

[1073] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 혈우병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 F9 유전자에 존재한다:

[1074] NM_000133.3(F9):c.835G>A (p.Ala279Thr)

[1075] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 F9 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 혈우병 B를 치료 또는 예방하기 위한 방법에 관한 것이다.

[1076] **혈우병 C**

[1077] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 혈우병 C와 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 F11 유전자에 존재한다:

[1078] NM_000128.3(F11):c.400C>T (p.Gln134Ter)
NM_000128.3(F11):c.1432G>A (p.Gly478Arg)
NM_000128.3(F11):c.1288G>A (p.Ala430Thr)
NM_000128.3(F11):c.326-1G>A

[1079] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 F11 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 혈우병 C를 치료 또는 예방하기 위한 방법에 관한 것이다.

[1080] **비스코트-알드리히 증후군**

[1081] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 비스코트-알드리히 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여 적어도 WAS 유전자에 존재한다:

[1082] NM_000377.2(WAS):c.37C>T (p.Arg13Ter)
NM_000377.2(WAS):c.257G>A (p.Arg86His)
NM_000377.2(WAS):c.777+1G>A

[1083] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 WAS 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 비스코트-알드리히 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1084] **간 질환**

[1085] 제한없이, 트랜스타이레틴 아밀로이드증, 알파-1-안티트립신 결핍증, 윌슨병, 및 페닐케톤뇨증을 포함하여, 다양한 간 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP가 ClinVar 데이터베이스에 보고되어 있고 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기에 논의되는 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1086] **트랜스타이레틴 아밀로이드증**

[1087] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 트랜스타이레틴 아밀로이드증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 TTR 유전자에 존재한다:

[1088] NM_000371.3(TTR):c.424G>A (p.Val142Ile)
NM_000371.3(TTR):c.148G>A (p.Val50Met)
NM_000371.3(TTR):c.118G>A (p.Val40Ile)

[1089] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 TTR 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 트랜스타이레틴 아밀로이드증을 치료 또는 예방하기 위한 방법에

관한 것이다.

[1090] **알파-1-안티트립신 결핍증**

[1091] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 알파-1-안티트립신 결핍증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 SERPINA1 유전자에 존재한다:

NM_000295.4(SERPINA1):c.538C>T (p.Gln180Ter)

NM_001127701.1(SERPINA1):c.1178C>T (p.Pro393Leu)

NM_001127701.1(SERPINA1):c.230C>T (p.Ser77Phe)

NM_001127701.1(SERPINA1):c.1096G>A (p.Glu366Lys)

NM_000295.4(SERPINA1):c.1177C>T (p.Pro393Ser)

[1092]

[1093] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 SERPINA1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 알파-1-안티트립신 결핍증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1094] **윌슨병**

[1095] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 윌슨병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 ATP7B 유전자에 존재한다:

NM_000053.3(ATP7B):c.2293G>A (p.Asp765Asn)

NM_000053.3(ATP7B):c.3955C>T (p.Arg1319Ter)

NM_000053.3(ATP7B):c.2865+1G>A

NM_000053.3(ATP7B):c.3796G>A (p.Gly1266Arg)

NM_000053.3(ATP7B):c.2621C>T (p.Ala874Val)

NM_000053.3(ATP7B):c.2071G>A (p.Gly691Arg)

NM_000053.3(ATP7B):c.2128G>A (p.Gly710Ser)

NM_000053.3(ATP7B):c.2336G>A (p.Trp779Ter)

NM_000053.3(ATP7B):c.4021G>A (p.Gly1341Ser)

NM_000053.3(ATP7B):c.3182G>A (p.Gly1061Glu)

NM_000053.3(ATP7B):c.4114C>T (p.Gln1372Ter)

NM_000053.3(ATP7B):c.1708-1G>A

NM_000053.3(ATP7B):c.865C>T (p.Gln289Ter)

[1096]

NM_000053.3(ATP7B):c.2930C>T (p.Thr977Met)

NM_000053.3(ATP7B):c.3659C>T (p.Thr1220Met)

NM_000053.3(ATP7B):c.2605G>A (p.Gly869Arg)

NM_000053.3(ATP7B):c.2975C>T (p.Pro992Leu)

NM_000053.3(ATP7B):c.2519C>T (p.Pro840Leu)

NM_000053.3(ATP7B):c.2906G>A (p.Arg969Gln)

[1097]

[1098] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 ATP7B 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 윌슨병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1099] **페닐케톤뇨증**

[1100] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 페닐케톤뇨증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 PAH 유전자에 존재한다:

NM_000277.1(PAH):c.1315+1G>A
NM_000277.1(PAH):c.1222C>T (p.Arg408Trp)
NM_000277.1(PAH):c.838G>A (p.Glu280Lys)
NM_000277.1(PAH):c.331C>T (p.Arg111Ter)
NM_000277.1(PAH):c.782G>A (p.Arg261Gln)
NM_000277.1(PAH):c.754C>T (p.Arg252Trp)
NM_000277.1(PAH):c.473G>A (p.Arg158Gln)
NM_000277.1(PAH):c.727C>T (p.Arg243Ter)
NM_000277.1(PAH):c.842C>T (p.Pro281Leu)
NM_000277.1(PAH):c.728G>A (p.Arg243Gln)
NM_000277.1(PAH):c.1066-11G>A
NM_000277.1(PAH):c.781C>T (p.Arg261Ter)
NM_000277.1(PAH):c.1223G>A (p.Arg408Gln)
NM_000277.1(PAH):c.1162G>A (p.Val388Met)
NM_000277.1(PAH):c.1066-3C>T
NM_000277.1(PAH):c.1208C>T (p.Ala403Val)
NM_000277.1(PAH):c.890G>A (p.Arg297His)

[1101]

NM_000277.1(PAH):c.926C>T (p.Ala309Val)
NM_000277.1(PAH):c.441+1G>A
NM_000277.1(PAH):c.526C>T (p.Arg176Ter)
NM_000277.1(PAH):c.688G>A (p.Val230Ile)
NM_000277.1(PAH):c.721C>T (p.Arg241Cys)
NM_000277.1(PAH):c.745C>T (p.Leu249Phe)
NM_000277.1(PAH):c.442-1G>A
NM_000277.1(PAH):c.842+1G>A
NM_000277.1(PAH):c.776C>T (p.Ala259Val)
NM_000277.1(PAH):c.1200-1G>A
NM_000277.1(PAH):c.912+1G>A
NM_000277.1(PAH):c.1065+1G>A
NM_000277.1(PAH):c.472C>T (p.Arg158Trp)
NM_000277.1(PAH):c.755G>A (p.Arg252Gln)
NM_000277.1(PAH):c.809G>A (p.Arg270Lys)

[1102]

[1103] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 PAH 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 페닐케톤뇨증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1104] **신장 질환**

[1105] 제한없이, 상염색체 열성 다낭성 신장병 및 신장 카르니틴 수송 결함증을 포함하여, 다양한 신장 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 ClinVar 데이타베이스에 보고되어 있고 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기에 논의되는 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1106] **상염색체 열성 다낭성 신장병**

[1107] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 상염색체 열성 다낭성 신장병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 PKHD1 유전자에 존재한다:

[1108] NM_138694.3(PKHD1):c.10444C>T (p.Arg3482Cys)

NM_138694.3(PKHD1):c.9319C>T (p.Arg3107Ter)
NM_138694.3(PKHD1):c.1480C>T (p.Arg494Ter)
NM_138694.3(PKHD1):c.707+1G>A
NM_138694.3(PKHD1):c.1486C>T (p.Arg496Ter)
NM_138694.3(PKHD1):c.8303-1G>A
NM_138694.3(PKHD1):c.2854G>A (p.Gly952Arg)
NM_138694.3(PKHD1):c.7194G>A (p.Trp2398Ter)
NM_138694.3(PKHD1):c.10219C>T (p.Gln3407Ter)
NM_138694.3(PKHD1):c.107C>T (p.Thr36Met)
NM_138694.3(PKHD1):c.8824C>T (p.Arg2942Ter)
NM_138694.3(PKHD1):c.982C>T (p.Arg328Ter)
NM_138694.3(PKHD1):c.4870C>T (p.Arg1624Trp)
NM_138694.3(PKHD1):c.1602+1G>A
NM_138694.3(PKHD1):c.1694-1G>A
NM_138694.3(PKHD1):c.2341C>T (p.Arg781Ter)
NM_138694.3(PKHD1):c.2407+1G>A
NM_138694.3(PKHD1):c.2452C>T (p.Gln818Ter)
NM_138694.3(PKHD1):c.5236+1G>A
NM_138694.3(PKHD1):c.6499C>T (p.Gln2167Ter)
NM_138694.3(PKHD1):c.2725C>T (p.Arg909Ter)
NM_138694.3(PKHD1):c.370C>T (p.Arg124Ter)
NM_138694.3(PKHD1):c.2810G>A (p.Trp937Ter)

[1109]

[1110] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 PKHD1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 상염색체 열성 다낭성 신장병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1111] **신장 카르니틴 수송 결함증**

[1112] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 신장 카르니틴 수송 결함증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 SLC22A5 유전자에 존재한다:

[1113] NM_003060.3(SLC22A5):c.760C>T (p.Arg254Ter)

NM_003060.3(SLC22A5):c.396G>A (p.Trp132Ter)

NM_003060.3(SLC22A5):c.844C>T (p.Arg282Ter)

NM_003060.3(SLC22A5):c.505C>T (p.Arg169Trp)

NM_003060.3(SLC22A5):c.1319C>T (p.Thr440Met)

NM_003060.3(SLC22A5):c.1195C>T (p.Arg399Trp)

NM_003060.3(SLC22A5):c.695C>T (p.Thr232Met)

NM_003060.3(SLC22A5):c.845G>A (p.Arg282Gln)

NM_003060.3(SLC22A5):c.1193C>T (p.Pro398Leu)

NM_003060.3(SLC22A5):c.1463G>A (p.Arg488His)

NM_003060.3(SLC22A5):c.338G>A (p.Cys113Tyr)

NM_003060.3(SLC22A5):c.136C>T (p.Pro46Ser)

NM_003060.3(SLC22A5):c.506G>A (p.Arg169Gln)

[1114]

[1115] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 SLC22A5 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 신장 카르니틴 수송 결함증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1116] **근육 질환**

[1117] 제한없이, 뒤센 근이영양증, 벡커 근이영양증, 지대근 이영양증, 에머리-드레이푸스 근이영양증, 및 얼굴 어깨팔 근이영양증을 포함한, 다양한 근육 질환과 연 관된 병원성 G>A 또는 C>T 돌연변이/SNP가 ClinVar 데이터베이스에 보고되어 있고, 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기에 논의되는 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1118] **뒤센 근이영양증**

[1119] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 뒤센 근이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 DMD 유전자에 존재한다:

NM_004006.2(DMD):c.2797C>T (p.Gln933Ter)

NM_004006.2(DMD):c.4870C>T (p.Gln1624Ter)

NM_004006.2(DMD):c.5551C>T (p.Gln1851Ter)

[1120]

NM_004006.2(DMD):c.3188G>A (p.Trp1063Ter)
NM_004006.2(DMD):c.8357G>A (p.Trp2786Ter)
NM_004006.2(DMD):c.7817G>A (p.Trp2606Ter)
NM_004006.2(DMD):c.7755G>A (p.Trp2585Ter)
NM_004006.2(DMD):c.5917C>T (p.Gln1973Ter)
NM_004006.2(DMD):c.5641C>T (p.Gln1881Ter)
NM_004006.2(DMD):c.5131C>T (p.Gln1711Ter)
NM_004006.2(DMD):c.4240C>T (p.Gln1414Ter)
NM_004006.2(DMD):c.3427C>T (p.Gln1143Ter)
NM_004006.2(DMD):c.2407C>T (p.Gln803Ter)
NM_004006.2(DMD):c.2368C>T (p.Gln790Ter)
NM_004006.2(DMD):c.1683G>A (p.Trp561Ter)
NM_004006.2(DMD):c.1663C>T (p.Gln555Ter)
NM_004006.2(DMD):c.1388G>A (p.Trp463Ter)
NM_004006.2(DMD):c.1331+1G>A
NM_004006.2(DMD):c.1324C>T (p.Gln442Ter)
NM_004006.2(DMD):c.355C>T (p.Gln119Ter)
NM_004006.2(DMD):c.94-1G>A
NM_004006.2(DMD):c.5506C>T (p.Gln1836Ter)
NM_004006.2(DMD):c.1504C>T (p.Gln502Ter)
NM_004006.2(DMD):c.5032C>T (p.Gln1678Ter)
NM_004006.2(DMD):c.457C>T (p.Gln153Ter)
NM_004006.2(DMD):c.1594C>T (p.Gln532Ter)
NM_004006.2(DMD):c.1150-1G>A
NM_004006.2(DMD):c.6223C>T (p.Gln2075Ter)
NM_004006.2(DMD):c.3747G>A (p.Trp1249Ter)
NM_004006.2(DMD):c.2861G>A (p.Trp954Ter)
NM_004006.2(DMD):c.9563+1G>A
NM_004006.2(DMD):c.4483C>T (p.Gln1495Ter)
NM_004006.2(DMD):c.4312C>T (p.Gln1438Ter)
NM_004006.2(DMD):c.8209C>T (p.Gln2737Ter)
NM_004006.2(DMD):c.4071+1G>A
NM_004006.2(DMD):c.2665C>T (p.Arg889Ter)

[1121]

NM_004006.2(DMD):c.2202G>A (p.Trp734Ter)
NM_004006.2(DMD):c.2077C>T (p.Gln693Ter)
NM_004006.2(DMD):c.1653G>A (p.Trp551Ter)
NM_004006.2(DMD):c.1061G>A (p.Trp354Ter)
NM_004006.2(DMD):c.8914C>T (p.Gln2972Ter)
NM_004006.2(DMD):c.6118-1G>A
NM_004006.2(DMD):c.4729C>T (p.Arg1577Ter)

[1122]

[1123] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 DMD 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 뒤센 근이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1124] **벡커 근이영양증**

[1125] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 벡커 근이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 DMD 유전자에 존재한다:

[1126]
NM_004006.2(DMD):c.3413G>A (p.Trp1138Ter)
NM_004006.2(DMD):c.358-1G>A
NM_004006.2(DMD):c.10108C>T (p.Arg3370Ter)
NM_004006.2(DMD):c.6373C>T (p.Gln2125Ter)
NM_004006.2(DMD):c.9568C>T (p.Arg3190Ter)
NM_004006.2(DMD):c.8713C>T (p.Arg2905Ter)
NM_004006.2(DMD):c.1615C>T (p.Arg539Ter)
NM_004006.2(DMD):c.3151C>T (p.Arg1051Ter)
NM_004006.2(DMD):c.3432+1G>A
NM_004006.2(DMD):c.5287C>T (p.Arg1763Ter)
NM_004006.2(DMD):c.5530C>T (p.Arg1844Ter)
NM_004006.2(DMD):c.8608C>T (p.Arg2870Ter)
NM_004006.2(DMD):c.8656C>T (p.Gln2886Ter)
NM_004006.2(DMD):c.8944C>T (p.Arg2982Ter)
NM_004006.2(DMD):c.5899C>T (p.Arg1967Ter)
NM_004006.2(DMD):c.10033C>T (p.Arg3345Ter)

NM_004006.2(DMD):c.10086+1G>A
NM_004019.2(DMD):c.1020G>A (p.Thr340=)
NM_004006.2(DMD):c.1261C>T (p.Gln421Ter)
NM_004006.2(DMD):c.1465C>T (p.Gln489Ter)
NM_004006.2(DMD):c.1990C>T (p.Gln664Ter)
NM_004006.2(DMD):c.2032C>T (p.Gln678Ter)
NM_004006.2(DMD):c.2332C>T (p.Gln778Ter)
NM_004006.2(DMD):c.2419C>T (p.Gln807Ter)
NM_004006.2(DMD):c.2650C>T (p.Gln884Ter)
NM_004006.2(DMD):c.2804-1G>A
NM_004006.2(DMD):c.3276+1G>A
NM_004006.2(DMD):c.3295C>T (p.Gln1099Ter)
NM_004006.2(DMD):c.336G>A (p.Trp112Ter)
NM_004006.2(DMD):c.3580C>T (p.Gln1194Ter)
NM_004006.2(DMD):c.4117C>T (p.Gln1373Ter)
NM_004006.2(DMD):c.649+1G>A
NM_004006.2(DMD):c.6906G>A (p.Trp2302Ter)
NM_004006.2(DMD):c.7189C>T (p.Gln2397Ter)
NM_004006.2(DMD):c.7309+1G>A
NM_004006.2(DMD):c.7657C>T (p.Arg2553Ter)
NM_004006.2(DMD):c.7682G>A (p.Trp2561Ter)
NM_004006.2(DMD):c.7683G>A (p.Trp2561Ter)
NM_004006.2(DMD):c.7894C>T (p.Gln2632Ter)
NM_004006.2(DMD):c.9361+1G>A
NM_004006.2(DMD):c.9564-1G>A
NM_004006.2(DMD):c.2956C>T (p.Gln986Ter)
NM_004006.2(DMD):c.883C>T (p.Arg295Ter)
NM_004006.2(DMD):c.31+36947G>A
NM_004006.2(DMD):c.10279C>T (p.Gln3427Ter)
NM_004006.2(DMD):c.433C>T (p.Arg145Ter)
NM_004006.2(DMD):c.9G>A (p.Trp3Ter)
NM_004006.2(DMD):c.10171C>T (p.Arg3391Ter)
NM_004006.2(DMD):c.583C>T (p.Arg195Ter)

[1127]

NM_004006.2(DMD):c.9337C>T (p.Arg3113Ter)

NM_004006.2(DMD):c.8038C>T (p.Arg2680Ter)

NM_004006.2(DMD):c.1812+1G>A

NM_004006.2(DMD):c.1093C>T (p.Gln365Ter)

NM_004006.2(DMD):c.1704+1G>A

NM_004006.2(DMD):c.1912C>T (p.Gln638Ter)

NM_004006.2(DMD):c.133C>T (p.Gln45Ter)

NM_004006.2(DMD):c.5868G>A (p.Trp1956Ter)

NM_004006.2(DMD):c.565C>T (p.Gln189Ter)

NM_004006.2(DMD):c.5089C>T (p.Gln1697Ter)

NM_004006.2(DMD):c.2512C>T (p.Gln838Ter)

NM_004006.2(DMD):c.10477C>T (p.Gln3493Ter)

NM_004006.2(DMD):c.93+1G>A

NM_004006.2(DMD):c.4174C>T (p.Gln1392Ter)

NM_004006.2(DMD):c.3940C>T (p.Arg1314Ter)

[1128]

[1129] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 DMD 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 벡커 근이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1130] **지대근 이영양증**

[1131] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 지대근 이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, SGCB, MYOT, LMNA, CAPN3, DYSF, SGCA, TTN, ANO5, TRAPPC11, LMNA, POMT1, 및 FKRP로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000232.4(SGCB):c.31C>T (p.Gln11Ter)

NM_006790.2(MYOT):c.164C>T (p.Ser55Phe)

NM_006790.2(MYOT):c.170C>T (p.Thr57Ile)

NM_170707.3(LMNA):c.1488+1G>A

NM_170707.3(LMNA):c.1609-1G>A

NM_000070.2(CAPN3):c.1715G>A (p.Arg572Gln)

NM_000070.2(CAPN3):c.2243G>A (p.Arg748Gln)

NM_000070.2(CAPN3):c.145C>T (p.Arg49Cys)

[1132]

NM_000070.2(CAPN3):c.1319G>A (p.Arg440Gln)
NM_000070.2(CAPN3):c.1343G>A (p.Arg448His)
NM_000070.2(CAPN3):c.1465C>T (p.Arg489Trp)
NM_000070.2(CAPN3):c.1714C>T (p.Arg572Trp)
NM_000070.2(CAPN3):c.2306G>A (p.Arg769Gln)
NM_000070.2(CAPN3):c.133G>A (p.Ala45Thr)
NM_000070.2(CAPN3):c.499-1G>A
NM_000070.2(CAPN3):c.439C>T (p.Arg147Ter)
NM_000070.2(CAPN3):c.1063C>T (p.Arg355Trp)
NM_000070.2(CAPN3):c.1250C>T (p.Thr417Met)
NM_000070.2(CAPN3):c.245C>T (p.Pro82Leu)
NM_000070.2(CAPN3):c.2242C>T (p.Arg748Ter)
NM_000070.2(CAPN3):c.1318C>T (p.Arg440Trp)
NM_000070.2(CAPN3):c.1333G>A (p.Gly445Arg)
NM_000070.2(CAPN3):c.1957C>T (p.Gln653Ter)
NM_000070.2(CAPN3):c.1801-1G>A
NM_000070.2(CAPN3):c.2263+1G>A
NM_000070.2(CAPN3):c.956C>T (p.Pro319Leu)
NM_000070.2(CAPN3):c.1468C>T (p.Arg490Trp)
NM_000070.2(CAPN3):c.802-9G>A
NM_000070.2(CAPN3):c.1342C>T (p.Arg448Cys)
NM_000070.2(CAPN3):c.1303G>A (p.Glu435Lys)
NM_000070.2(CAPN3):c.1993-1G>A
NM_003494.3(DYSF):c.3113G>A (p.Arg1038Gln)
NM_001130987.1(DYSF):c.5174+1G>A
NM_001130987.1(DYSF):c.159G>A (p.Trp53Ter)
NM_001130987.1(DYSF):c.2929C>T (p.Arg977Trp)
NM_001130987.1(DYSF):c.4282C>T (p.Gln1428Ter)
NM_001130987.1(DYSF):c.1577-1G>A
NM_003494.3(DYSF):c.5529G>A (p.Trp1843Ter)
NM_001130987.1(DYSF):c.1576+1G>A
NM_001130987.1(DYSF):c.4462C>T (p.Gln1488Ter)
NM_003494.3(DYSF):c.5429G>A (p.Arg1810Lys)

[1133]

NM_003494.3(DYSF):c.5077C>T (p.Arg1693Trp)
NM_001130978.1(DYSF):c.1813C>T (p.Gln605Ter)
NM_003494.3(DYSF):c.3230G>A (p.Trp1077Ter)
NM_003494.3(DYSF):c.265C>T (p.Arg89Ter)
NM_003494.3(DYSF):c.4434G>A (p.Trp1478Ter)
NM_003494.3(DYSF):c.3478C>T (p.Gln1160Ter)
NM_001130987.1(DYSF):c.1372G>A (p.Gly458Arg)
NM_003494.3(DYSF):c.4090C>T (p.Gln1364Ter)
NM_001130987.1(DYSF):c.2409+1G>A
NM_003494.3(DYSF):c.1708C>T (p.Gln570Ter)
NM_003494.3(DYSF):c.1956G>A (p.Trp652Ter)
NM_001130987.1(DYSF):c.5004-1G>A
NM_003494.3(DYSF):c.331C>T (p.Gln111Ter)
NM_001130978.1(DYSF):c.5776C>T (p.Arg1926Ter)
NM_003494.3(DYSF):c.6124C>T (p.Arg2042Cys)
NM_003494.3(DYSF):c.2643+1G>A
NM_003494.3(DYSF):c.4253G>A (p.Gly1418Asp)
NM_003494.3(DYSF):c.610C>T (p.Arg204Ter)
NM_003494.3(DYSF):c.1834C>T (p.Gln612Ter)
NM_003494.3(DYSF):c.5668-7G>A
NM_001130978.1(DYSF):c.3137G>A (p.Arg1046His)
NM_003494.3(DYSF):c.1053+1G>A
NM_003494.3(DYSF):c.1398-1G>A
NM_003494.3(DYSF):c.1481-1G>A
NM_003494.3(DYSF):c.2311C>T (p.Gln771Ter)
NM_003494.3(DYSF):c.2869C>T (p.Gln957Ter)
NM_003494.3(DYSF):c.4756C>T (p.Arg1586Ter)
NM_003494.3(DYSF):c.5509G>A (p.Asp1837Asn)
NM_003494.3(DYSF):c.5644C>T (p.Gln1882Ter)
NM_003494.3(DYSF):c.5946+1G>A
NM_003494.3(DYSF):c.937+1G>A
NM_003494.3(DYSF):c.5266C>T (p.Gln1756Ter)
NM_003494.3(DYSF):c.3832C>T (p.Gln1278Ter)

[1134]

NM_003494.3(DYSF):c.5525+1G>A

NM_003494.3(DYSF):c.3112C>T (p.Arg1038Ter)

NM_000023.3(SGCA):c.293G>A (p.Arg98His)

NM_000023.3(SGCA):c.850C>T (p.Arg284Cys)

NM_000023.3(SGCA):c.403C>T (p.Gln135Ter)

NM_000023.3(SGCA):c.409G>A (p.Glu137Lys)

NM_000023.3(SGCA):c.747+1G>A

NM_000023.3(SGCA):c.229C>T (p.Arg77Cys)

NM_000023.3(SGCA):c.101G>A (p.Arg34His)

NM_000023.3(SGCA):c.739G>A (p.Val247Met)

NM_001256850.1(TTN):c.87394C>T (p.Arg29132Ter)

NM_213599.2(ANO5):c.762+1G>A

NM_213599.2(ANO5):c.1213C>T (p.Gln405Ter)

NM_213599.2(ANO5):c.1639C>T (p.Arg547Ter)

NM_213599.2(ANO5):c.1406G>A (p.Trp469Ter)

NM_213599.2(ANO5):c.1210C>T (p.Arg404Ter)

NM_213599.2(ANO5):c.2272C>T (p.Arg758Cys)

NM_213599.2(ANO5):c.41-1G>A

NM_213599.2(ANO5):c.172C>T (p.Arg58Trp)

NM_213599.2(ANO5):c.1898+1G>A

NM_021942.5(TRAPPC11):c.1287+5G>A

NM_170707.3(LMNA):c.1608+1G>A

NM_007171.3(POMT1):c.1864C>T (p.Arg622Ter)

NM_024301.4(FKRP):c.313C>T (p.Gln105Ter)

[1135]

[1136] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 SGCB, MYOT, LMNA, CAPN3, DYSF, SGCA, TTN, ANO5, TRAPPC11, LMNA, POMT1, 및 FKRP로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 지대근 이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1137] **에머리-드레이푸스 근이영양증**

[1138] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 에머리-드레이푸스 근이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 EMD 또는 SYNE1 유전자에 존재한다:

NM_000117.2(EMD):c.3G>A (p.Met1Ile)

NM_033071.3(SYNE1):c.11908C>T (p.Arg3970Ter)

NM_033071.3(SYNE1):c.21721C>T (p.Gln7241Ter)

NM_000117.2(EMD):c.130C>T (p.Gln44Ter)

[1139]

[1140] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 EMD 또는 SYNE1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 에머리-드레이푸스 근이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1141] **얼굴 어깨 팔 근이영양증**

[1142] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 얼굴 어깨 팔 근이영양증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 SMCHD1 유전자에 존재한다:

[1143] NM_015295.2(SMCHD1):c.3801+1G>A
NM_015295.2(SMCHD1):c.1843-1G>A

[1144] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 SMCHD1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 얼굴 어깨 팔 근이영양증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1145] **선천성 대사 이상 (IEM)**

[1146] 제한없이, 원발성 과옥살산뇨증 1형, 아르기니노숙시네이트 리아제 결핍증, 오르니틴 카바모일트랜스퍼라제 결핍증, 및 메이플시럽 뇨증을 포함한, 다양한 IEM과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 ClinVar 데이터베이스에 보고되어 있고 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기에 논의되는 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1147] **원발성 과옥살산뇨증 1형**

[1148] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 원발성 과옥살산뇨증 1형과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 AGXT 유전자에 존재한다:

[1149] NM_000030.2(AGXT):c.245G>A (p.Gly82Glu)
NM_000030.2(AGXT):c.698G>A (p.Arg233His)
NM_000030.2(AGXT):c.466G>A (p.Gly156Arg)
NM_000030.2(AGXT):c.106C>T (p.Arg36Cys)
NM_000030.2(AGXT):c.346G>A (p.Gly116Arg)
NM_000030.2(AGXT):c.568G>A (p.Gly190Arg)
NM_000030.2(AGXT):c.653C>T (p.Ser218Leu)
NM_000030.2(AGXT):c.737G>A (p.Trp246Ter)
NM_000030.2(AGXT):c.1049G>A (p.Gly350Asp)
NM_000030.2(AGXT):c.473C>T (p.Ser158Leu)
NM_000030.2(AGXT):c.907C>T (p.Gln303Ter)
NM_000030.2(AGXT):c.996G>A (p.Trp332Ter)
NM_000030.2(AGXT):c.508G>A (p.Gly170Arg)

[1150] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 AGXT 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 원발성 과옥살산뇨증 1형을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1151] **아르기니노숙시네이트 리아제 결핍증**

[1152] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 아르기니노숙시네이트 리아제 결핍증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 ASL 유전자에 존재한다:

NM_001024943.1(ASL):c.1153C>T (p.Arg385Cys)
NM_000048.3(ASL):c.532G>A (p.Val178Met)
NM_000048.3(ASL):c.545G>A (p.Arg182Gln)
NM_000048.3(ASL):c.175G>A (p.Glu59Lys)
NM_000048.3(ASL):c.718+5G>A
NM_000048.3(ASL):c.889C>T (p.Arg297Trp)

[1153]

NM_000048.3(ASL):c.1360C>T (p.Gln454Ter)
NM_000048.3(ASL):c.1060C>T (p.Gln354Ter)
NM_000048.3(ASL):c.35G>A (p.Arg12Gln)
NM_000048.3(ASL):c.446+1G>A
NM_000048.3(ASL):c.544C>T (p.Arg182Ter)
NM_000048.3(ASL):c.1135C>T (p.Arg379Cys)

[1154]

[1155] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 ASL 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 아르기니노숙시네이트 리아제 결핍증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1156] **오르니틴 카바모일트랜스퍼라제 결핍증**

[1157] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 오르니틴 카바모일트랜스퍼라제 결핍증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 OTC 유전자에 존재한다:

NM_000531.5(OTC):c.119G>A (p.Arg40His)
NM_000531.5(OTC):c.422G>A (p.Arg141Gln)
NM_000531.5(OTC):c.829C>T (p.Arg277Trp)
NM_000531.5(OTC):c.674C>T (p.Pro225Leu)

[1158]

[1159] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 OTC 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 오르니틴 카바모일트랜스퍼라제 결핍증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1160] **메이플시럽 뇨증**

[1161] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 메이플시럽 뇨증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, BCKDHA, BCKDHB, DBT, 및 DLD로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000709.3(BCKDHA):c.476G>A (p.Arg159Gln)

[1162]

NM_183050.3(BCKDHB):c.3G>A (p.Met1Ile)
NM_183050.3(BCKDHB):c.554C>T (p.Pro185Leu)
NM_001918.3(DBT):c.1033G>A (p.Gly345Arg)
NM_000709.3(BCKDHA):c.940C>T (p.Arg314Ter)
NM_000709.3(BCKDHA):c.793C>T (p.Arg265Trp)
NM_000709.3(BCKDHA):c.868G>A (p.Gly290Arg)
NM_000108.4(DLD):c.1123G>A (p.Glu375Lys)
NM_000709.3(BCKDHA):c.1234G>A (p.Val412Met)
NM_000709.3(BCKDHA):c.288+1G>A
NM_000709.3(BCKDHA):c.979G>A (p.Glu327Lys)
NM_001918.3(DBT):c.901C>T (p.Arg301Cys)
NM_183050.3(BCKDHB):c.509G>A (p.Arg170His)
NM_183050.3(BCKDHB):c.799C>T (p.Gln267Ter)
NM_183050.3(BCKDHB):c.853C>T (p.Arg285Ter)
NM_183050.3(BCKDHB):c.970C>T (p.Arg324Ter)
NM_183050.3(BCKDHB):c.832G>A (p.Gly278Ser)
NM_000709.3(BCKDHA):c.1036C>T (p.Arg346Cys)
NM_000709.3(BCKDHA):c.288+9C>T
NM_000709.3(BCKDHA):c.632C>T (p.Thr211Met)
NM_000709.3(BCKDHA):c.659C>T (p.Ala220Val)
NM_000709.3(BCKDHA):c.964C>T (p.Gln322Ter)
NM_001918.3(DBT):c.1291C>T (p.Arg431Ter)
NM_001918.3(DBT):c.251G>A (p.Trp84Ter)
NM_001918.3(DBT):c.871C>T (p.Arg291Ter)
NM_000056.4(BCKDHB):c.1016C>T (p.Ser339Leu)
NM_000056.4(BCKDHB):c.344-1G>A
NM_000056.4(BCKDHB):c.633+1G>A
NM_000056.4(BCKDHB):c.952-1G>A

[1163]

[1164] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 BCKDHA, BCKDHB, DBT, 및 DLD로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 메이플시럽 뇨증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1165] **암-관련 질환**

[1166] 제한없이 유방-난소 암 및 린치 증후군을 포함하여, 다양한 암 및 암-관련 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 ClinVar 데이터베이스에 보고되어 있고 표 A에 개시되어 있다. 따라서, 본 발명의 일 양상은 하기에 논의되는 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1167] **유방-난소 암**

[1168] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 유방-난소 암과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 BRCA1 또는 BRCA2 유전자에 존재한다:

NM_007294.3(BRCA1):c.5095C>T (p.Arg1699Trp)

NM_000059.3(BRCA2):c.7558C>T (p.Arg2520Ter)

NM_007294.3(BRCA1):c.2572C>T (p.Gln858Ter)

NM_007294.3(BRCA1):c.3607C>T (p.Arg1203Ter)

NM_007294.3(BRCA1):c.5503C>T (p.Arg1835Ter)

NM_007294.3(BRCA1):c.2059C>T (p.Gln687Ter)

NM_007294.3(BRCA1):c.4675+1G>A

NM_007294.3(BRCA1):c.5251C>T (p.Arg1751Ter)

NM_007294.3(BRCA1):c.5444G>A (p.Trp1815Ter)

NM_000059.3(BRCA2):c.9318G>A (p.Trp3106Ter)

NM_000059.3(BRCA2):c.9382C>T (p.Arg3128Ter)

NM_000059.3(BRCA2):c.274C>T (p.Gln92Ter)

NM_000059.3(BRCA2):c.6952C>T (p.Arg2318Ter)

NM_007294.3(BRCA1):c.1687C>T (p.Gln563Ter)

NM_007294.3(BRCA1):c.2599C>T (p.Gln867Ter)

NM_007294.3(BRCA1):c.784C>T (p.Gln262Ter)

NM_007294.3(BRCA1):c.280C>T (p.Gln94Ter)

NM_007294.3(BRCA1):c.5542C>T (p.Gln1848Ter)

NM_007294.3(BRCA1):c.5161C>T (p.Gln1721Ter)

NM_007294.3(BRCA1):c.4573C>T (p.Gln1525Ter)

NM_007294.3(BRCA1):c.4270C>T (p.Gln1424Ter)

NM_007294.3(BRCA1):c.4225C>T (p.Gln1409Ter)

[1169]

NM_007294.3(BRCA1):c.4066C>T (p.Gln1356Ter)
NM_007294.3(BRCA1):c.3679C>T (p.Gln1227Ter)
NM_007294.3(BRCA1):c.1918C>T (p.Gln640Ter)
NM_007294.3(BRCA1):c.963G>A (p.Trp321Ter)
NM_007294.3(BRCA1):c.718C>T (p.Gln240Ter)
NM_000059.3(BRCA2):c.9196C>T (p.Gln3066Ter)
NM_000059.3(BRCA2):c.9154C>T (p.Arg3052Trp)
NM_007294.3(BRCA1):c.3991C>T (p.Gln1331Ter)
NM_007294.3(BRCA1):c.4097-1G>A
NM_007294.3(BRCA1):c.1059G>A (p.Trp353Ter)
NM_007294.3(BRCA1):c.1115G>A (p.Trp372Ter)
NM_007294.3(BRCA1):c.1138C>T (p.Gln380Ter)
NM_007294.3(BRCA1):c.1612C>T (p.Gln538Ter)
NM_007294.3(BRCA1):c.1621C>T (p.Gln541Ter)
NM_007294.3(BRCA1):c.1630C>T (p.Gln544Ter)
NM_007294.3(BRCA1):c.178C>T (p.Gln60Ter)
NM_007294.3(BRCA1):c.1969C>T (p.Gln657Ter)
NM_007294.3(BRCA1):c.2275C>T (p.Gln759Ter)
NM_007294.3(BRCA1):c.2410C>T (p.Gln804Ter)
NM_007294.3(BRCA1):c.2869C>T (p.Gln957Ter)
NM_007294.3(BRCA1):c.2923C>T (p.Gln975Ter)
NM_007294.3(BRCA1):c.3268C>T (p.Gln1090Ter)
NM_007294.3(BRCA1):c.3430C>T (p.Gln1144Ter)
NM_007294.3(BRCA1):c.3544C>T (p.Gln1182Ter)
NM_007294.3(BRCA1):c.4075C>T (p.Gln1359Ter)
NM_007294.3(BRCA1):c.4201C>T (p.Gln1401Ter)
NM_007294.3(BRCA1):c.4399C>T (p.Gln1467Ter)
NM_007294.3(BRCA1):c.4552C>T (p.Gln1518Ter)
NM_007294.3(BRCA1):c.5054C>T (p.Thr1685Ile)
NM_007294.3(BRCA1):c.514C>T (p.Gln172Ter)
NM_007294.3(BRCA1):c.5239C>T (p.Gln1747Ter)
NM_007294.3(BRCA1):c.5266C>T (p.Gln1756Ter)
NM_007294.3(BRCA1):c.5335C>T (p.Gln1779Ter)

[1170]

NM_007294.3(BRCA1):c.5345G>A (p.Trp1782Ter)
NM_007294.3(BRCA1):c.5511G>A (p.Trp1837Ter)
NM_007294.3(BRCA1):c.5536C>T (p.Gln1846Ter)
NM_007294.3(BRCA1):c.55C>T (p.Gln19Ter)
NM_007294.3(BRCA1):c.949C>T (p.Gln317Ter)
NM_007294.3(BRCA1):c.928C>T (p.Gln310Ter)
NM_007294.3(BRCA1):c.5117G>A (p.Gly1706Glu)
NM_007294.3(BRCA1):c.5136G>A (p.Trp1712Ter)
NM_007294.3(BRCA1):c.4327C>T (p.Arg1443Ter)
NM_007294.3(BRCA1):c.1471C>T (p.Gln491Ter)
NM_007294.3(BRCA1):c.1576C>T (p.Gln526Ter)
NM_007294.3(BRCA1):c.160C>T (p.Gln54Ter)
NM_007294.3(BRCA1):c.2683C>T (p.Gln895Ter)
NM_007294.3(BRCA1):c.2761C>T (p.Gln921Ter)
NM_007294.3(BRCA1):c.3895C>T (p.Gln1299Ter)
NM_007294.3(BRCA1):c.4339C>T (p.Gln1447Ter)
NM_007294.3(BRCA1):c.4372C>T (p.Gln1458Ter)
NM_007294.3(BRCA1):c.5153G>A (p.Trp1718Ter)
NM_007294.3(BRCA1):c.5445G>A (p.Trp1815Ter)
NM_007294.3(BRCA1):c.5510G>A (p.Trp1837Ter)
NM_007294.3(BRCA1):c.5346G>A (p.Trp1782Ter)
NM_007294.3(BRCA1):c.1116G>A (p.Trp372Ter)
NM_007294.3(BRCA1):c.1999C>T (p.Gln667Ter)
NM_007294.3(BRCA1):c.4183C>T (p.Gln1395Ter)
NM_007294.3(BRCA1):c.4810C>T (p.Gln1604Ter)
NM_007294.3(BRCA1):c.850C>T (p.Gln284Ter)
NM_007294.3(BRCA1):c.1058G>A (p.Trp353Ter)
NM_007294.3(BRCA1):c.131G>A (p.Cys44Tyr)
NM_007294.3(BRCA1):c.1600C>T (p.Gln534Ter)
NM_007294.3(BRCA1):c.3286C>T (p.Gln1096Ter)
NM_007294.3(BRCA1):c.3403C>T (p.Gln1135Ter)
NM_007294.3(BRCA1):c.34C>T (p.Gln12Ter)
NM_007294.3(BRCA1):c.4258C>T (p.Gln1420Ter)

[1171]

NM_007294.3(BRCA1):c.4609C>T (p.Gln1537Ter)
NM_007294.3(BRCA1):c.5154G>A (p.Trp1718Ter)
NM_007294.3(BRCA1):c.5431C>T (p.Gln1811Ter)
NM_007294.3(BRCA1):c.241C>T (p.Gln81Ter)
NM_007294.3(BRCA1):c.3331C>T (p.Gln1111Ter)
NM_007294.3(BRCA1):c.3967C>T (p.Gln1323Ter)
NM_007294.3(BRCA1):c.415C>T (p.Gln139Ter)
NM_007294.3(BRCA1):c.505C>T (p.Gln169Ter)
NM_007294.3(BRCA1):c.5194-12G>A
NM_007294.3(BRCA1):c.5212G>A (p.Gly1738Arg)
NM_007294.3(BRCA1):c.5332+1G>A
NM_007294.3(BRCA1):c.1480C>T (p.Gln494Ter)
NM_007294.3(BRCA1):c.2563C>T (p.Gln855Ter)
NM_007294.3(BRCA1):c.1066C>T (p.Gln356Ter)
NM_007294.3(BRCA1):c.3718C>T (p.Gln1240Ter)
NM_007294.3(BRCA1):c.3817C>T (p.Gln1273Ter)
NM_007294.3(BRCA1):c.3937C>T (p.Gln1313Ter)
NM_007294.3(BRCA1):c.4357+1G>A
NM_007294.3(BRCA1):c.5074+1G>A
NM_007294.3(BRCA1):c.5277+1G>A
NM_007294.3(BRCA1):c.2338C>T (p.Gln780Ter)
NM_007294.3(BRCA1):c.3598C>T (p.Gln1200Ter)
NM_007294.3(BRCA1):c.3841C>T (p.Gln1281Ter)
NM_007294.3(BRCA1):c.4222C>T (p.Gln1408Ter)
NM_007294.3(BRCA1):c.4524G>A (p.Trp1508Ter)
NM_007294.3(BRCA1):c.5353C>T (p.Gln1785Ter)
NM_007294.3(BRCA1):c.962G>A (p.Trp321Ter)
NM_007294.3(BRCA1):c.220C>T (p.Gln74Ter)
NM_007294.3(BRCA1):c.2713C>T (p.Gln905Ter)
NM_007294.3(BRCA1):c.2800C>T (p.Gln934Ter)
NM_007294.3(BRCA1):c.4612C>T (p.Gln1538Ter)
NM_007294.3(BRCA1):c.3352C>T (p.Gln1118Ter)
NM_007294.3(BRCA1):c.4834C>T (p.Gln1612Ter)

[1172]

NM_007294.3(BRCA1):c.4523G>A (p.Trp1508Ter)
NM_007294.3(BRCA1):c.5135G>A (p.Trp1712Ter)
NM_007294.3(BRCA1):c.1155G>A (p.Trp385Ter)
NM_007294.3(BRCA1):c.4987-1G>A
NM_000059.3(BRCA2):c.9573G>A (p.Trp3191Ter)
NM_000059.3(BRCA2):c.1945C>T (p.Gln649Ter)
NM_000059.3(BRCA2):c.217C>T (p.Gln73Ter)
NM_000059.3(BRCA2):c.523C>T (p.Gln175Ter)
NM_000059.3(BRCA2):c.2548C>T (p.Gln850Ter)
NM_000059.3(BRCA2):c.2905C>T (p.Gln969Ter)
NM_000059.3(BRCA2):c.4689G>A (p.Trp1563Ter)
NM_000059.3(BRCA2):c.4972C>T (p.Gln1658Ter)
NM_000059.3(BRCA2):c.1184G>A (p.Trp395Ter)
NM_000059.3(BRCA2):c.2137C>T (p.Gln713Ter)
NM_000059.3(BRCA2):c.3217C>T (p.Gln1073Ter)
NM_000059.3(BRCA2):c.3523C>T (p.Gln1175Ter)
NM_000059.3(BRCA2):c.4783C>T (p.Gln1595Ter)
NM_000059.3(BRCA2):c.5800C>T (p.Gln1934Ter)
NM_000059.3(BRCA2):c.6478C>T (p.Gln2160Ter)
NM_000059.3(BRCA2):c.7033C>T (p.Gln2345Ter)
NM_000059.3(BRCA2):c.7495C>T (p.Gln2499Ter)
NM_000059.3(BRCA2):c.7501C>T (p.Gln2501Ter)
NM_000059.3(BRCA2):c.7887G>A (p.Trp2629Ter)
NM_000059.3(BRCA2):c.8910G>A (p.Trp2970Ter)
NM_000059.3(BRCA2):c.9139C>T (p.Gln3047Ter)
NM_000059.3(BRCA2):c.9739C>T (p.Gln3247Ter)
NM_000059.3(BRCA2):c.582G>A (p.Trp194Ter)
NM_000059.3(BRCA2):c.7963C>T (p.Gln2655Ter)
NM_000059.3(BRCA2):c.8695C>T (p.Gln2899Ter)
NM_000059.3(BRCA2):c.8869C>T (p.Gln2957Ter)
NM_000059.3(BRCA2):c.1117C>T (p.Gln373Ter)
NM_000059.3(BRCA2):c.1825C>T (p.Gln609Ter)
NM_000059.3(BRCA2):c.2455C>T (p.Gln819Ter)

[1173]

NM_000059.3(BRCA2):c.2881C>T (p.Gln961Ter)
NM_000059.3(BRCA2):c.3265C>T (p.Gln1089Ter)
NM_000059.3(BRCA2):c.3283C>T (p.Gln1095Ter)
NM_000059.3(BRCA2):c.3442C>T (p.Gln1148Ter)
NM_000059.3(BRCA2):c.3871C>T (p.Gln1291Ter)
NM_000059.3(BRCA2):c.439C>T (p.Gln147Ter)
NM_000059.3(BRCA2):c.4525C>T (p.Gln1509Ter)
NM_000059.3(BRCA2):c.475+1G>A
NM_000059.3(BRCA2):c.5344C>T (p.Gln1782Ter)
NM_000059.3(BRCA2):c.5404C>T (p.Gln1802Ter)
NM_000059.3(BRCA2):c.5773C>T (p.Gln1925Ter)
NM_000059.3(BRCA2):c.5992C>T (p.Gln1998Ter)
NM_000059.3(BRCA2):c.6469C>T (p.Gln2157Ter)
NM_000059.3(BRCA2):c.7261C>T (p.Gln2421Ter)
NM_000059.3(BRCA2):c.7303C>T (p.Gln2435Ter)
NM_000059.3(BRCA2):c.7471C>T (p.Gln2491Ter)
NM_000059.3(BRCA2):c.7681C>T (p.Gln2561Ter)
NM_000059.3(BRCA2):c.7738C>T (p.Gln2580Ter)
NM_000059.3(BRCA2):c.7886G>A (p.Trp2629Ter)
NM_000059.3(BRCA2):c.8140C>T (p.Gln2714Ter)
NM_000059.3(BRCA2):c.8363G>A (p.Trp2788Ter)
NM_000059.3(BRCA2):c.8572C>T (p.Gln2858Ter)
NM_000059.3(BRCA2):c.8773C>T (p.Gln2925Ter)
NM_000059.3(BRCA2):c.8821C>T (p.Gln2941Ter)
NM_000059.3(BRCA2):c.9109C>T (p.Gln3037Ter)
NM_000059.3(BRCA2):c.9317G>A (p.Trp3106Ter)
NM_000059.3(BRCA2):c.9466C>T (p.Gln3156Ter)
NM_000059.3(BRCA2):c.9572G>A (p.Trp3191Ter)
NM_000059.3(BRCA2):c.8490G>A (p.Trp2830Ter)
NM_000059.3(BRCA2):c.5980C>T (p.Gln1994Ter)
NM_000059.3(BRCA2):c.7721G>A (p.Trp2574Ter)
NM_000059.3(BRCA2):c.196C>T (p.Gln66Ter)
NM_000059.3(BRCA2):c.7618-1G>A

[1174]

NM_000059.3(BRCA2):c.8489G>A (p.Trp2830Ter)
NM_000059.3(BRCA2):c.7857G>A (p.Trp2619Ter)
NM_000059.3(BRCA2):c.1261C>T (p.Gln421Ter)
NM_000059.3(BRCA2):c.1456C>T (p.Gln486Ter)
NM_000059.3(BRCA2):c.3319C>T (p.Gln1107Ter)
NM_000059.3(BRCA2):c.5791C>T (p.Gln1931Ter)
NM_000059.3(BRCA2):c.6070C>T (p.Gln2024Ter)
NM_000059.3(BRCA2):c.7024C>T (p.Gln2342Ter)
NM_000059.3(BRCA2):c.961C>T (p.Gln321Ter)
NM_000059.3(BRCA2):c.9380G>A (p.Trp3127Ter)
NM_000059.3(BRCA2):c.8364G>A (p.Trp2788Ter)
NM_000059.3(BRCA2):c.7758G>A (p.Trp2586Ter)
NM_000059.3(BRCA2):c.2224C>T (p.Gln742Ter)
NM_000059.3(BRCA2):c.5101C>T (p.Gln1701Ter)
NM_000059.3(BRCA2):c.5959C>T (p.Gln1987Ter)
NM_000059.3(BRCA2):c.7060C>T (p.Gln2354Ter)
NM_000059.3(BRCA2):c.9100C>T (p.Gln3034Ter)
NM_000059.3(BRCA2):c.9148C>T (p.Gln3050Ter)
NM_000059.3(BRCA2):c.9883C>T (p.Gln3295Ter)
NM_000059.3(BRCA2):c.1414C>T (p.Gln472Ter)
NM_000059.3(BRCA2):c.1689G>A (p.Trp563Ter)
NM_000059.3(BRCA2):c.581G>A (p.Trp194Ter)
NM_000059.3(BRCA2):c.6490C>T (p.Gln2164Ter)
NM_000059.3(BRCA2):c.7856G>A (p.Trp2619Ter)
NM_000059.3(BRCA2):c.8970G>A (p.Trp2990Ter)
NM_000059.3(BRCA2):c.92G>A (p.Trp31Ter)
NM_000059.3(BRCA2):c.9376C>T (p.Gln3126Ter)
NM_000059.3(BRCA2):c.93G>A (p.Trp31Ter)
NM_000059.3(BRCA2):c.1189C>T (p.Gln397Ter)
NM_000059.3(BRCA2):c.2818C>T (p.Gln940Ter)
NM_000059.3(BRCA2):c.2979G>A (p.Trp993Ter)
NM_000059.3(BRCA2):c.3166C>T (p.Gln1056Ter)
NM_000059.3(BRCA2):c.4285C>T (p.Gln1429Ter)

[1175]

NM_000059.3(BRCA2):c.6025C>T (p.Gln2009Ter)
NM_000059.3(BRCA2):c.772C>T (p.Gln258Ter)
NM_000059.3(BRCA2):c.7877G>A (p.Trp2626Ter)
NM_000059.3(BRCA2):c.3109C>T (p.Gln1037Ter)
NM_000059.3(BRCA2):c.4222C>T (p.Gln1408Ter)
NM_000059.3(BRCA2):c.7480C>T (p.Arg2494Ter)
NM_000059.3(BRCA2):c.7878G>A (p.Trp2626Ter)
NM_000059.3(BRCA2):c.9076C>T (p.Gln3026Ter)
NM_000059.3(BRCA2):c.1855C>T (p.Gln619Ter)
NM_000059.3(BRCA2):c.4111C>T (p.Gln1371Ter)
NM_000059.3(BRCA2):c.5656C>T (p.Gln1886Ter)
NM_000059.3(BRCA2):c.7757G>A (p.Trp2586Ter)
NM_000059.3(BRCA2):c.8243G>A (p.Gly2748Asp)
NM_000059.3(BRCA2):c.8878C>T (p.Gln2960Ter)
NM_000059.3(BRCA2):c.8487+1G>A
NM_000059.3(BRCA2):c.8677C>T (p.Gln2893Ter)
NM_000059.3(BRCA2):c.250C>T (p.Gln84Ter)
NM_000059.3(BRCA2):c.6124C>T (p.Gln2042Ter)
NM_000059.3(BRCA2):c.7617+1G>A
NM_000059.3(BRCA2):c.8575C>T (p.Gln2859Ter)
NM_000059.3(BRCA2):c.8174G>A (p.Trp2725Ter)
NM_000059.3(BRCA2):c.3187C>T (p.Gln1063Ter)
NM_000059.3(BRCA2):c.9381G>A (p.Trp3127Ter)
NM_000059.3(BRCA2):c.2095C>T (p.Gln699Ter)
NM_000059.3(BRCA2):c.1642C>T (p.Gln548Ter)
NM_000059.3(BRCA2):c.8608C>T (p.Gln2870Ter)
NM_000059.3(BRCA2):c.3412C>T (p.Gln1138Ter)
NM_000059.3(BRCA2):c.4246C>T (p.Gln1416Ter)
NM_000059.3(BRCA2):c.6475C>T (p.Gln2159Ter)
NM_000059.3(BRCA2):c.7366C>T (p.Gln2456Ter)
NM_000059.3(BRCA2):c.7516C>T (p.Gln2506Ter)
NM_000059.3(BRCA2):c.8969G>A (p.Trp2990Ter)
NM_000059.3(BRCA2):c.6487C>T (p.Gln2163Ter)

[1176]

NM_000059.3(BRCA2):c.2978G>A (p.Trp993Ter)
NM_000059.3(BRCA2):c.7615C>T (p.Gln2539Ter)
NM_000059.3(BRCA2):c.9106C>T (p.Gln3036Ter)

[1177]

[1178] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 BRCA1 또는 BRCA2 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 유방-난소 암을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1179] **린치 증후군**

[1180] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 린치 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하

기를 포함하여, MSH6, MSH2, EPCAM, PMS2, 및 MLH1로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000179.2(MSH6):c.1045C>T (p.Gln349Ter)
NM_000251.2(MSH2):c.1384C>T (p.Gln462Ter)
NM_002354.2(EPCAM):c.133C>T (p.Gln45Ter)
NM_002354.2(EPCAM):c.429G>A (p.Trp143Ter)
NM_002354.2(EPCAM):c.523C>T (p.Gln175Ter)
NM_000179.2(MSH6):c.2680C>T (p.Gln894Ter)
NM_000251.2(MSH2):c.350G>A (p.Trp117Ter)
NM_000179.2(MSH6):c.2735G>A (p.Trp912Ter)
NM_000179.2(MSH6):c.3556+1G>A
NM_000251.2(MSH2):c.388C>T (p.Gln130Ter)
NM_000535.6(PMS2):c.1912C>T (p.Gln638Ter)
NM_000535.6(PMS2):c.1891C>T (p.Gln631Ter)
NM_000249.3(MLH1):c.454-1G>A
NM_000251.2(MSH2):c.1030C>T (p.Gln344Ter)
NM_000179.2(MSH6):c.2330G>A (p.Trp777Ter)
NM_000179.2(MSH6):c.2191C>T (p.Gln731Ter)
NM_000179.2(MSH6):c.2764C>T (p.Arg922Ter)
NM_000179.2(MSH6):c.2815C>T (p.Gln939Ter)
NM_000179.2(MSH6):c.3020G>A (p.Trp1007Ter)

[1181]

NM_000179.2(MSH6):c.3436C>T (p.Gln1146Ter)
NM_000179.2(MSH6):c.3647-1G>A
NM_000179.2(MSH6):c.3772C>T (p.Gln1258Ter)
NM_000179.2(MSH6):c.3838C>T (p.Gln1280Ter)
NM_000179.2(MSH6):c.706C>T (p.Gln236Ter)
NM_000179.2(MSH6):c.730C>T (p.Gln244Ter)
NM_000249.3(MLH1):c.1171C>T (p.Gln391Ter)
NM_000249.3(MLH1):c.1192C>T (p.Gln398Ter)
NM_000249.3(MLH1):c.1225C>T (p.Gln409Ter)
NM_000249.3(MLH1):c.1276C>T (p.Gln426Ter)
NM_000249.3(MLH1):c.1528C>T (p.Gln510Ter)
NM_000249.3(MLH1):c.1609C>T (p.Gln537Ter)
NM_000249.3(MLH1):c.1613G>A (p.Trp538Ter)
NM_000249.3(MLH1):c.1614G>A (p.Trp538Ter)
NM_000249.3(MLH1):c.1624C>T (p.Gln542Ter)
NM_000249.3(MLH1):c.1684C>T (p.Gln562Ter)
NM_000249.3(MLH1):c.1731+1G>A
NM_000249.3(MLH1):c.1731+5G>A
NM_000249.3(MLH1):c.1732-1G>A
NM_000249.3(MLH1):c.1896G>A (p.Glu632=)
NM_000249.3(MLH1):c.1989+1G>A
NM_000249.3(MLH1):c.1990-1G>A
NM_000249.3(MLH1):c.1998G>A (p.Trp666Ter)
NM_000249.3(MLH1):c.208-1G>A
NM_000249.3(MLH1):c.2101C>T (p.Gln701Ter)
NM_000249.3(MLH1):c.2136G>A (p.Trp712Ter)
NM_000249.3(MLH1):c.2224C>T (p.Gln742Ter)
NM_000249.3(MLH1):c.230G>A (p.Cys77Tyr)
NM_000249.3(MLH1):c.256C>T (p.Gln86Ter)
NM_000249.3(MLH1):c.436C>T (p.Gln146Ter)
NM_000249.3(MLH1):c.445C>T (p.Gln149Ter)
NM_000249.3(MLH1):c.545G>A (p.Arg182Lys)
NM_000249.3(MLH1):c.731G>A (p.Gly244Asp)

[1182]

NM_000249.3(MLH1):c.76C>T (p.Gln26Ter)
NM_000249.3(MLH1):c.842C>T (p.Ala281Val)
NM_000249.3(MLH1):c.882C>T (p.Leu294=)
NM_000249.3(MLH1):c.901C>T (p.Gln301Ter)
NM_000251.2(MSH2):c.1013G>A (p.Gly338Glu)
NM_000251.2(MSH2):c.1034G>A (p.Trp345Ter)
NM_000251.2(MSH2):c.1129C>T (p.Gln377Ter)
NM_000251.2(MSH2):c.1183C>T (p.Gln395Ter)
NM_000251.2(MSH2):c.1189C>T (p.Gln397Ter)
NM_000251.2(MSH2):c.1204C>T (p.Gln402Ter)
NM_000251.2(MSH2):c.1276+1G>A
NM_000251.2(MSH2):c.1528C>T (p.Gln510Ter)
NM_000251.2(MSH2):c.1552C>T (p.Gln518Ter)
NM_000251.2(MSH2):c.1720C>T (p.Gln574Ter)
NM_000251.2(MSH2):c.1777C>T (p.Gln593Ter)
NM_000251.2(MSH2):c.1885C>T (p.Gln629Ter)
NM_000251.2(MSH2):c.2087C>T (p.Pro696Leu)
NM_000251.2(MSH2):c.2251G>A (p.Gly751Arg)
NM_000251.2(MSH2):c.2291G>A (p.Trp764Ter)
NM_000251.2(MSH2):c.2292G>A (p.Trp764Ter)
NM_000251.2(MSH2):c.2446C>T (p.Gln816Ter)
NM_000251.2(MSH2):c.2470C>T (p.Gln824Ter)
NM_000251.2(MSH2):c.2536C>T (p.Gln846Ter)
NM_000251.2(MSH2):c.2581C>T (p.Gln861Ter)
NM_000251.2(MSH2):c.2634G>A (p.Glu878=)
NM_000251.2(MSH2):c.2635C>T (p.Gln879Ter)
NM_000251.2(MSH2):c.28C>T (p.Gln10Ter)
NM_000251.2(MSH2):c.472C>T (p.Gln158Ter)
NM_000251.2(MSH2):c.478C>T (p.Gln160Ter)
NM_000251.2(MSH2):c.484G>A (p.Gly162Arg)
NM_000251.2(MSH2):c.490G>A (p.Gly164Arg)
NM_000251.2(MSH2):c.547C>T (p.Gln183Ter)
NM_000251.2(MSH2):c.577C>T (p.Gln193Ter)

[1183]

NM_000251.2(MSH2):c.643C>T (p.Gln215Ter)
NM_000251.2(MSH2):c.645+1G>A
NM_000251.2(MSH2):c.652C>T (p.Gln218Ter)
NM_000251.2(MSH2):c.754C>T (p.Gln252Ter)
NM_000251.2(MSH2):c.792+1G>A
NM_000251.2(MSH2):c.942G>A (p.Gln314=)
NM_000535.6(PMS2):c.949C>T (p.Gln317Ter)
NM_000249.3(MLH1):c.306+1G>A
NM_000249.3(MLH1):c.62C>T (p.Ala21Val)
NM_000251.2(MSH2):c.1865C>T (p.Pro622Leu)
NM_000179.2(MSH6):c.426G>A (p.Trp142Ter)
NM_000251.2(MSH2):c.715C>T (p.Gln239Ter)
NM_000249.3(MLH1):c.350C>T (p.Thr117Met)
NM_000251.2(MSH2):c.1915C>T (p.His639Tyr)
NM_000251.2(MSH2):c.289C>T (p.Gln97Ter)
NM_000251.2(MSH2):c.2785C>T (p.Arg929Ter)
NM_000249.3(MLH1):c.131C>T (p.Ser44Phe)
NM_000249.3(MLH1):c.1219C>T (p.Gln407Ter)
NM_000249.3(MLH1):c.306+5G>A
NM_000251.2(MSH2):c.1801C>T (p.Gln601Ter)
NM_000535.6(PMS2):c.1144+1G>A
NM_000251.2(MSH2):c.1984C>T (p.Gln662Ter)
NM_000249.3(MLH1):c.381-1G>A
NM_000535.6(PMS2):c.631C>T (p.Arg211Ter)
NM_000251.2(MSH2):c.790C>T (p.Gln264Ter)
NM_000251.2(MSH2):c.366+1G>A
NM_000249.3(MLH1):c.298C>T (p.Arg100Ter)
NM_000179.2(MSH6):c.3013C>T (p.Arg1005Ter)
NM_000179.2(MSH6):c.694C>T (p.Gln232Ter)
NM_000179.2(MSH6):c.742C>T (p.Arg248Ter)
NM_000249.3(MLH1):c.1039-1G>A
NM_000249.3(MLH1):c.142C>T (p.Gln48Ter)
NM_000249.3(MLH1):c.1790G>A (p.Trp597Ter)

[1184]

NM_000249.3(MLH1):c.1961C>T (p.Pro654Leu)
NM_000249.3(MLH1):c.2103+1G>A
NM_000249.3(MLH1):c.2135G>A (p.Trp712Ter)
NM_000249.3(MLH1):c.588+5G>A
NM_000249.3(MLH1):c.790+1G>A
NM_000251.2(MSH2):c.1035G>A (p.Trp345Ter)
NM_000251.2(MSH2):c.1255C>T (p.Gln419Ter)
NM_000251.2(MSH2):c.1861C>T (p.Arg621Ter)
NM_000251.2(MSH2):c.226C>T (p.Gln76Ter)
NM_000251.2(MSH2):c.2653C>T (p.Gln885Ter)
NM_000251.2(MSH2):c.508C>T (p.Gln170Ter)
NM_000251.2(MSH2):c.862C>T (p.Gln288Ter)
NM_000251.2(MSH2):c.892C>T (p.Gln298Ter)
NM_000251.2(MSH2):c.970C>T (p.Gln324Ter)
NM_000179.2(MSH6):c.4001G>A (p.Arg1334Gln)
NM_000251.2(MSH2):c.1662-1G>A
NM_000535.6(PMS2):c.1882C>T (p.Arg628Ter)
NM_000535.6(PMS2):c.2174+1G>A
NM_000535.6(PMS2):c.2404C>T (p.Arg802Ter)
NM_000179.2(MSH6):c.3991C>T (p.Arg1331Ter)
NM_000179.2(MSH6):c.2503C>T (p.Gln835Ter)
NM_000179.2(MSH6):c.718C>T (p.Arg240Ter)
NM_000249.3(MLH1):c.1038G>A (p.Gln346=)
NM_000249.3(MLH1):c.245C>T (p.Thr82Ile)
NM_000249.3(MLH1):c.83C>T (p.Pro28Leu)
NM_000249.3(MLH1):c.884G>A (p.Ser295Asn)
NM_000249.3(MLH1):c.982C>T (p.Gln328Ter)
NM_000251.2(MSH2):c.1046C>T (p.Pro349Leu)
NM_000251.2(MSH2):c.1120C>T (p.Gln374Ter)
NM_000251.2(MSH2):c.1285C>T (p.Gln429Ter)
NM_000251.2(MSH2):c.1477C>T (p.Gln493Ter)
NM_000251.2(MSH2):c.2152C>T (p.Gln718Ter)
NM_000535.6(PMS2):c.703C>T (p.Gln235Ter)

[1185]

NM_000249.3(MLH1):c.2141G>A (p.Trp714Ter)
NM_000251.2(MSH2):c.1009C>T (p.Gln337Ter)
NM_000251.2(MSH2):c.1216C>T (p.Arg406Ter)
NM_000179.2(MSH6):c.3202C>T (p.Arg1068Ter)
NM_000251.2(MSH2):c.1165C>T (p.Arg389Ter)
NM_000249.3(MLH1):c.1943C>T (p.Pro648Leu)
NM_000249.3(MLH1):c.200G>A (p.Gly67Glu)
NM_000249.3(MLH1):c.793C>T (p.Arg265Cys)
NM_000249.3(MLH1):c.2059C>T (p.Arg687Trp)
NM_000249.3(MLH1):c.677G>A (p.Arg226Gln)
NM_000249.3(MLH1):c.2041G>A (p.Ala681Thr)
NM_000249.3(MLH1):c.1942C>T (p.Pro648Ser)
NM_000249.3(MLH1):c.676C>T (p.Arg226Ter)
NM_000251.2(MSH2):c.2038C>T (p.Arg680Ter)
NM_000179.2(MSH6):c.1483C>T (p.Arg495Ter)
NM_000179.2(MSH6):c.2194C>T (p.Arg732Ter)
NM_000179.2(MSH6):c.3103C>T (p.Arg1035Ter)
NM_000179.2(MSH6):c.892C>T (p.Arg298Ter)
NM_000249.3(MLH1):c.1459C>T (p.Arg487Ter)
NM_000249.3(MLH1):c.1731G>A (p.Ser577=)
NM_000249.3(MLH1):c.184C>T (p.Gln62Ter)
NM_000249.3(MLH1):c.1975C>T (p.Arg659Ter)
NM_000249.3(MLH1):c.199G>A (p.Gly67Arg)
NM_000251.2(MSH2):c.1076+1G>A
NM_000251.2(MSH2):c.1147C>T (p.Arg383Ter)
NM_000251.2(MSH2):c.181C>T (p.Gln61Ter)
NM_000251.2(MSH2):c.212-1G>A
NM_000251.2(MSH2):c.2131C>T (p.Arg711Ter)
NM_000535.6(PMS2):c.697C>T (p.Gln233Ter)
NM_000535.6(PMS2):c.1261C>T (p.Arg421Ter)
NM_000251.2(MSH2):c.2047G>A (p.Gly683Arg)
NM_000535.6(PMS2):c.400C>T (p.Arg134Ter)
NM_000535.6(PMS2):c.1927C>T (p.Gln643Ter)

[1186]

NM_000179.2(MSH6):c.1444C>T (p.Arg482Ter)
NM_000179.2(MSH6):c.2731C>T (p.Arg911Ter)
NM_000535.6(PMS2):c.943C>T (p.Arg315Ter)

[1187]

[1188] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 BCKDHA, BCKDHB, DBT, 및 DLD로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 린치 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1189] **기타 유전 질환**

[1190] 제한없이, 마르판 증후군, 후를러 증후군, 당원 축적병, 및 낭성 섬유증을 포함한 추가적인 유전자 질환과 연관된 병원성 G>A 또는 C>T 돌연변이/SNP는 또한 ClinVar 데이터베이스에 보고되어 있고, 표 A에 개시되어 있다.

따라서, 본 발명의 일 양상은 하기에 논의되는 바와 같이, 임의의 이들 질환과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하기 위한 방법에 관한 것이다.

[1191] **마르판 증후군**

[1192] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 마르판 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 FBN1 유전자에 존재한다:

NM_000138.4(FBN1):c.1879C>T (p.Arg627Cys)

NM_000138.4(FBN1):c.1051C>T (p.Gln351Ter)

NM_000138.4(FBN1):c.184C>T (p.Arg62Cys)

NM_000138.4(FBN1):c.2855-1G>A

NM_000138.4(FBN1):c.3164G>A (p.Cys1055Tyr)

NM_000138.4(FBN1):c.368G>A (p.Cys123Tyr)

NM_000138.4(FBN1):c.4955G>A (p.Cys1652Tyr)

NM_000138.4(FBN1):c.7180C>T (p.Arg2394Ter)

NM_000138.4(FBN1):c.8267G>A (p.Trp2756Ter)

NM_000138.4(FBN1):c.1496G>A (p.Cys499Tyr)

NM_000138.4(FBN1):c.6886C>T (p.Gln2296Ter)

NM_000138.4(FBN1):c.3373C>T (p.Arg1125Ter)

[1193] NM_000138.4(FBN1):c.640G>A (p.Gly214Ser)

NM_000138.4(FBN1):c.5038C>T (p.Gln1680Ter)
NM_000138.4(FBN1):c.434G>A (p.Cys145Tyr)
NM_000138.4(FBN1):c.2563C>T (p.Gln855Ter)
NM_000138.4(FBN1):c.7466G>A (p.Cys2489Tyr)
NM_000138.4(FBN1):c.2089C>T (p.Gln697Ter)
NM_000138.4(FBN1):c.592C>T (p.Gln198Ter)
NM_000138.4(FBN1):c.6695G>A (p.Cys2232Tyr)
NM_000138.4(FBN1):c.6164-1G>A
NM_000138.4(FBN1):c.5627G>A (p.Cys1876Tyr)
NM_000138.4(FBN1):c.4061G>A (p.Trp1354Ter)
NM_000138.4(FBN1):c.1982G>A (p.Cys661Tyr)
NM_000138.4(FBN1):c.6784C>T (p.Gln2262Ter)
NM_000138.4(FBN1):c.409C>T (p.Gln137Ter)
NM_000138.4(FBN1):c.364C>T (p.Arg122Cys)
NM_000138.4(FBN1):c.3217G>A (p.Glu1073Lys)
NM_000138.4(FBN1):c.4460-8G>A
NM_000138.4(FBN1):c.4786C>T (p.Arg1596Ter)
NM_000138.4(FBN1):c.7806G>A (p.Trp2602Ter)
NM_000138.4(FBN1):c.247+1G>A
NM_000138.4(FBN1):c.2495G>A (p.Cys832Tyr)
NM_000138.4(FBN1):c.493C>T (p.Arg165Ter)
NM_000138.4(FBN1):c.5504G>A (p.Cys1835Tyr)
NM_000138.4(FBN1):c.5863C>T (p.Gln1955Ter)
NM_000138.4(FBN1):c.6658C>T (p.Arg2220Ter)
NM_000138.4(FBN1):c.7606G>A (p.Gly2536Arg)
NM_000138.4(FBN1):c.7955G>A (p.Cys2652Tyr)
NM_000138.4(FBN1):c.3037G>A (p.Gly1013Arg)
NM_000138.4(FBN1):c.8080C>T (p.Arg2694Ter)
NM_000138.4(FBN1):c.1633C>T (p.Arg545Cys)
NM_000138.4(FBN1):c.7205-1G>A
NM_000138.4(FBN1):c.4621C>T (p.Arg1541Ter)
NM_000138.4(FBN1):c.1090C>T (p.Arg364Ter)
NM_000138.4(FBN1):c.1585C>T (p.Arg529Ter)

[1194]

NM_000138.4(FBN1):c.4781G>A (p.Gly1594Asp)

NM_000138.4(FBN1):c.643C>T (p.Arg215Ter)

NM_000138.4(FBN1):c.3668G>A (p.Cys1223Tyr)

NM_000138.4(FBN1):c.8326C>T (p.Arg2776Ter)

NM_000138.4(FBN1):c.6354C>T (p.Ile2118=)

NM_000138.4(FBN1):c.1468+5G>A

NM_000138.4(FBN1):c.1546C>T (p.Arg516Ter)

NM_000138.4(FBN1):c.4615C>T (p.Arg1539Ter)

NM_000138.4(FBN1):c.5368C>T (p.Arg1790Ter)

NM_000138.4(FBN1):c.1285C>T (p.Arg429Ter)

[1195]

[1196] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 FBN1 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 마르판 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1197] **후를러 증후군**

[1198] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 후를러 증후군과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, 적어도 IDUA 유전자에 존재한다:

NM_000203.4(IDUA):c.972+1G>A

NM_000203.4(IDUA):c.1855C>T (p.Arg619Ter)

NM_000203.4(IDUA):c.152G>A (p.Gly51Asp)

NM_000203.4(IDUA):c.1205G>A (p.Trp402Ter)

NM_000203.4(IDUA):c.208C>T (p.Gln70Ter)

NM_000203.4(IDUA):c.1045G>A (p.Asp349Asn)

NM_000203.4(IDUA):c.1650+5G>A

[1199]

[1200] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 IDUA 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 후를러 증후군을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1201] **당원 축적병**

[1202] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 당원 축적병과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, GAA, AGL, PHKB, PRKAG2, G6PC, PGAM2, GBE1, PYGM, 및 PFKM로부터 선택되는 적어도 하나의 유전자에 존재한다:

NM_000152.4(GAA):c.1927G>A (p.Gly643Arg)
NM_000152.4(GAA):c.2173C>T (p.Arg725Trp)
NM_000642.2(AGL):c.3980G>A (p.Trp1327Ter)
NM_000642.2(AGL):c.16C>T (p.Gln6Ter)
NM_000642.2(AGL):c.2039G>A (p.Trp680Ter)
NM_000293.2(PHKB):c.1546C>T (p.Gln516Ter)
NM_016203.3(PRKAG2):c.1592G>A (p.Arg531Gln)
NM_000151.3(G6PC):c.248G>A (p.Arg83His)
NM_000151.3(G6PC):c.724C>T (p.Gln242Ter)
NM_000151.3(G6PC):c.883C>T (p.Arg295Cys)
NM_000151.3(G6PC):c.247C>T (p.Arg83Cys)
NM_000151.3(G6PC):c.1039C>T (p.Gln347Ter)
NM_000152.4(GAA):c.1561G>A (p.Glu521Lys)
NM_000642.2(AGL):c.2590C>T (p.Arg864Ter)
NM_000642.2(AGL):c.3682C>T (p.Arg1228Ter)
NM_000642.2(AGL):c.118C>T (p.Gln40Ter)
NM_000642.2(AGL):c.256C>T (p.Gln86Ter)
NM_000642.2(AGL):c.2681+1G>A
NM_000642.2(AGL):c.2158-1G>A
NM_000290.3(PGAM2):c.233G>A (p.Trp78Ter)
NM_000152.4(GAA):c.1548G>A (p.Trp516Ter)
NM_000152.4(GAA):c.2014C>T (p.Arg672Trp)
NM_000152.4(GAA):c.546G>A (p.Thr182=)
NM_000152.4(GAA):c.1802C>T (p.Ser601Leu)
NM_000152.4(GAA):c.1754+1G>A
NM_000152.4(GAA):c.1082C>T (p.Pro361Leu)
NM_000152.4(GAA):c.2560C>T (p.Arg854Ter)

[1203]

NM_000152.4(GAA):c.655G>A (p.Gly219Arg)
NM_000152.4(GAA):c.1933G>A (p.Asp645Asn)
NM_000152.4(GAA):c.1979G>A (p.Arg660His)
NM_000152.4(GAA):c.1465G>A (p.Asp489Asn)
NM_000152.4(GAA):c.2512C>T (p.Gln838Ter)
NM_000158.3(GBE1):c.1543C>T (p.Arg515Cys)
NM_005609.3(PYGM):c.1726C>T (p.Arg576Ter)
NM_005609.3(PYGM):c.1827G>A (p.Lys609=)
NM_005609.3(PYGM):c.148C>T (p.Arg50Ter)
NM_005609.3(PYGM):c.613G>A (p.Gly205Ser)
NM_005609.3(PYGM):c.1366G>A (p.Val456Met)
NM_005609.3(PYGM):c.1768+1G>A
NM_001166686.1(PFKM):c.450+1G>A

[1204]

[1205] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 GAA, AGL, PHKB, PRKAG2, G6PC, PGAM2, GBE1, PYGM, 및 PFKM로부터 선택되는 적어도 하나의 유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 당원 축적병을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1206] **낭성 섬유증**

[1207] 일부 실시형태에서, 본 명세서에 기술된 방법, 시스템 및 조성물은 낭성 섬유증과 연관된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하는데 사용된다. 일부 실시형태에서, 병원성 돌연변이/SNP는 적어도 하기를 포함하여, CFTR 유전자에 존재한다:

NM_000492.3(CFTR):c.3712C>T (p.Gln1238Ter)

NM_000492.3(CFTR):c.3484C>T (p.Arg1162Ter)

NM_000492.3(CFTR):c.1766+1G>A

NM_000492.3(CFTR):c.1477C>T (p.Gln493Ter)

NM_000492.3(CFTR):c.2538G>A (p.Trp846Ter)

NM_000492.3(CFTR):c.2551C>T (p.Arg851Ter)

NM_000492.3(CFTR):c.3472C>T (p.Arg1158Ter)

NM_000492.3(CFTR):c.1475C>T (p.Ser492Phe)

NM_000492.3(CFTR):c.1679G>A (p.Arg560Lys)

[1208]

NM_000492.3(CFTR):c.3197G>A (p.Arg1066His)
NM_000492.3(CFTR):c.3873+1G>A
NM_000492.3(CFTR):c.3196C>T (p.Arg1066Cys)
NM_000492.3(CFTR):c.2490+1G>A
NM_000492.3(CFTR):c.3718-1G>A
NM_000492.3(CFTR):c.171G>A (p.Trp57Ter)
NM_000492.3(CFTR):c.3937C>T (p.Gln1313Ter)
NM_000492.3(CFTR):c.274G>A (p.Glu92Lys)
NM_000492.3(CFTR):c.1013C>T (p.Thr338Ile)
NM_000492.3(CFTR):c.3266G>A (p.Trp1089Ter)
NM_000492.3(CFTR):c.1055G>A (p.Arg352Gln)
NM_000492.3(CFTR):c.1654C>T (p.Gln552Ter)
NM_000492.3(CFTR):c.2668C>T (p.Gln890Ter)
NM_000492.3(CFTR):c.3611G>A (p.Trp1204Ter)
NM_000492.3(CFTR):c.1585-8G>A
NM_000492.3(CFTR):c.223C>T (p.Arg75Ter)
NM_000492.3(CFTR):c.1680-1G>A
NM_000492.3(CFTR):c.349C>T (p.Arg117Cys)
NM_000492.3(CFTR):c.1203G>A (p.Trp401Ter)
NM_000492.3(CFTR):c.1240C>T (p.Gln414Ter)
NM_000492.3(CFTR):c.1202G>A (p.Trp401Ter)
NM_000492.3(CFTR):c.1209+1G>A
NM_000492.3(CFTR):c.115C>T (p.Gln39Ter)
NM_000492.3(CFTR):c.1116+1G>A
NM_000492.3(CFTR):c.1393-1G>A
NM_000492.3(CFTR):c.1573C>T (p.Gln525Ter)
NM_000492.3(CFTR):c.164+1G>A
NM_000492.3(CFTR):c.166G>A (p.Glu56Lys)
NM_000492.3(CFTR):c.170G>A (p.Trp57Ter)
NM_000492.3(CFTR):c.2053C>T (p.Gln685Ter)
NM_000492.3(CFTR):c.2125C>T (p.Arg709Ter)
NM_000492.3(CFTR):c.2290C>T (p.Arg764Ter)
NM_000492.3(CFTR):c.2353C>T (p.Arg785Ter)

[1209]

NM_000492.3(CFTR):c.2374C>T (p.Arg792Ter)
NM_000492.3(CFTR):c.2537G>A (p.Trp846Ter)
NM_000492.3(CFTR):c.292C>T (p.Gln98Ter)
NM_000492.3(CFTR):c.2989-1G>A
NM_000492.3(CFTR):c.3293G>A (p.Trp1098Ter)
NM_000492.3(CFTR):c.4144C>T (p.Gln1382Ter)
NM_000492.3(CFTR):c.4231C>T (p.Gln1411Ter)
NM_000492.3(CFTR):c.4234C>T (p.Gln1412Ter)
NM_000492.3(CFTR):c.579+5G>A
NM_000492.3(CFTR):c.595C>T (p.His199Tyr)
NM_000492.3(CFTR):c.613C>T (p.Pro205Ser)
NM_000492.3(CFTR):c.658C>T (p.Gln220Ter)
NM_000492.3(CFTR):c.1117-1G>A
NM_000492.3(CFTR):c.3294G>A (p.Trp1098Ter)
NM_000492.3(CFTR):c.1865G>A (p.Gly622Asp)
NM_000492.3(CFTR):c.743+1G>A
NM_000492.3(CFTR):c.1679+1G>A
NM_000492.3(CFTR):c.1657C>T (p.Arg553Ter)
NM_000492.3(CFTR):c.1675G>A (p.Ala559Thr)
NM_000492.3(CFTR):c.165-1G>A
NM_000492.3(CFTR):c.200C>T (p.Pro67Leu)
NM_000492.3(CFTR):c.2834C>T (p.Ser945Leu)
NM_000492.3(CFTR):c.3846G>A (p.Trp1282Ter)
NM_000492.3(CFTR):c.1652G>A (p.Gly551Asp)
NM_000492.3(CFTR):c.4426C>T (p.Gln1476Ter)
NM_000492.3:c.3718-2477C>T
NM_000492.3(CFTR):c.2988+1G>A
NM_000492.3(CFTR):c.2657+5G>A
NM_000492.3(CFTR):c.2988G>A (p.Gln996=)
NM_000492.3(CFTR):c.274-1G>A
NM_000492.3(CFTR):c.3612G>A (p.Trp1204Ter)
NM_000492.3(CFTR):c.1646G>A (p.Ser549Asn)
NM_000492.3(CFTR):c.3752G>A (p.Ser1251Asn)

[1210]

NM_000492.3(CFTR):c.4046G>A (p.Gly1349Asp)
NM_000492.3(CFTR):c.532G>A (p.Gly178Arg)
NM_000492.3(CFTR):c.3731G>A (p.Gly1244Glu)
NM_000492.3(CFTR):c.1651G>A (p.Gly551Ser)
NM_000492.3(CFTR):c.1585-1G>A
NM_000492.3(CFTR):c.1000C>T (p.Arg334Trp)
NM_000492.3(CFTR):c.254G>A (p.Gly85Glu)
NM_000492.3(CFTR):c.1040G>A (p.Arg347His)
NM_000492.3(CFTR):c.273+1G>A

[1211]

[1212] 표 A를 참조한다. 따라서, 본 발명의 일 양상은 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 특히 CFTR

유전자에 존재하는 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP, 및 보다 특히 상기 기술된 하나 이상의 병원성 G>A 또는 C>T 돌연변이/SNP를 교정하여 낭성 섬유증을 치료 또는 예방하기 위한 방법에 관한 것이다.

[1213] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 2 유방-난소암과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA2 유전자에 위치된다 (HGVS: U43746.1:n.7829+1G>A). 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 2 유방-난소암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1214] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 유전성 인자 IX 결핍증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 F9 유전자에서 GRCh38: ChrX: 139537145에 위치되어 Arg의 Gln으로의 치환을 초래한다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 유전성 인자 IX 결핍증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1215] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 베타-플러스-탈라세미아, 베타 탈라세미아, 및 주요 베타 탈라세미아와 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 HBB 유전자 내 GRCh38: Chr11: 5226820에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 베타-플러스-탈라세미아, 베타 탈라세미아, 및 주요 베타 탈라세미아를 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1216] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 마르판 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 문헌 [Yamamoto *et al*. J Hum Genet. 2000;45(2):115-8]에 보고되는 바와 같이, 병원성 A>G 돌연변이 또는 SNP는 FBN1 유전자(IVS2DS, G-A, +1)에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 마르판 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1217] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 비스코트-알드리히 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 문헌 [Kwan *et al*. (1995)]에 보고되는 바와 같이, WAS 유전자(IVS6AS, G-A, -1)의 인트론 6의 -1번 위치에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 비스코트-알드리히 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1218] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7:117590440에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1219] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증 및 유전성 췌장염과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7:117606754에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증 및 유전성 췌장염을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1220] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7: 117587738에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1221] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 터코트 증후군 및 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MSH2 유전자에서 GRCh38: Chr2:47470964에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 터코트 증후군 및 린치 증후군을 치료하거나

또는 예방하기 위한 방법에 관한 것이다.

[1222] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7: 117642437에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1223] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군 II 및 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MLH1 유전자에서 GRCh38: Chr3:37001058에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군 II 및 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1224] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7: 117642594에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1225] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7: 117592658에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1226] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군, 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43057051에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1227] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 디히드로피리미딘 디히드로게나제 결핍증, 히르슈슈프룽병 1, 플루오로우라실 반응, 피리미딘 유사체 반응 - 독성/ADR, 카펙시타빈 반응 - 독성/ADR, 플루오로우라실 반응 - 독성/ADR, 테가푸르 반응 - 독성/ADR과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 DPYD 유전자에서 GRCh38: Chr1:97450058에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 디히드로피리미딘 디히드로게나제 결핍증, 히르슈슈프룽병 1, 플루오로우라실 반응, 피리미딘 유사체 반응 - 독성/ADR, 카펙시타빈 반응 - 독성/ADR, 플루오로우라실 반응 - 독성/ADR, 테가푸르 반응 - 독성/ADR을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1228] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MSH2 유전자에서 GRCh38: Chr2:47478520에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1229] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MLH1 유전자에서 GRCh38: Chr3:37011819에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1230] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP

는 MLH1 유전자에서 GRCh38: Chr3: 37014545에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1231] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MLH1 유전자에서 GRCh38: Chr3: 37011867에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1232] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MLH1 유전자에서 GRCh38: Chr3: 37025636에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1233] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MLH1 유전자에서 GRCh38: Chr3: 37004475에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1234] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MSH2 유전자에서 GRCh38: Chr2:47416430에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1235] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MSH2 유전자에서 GRCh38: Chr2: 47408400에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1236] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 린치 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MLH1 유전자에서 GRCh38: Chr3:36996710에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 린치 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1237] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43067696에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1238] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 2 유방-난소암 및 유전성 유방 및 난소암 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA2 유전자에서 GRCh38: Chr13:32356610에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 2 유방-난소암 및 유전성 유방 및 난소암 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1239] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 원발성 확장성 심장근육병증 및 원발성 가족성 비후성 심근증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MYH7 유전자에서 GRCh38: Chr14:23419993에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 원발성 확장성

심장근육병증 및 원발성 가족성 비대성 심장근육병증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1240] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 원발성 가족성 비후성 심근증, 캠토코미아(camptocormism), 및 비후성 심근증과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MYH7 유전자에서 GRCh38: Chr14:23415225에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 원발성 가족성 비후성 심근증, 캠토코미아 및 비대성 심장근육병증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1241] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 유방암, 가족성 2 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA2 유전자에서 GRCh38: Chr13:32357741에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 유방암, 가족성 2 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1242] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 원발성 확장성 심장근육병증, 비후성 심근증, 심장근육병증 및 좌심실비움과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 MYH7 유전자에서 GRCh38: Chr14:23431584에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 원발성 확장성 심장근육병증, 비후성 심근증, 심장근육병증 및 좌심실비움을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1243] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43067607에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1244] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군, 유전성 암-소인 증후군 및 유방암과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43047666에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군, 유전성 암-소인 증후군 및 유방암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1245] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 2 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA2 유전자에서 GRCh38: Chr13:32370558에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 2 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1246] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군, 유전성 암-소인 증후군 및 유방 암과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43074330에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군, 유전성 암-소인 증후군 및 유방암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1247] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 A-G (A-to-G) (A>G) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 A>G 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17: 43082403에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 A>G 돌연변이 또는 SNP를 교

정함으로써 가족성 1 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1248] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 낭성 섬유증 및 유전성 췌장염과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 CFTR 유전자에서 GRCh38: Chr7:117639961에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 낭성 섬유증 및 유전성 췌장염을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1249] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 2 유방-난소암과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 BRCA2 유전자에서 GRCh38: Chr13:32336492에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 2 유방-난소암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1250] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43063365에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1251] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 1 유방-난소암과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43093613에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 1 유방-난소암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1252] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 유방암 및 가족성 1 유방-난소암과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 BRCA1 유전자에서 GRCh38: Chr17:43093931에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 유방암 및 가족성 1 유방-난소암을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1253] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 비대성 심장근육병증1, 원발성 가족성 비후성 심근증 및 비후성 심근증과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 MYH7 유전자에서 GRCh38: Chr14:23429279에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 비대성 심장근육병증1, 원발성 가족성 비후성 심근증 및 비대성 심장근육병증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1254] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 2 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP을 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 BRCA2 유전자에서 GRCh38: Chr13:32356472에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 2 유방-난소암, 유전성 유방 및 난소암 증후군 및 유전성 암-소인 증후군을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1255] 일부 실시형태에서, 본 명세서에 기재된 방법, 시스템 및 조성물은 가족성 비대성 심장근육병증1, 원발성 가족성 비후성 심근증, 가족성 제한성 심장근육병증 및 비후성 심근증과 연관되는 것으로 여겨지는 병원성 C-T (C-to-T) (C>T) 돌연변이 또는 SNP를 교정하기 위해 사용되고, 여기서 병원성 C>T 돌연변이 또는 SNP는 MYH7 유전자에서 GRCh38: Chr14:23429005에 위치된다. 따라서, 본 발명의 추가적인 양상은 앞서 언급된 병원성 C>T 돌연변이 또는 SNP를 교정함으로써 가족성 비대성 심장근육병증1, 원발성 가족성 비후성 심근증, 가족성 제한성 심장근육병증, 및 비대성 심장근육병증을 치료하거나 또는 예방하기 위한 방법에 관한 것이다.

[1256] 추가적인 병원성 A>G 돌연변이 및 SNP는 ClinVar 데이터베이스에서 확인된다. 따라서, 본 개시내용의 추가적인 양상은 ClinVar에 열거된 병원성 A>G 돌연변이 또는 SNP의 교정과 관련된 질환 또는 병태를 치료하거나 또

는 예방하기 위해 본 명세서에 기재된 방법, 시스템을 이용하는 상기 돌연변이의 교정에 관한 것이다.

[1257] 추가적인 병원성 C>T 돌연변이 및 SNP가 또한 ClinVar 데이터베이스에서 확인된다. 따라서, 본 개시내용의 추가적인 양상은 이와 관련된 질환 또는 병태를 치료하거나 또는 예방하기 위해 본 명세서에 기재된 방법, 시스템 및 조성물을 이용하는 ClinVar에 열거된 병원성 C>T 돌연변이 또는 SNP의 교정에 관한 것이다. 본 명세서에 개시된 실시형태를 사용하여 해결할 수 있는 다른 T 돌연변이 또는 SNP는 "Clin_var_pathogenic_SNPS_TC_txt"의 명칭으로 여기에 함께 제출된 ASCII 텍스트에 존재하는 표에 열거되어 있다.

[1258] 본 발명은 하기 실시예를 통해 더욱 설명되지만, 청구항에 기술된 본 발명의 범주를 제한하는 것은 아니다.

[1259] **실시예**

[1260] **실시예 1**

[1261] 아데닌 데아미나제 (AD)는 이중 가닥 RNA의 특정 부위에서 아데닌을 탈아미노화시킬 수 있다.

[1262] 일부 AD가 DNA-RNAn RNA 듀플렉스 상의 아데닌 탈아미노화를 달성할 수 있다는 사실 (예를 들어, Zheng et al., Nucleic Acids Research 2017)은 불활성 Cas13에 의한 RNA-가이드된 DNA 결합 동안 형성된 R-루프에서 가이드 RNA와 그의 상보성 DNA 사이에 형성된 RNA 듀플렉스를 이용함으로써 RNA 가이드된 AD를 개발할 고유한 기회를 제시한다. AD를 동원하기 위해 불활성 Cas13을 이용함으로써, AD 효소는 RNA-DNAn RNA 듀플렉스 내 아데닌 상에서 작용하게 될 것이다.

[1263] 일 실시형태에서, 불활성 Cas13, 예컨대 Cas13b는 다음의 돌연변이를 사용하여 얻어진다: R116A, H121A, R1177A 및 H1182A. AD에 의한 편집 효율을 증가시키기 위해, 돌연변이된 ADAR, 예컨대 돌연변이 E488Q를 포함하는 돌연변이된 hADAR2d가 사용된다.

[1264] 특정 유전자좌로 AD의 동원을 위한 디자인:

[1265] 1. NLS-태그된 불활성 Cas13을 N-말단부 또는 C-말단부 상에서 AD와 융합시킨다. 가요성 링커, 예컨대 GSG5 또는 덜 가요성인 링커 예컨대 LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR (SEQ ID No. 11)를 비롯한, 다양한 링커가 사용된다.

[1266] 2. 가이드 RNA 스캐폴드는 압타머, 예컨대, MS2 결합 부위로 변형된다 (예를 들어, Konermann et al., Nature 2015). NLS-태그된 AD-MS2 결합 단백질 융합은 (NLS-태그된 불활성 또는 Cas13) 및 상응하는 가이드 RNA와 함께 표적 세포 내로 공동 도입된다.

[1267] 3. AD는 NLS-태그된 불활성 또는 닉카제 Cas13의 내부 루프로 삽입된다.

[1268] RNA 가이드에 대한 디자인:

[1269] 1. 관심 RNA를 표적화하기 위해 Cas13 단백질의 천연 가이드 서열의 길이에 상응하는 길이의 가이드 서열을 디자인한다.

[1270] 2. 정규 길이보다 더 긴 RNA 가이드를 사용하여 단백질-가이드 RNA-표적 DNA 복합체 밖에 RNA 듀플렉스를 형성시킨다.

[1271] 이들 RNA 가이드 디자인의 각각에 대해, 표적 RNA 가닥 상에서 아데닌 반대쪽 RNA 상의 염기는 U 반대쪽 C로서 명시될 것이다.

[1272] AD의 선택 및 디자인:

[1273] 다수의 AD를 사용하고, 각각은 다양한 수준의 활성을 가지게 될 것이다. 이들 AD는 하기를 포함한다:

[1274] 1. 인간 ADAR (hADAR1, hADAR2, hADAR3)

[1275] 2. 오징어 옥토푸스 불가리스 ADAR

[1276] 3. 오징어 세피아 (Squid Sepia) ADARS; 도리테우스티스 오팔레센스 (Doryteusthis opalescens) ADARS

[1277] ADAT (인간 ADAT, 초파리 ADAT)

[1278] DNA-RNAn RNA 듀플렉스에 대한 ADAR 반응 활성을 증가시키기 위해 돌연변이를 또한 사용할 수 있다. 예를 들어, 인간 ADAR 유전자에 대해, hADAR1d (E1008Q) 또는 hADAR2d (E488Q) 돌연변이를 사용하여 DNA-RNA 표적에

대한 그들의 활성을 증가시킨다.

[1279] 각각의 ADAR은 다양한 수준의 서열 내용 요건을 갖는다. 예를 들어, hADAR1d (E1008Q)에 대해, tAg 및 aAg 부위는 효율적으로 탈아미노화되는 반면, aAt 및 cAc는 덜 효율적으로 편집되고, gAa 및 gAc는 훨씬 덜 편집된다. 그러나, 내용 요건은 상이한 ADAR에 대해 다양할 것이다.

[1280] 시스템의 한 형태를 도시한 개략도를 도 1에 제공한다. 예시적인 AD 단백질의 아미노산 서열을 도 4에 제공한다.

[1281] 본 명세서에 예시적으로 기술된 실시형태들은 본 명세서에 특별히 개시되지 않은, 임의의 구성요소 또는 구성요소들, 제한 또는 제한들 없이 적합하게 실시될 수 있다. 따라서, 예를 들어, "포함하는", "포괄되는", "함유하는" 등은 확장적으로 제한없이 읽어야 한다. 추가적으로, 본 명세서에 적용되는 용어 및 표현은 제한이 아닌 서술적 용어로서 사용되었고, 도시되고 기술된 특성의 임의 균등물 또는 이의 일부분을 배제하는 그러한 용어 및 표현의 사용을 의도하는 것은 아니지만, 다양한 변형이 청구된 기술 범주 내에서 가능하다는 것을 인식한다. 추가적으로, 어구 "~로 본질적으로 이루어지는"은 청구된 기술의 기본적이고 신규한 특징에 구체적으로 영향을 미치지 않는 특별히 언급된 구성요소 및 추가 구성요소를 포함한다는 것을 이해하게 될 것이다. 어구 "~로 이루어지는"은 명시되지 않은 임의의 구성요소는 배제한다.

[1282] 본 개시는 본 출원에 기술된 특정한 실시형태의 관점으로 제한되는 것이 아니다. 당업자에게 자명하게 되는 바와 같이, 이의 범주 및 사조를 벗어나지 않고 많은 변형 및 변동을 만들 수 있다. 본 명세서에 열거된 것들 이외에도, 본 개시의 범주 내에서 기능적으로 균등한 방법 및 조성물은 전술한 설명을 통해서 당업자에게 자명하게 될 것이다. 이러한 변형 및 변동은 첨부된 청구항의 범주 내에 포함시키고자 한다. 본 개시는 이러한 청구항에 부여되는 동등한 완전한 범주와 함께, 첨부된 청구항의 용어에 의해서만 제한된다. 특정한 방법, 시약, 화합물, 조성물 또는 생물학적 시스템에 국한되지 않고, 역시 다양할 수 있다는 것을 이해한다. 본 명세서에서 사용되는 전문용어는 오직 특정한 실시형태를 설명하려는 목적을 위한 것이고, 제한하려는 의도가 아니라는 것을 또한 이해해야 한다.

[1283] 또한, 본 개시의 특성 또는 양상이 마쿠쉬 그룹으로 설명되는 경우에, 본 개시가 역시 마쿠쉬 그룹의 임의의 개별 구성원 또는 구성원의 하위 그룹으로 설명된다는 것을 인지하게 될 것이다.

[1284] 당업자가 이해하게 되는 바와 같이, 임의의 모든 목적, 특히 서면 설명을 제공한다는 관점에서, 본 명세서에 개시된 모든 범위는 또한 임의의 모든 가능한 하위 범위 및 이의 하위 범위의 조합을 포괄한다. 임의의 열거된 범위는 적어도 절반, 삼분의 일, 사분의 일, 오분의 일, 십분의 일 등으로 충분히 설명하고 나눌 수 있다는 것을 쉽게 인식할 수 있다. 비제한적인 예로서, 본 명세서에 기술된 각 범위는 하위 삼분의 일, 중간 삼분의 일 및 상위 삼분의 일 등으로 쉽게 나눌 수 있다. 또한 당업자가 이해하게 되는 바와 같이, 모든 언어 예컨대 "최대," "적어도," "초과", "미만" 등은 열거된 수치를 포함하고 상기 기술된 바와 같이 이후에 하위 범위로 나눌 수 있는 범위를 언급한다. 마지막으로, 당업자가 이해하게 되는 바와 같이, 범위는 각 개별 구성원을 포함한다.

[1285] 본 명세서에서 언급되는 모든 출판물, 특허 출원, 특허 공보, 및 다른 문서들은 각각의 개별 출판물, 특허 출원, 특허 공보, 또는 다른 문서가 특별히 개별적으로 그 전문이 참조로 편입된다고 표시된 대로 참조로 본 명세서에 편입된다. 참조로 편입된 참고서에 함유된 정의는 본 개시의 정의에 모순되는 정도까지 배제된다.

[1286] **실시예 2**

[1287] Cas13a/Cas13b 간섭을 위한 Cluc/Gluc 타일링

[1288] Cas13a와 Cas13b 사이의 녹다운 효율을 비교하기 위해, 시프리디나 및 가우시아 루시퍼라제 유전자를 각각 24 또는 96개의 가이드로 타일링하였다 (도 10). 가이드를 Cas13a 및 Cas13b에 대해 매칭시켰고, Cas13b에 대한 증가된 효율을 나타내었으며, 각각의 유전자에 대해 하나의 가이드를 제외한 모두는 Cas13b에 대해 더 높은 효율을 나타낸다.

[1289] NGS에 의한 ADAR 편집 정량화

[1290] 루시퍼라제의 발현을 방지하는 조기 정지 코돈 UAG로 루시퍼라제 리포터를 디자인함으로써 Cas13b-ADAR2 RNA 편집 효율을 시험하였다 (도 11A-도 11C). 길이가 다양한 7개의 가이드를 디자인하였고, UAG에서 A에 대한 C 미스매치를 모두 함유하는 UAG 정지 코돈에 대해 위치시켰다. C 미스매치는 ADAR 촉매적 도메인에 의해 선호되는 편집 부위에서 버블을 생성하는 것으로 알려져 있다. Cas13b-ADAR2에 의한 RNA 편집은 UAG를

UIG(UGG)로 전환시키는데, 이는 정지 코돈 대신 트립토판을 도입하고, 번역이 진행되도록 한다. HEK293FT 세포에서 가이드 및 Cas13b12-ADAR2 융합의 발현은 가이드 5에 대해 생기는 가장 큰 복원에 의해 루시퍼라제 발현을 다양한 수준으로 복원시킨다 (도 11B). 일반적으로, 편집 부위가 직접 반복부가 있고 그에 따라 단백질이 결합하는 crRNA의 3' 말단부로부터 더 멀리 이동함에 따라 가이드 1 내지 5로부터의 편집이 증가된다. 이는 단백질에 의해 결합되는 crRNA:표적 듀플렉스의 부분은 ADAR 촉매적 도메인에 대해 접근 가능하지 않다는 것을 나타낼 가능성이 있다. 편집 부위가 DR/단백질 결합 영역으로부터 떨어진 가이드의 멀리 있는 말단부상에 있기 때문에 그리고 그들의 가이드는 훨씬 더 길기 때문에 가이드 5, 6 및 7은 가장 큰 양의 활성을 나타내어, ADAR에 의해 선호되는 더 긴 RNA 듀플렉스를 생성한다. 편집 부위가 RNA 듀플렉스의 중간에 있을 때 ADAR 활성은 최적이다. 루시퍼라제 활성의 상대적 발현은 비표적화 가이드 조건에 대해 정규화된다.

[1291] RNA 편집 효율을 정확하게 정량화하기 위해 이들 샘플을 시퀀싱하였다 (도 11C). 편집 효율은 가이드 표지 다음에 삽입 어구로 열거된다. 전반적으로 시퀀싱에 의해 확인된 편집 백분율은 도 11B에서 보이는 루시퍼라제 발현 복원의 상대적 수준과 매칭되었다. 가이드 5는 대부분의 RNA 편집을 온-표적 A 대 G의 45% 전환율로 나타내었다. 일부 예에서, 영역에 소량의 오프-표적 A-G 편집이 있다. 이들은 ADAR 촉매적 도메인에 의해 선호되지 않는 가이드 서열에 G 미스매치를 도입함으로써 감소될 수 있다.

[1292] 루시퍼라제 리포터 전사체의 편집에 추가로, 가이드는 KRAS 및 PPIB 전사체 내 프레임 밖 UAG 부위를 편집하도록 디자인하였고, 두 가이드는 각각 전사물을 표적화한다 (도 12). 가이드는 상기 가이드 5와 동일한 원칙으로 디자인하였다 (편집 부위 아데노신에 대한 C 미스매치 및 3' DR로부터 27 nt 떨어진 편집 부위를 갖는 45 nt 스페이서). KRAS 가이드는 온-표적 아데노신에서 6.5% 및 13.7% 편집을 달성할 수 있었고, PPIB 가이드는 7.7% 및 9.2% 편집을 달성할 수 있었다. 또한 근처에서 가능한 오프-표적 아데노신에 대해 스페이서 내 G 미스매치를 디자인하여 감소될 수 있는 이들 가이드 중 일부에 존재하는 일부 오프-표적이 존재한다. 오프-표적은 표적 영역 3'의 듀플렉스 영역에 의해 일어나는 것으로 여겨진다.

[1293] RNA Seq로부터의 Cas13a/b + shRNA 특이성

[1294] Cas13b12 녹다운의 특이성을 결정하기 위해, RNA 시퀀싱을 전사체에 걸쳐 모든 mRNA 상에서 수행하였다 (도 13A). Gluc 및 KRAS를 표적화하는 가이드의 녹다운을 비-표적화 가이드와 비교하였고, Cas13a2 및 Cas13b12가 표적 전사물의 특이적 녹다운을 갖는 한편 (도 13A에서 적색점), shRNA가 분포에서 더 큰 변량에 의해 입증되는 바와 같이 많은 오프-표적을 갖는다는 것을 발견하였다. 각각의 이들 조건에 대한 상당한 오프-표적의 수를 도 13B에 도시한다. 2배 초과 또는 0.8배 미만으로 변하는 임의의 오프-표적 전사물에 대한 FDR 교정과 함께 t-검정에 의해 유의한 오프-표적을 측정한다 ($p < 0.01$). Cas13a 및 Cas13b 조건은 shRNA 조건에 대해 발견된 수백개의 오프-표적에 비해 매우 소수의 오프-표적을 갖는다. 조건 각각에 대한 녹다운 효율을 도 13C에 도시한다.

[1295] 오프 표적을 감소시키기 위한 미스매치 특이성 (A:A 또는 A:G)

[1296] 표적 아데노신 편집 부위 근처의 아데노신에서 오프-표적을 감소시키기 위해, 잠재적 오프-표적 아데노신에 대해 G 또는 A 미스매치를 갖는 가이드를 디자인하였다 (도 14 및 이하의 표). G 또는 A와의 미스매치는 ADAR 촉매적 도메인에 의한 활성에 바람직하지 않다.

[1297]

명칭	가이드
C 미스매치 존재의 루시퍼라제 가이드 WT (SEQ ID No. 162)	catagaatgttctaaaCCAtcctgcggcctctactctgcattcaa
C 미스매치 및 1 G MM 존재의 루시퍼라제 가이드 WT (SEQ ID No. 163)	catagaatgttcGaaaCCAtcctgcggcctctactctgcattcaa
C 미스매치 및 2 G MM 존재의 루시퍼라제 가이드 WT (SEQ ID No. 164)	catagaatgGtcGaaaCCAtcctgcggcctctactctgcattcaa
C 미스매치 및 1 G MM 존재의 루시퍼라제 가이드 WT (SEQ ID No. 165)	catagaatgttcAaaaCCAtcctgcggcctctactctgcattcaa

C 미스매치 및 2 G MM 존재의 루시퍼라제 가이드 WT (SEQ ID No. 166)	catagaatgAtcAaaaCCAtcctgcggcctctactctgcattcaa
C미스매치 존재의 KRAS 가이드 WT (SEQ ID No. 167)	ggtttctccatcaattacCacttgcttcctgtaggaatcctctatt
C 미스매치 및 1 G MM 존재의 KRAS 가이드 (SEQ ID No. 168)	ggtttctccatcaatGacCacttgcttcctgtaggaatcctctatt
C 미스매치 및 2 G MM 존재의 KRAS 가이드 (SEQ ID No. 169)	ggtttctccatcaaGGacCacttgcttcctgtaggaatcctctatt
C 미스매치 및 1 A MM 존재의 KRAS 가이드 (SEQ ID No. 170)	ggtttctccatcaatAacCacttgcttcctgtaggaatcctctatt
C 미스매치 및 2 A MM 존재의 KRAS 가이드 (SEQ ID No. 171)	ggtttctccatcaaAAacCacttgcttcctgtaggaatcctctatt
C 미스매치 존재의 PPIB 가이드 WT (SEQ ID No. 172)	gcctttctctcctgtagcCaaggccacaaaattatccactgttttt
C 미스매치 및 1 G MM 존재의 PPIB 가이드 WT (SEQ ID No. 173)	gcctttctctcctgGagcCaaggccacaaaattatccactgtttttt
C 미스매치 및 1 A MM 존재의 PPIB 가이드 WT (SEQ ID No. 174)	gcctttctctcctgAagcCaaggccacaaaattatccactgtttttt

[1298]

[1299] 상기 표에서 가이드는 편집하려는 온-표적 아데노신에 대해 C 미스매치 및 기지의 오프-표적 부위에 대해 G 또는 A 미스매치를 갖도록 디자인하였다 (상기로부터의 RNA 시퀀싱에 기반함). 스페이서 서열 내 미스매치는 대문자로 표시한다.

[1300] 온-표적 활성에 대한 미스매치

[1301] ADAR2의 촉매적 도메인에 대한 사전 연구는 표적 A 반대쪽의 상이한 염기가 이노신 편집량에 영향을 미칠수 있다는 것을 입증하였다 (Zheng et al. (2017), Nucleic Acid Research, 45(6):3369-3377). 구체적으로, U 및 C는 반대쪽 천연 ADAR-편집 A에서 발견된 반면, G 및 A는 그렇지 않다. 편집된 A와 A 및 G 미스매치가 ADAR 활성을 억제하는 데 사용될 수 있는지의 여부를 시험하기 위해, C 미스매치에 의해 활성인 것으로 알려진 가이드를 루시퍼라제 리포터 어세이를 기반으로 모든 다른 3개의 가능한 염기로 시험한다 (도 16). 상대적 활성은 루시퍼라제 활성을 평가하여 정량된다. 가이드 서열을 이하의 표에 제공한다.

미스매치	가이드 서열
미스매치-C (SEQ ID No. 175)	GcatagaatgttctaaaCCAtcctgcggcctctactctgcattcaa
미스매치-G (SEQ ID No. 176)	GcatagaatgttctaaaCGAtcctgcggcctctactctgcattcaa
미스매치-T (SEQ ID No. 177)	GcatagaatgttctaaaCTAtcctgcggcctctactctgcattcaa
미스매치-A (SEQ ID No. 178)	GcatagaatgttctaaaCAAtcctgcggcctctactctgcattcaa

[1302]

[1303] gRNA의 화학적 변형에 의한 오프-표적 변형의 편집 및 감소의 개선

[1304] 오프-표적 활성을 감소시키고 온-표적 효율을 개선시키기 위해 문헌 [Vogel et al. (2014), Angew Chem Int Ed, 53:6267-6271, doi:10.1002/anie.201402634)]에 예시된 바와 같이 gRNA를 화학적으로 변형시킨다. 2'-O-메틸 및 포스포티오에이트 변형된 가이드 RNA는 일반적으로 세포에서 편집 효율을 개선시킨다.

[1305] 모티프 선호도

[1306] ADAR은 편집된 A의 측면 중 하나 상에 이웃하는 뉴클레오티드에 대한 선호도를 입증하는 것으로 알려져 있다 (www.nature.com/nsmb/journal/v23/n5/full/nsmb.3203.html, Matthews et al. (2017), Nature Structural Mol Biol, 23(5): 426-433). 선호도는 표적화된 A 주변의 다양한 염기로 시프리디나 루시퍼라제 전사물를 표적화함으로써 체계적으로 시험한다 (도 17).

[1307] RNA 편집 효율을 증강시키기 위한 거대한 버블

[1308] 비선호된 5' 또는 3' 이웃 염기에 대한 RNA 편집 효율을 증강시키기 위해,이웃하는 염기에 의도적으로 임스매치를 도입시켜서, 시험관내에서 비선호 모티프의 편집이 허용된다는 것을 입증하였다 (https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gku272; Schneider et al (2014), Nucleic Acid Res, 42(10):e87); Fukuda et al. (2017), Scienticic Reports, 7, doi:10.1038/srep41478). 추가 미스매치, 예컨대 구아노신 치환을, 그들이 자연적 선호도를 감소시키는지의 여부를 알아보기 위해 시험한다 (도 18).

[1309] 전사물 내 다중 A의 편집

[1310] 결과는 ADAR 데아미나제 도메인의 표적화 창에서 C 반대쪽 A가 다른 염기에 비해 우선적으로 편집된다는 것을 시사한다. 추가적으로, 표적화된 염기의 소수 염기 내에서 U와 염기쌍 형성한 A는 Cas13b-ADAR 융합에 의한 낮은 편집 수준을 보여서, 다중 A를 편집하기 위해 효소에 탄력성이 존재한다는 것을 시사한다 (도 19). 이들 두 관찰은 Cas13b-ADAR 융합의 활성 창에서 다중 A는 C로 편집하려는 모든 A를 미스매칭시켜 편집을 위해 특정할 수 있다는 것을 시사한다. 이를 시험하기 위해, 최적화 실험에서 대부분의 유망한 가이드를 선택하고 활성 창 내에서 다수 A:C 미스매치는 다수 A:I 편집을 생성시킬 확률을 시험하도록 디자인된다. 이 실험에 대한 편집율은 NGS를 이용하여 정량된다. 활성창에서 잠재적 오프-표적 편집을 억제하기 위해, 비-표적 A는 A 또는 G와 쌍형성된다 (염기 선호도 실험 결과에 따름).

[1311] RNA 편집을 위한 가이드 길이 적정

[1312] ADAR는 > 20bp 길이의 분자간 또는 분자내 RNA 듀플렉스 상에서 자연적으로 작동한다 (또한 문헌 [Nishikura et al. (2010), Annu Rev Biochem, 79:321-349] 참조). 결과는 crRNA가 길수록 더 긴 듀플렉스가 생성되어, 더 높은 활성을 갖는다는 것이 입증되었다. RNA 편집 활성에 대한 상이한 길이의 가이드 활성을 체계적으로 비교하기 위해, 본 발명자들은 우리의 루시퍼라제 리포터 어세이에서 정지 코돈을 교정하도록 30, 50, 70 및 84 개 염기의 가이드를 디자인하였다 (도 20A-20F 및 이하의 표). 우리는 편집된 A의 위치가 crRNA의 특이성 결정 영역의 3' 말단에 대해 mRNA:crRNA 듀플렉스 내의 모든 가능한 균일한 거리 (즉, +2, +4 등)에 존재하도록 이들 가이드를 디자인하였다.

[1313]

서열 명칭	서열
가이드_Cas13bC-luc_30_ADAR0Top (SEQ ID No. 179)	GCATCCTGCGGCCTCTACTCTGCATTCAATT
가이드_Cas13bC-luc_30_ADAR1Top (SEQ ID No. 180)	GACCATCCTGCGGCCTCTACTCTGCATTCAA

가이드_Cas13bC-luc_30_ADAR2Top (SEQ ID No. 181)	GAAACCATCCTGCGGCCTCTACTCTGCATTC
가이드_Cas13bC-luc_30_ADAR3Top (SEQ ID No. 182)	GCTAAACCATCCTGCGGCCTCTACTCTGCAT
가이드_Cas13bC-luc_30_ADAR4Top (SEQ ID No. 183)	GTTCTAAACCATCCTGCGGCCTCTACTCTGC
가이드_Cas13bC-luc_30_ADAR5Top (SEQ ID No. 184)	GTGTTCTAAACCATCCTGCGGCCTCTACTCT
가이드_Cas13bC-luc_30_ADAR6Top (SEQ ID No. 185)	GAATGTTCTAAACCATCCTGCGGCCTCTACT
가이드_Cas13bC-luc_30_ADAR7Top (SEQ ID No. 186)	GAGAATGTTCTAAACCATCCTGCGGCCTCTA
가이드_Cas13bC-luc_30_ADAR8Top (SEQ ID No. 187)	GATAGAATGTTCTAAACCATCCTGCGGCCTC
가이드_Cas13bC-luc_30_ADAR9Top (SEQ ID No. 188)	GCCATAGAATGTTCTAAACCATCCTGCGGCC
가이드_Cas13bC-luc_30_ADAR10Top(SEQ ID No. 189)	GTTCCATAGAATGTTCTAAACCATCCTGCGG
가이드_Cas13bC-luc_30_ADAR11Top(SEQ ID No. 190)	GCTTTCCATAGAATGTTCTAAACCATCCTGC
가이드_Cas13bC-luc_30_ADAR12Top(SEQ ID No. 191)	GCTCTTTCCATAGAATGTTCTAAACCATCCT
가이드_Cas13bC-luc_30_ADAR13Top(SEQ ID No. 192)	GATCTCTTTCCATAGAATGTTCTAAACCATC
가이드_Cas13bC-luc_30_ADAR14Top(SEQ ID No. 193)	GGAATCTCTTTCCATAGAATGTTCTAAACCA
가이드_Cas13bC-luc_50_ADAR0Top (SEQ ID No. 194)	GCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCA
가이드_Cas13bC-luc_50_ADAR1Top (SEQ ID No. 195)	GACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGG

[1314]

가이드_Cas13bC-luc_50_ADAR2Top (SEQ ID No. 196)	GAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTC
가이드_Cas13bC-luc_50_ADAR3Top (SEQ ID No. 197)	GCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACAT
가이드_Cas13bC-luc_50_ADAR4Top (SEQ ID No. 198)	GTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACAC
가이드_Cas13bC-luc_50_ADAR5Top (SEQ ID No. 199)	GTGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGAC
가이드_Cas13bC-luc_50_ADAR6Top (SEQ ID No. 200)	GAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTG
가이드_Cas13bC-luc_50_ADAR7Top (SEQ ID No. 201)	GAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATAC
가이드_Cas13bC-luc_50_ADAR8Top (SEQ ID No. 202)	GATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACAT
가이드_Cas13bC-luc_50_ADAR9Top (SEQ ID No. 203)	GCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTAC
가이드_Cas13bC-luc_50_ADAR10Top(SEQ ID No. 204)	GTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATT
가이드_Cas13bC-luc_50_ADAR11Top(SEQ ID No. 205)	GCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAA
가이드_Cas13bC-luc_50_ADAR12Top(SEQ ID No. 206)	GCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTC
가이드_Cas13bC-luc_50_ADAR13Top(SEQ ID No. 207)	GATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCAT
가이드_Cas13bC-luc_50_ADAR14Top(SEQ ID No. 208)	GGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGC
가이드_Cas13bC-luc_50_ADAR15Top(SEQ ID No. 209)	GTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCT
가이드_Cas13bC-luc_50_ADAR16Top(SEQ ID No. 210)	GACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACT

[1315]

가이드_Cas13bC-luc_50_ADAR17Top(SEQ ID No. 211)	GGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA
가이드_Cas13bC-luc_50_ADAR18Top(SEQ ID No. 212)	GTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTC
가이드_Cas13bC-luc_50_ADAR19Top(SEQ ID No. 213)	GCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCC
가이드_Cas13bC-luc_50_ADAR20Top(SEQ ID No. 214)	GTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGG
가이드_Cas13bC-luc_50_ADAR21Top(SEQ ID No. 215)	GGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGC
가이드_Cas13bC-luc_50_ADAR22Top(SEQ ID No. 216)	GCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCT
가이드_Cas13bC-luc_50_ADAR23Top(SEQ ID No. 217)	GACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATC
가이드_Cas13bC-luc_50_ADAR24Top(SEQ ID No. 218)	GGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCA
가이드_Cas13bC-luc_70_ADAR0Top(SEQ ID No. 219)	GCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTAT
가이드_Cas13bC-luc_70_ADAR1Top(SEQ ID No. 220)	GACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTT
가이드_Cas13bC-luc_70_ADAR2Top(SEQ ID No. 221)	GAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGT
가이드_Cas13bC-luc_70_ADAR3Top(SEQ ID No. 222)	GCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTG
가이드_Cas13bC-luc_70_ADAR4Top(SEQ ID No. 223)	GTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCC
가이드_Cas13bC-luc_70_ADAR5Top(SEQ ID No. 224)	GTGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTT
가이드_Cas13bC-luc_70_ADAR6Top(SEQ ID No. 225)	GAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTT

[1316]

가이드_Cas13bC-luc_70_ADAR7Top (SEQ ID No. 226)	GAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGT
가이드_Cas13bC-luc_70_ADAR8Top (SEQ ID No. 227)	GATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACAT
가이드_Cas13bC-luc_70_ADAR9Top (SEQ ID No. 228)	GCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAAC
가이드_Cas13bC-luc_70_ADAR10Top(SEQ ID No. 229)	GTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCA
가이드_Cas13bC-luc_70_ADAR11Top(SEQ ID No. 230)	GCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGG
가이드_Cas13bC-luc_70_ADAR12Top(SEQ ID No. 231)	GCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTC
가이드_Cas13bC-luc_70_ADAR13Top(SEQ ID No. 232)	GATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACAT
가이드_Cas13bC-luc_70_ADAR14Top(SEQ ID No. 233)	GGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACAC
가이드_Cas13bC-luc_70_ADAR15Top(SEQ ID No. 234)	GTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGAC
가이드_Cas13bC-luc_70_ADAR16Top(SEQ ID No. 235)	GACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTG
가이드_Cas13bC-luc_70_ADAR17Top(SEQ ID No. 236)	GGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATAC
가이드_Cas13bC-luc_70_ADAR18Top(SEQ ID No. 237)	GTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACAT
가이드_Cas13bC-luc_70_ADAR19Top(SEQ ID No. 238)	GCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTAC
가이드_Cas13bC-luc_70_ADAR20Top(SEQ ID No. 239)	GTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATT
가이드_Cas13bC-luc_70_ADAR21Top(SEQ ID No. 240)	GGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAA

[1317]

가이드_Cas13bC-luc_70_ADAR22Top(SEQ ID No. 241)	GCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTC
가이드_Cas13bC-luc_70_ADAR23Top(SEQ ID No. 242)	GACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCAT
가이드_Cas13bC-luc_70_ADAR24Top(SEQ ID No. 243)	GGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGC
가이드_Cas13bC-luc_70_ADAR25Top(SEQ ID No. 244)	GATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCT
가이드_Cas13bC-luc_70_ADAR26Top(SEQ ID No. 245)	GGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACT
가이드_Cas13bC-luc_70_ADAR27Top(SEQ ID No. 246)	GACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA
가이드_Cas13bC-luc_70_ADAR28Top(SEQ ID No. 247)	GACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTC
가이드_Cas13bC-luc_70_ADAR29Top(SEQ ID No. 248)	GCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCC
가이드_Cas13bC-luc_70_ADAR30Top(SEQ ID No. 249)	GCCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGG
가이드_Cas13bC-luc_70_ADAR31Top(SEQ ID No. 250)	GGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGC
가이드_Cas13bC-luc_70_ADAR32Top(SEQ ID No. 251)	GTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCT
가이드_Cas13bC-luc_70_ADAR33Top(SEQ ID No. 252)	GCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATC
가이드_Cas13bC-luc_70_ADAR34Top(SEQ ID No. 253)	GTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCA
가이드_Cas13bC-luc_84_ADAR0Top(SEQ ID No. 254)	GCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACACAGTCCA
가이드_Cas13bC-luc_84_ADAR1Top(SEQ ID No. 255)	GACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACACAGTC

[1318]

가이드_Cas13bC-luc_84_ADAR2Top (SEQ ID No. 256)	GAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACACAG
가이드_Cas13bC-luc_84_ADAR3Top (SEQ ID No. 257)	GCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACAC
가이드_Cas13bC-luc_84_ADAR4Top (SEQ ID No. 258)	GTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCAC
가이드_Cas13bC-luc_84_ADAR5Top (SEQ ID No. 259)	GTGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTC
가이드_Cas13bC-luc_84_ADAR6Top (SEQ ID No. 260)	GAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTT
가이드_Cas13bC-luc_84_ADAR7Top (SEQ ID No. 261)	GAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTAT
가이드_Cas13bC-luc_84_ADAR8Top (SEQ ID No. 262)	GATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTT
가이드_Cas13bC-luc_84_ADAR9Top (SEQ ID No. 263)	GCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGT
가이드_Cas13bC-luc_84_ADAR10Top(SEQ ID No. 264)	GTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTG
가이드_Cas13bC-luc_84_ADAR11Top(SEQ ID No. 265)	GCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCC
가이드_Cas13bC-luc_84_ADAR12Top(SEQ ID No. 266)	GCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTT
가이드_Cas13bC-luc_84_ADAR13Top(SEQ ID No. 267)	GATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTT
가이드_Cas13bC-luc_84_ADAR14Top(SEQ ID No. 268)	GGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGT
가이드_Cas13bC-luc_84_ADAR15Top(SEQ ID No. 269)	GTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACAT
가이드_Cas13bC-luc_84_ADAR16Top(SEQ ID No. 270)	GACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAAC

[1319]

가이드_Cas13bC-luc_84_ADAR17Top(SEQ ID No. 271)	GGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCA
가이드_Cas13bC-luc_84_ADAR18Top(SEQ ID No. 272)	GTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGG
가이드_Cas13bC-luc_84_ADAR19Top(SEQ ID No. 273)	GCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTC
가이드_Cas13bC-luc_84_ADAR20Top(SEQ ID No. 274)	GTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACAT
가이드_Cas13bC-luc_84_ADAR21Top(SEQ ID No. 275)	GGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACAC
가이드_Cas13bC-luc_84_ADAR22Top(SEQ ID No. 276)	GCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGAC
가이드_Cas13bC-luc_84_ADAR23Top(SEQ ID No. 277)	GACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTG
가이드_Cas13bC-luc_84_ADAR24Top(SEQ ID No. 278)	GGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATAC
가이드_Cas13bC-luc_84_ADAR25Top(SEQ ID No. 279)	GATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACAT
가이드_Cas13bC-luc_84_ADAR26Top(SEQ ID No. 280)	GGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTAC
가이드_Cas13bC-luc_84_ADAR27Top(SEQ ID No. 281)	GACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATT
가이드_Cas13bC-luc_84_ADAR28Top(SEQ ID No. 282)	GACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAA
가이드_Cas13bC-luc_84_ADAR29Top(SEQ ID No. 283)	GCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTC
가이드_Cas13bC-luc_84_ADAR30Top(SEQ ID No. 284)	GCCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCAT
가이드_Cas13bC-luc_84_ADAR31Top(SEQ ID No. 285)	GGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGC

[1320]

가이드_Cas13bC-luc_84_ADAR32Top(SEQ ID No. 286)	GTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCT
가이드_Cas13bC-luc_84_ADAR33Top(SEQ ID No. 287)	GCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACT
가이드_Cas13bC-luc_84_ADAR34Top(SEQ ID No. 288)	GTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA
가이드_Cas13bC-luc_84_ADAR35Top(SEQ ID No. 289)	GGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTC
가이드_Cas13bC-luc_84_ADAR36Top(SEQ ID No. 290)	GTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCC
가이드_Cas13bC-luc_84_ADAR37Top(SEQ ID No. 291)	GCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGG
가이드_Cas13bC-luc_84_ADAR38Top(SEQ ID No. 292)	GGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGC
가이드_Cas13bC-luc_84_ADAR39Top(SEQ ID No. 293)	GCCGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCT
가이드_Cas13bC-luc_84_ADAR40Top(SEQ ID No. 294)	GCGCCGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATC
가이드_Cas13bC-luc_84_ADAR41Top(SEQ ID No. 295)	GGTCGCCGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCA

[1321]

[1322] 인과적 질환 돌연변이의 반전

[1323] 유전자에 사전-종결 정지 부위를 도입시킨 병원성 G>A 돌연변이를 갖게 하기 표의 3개 유전자를 합성하고 그들을 비인간 세포주에 통합시킨다. 정지 코돈 UAG를 UIG (UGG)로 바꾸어서 전사물을 교정하고 따라서 단백질 번역을 복원시키는 Cas13b12-ADAR2의 능력을 시험한다.

[1324]

단백질 길이 (aa)	**전체 길이 후보**	**유전자**	**질환**
498	NM_004992.3(MECP2):c.311G>A (p.Trp104Ter)	MECP2	레트 증후군
414	NM_007375.3(TARDBP):c.943G>A (p.Ala315Thr)	TARDBP	근위축성 측삭 경화증 10형
393	NM_000546.5(TP53):c.273G>A (p.Trp91Ter)	TP53	리-프라우메니 증후군|유전성 암-소인 증후군

[1325]

[1326] 48개의 보다 병원성인 G>A 돌연변이는 동반 질환과 함께 하기 표 5에 표시된다. 전체 유전자보다는, 이들 돌연변이 주변 200 bp 단편을 합성하고, GFP 앞에서 클로닝시킨다. 사전-종결 부위가 복원되었을 때, 이것은 GFP의 번역을 허용하게 될 것이고 교정은 RNA 시퀀싱 이외에도 고속대량으로 형광을 통해 측정할 수 있다.

[1327] 표 5

	후보	유전자	질환
1	NM_004006.2(DMD):c.3747G>A (p.Trp1249Ter)	DMD	뒤센 근이영양증
2	M_000344.3(SMN1):c.305G>A (p.Trp102Ter)	SMN1	척수 근위축증, II형|쿠겔베르그-웰란더병
3	NM_000492.3(CFTR):c.3846G>A (p.Trp1282Ter)	CFTR	낭성 섬유증|유전성 췌장염|미제공됨|아탈루렌 반응 - 효능
4	NM_004562.2(PRKN):c.1358G>A (p.Trp453Ter)	PRKN	파킨슨병 2
5	NM_017651.4(AHI1):c.2174G>A (p.Trp725Ter)	AHI1	주버트 증후군 3
6	NM_000238.3(KCNH2):c.3002G>A (p.Trp1001Ter)	KCNH2	긴 QT 증후군|미제공됨
7	NM_000136.2(FANCC):c.1517G>A (p.Trp506Ter)	FANCC|C9orf3	판코니 빈혈, 상보군 C
8	NM_001009944.2(PKD1):c.12420G>A (p.Trp4140Ter)	PKD1	다낭성 신장병, 성인형
9	NM_177965.3(C8orf37):c.555G>A (p.Trp185Ter)	C8orf37	망막 색소변성증 64
10	NM_000833.4(GRIN2A):c.3813G>A (p.Trp1271Ter)	GRIN2A	언어 장애 수반 및 정신 지체 수반 또는 미수반의 초점 뇌전증

[1328]

11	NM_000548.4(TSC2):c.2108G>A (p.Trp703Ter)	TSC2	결절성 경화증 2| 결절성 경화증 증후군
12	NM_000267.3(NF1):c.7044G>A (p.Trp2348Ter)	NF1	신경섬유종증, 1형
13	NM_000520.5(HEXA):c.1454G>A (p.Trp485Ter)	HEXA	테이-삭스병
14	NM_130838.1(UBE3A):c.2304G>A (p.Trp768Ter)	UBE3A	엔젤만 증후군
15	NM_000543.4(SMPD1):c.168G>A (p.Trp56Ter)	SMPD1	니만-피크병, A형
16	NM_000218.2(KCNQ1):c.1175G>A (p.Trp392Ter)	KCNQ1	긴 QT 증후군
17	NM_000256.3(MYBPC3):c.3293G>A (p.Trp1098Ter)	MYBPC3	원발성 가족성 비대성 심근병증
18	NM_000038.5(APC):c.1262G>A (p.Trp421Ter)	APC	가족성 샘종 폴립증 1
19	NM_000249.3(MLH1):c.1998G>A (p.Trp666Ter)	MLH1	린치 증후군
20	NM_000054.4(AVPR2):c.878G>A (p.Trp293Ter)	AVPR2	X-연관 신장성 요붕증
21	NM_001204.6(BMPR2):c.893G>A (p.W298*)	BMPR2	원발성 폐 고혈압
22	NM_004560.3(ROR2):c.2247G>A (p.Trp749Ter)	ROR2	B1 단지증
23	NM_000518.4(HBB):c.114G>A (p.Trp38Ter)	HBB	베타 탈라세미아| 베타 탈라세미아
24	NM_024577.3(SH3TC2):c.920G>A (p.Trp307Ter)	SH3TC2	샤르코-마리-투스 질환, 4C형
25	NM_206933.2(USH2A):c.9390G>A (p.Trp3130Ter)	USH2A	어셔 증후군, 2A형
26	NM_000179.2(MSH6):c.3020G>A (p.Trp1007Ter)	MSH6	린치 증후군
27	NM_002977.3(SCN9A):c.2691G>A (p.Trp897Ter)	SCN9A|LOC101929680	선천성 통각 무반응증, 상염색체 열성
28	NM_000090.3(COL3A1):c.30G>A (p.Trp10Ter)	COL3A1	엘러스-단로스 증후군, 4형

[1329]

29	NM_000551.3(VHL):c.263G>A (p.Trp88Ter)	VHL	폰 히펠-린다우 증후군|미제공됨
30	NM_015627.2(LDLRAP1):c.65G>A (p.Trp22Ter)	LDLRAP1	고콜레스테롤혈증, 상염색체 열성
31	NM_000132.3(F8):c.3144G>A (p.Trp1048Ter)	F8	유전성 인자 VIII 결핍 질환
32	NM_002185.4(IL7R):c.651G>A (p.Trp217Ter)	IL7R	중증 복합 면역결핍증, 상염색체 열성, T 세포-음성, B 세포-양성, NK 세포-양성
33	NM_000527.4(LDLR):c.1449G>A (p.Trp483Ter)	LDLR	가족성 고콜레스테롤혈증
34	NM_002294.2(LAMP2):c.962G>A (p.Trp321Ter)	LAMP2	다논병
35	NM_000271.4(NPC1):c.1142G>A (p.Trp381Ter)	NPC1	니만-피크병 C1형
36	NM_000267.3(NF1):c.1713G>A (p.Trp571Ter)	NF1	신경섬유종증, 1형
37	NM_000035.3(ALDOB):c.888G>A (p.Trp296Ter)	ALDOB	유전성 프룩토스뇨증
38	NM_000090.3(COL3A1):c.3833G>A (p.Trp1278Ter)	COL3A1	엘러스-단로스 증후군, 4형
39	NM_001369.2(DNAH5):c.8465G>A (p.Trp2822Ter)	DNAH5	원발성 섬모운동 이상증
40	NM_178443.2(FERMT3):c.48G>A (p.Trp16Ter)	FERMT3	백혈구 부착 결핍증, III형
41	NM_005359.5(SMAD4):c.906G>A (p.Trp302Ter)	SMAD4	소아 폴립증 증후군
42	NM_032119.3(ADGRV1):c.7406G>A (p.Trp2469Ter)	ADGRV1	어셔 증후군, 2C형
43	NM_000206.2(IL2RG):c.710G>A (p.Trp237Ter)	IL2RG	X-연관 중증 복합 면역결핍증
44	NM_007294.3(BRCA1):c.5511G>A (p.Trp1837Ter)	BRCA1	가족성 유방암|유방-난소암, 가족성 1
45	NM_130799.2(MEN1):c.1269G>A (p.Trp423Ter)	MEN1	유전성 암-소인 증후군
46	NM_000071.2(CBS):c.162G>A (p.Trp54Ter)	CBS	CBS 결핍에 의한 호모시스틴뇨증

[1330]

47	NM_000059.3(BRCA2):c.582G>A (p.Trp194Ter)	BRCA2	가족성 유방암|유방-난소암, 가족성 2
48	NM_000053.3(ATP7B):c.2336G>A (p.Trp779Ter)	ATP7B	윌슨병

[1331]

[1332] **실시예 3**

[1333] 효율적이고 정확한 핵산 편집은 특히 RNA 수준에서 유전자 질환을 치료하기 위한 상당한 장래성을 보유하며, 여기서 질환-관련 전사물은 기능성 단백질 생성물을 수득하도록 구제될 수 있다. VI형 CRISPR-Cas 시스템은 프로그램 가능한 단일-이펙터 RNA-가이드 RNases Cas13을 함유한다. 본 명세서에서, 본 발명자들은 강력한 녹다운을 할 수 있는 Cas13 오솔로그를 조작하기 위한 VI형 시스템의 다양성을 프로파일링하고 포유동물 세포 내 전사물에게로 아데노신 데아미나제 활성을 유도하기 위해 촉매적 불활성인 Cas13 (dCas13)을 사용하여 RNA 편집을 입증한다. ADAR2 데아미나제 도메인을 dCas13에 융합시키고 가이드 RNA를 조작하여 최적의 RNA 듀플렉스 기질을 생성함으로써, 우리는 대체로 20 내지 40% 및 최대 89% 범위의 효율로 특정 단일 아데노신의 이노신 (번역 동안 구아노신으로서 판독됨)으로의 표적화 편집을 달성한다. REPAIR (Edit for Programmable A

to I Replacement)라고 하는 이 시스템은 고특이성을 획득하도록 더욱 조작될 수 있다. 조작된 변이체, REPAIRv2는 강력한 온-표적 A>I 편집을 유지하면서 170배 초과의 특이성 증가를 나타낸다. 우리는 기지의 병원성 돌연변이를 함유하는 전체 길이 전사물을 편집하기 위해 REPAIRv2를 사용하고, 아데노-연합 바이러스 (AAV) 벡터 내 패키징에 적합한 기능성 절단 형태를 생성한다. REPAIR은 연구, 치료 제 및 생체기술을 위한 광범위 적용에 유망한 RNA 편집 플랫폼을 제공한다. 정확한 핵산 편집 기술은 세포 기능을 연구하기 위해서, 그리고 신규 치료제로서 가치가 있다. 현재의 편집 도구, 예컨대 Cas9 뉴클레아제는 게놈 유전자좌의 프로그램 가능한 변형을 달성할 수 있지만, 편집은 종종 삽입 또는 결실에 기인하여 이질적이거나 또는 정확한 편집을 위한 도너 주형을 필요로 한다. 염기 편집자, 예컨대 dCas9-APOBEC 융합은 이중 가닥 파손을 생성하는 일 없이 편집을 가능하게 하지만, 표적창에서 임의의 시티딘을 편집하는 시티딘 데아미나제의 특성에 기인하여 정확성이 결여될 수 있다. 더 나아가, 프로토스페이서 인접 모티프(PAM)에 대한 필요는 가능한 편집 부위의 수를 제한한다. 여기서, 우리는 VI형 원핵생물 클러스터링된 주기적 간격으로 분포하는 짧은 회문구조 반복부 (CRISPR) 적응 면역계로부터의 RNA-가이드된 표적화 Cas13 효소를 사용한 정확하고 탄력적인 RNA 염기 편집 도구의 개발을 기술한다.

[1334] 정확한 핵산 편집 기술은 세포 기능을 연구하는데 그리고 신규한 치료제로서 가치있다. 원핵생물 클러스터링된 주기적 간격으로 분포하는, 짧은 회문구조 반복부(CRISPR)-결합 뉴클레아제 Cas9(1-4) 또는 Cpf1(5)과 같은 프로그램 가능한 뉴클레아제에 기반한 현재의 편집 도구는 표적화된 DNA 절단을 매개하고, 결국 비상동성 말단 결합 (NHEJ)을 통한 표적화된 유전자 붕괴 또는 주형-의존적 상동정 지정 복구 (HDR)를 통한 정확한 유전자 편집을 유도하기 위해 널리 채택되었다 (6). NHEJ는 분열 세포와 유사분열 후 세포 둘 모두에서 활성인 숙주 기구를 이용하며, 단백질 코딩 유전자에서 프레임 시프트를 야기할 수 있는 삽입 또는 결실의 혼합 (indel) 돌연변이를 생성시켜 효율적인 유전자 붕괴를 제공한다. 대조적으로, HDR은 그 발현이 대체로 복제 세포에 국한되는 숙주 기구를 통해 매개된다. 이와 같이, 유사분열 후 세포에서 유전자-편집 능력의 개발은 주된 문제로 남아있다. 최근에, DNA 염기 편집자, 예컨대 표적창 내에서 시토신을 티민으로의 전환을 실시하도록 특별한 게놈 표적으로 시티딘 데아미나제 활성을 표적화하기 위한 촉매적 불활성 Cas9 (dCas9)의 사용은 DNA 이중 가닥 파괴를 생성시키지 않고 편집을 허용하고 유의하게 indel의 형성을 감소시킨다 (7, 8). 그러나, DNA 염기 편집자의 표적화 범위는 편집 부위에서 프로토스페이서 인접 모티프 (PAM)에 대한 Cas9의 요건으로 인해 제한적이다(9). 여기서, 우리는 VI형 CRISPR-결합 RNA-가이드된 RNase Cas13을 사용하는 정확하고 탄력적인 RNA 염기 편집 기술의 개발을 기술한다 (10-13).

[1335] Cas13 효소는 정확한 RNA 절단을 매개하는 2종의 고등 진핵생물 및 원핵생물 뉴클레오티드-결합 (HEPN) 엔도RNase 도메인을 갖는다 (10, 11). 3종의 Cas13 단백질 패밀리가 지금까지 확인되었다: Cas13a (이전에 C2c2로서 알려짐), Cas13b 및 Cas13c (12, 13). 우리는 최근에 Cas13a 효소가 핵산 검출을 위한 도구 (14)를 비롯하여 포유동물 및 식물 세포 RNA 녹다운 및 전사물 추적 (15)에 적합할 수 있다는 것을 보고하였다. Cas13 효소의 RNA-가이드 특성은 그들을 RNA 결합 및 섭동 적용을 위한 매력적인 도구로 만든다.

[1336] RNA 상에서 작용하는 아데노신 데아미나제 (ADAR)의 효소 패밀리는 번역 및 스플라이싱에서 구아닌과 기능적으로 동등한 핵염기인 이노신으로 아데노신의 가수분해적 탈아미노화를 통해 전사물의 내생성 편집을 매개한다 (16). N-말단 이중 가닥 RNA-결합 도메인 및 C-말단 촉매적 탈아미노화 도메인으로 이루어진 2종의 기능성 인간 ADAR 오솔로그, 즉, ADAR1 및 ADAR2가 존재한다. ADAR1 및 ADAR2의 내생성 표적 부위는 실질적인 이중 가닥 동일성을 함유하며, 촉매적 도메인은 시험관내 및 생체내에서 효율적인 편집을 위해 듀플렉스 영역을 필요로 한다 (17, 18). ADAR 단백질은 표적화의 잠재적 탄력성을 제한할 수 있는 편집을 위한 바람직한 모티프를 갖지만, 과활성 돌연변이체, 예컨대 ADAR(E488Q)(19)은 서열 제약을 풀어주고, 아데노신의 이노신으로의 편집율을 개선시킨다. ADAR은 RNA 듀플렉스에서 시티딘 염기 반대쪽 아데노신을 우선적으로 탈아미노화시켜 (20), 정확한 염기 편집을 위한 유망한 기회를 제공한다. 정확한 접근법은 RNA 가이드를 통한 표적화된 ADAR 융합을 조작하였지만 (21-24), 이들 접근법의 특이성은 보고되지 않았고, 그들 각각의 표적화 기전은 표적 인식 및 엄격성을 증강시킬 수 있는 단백질 파트너의 보조없이 RNA-RNA 하이브리드화에 의존한다.

[1337] 여기서 우리는 포유동물 세포에서 RNA 녹다운 활성을 위한 Cas13 효소의 전체 패밀리를 분석하고 포유동물 세포 적용을 위해 가장 효율적이고 특이적인 것으로 프레보텔라 종 P5-125 (PspCas13b) 유래의 Cas13b 오솔로그를 확인한다. 이어서, 우리는 ADAR2 데아미나제 도메인 (ADARDD)을 촉매적 불활성 PspCas13b에 융합시키고, 리포터 및 내생성 전사물뿐만 아니라 질환-관련 돌연변이의 프로그램 가능한 A → I(G) 대체(REPAIR)를 위한 RNA 편집을 입증한다. 마지막으로, 우리는 170배 초과로 특이성이 증가되면서 REPAIRv2를 생성하기 위해 dCas13b-ADAR2DD 융합의 특이성을 개선시키기 위한 합리적인 돌연변이유발 계획을 사용한다.

[1338] 방법 디자인 및 박테리아 구성체의 클로닝

[1339] 포유동물 코돈 최적화된 Cas13b 구성체를 Lac 프로모터의 제어 하에 클로람페니콜 내성 pACYC184 벡터에 클로닝시켰다. 이어서, BsaI 제한 부위에 의해 분리되는 두 상응하는 직접 반복부 (DR) 서열을 pJ23119 프로모터의 제어 하에, Cas13b 하류에 삽입하였다. 마지막에, 스페이서 표적 용 올리고는 T4 PNK (New England Biolabs)를 이용하여 인산화시켰고, 어닐링시켰으며 T7 리가제 (Enzymatics)를 이용하여 BsaI 분해 벡터에 결찰시켜서 표적화 Cas13b 벡터를 생성하였다.

[1340] 박테리아 PFS 스크리닝

[1341] PFS 스크린을 위한 암피실린 내성 플라스미드는 NEB 깁슨 어셈블리 (Gibson Assembly)(New England Biolabs)를 사용하여 암피실린 내성 유전자 bla의 출발 코돈의 바로 하류의 표적 부위를 통해 분리된 3 4' 무작위화 뉴클레오티드 및 2 5' 무작위화 뉴클레오티드와 Cas13b 표적을 함유하는 생성물을 삽입시켜 클로닝하였다. 이어서, 100 ng의 암피실린-내성 표적 플라스미드는 Endura Electrocompetent 세포 내로 65 내지 100 ng의 클로람페니콜-내성 Cas13b 표적화 플라스미드와 함께 전기천공시켰다. 플라스미드를 세포에 첨가하고 나서, 얼음 상에서 15분 동안 인큐베이션시키고, 제조자의 프로토콜을 이용하여 전기천공시키고, 이어서, 950 ㎕의 회수 배지를 세포에 첨가한 후에 37℃에서 1시간 증식시켰다. 증식물을 클로람페니콜 및 암피실린 이중 선택 플레이트 상에 플레이팅하였다. 증식물의 연속 희석물을 사용하여 cfu/ng DNA를 추정하였다. 플레이팅 후 16시간에, 세포를 플레이트로부터 모아서 Qiagen Plasmid Plus Maxi Kit (Qiagen)을 이용하여 생존 플라스미드 DNA를 채취하였다. 생존 Cas13b 표적 서열 및 그들의 측접 영역을 PCR에 의해 증폭시키고, Illumina NextSeq를 이용하여 시퀀싱하였다. PFS 선호도를 평가하기 위해, 본래 라이브러리 내 무작위화 뉴클레오티드를 함유하는 위치를 추출하고 나서, 생복제물 둘 모두에 존재하는 벡터 단독 조건에 대해 감손된 서열을 맞춤 python script를 이용하여 추출하였다. 이어서, 벡터 단독 조건과 비교되는 Cas13b 조건에서의 PFS 존재비의 -log2를 사용하여 바람직한 모티프를 계산하였다. 구체적으로, 특정 역치 이상의 -log2(샘플/벡터) 감손비를 갖는 모든 서열을 사용하여 서열 모티프의 웹로고를 생성하였다 (weblogo.berkeley.edu). 각각의 Cas13b 오솔로그에 대한 웹로고를 생성하기 위해 사용한 특정 감손 비율 값을 표 9에 열거한다.

[1342] RNA 간섭을 위한 포유동물 구성체의 디자인 및 클로닝

[1343] 포유동물 세포에서 Cas13 오솔로그를 시험하기 위한 벡터를 생성하기 위해, 포유동물 코돈 최적화된 Cas13a, Cas13b 및 Cas13c 유전자를 PCR 증폭시키고 나서, EF1알파 프로모터의 제어 하에 이중 NLS 서열 및 C-말단의 msfGFP를 함유하는 포유동물 발현 벡터에 골든-게이트 클로닝시켰다. 추가 최적화를 위해, EF1알파 프로모터의 제어 하에 상이한 C-말단의 국소화 태그를 함유하는 목적 벡터에 Cas13 오솔로그를 골든 게이트 클로닝시켰다.

[1344] 이중 루시퍼라제 리포터는 가우시아 및 시프리디니아 루시퍼라제 코딩 DNA, EF1알파 및 CMV 프로모터를 PCR 증폭시켜 클로닝하였고 NEB 깁슨 어셈블리 (New England Biolabs)를 사용해 조립하였다.

[1345] Cas13a, Cas13b 또는 Cas13c 오솔로그에 대한 포유동물 가이드 RNA의 발현을 위해서, 상응하는 직접 반복부 서열은 골든-게이트 수용자 부위로 합성시켰고 제한효소 분해 클로닝을 통해 U6 발현 하에 클로닝하였다. 그 다음으로 개별 가이드는 골든 게이트 클로닝을 통해 각 오솔로그를 위한 상응하는 발현 골격에 클로닝시켰다.

[1346] Cas13 간섭 특이성을 위해 풀링된 미스매치 라이브러리의 클로닝

[1347] 풀링된 미스매치 라이브러리 표적 부위를 PCR에 의해 생성시켰다. 표적 내 각 위치에서 94%의 정확한 염기와 2%의 각각의 부정확한 염기의 혼합을 함유하는 G-루시퍼라제 내 반-축퇴성 표적 서열을 함유하는 올리고를 1개의 프라이머로서 사용하고, G-루시퍼라제의 비표적화된 영역에 상응하는 올리고를 PCR 반응에서 제2 프라이머로서 사용하였다. 이어서, 미스매치 라이브러리 표적은 NEB 깁슨 어셈블리 (New England Biolabs)를 이용하여 야생형 G-루시퍼라제 대신 이중 루시퍼라제 리포터에 클로닝시켰다.

[1348] RNA 편집을 위한 포유동물 구성체의 디자인 및 클로닝

[1349] PspCas13b는 HEPN 도메인의 촉매적 부위에서 2개 히스티딘의 알라닌으로의 돌연변이 (H133A/H1058A)를 통해 촉매적으로 불활성하게 만들었다 (dPspCas13b). 인간 ADAR1 및 ADAR2의 데아미나제 도메인을 합성하였고, pcDNA-CMV 벡터 골격에 깁슨 클로닝을 위해 PCR 증폭시키고, GS 또는 GSGGGGS (SEQ ID No. 296) 링커를 통해 C-말단에서 dPspCas13b에 융합시켰다. 우리가 상이한 링커를 시험한 실험의 경우에서, 우리는 dPspCas13b와 ADAR2dd 사이에 다음의 추가적인 링커를 클로닝시켰다: GGGGSGGGGSGGGGS, EAAAK (SEQ ID No. 297),

GGSGGSGGSGGSGGSGGS (SEQ ID No. 298) 및 SGSETPGTSESATPES (XTEN)(SEQ ID No. 299) (XTEN). 특이성 돌연변이체는 적절한 돌연변이체를 dPspCas13b-GSGGGGS (SEQ ID No. 761)골격에 깁슨 클로닝시켜서 생성시켰다.

[1350] RNA 편집 활성을 측정하기 위한 루시퍼라제 리포터 벡터는 녹다운 실험에 사용된 사용한 루시퍼라제 리포터 벡터에 W85X 돌연변이 (TGG>TAG)를 생성시켜서 생성시켰다. 이 리포터 벡터는 정규화 대조군으로서 기능성 Gluc를 발현시키지만, 사전종결 부위의 첨가에 기인하는 결함성 Cluc를 발현한다. ADAR 편집 모티프 선호도를 시험하기 위해, 우리는 Cluc의 코돈 85(XAX)의 아데노신 주변의 모든 가능한 모티프를 클로닝시켰다.

[1351] REPAIR의 PFS 선호도를 시험하기 위해, 본 발명자들은 표적 영역과 아데노신 편집 부위 상류의 6개의 염기쌍 축퇴 PFS 서열을 함유하는 풀링된 플라스미드 라이브러리를 클로닝시켰다. Integrated DNA Technology (IDT)의 ultramer로서 라이브러리를 합성하였고, 프라이머 어닐링을 통해 이중 가닥을 생성하고, DNA 중합효소 I의 Klenow 단편 (New England Biolabs)이 서열을 채우게 된다. 이어서, 축퇴 서열을 함유하는 이러한 dsDNA 단편을 분해된 리포터 벡터에 깁슨 클로닝시켰고, 이소프로판올 침전시켰고, 정제하였다. 이어서, 클로닝된 라이브러리를 Endura competent 이. 콜라이 세포 (Lucigen)에 전기천공시키고 나서, 245 ㎜ x 245㎜ 사각 바이오어세이 플레이트 (Nunc) 상에 플레이팅시켰다. 16시간 후에, 콜로니를 채취하고 나서, 무 엔도톡신 MACHEREY-NAGEL 미디프렙 키트를 이용하여 미디프렙하였다. 차세대 시퀀싱에 의해 클로닝된 라이브러리를 확인하였다.

[1352] REPAIR 활성을 시험하기 위한 질환-관련 돌연변이를 클로닝시키기 위해, ClinVar에 정의된 바와 같은 질환 발병과 관련된 34개의 G>A 돌연변이를 선택하고 나서, 이들 돌연변이 주변의 200bp 영역을 CMV 프로모터 하에 mScarlett와 EGFP 사이에 골든 게이트 클로닝시켰다. EF1알파의 발현 하에 전체 게놈 서열의 깁슨 클로닝을 위해 AVPR2 및 FANCC에서 2개의 추가적인 G>A 돌연변이를 선택하였다.

[1353] REPAIR를 위한 포유동물 가이드 RNA의 발현을 위해, PspCas13b 직접 반복 서열을 골든-게이트 수용자 부위로 합성하고 나서, 제한 분해 클로닝을 통해 U6 발현 하에 클로닝시켰다. 이어서, 개개 가이드를 골든 게이트 클로닝에 의해 이 발현 골격에 클로닝시켰다.

[1354] 포유동물 세포 배양물

[1355] 포유동물 세포 배양 실험은 고 글루코스 소듐 피루베이트, 및 GlutaMAX (Thermo Fisher Scientific)가 존재하고, 추가적으로 1X 페니실린-스트렙토마이신 (Thermo Fisher Scientific) 및 10% 소 태아 혈청 (VWR Seradigm)을 보충한 Dulbecco의 변형 Eagle 배지에서 성장시킨, HEK293FT 세포주 (American Type Culture Collection (ATCC))에서 수행하였다. 세포를 80% 이하의 합류점으로 유지시켰다.

[1356] 달리 언급되지 않는 한, 모든 형질감염은 폴리-D-리신 (BD Biocoat)으로 코팅된 96-웰 플레이트에서 Lipofectamine 2000(Thermo Fisher Scientific)을 이용하여 수행하였다. 형질감염 시 90% 합류를 보장하기 위해 형질감염 전 16시간에 세포를 대략 20,000개 세포/웰로 플레이팅시켰다. 플레이트 상의 각각의 웰에 대해, 형질감염 플라스미드를 Opti-MEM I Reduced Serum Medium)(Thermo Fisher)와 총 25㎕로 배합하였다. 별개로, 24.5㎕의 Opti-MEM을 0.5㎕의 Lipofectamine 2000과 배합하였다. 이어서, 플라스미드 및 Lipofectamine 용액을 배합하고 나서, 5분 동안 인큐베이션시키고 나서, 이 후에, 그들을 세포에 피펫팅하였다.

[1357] RNA 녹다운 포유동물 세포 어세이

[1358] 리포터 구성체를 이용하여 포유동물 세포에서 RNA 표적화를 평가하기 위해, 150 ng의 Cas13 구성체를 300 ng의 가이드 발현 플라스미드와 12.5 ng의 녹다운 리포터 구성체와 공동형질감염시켰다. 형질감염 후 48시간에, 분비된 루시퍼라제를 함유하는 배지를 세포에서 제거하고, PBS에서 1:5로 희석시키고, 주입 프로토콜을 이용하여 플레이트 판독기 (Biotek Synergy Neo2) 상에서 BioLux 시프리디니아 및 바이오룩스 가우시아 루시퍼라제 분석 키트 (New England Biolabs)로 활성에 대해 측정하였다. 수행한 모든 복제물은 생물학적 복제물이다.

[1359] 내생성 유전자의 표적화를 위해, 150 ng의 Cas13 구성체를 300 ng의 가이드 발현 플라스미드와 공동 형질감염시켰다. 형질감염후 48시간에, 세포를 용해시키고, RNA를 채취하고 나서, Cells-to-Ct (Thermo Fisher Scientific)의 이전에 기술된 [CITE PROTOCOLS] 변형을 이용하여 역전사시켰다. KRAS 전사물에 대해 TaqMan qPCR 프로브 (Thermo Fisher Scientific), GAPDH 대조군 프로브 (Thermo Fisher Scientific) 및 Fast Advanced Master Mix (Thermo Fisher Scientific)를 사용하는 qPCR을 통해 cDNA 발현을 측정하였다. 384-웰 형식으로 4개의 5 ㎕ 기술적 복제물을 이용하여 LightCycler 480 장비 II (Roche) 상에서 qPCR을 판독하였다.

[1360] 미스매치된 표적의 풀링된 라이브러리를 사용한 RNA 특이성의 평가

[1361] 6개의 웰 플레이트에 파종시킨 HEK293FT 세포를 이용하여 미스매치된 표적 라이브러리를 간섭하는 Cas13의 능력을 시험하였다. 2400 ng의 Cas13 벡터, 4800 ng의 가이드 및 240 ng의 미스매치된 표적 라이브러리를 이용하여 ~70% 합류점 세포를 형질감염시켰다. 형질감염 후 48시간에, 세포를 채취하고 나서, QIAshredder (Qiagen) 및 Qiagen RNeasy 미니 키트를 이용하여 RNA 추출하였다. 제조사의 유전자-특이적 프라이밍 프로토콜에 따라 qScript Flex cDNA 합성 키트 (Quantabio) 및 Gluc 특이적 RT 프라이머를 이용하여 1 ㎍의 추출된 RNA를 역전사시켰다. 이어서, cDNA를 증폭시키고, Illumina NextSeq 상에서 시퀀싱하였다.

[1362] 서열 당 판독을 계측하여 시퀀싱을 분석하였고 감손 점수는 log2 (-판독 계수 비) 값을 결정하여 계산하였고, 여기서 판독 계측비는 표적화 가이드 조건 대 비표적화 가이드 조건에서의 판독 계측치의 비율이다. 이런 점수 값은 서열 상의 Cas13 활성 수준을 나타내며 값이 높을수록 더 강한 감손을 나타내고, 따라서 더 큰 Cas13 절단 활성을 나타낸다. 단일 미스매치 및 이중 미스매치 서열에 대한 별개의 분포를 결정하고 나서, 각각의 미스매치 정체에 대한 감손 점수에 따라 히트맵으로서 그래프화하였다. 이중 미스매치 서열의 경우 소정 위치에서 모든 가능한 이중 미스매치의 평균을 그래프화하였다.

[1363] RNA 시퀀싱에 의한 포유동물 세포에서의 Cas13의 전사체-와이드 프로파일링

[1364] 전사체-와이드 특이성의 측정을 위해, 150 ng의 Cas13 구성체, 300 ng의 가이드 발현 플라스미드 및 15 ng의 녹다운 리포터 구성체를 공동형질감염시켰고; shRNA 조건의 경우, 300 ng의 shRNA 표적화 플라스미드, 15 ng의 녹다운 리포터 구성체 및 150 ng의 EF1-알파 구동 mCherry(리포터 로딩와 균형을 위함)를 공동형질감염시켰다. 형질감염 후 48시간에, RNA를 RNeasy Plus 미니 키트 (Qiagen)로 정제하고 나서, NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs)을 사용하기 위해 mRNA를 선택하고 나서, 시퀀싱을 위해 Illumina를 위한 EBNext Ultra RNA Library Prep Kit (New England Biolabs)를 사용해 제조하였다. 이어서, RNA 시퀀싱 라이브러리를 NextSeq(Illumina) 상에서 시퀀싱하였다.

[1365] 전사체-와이드 시퀀싱 데이터를 분석하기 위해, 판독은 디폴트 매개변수: [CITE RSEM: 기준 게놈 존재 또는 부재의 RNA-Seq 데이터로부터의 정확한 전사물 정량]로 Bowtie 및 RSEM version 1.2.3.1을 사용하여 정렬된 RefSeq GRCh39 어셈블리였다. 전사물 발현은 log2(TPM + 1)로서 정량하고였고, 유전자는 log2(TPM + 1) >2.5에 대해 필터링하였다. 차등적으로 발현된 유전자의 선택을 위해, >2 또는 < .75의 차등 변화가 있는 유전자만을 고려하였다. 3개의 표적화 복제물 대 비표적화 복제물에 대한 스튜던트 T-검정으로 차등적 발현의 통계적 유의성을 평가하였고, 벤자미니-호크버그 (Benjamini-Hochberg) 절차에 의해 < 0.01%의 거짓 발견율에 대해 필터링하였다.

[1366] 포유동물 세포 형질감염에서의 ADAR RNA 편집

[1367] 포유동물 세포에서 REPAIR 활성을 평가하기 위해, 우리는 150 ng의 REPAIR 벡터, 300 ng의 가이드 발현 플라스미드, 및 40 ng의 RNA 편집 리포터를 형질감염시켰다. 48시간 후에, 세포로부터의 RNA를 채취하고 나서, 유전자 특이적 역전사 프라이머를 사용해 앞서 기재한 방법 [cite JJ] 을 이용하여 역전사시켰다. 이어서, NEBNext High-Fidelity 2X PCR 마스터 믹스를 이용하여 Illumina 어댑터 및 샘플 바코드를 첨가하기 위해 추출된 cDNA에 대해 2 라운드의 PCR를 실시하였다. 이어서, Illumina NextSeq 또는 MiSeq 상에서 라이브러리에 대해 차세대 시퀀싱을 실시하였다. 이어서, 시퀀싱창 내의 모든 아데노신에서 RNA 편집율을 평가하였다.

[1368] 루시퍼라제 리포터를 RNA 편집에 대해 표적화한 실험에서, 우리는 또한 RNA 회수 이전에 분비된 루시퍼라제가 존재하는 배지를 회수하였다. 이 경우에, 보정된 Cluc은 저수준일 수 있기 때문에, 우리는 배지를 희석시키지 않았다. 우리는 주입 프로토콜에 따라서 플레이트 판독기 (Bilteck Synergy Neo2) 상에서 BioLux 시프리디니아 및 Biolux 가우시아 루시퍼라제 분석 키트 (New England Biolabs)로 루시퍼라제 활성을 측정하였다. 수행한 모든 복제물은 생물학적 복제물이다.

[1369] PFS 결합 포유동물 스크린

[1370] 편집 효율에 대한 PFS의 기여도를 결정하기 위해, 625 ng의 PFS 표적 라이브러리, 4.7 ㎍의 가이드 및 2.35 ㎍의 REPAIR를 225 cm2 플라스크에 플레이팅된 HEK293FT 세포 상에서 공동형질감염시켰다. 플라스미드를 33 ㎕의 PLUS 시약 (Thermo Fisher Scientific)과 혼합하고 나서, Opti-MEM으로 533 ㎕가 되게 하고, 5분 동안 인큐베이션시키고 나서, 30 ㎕의 Lipofectamine 2000 및 500㎕의 Opti-MEM과 배합하고, 추가 5분 동안 인큐베이션시키고, 세포 상에 피펫팅시켰다. 형질감염 후 48시간에, RNeasy Plus 미니 키트 (Qiagen)로 RNA를 채취

하고 나서, 유전자 특이적 프라이머를 이용하여 qScript Flex (Quantabio)로 역전사시키고, NEBNext High-Fidelity 2X PCR 마스터 믹스 (New England Biolabs)를 사용하여 2 라운드의 PCR로 증폭시켜 Illumina 어댑터 및 샘플 바코드를 첨가하였다. 라이브러리를 Illumina NextSeq 상에서 시퀀싱하고 나서, 표적 아데노신에서의 RNA 편집율을 PFS 정체에 대해 맵핑하였다. 적용범위를 증가시키기 위해, PFS를 4개의 뉴클레오티드에 대해 컴퓨터로 붕괴시켰다. 각각의 PFS에 대해 REPAIR 편집율을 계산하였고, 차감된 상응하는 PFS에 대해 비표적화율로 생물학적 복제물에 대해 평균 내었다.

[1371] ADAR 편집 특이성을 평가하기 위한 전체-전사체 시퀀싱

[1372] 전사체 전반에서 오프-표적 RNA 편집 부위를 분석하기 위해, 우리는 RNeasy Plus 미니프렙 키트(Qiagen)를 사용하여 형질감염 후 48시간에 세포로부터 전체 RNA를 채취하였다. 이어서, NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB)을 이용하여 mRNA 분획을 농축시키고, 이어서 NEBNext Ultra RNA Library Prep Kit for Illumina (NEB)을 사용하여 시퀀싱을 위한 이러한 RNA를 제조한다. 이어서, 라이브러리를 Illumina NextSeq 상에서 시퀀싱하고, 샘플 당 적어도 5백만개의 판독이 생기도록 로딩하였다.

[1373] 표적화된 그리고 전사체 와이드 실험을 위한 RNA 편집

[1374] 전사체-와이드 RNA 편집 RNA 시퀀싱 데이터의 분석을 위해, 서열 파일을 5백만개의 판독에 대해 무작위로 다운 샘플링하였다. Gluc 및 Cluc 서열이 첨가된 RefSeq GRCh38 어셈블리를 사용하여 인덱스를 생성시켰고 판독치를 Bowtie/RSEM version 1.3.0을 이용하여 정렬 및 정량하였다. 이어서, 정렬 BAM을 분류하고, 하기 매개변수로 REDitools)[cite]를 이용하여 RNA 편집 부위에 대해 분석하였다: -t 8 -e -d -l -U [AG 또는 TC] -p -u -m20 -T6-0 -W -v 1 -n 0.0. 비형질감염 또는 EGFP-형질감염 조건에서 발견된 임의의 유의한 편집은 SNP 또는 형질감염의 인공물인 것으로 간주하고, 오프-표적 분석에서 필터링하였다. 피셔의 정확도 검정이 0.5 미만의 p-값을 산출하고 3개 생물학적 복제물 중 적어도 2개가 편집 부위를 식별하였으면 오프-표적을 유의하다고 간주하였다.

[1375] 각각의 오프-표적의 예측되는 변이체 효과를 분석하기 위해서, SIFT 및 PolyPhen-2 주석을 이용하는 도구의 UCSC 게놈 브라우저 스위트의 일부분으로서 변이체 주석 통합기 (integrator)(https://genome.ucsc.edu/cgi-bin/hgVai)를 이용하여 오프-표적 편집 부위의 목록을 분석하였다. 오프-표적 유전자가 종양원성인지 여부를 명확히 하기 위해서, 암의 체세포 돌연변이의 COSMIC 카탈로그로부터의 종양유전자 주석의 데이터베이스를 참조한다 (cancer.sanger.ac.uk).

[1376] REPAIR 구성체가 RNA 수준을 동요시키는지 여부를 분석하기 위해서, RSEM 분석으로부터의 백만개당 전사물 (TMP) 값 출력치를 발현 계측에 대해 사용하였고 log2(TPM+1)를 취하여 log-스페이스로 전환시켰다. 차등적으로 조절된 유전자를 찾기 위해서, 3개의 표적화 가이드 복제물 대 3개의 비표적화 가이드 복제물에 대해 스튜던트 t-검정을 수행하였다. 2.5 초과의 log2(TPM+1) 값으로 유전자에 대해 통계학적 분석만을 수행하였고, 그들이 2 초과 또는 0.8 미만의 배수 벼화를 가지면, 유전자는 단지 차등적으로 조절되는 것으로 간주되었다. 유전자는 그들이 0.01 미만의 거짓 발견율을 갖는지 여부가 보고되었다.

[1377] 결과

[1378] 포유동물 세포에서 Cas13 패밀리 구성원의 종합적 특징규명

[1379] 본 발명자들은 이전에 포유동물 녹다운 적용을 위해 LwaCas13a를 개발하였지만, 효율적인 녹다운을 위해 msfGFP 안정화 도메인을 필요로 하였고, 특이성이 높음에도 불구하고, 녹다운 효율은 일관적으로 50% 미만이 아니었다 (*15*). 우리는 포유동물 세포에서 그들의 RNA 녹다운 활성을 평가하기 위해 Cas13 패밀리 멤버의 유전적으로 다양한 세트를 특징규명함으로써 더 강한 RNA-표적화 CRISPR 시스템을 확인하고자 추구하였다 (도 49A). 우리는 단백질 안정성을 증강시키기 위해 N-말단 및 C-말단의 핵 유출 신호 (NES) 서열 및 C-말단의 msfGFP를 갖는 발현 벡터에 21개의 Cas13a, 15개의 Cas13b 및 7개의 Cas13c 포유동물 코돈-최적화된 오솔로그 (표 6)를 클로닝시켰다. 포유동물 세포에서 간섭을 분석하기 위해, 우리는 별개의 프로모터 하에 직교성 가우시아 (Gluc) 및 시프리디니아 (Cluc) 루시퍼라제를 발현시키는 이중 리포터를 디자인하였고 이것은 한 루시퍼라제가 Cas13 간섭 활성의 측정으로서 작용하도록 하고, 다른 루시퍼라제는 내부 대조군으로서 작용하게 한다. 각각의 오솔로그에 대해, 우리는 암피실린 간섭 분석으로부터 유래된 Cas13b PFS 모티프 (도 55A-55C; 표 7; 보충주 1) 및 Cas13a 활성의 이전의 보고로부터의 3' H PFS(*10*)을 이용하여, PFS-적합 가이드 RNA를 디자인하였다.

[1380] 우리는 Cas13 발현, 가이드 RNA 및 리포터 플라스미드로 HEK293FT 세포를 형질감염시키고 나서, 48시간 후에 표적화된 Gluc의 수준을 정량하였다. 최근에 특징규명된 LwaCas13a에 비해 양쪽 가이드 RNA 전반에서 유사하

거나 또는 증가된 간섭을 갖는 5종의 Cas13b 오솔로그를 포함하여, 각각 Cas13 오솔로그에 대한 2개 가이드 RNA의 시험은 다양한 활성 수준을 나타내었다 (도 49B). 우리는 추가 조작을 위해 이들 5종 Cas13b 오솔로그뿐만 아니라 상위 2종 Cas13a 오솔로그를 선택하였다.

[1381] 우리는 다음에 강한 활성에 대해 안정화 도메인을 필요로 하지 않는 오솔로그를 선택하기 위해 msfGFP 없이 Gluc의 Cas13-매개된 녹다운을 시험하였다. 우리는 msfGFP에 추가로, Cas13 활성이 LwaCas13a의 최적화에 대해 이전에 보고된 바와 같이 (*15*) 세포하 국재화에 의해 영향받을 수 있다는 가설을 세웠다. 따라서, 본 발명자들은 msfGFP 없이 6개의 상이한 국재화 태그 중 하나에 C-말단이 융합된 7종의 선택된 Cas13 오솔로그의 간섭 활성을 시험하였다. 루시퍼라제 리포터 분석을 이용하여, 본 발명자들은 HIV Rev 유전자 NES에 C-말단이 융합된 PspCas13b 및 PguCas13b 및 MAPK NES에 C-말단이 융합된 RanCas13b가 가장 고수준의 간섭 활성을 갖는다는 것을 발견하였다 (도 56A). 상위 오솔로그의 활성 수준을 추가로 구별하기 위해, 우리는 위치-매칭된 가이드를 이용하여 *KRAS* 전사물을 녹다운시키는 능력에 대해 3종의 최적화된 Cas13b 구성체를 최적의 LwaCas13a-msfGFP 융합체 및 shRNA와 비교하였다 (도 56B). 우리는 PspCas13b에 대해 최고 수준의 간섭을 관찰하였고 (평균 녹다운 62.9%) 따라서 LwaCas13a와의 추가 비교를 위해 이것을 선택하였다.

[1382] PspCas13b 및 LwaCas13a의 활성 수준을 더 엄격하게 정하기 위해, 우리는 Gluc와 Cluc 둘 모두를 따라 위치 매칭된 가이드 타일링을 디자인하였고, 우리의 루시퍼라제 리포터 어세이를 사용해 그들 활성을 검정하였다. 우리는 Gluc 및 Cluc를 각각 표적화하는 93 및 20 위치 매칭된 가이드를 시험하였고, PspCas13b가 LwaCas13a에 비해 녹다운 수준을 일관적으로 증가시켰다는 것을 확인하였다 (PspCas13b에 대해 평균 92.3% 대 LwaCas13a에 대해 평균 40.1% 녹다운) (도 49C, 49D).

[1383] Cas13 포유동물 간섭 활성의 특이성

[1384] PspCas13b 및 LwaCas13a의 간섭 특이성을 특징규명하기 위해, 우리는 표적 서열 전반에 단일 미스매치 및 이중 미스매치 및 3개의 측접하는 5' 및 3' 염기쌍을 함유하는 루시퍼라제 표적의 플라스미드 라이브러리를 디자인하였다 (도 56C). 우리는 LwaCas13a 또는 PspCas13b, 비변형 표적 서열을 표적화하는 고정된 가이드 RNA, 및 적절한 시스템에 상응하는 미스매치된 표적 라이브러리로 HEK293FT 세포를 형질감염시켰다. 이어서, 우리는 미스매치된 표적 서열의 감손을 정량하기 위해 비절단 전사물의 표적화된 RNA 시퀀싱을 수행하였다. 우리는 LwaCas13a 및 PspCas13b가 PspCas13b 표적의 경우에 12 내지 26 염기쌍 및 LwaCas13a 표적의 경우 13 내지 24 염기쌍으로부터 연장된, 단일 미스매치에 대해 비교적 비용인된 중심 영역을 갖는다는 것을 확인하였다 (도 56D). 이중 미스매치는 단일 돌연변이에 비해 훨씬 덜 용인되었는데, 그들 개별 표적에서 PspCas13b의 경우 12 내지 29 염기쌍 및 LwaCas13a의 경우 8 내지 27 염기쌍으로 연장되는, 더 큰 창에서 관찰된 녹다움 활성이 거의 없었다 (도 56E). 추가적으로, 표적 서열의 5' 및 3' 말단부에 측접하는 3개의 뉴클레오티드에 포함되는 미스매치가 있기 때문에, 우리는 Cas13 녹다운 활성에 대한 PFS 제약을 평가할 수 있었다. 시퀀싱은 거의 모든 PFS 조합이 강한 녹다운을 허용한다는 것을 보여주어서, 포유동물 세포에서 간섭에 대한 PFS 제약은 아마도 시험된 각 효소의 경우에는 존재하지 않을 수도 있다는 것을 의미한다. 이들 결과는 Cas13a 및 Cas13b가 유사한 서열 제약 및 미스매치에 대한 민감도를 나타낸다는 것을 의미한다.

[1385] 우리는 다음으로 전사체의 mRNA 분획 전반에서 PspCas13b 및 LwaCas13a의 간섭 특이성을 특징규명하였다. 우리는 상당히 차등적으로 발현되는 유전자를 검출하기 위해 전사체-와이드 mRNA 시퀀싱을 수행하였다. LwaCas13a 및 PspCas13b는 Gluc의 강한 녹다운을 입증하였고 (도 49E, 49F), 위치-매칭된 shRNA와 비교하여 고도로 특이적이었고, 수백개의 오프-표적을 확인하였다 (도 49G).

[1386] Cas13-ADAR 융합체는 표적화 RNA 편집을 가능하게 한다

[1387] PspCas13b가 포유동물 세포에서 mRNA의 일관적이고, 강건하며 특이적인 녹다운을 획득한다는 것을 고려하면, 우리는 이것이 프로그램 가능한 RNA 편집을 위해 ADAR의 데아미나제 도메인 ($ADAR_{DD}$)을 동원하기 위한 RNA 결합 플랫폼으로서 개조될 수 있다고 계획하였다. 뉴클레아제 활성이 결여된 PspCas13b를 조작하기 위해 (dPspCas13b, 여기서는 dCas13b라고 함), 우리는 HEPN 도메인의 보존된 촉매 잔기를 돌연변이시켰고, 루시퍼라제 RNA 녹다운 활성의 결실을 관찰하였다 (도 57A). 우리는 dCas13b-$ADAR_{DD}$ 융합체가 ADAR 활성을 위해 필요한 듀플렉스 기질을 생성하는 하이브리드화된 RNA와 함께, 아데노신을 표적화하기 위해 가이드 RNA를 통해 동원될 수 있다고 가정하였다 (도 50A). 표적 아데노신 탈아미노화 비율을 증강시키기 위해, 우리는 우리의 초기 RNA 편집 디자인에 2개의 추가 변형을 도입하였다: 우리는 이전에 탈아미노화 빈도를 증가시킨다고 보고된 표적 아데노신 반대쪽 미스매치된 시티딘을 도입시켰고, 촉매 활성을 증강시키기 위해 과활성화 돌연변이

를 함유하는 인간 ADAR1 또는 ADAR2의 데아미나제 도메인에 dCas13b를 융합시켰다 (ADAR1$_{DD}$(E1008Q)(25) 또는 ADAR2$_{DD}$(E488Q)(19)).

[1388] dCas13b-ADAR$_{DD}$의 활성을 시험하기 위해, 우리는 넌센스 돌연변이(W85X(UGG->UAG))를 도입시켜 Clu 상에 RNA-편집 리포터를 생성시켰는데, 이것은 A->I 편집을 통해 야생형 코돈에 대해 기능적으로 복구될 수 있고 (도 50B), Cluc 발광성의 복원으로서 검출될 수 있다. 우리는 최적 가이드 배치 및 디자인을 결정하기 위해 표적 아데노신에 걸쳐 30, 50, 70 또는 84개 뉴클레오티드 길이의 스페이서로 가이드를 균일하게 타일링하였다 (도 50C).

우리는 Cluc 리포터를 복구하기 위해 dCas13b-ADAR1$_{DD}$가 더 긴 가이드를 필요로 하는 한편, dCas13b-ADAR2$_{DD}$는 시험된 모든 가이드 길이에서 기능성이라는 것을 발견하였다 (도 50C). 우리는 또한 야생형 ADAR2$_{DD}$에 의한 루시퍼라제 복원이 감소됨에 따라 과활성 E488Q 돌연변이가 편집 효율을 개선시켰다는 것을 발견하였다 (도 57B). 이러한 활성 입증을 통해서, 우리는 추가 특징규명을 위해 dCas13b-ADAR2$_{DD}$(E488Q)를 선택하였고, 이 접근법을 REPAIRv1 (RNA Edit for Programmable A to I Replacement version 1)이라고 명명하였다.

[1389] 루시퍼라제 활성의 복원이 진실된 편집 사건에 기인한다는 것을 입증하기 위해, 우리는 역전사 및 표적화 차세대 시퀀싱을 통해 직접적으로 REPAIRv1에 대한 Cluc 전사물의 편집을 측정하였다. 우리는 표적 부위 주변 30-nt 및 50-nt 스페이서를 시험하였고, 양쪽 가이드 길이가 예상되는 A > I 편집을 일으켰다는 것을 확인하였는데, 50-nt 스페이서가 더 높은 편집율을 획득하였다 (도 50D, 50E, 도 57C). 우리는 또한 50-nt 스페이서가 아마도 듀플렉스 RNA의 증가된 영역에 기인하여, 비표적화된 아데노신의 편집에 높은 경향을 가졌다는 것을 관찰하였다 (도 50E, 도 57C).

[1390] 우리는 다음으로 내생성 유전자인 *PPIB*를 표적화하였다. 우리는 *PPIB*를 타일링하는 50-nt 스페이서를 디자인하였고, 최대 28%의 편집 효율로 *PPIB* 전사물을 편집할 수 있다는 것을 확인하였다 (도 57D). REPAIR를 더욱 최적화할 수 있는가를 시험하기 위해서, 우리는 dCas13b와 ADAR2$_{DD}$(E488Q) 사이의 링커를 변형시켰고 (도 57E, 표 8) 링커 선택이 루시퍼라제 활성 복원에 중간정도로 영향을 미친다는 것을 확인하였다.

[1391] RNA 편집을 위한 서열 매개변수의 정의

[1392] 시험 부위에서 정확한 RNA 편집을 달성할 수 있다는 것을 고려하면, 우리는 전사체 내 임의 RNA 표적에 대해 시스템을 프로그래밍하기 위한 서열 제약을 특징규명하고 싶었다. 서열 제약은 dCas13b 표적화 한계, 예컨대 PFS, 또는 ADAR 서열 선호도에서 발생될 수 있었다 (*26*). REPAIRv1에 대한 PFS 제약을 조구하기 위해, 우리는 Cluc 전사물 상의 표적 부위의 5' 말단에 일련의 4종의 무작위 뉴클레오티드를 운반하는 플라스미드 라이브러리를 디자인하였다 (도 51A). 우리는 UAG 또는 AAC 모티프 내 중심 아데노신을 표적화하였고 양쪽 모티프 경우에, 모든 PFS가 검출가능한 수준의 RNA 편집을 입증하였다는 것을 확인하였으며, 대부분으 PFS는 표적 부위에서 50% 초과의 편집율을 가졌다 (도 51B). 다음으로, 우리는 ADAR2$_{DD}$에 대해 이전에 보고된 바와 같이 (*26*), REPAIRv1의 ADAR2$_{DD}$가 표적화된 염기에 바로 측접하는 임의의 서열 제약을 갖는가를 결정하고자 하였다.

우리는 표적 아데노신을 직접적으로 둘러싼 5' 및 3' 측접 뉴클레오티드의 모든 가능한 조합을 시험하였고 (도 51C), REPAIRv1이 모든 모티프를 편집할 수 있었다는 것을 발견하였다 (도 51D). 마지막으로, 우리는 스페이서 서열 내 표적 A 반대쪽 염기의 정체가 편집 효율에 영향을 미치는지 여부를 분석하였고 A-C 미스매치가 극적으로 감소된 REPAIRv1 활성을 갖는 A-G, A-U 및 A-A에 의해 가장 높은 루시퍼라제 복원성을 가졌다는 것을 확인하였다 (도 57F).

[1393] REPAIRv1을 사용한 질환-관련 인간 돌연변이의 교정

[1394] 포유동물 세포에서 RNA 편집을 위한 REPAIRv1 시스템의 광범위 적용가능성을 입증하기 위해, 우리는 하기의 2종 질환 관련 돌연변이에 대해 REPAIRv1 가이드를 디자인하였다: X-연관 신장성 요붕증의 878G>A (*AVPR2* W293X) 및 판코니 빈혈증의 1517G>A (*FANCC* W506X). 우리는 이들 돌연변이를 운반하는 유전자의 cDNA를 위한 발현 구성체로 HEK293FT 세포를 형질감염시켰고, REPAIRv1이 돌연변이를 교정할 수 있는지의 여부를 시험하였다. 50-nt 스페이서를 함유하는 가이드 RNA를 사용하여, 우리는 *AVPR2*의 35% 교정 및 *FANCC*의 23% 교정을 달성할 수 있었다 (도 52A-도 52D). 다음으로 우리는 34개의 상이한 질환-관련 G>A 돌연변이를 교정하는 REPAIRv1의 능력을 시험하였고 (표 9) 우리가 최대 28%의 편집 효율로 33개 부위에서 유의한 편집을 달성할 수 있었다는 것을 확인하였다 (도 52E). 우리가 선택한 돌연변이는 ClinVar 데이터베이스에서 병원성 G>A 돌연변이의 유일

한 부분 (5,739)이며, 또한 추가 11,943개 G>A 변이체를 포함한다 (도 52F 및 도 58). 서열 제약이 존재하지 않기 때문에, 특히 우리가 표적 모티프와 무관하게 유의한 편집을 관찰하였다는 것을 고려하면, REPAIRv1는 모든 이들 질환 관련 돌연변이를 잠재적으로 편집할 수 있다 (도 51C 및 도 52G).

[1395] 질환 세포에 REPAIRv1 시스템의 전달은 치료적 용도를 위한 전제 조건이므로, 우리는 치료적으로 적절한 바이러스 벡터, 예컨대 아데노-연합 바이러스 (AAV) 벡터에 패키징될 수 있는 REPAIRv1 구성체를 디자인하고자 하였다. AAV 벡터는 4.7kb의 패키징 한계를 가져서, 프로모터 및 발현 조절 구성요소와 함께 거대 크기의 dCas13b-$ADAR_{DD}$(4473 bp)를 수용할 수 없다. 크기를 감소시키기 위해, 우리는 RNA 편집 활성을 위해 $ADAR2_{DD}$(E488Q)에 융합된 dCas13의 다양한 N-말단 및 C-말단의 절단을 시험하였다. 우리는 시험한 모든 C-말단 절단이 여전히 기능성이었고 루시퍼라제 신호를 복원할 수 있으며 (도 59), 최대 절단, C-말단 Δ984-1090 (총 4,152bp 융합단백질 크기)이 AAV 벡터의 패키징 한계 내에 맞추기에 충분히 작다는 것을 확인하였다.

[1396] <u>REPAIRv1의 전사체-와이드 특이성</u>

[1397] PspCas13b에 의한 RNA 녹다운이 고도로 특이적이지만, 우리의 루시퍼라제 타일링 실험에서, 우리는 가이드:표적 듀플렉스 내에서 오프-표적 아데노신 편집을 관찰하였다 (도 50E). 이것이 널리 퍼진 현상인지의 여부를 알아보기 위해, 우리는 내생성 전사물, *KRAS*를 타일링하였고, 표적 아데노신 근처에서 오프-표적 편집 정도를 측정하였다 (도 53A). 우리는 *KRAS* 경우에, 온-표적 편집율이 23%였지만, 또한 검출 가능한 A>G 편집을 갖는 표적 주위 주변에 많은 부위가 존재하였다는 것을 확인하였다 (도 53B).

[1398] 가이드:표적 듀플렉스 내에서 관찰된 오프-표적 편집 때문에, 우리는 모든 mRNA에 대해 RNA 서열을 수행하여 모든 가능한 전사체 오프-표적을 평가하였다. RNA 시퀀싱은 상당한 수의 A>G 오프-표적 사건이 존재하였다는 것을 밝혀주었는데, 표적화 조건에서 1,732개의 오프-표적 및 비-표적화 조건에서 925개의 오프-표적이 있었고, 828개 오프-표적이 중복되었다 (도 53C, 53D). 전사체 전반에서 모든 편집 부위 중에, 온-표적 편집 부위는 최고 편집율을 가졌는데, 89%의 A>G 전환율이었다.

[1399] Cas13 표적화의 높은 특이성을 고려하면, 우리는 오프-표적이 ADAR로부터 발생될 수 있다고 추론하였다. 두 RNA-가이드 ADAR 시스템이 이전에 기재된 바 있다 (도 60A). 제1 디자인은 BoxB-λ RNA 헤어핀에 결합하는 소형 바이러스 단백질 람다N (λN)과 $ADAR2_{DD}$의 융합을 이용한다 (*22*). 이중 BoxB-λ 헤어핀을 갖는 가이드 RNA는 가이드 RNA에 코딩된 부위를 편집하기 위해 $ADAR2_{DD}$를 가이드한다 (*23*). 제2 디자인은 ADAR2의 이중 가닥 RNA 결합 도메인 (dsRBD)이 인식하는 헤어핀을 갖는 가이드 RNA 및 전체 길이 ADAR2(ADAR2)를 이용한다 (*21, 24*). 우리는 REPAIRv1과 비교하여 이들 두 시스템의 편집 효율을 분석하였고, BoxB-ADAR2 및 ADAR2 시스템이 REPAIRv1에 의해 달성된 89% 편집율과 비교하여 63% 및 36% 편집율을 입증하였다는 것을 발견하였다 (도 60B-60E). 추가적으로, BoxB 및 ADAR2 시스템은 REPAIRv1 표적화 가이드 조건에서 1,229개의 오프-표적과 비교하여, 표적화 가이드 조건에서 각각 2018 및 174개의 관찰된 오프-표적을 생성하였다. 특히, 2개의 $ADAR2_{DD}$-기반 시스템 (REPAIRv1 및 BoxB)에 의한 모든 조건은 그들의 오프-표적에서 높은 중복율을보인데 반해서 ADAR2 시스템은 대체로 구별되는 오프-표적 세트를 가졌다 (도 60F). 표적화 조건과 비표적화 조건 사이 및REPAIRv1과 BoxB 조건 사이의 오프-표적 중복은 $ADAR2_{DD}$가 dCas13 표적화와 독립적으로 오프-표적을 구동한다는 것을 시사한다 (도 60F).

[1400] <u>합리적 단백질 조작을 통한 REPAIRv1의 특이성 개선</u>

[1401] REPAIR의 특이성을 개선시키기 위해, 우리는 $ADAR2_{DD}$(E488Q)의 구조-가이드된 단백질 조작을 사용하였다. 오프-표적의 가이드-독립적 특성 때문에, 우리는 $ADAR2_{DD}$(E488Q)-RNA 결합의 탈안정화가 오프-표적 편집을 선택적으로 감소시키지만, $ADAR2_{DD}$(E488Q)의 dCas13b 속박으로부터 표적 부위까지의 증가된 국소 농도로 인해 온-표적 편집을 유지한다고 가정하였다. 우리는 $ADAR2_{DD}$(E488Q) 배경에서 표적 RNA의 듀플렉스 영역과 접촉시키기 위해 이전에 결정한 $ADAR2_{DD}$(E488Q) 잔기를 돌연변이시켰다 (도 54A)(18). 효율 및 특이성을 평가하기 위해, 우리는 검출된 비표적화 조건에서 배경 루시퍼라제 복원이 더 넓은 오프-표적 활성을 의미한다는 가정 하에, 표적화 및 비표적화 가이드로 17개 단일 돌연변이체를 시험하였다. 우리는 선택된 잔기의 돌연변이가 표적화 및 비표적화 가이드에 대한 루시퍼라제 활성에 대해 상당한 효과를 갖는다는 것을 발견하였다 (도 54A, 54B, 도 61A). 대다수의 돌연변이체는 표적화 가이드에 대한 루시퍼라제 활성을 상당히 개선시키거나 또는 우리가 특이성 점수라고 한 표적화 대 비표적화 가이드 활성의 비율을 증가시켰다 (도 54A, 54B). 우리는

차세대 시퀀싱에 의한 전사체-와이드 특이성 프로파일링을 위해 이들 돌연변이체의 서브세트를 선택하였다 (도 54B). 예상대로, 전사체-와이드 시퀀싱에서 측정된 오프-표적은 돌연변이체에 대한 우리의 특이성 점수와 상관성이 있었다 (도 61B). 우리는 $ADAR2_{DD}$(E488Q/R455E)를 제외하고, 모든 시퀀싱된 REPAIRv1 돌연변이체가 리포터 전사물을 효과적으로 편집할 수 있다는 것을 확인하였고 (도 54C), 많은 돌연변이체가 오프-표적 개수의 감소를 보였다 (도 61C, 62). 우리는 특이성 돌연변이체에 대한 오프-표적의 주변 모티프를 추가로 검토하였고, REPAIRv1 및 대부분의 조작된 돌연변이체가 특징규명된 ADAR2 모티프와 일치되게, 그들의 편집을 위해 강력한 3' G 선호도를 나타내었다는 것을 발견하였다 (도 63A)(*26*). 우리는 향후 실험을 위해 돌연변이체 돌연변이체 $ADAR2_{DD}$(E488Q/T375G)를 선택하였는데, 이것이 최저 개수의 전사체-와이드 오프-표적을 가지면서 4개 돌연변이체 중에서 최고 편집율을 가지기 때문이며, 이것을 REPAIRv2라고 명명하였다. REPAIRv1과 비교하여, REPAIRv2는 1732내지 10의 전사체 오프-표적 감소와 함께 증가된 특이성을 나타내었다 (도 54D). Cluc 내 표적화된 아데노신을 둘러싼 영역에서, REPAIRv2는 오프-표적 편집을 감소시켰는데, 이것은 시퀀싱 추적으로 가시화된다 (도 54E). 차세대 시퀀싱으로부터 유래된 모티프에서, REPAIRv1은 3' G쪽으로 강력한 선호도를 나타내었지만, 모든 모티프에 대해 오프-표적 편집을 보였으며 (도 63B); 대조적으로, REPAIRv2는 오직 가장 강력한 오프-표적 모티프만을 편집하였다 (도 63C). 전사물 상에서 편집 분포는 심하게 왜곡되었는데, 고도의-편집 유전자는 60개 초과 편집을 갖는 반면 (도 64A, 64B), REPAIRv2는 오직 하나의 전사물 (*EEF1A1*)를 다수회 편집하였다 (도 64D-64F). REPAIRv1 오프-표적 편집은 93회의 종양원성 사건 (도 64D)과 함께, 1000개 미스센스 돌연변이 (도 64C)를 포함하는, 수많은 변이체를 생성시킨 것으로 예측되었다. 대조적으로, REPAIRv2는 오직 6개 미스센스 돌연변이 (도 64E)를 갖지만, 이중 어떠한 것도 종양원성 결과를 갖지 않았다 (도 64F). 예측된 오프-표적 효과의 이러한 감소는 REPAIRv2를 다른 RNA 편집 접근법과 구별한다. 상이한 용량의 가이드 RNA를 사용한 실험은 용량 반응이 오프 표적 활성을 감소시킬 수 있다는 것을 시사한다 (도 68).

[1402] 우리는 특이성-증강 돌연변이가 온-표적 편집에 대해 고효율을 유지하면서 표적 전사물 내 근처 편집을 감소시켰는지 여부를 시험하기 위해 REPAIRv2를 내생성 유전자에 대해 표적화시켰다. *KRAS* 또는 *PPIB* 중 어느 하나를 표적화하는 가이드의 경우, 우리는 REPAIRv2가 REPAIRv1과 달리, 검출 가능한 오프-표적 편집을 갖지 않으며, 각각 27.1% 및 13%로 온-표적 아데노신을 효율적으로 편집할 수 있었다는 것을 확인하였다 (도 54F, 54G). 이러한 특이성은 다른 *KRAS* 또는 *PPIB* 표적 부위와 같이, REPAIRv1에 대한 비표적화 조건에서 높은 수준의 배경을 입증하는 영역을 포함하여, 추가 표적 부위로 연장되었다 (도 65A-65C). 전체적으로, REPAIRv2는 편집된 아데노신 주변의 듀플렉스 영역에서 오프-표적을 제거하였고 극적으로 증강된 전사체-와이드 특이성을 보였다.

[1403] 결론

[1404] 우리는 본 명세서에서여기서 RNA-가이드된 RNA-표적화 VI-B형 이펙터 Cas13b이 고도로 효율적이고 특이적인 RNA 녹다운을 가능하게 하여, 필수 유전자 및 넌코딩 RNA의 정보를 얻고 전사체 수준에서 세포 과정을 제어하기 위한 개선된 도구의 기초를 제공한다는 것을 확인하였다. 촉매적 불활성 Cas13b (dCas13b)가 프로그램 가능한 RNA 결합 능력을 보유하는데, 여기에 우리는 우리가 REPAIRv1(RNA Editing for Programmable A to I Replacement version 1)라고 명명한 시스템, 정확한 A>I 편집을 달성하기 위해 dCas13b를 아데노신 데아미나제 ADAR2에 융합시켜서 영향을 주었다. 시스템의 추가적인 조작은 REPAIRv2를 생성하였고, 현행 편집 플랫폼에 비해서 비슷하거나 또는 증가된 활성을 갖는 방법은 극적으로 특이성이 개선되었다.

[1405] Cas13b는 높은 충실도를 나타내지만, dCas13b-$ADAR2_{DD}$ 융합에 의한 우리의 초기 결과는 수천개의 오프-표적을 밝혀내었다. 이것을 해결하기 위해서, 우리는 RNA 듀플렉스과 접촉되는 $ADAR2_{DD}$ 잔기를 다양화시키도록 합리적인 돌연변이유발 전략을 적용하였고, dCas13b와 융합했을 때, 정확하고, 효율적으로, 매우 특이적인 편집을 할 수 있는 변이체, $ADAR2_{DD}$(E488Q/T375G)를 동정하였다. 이 변이체에 의한 편집 효율은 2종의 현재 이용가능한 시스템, BoxB-$ADAR_{DD}$ 또는 ADAR2 편집으로 획득된 것에 비해 비슷하거나 또는 더 양호하였다. REPAIRv2 시스템은 대안적인 양쪽 편집 기술에 비해서 적어도 한자릿수가 더 양호하게, 전체 전사체에서 오직 10개의 관찰가능한 오프-표적을 생성시켰다.

[1406] REPAIR 시스템은 다른 핵산 편집 도구와 비교하여 다수의 장점을 제공한다. 첫째로, ADAR 편집 활성에 이상적인 바람직한 A-C 미스매치를 생성시키기 위해 바람직한 아데노신으로부터 가이드 연장부 내에 시티딘을 위치시킴으로써 정확한 표적 부위가 가이드 내에서 코딩될 수 있다. 둘째로, Cas13은 표적화 서열 제약, 예컨대

PFS 또는 PAM을 갖지 않고, 표적 아데노신 주변 모티프 선호도를 갖지 않아서, 전사체 내 임의 아데노신이 REPAIR 시스템에 의해 강력하게 표적화될 수 있게 한다. 그러나, 우리는 DNA 염기 편집자가 센스 또는 안티센스 가닥을 표적화할 수 있는 한편, REPAIR 시스템이 전사된 서열로 제한됨으로써 우리가 표적화할 수 있는 가능한 편집 부위의 총 개수에 제약이 있을 수 있다는 것을 유의한다. 하지만, REPAIR에 의한 표적화의 보다 탄력적인 특성에 기인하여, 이 시스템은 Cas9-DNA 염기 편집자에 비해서 ClinVar (도 52C) 내에서 더 많은 편집을 실시할 수 있다. 셋째로, REPAIR 시스템은 표적 아데노신을 이노신으로 직접적으로 탈아미노화시키고, 목적하는 편집 결과를 생성하기 위해 내생성 복구 경로, 예컨대 염기-절단 또는 미스매치 복구에 의존하지 않는다. 따라서, REPAIR은 다른 편집 형태를 지원할 수 없는 비-분열 세포에서 가능해야 한다. 넷째로, RNA 편집은 일시적일 수 있어서, 편집 결과에 시간적 제어 가능성을 허용한다. 이러한 특성은 아마도 세포 상태의 일시적 변화에 의해 야기되는 질환, 예컨대 국소 염증을 치료하는 데 유용할 수 있을 것이다.

[1407] REPAIR 시스템은 추가 조작을 위한 다수 기회를 제공한다. Cas13b는 전-crRNA 가공 활성을 가져서 (*13*), 단독으로 질환 위험을 변경시키지 않을 수 있지만, 함께 가산 효과 및 질환-변형 가능성을 가질 수 있는, 다수 변이체의 다중복합 편집을 가능하게 한다. 우리의 합리적 디자인 접근의 확장, 예컨대 유망한 돌연변이의 조합은 시스템의 특이성 및 효율을 더 증가시킬 수 있는 한편, 비편향된 스크리닝 접근법은 REPAIR 활성 및 특이성을 개선시키기 위한 추가 잔기를 동정할 수 있었다.

[1408] 현재, REPAIR에 의해 획득가능한 염기 전환은 아데노신에서 이노신의 생성에 국한되며; 다른 촉매적 RNA 편집 도메인, 예컨대 APOBEC와 dCas13의 추가 융합은 시티딘에서 우리딘으로의 편집을 가능하게 할 수 있었다. 추가적으로, ADAR의 돌연변이유발은 표적 시티딘에 대한 기질 선호도를 완화시켜서, C->U 편집자에 의해 이용되는 듀플렉스 RNA 기질 요건에 의해 부여되는 증강된 특이성을 가능하게 한다. DNA 기질 상에서의 아데닌의 이노신으로의 편집은 DNA-RNA 이형 듀플렉스 표적의 형성 (*27*) 또는 ADAR 도메인의 돌연변이유발을 통해, 촉매적 불활성 DNA-표적화 CRISPR 이펙터, 예컨대 dCas9 또는 dCpf1에 의해 가능할 수 있다.

[1409] REPAIR은 다양한 치료적 적응증에 대해 적용될 수 있으며, 여기서는 A >I (A > G) 편집이 질환 진행을 반전시키거나 또는 늦출 수 있다 (도 66A-66G). 먼저, 질환-관련 조직에서 원인성 멘델 G >A 돌연변이를 표적화하기 위한 REPAIR의 발현은 유해한 돌연변이를 반전시키고 질환을 치료하는데 사용될 수 있었다. 예를 들어, 뇌 조직에서 AAV를 통한 안정한 REPAIR 발현은 초점 뇌전증을 야기하는 *GRIN2A* 미스센스 돌연변이 c.2191G>A (Asp731Asn)(*28*) 또는 조기 발병 알츠하이머병을 야기하는 *APP* 미스센스 돌연변이 c.2149G>A(Val717Ile)(*29*)를 교정하기 위해 사용할 수 있었다. 둘째로, REPAIR은 질환-관련 신호전달에 관여되는 단백질의 기능을 변형시킴으로써 질환을 치료하는 데 사용될 수 있었다. 예를 들어, REPAIR 편집은 키나제의 표적인 일부 세린, 트레오닌 및 티로신 잔기의 재-코딩를 허용할 것이다 (도 66A-66G). 질환-관련 단백질에서 이들 잔기의 인산화는 알츠하이머병 및 다발성 신경퇴행성 병태를 포함한 많은 장애에 대한 질환 진행에 영향을 미친다 (*30*).

셋째로, REPAIR는 환자가 질환 상태로 들어갈 기회를 선제적으로 감소시키기 위해 발현된, 위험-변형 G> A 변이체 서열을 변화시키는 데 사용할 수 있었다. 가장 흥미로운 경우는 질환 상태로 들어갈 기회를 극적으로 감소시키고, 일부 경우에, 추가의 건강 이득을 부여하는, '보호적' 위험-변형 대립유전자이다. 예를 들어, REPAIR은 심혈관계 질환 및 건선성 관절염에 대해 각각 보호하는 *PCSK9* 및 *IFIH1*의 A> G 대립유전자를 기능적으로 모방하기 위해 사용할 수 있었다 (*31*). 마지막으로, REPAIR는 엑손 조절 요법을 위한 스플라이스 억셉터 및 도너 부위를 치료적으로 변형시키는 데 사용될 수 있다. REPAIR는 각각 공통 5' 스플라이스 도너 또는 3' 스플라이스 억셉터 부위와 기능적으로 동등한, AU > IU 또는 AA > AI를 변화시켜서, 새로운 스플라이스 접합부를 생성할 수 있다. 추가적으로, REPAIR 편집은 DMD의 치료를 위해 상당한 관심을 받은 치료 전략인, 인접한 하류 엑손의 스키핑을 촉진시키기 위해 AG->IG의 공통 3' 스플라이스 억셉터 부위를 돌연변이시킬 수 있다. 스플라이스 부위의 조절은, 예컨대 척수성 근위축증의 치료를 위한 SMN2 스플라이싱의 조절을 위해 안티센스 올리고가 일부 성공적이었던 질환에서 광범위한 적용성을 가질 수 있었다 (*32*).

[1410] 우리는 RNA 녹다운과 RNA 편집 도구 둘 모두로서 PspCas13b 효소의 용도를 입증하였다. 프로그램가능한 RNA 결합을 위한 dCas13b 플랫폼은 라이브 전사물 영상화, 스플라이싱 변형, 전사물의 표적화 국재화, RNA-결합 단백질의 풀 다운 및 후성전사체 변형을 포함한, 많은 적용성을 가진다. 여기서 우리는 핵산 편집 기술의 기존 세트에 더하여, dCas13을 사용하여 REPAIR를 생성시켰다. REPAIR은 유전 질환을 치료하거나 보호 대립유전자를 모방하기 위한 새로운 접근을 제공하며, 유전자 기능을 변형시키기 위한 유용한 도구로서 RNA 편집을 확립한다.

[1411] 표 6. 본 연구에서 사용된 Cas13 오솔로그

Cas13 ID	Cas13 축약형	숙주 유기체	단백질 수탁번호
Cas13a1	LshCas13a	렙토트리키아 샤히	WP_018451595.1
Cas13a2	LwaCas13a	렙토트리키아 웨이데이(Lw2)	WP_021746774.1
Cas13a3	LseCas13a	리스테리아 실리게리	WP_012985477.1
Cas13a4	LbmCas13a	라크노스피라과 박테리아 MA2020	WP_044921188.1
Cas13a5	LbnCas13a	라크노스피라과 박테리아 NK4A179	WP_022785443.1
Cas13a6	CamCas13a	[클로스트리듐] 아미노필럼 DSM 10710	WP_031473346.1
Cas13a7	CgaCas13a	카르노박테리움 갈리나럼 DSM 4847	WP_034560163.1
Cas13a8	Cga2Cas13a	카르노박테리움 갈리나럼 DSM 4847	WP_034563842.1
Cas13a9	Pprcas13a	팔루디박터 프로피오니시게네스 WB4	WP_013443710.1
Cas13a10	LweCas13a	리스테리아 웨이헨스테파넨시스 FSL R9-0317	WP_036059185.1
Cas13a11	LbfCas13a	리스테리아과 박테리아 FSL M6-0635	WP_036091002.1
Cas13a12	Lwa2cas13a	렙토트리키아 웨이데이 F0279	WP_021746774.1
Cas13a13	RcsCas13a	로도박터 캡슐라터스 SB 1003	WP_013067728.1
Cas13a14	RcrCas13a	로도박터 캡슐라터스 R121	WP_023911507.1
Cas13a15	RcdCas13a	로도박터 캡슐라터스 DE442	WP_023911507.1
Cas13a16	LbuCas13a	렙토트리키아 부칼리스 C-1013-b	WP_015770004.1
Cas13a17	HheCas13a	허비닉스 헤미셀룰로실리티카	CRZ35554.1
Cas13a18	EreCas13a	[유박테리움] 렉탈레	WP_055061018.1
Cas13a19	EbaCas13a	유박테리아과 박테리아 CHKCI004	WP_090127496.1
Cas13a20	BmaCas13a	플라우티아 sp. 마르세이유-P2398	WP_062808098.1

[1412]

Cas13a21	LspCas13a	렙토트리키아 sp. 경구 분류군 879 str. F0557	WP_021744063.1
Cas13b1	BzoCas13b	베르게이엘라 주헬컴	WP_002664492
Cas13b2	PinCas13b	프레보텔라 인터메디아	WP_036860899
Cas13b3	PbuCas13b	프레보텔라 부카에	WP_004343973
Cas13b4	AspCas13b	알리스티페스 sp. ZOR0009	WP_047447901
Cas13b5	PsmCas13b	프레보텔라 sp. MA2016	WP_036929175
Cas13b6	RanCas13b	리에메렐라 아나티페스티퍼	WP_004919755
Cas13b7	PauCas13b	프레보텔라 아우란티아카	WP_025000926
Cas13b8	PsaCas13b	프레보텔라 사카롤리티카	WP_051522484
Cas13b9	Pin2Cas13b	프레보텔라 인터메디아	WP_061868553
Cas13b10	CcaCas13b	카프노사이토파가 카니몰서스	WP_013997271
Cas13b11	PguCas13b	포르피로모나스 굴라에	WP_039434803
Cas13b12	PspCas13b	프레보텔라 sp. P5-125	WP_044065294
Cas13b13	FbrCas13b	플라보박테리움 브란키오필럼	WP_014084666
Cas13b14	PgiCas13b	포르피로모나스 진지발리스	WP_053444417
Cas13b15	Pin3Cas13b	프레보텔라 인터메디아	WP_050955369
Cas13c1	FnsCas13c	푸소박테리움 네크로포럼 subsp. 푼둘리포르메 ATCC 51357 contig00003	WP_005959231.1
Cas13c2	FndCas13c	푸소박테리움 네크로포럼 DJ-2 contig0065, 전체 게놈 샷건 서열	WP_035906563.1
Cas13c3	FnbCas13c	푸소박테리움 네크로포럼 BFTR-1 contig0068	WP_035935671.1
Cas13c4	FnfCas13c	푸소박테리움 네크로포럼 subsp. 푼둘리포르메 1_1_36S cont1.14	EHO19081.1
Cas13c5	FpeCas13c	푸소박테리움 퍼포에텐스 ATCC 29250 T364DRAFT_scaffold00009.9_C	WP_027128616.1
Cas13c6	FulCas13c	푸소박테리움 울세란스 ATCC 49185 cont2.38	WP_040490876.1
Cas13c7	AspCas13c	아나에로살리박터 sp. ND1 게놈 어셈블리 아나에로살리박터 마실리엔시스 ND1	WP_042678931.1

[1413]

[1414] 표 7. 박테리아 스크린에서의 PFS 컷오프

Cas13b 오솔로그	키	**PFS** 모티프 생성에 사용된 **-Log$_2$** 감손 점수
베르게이엘라 주헬컴	1	2
프레보텔라 인터메디아 유전자좌 1	2	1
프레보텔라 부카에	3	3
알리스티페스 sp. ZOR0009	4	1
프레보텔라 sp. MA2016	5	2
리에메렐라 아나티페스티퍼	6	4
프레보텔라 아우란티아카	7	1
프레보텔라 사카롤리티카	8	0
프레보텔라 인터메디아 유전자좌 2	9	0
카프노사이토파가 카니몰서스	10	3
포르피로모나스 굴라에	11	4
프레보텔라 sp. P5-125	12	2.1
플라보박테리움 브란키오필럼	13	1
포르피로모나스 진지발리스	14	3
프레보텔라 인터메디아 유전자좌 2	15	4

[1415]

[1416] 표 8. 포유동물 세포에서 RNA 편집을 위해 본 연구에서 사용된 dCas13b-ADAR 링커 서열

도면	링커
50C	GSGGGGS (SEQ ID No. 296)
50E	GS
57B	GSGGGGS (SEQ ID No. 296)
57C	GS
57D	GS
57E: GS	GS
57E: GSGGGGS	GSGGGGS
57E: (GGGS)3	GGGGSGGGGSGGGGS (SEQ ID No. 802)
57E: 경성 (Rigid)	EAAAK (SEQ ID No. 803)
57E: (GGS)6	GGSGGSGGSGGSGGSGGS (SEQ ID No. 804)
57E: XTEN	SGSETPGTSESATPES (SEQ ID No. 805)
51B	GS
57F	GS
51C	GS
52B	GS

[1417]

52D	GS
52E	GS
51A: Δ984-1090, Δ1026-1090, Δ1053-1090	GS
51A: Δ1-125, Δ1-88, Δ1-72	GSGGGGS
53B	GS
53C	GS
53D	GS
60A	GS
60C	GS
60D	GS
61D	GS
54A	GS
62A	GS
54B	GS
62B	GS
62C	GS
63A	GS
63B	GS
54C	GS
54D	GS
54E	GS
54F	GS
66A	GS
66A	GS

[1418]

[1419] **표 9. 질환-관련 돌연변이에 대한 질환 정보**

전체 길이 후보	**유전자**	**질환**
NM_000054.4(AVPR2):c.878G>A (p.Trp293Ter)	AVPR2	X-연관 신장성 요붕증
NM_000136.2(FANCC):c.1517G>A (p.Trp506Ter)	FANCC	판코니 빈혈, 상보군 C
추가 모의 후보		
후보	**유전자**	**질환**
NM_000206.2(IL2RG):c.710G>A (p.Trp237Ter)	IL2RG	X-연관 중증 복합 면역결핍증
NM_000132.3(F8):c.3144G>A (p.Trp1048Ter)	F8	유전성 인자 VIII 결핍 질환
NM_000527.4(LDLR):c.1449G>A (p.Trp483Ter)	LDLR	가족성 고콜레스테롤혈증

[1420]

NM_000071.2(CBS):c.162G>A (p.Trp54Ter)	CBS	CBS 결핍에 의한 호모시스틴뇨증
NM_000518.4(HBB):c.114G>A (p.Trp38Ter)	HBB	베타^0^ 탈라세미아|베타 탈라세미아
NM_000035.3(ALDOB):c.888G>A (p.Trp296Ter)	ALDOB	유전성 프룩토스뇨증
NM_004006.2(DMD):c.3747G>A (p.Trp1249Ter)	DMD	뒤센 근이영양증
NM_005359.5(SMAD4):c.906G>A (p.Trp302Ter)	SMAD4	소아 폴립증 증후군
NM_000059.3(BRCA2):c.582G>A (p.Trp194Ter)	BRCA2	가족성 유방암|유방-난소암, 가족성 2
NM_000833.4(GRIN2A):c.3813G>A (p.Trp1271Ter)	GRIN2A	언어 장애 수반 및 정신 지체 수반 또는 미수반의 초점 간질
NM_002977.3(SCN9A):c.2691G>A (p.Trp897Ter)	SCN9A	선천성 통각 무반응증, 상염색체 열성
NM_007375.3(TARDBP):c.943G>A (p.Ala315Thr)	TARDBP	근위축성 측삭 경화증 10형
NM_000492.3(CFTR):c.3846G>A (p.Trp1282Ter)	CFTR	낭성 섬유증|유전성 췌장염|미제공됨|아탈루렌 반응 - 효능
NM_130838.1(UBE3A):c.2304G>A (p.Trp768Ter)	UBE3A	엔젤만 증후군
NM_000543.4(SMPD1):c.168G>A (p.Trp56Ter)	SMPD1	니만-피크병, A형
NM_206933.2(USH2A):c.9390G>A (p.Trp3130Ter)	USH2A	어셔 증후군, 2A형
NM_130799.2(MEN1):c.1269G>A (p.Trp423Ter)	MEN1	유전성 암-소인 증후군
NM_177965.3(C8orf37):c.555G>A (p.Trp185Ter)	C8orf37	망막 색소변성증 64
NM_000249.3(MLH1):c.1998G>A (p.Trp666Ter)	MLH1	린치 증후군
NM_000548.4(TSC2):c.2108G>A (p.Trp703Ter)	TSC2	결절성 경화증 2|결절성 경화증 증후군
NM_000267.3(NF1):c.7044G>A (p.Trp2348Ter)	NF1	신경섬유종증, 1형
NM_000179.2(MSH6):c.3020G>A (p.Trp1007Ter)	MSH6	린치 증후군
NM_000344.3(SMN1):c.305G>A (p.Trp102Ter)	SMN1	척수 근위축증, II형|쿠겔베르그-웰란더병
NM_024577.3(SH3TC2):c.920G>A (p.Trp307Ter)	SH3TC2	샤르코-마리-투스 질환, 4C형

[1421]

NM_001369.2(DNAH5):c.8465G>A (p.Trp2822Ter)	DNAH5	원발성 섬모운동 이상증
NM_004992.3(MECP2):c.311G>A (p.Trp104Ter)	MECP2	레트 증후군
NM_032119.3(ADGRV1):c.7406G>A (p.Trp2469Ter)	ADGRV1	어셔 증후군, 2C형
NM_017651.4(AHI1):c.2174G>A (p.Trp725Ter)	AHI1	주버트 증후군 3
NM_004562.2(PRKN):c.1358G>A (p.Trp453Ter)	PRKN	파킨슨병 2
NM_000090.3(COL3A1):c.3833G>A (p.Trp1278Ter)	COL3A1	엘러스-단로스 증후군, 4형
NM_007294.3(BRCA1):c.5511G>A (p.Trp1837Ter)	BRCA1	가족성 유방암|유방-난소암, 가족성 1
NM_000256.3(MYBPC3):c.3293G>A (p.Trp1098Ter)	MYBPC3	원발성 가족성 비대성 심근병증
NM_000038.5(APC):c.1262G>A (p.Trp421Ter)	APC	가족성 샘종 폴립증 1
NM_001204.6(BMPR2):c.893G>A (p.W298*)	BMPR2	원발성 폐 고혈압

[1422]

[1423] **표 10. 본 연구에서 사용된 핵심 플라스미드**

플라스미드 명칭	설명
pAB0006	CMV-C루시퍼라제-폴리 A EF1a-G-루시퍼라제-폴리 A
pAB0040	CMV-C루시퍼라제 (STOP85)-폴리A EF1a-G-루시퍼라제-폴리A
pAB0048	pCDNA-ADAR2-N-말단 B12-HIV NES
pAB0050	pAB0001-A02 (crRNA 골격)
pAB0051	pAB0001-B06 (crRNA 골격)
pAB0052	pAB0001-B11 (crRNA 골격)
pAB0053	pAB0001-B12 (crRNA 골격)
pAB0014.B6	EF1a-BsiWI-Cas13b6-NES-mapk
pAB0013.B11	EF1a-BsiWI-Cas13b11-NES-HIV
pAB0013.B12	EF1a-BsiWI-Cas13b12-NES-HIV
pAB0051	pAB0001-B06 (crRNA 골격)
pAB0052	pAB0001-B11 (crRNA 골격)
pAB0053	pAB0001-B12 (crRNA 골격)
pAB0079	pCDNA-ADAR1hu-EQ 돌연변이체-N-말단 목적 벡터
pAB0085	pCDNA-ADAR2 (E488Q)hu-EQ 돌연변이체-N-말단 목적 벡터
pAB0095	EF1a-BsiWI-Cas13-B12-NES-HIV, 이중 H HEPN 돌연변이체 존재
pAB0114	pCDNA-wtADAR2hu-EQ 돌연변이체-N-말단 목적 벡터

[1424]

pAB0120	루시퍼라제 ADAR 가이드 최적 (wC0054 의 가이드 24)
pAB0122	ADAR 실험용 pAB0001-B12 NT 가이드
pAB0151	pCDNA-ADAR2hu-EQ 돌연변이체-N-말단 목적 벡터 C-말단 델타 984-1090
pAB0180	T375G 특이성 돌연변이체
pAB0181	T375G Cas13b C-말단 델타 984-1090

[1425]

[1426] 표 11. 포유동물 세포에서 녹다운을 위해 본 실험에서 사용된 가이드/shRNA 서열

명칭	스페이서 서열	간섭 기전	주	제1 도면
박테리아 PFS 가이드	GCCAGCUUUCCGGGCAUU GGCUUCCAUC (SEQ ID No. 300)	Cas13b	모든 오솔로그에 사용	
Cas13a-Gluc 가이드1	GCCAGCTTTCCGGGCATTG GCTTCCATC (SEQ ID No. 301)	Cas13a	모든 Cas13a 오솔로그에 사용	도 49B
Cas13a-Gluc 가이드2	ACCCAGGAATCTCAGGAA TGTCGACGAT (SEQ ID No. 302)	Cas13a	모든 Cas13a 오솔로그에 사용	도 49B
Cas13a-비표적화 가이드 (LacZ)	AGGGTTTTCCCAGTCACGA CGTTGTAAA (SEQ ID No. 303)	Cas13a	모든 Cas13a 오솔로그에 사용	도 49B
Cas13b-Gluc 가이드1.1	GGGCATTGGCTTCCATCTC TTTGAGCACCT (SEQ ID No. 304)	Cas13b	오솔로그 1-3, 6, 7, 10, 11, 12, 14, 15 에 사용	도 49B
Cas13b-Gluc 가이드1.2	GUGCAGCCAGCUUUCCGG GCAUUGGCUUCC (SEQ ID No. 305)	Cas13b	오솔로그 4에 사용	도 49B
Cas13b-Gluc 가이드1.3	GCAGCCAGCUUUCCGGGC AUUGGCUUCCAU (SEQ ID No. 306)	Cas13b	오솔로그 5에 사용	도 49B
Cas13b-Gluc 가이드1.4	GGCUUCCAUCUCUUUGAG CACCUCCAGCGG (SEQ ID No. 307)	Cas13b	오솔로그 8, 9에 사용	도 49B
Cas13b-Gluc 가이드1.5	GGAAUGUCGACGAUCGCC UCGCCUAUGCCG (SEQ ID No. 308)	Cas13b	오솔로그 13에 사용	도 49B
Cas13b-	GAAUGUCGACGAUCGCCU	Cas13b	오솔로그	도 49B

[1427]

Gluc 가이드2.1	CGCCUAUGCCGC (SEQ ID No. 309)		1-3, 6, 7, 10, 11, 14, 15 에 사용	
Cas13b-Gluc 가이드2.2	GACCUGUGCGAUGAACUG CUCCAUGGGCUC (SEQ ID No. 310)	Cas13b	오솔로그 12에 사용	도 49B
Cas13b-Gluc 가이드2.2	GUGUGGCAGCGUCCUGGG AUGAACUUCUUC (SEQ ID No. 311)	Cas13b	오솔로그 4에 사용	도 49B
Cas13b-Gluc 가이드2.3	GUGGCAGCGUCCUGGGAU GAACUUCUUCAU (SEQ ID No. 312)	Cas13b	오솔로그 5에 사용	도 49B
Cas13b-Gluc 가이드2.4	GCUUCUUGCCGGGCAACU UCCCGCGGUCAG (SEQ ID No. 313)	Cas13b	오솔로그 8, 9에 사용	도 49B
Cas13b-Gluc 가이드2.6	GCAGGGUUUUCCCAGUCA CGACGUUGUAAAA (SEQ ID No. 314)	Cas13b	오솔로그 13에 사용	도 49B
Cas13b-비표적화 가이드	GCAGGGUUUUCCCAGUCA CGACGUUGUAAAA (SEQ ID No. 315)	Cas13b	모든 오솔로그에 사용	도 49B
Cas13a-Gluc 가이드-RNASeq	ACCCAGGAAUCUCAGGAA UGUCGACGAU (SEQ ID No. 316)	Cas13a		도 49E
shRNA-Gluc 가이드	CAGCTTTCCGGGCATTGGC TT (SEQ ID No. 317)	shRNA		도 49F
Cas13b-Gluc 가이드-RNASeq	CCGCUGGAGGUGCUCAAA GAGAUGGAAGCC (SEQ ID No. 318)	Cas13b		도 49F
Cas13a-Gluc-가이드-1	GCCAGCTTTCCGGGCATTG GCTTCCATC (SEQ ID No. 319)	Cas13a		도 56A
Cas13a-Gluc-가이드-2	ACCCAGGAATCTCAGGAA TGTCGACGAT (SEQ ID No. 320)	Cas13a		도 56A
Cas13b-Gluc-opt-가이드-1	GGGCATTGGCTTCCATCTC TTTGAGCACCT (SEQ ID No. 321)	Cas13b		도 56A
Cas13b-Gluc-opt-가이드-2	GAAUGUCGACGAUCGCCU CGCCUAUGCCGC (SEQ ID No. 322)	Cas13b		도 56A

[1428]

Cas13a KRAS 가이드 1	CAAGGCACTCTTGCCTACG CCACCAGCT (SEQ ID No. 323)	Cas13a		도 56B
Cas13a KRAS 가이드 2	TCATATTCGTCCACAAAAT GATTCTGAA (SEQ ID No. 324)	Cas13a		도 56B
Cas13a KRAS 가이드 3	ATTATTTATGGCAAATACA CAAAGAAAG (SEQ ID No. 325)	Cas13a		도 56B
Cas13a KRAS 가이드 4	GAATATCTTCAAATGATTT AGTATTATT (SEQ ID No. 326)	Cas13a		도 56B
Cas13a KRAS 가이드 5	ACCATAGGTACATCTTCAG AGTCCTTAA (SEQ ID No. 327)	Cas13a		도 56B
Cas13b KRAS 가이드 1	GTCAAGGCACTCTTGCCTA CGCCACCAGCT (SEQ ID No. 328)	Cas13b		도 56B
Cas13b KRAS 가이드 2	GATCATATTCGTCCACAAA ATGATTCTGAA (SEQ ID No. 329)	Cas13b		도 56B
Cas13b KRAS 가이드 3	GTATTATTTATGGCAAATA CACAAAGAAAG (SEQ ID No. 330)	Cas13b		도 56B
Cas13b KRAS 가이드 4	GTGAATATCTTCAAATGAT TTAGTATTATT (SEQ ID No. 331)	Cas13b		도 56B
Cas13b KRAS 가이드 5	GGACCATAGGTACATCTTC AGAGTCCTTAA (SEQ ID No. 332)	Cas13b		도 56B
shRNA KRAS 가이드 1	aagagtgccttgacgatacagcCTCGA Ggctgtatcgtcaaggcactctt (SEQ ID No. 333)	shRNA		도 56B
shRNA KRAS 가이드 2	aatcattttgtggacgaatatCTCGAGa tattcgtccacaaaatgatt (SEQ ID No. 334)	shRNA		도 56B
shRNA KRAS 가이드 3	aaataatactaaatcatttgaCTCGAGt caaatgatttagtattattt (SEQ ID No. 335)	shRNA		도 56B
shRNA KRAS 가이드 4	aataatactaaatcatttgaaCTCGAGtt caaatgatttagtattatt (SEQ ID No. 336)	shRNA		도 56B
shRNA KRAS 가이드 5	aaggactctgaagatgtacctCTCGAG aggtacatcttcagagtcctt (SEQ ID No. 337)	shRNA		도 56B

[1429]

[1430] 표 12. Gluc 녹다운에 사용된 가이드 서열

명칭	스페이서 서열	위치	주	제1 도면
Gluc 타일링 가이드1	GAGATCAGGGCAAACAGAACTTTGACTCCC (SEQ ID No. 338)	2	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드2	GGATGCAGATCAGGGCAAACAGAACTTTGA (SEQ ID No. 339)	7	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드3	GCACAGCGATGCAGATCAGGGCAAACAGA A (SEQ ID No. 340)	13	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드4	GCTCGGCCACAGCGATGCAGATCAGGGCAA (SEQ ID No. 341)	19	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드5	GGGGCTTGGCCTCGGCCACAGCGATGCAGA (SEQ ID No. 342)	28	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc	GTGGGCTTGGCCTCGGCCACAGCGATGCAG	29	Cas13a	49C

[1431]

타일링 가이드 6	(SEQ ID No. 343)		스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 7	GTCTCGGTGGGCTTGGCCTCGGCCACAGCG (SEQ ID No. 344)	35	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 8	GTTCGTTGTTCTCGGTGGGCTTGGCCTCGG (SEQ ID No. 345)	43	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 9	GGAAGTCTTCGTTGTTCTCGGTGGGCTTGG (SEQ ID No. 346)	49	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드10	GATGTTGAAGTCTTCGTTGTTCTCGGTGGG (SEQ ID No. 347)	54	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드11	GCGGCCACGATGTTGAAGTCTTCGTTGTTC (SEQ ID No. 348)	62	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C

[1432]

Gluc 타일링 가이드 12	GTGGCCACGGCCACGATGTTGAAGTCTTCG (SEQ ID No. 349)	68	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 13	GGTTGCTGGCCACGGCCACGATGTTGAAGT (SEQ ID No. 350)	73	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 14	GTCGCGAAGTTGCTGGCCACGGCCACGATG (SEQ ID No. 351)	80	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 15	GCCGTGGTCGCGAAGTTGCTGGCCACGGCC (SEQ ID No. 352)	86	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 16	GCGAGATCCGTGGTCGCGAAGTTGCTGGCC (SEQ ID No. 353)	92	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 17	GCAGCATCGAGATCCGTGGTCGCGAAGTTG (SEQ ID No. 354)	98	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오	49C

[1433]

			티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 18	GGGTCAGCATCGAGATCCGTGGTCGCGAAG (SEQ ID No. 355)	101	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 19	GCTTCCCGCGGTCAGCATCGAGATCCGTGG (SEQ ID No. 356)	109	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 20	GGGGCAACTTCCCGCGGTCAGCATCGAGAT (SEQ ID No. 357)	115	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 21	GTCTTGCCGGGCAACTTCCCGCGGTCAGCA (SEQ ID No. 358)	122	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 22	GGCAGCTTCTTGCCGGGCAACTTCCCGCGG (SEQ ID No. 359)	128	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 23	GCCAGCGGCAGCTTCTTGCCGGGCAACTTC (SEQ ID No. 360)	134	Cas13a 스페이서는 5' 말단에서	49C

[1434]

			2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 24	GCACCTCCAGCGGCAGCTTCTTGCCGGGCA (SEQ ID No. 361)	139	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 25	GCTTTGAGCACCTCCAGCGGCAGCTTCTTG (SEQ ID No. 362)	146	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 26	GCATCTCTTTGAGCACCTCCAGCGGCAGCT (SEQ ID No. 363)	151	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 27	GTCCATCTCTTTGAGCACCTCCAGCGGCAG (SEQ ID No. 364)	153	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 28	GGGCATTGGCTTCCATCTCTTTGAGCACCT (SEQ ID No. 365)	163	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc	GTCCGGGCATTGGCTTCCATCTCTTTGAGC	167	Cas13a	49C

[1435]

타일링 가이드 29	(SEQ ID No. 366)		스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 30	GGCCAGCTTTCCGGGCATTGGCTTCCATCT (SEQ ID No. 367)	175	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 31	GGGTGCAGCCAGCTTTCCGGGCATTGGCTT (SEQ ID No. 368)	181	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 32	GAGCCCCTGGTGCAGCCAGCTTTCCGGGCA (SEQ ID No. 369)	188	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 33	GATCAGACAGCCCCTGGTGCAGCCAGCTTT (SEQ ID No. 370)	195	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 34	GGCAGATCAGACAGCCCCTGGTGCAGCCAG (SEQ ID No. 371)	199	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C

[1436]

[1437]

Gluc 타일링 가이드 35	GACAGGCAGATCAGACAGCCCCTGGTGCAG (SEQ ID No. 372)	203	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 36	GTGATGTGGGACAGGCAGATCAGACAGCCC (SEQ ID No. 373)	212	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 37	GACTTGATGTGGGACAGGCAGATCAGACAG (SEQ ID No. 374)	215	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 38	GGGGCGTGCACTTGATGTGGGACAGGCAGA (SEQ ID No. 375)	223	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 39	GCTTCATCTTGGGCGTGCACTTGATGTGGG (SEQ ID No. 376)	232	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 40	GTGAACTTCTTCATCTTGGGCGTGCACTTG (SEQ ID No. 377)	239	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단	49C

			된다는 것을 유의한다	
Gluc 타일링 가이드 41	GGGATGAACTTCTTCATCTTGGGCGTGCAC (SEQ ID No. 378)	242	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 42	GTGGGATGAACTTCTTCATCTTGGGCGTGC (SEQ ID No. 379)	244	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 43	GGGCAGCGTCCTGGGATGAACTTCTTCATC (SEQ ID No. 380)	254	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 44	GGGTGTGGCAGCGTCCTGGGATGAACTTCT (SEQ ID No. 381)	259	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 45	GTTCGTAGGTGTGGCAGCGTCCTGGGATGA (SEQ ID No. 382)	265	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 46	GCGCCTTCGTAGGTGTGGCAGCGTCCTGGG (SEQ ID No. 383)	269	Cas13a 스페이서는 5' 말단에서	49C

[1438]

			2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 47	GTCTTTGTCGCCTTCGTAGGTGTGGCAGCG (SEQ ID No. 384)	276	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 48	GCTTTGTCGCCTTCGTAGGTGTGGCAGCGT (SEQ ID No. 385)	275	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 49	GTGCCGCCCTGTGCGGACTCTTTGTCGCCT (SEQ ID No. 386)	293	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 50	GTATGCCGCCCTGTGCGGACTCTTTGTCGC (SEQ ID No. 387)	295	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 51	GCCTCGCCTATGCCGCCCTGTGCGGACTCT (SEQ ID No. 388)	302	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc	GGATCGCCTCGCCTATGCCGCCCTGTGCGG	307	Cas13a	49C

[1439]

[1440]

타일링 가이드 52	(SEQ ID No. 389)		스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 53	GATGTCGACGATCGCCTCGCCTATGCCGCC (SEQ ID No. 390)	315	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 54	GCAGGAATGTCGACGATCGCCTCGCCTATG (SEQ ID No. 391)	320	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 55	GAATCTCAGGAATGTCGACGATCGCCTCGC (SEQ ID No. 392)	325	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 56	GCCCAGGAATCTCAGGAATGTCGACGATCG (SEQ ID No. 393)	331	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 57	GCCTTGAACCCAGGAATCTCAGGAATGTCG (SEQ ID No. 394)	338	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C

Gluc 타일링 가이드 58	GCCAAGTCCTTGAACCCAGGAATCTCAGGA (SEQ ID No. 395)	344	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 59	GTGGGCTCCAAGTCCTTGAACCCAGGAATC (SEQ ID No. 396)	350	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 60	GCCATGGGCTCCAAGTCCTTGAACCCAGGA (SEQ ID No. 397)	353	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 61	GGAACTGCTCCATGGGCTCCAAGTCCTTGA (SEQ ID No. 398)	361	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 62	GTGCGATGAACTGCTCCATGGGCTCCAAGT (SEQ ID No. 399)	367	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 63	GGACCTGTGCGATGAACTGCTCCATGGGCT (SEQ ID No. 400)	373	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단	49C

[1441]

			된다는 것을 유의한다	
Gluc 타일링 가이드 64	GACAGATCGACCTGTGCGATGAACTGCTCC (SEQ ID No. 401)	380	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 65	GACACACAGATCGACCTGTGCGATGAACTG (SEQ ID No. 402)	384	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 66	GTGCAGTCCACACACAGATCGACCTGTGCG (SEQ ID No. 403)	392	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 67	GCCAGTTGTGCAGTCCACACACAGATCGAC (SEQ ID No. 404)	399	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 68	GGGCAGCCAGTTGTGCAGTCCACACACAGA (SEQ ID No. 405)	404	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 69	GTTTGAGGCAGCCAGTTGTGCAGTCCACAC (SEQ ID No. 406)	409	Cas13a 스페이서는 5' 말단에서	49C

[1442]

			2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 70	GAAGCCCTTTGAGGCAGCCAGTTGTGCAGT (SEQ ID No. 407)	415	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 71	GCACGTTGGCAAGCCCTTTGAGGCAGCCAG (SEQ ID No. 408)	424	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 72	GACTGCACGTTGGCAAGCCCTTTGAGGCAG (SEQ ID No. 409)	428	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 73	GGGTCAGAACACTGCACGTTGGCAAGCCCT (SEQ ID No. 410)	437	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 74	GCAGGTCAGAACACTGCACGTTGGCAAGCC (SEQ ID No. 411)	439	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc	GAGCAGGTCAGAACACTGCACGTTGGCAAG	441	Cas13a	49C

[1443]

타일링 가이드 75	(SEQ ID No. 412)		Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 76	GGCCACTTCTTGAGCAGGTCAGAACACTGC (SEQ ID No. 413)	452	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 77	GCGGCAGCCACTTCTTGAGCAGGTCAGAAC (SEQ ID No. 414)	457	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 78	GTGCGGCAGCCACTTCTTGAGCAGGTCAGA (SEQ ID No. 415)	459	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 79	GAGCGTTGCGGCAGCCACTTCTTGAGCAGG (SEQ ID No. 416)	464	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 80	GAAAGGTCGCACAGCGTTGCGGCAGCCACT (SEQ ID No. 417)	475	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C

[1444]

Gluc 타일링 가이드 81	GCTGGCAAAGGTCGCACAGCGTTGCGGCAG (SEQ ID No. 418)	480	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 82	GGGCAAAGGTCGCACAGCGTTGCGGCAGCC (SEQ ID No. 419)	478	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 83	GTGGATCTTGCTGGCAAAGGTCGCACAGCG (SEQ ID No. 420)	489	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 84	GCACCTGGCCCTGGATCTTGCTGGCAAAGG (SEQ ID No. 421)	499	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 85	GTGGCCCTGGATCTTGCTGGCAAAGGTCGC (SEQ ID No. 422)	495	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드 86	GTGATCTTGTCCACCTGGCCCTGGATCTTG (SEQ ID No. 423)	509	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단	49C

[1445]

			된다는 것을 유의한다	
Gluc 타일링 가이드87	GCCCCTTGATCTTGTCCACCTGGCCCTGGA (SEQ ID No. 424)	514	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드88	GCCCTTGATCTTGTCCACCTGGCCCTGGAT (SEQ ID No. 425)	513	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드89	GCCTTGATCTTGTCCACCTGGCCCTGGATC (SEQ ID No. 426)	512	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드90	GGCAAAGGTCGCACAGCGTTGCGGCAGCCA (SEQ ID No. 427)	477	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드91	GCAAAGGTCGCACAGCGTTGCGGCAGCCAC (SEQ ID No. 428)	476	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
Gluc 타일링 가이드92	GAAGGTCGCACAGCGTTGCGGCAGCCACTT (SEQ ID No. 429)	474	Cas13a 스페이서는 5' 말단에서	49C

[1446]

			2개 뉴클레오티드가 절단된다는 것을 유의한다	
Gluc 타일링 가이드 93	GAGGTCGCACAGCGTTGCGGCAGCCACTTC (SEQ ID No. 430)	473	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
비표적화 가이드 1	GGTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID No. 431)	N/A	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
비표적화 가이드 2	GGGAACCTTGGCCGTTATAAAGTCTGACCA G (SEQ ID No. 432)	N/A	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C
비표적화 가이드 3	GGAGGGTGAGAATTTAGAACCAAGATTGTT G (SEQ ID No. 433)	N/A	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49C

[1447]

[1448] 표 13. Cluc 녹다운에 사용된 가이드 서열

명칭	스페이서 서열	위치	주	제1 도면
Cluc 타일링 가이드1	GAGTCCTGGCAATGAACAGTGGCGCAGTAG (SEQ ID No. 434)	32	Cas13a 스페이서는 5' 말단에서	49D

[1449]

			2개 뉴클레오티드가 절단된다는 것을 유의한다	
Cluc 타일링 가이드 2	GGGTGCCACAGCTGCTATCAATACATTCTC (SEQ ID No. 435)	118	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 3	GTTACATACTGACACATTCGGCAACATGTT (SEQ ID No. 436)	197	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 4	GTATGTACCAGGTTCCTGGAACTGGAATCT (SEQ ID No. 437)	276	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 5	GCCTTGGTTCCATCCAGGTTCTCCAGGGTG (SEQ ID No. 438)	350	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 6	GCAGTGATGGGATTCTCAGTAGCTTGAGCG (SEQ ID No. 439)	431	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc	GAGCCTGGCATCTCAACAACAGCGATGGTG	512	Cas13a	49D

[1450]

타일링 가이드7	(SEQ ID No. 440)		스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	
Cluc 타일링 가이드8	GTGTCTGGGGCGATTCTTACAGATCTTCCT (SEQ ID No. 441)	593	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드9	GCTGGATCTGAAGTGAAGTCTGTATCTTCC (SEQ ID No. 442)	671	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드10	GGCAACGTCATCAGGATTTCCATAGAGTGG (SEQ ID No. 443)	747	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드11	GAGGCGCAGGAGATGGTGTAGTAGTAGAAG (SEQ ID No. 444)	830	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드13	GAGGGACCCTGGAATTGGTATCTTGCTTTG (SEQ ID No. 445)	986	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D

[1451]

Cluc 타일링 가이드 14	GGTAAGAGTCAACATTCCTGTGTGAAACCT (SEQ ID No. 446)	1066	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 15	GACCAGAATCTGTTTTCCATCAACAATGAG (SEQ ID No. 447)	1143	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 16	GATGGCTGTAGTCAGTATGTCACCATCTTG (SEQ ID No. 448)	1227	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 17	GTACCATCGAATGGATCTCTAATATGTACG (SEQ ID No. 449)	1304	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 18	GAGATCACAGGCTCCTTCAGCATCAAAAGA (SEQ ID No. 450)	1380	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 19	GCTTTGACCGGCGAAGAGACTATTGCAGAG (SEQ ID No. 451)	1461	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단	49D

[1452]

			된다는 것을 유의한다	
Cluc 타일링 가이드 20	GCCCCTCAGGCAATACTCGTACATGCATCG (SEQ ID No. 452)	1539	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
Cluc 타일링 가이드 21	GCTGGTACTTCTAGGGTGTCTCCATGCTTT (SEQ ID No. 453)	1619	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
비표적화 가이드 1	GGTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID No. 454)	N/A	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
비표적화 가이드 2	GGGAACCTTGGCCGTTATAAAGTCTGACCA G (SEQ ID No. 455)	N/A	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D
비표적화 가이드 3	GGAGGGTGAGAATTTAGAACCAAGATTGTT G (SEQ ID No. 456)	N/A	Cas13a 스페이서는 5' 말단에서 2개 뉴클레오티드가 절단된다는 것을 유의한다	49D

[1453]

[1454] **표 14. 포유동물 세포에서 RNA 편집을 위해 본 연구에서 사용된 가이드 서열. 미스매치된 염기 플립은 대문자로 표시됨**

명칭	스페이서 서열	주	제1 도면
타일링 30 nt 30 미스매치 거리	gCatcctgcggcctctactctgcattcaatt (SEQ ID No. 457)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 28 미스매치 거리	gacCatcctgcggcctctactctgcattcaa (SEQ ID No. 458)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 26 미스매치 거리	gaaacCatcctgcggcctctactctgcattc (SEQ ID No. 459)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 24 미스매치 거리	gctaaacCatcctgcggcctctactctgcat (SEQ ID No. 460)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 22 미스매치 거리	gttctaaacCatcctgcggcctctactctgc (SEQ ID No. 461)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 20 미스매치 거리	gtgttctaaacCatcctgcggcctctactct (SEQ ID No. 462)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 18 미스매치 거리	gaatgttctaaacCatcctgcggcctctact (SEQ ID No. 463)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 16 미스매치 거리	gagaatgttctaaacCatcctgcggcctcta (SEQ ID No. 464)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 14 미스매치 거리	gatagaatgttctaaacCatcctgcggcctc (SEQ ID No. 465)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 12 미스매치 거리	gccatagaatgttctaaacCatcctgcggcc (SEQ ID No. 466)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 10 미스매치 거리	gttccatagaatgttctaaacCatcctgcgg (SEQ ID No. 467)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 8 미스매치 거리	gctttccatagaatgttctaaacCatcctgc (SEQ ID No. 468)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 6 미스매치 거리	gctctttccatagaatgttctaaacCatcct (SEQ ID No. 469)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 4 미스매치 거리	gatctctttccatagaatgttctaaacCatc (SEQ ID No. 470)	U6 발현을 위해 5' G를 갖는다	50C
타일링 30 nt 2 미스매치 거리	ggaatctctttccatagaatgttctaaacCa (SEQ ID No. 471)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 50 미스매치 거리	gCatcctgcggcctctactctgcattcaattac atactgacacattcggca (SEQ ID No. 472)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 48 미스매치 거리	gacCatcctgcggcctctactctgcattcaatt acatactgacacattcgg (SEQ ID No. 473)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 46 미스매치 거리	gaaacCatcctgcggcctctactctgcattcaa ttacatactgacacattc (SEQ ID No. 474)	U6 발현을 위해 5' G를 갖는다	50C

[1455]

타일링 50 nt 44 미스매치 거리	gctaaacCatcctgcggcctctactctgcattc aattacatactgacacat (SEQ ID No. 475)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 42 미스매치 거리	gttctaaacCatcctgcggcctctactctgcatt caattacatactgacac (SEQ ID No. 476)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 40 미스매치 거리	gtgttctaaacCatcctgcggcctctactctgc attcaattacatactgac (SEQ ID No. 477)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 38 미스매치 거리	gaatgttctaaacCatcctgcggcctctactct gcattcaattacatactg (SEQ ID No. 478)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 36 미스매치 거리	gagaatgttctaaacCatcctgcggcctctact ctgcattcaattacatac (SEQ ID No. 479)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 34 미스매치 거리	gatagaatgttctaaacCatcctgcggcctcta ctctgcattcaattacat (SEQ ID No. 480)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 32 미스매치 거리	gccatagaatgttctaaacCatcctgcggcctc tactctgcattcaattac (SEQ ID No. 481)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 30 미스매치 거리	gttccatagaatgttctaaacCatcctgcggcc tctactctgcattcaatt (SEQ ID No. 482)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 28 미스매치 거리	gctttccatagaatgttctaaacCatcctgcgg cctctactctgcattcaa (SEQ ID No. 483)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 26 미스매치 거리	gctctttccatagaatgttctaaacCatcctgcg gcctctactctgcattc (SEQ ID No. 484)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 24 미스매치 거리	gatctctttccatagaatgttctaaacCatcctg cggcctctactctgcat (SEQ ID No. 485)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 22 미스매치 거리	ggaatctctttccatagaatgttctaaacCatcc tgcggcctctactctgc (SEQ ID No. 486)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 20 미스매치 거리	gtggaatctctttccatagaatgttctaaacCat cctgcggcctctactct (SEQ ID No. 487)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 18 미스매치 거리	gactggaatctctttccatagaatgttctaaacC atcctgcggcctctact (SEQ ID No. 488)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 16 미스매치 거리	ggaactggaatctctttccatagaatgttctaaa cCatcctgcggcctcta (SEQ ID No. 489)	U6 발현을 위해 5' G를 갖는다	50C

[1456]

타일링 50 nt 14 미스매치 거리	gtggaactggaatctctttccatagaatgttctaaacCatcctgcggcct c(SEQ ID No. 490)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 12 미스매치 거리	gcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcc (SEQ ID No. 491)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 10 미스매치 거리	gttcctggaactggaatctctttccatagaatgttctaaacCatcctgcgg (SEQ ID No. 492)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 8 미스매치 거리	gggttcctggaactggaatctctttccatagaatgttctaaacCatcctgc (SEQ ID No. 493)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 6 미스매치 거리	gcaggttcctggaactggaatctctttccatagaatgttctaaacCatcct (SEQ ID No. 494)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 4 미스매치 거리	gaccaggttcctggaactggaatctctttccatagaatgttctaaacCatc (SEQ ID No. 495)	U6 발현을 위해 5' G를 갖는다	50C
타일링 50 nt 2 미스매치 거리	ggtaccaggttcctggaactggaatctctttccatagaatgttctaaacCa (SEQ ID No. 496)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 70 미스매치 거리	gCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggttat (SEQ ID No. 497)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 68 미스매치 거리	gacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggttt (SEQ ID No. 498)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 66 미스매치 거리	gaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggt (SEQ ID No. 499)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 64 미스매치 거리	gctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctg (SEQ ID No. 500)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 62 미스매치 거리	gttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcc (SEQ ID No. 501)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 60 미스매치 거리	gtgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttt (SEQ ID No. 502)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 58 미스매치 거리	gaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttt (SEQ ID No. 503)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 56 미스매치 거리	gagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgt (SEQ ID No. 504)	U6 발현을 위해 5' G를 갖는다	50C

[1457]

타일링 70 nt 54 미스매치 거리	gatagaatgttctaaacCatcctgcggcctcta ctctgcattcaattacatactgacacattcggca acat (SEQ ID No. 505)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 52 미스매치 거리	gccatagaatgttctaaacCatcctgcggcctc tactctgcattcaattacatactgacacattcgg caac (SEQ ID No. 506)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 50 미스매치 거리	gttccatagaatgttctaaacCatcctgcggcc tctactctgcattcaattacatactgacacattcg gca (SEQ ID No. 507)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 48 미스매치 거리	gctttccatagaatgttctaaacCatcctgcgg cctctactctgcattcaattacatactgacacatt cgg (SEQ ID No. 508)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 46 미스매치 거리	gctctttccatagaatgttctaaacCatcctgcg gcctctactctgcattcaattacatactgacacat tc (SEQ ID No. 509)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 44 미스매치 거리	gatctctttccatagaatgttctaaacCatcctg cggcctctactctgcattcaattacatactgaca cat (SEQ ID No. 510)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 42 미스매치 거리	ggaatctctttccatagaatgttctaaacCatcc tgcggcctctactctgcattcaattacatactga cac (SEQ ID No. 511)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 40 미스매치 거리	gtggaatctctttccatagaatgttctaaacCat cctgcggcctctactctgcattcaattacatact gac (SEQ ID No. 512)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 38 미스매치 거리	gactggaatctctttccatagaatgttctaaacC atcctgcggcctctactctgcattcaattacata ctg (SEQ ID No. 513)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 36 미스매치 거리	ggaactggaatctctttccatagaatgttctaaa cCatcctgcggcctctactctgcattcaattaca tac (SEQ ID No. 514)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 34 미스매치 거리	gtggaactggaatctctttccatagaatgttcta aacCatcctgcggcctctactctgcattcaatta cat (SEQ ID No. 515)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 32 미스매치 거리	gcctggaactggaatctctttccatagaatgttc taaacCatcctgcggcctctactctgcattcaat tac (SEQ ID No. 516)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 30 미스매치 거리	gttcctggaactggaatctctttccatagaatgtt ctaaacCatcctgcggcctctactctgcattca att (SEQ ID No. 517)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 28 미스매치 거리	gggttcctggaactggaatctctttccatagaat gttctaaacCatcctgcggcctctactctgcatt caa (SEQ ID No. 518)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 26 미스매치 거리	gcaggttcctggaactggaatctctttccataga atgttctaaacCatcctgcggcctctactctgc attc (SEQ ID No. 519)	U6 발현을 위해 5' G를 갖는다	50C

[1458]

타일링 70 nt 24 미스매치 거리	gaccaggttcctggaactggaatctctttccata gaatgttctaaacCatcctgcggcctctactct gcat (SEQ ID No. 520)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 22 미스매치 거리	ggtaccaggttcctggaactggaatctctttcca tagaatgttctaaacCatcctgcggcctctact ctgc (SEQ ID No. 521)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 20 미스매치 거리	gatgtaccaggttcctggaactggaatctctttc catagaatgttctaaacCatcctgcggcctcta ctct (SEQ ID No. 522)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 18 미스매치 거리	ggtatgtaccaggttcctggaactggaatctctt tccatagaatgttctaaacCatcctgcggcctc tact (SEQ ID No. 523)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 16 미스매치 거리	gacgtatgtaccaggttcctggaactggaatct ctttccatagaatgttctaaacCatcctgcggc ctcta (SEQ ID No. 524)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 14 미스매치 거리	gacacgtatgtaccaggttcctggaactggaat ctctttccatagaatgttctaaacCatcctgcgg cctc (SEQ ID No. 525)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 12 미스매치 거리	gcaacacgtatgtaccaggttcctggaactgg aatctctttccatagaatgttctaaacCatcctg cggcc (SEQ ID No. 526)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 10 미스매치 거리	gcccaacacgtatgtaccaggttcctggaact ggaatctctttccatagaatgttctaaacCatcc tgcgg (SEQ ID No. 527)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 8 미스매치 거리	ggacccaacacgtatgtaccaggttcctggaa ctggaatctctttccatagaatgttctaaacCat cctgc (SEQ ID No. 528)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 6 미스매치 거리	gttgacccaacacgtatgtaccaggttcctgga actggaatctctttccatagaatgttctaaacCa tcct (SEQ ID No. 529)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 4 미스매치 거리	gccttgacccaacacgtatgtaccaggttcctg gaactggaatctctttccatagaatgttctaaac Catc (SEQ ID No. 530)	U6 발현을 위해 5' G를 갖는다	50C
타일링 70 nt 2 미스매치 거리	gttccttgacccaacacgtatgtaccaggttcct ggaactggaatctctttccatagaatgttctaaa cCa (SEQ ID No. 531)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 84 미스매치 거리	gCatcctgcggcctctactctgcattcaattac atactgacacattcggcaacatgtttttcctggtt tattttcacacagtcca (SEQ ID No. 532)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 82 미스매치 거리	gacCatcctgcggcctctactctgcattcaatt acatactgacacattcggcaacatgtttttcctg gtttattttcacacagtc (SEQ ID No. 533)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 80 미스매치 거리	gaaacCatcctgcggcctctactctgcattcaa ttacatactgacacattcggcaacatgtttttcct	U6 발현을 위해 5' G를 갖는다	50C

[1459]

	ggtttattttcacacag (SEQ ID No. 534)		
타일링 84 nt 78 미스매치 거리	gctaaacCatcctgcggcctctactctgcattc aattacatactgacacattcggcaacatgtttttc ctggtttattttcacac (SEQ ID No. 535)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 76 미스매치 거리	gttctaaacCatcctgcggcctctactctgcatt caattacatactgacacattcggcaacatgttttt cctggtttattttcac (SEQ ID No. 536)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 74 미스매치 거리	gtgttctaaacCatcctgcggcctctactctgc attcaattacatactgacacattcggcaacatgt ttttcctggtttattttc (SEQ ID No. 537)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 72 미스매치 거리	gaatgttctaaacCatcctgcggcctctactct gcattcaattacatactgacacattcggcaacat gttttcctggtttattt (SEQ ID No. 538)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 70 미스매치 거리	gagaatgttctaaacCatcctgcggcctctact ctgcattcaattacatactgacacattcggcaac atgttttcctggtttat (SEQ ID No. 539)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 68 미스매치 거리	gatagaatgttctaaacCatcctgcggcctcta ctctgcattcaattacatactgacacattcggca acatgttttcctggttt (SEQ ID No. 540)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 66 미스매치 거리	gccatagaatgttctaaacCatcctgcggcctc tactctgcattcaattacatactgacacattcgg caacatgttttcctggt (SEQ ID No. 541)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 64 미스매치 거리	gttccatagaatgttctaaacCatcctgcggcc tctactctgcattcaattacatactgacacattcg gcaacatgttttcctg (SEQ ID No. 542)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 62 미스매치 거리	gctttccatagaatgttctaaacCatcctgcgg cctctactctgcattcaattacatactgacacatt cggcaacatgttttcc (SEQ ID No. 543)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 60 미스매치 거리	gctctttccatagaatgttctaaacCatcctgcg gcctctactctgcattcaattacatactgacacat tcggcaacatgttttt (SEQ ID No. 544)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 58 미스매치 거리	gatctctttccatagaatgttctaaacCatcctg cggcctctactctgcattcaattacatactgaca cattcggcaacatgttt (SEQ ID No. 545)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 56 미스매치 거리	ggaatctctttccatagaatgttctaaacCatcc tgcggcctctactctgcattcaattacatactga cacattcggcaacatgt (SEQ ID No. 546)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 54 미스매치 거리	gtggaatctctttccatagaatgttctaaacCat cctgcggcctctactctgcattcaattacatact	U6 발현을 위해 5' G를 갖는다	50C

[1460]

	gacacattcggcaacat (SEQ ID No. 547)		
타일링 84 nt 52 미스매치 거리	gactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaac (SEQ ID No. 548)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 50 미스매치 거리	ggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggca (SEQ ID No. 549)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 48 미스매치 거리	gtggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcgg (SEQ ID No. 550)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 46 미스매치 거리	gcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattc (SEQ ID No. 551)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 44 미스매치 거리	gttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacat (SEQ ID No. 552)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 42 미스매치 거리	gggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacac (SEQ ID No. 553)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 40 미스매치 거리	gcaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgac (SEQ ID No. 554)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 38 미스매치 거리	gaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactg (SEQ ID No. 555)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 36 미스매치 거리	ggtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatac (SEQ ID No. 556)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 34 미스매치 거리	gatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacat (SEQ ID No. 557)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 32 미스매치 거리	ggtatgtaccaggttcctggaactggaatctcttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattac (SEQ ID No. 558)	U6 발현을 위해 5' G를 갖는다	50C

[1461]

타일링 84 nt 30 미스매치 거리	gacgtatgtaccaggttcctggaactggaatct ctttccatagaatgttctaaacCatcctgcggc ctctactctgcattcaatt (SEQ ID No. 559)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 28 미스매치 거리	gacacgtatgtaccaggttcctggaactggaat ctctttccatagaatgttctaaacCatcctgcgg cctctactctgcattcaa (SEQ ID No. 560)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 26 미스매치 거리	gcaacacgtatgtaccaggttcctggaactgg aatctctttccatagaatgttctaaacCatcctg cggcctctactctgcattc (SEQ ID No. 561)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 24 미스매치 거리	gcccaacacgtatgtaccaggttcctggaact ggaatctctttccatagaatgttctaaacCatcc tgcggcctctactctgcat (SEQ ID No. 562)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 22 미스매치 거리	ggacccaacacgtatgtaccaggttcctggaa ctggaatctctttccatagaatgttctaaacCat cctgcggcctctactctgc (SEQ ID No. 563)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 20 미스매치 거리	gttgacccaacacgtatgtaccaggttcctgga actggaatctctttccatagaatgttctaaacCa tcctgcggcctctactct (SEQ ID No. 564)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 18 미스매치 거리	gccttgacccaacacgtatgtaccaggttcctg gaactggaatctctttccatagaatgttctaaac Catcctgcggcctctact (SEQ ID No. 565)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 16 미스매치 거리	gttccttgacccaacacgtatgtaccaggttcct ggaactggaatctctttccatagaatgttctaaa cCatcctgcggcctcta (SEQ ID No. 566)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 14 미스매치 거리	gggttccttgacccaacacgtatgtaccaggtt cctggaactggaatctctttccatagaatgttcta aacCatcctgcggcctc (SEQ ID No. 567)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 12 미스매치 거리	gttggttccttgacccaacacgtatgtaccaggt tcctggaactggaatctctttccatagaatgttct aaacCatcctgcggcc (SEQ ID No. 568)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 10 미스매치 거리	gccttggttccttgacccaacacgtatgtacca ggttcctggaactggaatctctttccatagaatg ttctaaacCatcctgcgg (SEQ ID No. 569)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 8 미스매치 거리	ggcccttggttccttgacccaacacgtatgtac caggttcctggaactggaatctctttccatagaa	U6 발현을 위해 5' G를 갖는다	50C

[1462]

	tgttctaaacCatcctgc (SEQ ID No. 570)		
타일링 84 nt 6 미스매치 거리	gccgcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcct (SEQ ID No. 571)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 4 미스매치 거리	gcgccgcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatc (SEQ ID No. 572)	U6 발현을 위해 5' G를 갖는다	50C
타일링 84 nt 2 미스매치 거리	ggtcgccgcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCa (SEQ ID No. 573)	U6 발현을 위해 5' G를 갖는다	50C
ADAR 비표적화 가이드	GTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID No. 574)	U6 발현을 위해 5' G를 갖는다	50C
TAG 모티프 용 PFS 결합 스크린 가이드	gaaaacgcaggttcctcCagtttcgggagcagcgcacgtctccctgtagtc (SEQ ID No. 575)	U6 발현을 위해 5' G를 갖는다	51B
AAC 모티프 용 PFS 결합 스크린 가이드	gacgcaggttcctctagCttcgggagcagcgcacgtctccctgtagtcaag (SEQ ID No. 576)	U6 발현을 위해 5' G를 갖는다	51B
PFS 결합 스크린 비표적화	GTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID No. 577)	U6 발현을 위해 5' G를 갖는다	51B
모티프 선호도 표적화 가이드	gatagaatgttctaaacCatcctgcggcctctactctgcattcaattacat (SEQ ID No. 578)	U6 발현을 위해 5' G를 갖는다	51C
모티프 선호도 비표적화 가이드	GTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID No. 579)	U6 발현을 위해 5' G를 갖는다	51C
PPIB 타일링 가이드 50 미스매치 거리	gCaaggccacaaaattatccactgtttttggaacagtctttccgaagagac (SEQ ID No. 580)	U6 발현을 위해 5' G를 갖는다	57D
PPIB 타일링 가이드 42 미스매치 거리	gcctgtagcCaaggccacaaaattatccactgtttttggaacagtctttcc (SEQ ID No. 581)	U6 발현을 위해 5' G를 갖는다	57D
PPIB 타일링 가이드 34 미스매치 거리	gctttctctcctgtagcCaaggccacaaaattatccactgtttttggaaca (SEQ ID No. 582)	U6 발현을 위해 5' G를 갖는다	57D
PPIB 타일링 가이드 26 미스매치 거리	ggccaaatcctttctctcctgtagcCaaggccacaaaattatccactgttt (SEQ ID No. 583)	U6 발현을 위해 5' G를 갖는다	57D
PPIB 타일링 가이드 18 미스매치 거리	gtttttgtagccaaatcctttctctcctgtagcCaaggccacaaaattatc (SEQ ID No.	U6 발현을 위해 5' G를 갖는다	57D

[1463]

	584)		
PPIB 타일링 가이드 10 미스매치 거리	gatttgctgtttttgtagccaaatcctttctctcct gtagcCaaggccaca (SEQ ID No. 585)	U6 발현을 위해 5' G를 갖는다	57D
PPIB 타일링 가이드 2 미스매치 거리	gacgatggaatttgctgtttttgtagccaaatcct ttctctcctgtagcCa (SEQ ID No. 586)	U6 발현을 위해 5' G를 갖는다	57D
표적화 가이드, 반대쪽 염기 G	gatagaatgttctaaacGatcctgcggcctcta ctctgcattcaattacat (SEQ ID No. 587)	U6 발현을 위해 5' G를 갖는다	57D
표적화 가이드, 반대쪽 염기 A	gatagaatgttctaaacAatcctgcggcctcta ctctgcattcaattacat (SEQ ID No. 588)	U6 발현을 위해 5' G를 갖는다	57D
표적화 가이드, 반대쪽 염기 C	gatagaatgttctaaacTatcctgcggcctcta ctctgcattcaattacat (SEQ ID No. 589)	U6 발현을 위해 5' G를 갖는다	57D
AVPR2 가이드 37 미스매치 거리	ggtcccacgcggccCacagctgcaccagga agaagggtgcccagcacagca (SEQ ID No. 590)	U6 발현을 위해 5' G를 갖는다	52A
AVPR2 가이드 35 미스매치 거리	ggggtcccacgcggccCacagctgcaccag gaagaagggtgcccagcacag (SEQ ID No. 591)	U6 발현을 위해 5' G를 갖는다	52A
AVPR2 가이드 33 미스매치 거리	gccgggtcccacgcggccCacagctgcacc aggaagaagggtgcccagcac (SEQ ID No. 592)	U6 발현을 위해 5' G를 갖는다	52A
FANCC 가이드 37 미스매치 거리	gggtgatgacatccCaggcgatcgtgtggcc tccaggagcccagagcagga (SEQ ID No. 593)	U6 발현을 위해 5' G를 갖는다	52B
FANCC 가이드 35 미스매치 거리	gagggtgatgacatccCaggcgatcgtgtgg cctccaggagcccagagcag (SEQ ID No. 594)	U6 발현을 위해 5' G를 갖는다	52B
FANCC 가이드 32 미스매치 거리	gatcagggtgatgacatccCaggcgatcgtgt ggcctccaggagcccagag (SEQ ID No. 595)	U6 발현을 위해 5' G를 갖는다	52B
합성 질환 유전자 표적 IL2RG	ggtggctccattcactcCaatgctgagcacttc cacagagtgggttaaagc (SEQ ID No. 596)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 F8	gtttctaatatattttgCcagactgatggactatt ctcaattaataatgat (SEQ ID No. 597)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 LDLR	gagatgttgctgtggatCcagtccacagccag cccgtcgggggcctggatg (SEQ ID No. 598)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 CBS	gcaggccggcccagctgCcaggtgcacctg ctcggagcatcgggccggatc (SEQ ID No. 599)	U6 발현을 위해 5' G를 갖는다	52E

[1464]

합성 질환 유전자 표적 HBB	gcaaagaacctctgggtCcaagggtagaccaccagcagcctgcccagggcc (SEQ ID No. 600)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 ALDOB	gaagagaaacttagtttCcagggctttggtagagggcaaaggttgatagca (SEQ ID No. 601)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 DMD	gtcagcctagtgcagagCcactggtagttggtggttagagtttcaagttcc (SEQ ID No. 602)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 SMAD4	ggctcattgtgaacaggCcagtaatgtccgggatggggcggcataggcggg (SEQ ID No. 603)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 BRCA2	gtagctaaagaacttgaCcaagacatatcaggatccacctcagctcctaga (SEQ ID No. 604)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 GRIN2A	ggggcattgttctgtgcCcagtcctgctggtagacctgctccccggtggct (SEQ ID No. 605)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 SCN9A	gagaagtcgttcatgtgCcaccgtgggagcgtacagtcatcattgatcttg (SEQ ID No. 606)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 TARDBP	gggattaatgctgaacgCaccaaagttcatccccaccacccatattactacc (SEQ ID No. 607)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 CFTR	gctccaaaggctttcctCcactgttgcaaagttattgaatcccaagacaca (SEQ ID No. 608)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 UBE3A	gatgaatgaacgatttcCcagaactccctaatcagaacagagtccctggta (SEQ ID No. 609)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 SMPD1	ggagcctctgccggagcCcagagaacccgagagtcagacagagccagcgcc (SEQ ID No. 610)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 USH2A	ggcttccgtggagacacCcaatcaatttgaagagatcttgaagtgatgcca (SEQ ID No. 611)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 MEN1	gtgggactgccctcctcCcatttgcagatgccgtcgtagaatcgcagcagg (SEQ ID No. 612)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 C8orf37	gcttcttcaatagttctCcagctacactggcaggcatatgcccgtgttcct (SEQ ID No. 613)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 MLH1	gattccttttcttcgtcCcaattcacctcagtggctagtcgaagaatgaag (SEQ ID No. 614)	U6 발현을 위해 5' G를 갖는다	52E

[1465]

합성 질환 유전자 표적 TSC2	gcagcttcagcaccttcCagtcagactcctgctcaagcactgcagcagga (SEQ ID No. 615)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 NF1	gccatttgcttgcagtgCcactccagaggattccggattgccataaatact (SEQ ID No. 616)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 MSH6	gttcaatagttttggtcCagtatcgtttacagccctcttggtagatttca (SEQ ID No. 617)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 SMN1	ggcaaccgtcttctgacCaaatggcagaacatttgtccccaactttccact (SEQ ID No. 618)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 SH3TC2	gcgactttccaatgaacCactgaagcccaggtatgacaaagccgatgatct (SEQ ID No. 619)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 DNAH5	gtttacactcatgcttcCacagctttaacagatcatttggttccttgatga (SEQ ID No. 620)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 MECP2	gcttaagcttccgtgtcCagccttcaggcagggtggggtcatcatacatgg (SEQ ID No. 621)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 ADGRV1	ggacagctgggctgatcCatgatgtcatccagaaacactggggaccctcag (SEQ ID No. 622)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 AHI1	gtctcatctcaactttcCatatccgtatcatggaatcatagcatcctgtaa (SEQ ID No. 623)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 PRKN	gcatgcagacgcggttcCactcgcagccacagttccagcaccactcgagcc (SEQ ID No. 624)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 COL3A1	gttggttagggtcaaccCagtattctccactcttgagttcaggatggcaga (SEQ ID No. 625)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 BRCA1	gctacactgtccaacacCcactctcgggtcaccacaggtgcctcacacatc (SEQ ID No. 626)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 MYBPC3	gctgcactgtgtaccccCagagctccgtgttgccgacatcctggggtggct (SEQ ID No. 627)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 APC	gagcttcctgccactccCaacaggtttcacagtaagcgcgtatctgttcca (SEQ ID No. 628)	U6 발현을 위해 5' G를 갖는다	52E
합성 질환 유전자 표적 BMPR2	gacggcaagagcttaccCagtcacttgtgtggagacttaaatacttgcata (SEQ ID No. 629)	U6 발현을 위해 5' G를 갖는다	52E

[1466]

KRAS 타일링 가이드 50 미스매치 거리	gCaaggccacaaaattatccactgtttttggaacagtctttccgaagagac (SEQ ID No. 630)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 가이드 42 미스매치 거리	gcctgtagcCaaggccacaaaattatccactgtttttggaacagtctttcc (SEQ ID No. 631)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 가이드 34 미스매치 거리	gctttctctcctgtagcCaaggccacaaaattatccactgtttttggaaca (SEQ ID No. 632)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 가이드 26 미스매치 거리	ggccaaatcctttctctcctgtagcCaaggccacaaaattatccactgttt (SEQ ID No. 633)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 가이드 18 미스매치 거리	gtttttgtagccaaatcctttctctcctgtagcCaaggccacaaaattatc (SEQ ID No. 634)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 가이드 10 미스매치 거리	gatttgctgtttttgtagccaaatcctttctctcctgtagcCaaggccaca (SEQ ID No. 635)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 가이드 2 미스매치 거리	gacgatggaatttgctgtttttgtagccaaatccttctctcctgtagcCa (SEQ ID No. 636)	U6 발현을 위해 5' G를 갖는다	53A
KRAS 타일링 비표적화 가이드	GTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID No. 637)	U6 발현을 위해 5' G를 갖는다	53A
전사체 특이성 용 루시퍼라제 W85X 표적화 가이드	gatagaatgttctaaacCatcctgcggcctctactctgcattcaattacat (SEQ ID No. 638)	U6 발현을 위해 5' G를 갖는다	53B
전사체 특이성 용 비표적화 가이드	GCAGGGTTTTCCCAGTCACGACGTTGTAAAGTTG (SEQ ID No. 639)	U6 발현을 위해 5' G를 갖는다	53C
내생성 KRAS 가이드 2	gtcaaggcactcttgccCacgccaccagctccaactaccacaagtttatat (SEQ ID No. 640)	U6 발현을 위해 5' G를 갖는다	54F
내생성 PPIB 가이드 1	gcaaagatcacccggccCacatcttcatctccaattcgtaggtcaaaatac (SEQ ID No. 641)	U6 발현을 위해 5' G를 갖는다	54G
내생성 KRAS 가이드 1	GcgccaccagctccaacCaccacaagtttatattcagtcattttcagcagg (SEQ ID No. 642)	U6 발현을 위해 5' G를 갖는다	54F
내생성 KRAS 가이드 3	GtttctccatcaattacCacttgcttcctgtaggaatcctctattGTtgga (SEQ ID No. 643)	U6 발현을 위해 5' G를 갖는다	54F
내생성 PPIB 가이드 2	GctttctctcctgtagcCaaggccacaaaattatccactgtttttggaaca (SEQ ID No. 644)	U6 발현을 위해 5' G를 갖는다	54G

[1467]

내생성 비표적화 가이드	GTAATGCCTGGCTTGTCGAC GCATAGTCTG (SEQ ID No. 645)	U6 발현을 위해 5' G를 갖는다	54F
BoxB Cluc 가이드	tctttccataGGCCCTGAAAAAGG GCCtgttctaaacCatcctgcggcctctactc GGCCCTGAAAAAGGGCCattca attac (SEQ ID No. 646)	U6 발현을 위해 5' G를 갖는다	62B
BoxB 비표적화 가이드	cagctggcgaGGCCCTGAAAAAG GGCCggggatgtgcCgcaaggcgattaa gttggGGCCCTGAAAAAGGGC Cacgccagggt (SEQ ID No. 647)	U6 발현을 위해 5' G를 갖는다	62B
Stafforst 전체 길이 ADAR2 가이드1	GTGGAATAGTATAACAATAT GCTAAATGTTGTTATAGTAT CCCACtctaaaCCAtcctgcgGGGC CCTCTTCAGGGCCC (SEQ ID No. 648)	U6 발현을 위해 5' G를 갖는다	62C
Stafforst 전체 길이 ADAR2 비표적화 가이드	GTGGAATAGTATAACAATAT GCTAAATGTTGTTATAGTAT CCCACaccctggcgttacccaGGGCC CTCTTCAGGGCCC (SEQ ID No. 649)	U6 발현을 위해 5' G를 갖는다	62C

[1468]

[1469] 참조문헌

1. P. D. Hsu, E. S. Lander, F. Zhang, Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262-1278 (2014).
2. A. C. Komor, A. H. Badran, D. R. Liu, CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. Cell 168, 20-36 (2017).
3. L. Cong et al., Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823 (2013).
4. P. Mali et al., RNA-guided human genome engineering via Cas9. Science 339, 823-826 (2013).
5. B. Zetsche et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759-771 (2015).
6. H. Kim, J. S. Kim, A guide to genome engineering with programmable nucleases. Nat Rev Genet 15, 321-334 (2014).
7. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420-424 (2016).
8. K. Nishida et al., Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 353, (2016).
9. Y. B. Kim et al., Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol 35, 371-376 (2017).
10. O. O. Abudayyeh et al., C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353, aaf5573 (2016).

[1470]

11. S. Shmakov et al., Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol Cell 60, 385-397 (2015).
12. S. Shmakov et al., Diversity and evolution of class 2 CRISPR-Cas systems. Nat Rev Microbiol 15, 169-182 (2017).
13. A. A. Smargon et al., Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. Mol Cell 65, 618-630 e617 (2017).
14. J. S. Gootenberg et al., Nucleic acid detection with CRISPR-Cas13a/C2c2. Science 356, 438-442 (2017).
15. O. O. Abudayyeh et al., RNA targeting with CRISPR-Cas13a. Nature in press, (2017).
16. K. Nishikura, Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79, 321-349 (2010).
17. B. L. Bass, H. Weintraub, An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55, 1089-1098 (1988).
18. M. M. Matthews et al., Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. Nat Struct Mol Biol 23, 426-433 (2016).
19. A. Kuttan, B. L. Bass, Mechanistic insights into editing-site specificity of ADARs. Proc Natl Acad Sci U S A 109, E3295-3304 (2012).
20. S. K. Wong, S. Sato, D. W. Lazinski, Substrate recognition by ADAR1 and ADAR2. RNA 7, 846-858 (2001).
21. M. Fukuda et al., Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. Sci Rep 7, 41478 (2017).
22. M. F. Montiel-Gonzalez, I. Vallecillo-Viejo, G. A. Yudowski, J. J. Rosenthal, Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing. Proc Natl Acad Sci U S A 110, 18285-18290 (2013).
23. M. F. Montiel-Gonzalez, I. C. Vallecillo-Viejo, J. J. Rosenthal, An efficient system for selectively altering genetic information within mRNAs. Nucleic Acids Res 44, e157 (2016).
24. J. Wettengel, P. Reautschnig, S. Geisler, P. J. Kahle, T. Stafforst, Harnessing human ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy. Nucleic Acids Res 45, 2797-2808 (2017).
25. Y. Wang, J. Havel, P. A. Beal, A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1. ACS Chem Biol 10, 2512-2519 (2015).
26. K. A. Lehmann, B. L. Bass, Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. Biochemistry 39, 12875-12884 (2000).
27. Y. Zheng, C. Lorenzo, P. A. Beal, DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. Nucleic Acids Res 45, 3369-3377 (2017).
28. K. Gao et al., A de novo loss-of-function GRIN2A mutation associated with childhood focal epilepsy and acquired epileptic aphasia. PLoS One 12, e0170818 (2017).
29. H. M. Lanoiselee et al., APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases. PLoS Med 14, e1002270 (2017).
30. C. Ballatore, V. M. Lee, J. Q. Trojanowski, Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. Nat Rev Neurosci 8, 663-672 (2007).
31. Y. Li et al., Carriers of rare missense variants in IFIH1 are protected from psoriasis. J Invest Dermatol 130, 2768-2772 (2010).

[1471]

32. R. S. Finkel et al., Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. Lancet 388, 3017-3026 (2016).

[1472]

[1473] **실시예 4**

[1474] ADAR 활성에 영향을 미치는 ADAR 내 돌연변이는 효모 스크리닝을 사용하여 스크리닝하였다. 스크린은 다수 라운드로 수행되었다. 각 라운드의 스크리닝은 한 세트의 후보 돌연변이를 산출하였다. 그 다음으로 후보 돌연변이는 포유동물 세포에서 검증하였다. 상위-수행 돌연변이를 마지막 형태의 돌연변이에 첨가하였고 재스크리닝하였다. 8라운드에서 스크리닝된 돌연변이는 하기 표에 표시되어 있다. n 라운드에서 동정된 돌연변이체는 "RESCUE vn"로 표시하였다. 본 명세서에서 논의되는 바와 같이 RESCUE는 아데노신 데아미나제 활성을 시티딘 데아미나제 활성으로 전환시킨 돌연변이를 의미한다.

RESCUE 라운드	ADAR 돌연변이	플라스미드 번호
RESCUEv0	E488Q	pAB0048
RESCUEv1	E488Q, V351G	
RESCUEv2	E488Q, V351G, S486A	
RESCUEv3	E488Q, V351G, S486A, T375S	pAB0642
RESCUEv4	E488Q, V351G, S486A, T375S, S370C,	pAB1072
RESCUEv5	E488Q, V351G, S486A, T375S, S370C, P462A	pAB1135
RESCUEv6	E488Q, V351G, S486A, T375S, S370C, P462A, N597I	pAB1146

[1475]

[1476] 추가의 ADAR 돌연변이를 또한 스크리닝하였다. 이들 돌연변이의 활성을 시험하였다 (도 69). 스크리닝에서, 돌연변이 P462A가 모든 시험된 모티프에 대해 훨씬 양호하게 수행되었다 (도 70). 돌연변이 P462는 모티프를 개선시켰다 (도 71).

[1477] 출원인은 일부 활성이 증강된, 다른 돌연변이, N579I를 더 동정하였다 (도 72). N579I는 이 돌연변이를 RESCUEv3 (즉, pAB0642), RESCUEv4 (즉, pAB1072), RESCUEv5 (즉, pAB1135), 및 RESCUEv6 (즉, pAB1146)에 첨가하여 검증하였다. 플라스미드 pAB1143 및 pAB1144는 각각 RESCUEv4 및 RESCUEv5 상에서 G333A였다 (도 73).

[1478] 출원인은 또한 30 bp 가이드로 RESCUEv3의 성능을 시험하였다 (도 74). 30 bp 가이드는 다음과 같이 명명하였다: 첫번째 번호는 가이드의 길이 (즉, 30bp)이고, 두번째 번호는 가이드의 5' 말단으로부터 미스매치의 위치이다.

[1479] 그 다음으로 출원인은 50 bp 가이드를 2개 30 bp 가이드 (30_5 및 30_11)와 비교하였다. 다양한 모티프 상에서 이들 가이드에 의한 RESCUE 돌연변이체의 활성은 도 75에 도시되어 있다. 결과는 RESCUEv6가 모든 5' 모티프 (TC, GC, AC, 및 CC) 상에서 모든 다른 구성체를 능가하여서, 각각의 가능한 모티프 상에서 10% 초과의 편집을 가능하게 한다는 것을 보여준다. GC에 대한 최적 가이드 디자인은 TC, AC, 및 CC 디자인과 약간 상이하였다는 것을 유의한다. 따라서 본 명세서에서는 각 실시형태에 대해 가이드 디자인을 최적화하기 위한 방법을 제공한다.

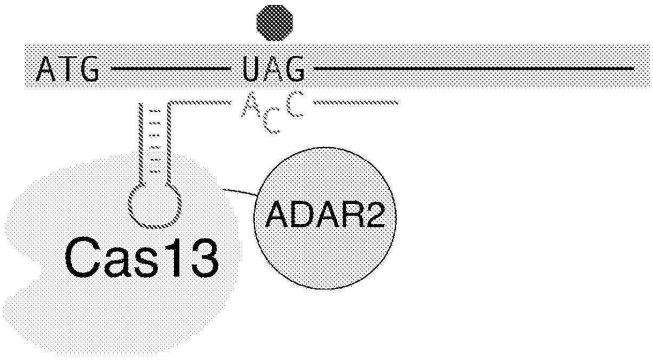
[1480] 일정 실시형태에서, 데아미나제 도메인 단독 또는 표적화 Cas13 가이드와 전달된 전체-길이 ADAR 상에서 RESCUE 돌연변이의 위치는 편집을 허용하지 않았다. 데아미나제 도메인은 프로그램가능한 활성을 위해 Cas13과 융합되었을 때 활성을 보였다 (도 76).

[1481] ***

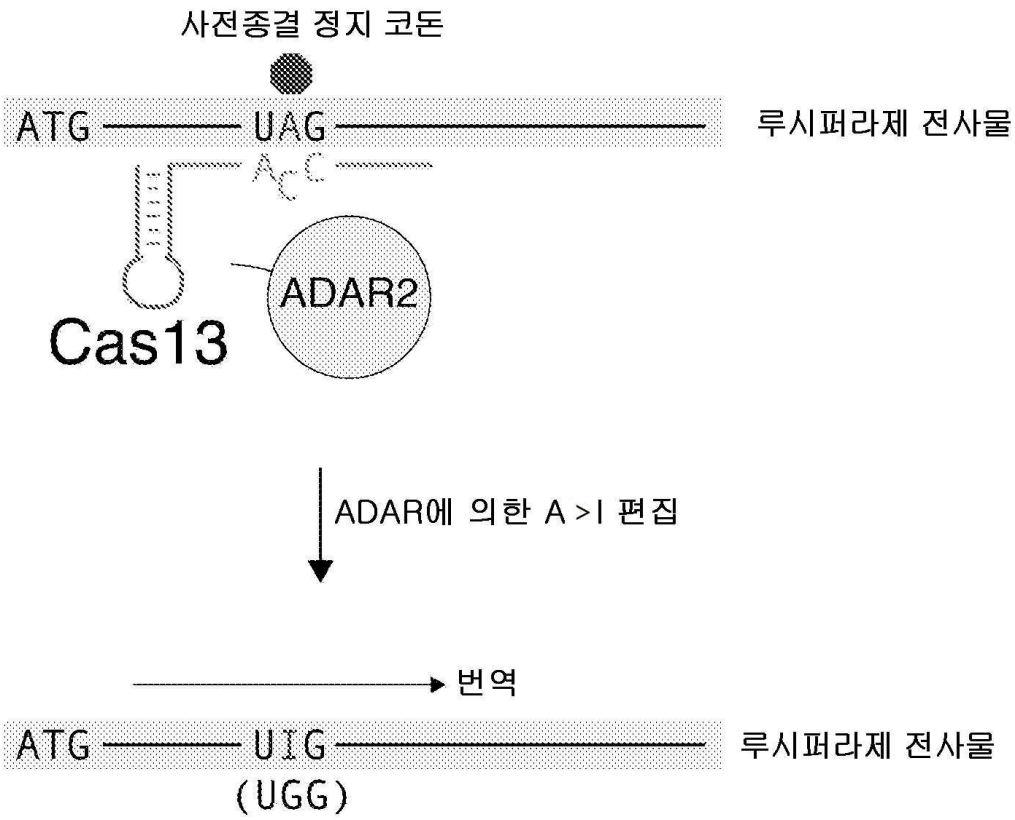
[1482] 기술된 본 발명의 방법, 약학 조성물, 및 키트의 다양한 변형 및 변동은 본 발명의 범주 및 사조를 벗어나지 않고 당업자에게 자명해질 것이다. 본 발명을 특별한 실시형태와 함께 기술하였지만, 추가로 변형할 수 있고 청구된 발명을 이러한 특별한 실시형태에 과도하게 제한해서는 안된다는 것을 이해하게 될 것이다. 실제로, 당업자에게 명확한 본 발명의 수행을 위해 기술된 방식의 다양한 변형을 본 발명의 범주 내에 두고자 한다. 본 출원은 일반적으로 본 발명의 원리에 따라서 본 발명의 임의의 변동, 용도, 또는 개조를 포괄하고자 하고 본 개시로부터의 이러한 이탈의 포함은 본 발명이 속하는 분야에서 공지된 통상의 관례 내에 속하며 앞서 기재된 필수 특성에 적용될 수 있다.

도면

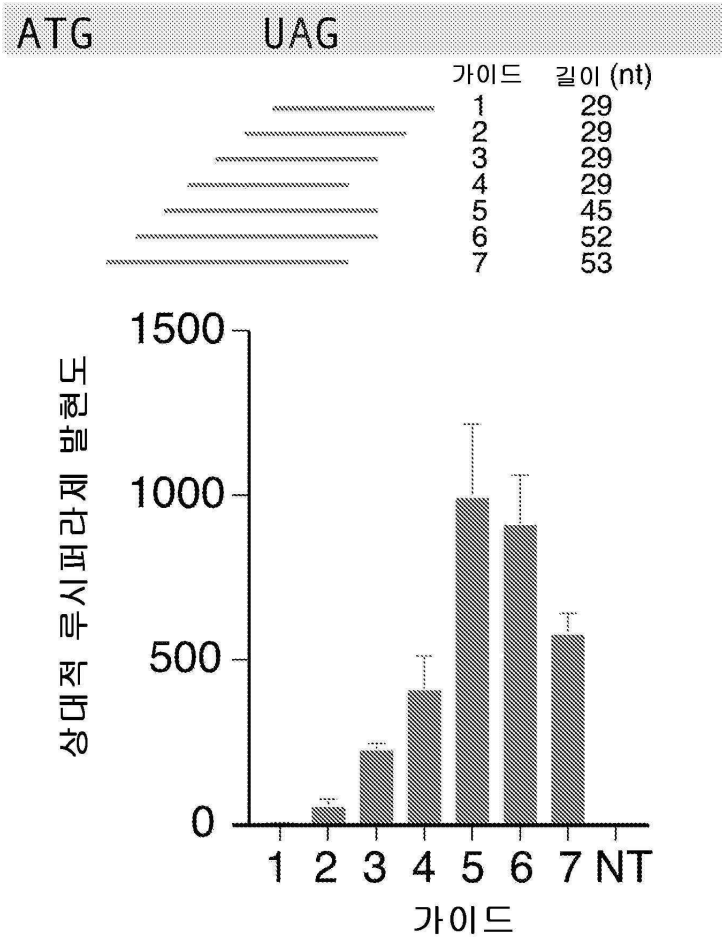
도면1


ATG
UAG
ADAR2
Cas13

도면2


사전종결 정지 코돈
ATG
UAG
ADAR2
Cas13
루시퍼라제 전사물
ADAR에 의한 A >I 편집
번역
ATG
UIG
(UGG)
루시퍼라제 전사물

도면3


ATG
UAG
가이드
길이 (nt)
1 29
2 29
3 29
4 29
5 45
6 52
7 53
상대적 루시퍼라제 발현도
1500
1000
500
0
1 2 3 4 5 6 7 NT
가이드

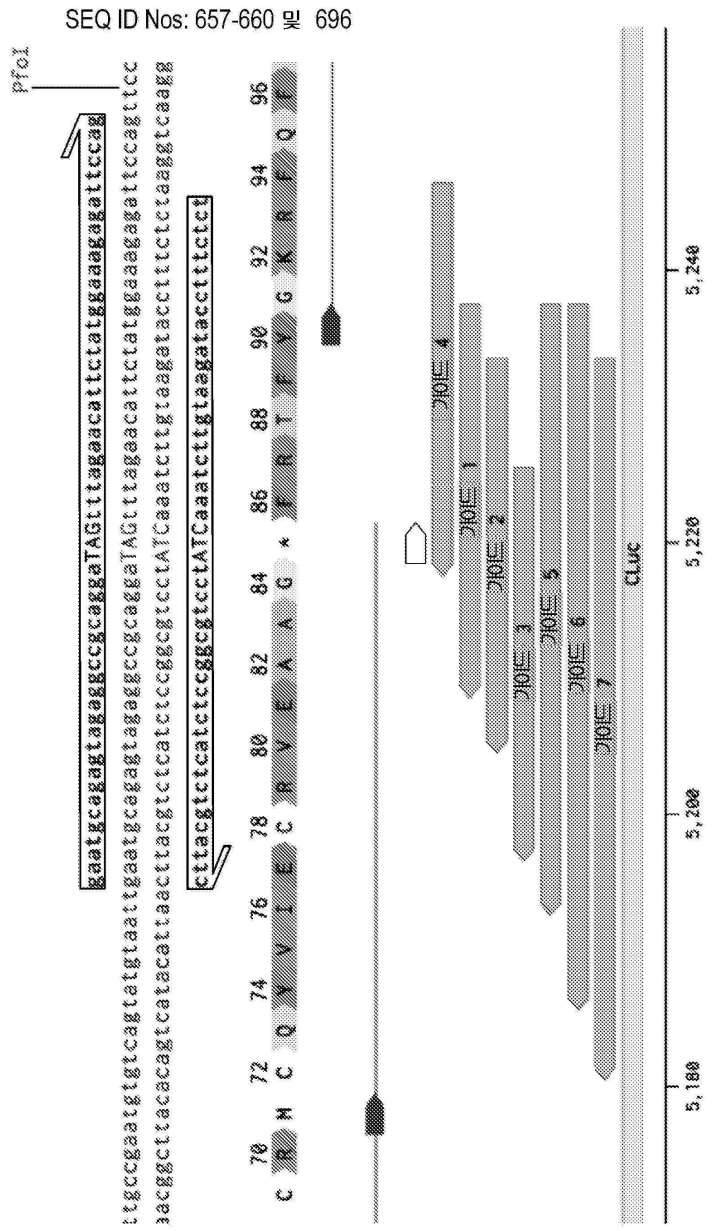
도면4a

ADAR 명칭	단백질 서열
호모 사피언스_ADAR2_E_Q_돌연변이체 (SEQ ID NO : 650)	QLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMT TGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSL LRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYIST SPCGDARIFSPHEPILEEPADRHPNRKARGQLRTKIESGQGTIPVR SNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEP IYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISN AEARQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKH ALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKA RLFTAFIKAGLGAWVEKPTEQDQFSLT*
호모 사피언스_ADAR1_E_Q 돌연변이체 (SEQ ID NO : 651)	SLGTGNRCVKGDSLSLKGETVNDCHAEIISRRGFIRFLYSELMKY NSQTAKDSIFEPAKGGEKLQIKKTVSFHLYISTAPCGDGALFDKS CSDRAMESTESRHYPVFENPKQGKLRTKVENGQGTIPVESSDIVP TWDGIRLGERLRTMSCSDKILRWNVLGLQGALLTHFLQPIYLKS VTLGYLFSQGHLTRAICCRVTRDGSAFEDGLRHPFIVNHPKVGR VSIYDSKRQSGKTKETSVNWCLADGYDLEILDGTRGTVDGPRNE LSRVSKKNIFLLFKKLCSFRYRRDLLRLSYGEAKKAARDYETAK NYFKKGLKDMGYGNWISKPQEEKNF
옥토푸스 불가리스_ADAR1_E_Q_돌연변이체 (SEQ ID NO : 652)	SVGTGNRCLTGDHLSLEGNSVNDSHAEMITRRGFLRYLYRHLLE YDAEVPNDLFEKGERSICRIKTNITFHLYISTAPCGDGALFSPRDT DSSNAKMEEENKHIHNPTFSSSVQGLLRTKVEGGQGTIPIDADFT EQTWDGIQRGERLRTMSCSDKICRWNVVGLQGALLSHFIEPIYL DSLTLGYLYDHGHLARAVCCRIERGEASVNQLLPEGYRLNHPW LGRVTACDPPRETQKTKSLSINWCYDDEKSEVLDGTAGICYTAIE KNLFSRLTKHNLYEEFKRVCRKFDRNDLLTAPSYNKAKMMATP FQTAKNVMLKKLKENNCGTWVSKPIEEEMF
세피아_ADAR1_E_Q_돌연변이체 (SEQ ID NO : 653)	SVGTGNRCLTGDRLSLEGNSVNDSHAEMVTRRGFLRYLYKHLL EYDPEKPHDLFEKGERSLCRIKTNITFHLYISTAPCGDGALFSPRD TDSSNVKVDEENKHVHNPTFSSSVQGLLRTKVEGGQGTIPIDADF TEQTWDGIQRGERLRTMSCSDKICRWNVVGLQGALLSHFVEPIY LESLTLGYLYDHGHLARAVCCRIERGEASVNQLLPEGYRLNHPW LGRVTACDPPRETQKTKSLSINWCYDDEKSEVLDGTAGICYTAIE KNLFSRLTKHSLYEEFKKVCQKFEREDLLNVTSYNKAKMMAIPF QTAKNVMLKKLKENNCGTWVSKPIEEEMF

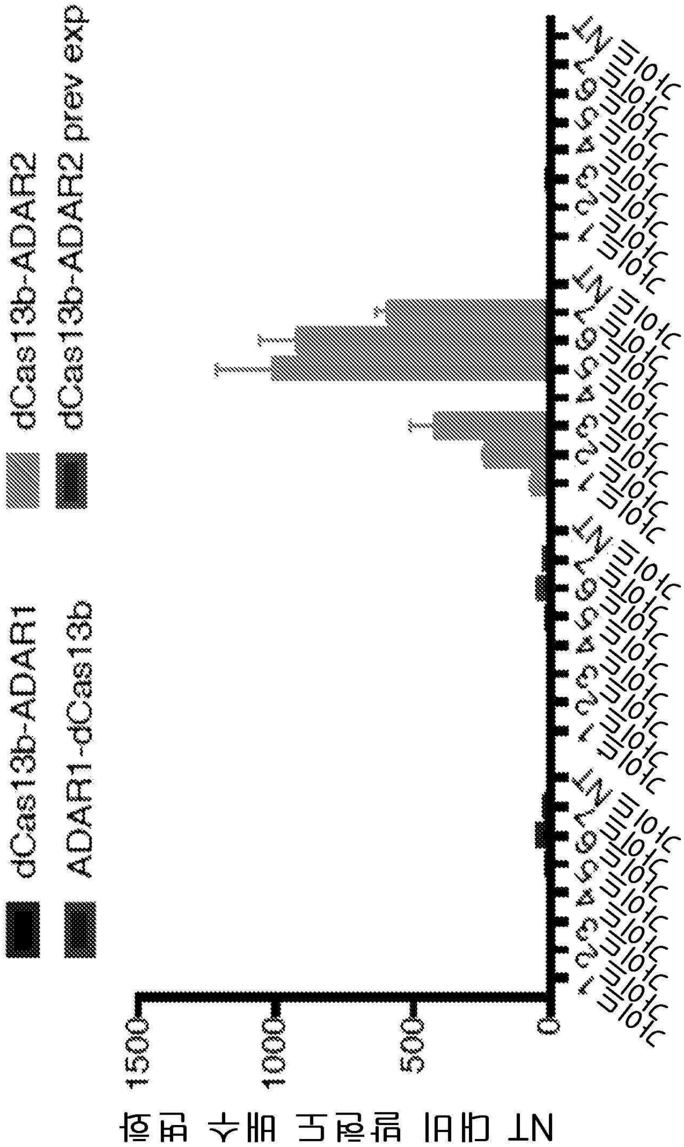
도면4b

옥토푸스 불가리스_ADAR2_E_Q_ 돌연변이체 (SEQ ID NO : 654)	GIGTGTKCINGEHMSDRGFGVNDCHAEIIARRCFLRYIYDQLELH LSDNSDVRNSSIFELRDKGGYQLKENIQFHLYISTAPCGDARIFSP HGQDVETGDRHPNRKARGQLRTKIESGQGTIPVRTSGVIQTWDG VLEGERLLTMSCSDKIARWNVLGIQGSLLSHFMNPIYLESIILGSL YHSDHLSRAMYSRISIIENLPEPFHLNRPFLSGISSPESRQPGKAPN FGINWRKEDETFEVINAMTGRVEGGSVSRICKQALFGRFMSLYG KLSSLTGQSVTTRPTHYSDAKAAVMEYQLAKQCVFQAFQKAGL GNWVQKPIEQDQF
세피아_ADAR2_E_Q_ 돌연변이체(SEQ ID NO : 655)	GIGTGTKCINGEYMNDRGFAVNDCHAEIIARRCFLRFIYDQLEMH LSEDPEVRGQSVFELRDGGGYKLKPNIHFHLYISTAPCGDARIFSP HGQDVETGDRHPNRKARGQLRTKIESGQGTIPVRSSGFIQTWDG VLEGERLLTMSCSDKIARWNVLGIQGALLCHFMHPIYLESIILGSL YHSDHLSRAVYCRIASIENLPDLFQLNRPFLSGISSPESRQPGKAP NFGINWRRNDDTFEVINAMTGRVEGGNMSRICKQALFDRFMNL YGRLSSLTGQSVTTRPTLYSEAKAAVMEYQLAKQCVFQAFQKA GLGNWVQKPIEQDQF
도리테우스티스 오팔레센스_ADAR2_E_Q_ 돌연변이체 (SEQ ID NO : 656)	GIGTGTKCINGEYMNDRGFAVNDCHAEIIARRCFLRFIYDQLELH LSDNAEVRGQSIFELRDAGGYKLKPNIQFHLYISTAPCGDARIFSP HGQDVETGDRHPNRKARGQLRTKIESGQGTIPVRSSGFIQTWDG VLEGERLLTMSCSDKIARWNVLGVQGALLCHFMHPIYLESIILGS LYHSDHLSRAVYCRIAAIENLPDLFRLNRPFLSGISSPESRQPGKA PNFGINWRRNDDSFEVINAMTGRVEGGSMSRICKQALFDRFMNL YGKLSSLTGQSVTTRPALYSEAKATVMEYQLAKQCVFQAFQKA GLGNWVQKPIEQDQF

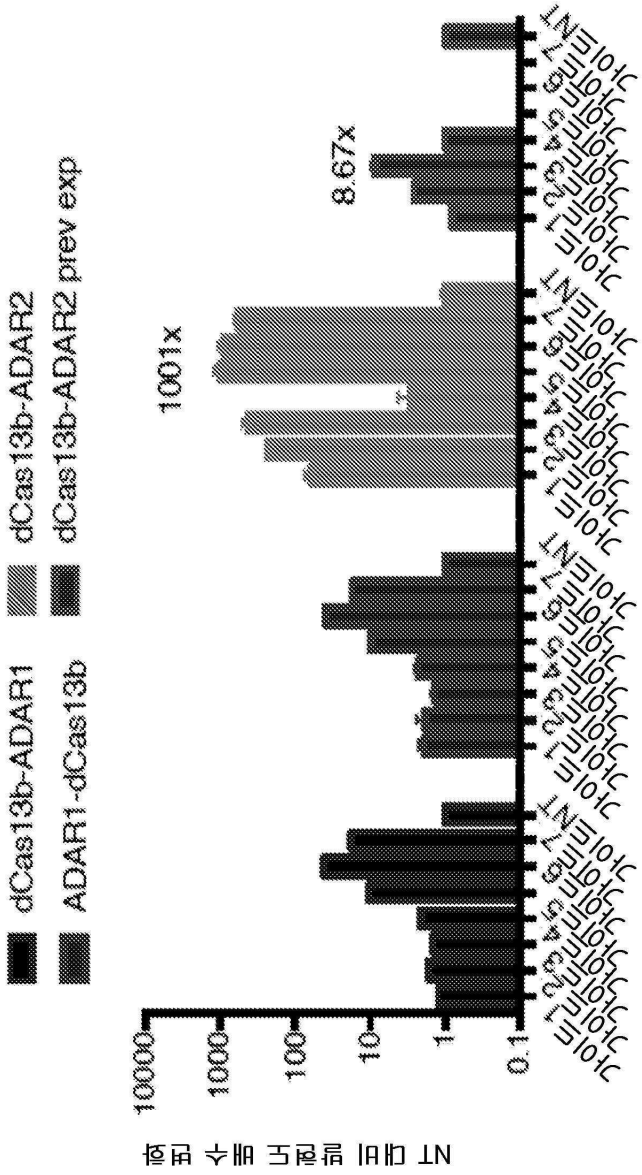
도면5


SEQ ID Nos: 657-660 및 696
PfoI
5,180
5,200
5,220
5,240
CLuc
카이트 1
카이트 2
카이트 3
카이트 4
카이트 5
카이트 6
카이트 7

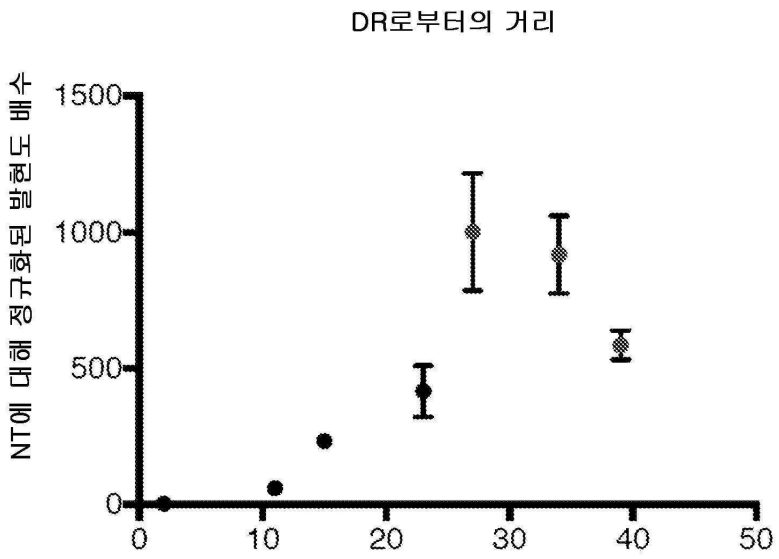
도면6


dCas13b-ADAR1
ADAR1-dCas13b
dCas13b-ADAR2
dCas13b-ADAR2 prev exp
NT 대비 발광도 배수 변화
1500
1000
500
0

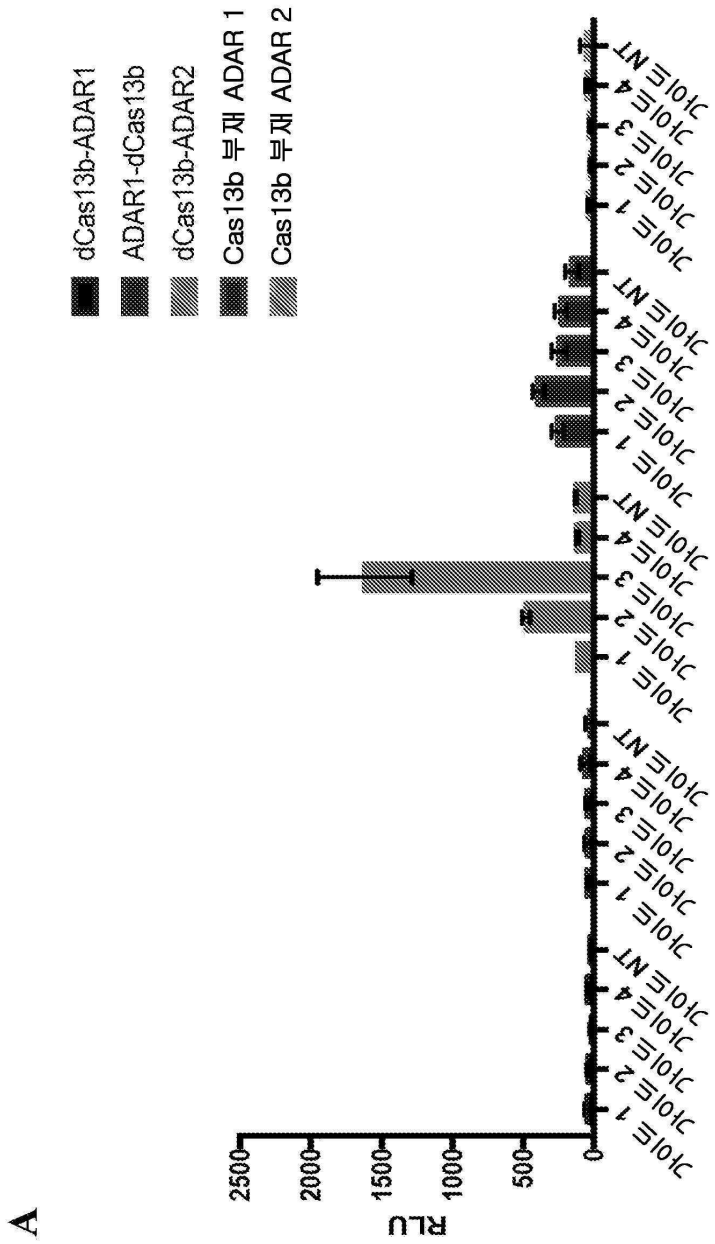
도면7


dCas13b-ADAR1
ADAR1-dCas13b
dCas13b-ADAR2
dCas13b-ADAR2 prev exp
1001x
8.67x
NT 대비 발현도 배수 변화
10000
1000
100
10
1
0.1
NT
G4C2 1
G4C2 2
G4C2 3
G4C2 4
G4C2 5
G4C2 6
G4C2 7
G4C2 8
G4C2 9
G4C2 10

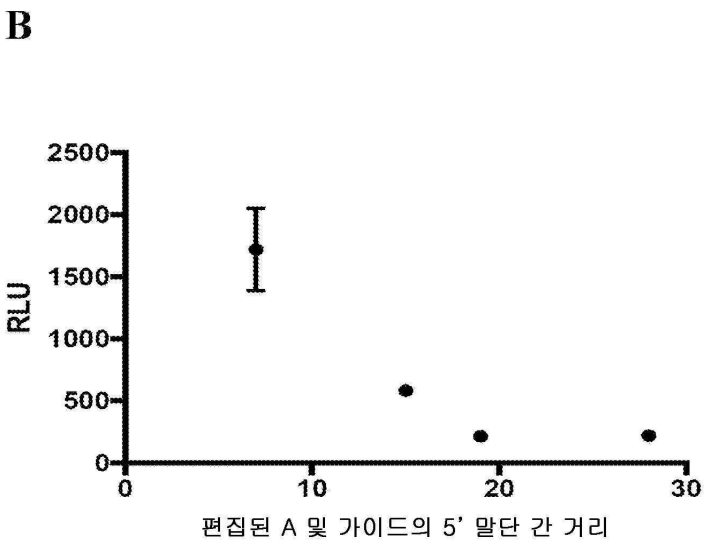
도면8



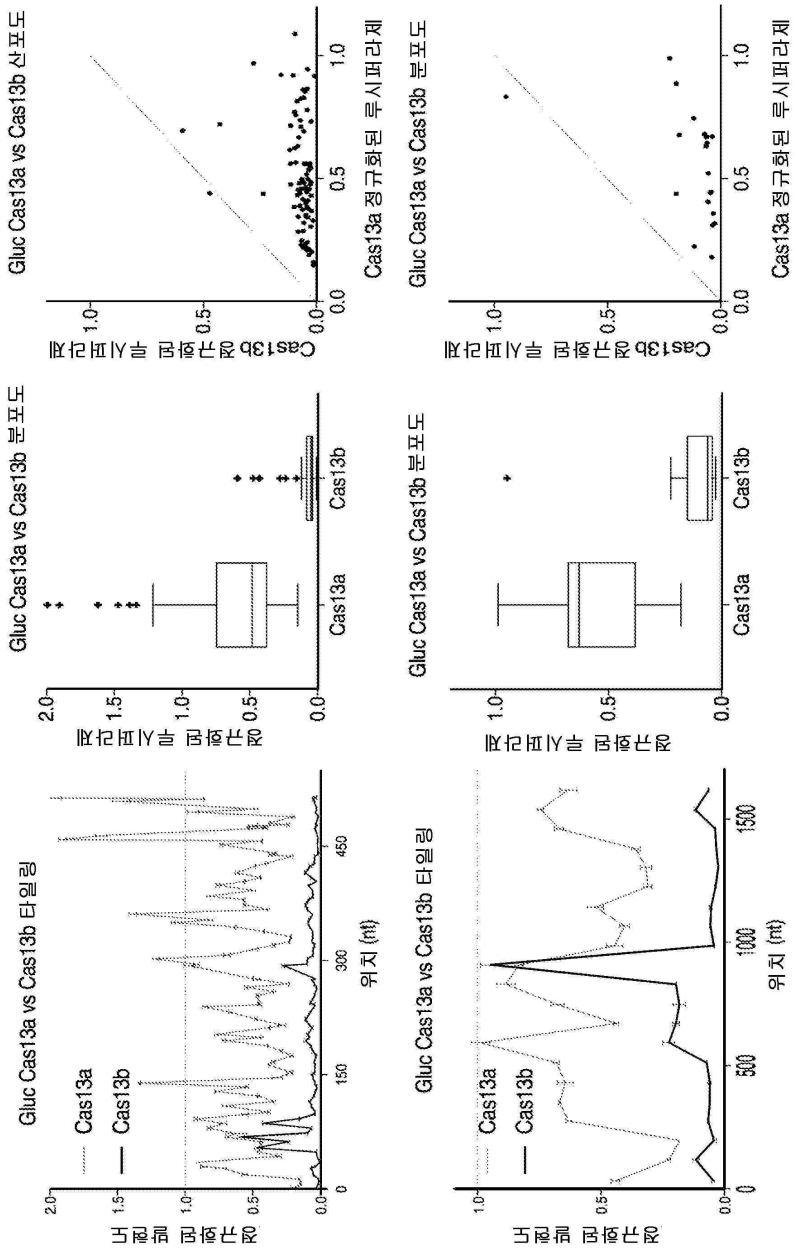
도면9a


A
dCas13b-ADAR1
ADAR1-dCas13b
dCas13b-ADAR2
Cas13b 부재 ADAR 1
Cas13b 부재 ADAR 2
RLU
2500
2000
1500
1000
500
0
가이드 1
가이드 2
가이드 3
가이드 4
가이드 NT

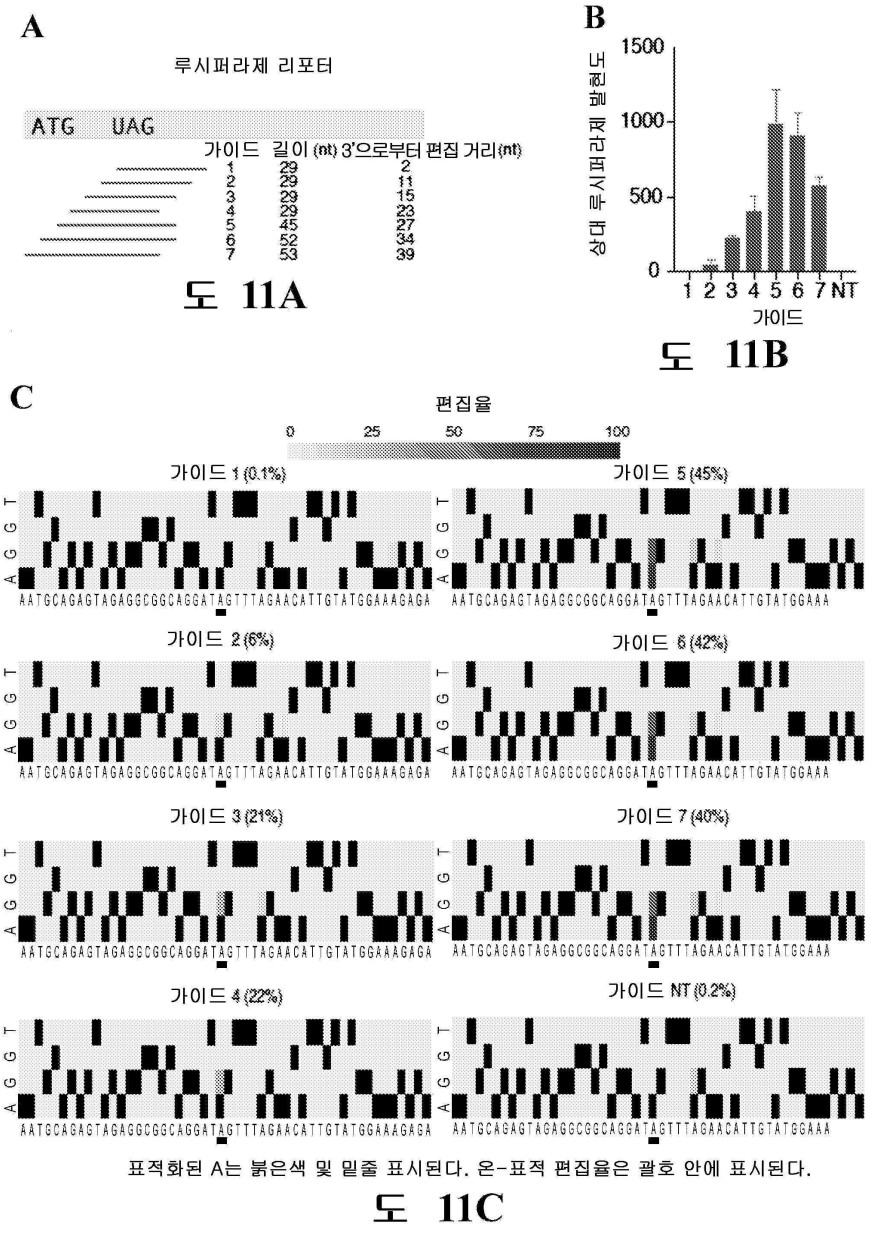
도면9b



도면10


Gluc Cas13a vs Cas13b 타일링
Cas13a
Cas13b
정규화된 발현도
위치 (nt)
Gluc Cas13a vs Cas13b 분포도
정규화된 루시퍼라제
Cas13a
Cas13b
Gluc Cas13a vs Cas13b 산포도
Cas13b 정규화된 루시퍼라제
Cas13a 정규화된 루시퍼라제

도면11



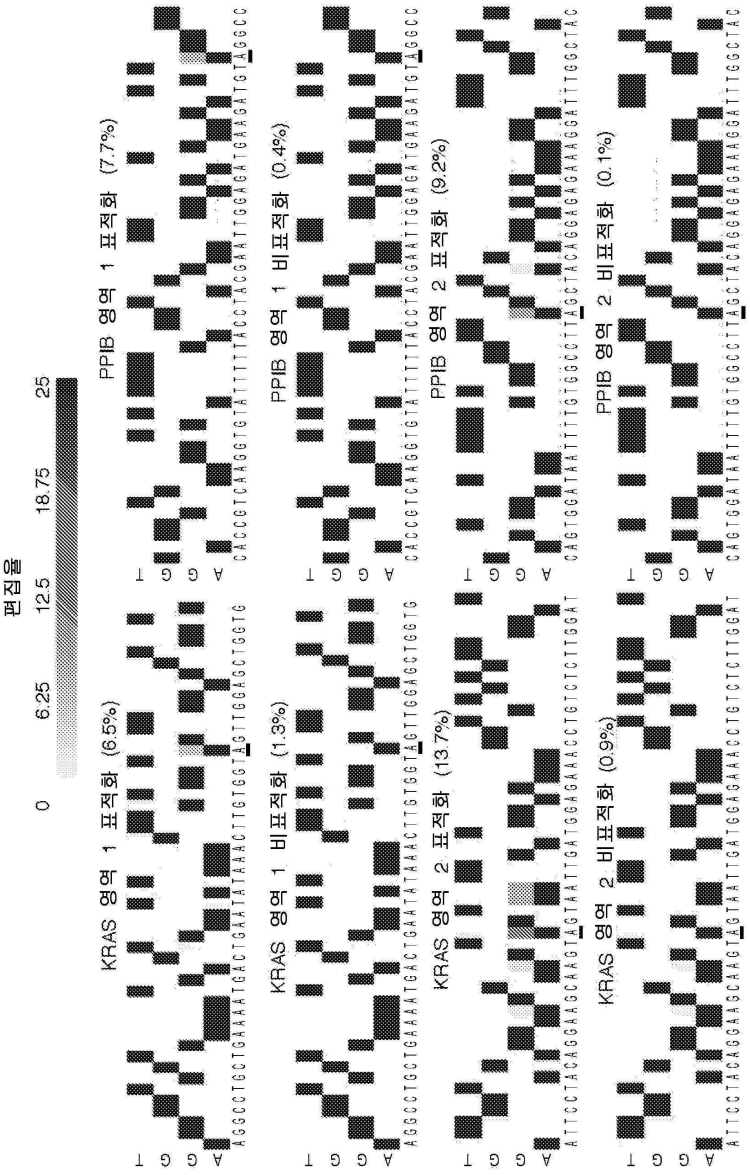
도 11A

도 11B

표적화된 A는 붉은색 및 밑줄 표시된다. 온-표적 편집율은 괄호 안에 표시된다.

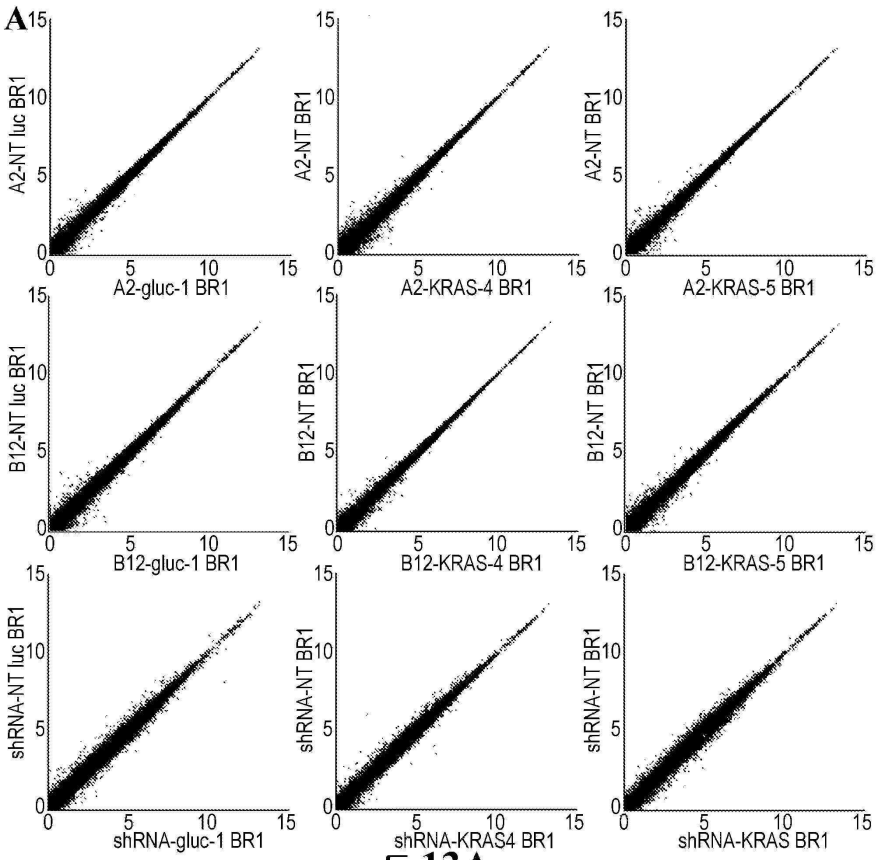
도 11C

도면12

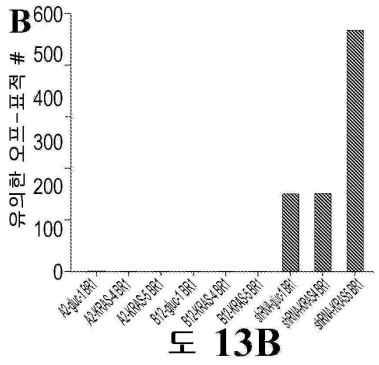


표적화된 A는 붉은색 및 밑줄표시된다. 온-표적 편집율은 괄호 안에 표시된다

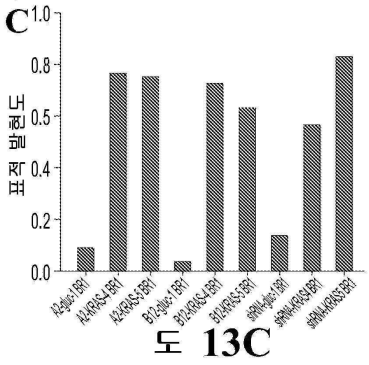
도면13



도 13A

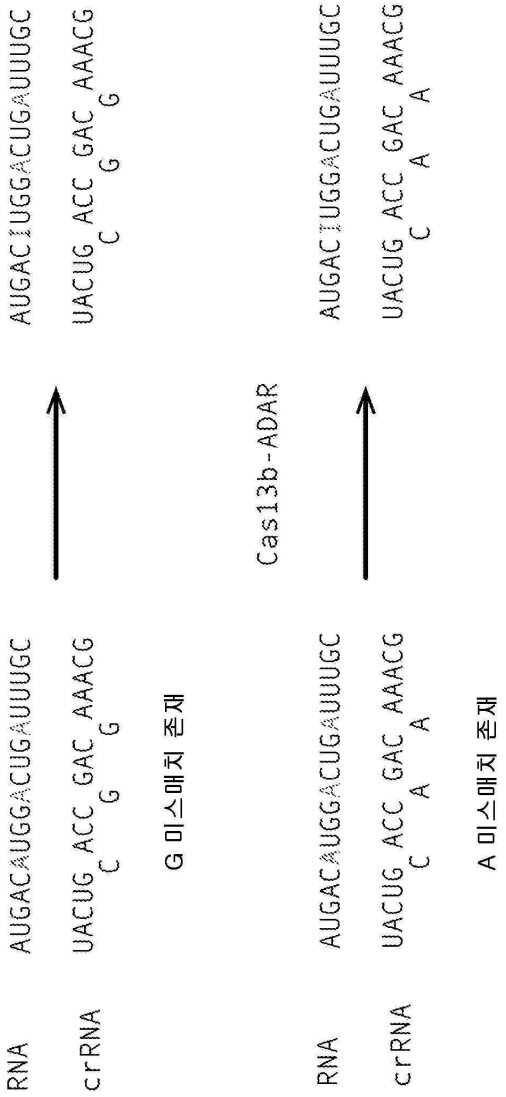


도 13B

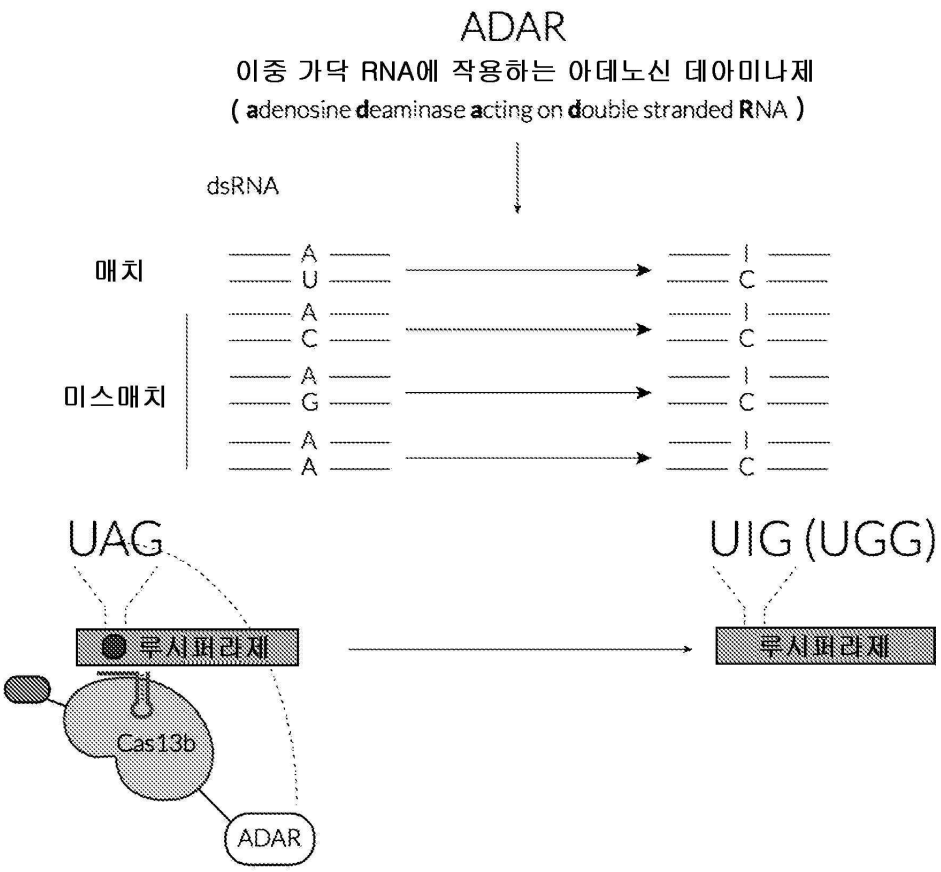


도 13C

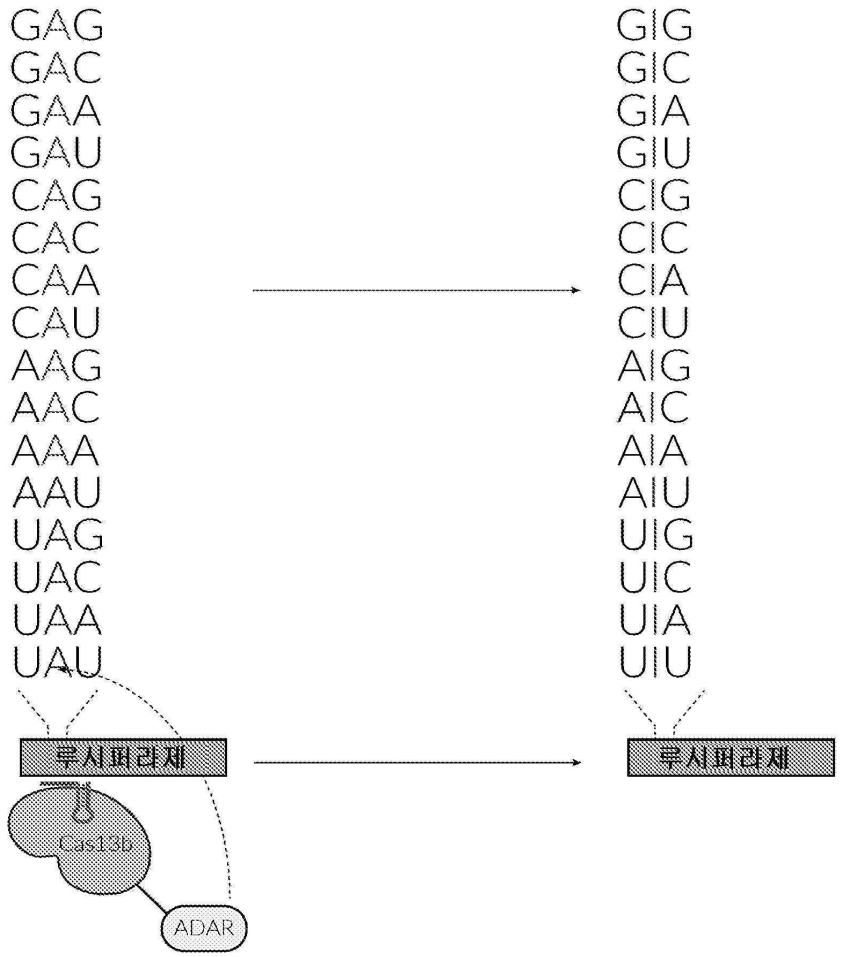
도면14


RNA
AUGACAUGGACUGAUUUGC
crRNA
UACUG ACC GAC AAACG
C G G
G 미스매치 존재
Cas13b-ADAR
RNA
AUGACIUGGACUGAUUUGC
crRNA
UACUG ACC GAC AAACG
C A A
A 미스매치 존재

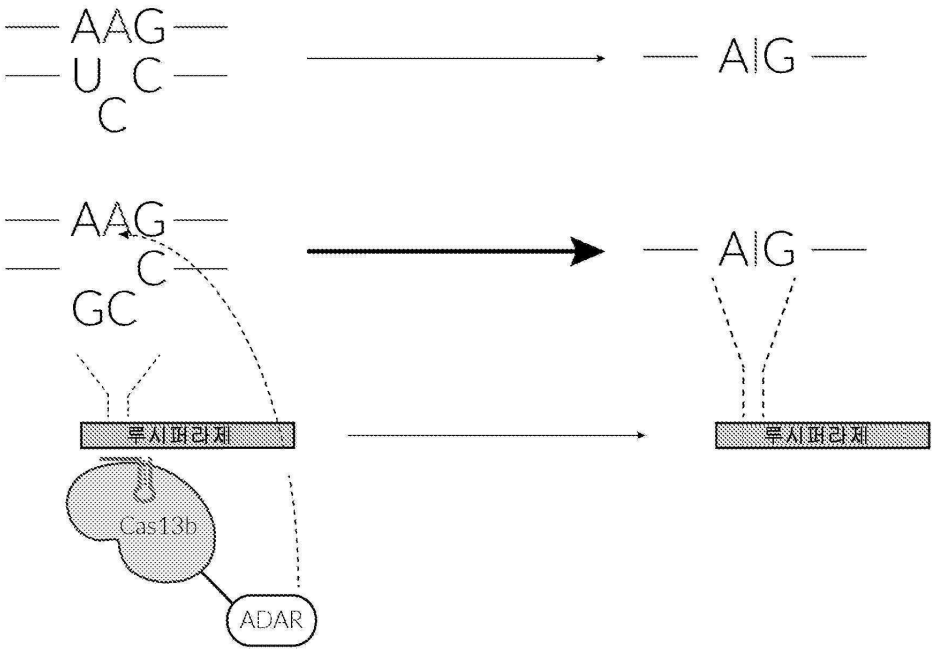
도면15


ADAR
이중 가닥 RNA에 작용하는 아데노신 데아미나제
(adenosine deaminase acting on double stranded RNA)
dsRNA
매치
미스매치
A
U
I
C
A
C
I
C
A
G
I
C
A
A
I
C
UAG
UIG (UGG)
루시퍼라제
루시퍼라제
Cas13b
ADAR

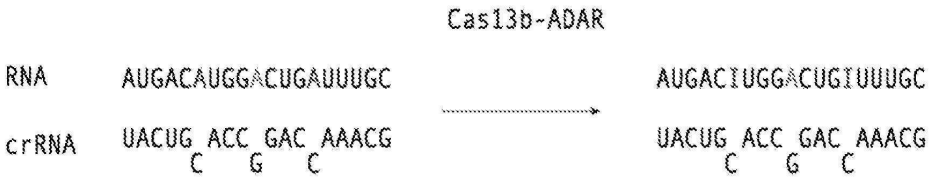
도면16


GAG
GAC
GAA
GAU
CAG
CAC
CAA
CAU
AAG
AAC
AAA
AAU
UAG
UAC
UAA
UAU
G|G
G|C
G|A
G|U
C|G
C|C
C|A
C|U
A|G
A|C
A|A
A|U
U|G
U|C
U|A
U|U
루시퍼라제
루시퍼라제
Cas13b
ADAR

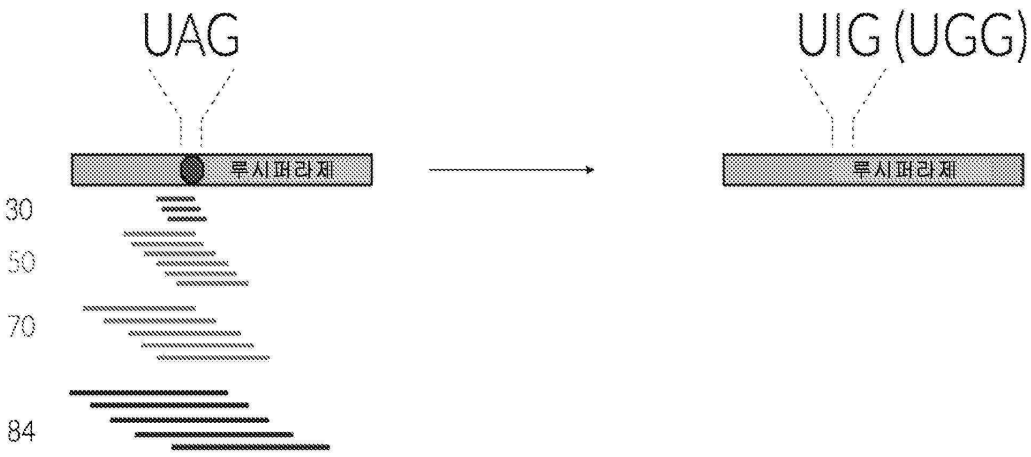
도면17


AAG
U C
C
A|G
AAG
C
GC
A|G
루시퍼라제
루시퍼라제
Cas13b
ADAR

도면18

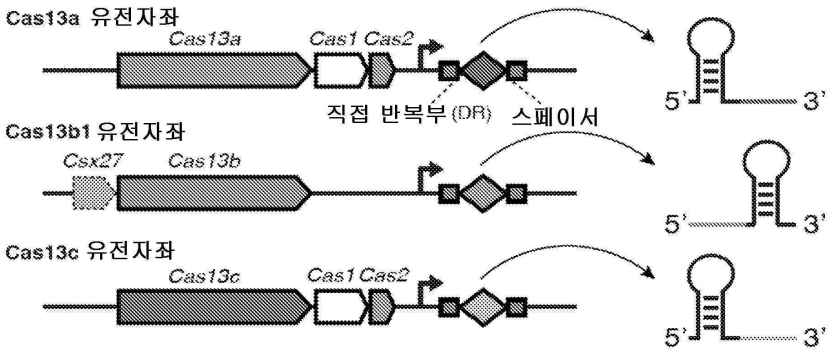

Cas13b-ADAR
RNA
AUGACAUGGACUGAUUUGC
AUGACIUGGACUGIUUUGC
crRNA
UACUG ACC GAC AAACG
C G C
UACUG ACC GAC AAACG
C G C

도면19


UAG
UIG (UGG)
루시퍼라제
루시퍼라제
30
50
70
84

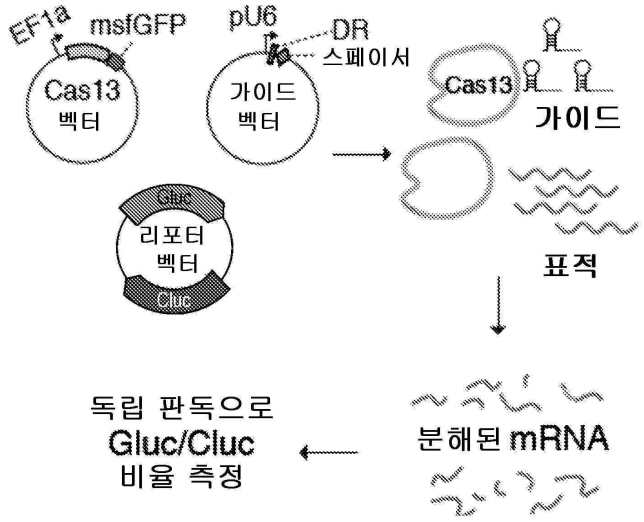
도면20ab

A



도 20A

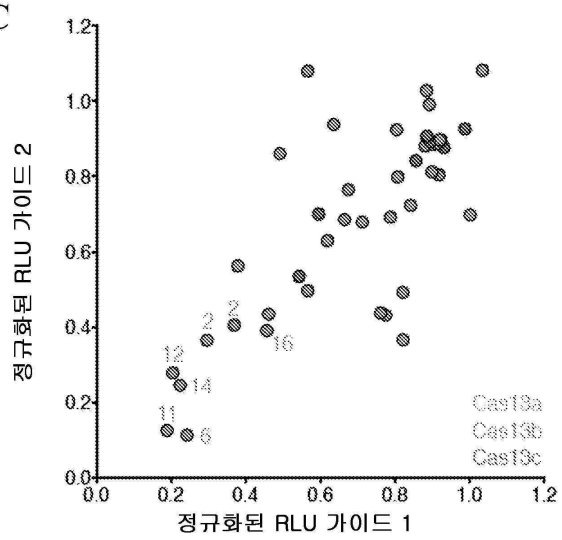
B



도 20B

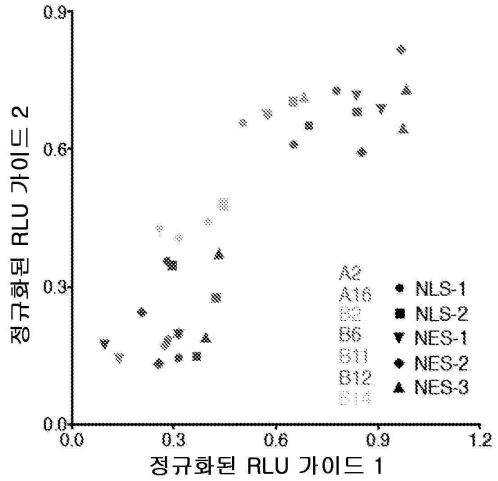
도면20cd

C



도 20C

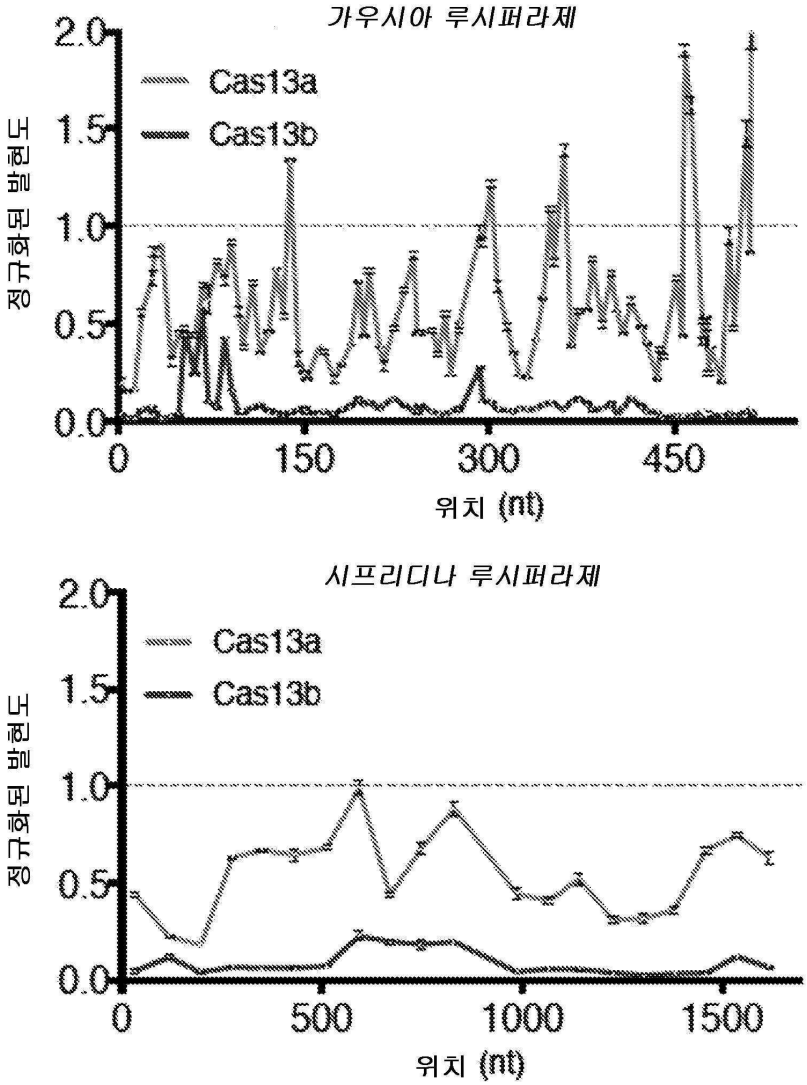
D



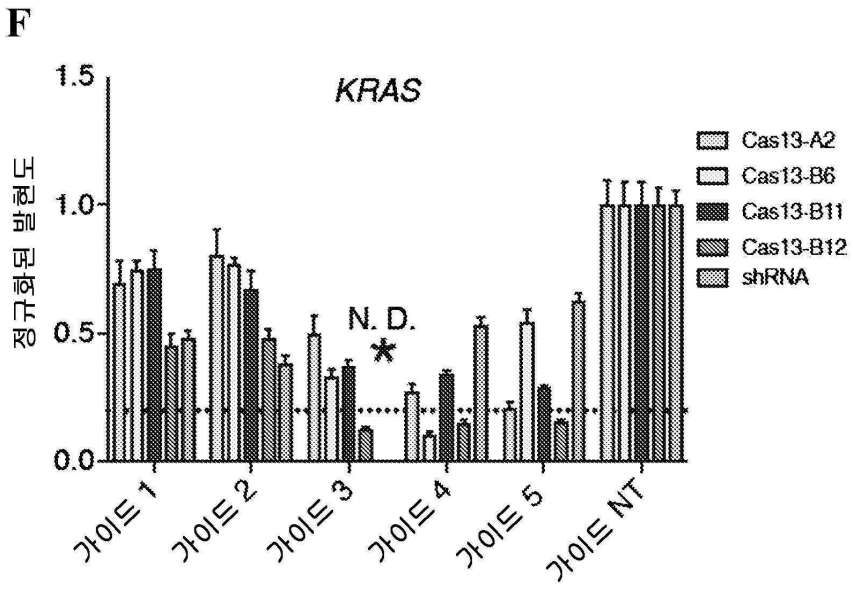
도 20D

도면20e

E


가우시아 루시퍼라제
Cas13a
Cas13b
정규화된 발현도
2.0
1.5
1.0
0.5
0.0
0
150
300
450
위치 (nt)
시프리디나 루시퍼라제
Cas13a
Cas13b
정규화된 발현도
2.0
1.5
1.0
0.5
0.0
0
500
1000
1500
위치 (nt)

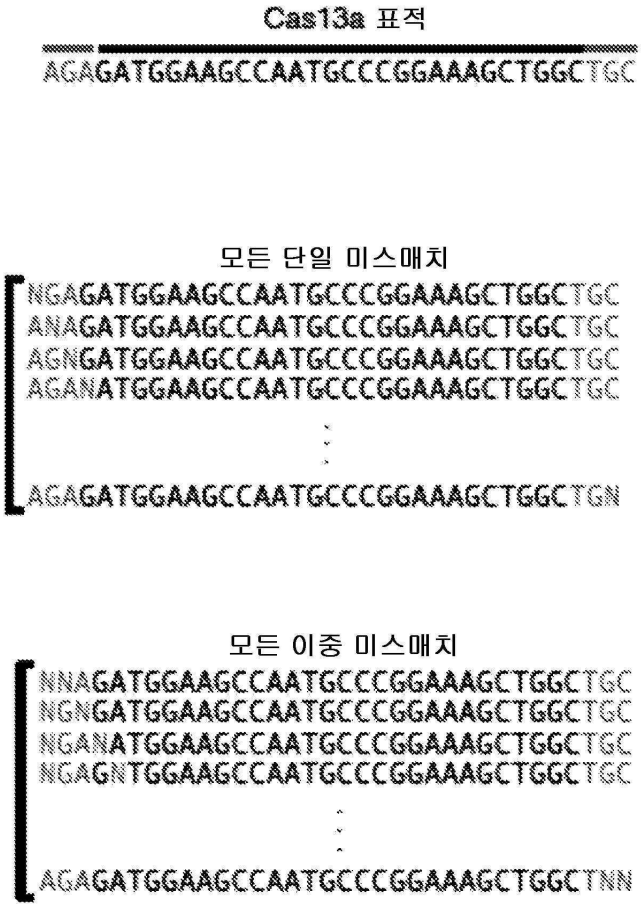
도면20f



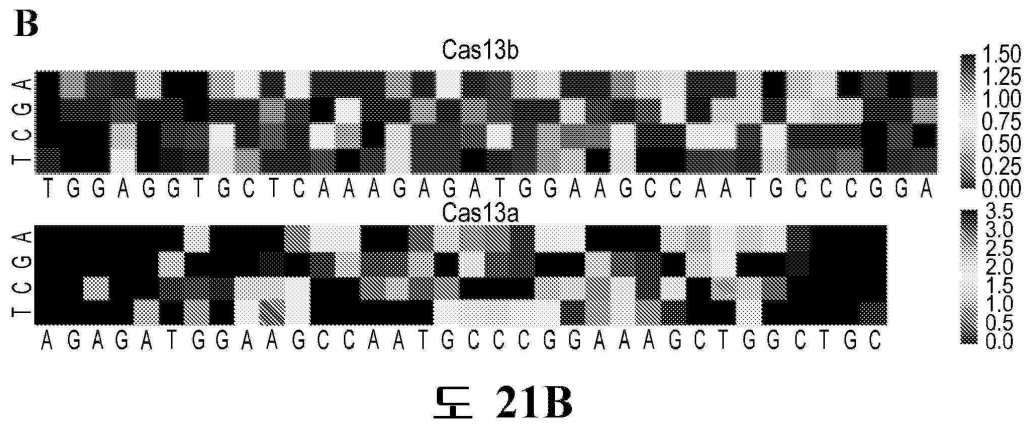
도면21aa



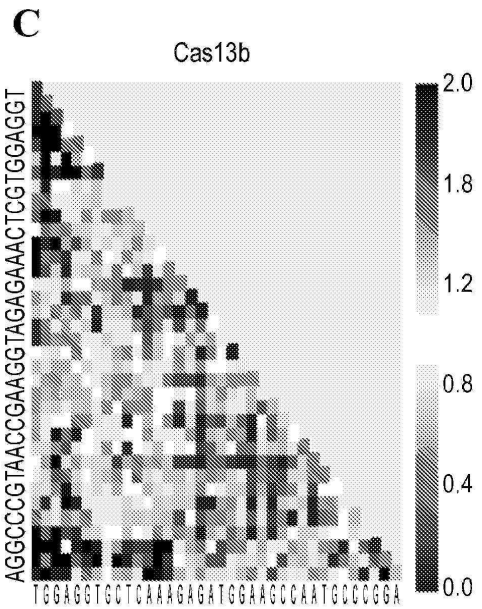
도면21ab


Cas13a 표적
AGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGC
모든 단일 미스매치
NGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGC
ANAGATGGAAGCCAATGCCCGGAAAGCTGGCTGC
AGNGATGGAAGCCAATGCCCGGAAAGCTGGCTGC
AGANATGGAAGCCAATGCCCGGAAAGCTGGCTGC
AGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGN
모든 이중 미스매치
NNAGATGGAAGCCAATGCCCGGAAAGCTGGCTGC
NGNGATGGAAGCCAATGCCCGGAAAGCTGGCTGC
NGANATGGAAGCCAATGCCCGGAAAGCTGGCTGC
NGAGNTGGAAGCCAATGCCCGGAAAGCTGGCTGC
AGAGATGGAAGCCAATGCCCGGAAAGCTGGCTNN

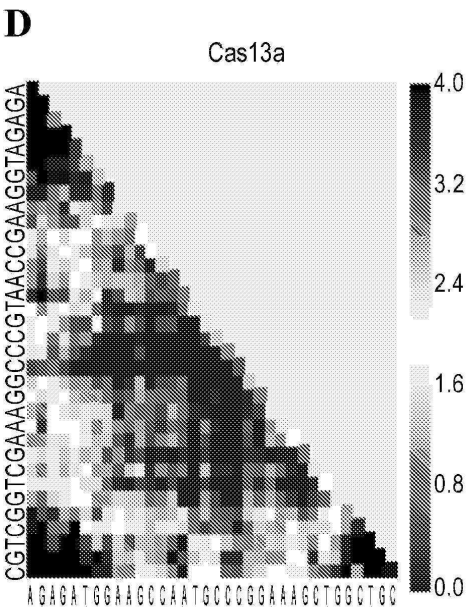
도면21bcd



도 21B

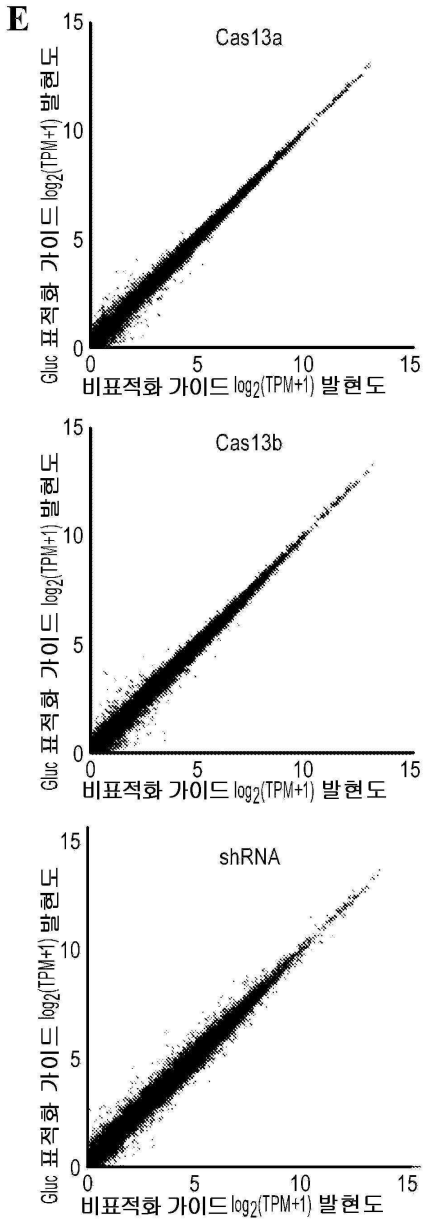


도 21C

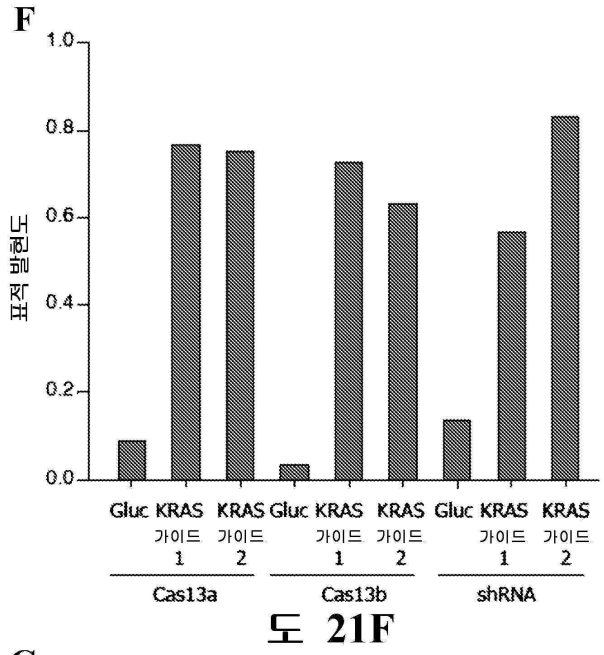


도 21D

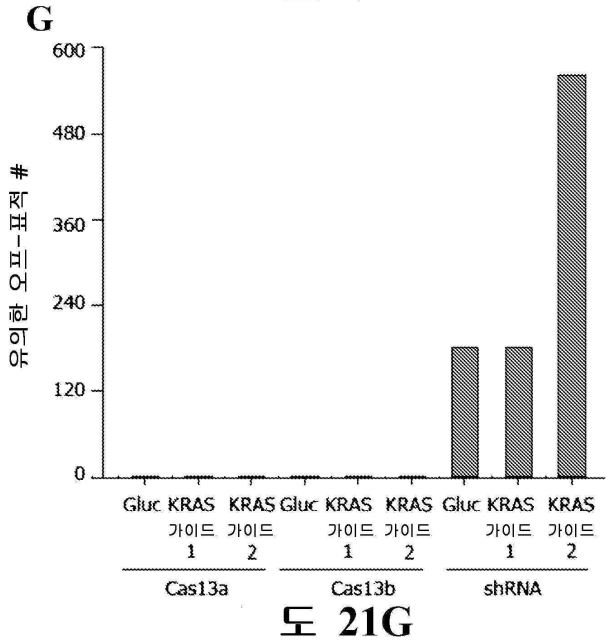
도면21e


E
Cas13a
Cas13b
shRNA
Gluc 표적화 가이드 log2(TPM+1) 발현도
비표적화 가이드 log2(TPM+1) 발현도
0
5
10
15

도면21fg

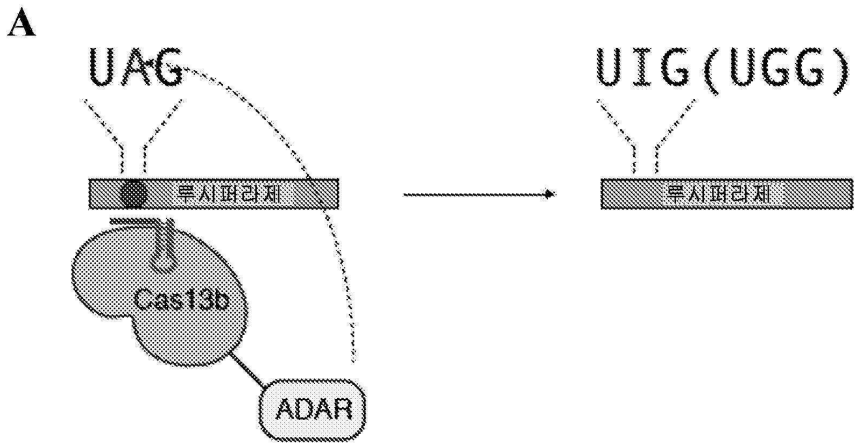


도 21F

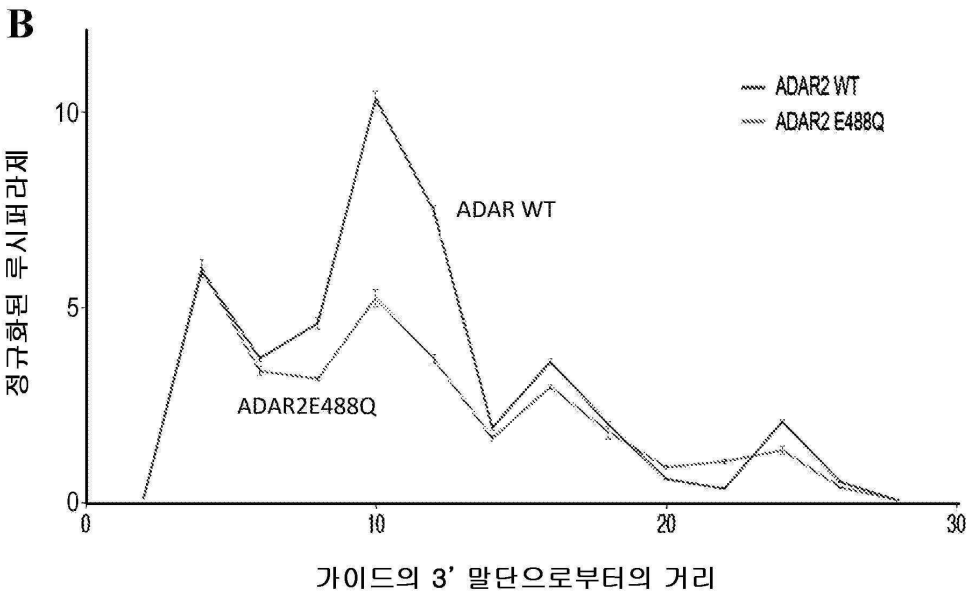


도 21G

도면22ab

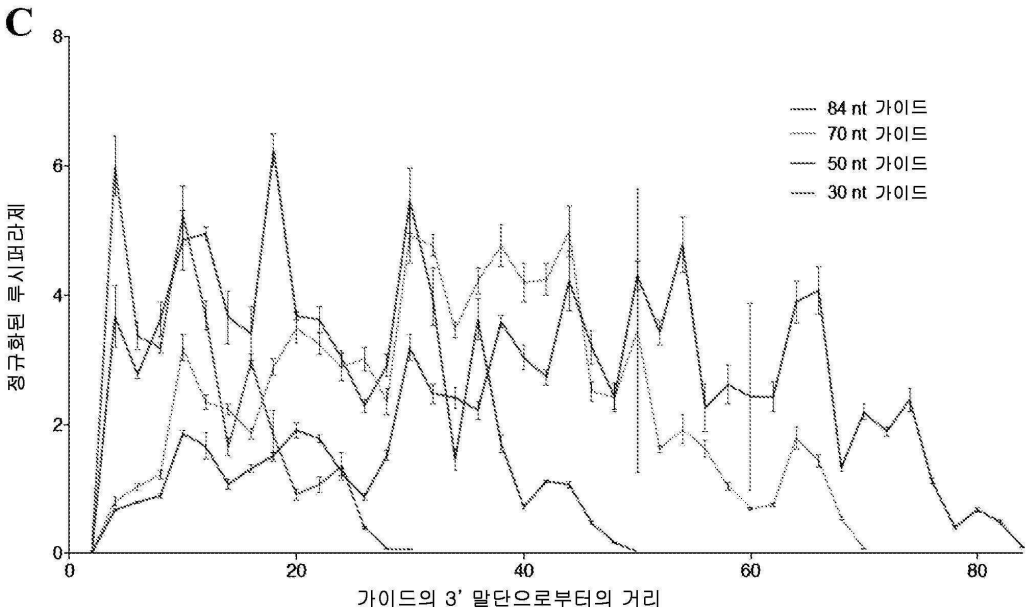


도 22A

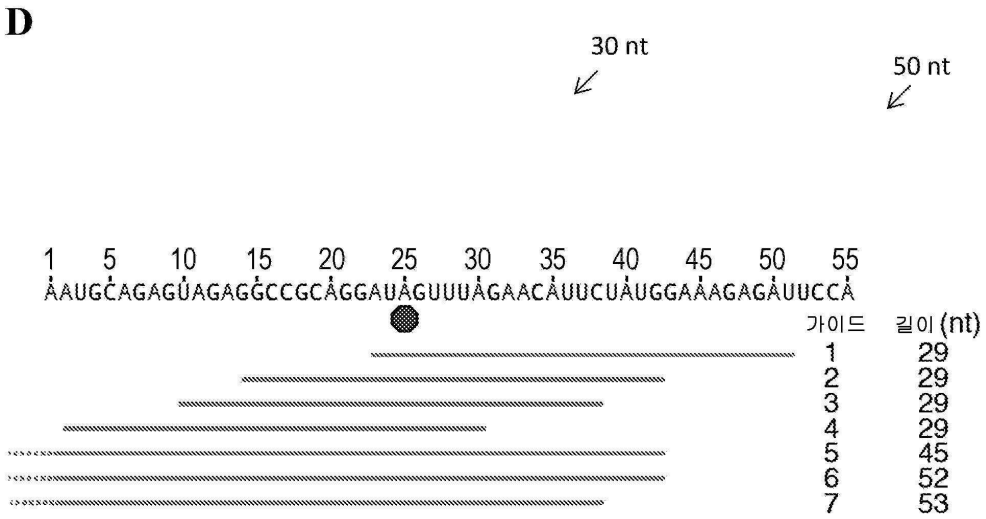


도 22B

도면22cd

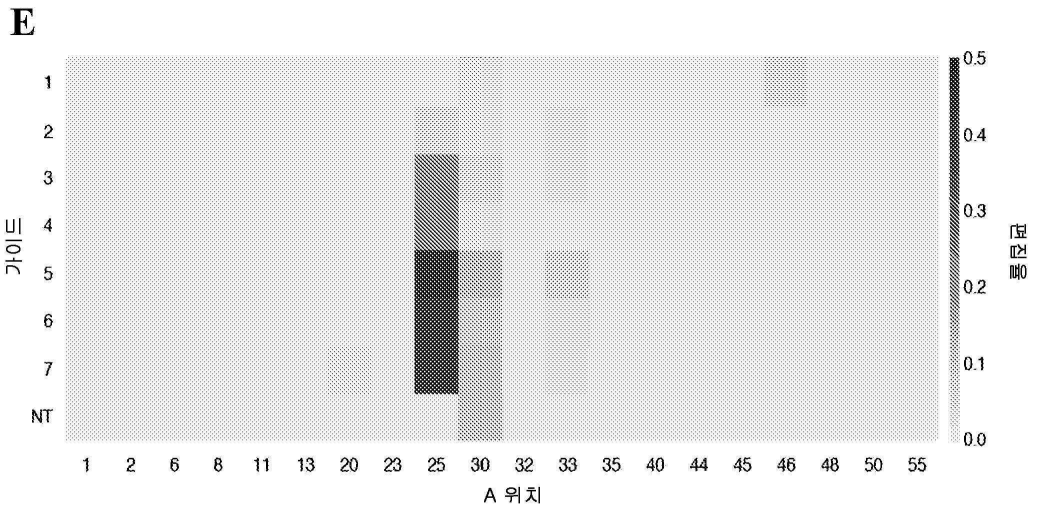


도 22C

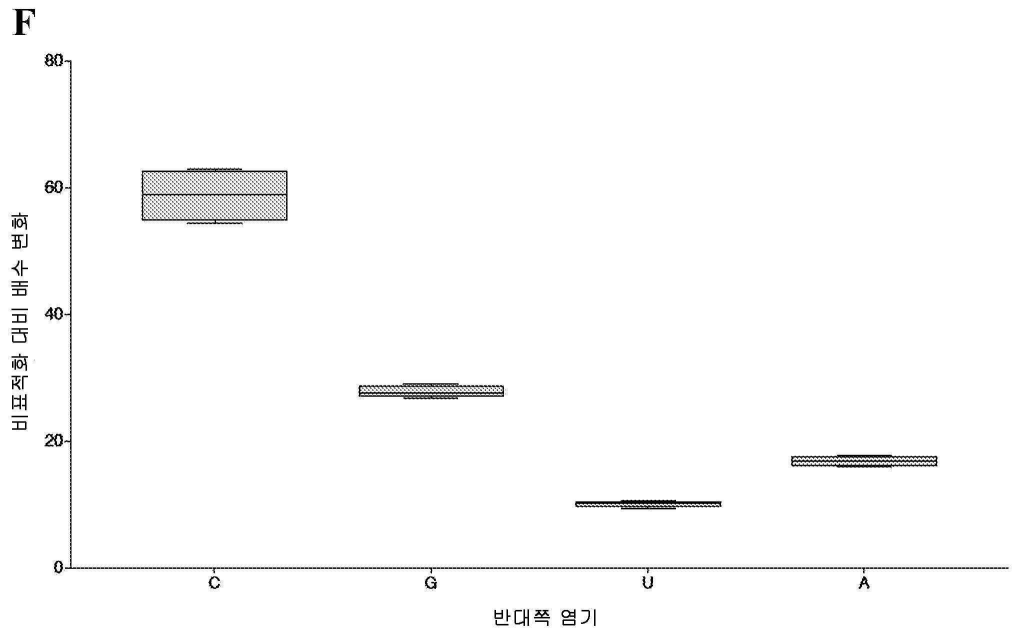


도 22D

도면22ef

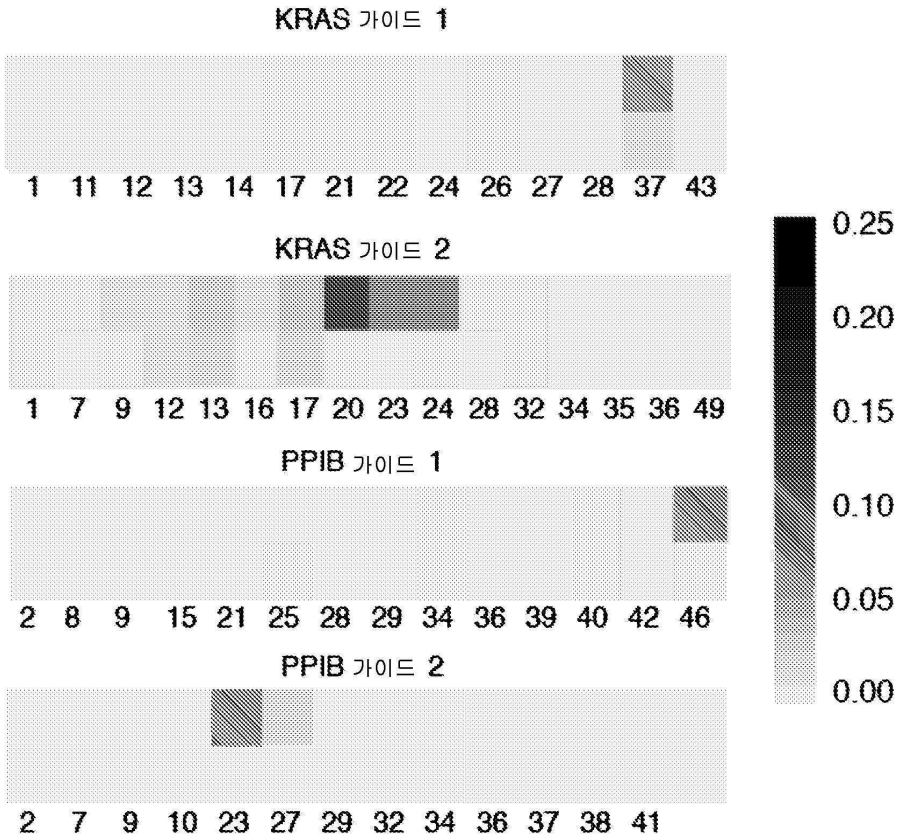


도 22E

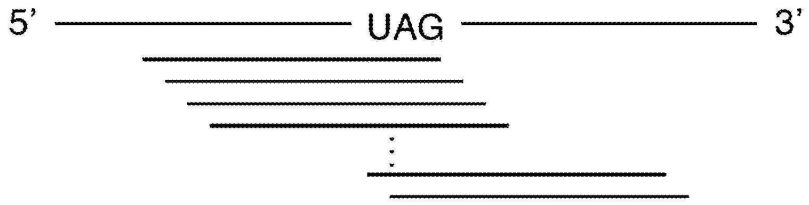


도 22F

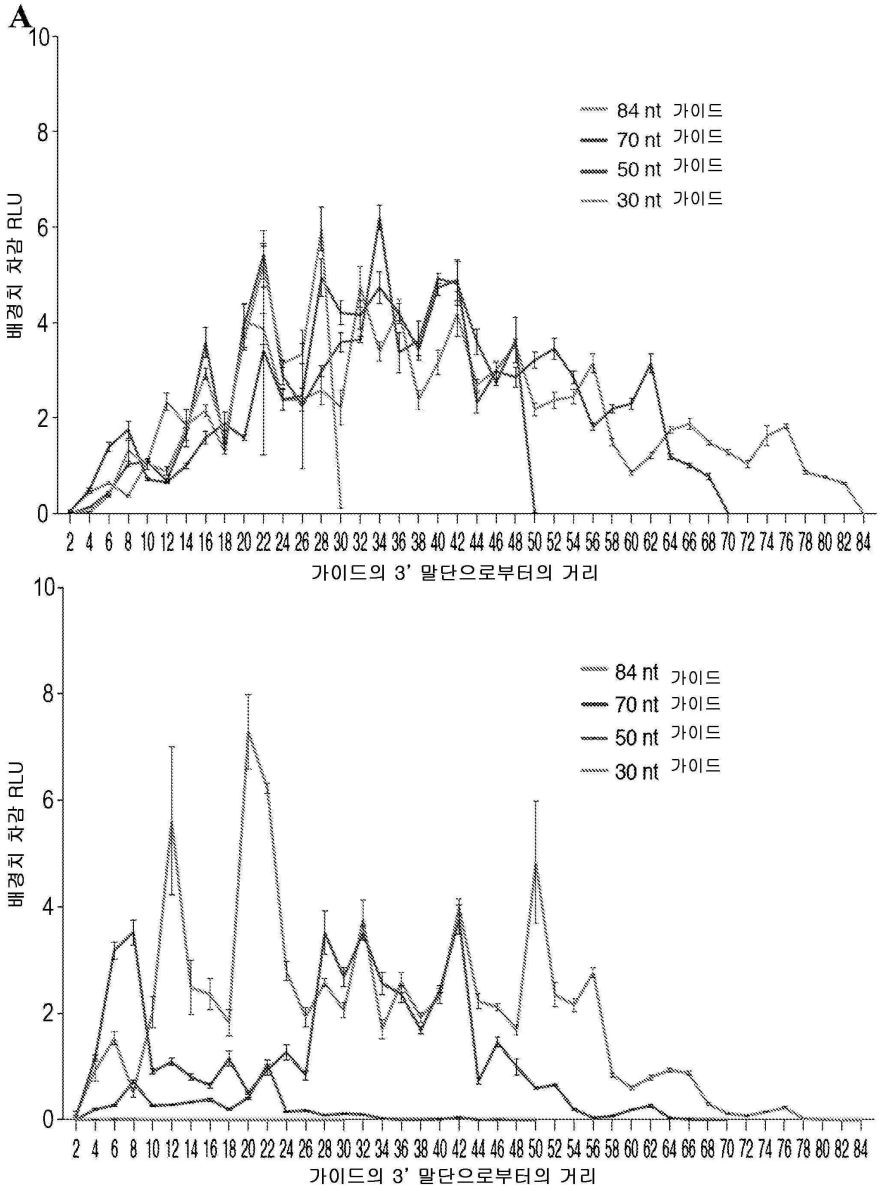
도면23


KRAS 가이드 1
1 11 12 13 14 17 21 22 24 26 27 28 37 43
KRAS 가이드 2
1 7 9 12 13 16 17 20 23 24 28 32 34 35 36 49
PPIB 가이드 1
2 8 9 15 21 25 28 29 34 36 39 40 42 46
PPIB 가이드 2
2 7 9 10 23 27 29 32 34 36 37 38 41
0.25
0.20
0.15
0.10
0.05
0.00

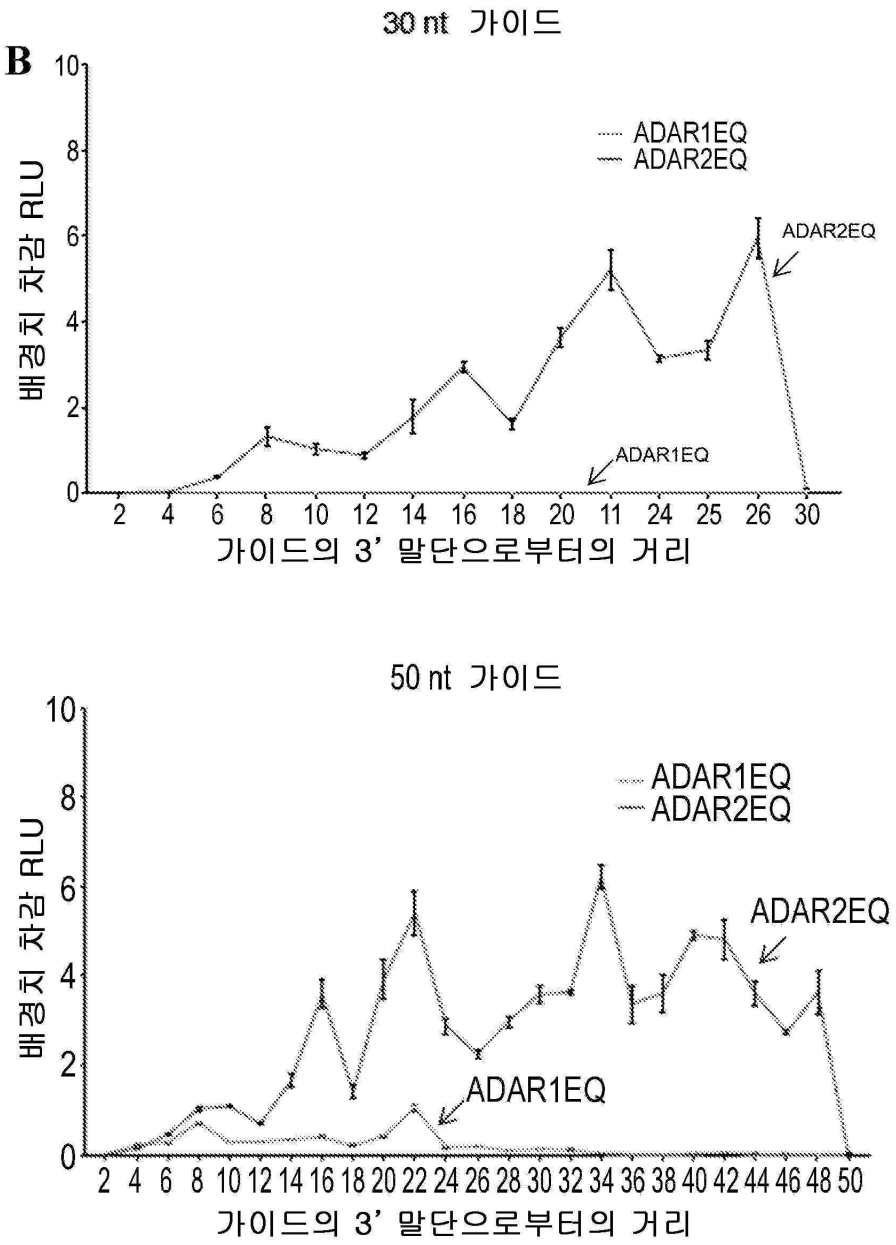
도면24


5'
UAG
3'

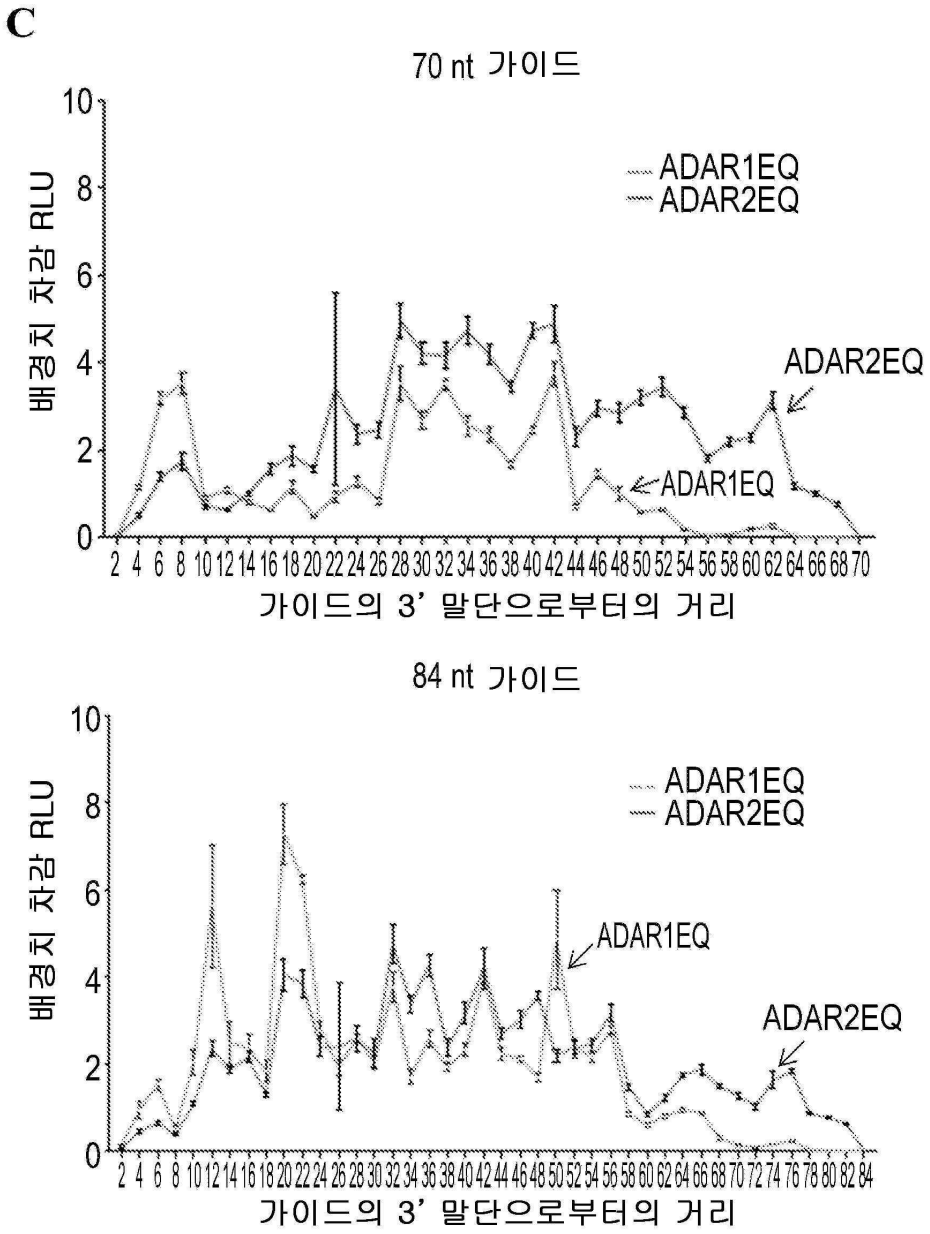
도면25a


A
84 nt 가이드
70 nt 가이드
50 nt 가이드
30 nt 가이드
배경치 차감 RLU
가이드의 3' 말단으로부터의 거리
84 nt 가이드
70 nt 가이드
50 nt 가이드
30 nt 가이드
배경치 차감 RLU
가이드의 3' 말단으로부터의 거리

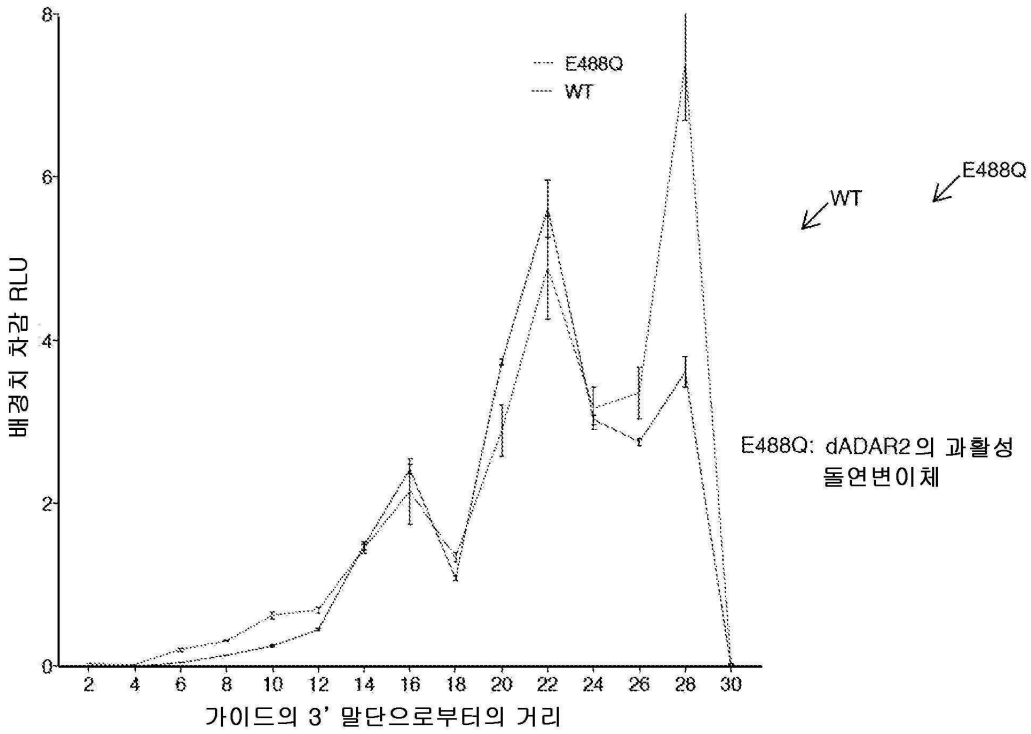
도면25b


B
30 nt 가이드
배경치 차감 RLU
ADAR1EQ
ADAR2EQ
ADAR2EQ
ADAR1EQ
가이드의 3' 말단으로부터의 거리
50 nt 가이드
배경치 차감 RLU
ADAR1EQ
ADAR2EQ
ADAR2EQ
ADAR1EQ
가이드의 3' 말단으로부터의 거리

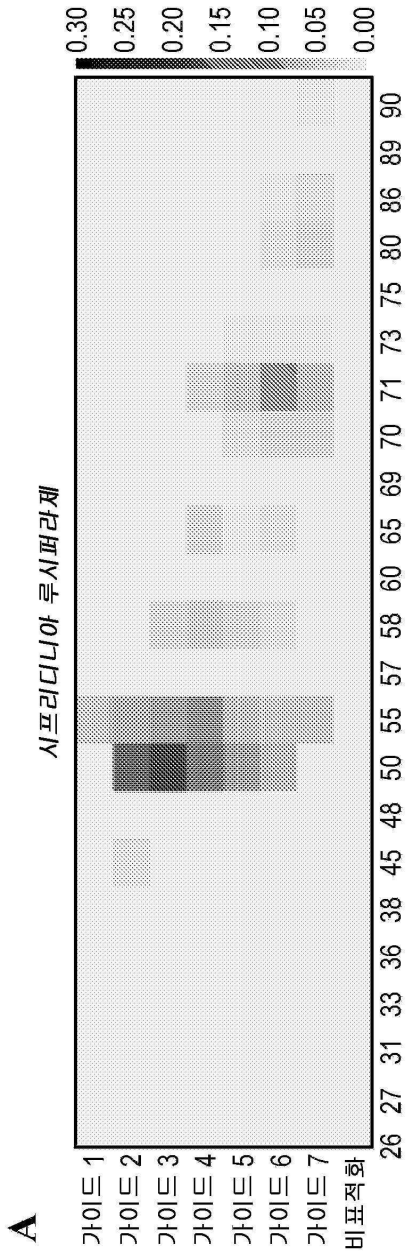
도면25c


C
70 nt 가이드
배경치 차감 RLU
ADAR1EQ
ADAR2EQ
ADAR2EQ
ADAR1EQ
가이드의 3' 말단으로부터의 거리
84 nt 가이드
배경치 차감 RLU
ADAR1EQ
ADAR2EQ
ADAR1EQ
ADAR2EQ
가이드의 3' 말단으로부터의 거리

도면26


E488Q
WT
배경치 차감 RLU
가이드의 3' 말단으로부터의 거리
WT
E488Q
E488Q: dADAR2의 과활성 돌연변이체

도면27a

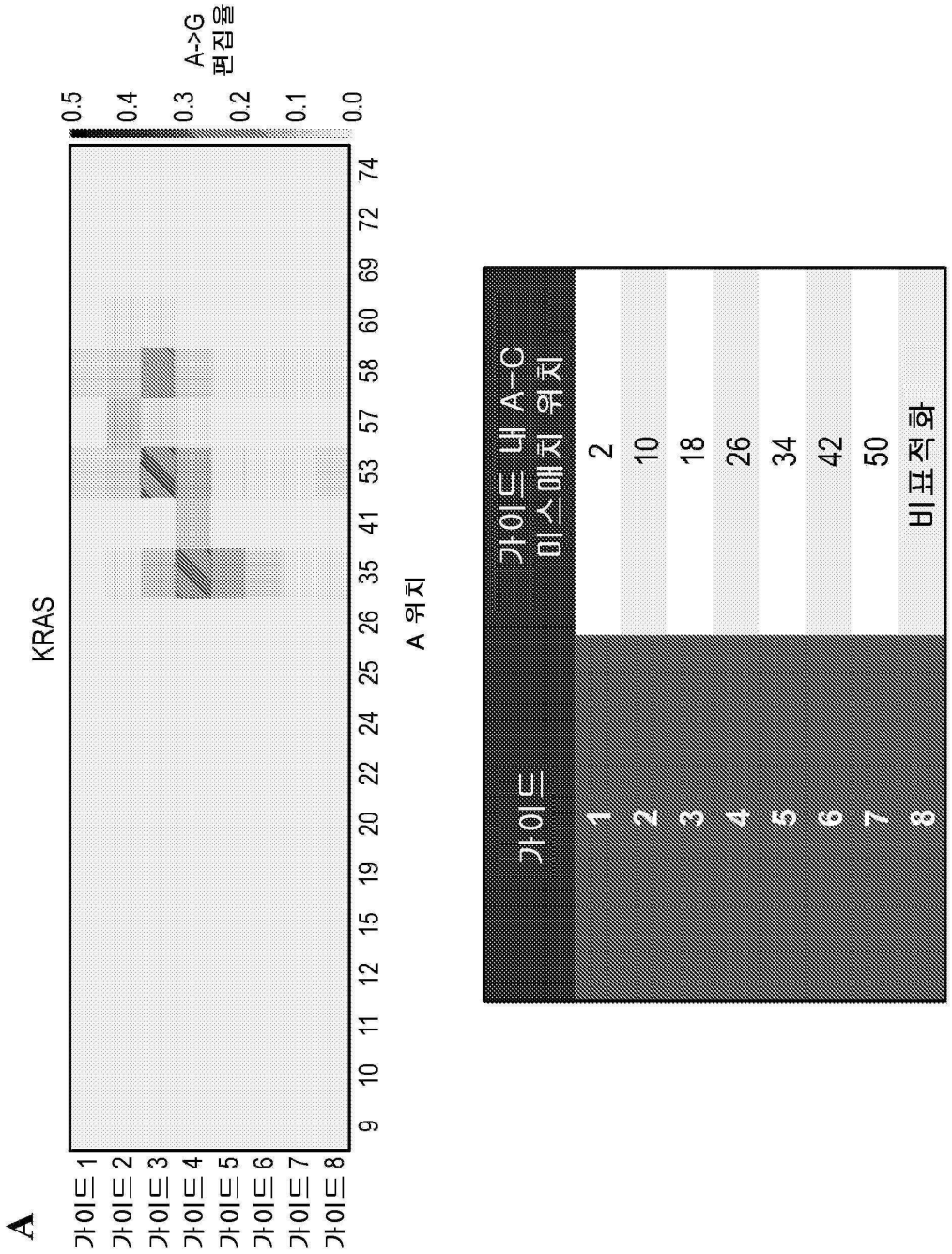

A
시프리디니아 루시퍼라제
0.30
0.25
0.20
0.15
0.10
0.05
0.00
가이드 1
가이드 2
가이드 3
가이드 4
가이드 5
가이드 6
가이드 7
비표적화
26 27 31 33 36 38 45 48 50 55 57 58 60 65 69 70 71 73 75 80 86 89 90

도면27b

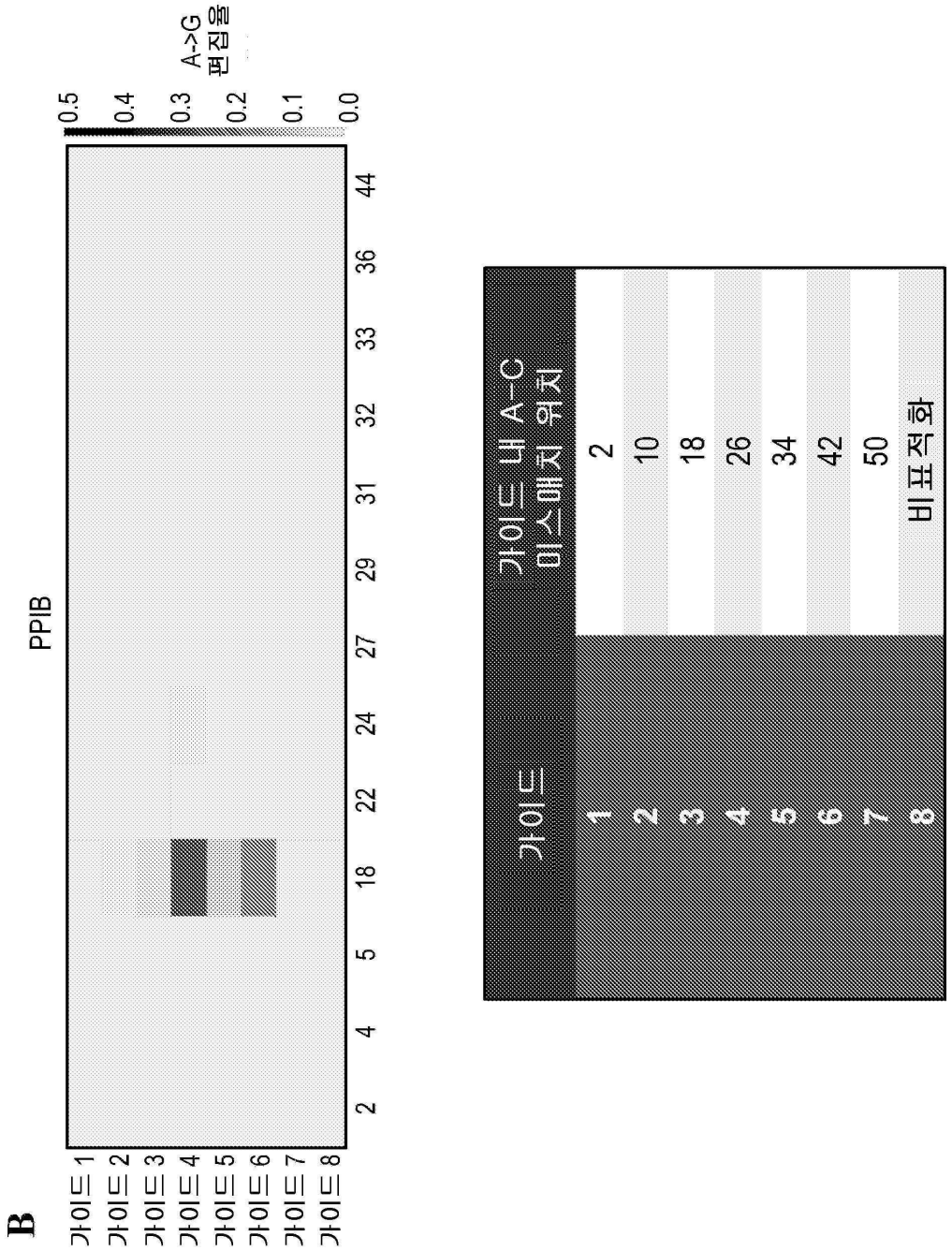
B

가이드	3' 출발 위치	5' 출발 위치	A-C 위치
1	1	50	50
2	9	59	50
3	17	67	50
4	25	75	50
5	33	83	50
6	41	90	50
7	49	99	50
NT			

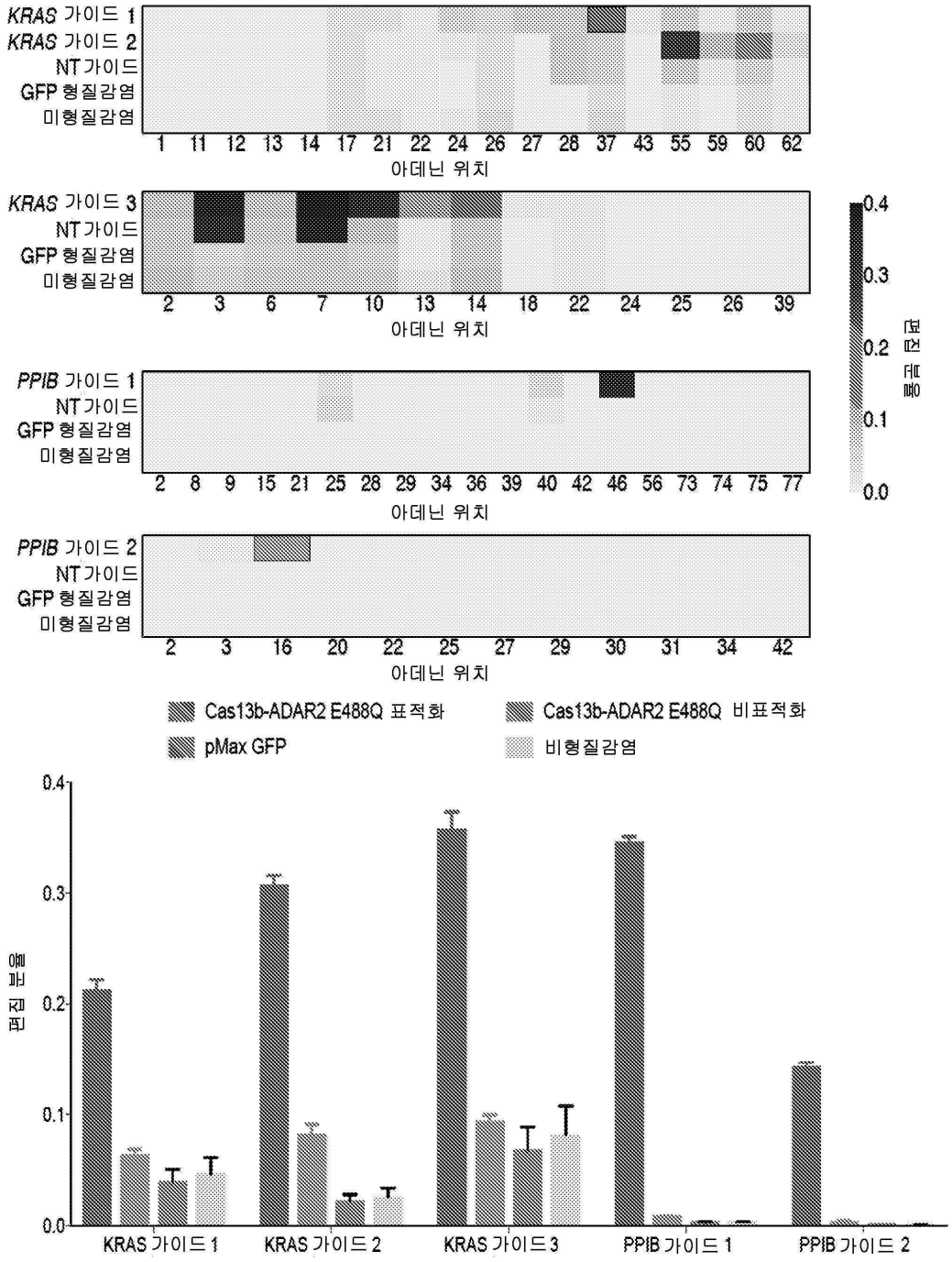
도면28a


A
KRAS
가이드 1
가이드 2
가이드 3
가이드 4
가이드 5
가이드 6
가이드 7
가이드 8
9 10 11 12 15 19 20 22 24 25 26 35 41 53 57 58 60 69 72 74
A 위치
0.5 0.4 0.3 0.2 0.1 0.0
A->G 편집율
가이드
가이드 내 A-C 미스매치 위치
1 2 3 4 5 6 7 8
2 10 18 26 34 42 50
비표적화

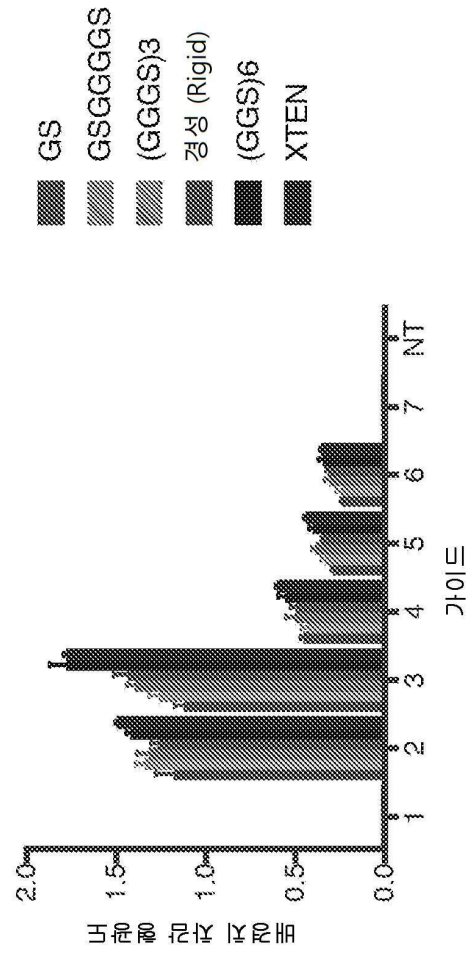
도면28b


B
PPIB
A->G 편집율
0.5
0.4
0.3
0.2
0.1
0.0
가이드 1
가이드 2
가이드 3
가이드 4
가이드 5
가이드 6
가이드 7
가이드 8
2
4
5
18
22
24
27
29
31
32
33
36
44
가이드
가이드 내 A-C 미스매치 위치
1
2
3
4
5
6
7
8
2
10
18
26
34
42
50
비표적화

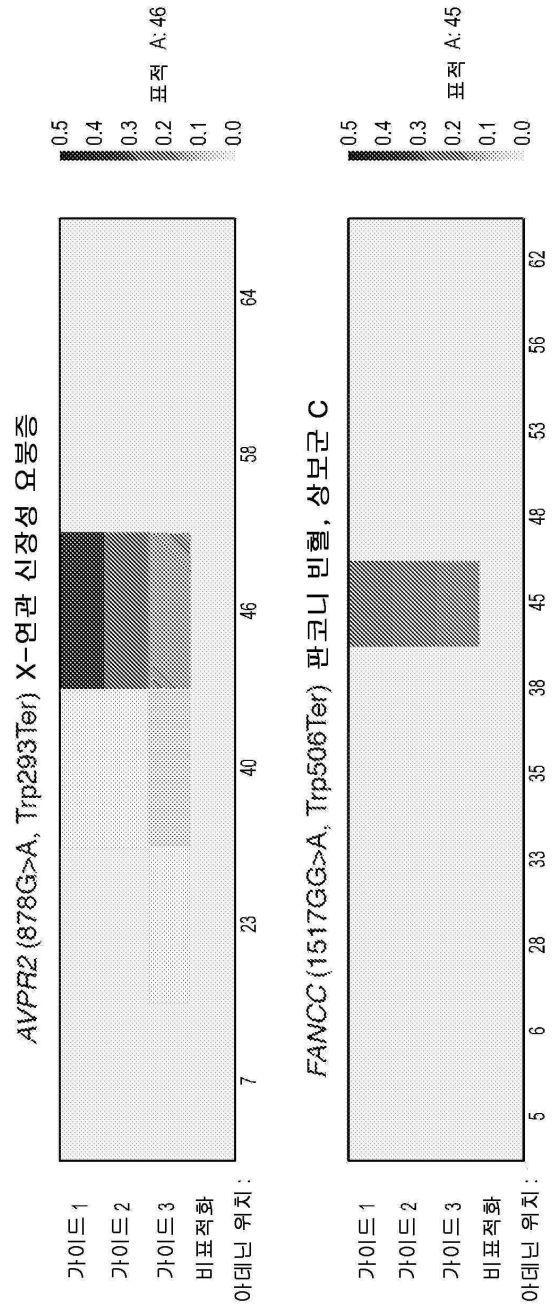
도면29


KRAS 가이드 1
KRAS 가이드 2
NT 가이드
GFP 형질감염
미형질감염
1 11 12 13 14 17 21 22 24 26 27 28 37 43 55 59 60 62
아데닌 위치
KRAS 가이드 3
NT 가이드
GFP 형질감염
미형질감염
2 3 6 7 10 13 14 18 22 24 25 26 39
아데닌 위치
PPIB 가이드 1
NT 가이드
GFP 형질감염
미형질감염
2 8 9 15 21 25 28 29 34 36 39 40 42 46 56 73 74 75 77
아데닌 위치
PPIB 가이드 2
NT 가이드
GFP 형질감염
미형질감염
2 3 16 20 22 25 27 29 30 31 34 42
아데닌 위치
0.4
0.3
0.2
0.1
0.0
편집 퍼센트
Cas13b-ADAR2 E488Q 표적화
Cas13b-ADAR2 E488Q 비표적화
pMax GFP
비형질감염
0.4
0.3
0.2
0.1
0.0
편집 퍼센트
KRAS 가이드 1
KRAS 가이드 2
KRAS 가이드 3
PPIB 가이드 1
PPIB 가이드 2

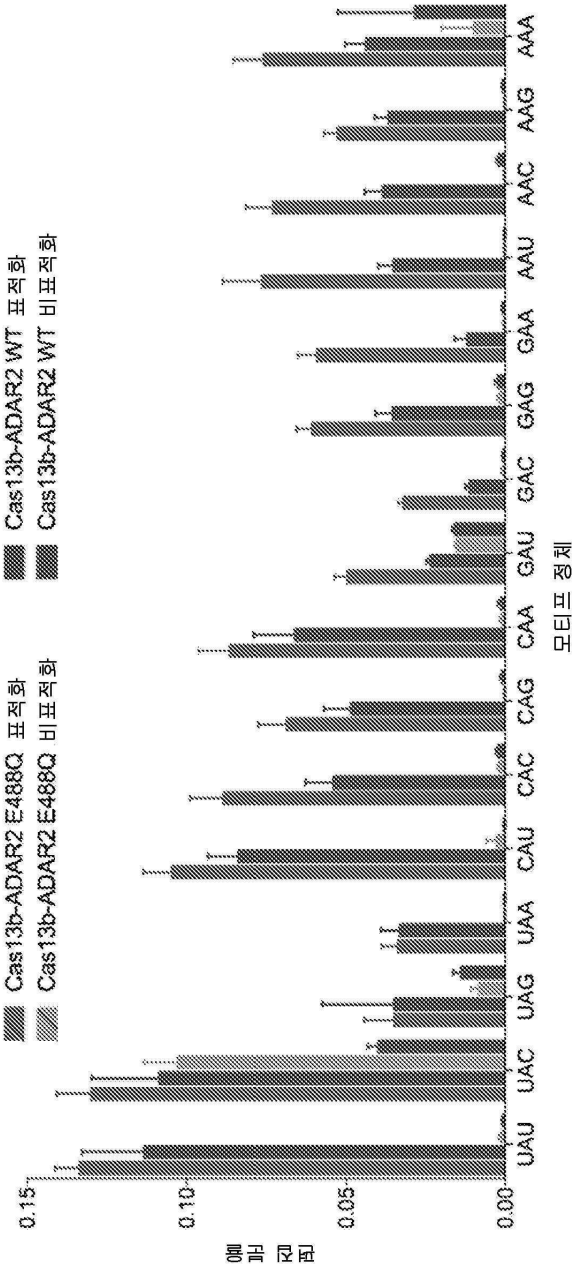
도면30


GS
GSGGGGS
(GGGS)3
경성 (Rigid)
(GGS)6
XTEN
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
NT
가이드
배경치 차감 형광도

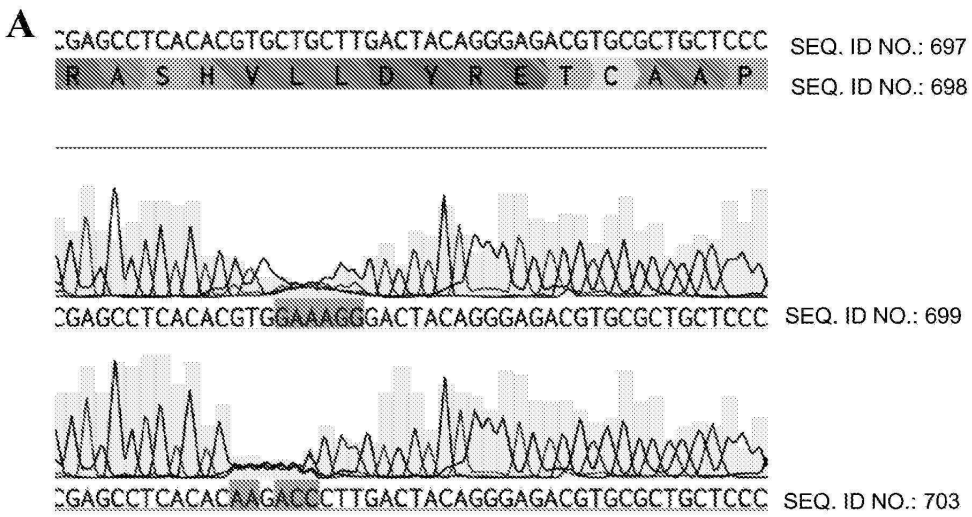
도면31


AVPR2 (878G>A, Trp293Ter) X-연관 신장성 요붕증
표적 A: 46
0.5
0.4
0.3
0.2
0.1
0.0
가이드 1
가이드 2
가이드 3
비표적화
아데닌 위치 :
7
23
40
46
58
64
FANCC (1517GG>A, Trp506Ter) 판코니 빈혈, 상보군 C
표적 A: 45
0.5
0.4
0.3
0.2
0.1
0.0
가이드 1
가이드 2
가이드 3
비표적화
아데닌 위치 :
5
6
28
33
35
38
45
48
53
56
62

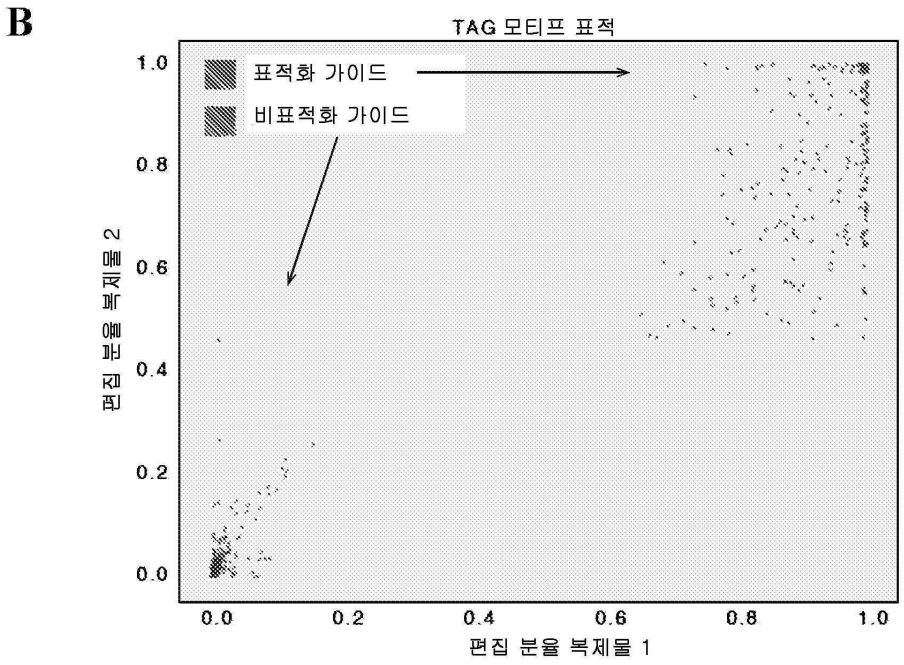
도면32


Cas13b-ADAR2 E488Q 표적화
Cas13b-ADAR2 E488Q 비표적화
Cas13b-ADAR2 WT 표적화
Cas13b-ADAR2 WT 비표적화
0.15
0.10
0.05
0.00
편집 분율
UAU
UAC
UAG
UAA
CAU
CAC
CAG
CAA
GAU
GAC
GAG
GAA
AAU
AAC
AAG
AAA
모티프 정체

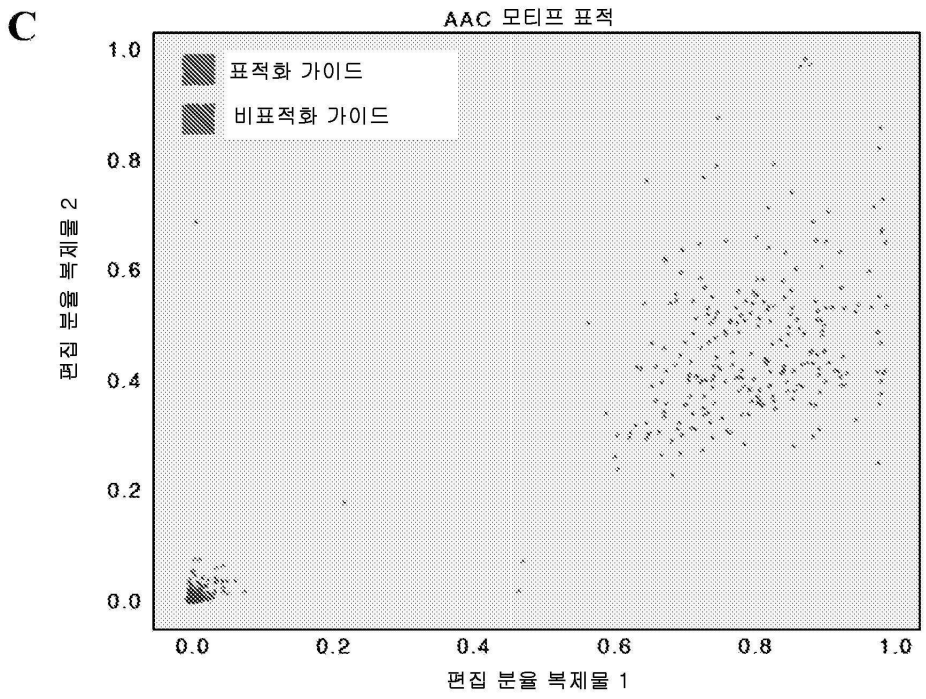
도면33a


A
CGAGCCTCACACGTGCTGCTTGACTACAGGGAGACGTGCGCTGCTCCC
R A S H V L L D Y R E T C A A P
SEQ. ID NO.: 697
SEQ. ID NO.: 698
CGAGCCTCACACGTGGAAAGGGACTACAGGGAGACGTGCGCTGCTCCC
SEQ. ID NO.: 699
CGAGCCTCACACAAGACCCTTGACTACAGGGAGACGTGCGCTGCTCCC
SEQ. ID NO.: 703

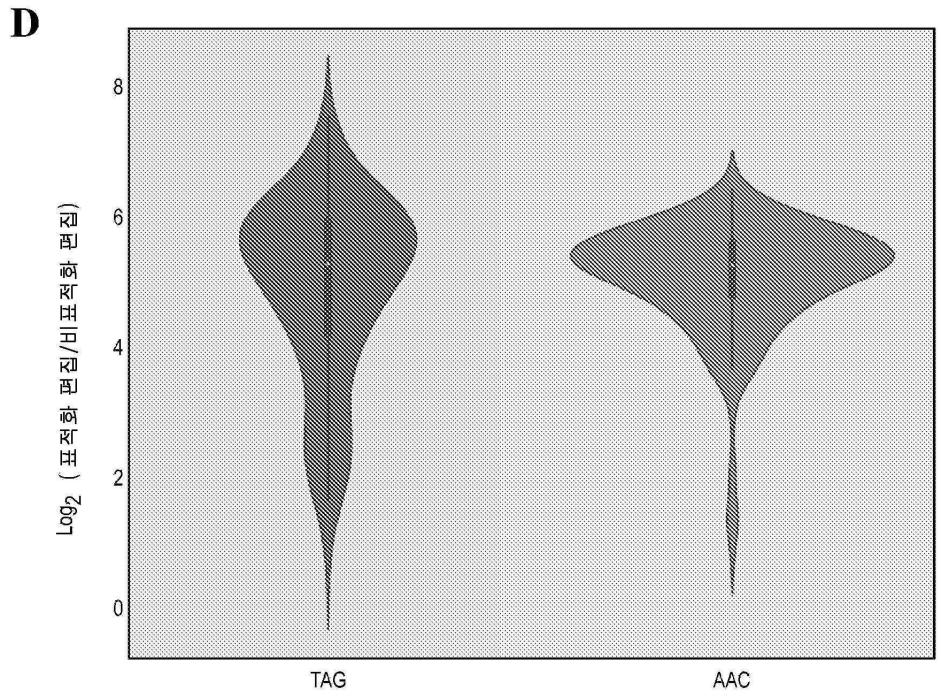
도면33b


B
TAG 모티프 표적
표적화 가이드
비표적화 가이드
편집 분율 복제물 2
편집 분율 복제물 1
0.0
0.2
0.4
0.6
0.8
1.0

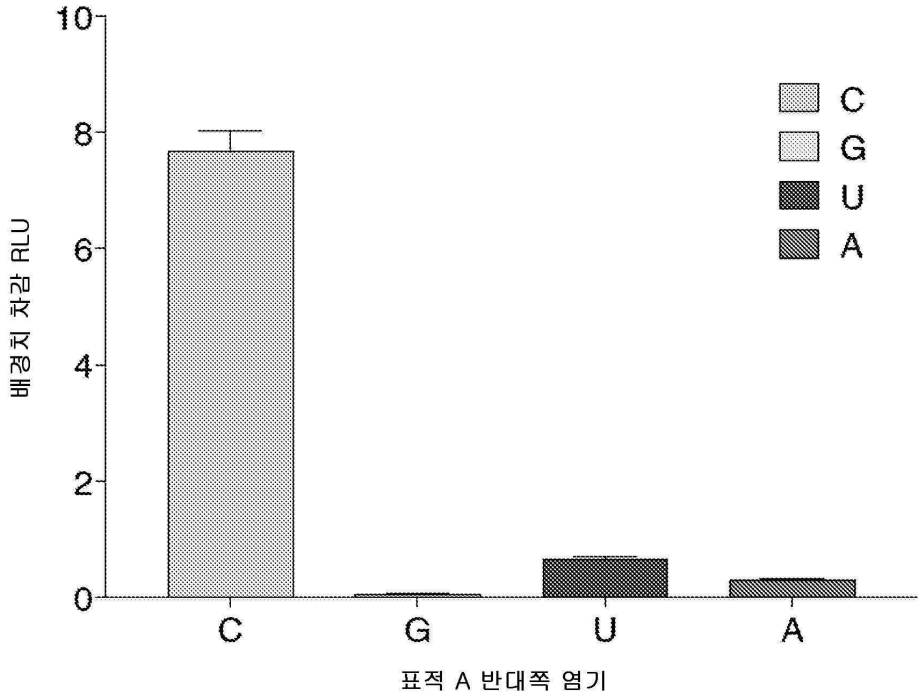
도면33c


C
AAC 모티프 표적
표적화 가이드
비표적화 가이드
편집 분율 복제물 2
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
편집 분율 복제물 1

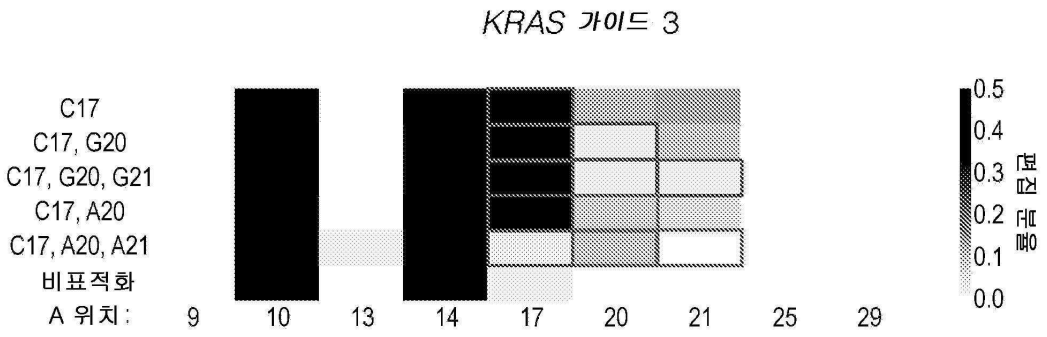
도면33d


D
Log2 (표적화 편집/비표적화 편집)
8
6
4
2
0
TAG
AAC

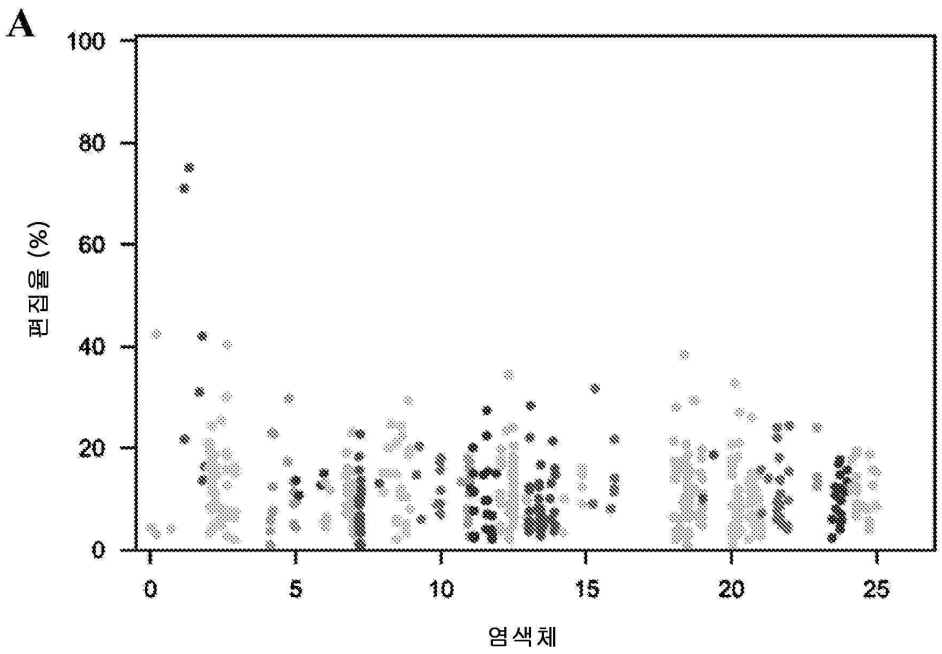
도면34


10
8
6
4
2
0
배경치 차감 RLU
C
G
U
A
C
G
U
A
표적 A 반대쪽 염기

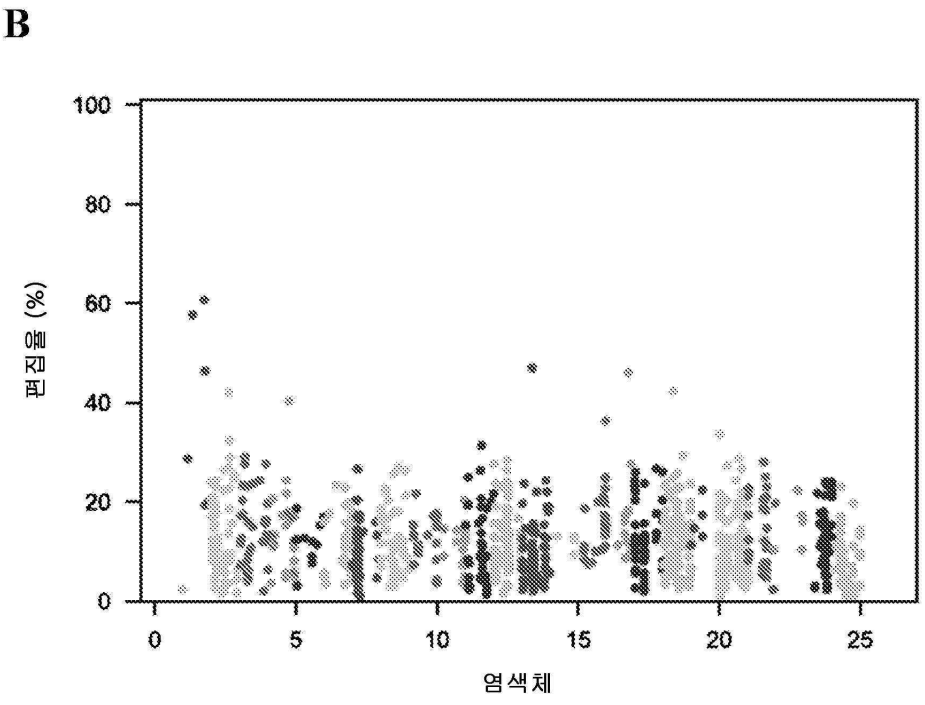
도면35


KRAS 가이드 3
C17
C17, G20
C17, G20, G21
C17, A20
C17, A20, A21
비표적화
A 위치 :
9
10
13
14
17
20
21
25
29
0.5
0.4
0.3
0.2
0.1
0.0
편집 빈도

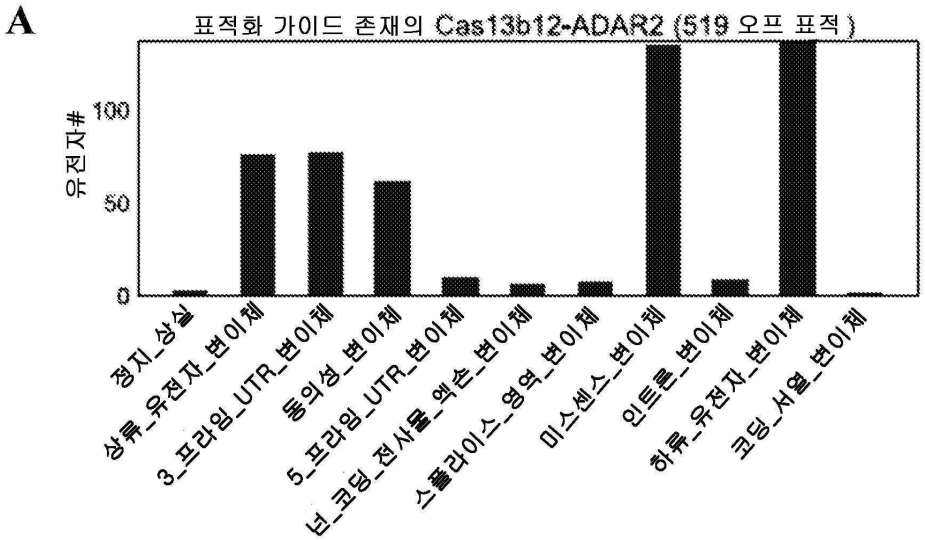
도면36a

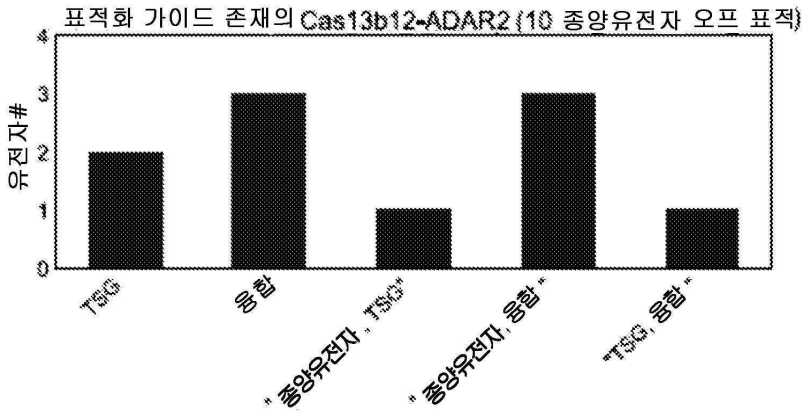

A
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

도면36b

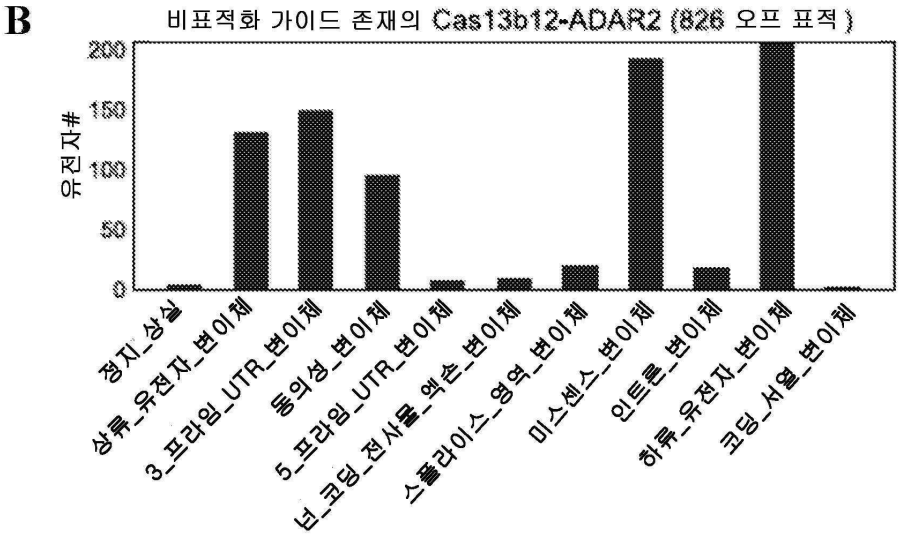

B
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

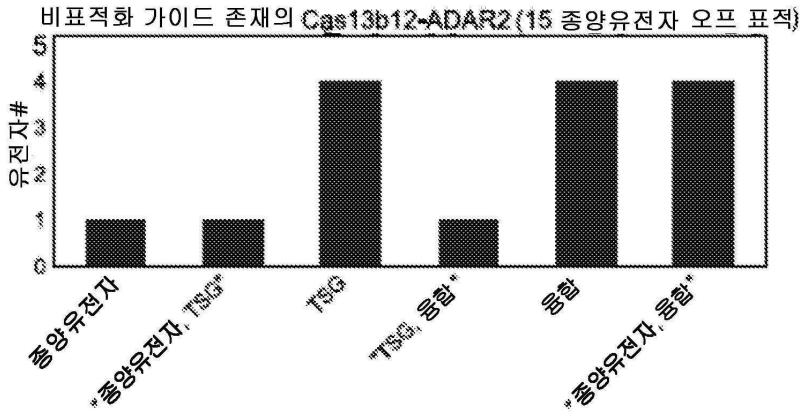
도면37a


A
표적화 가이드 존재의 Cas13b12-ADAR2 (519 오프 표적)
유전자#
100
50
0
정지_상실
상류_유전자_변이체
3_프라임_UTR_변이체
동의성_변이체
5_프라임_UTR_변이체
넌_코딩_전사물_엑손_변이체
스플라이스_영역_변이체
미스센스_변이체
인트론_변이체
하류_유전자_변이체
코딩_서열_변이체

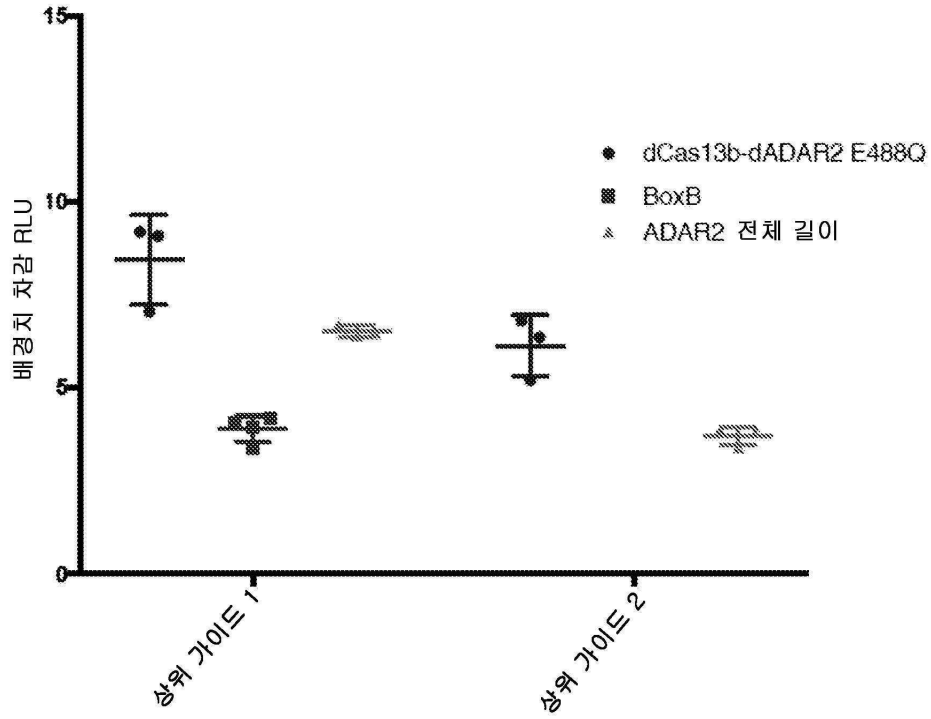

표적화 가이드 존재의 Cas13b12-ADAR2 (10 종양유전자 오프 표적)
유전자#
4
3
2
1
0
TSG
융합
"종양유전자, TSG"
"종양유전자, 융합"
"TSG, 융합"

도면37b

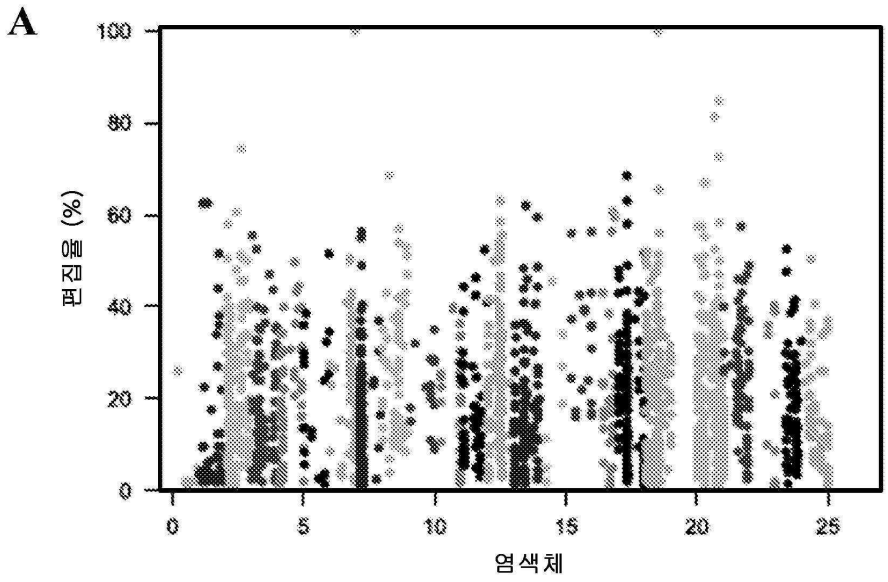

B
비표적화 가이드 존재의 Cas13b12-ADAR2 (826 오프 표적)
유전자#
200
150
100
50
0
정지_상실
상류_유전자_변이체
3_프라임_UTR_변이체
동의성_변이체
5_프라임_UTR_변이체
넌_코딩_전사물_엑손_변이체
스플라이스_영역_변이체
미스센스_변이체
인트론_변이체
하류_유전자_변이체
코딩_서열_변이체


비표적화 가이드 존재의 Cas13b12-ADAR2 (15 종양유전자 오프 표적)
유전자#
5
4
3
2
1
0
종양유전자
종양유전자, TSG
TSG
TSG, 융합
융합
종양유전자, 융합

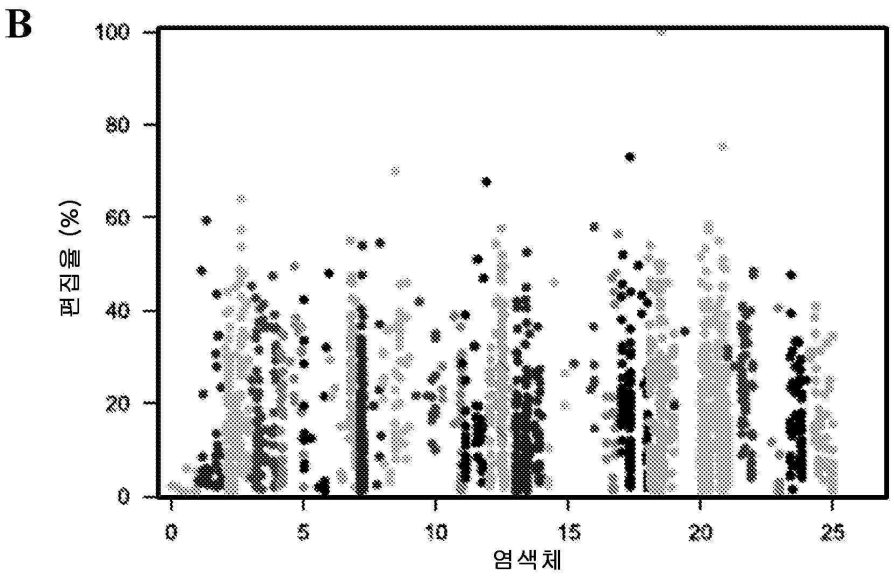
도면38


15
10
5
0
배경치 차감 RLU
dCas13b-dADAR2 E488Q
BoxB
ADAR2 전체 길이
상위 가이드 1
상위 가이드 2

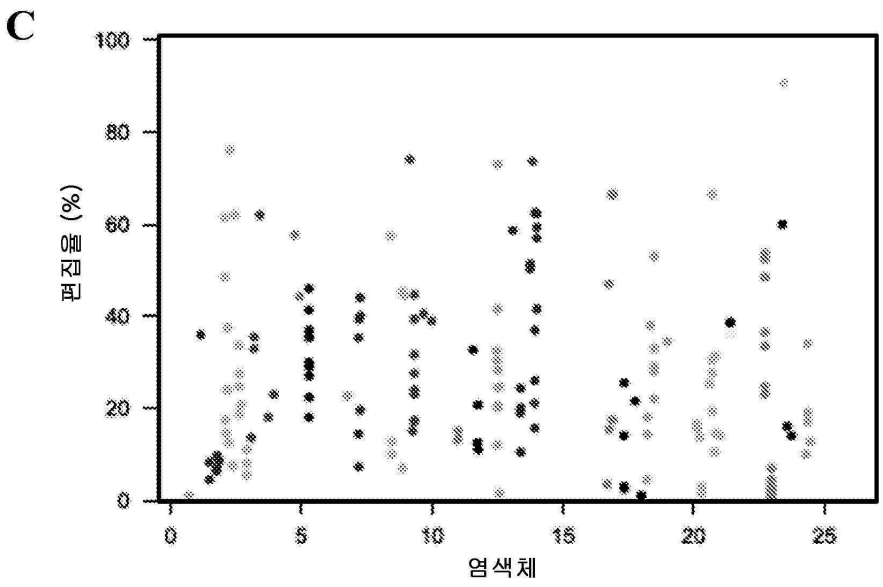
도면39a


A
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

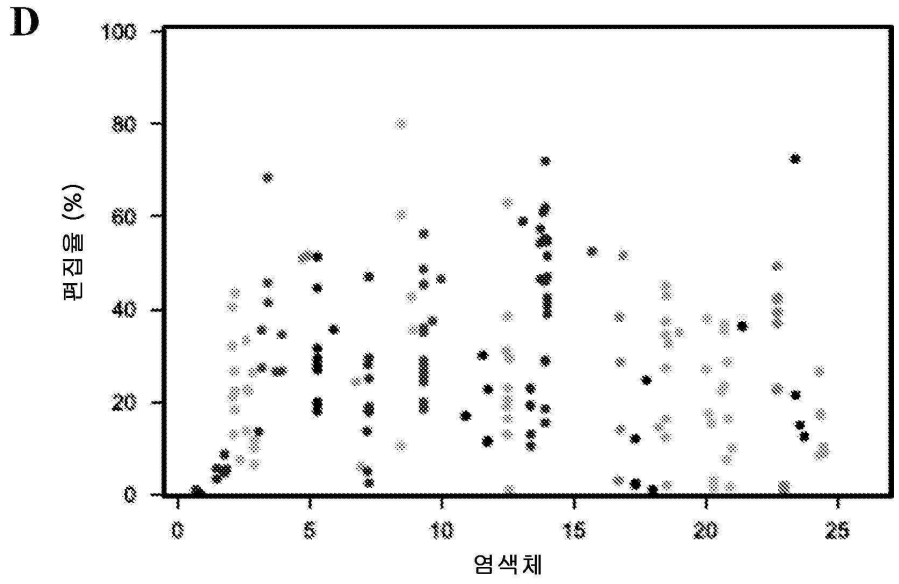
도면39b


B
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

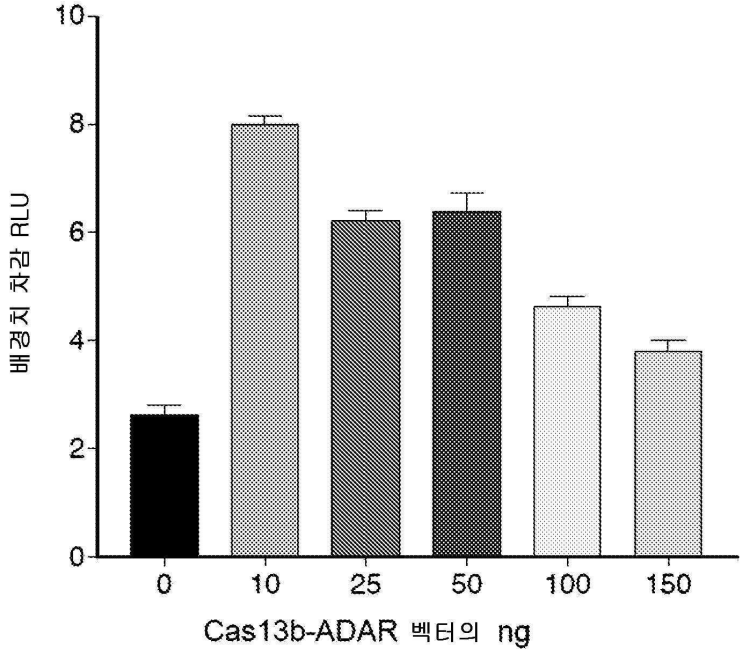
도면39c


C
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

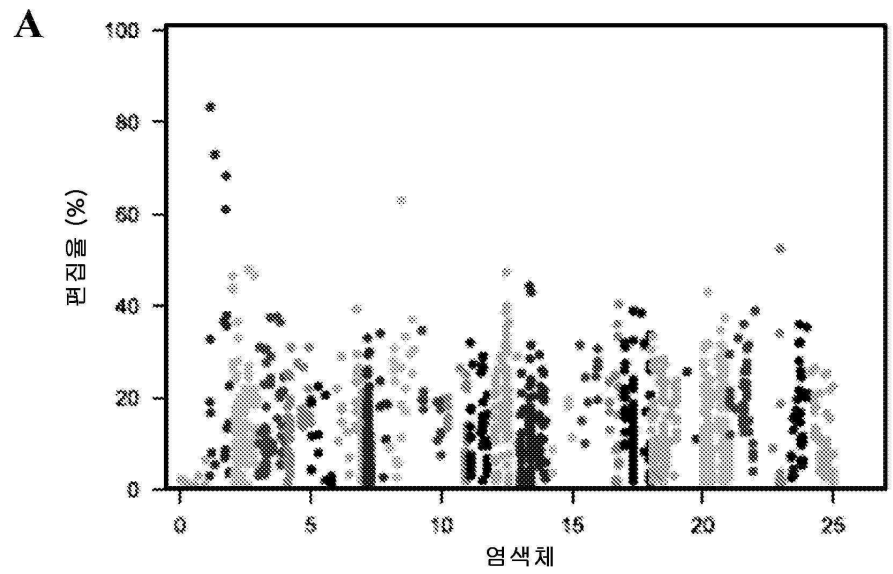
도면39d


D
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

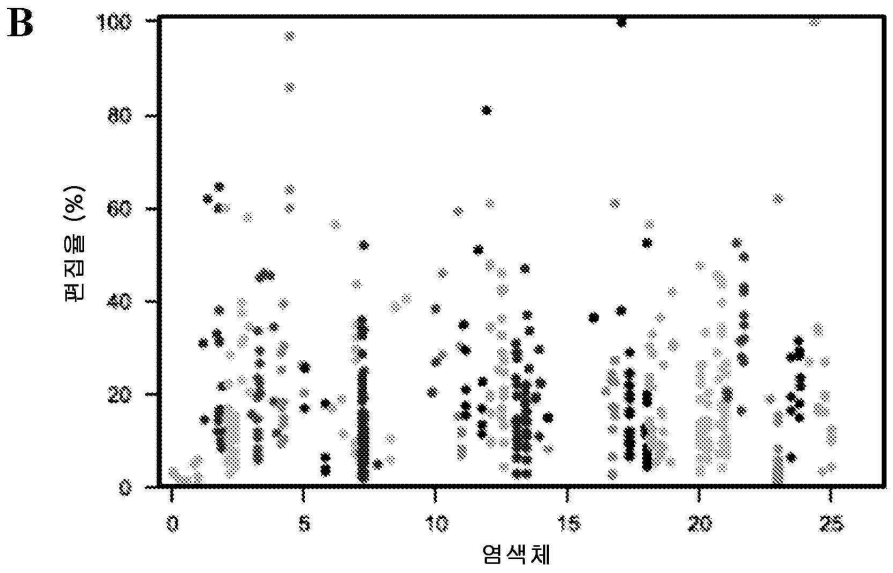
도면40


10
8
6
4
2
0
배경치 차감 RLU
0
10
25
50
100
150
Cas13b-ADAR 벡터의 ng

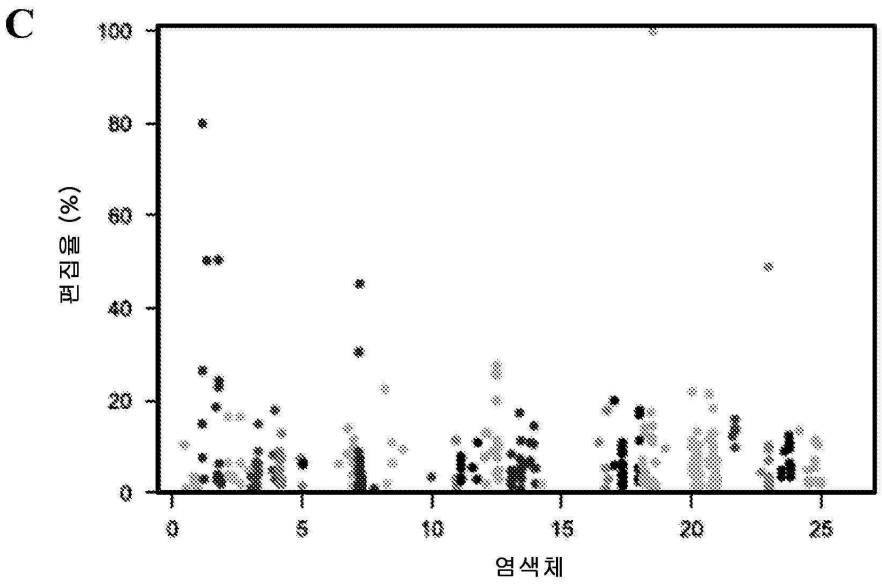
도면41a


A
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

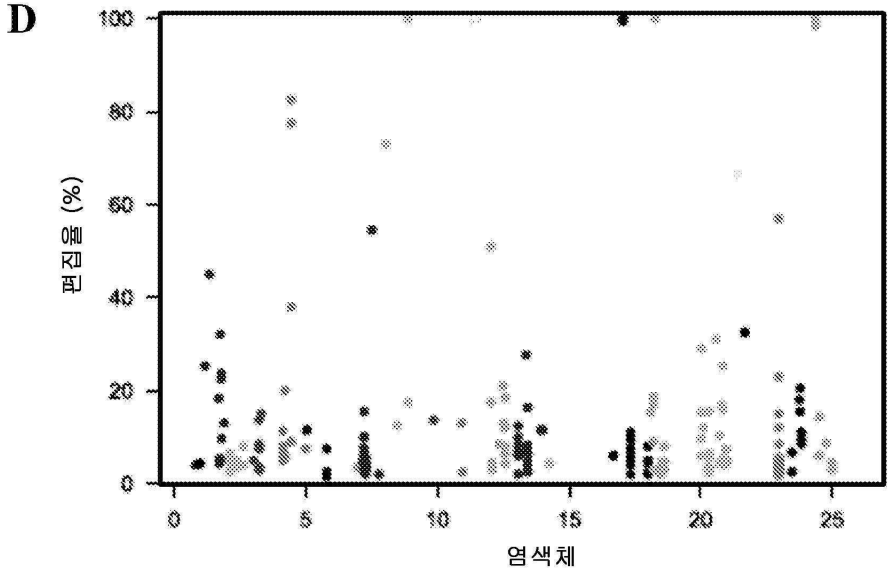
도면41b


B
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

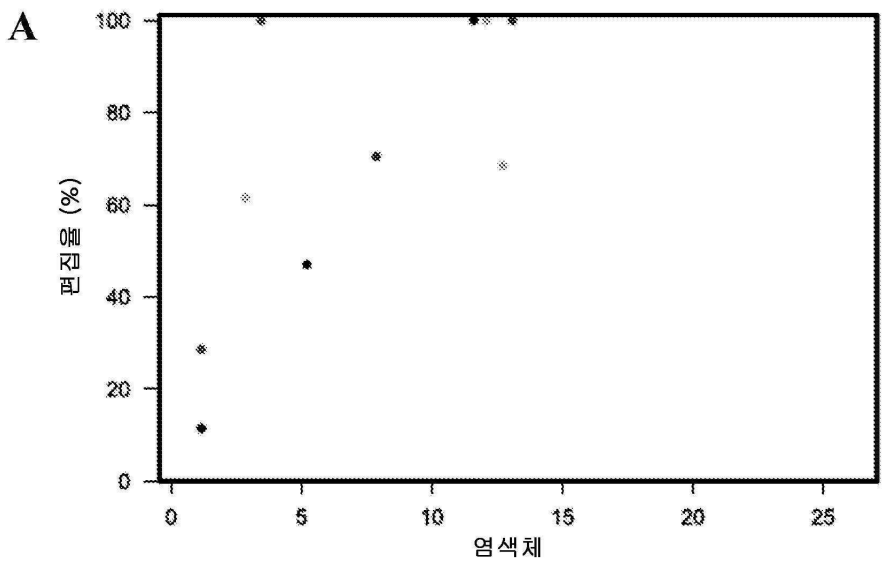
도면41c


C
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

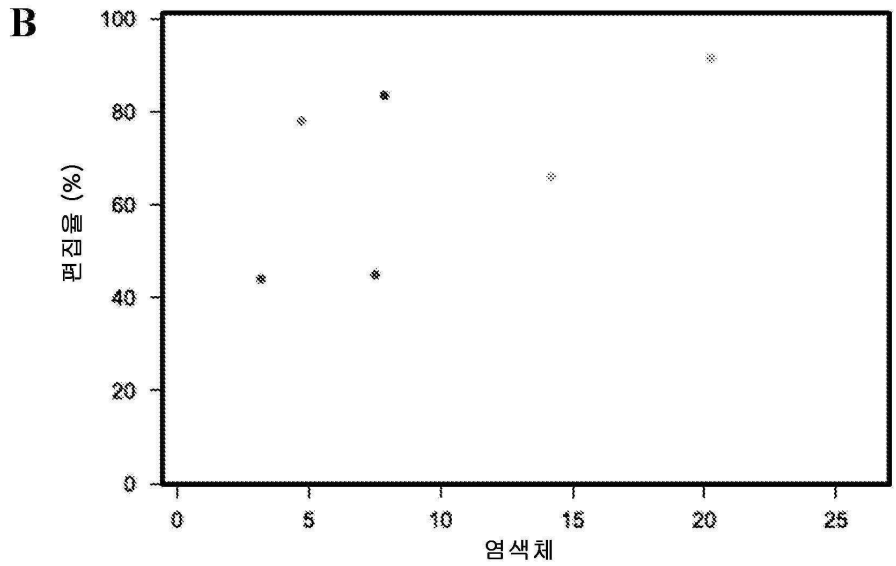
도면41d


D
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

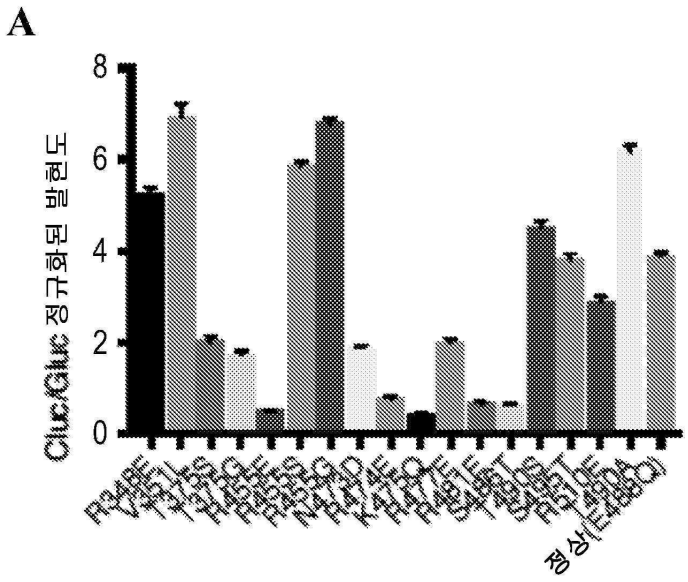
도면42a


A
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

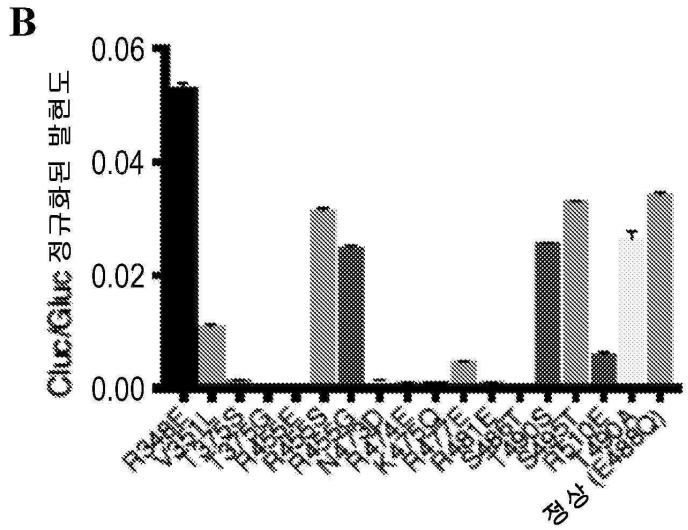
도면42b


B
100
80
60
40
20
0
편집율 (%)
0
5
10
15
20
25
염색체

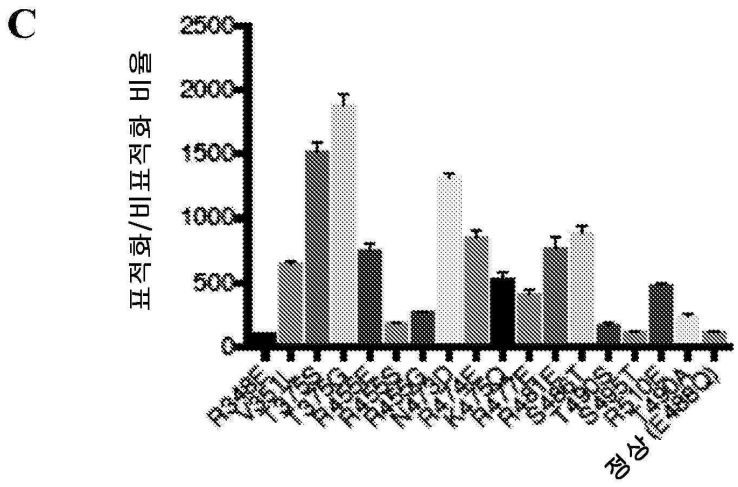
도면43a


A
8
6
4
2
0
Cluc/Gluc 정규화된 발현도
정상(E4380)

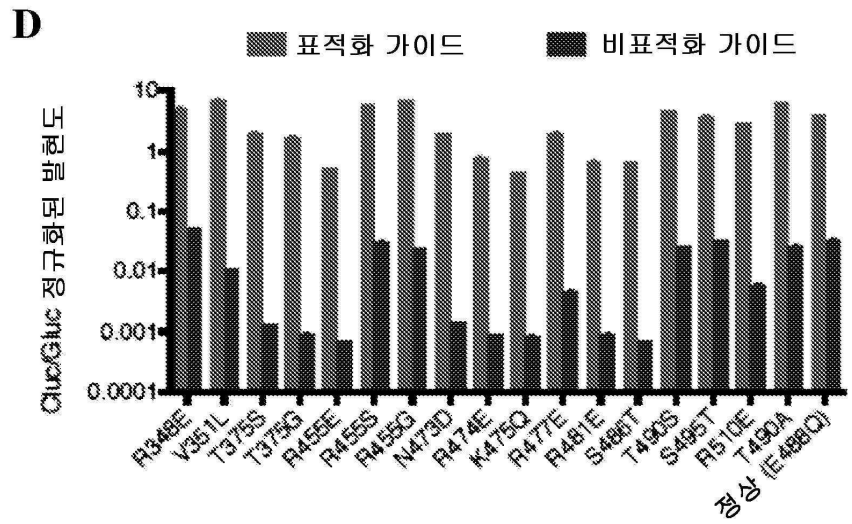
도면43b


B
0.06
0.04
0.02
0.00
Cluc/Gluc 정규화된 발현도
정상 (E4380)

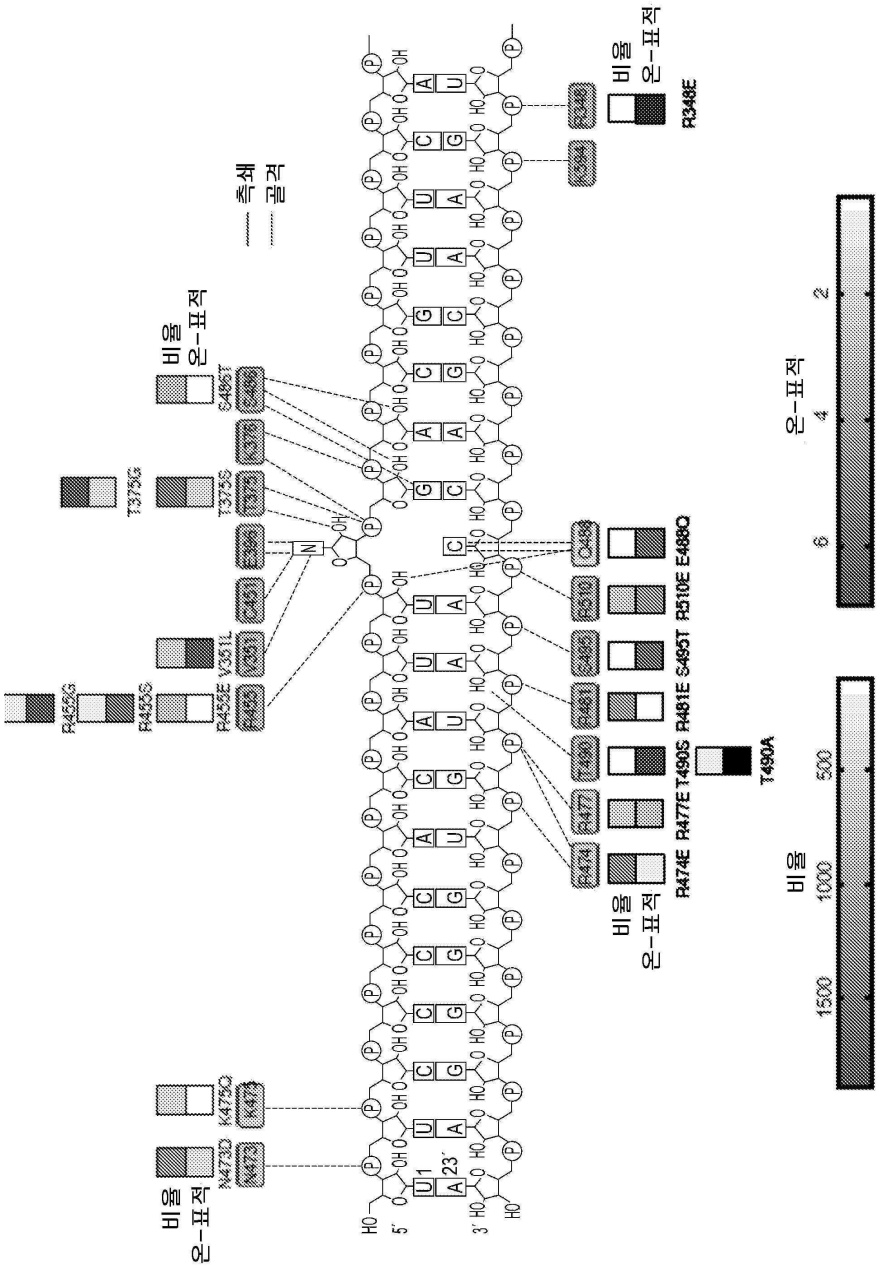
도면43c


C
표적화/비표적화 비율
2500
2000
1500
1000
500
0
정상 (E488Q)

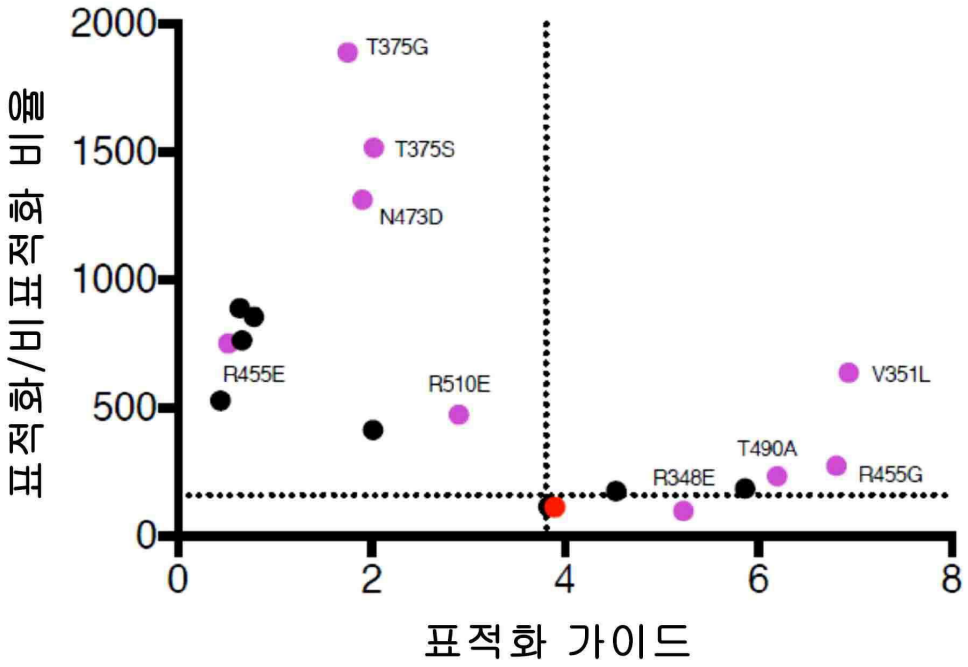
도면43d


D
표적화 가이드
비표적화 가이드
GucGluc 정규화된 발현도
10
1
0.1
0.01
0.001
0.0001
R348E
V351L
T375S
T375G
R455E
R455S
R455G
N473D
R474E
K475Q
R477E
R481E
S486T
T490S
S495T
R510E
T490A
정상 (E488Q)

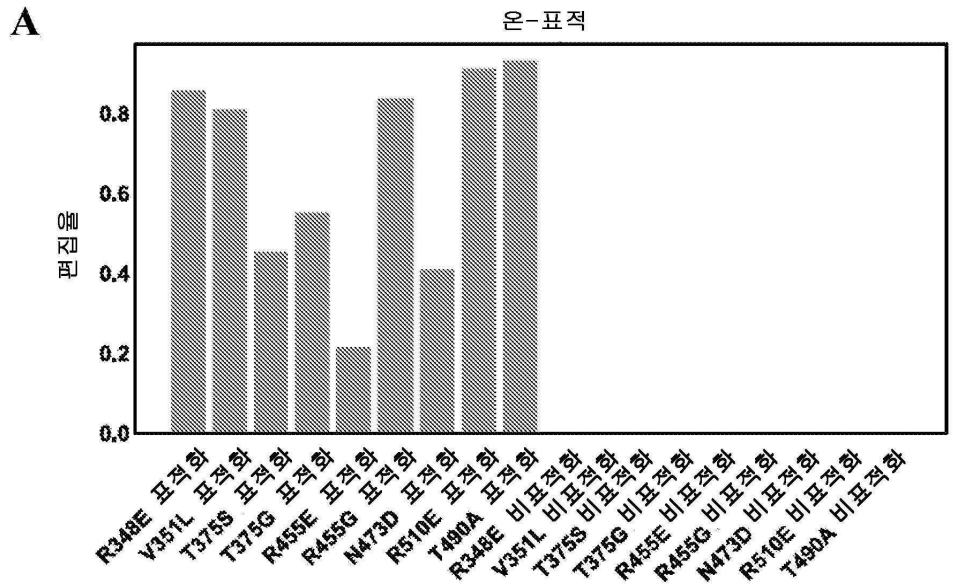
도면44


측쇄
골격
비율
온-표적
R348E
R474E
R477E
T490S
R481E
S495T
R510E
E488Q
T490A
R455G
R455S
R452E
V351L
T375G
T375S
S486T
K475Q
K473D
온-표적
2
4
6
비율
500
1000
1500

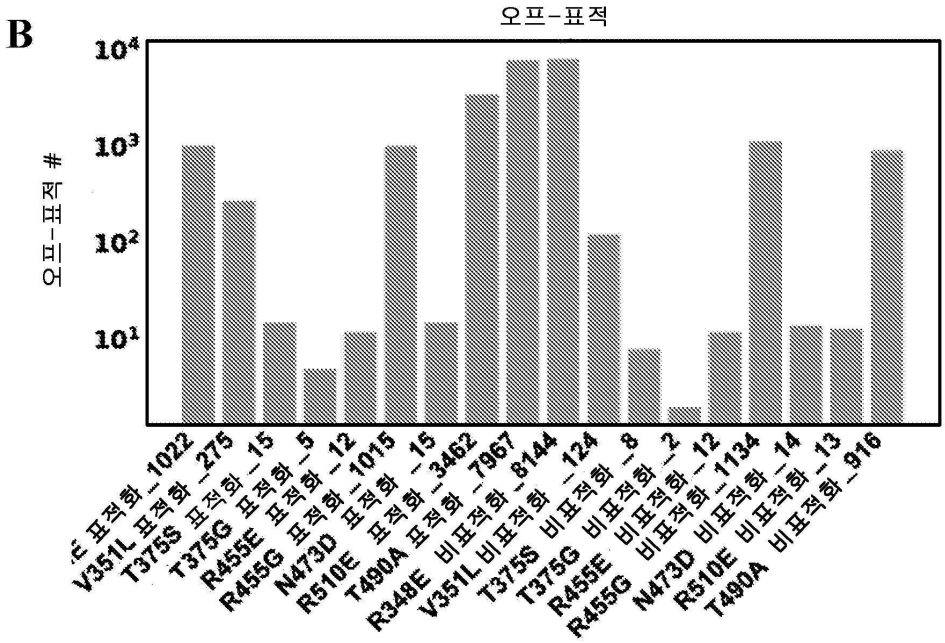
도면45


2000
1500
1000
500
0
표적화/비표적화 비율
T375G
T375S
N473D
R455E
R510E
V351L
T490A
R348E
R455G
0
2
4
6
8
표적화 가이드

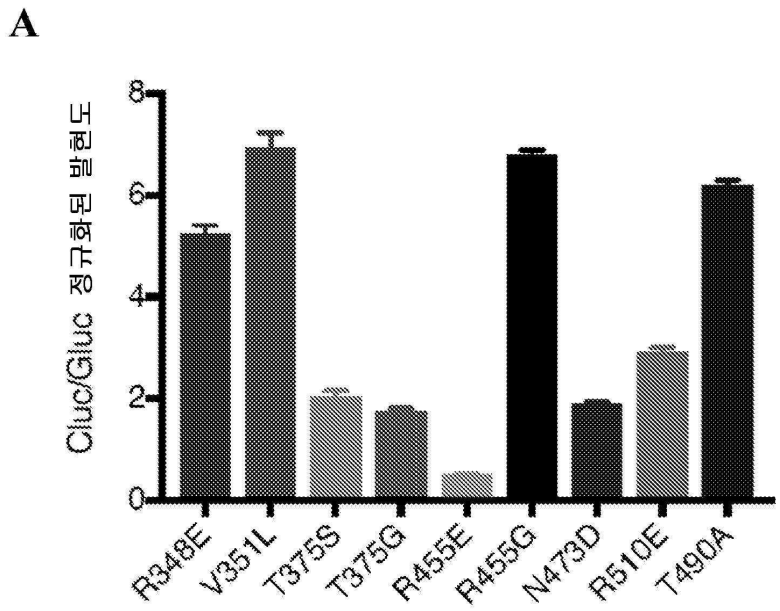
도면46a


A
온-표적
편집율
0.0
0.2
0.4
0.6
0.8
R348E 표적화
V351L 표적화
T375S 표적화
T375G 표적화
R455E 표적화
R455G 표적화
N473D 표적화
R510E 표적화
T490A 표적화
R348E 비표적화
V351L 비표적화
T375S 비표적화
T375G 비표적화
R455E 비표적화
R455G 비표적화
N473D 비표적화
R510E 비표적화
T490A 비표적화

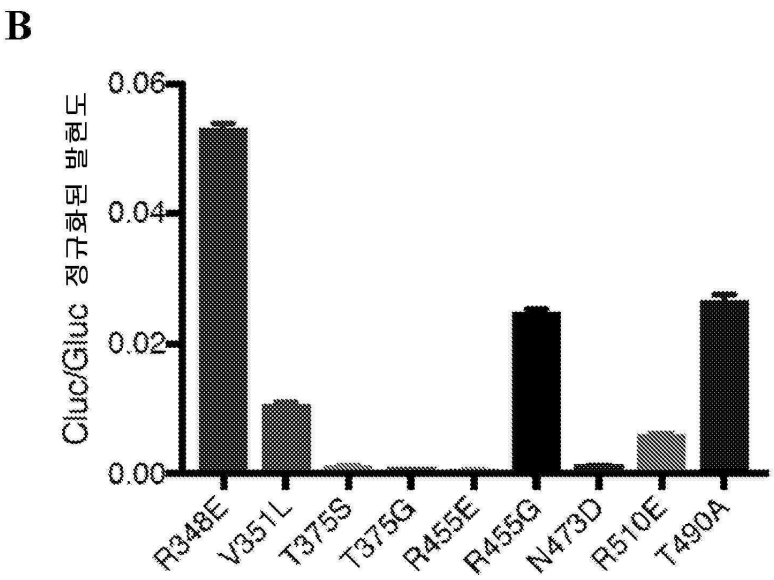
도면46b


B
오프-표적
오프-표적 #
10⁴
10³
10²
10¹
표적화_1022
V351L 표적화_275
T375S 표적화_15
T375G 표적화_5
R455E 표적화_12
R455G 표적화_1015
N473D 표적화_15
R510E 표적화_3462
T490A 표적화_7967
R348E 비표적화_8144
V351L 비표적화_124
T375S 비표적화_8
T375G 비표적화_2
R455E 비표적화_12
R455G 비표적화_1134
N473D 비표적화_14
R510E 비표적화_13
T490A 비표적화_916

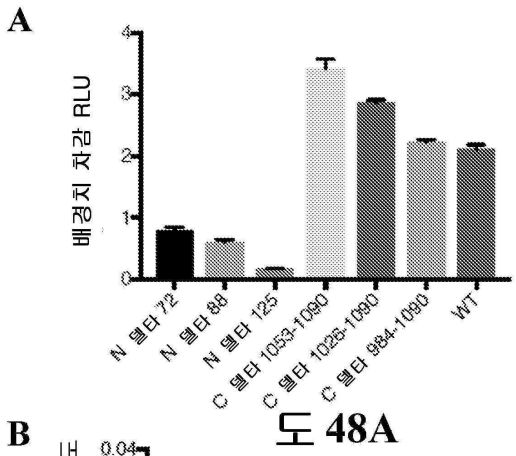
도면47a


A
Cluc/Gluc 정규화된 발현도
8
6
4
2
0
R348E
V351L
T375S
T375G
R455E
R455G
N473D
R510E
T490A

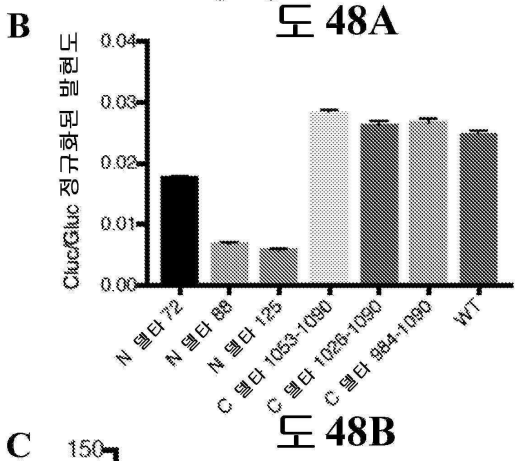
도면47b


B
Cluc/Gluc 정규화된 발현도
0.06
0.04
0.02
0.00
R348E
V351L
T375S
T375G
R455E
R455G
N473D
R510E
T490A

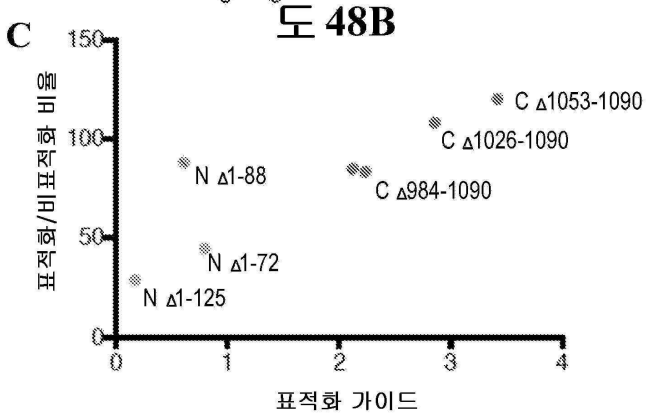
도면48



도 48A

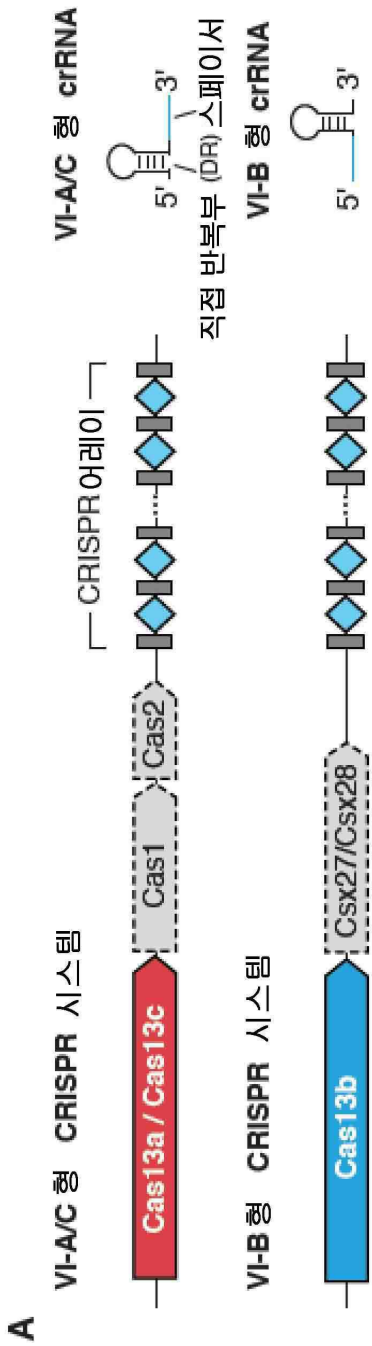


도 48B

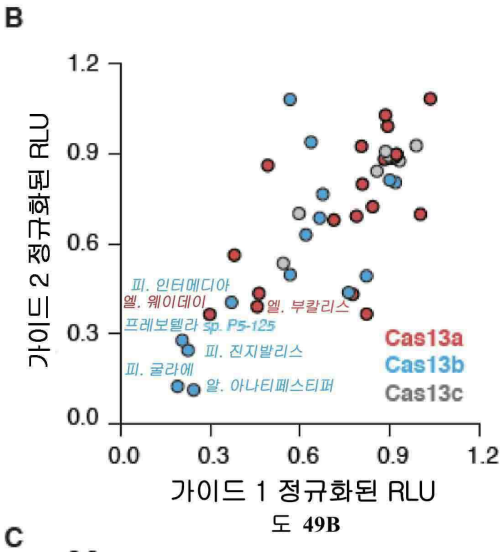


도 48C

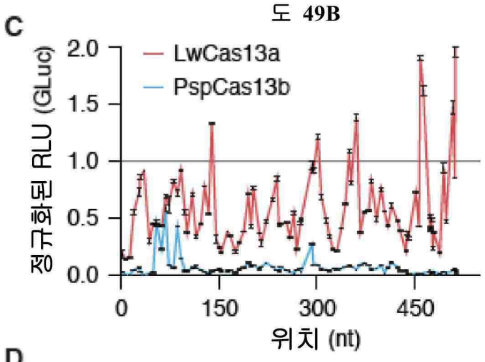
도면49a


A
VI-A/C 형 CRISPR 시스템
Cas13a / Cas13c
Cas1
Cas2
CRISPR 어레이
VI-A/C 형 crRNA
5'
3'
직접 반복부 (DR)
스페이서
VI-B 형 CRISPR 시스템
Cas13b
Csx27/Csx28
VI-B 형 crRNA
5'
3'

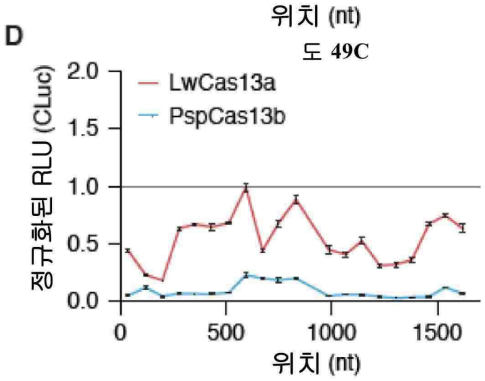
도면49bcd



도 49B

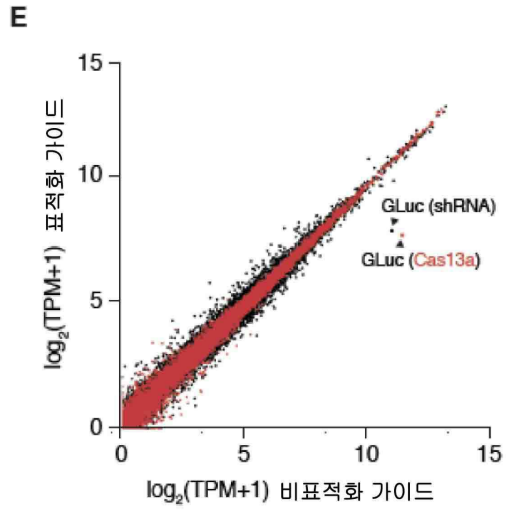


도 49C

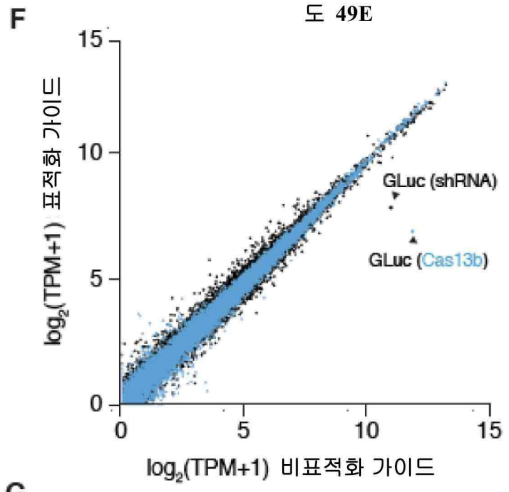


도 49D

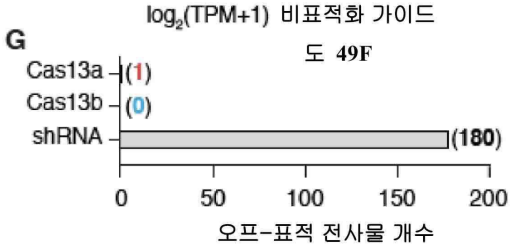
도면49efg



도 49E

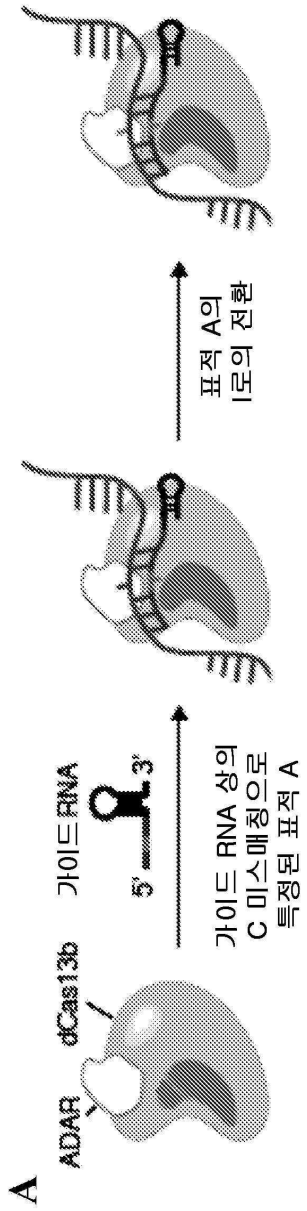


도 49F



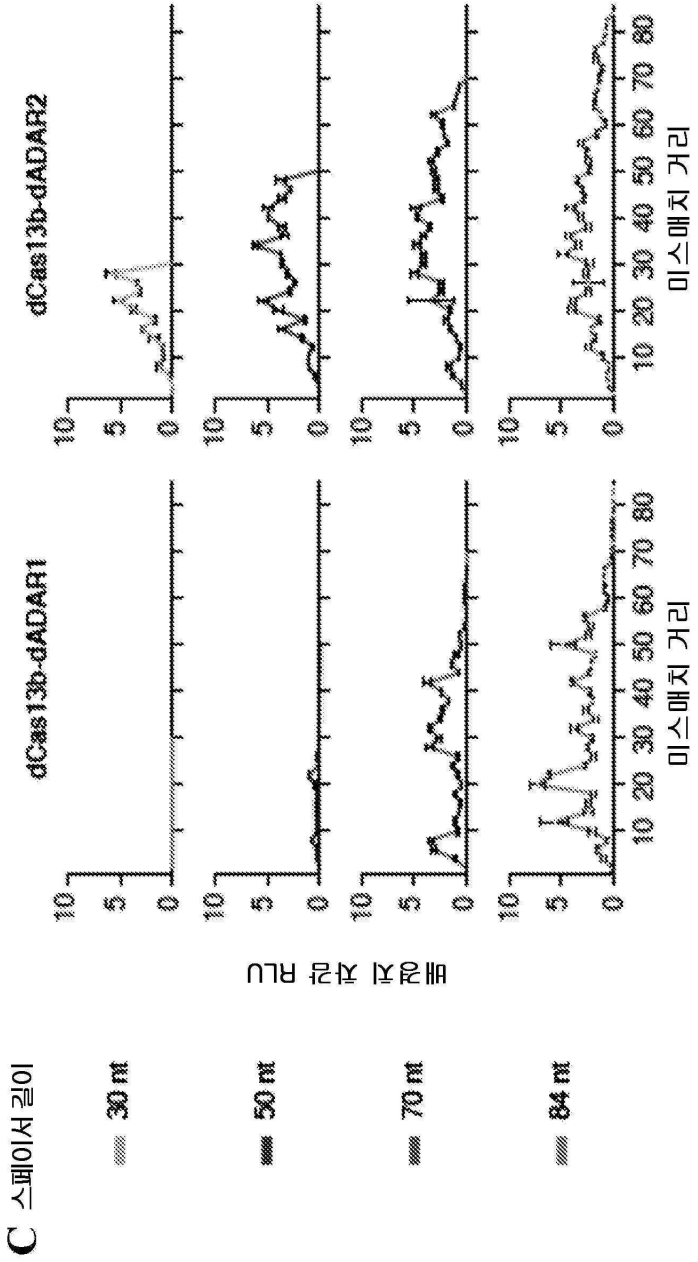
도 49G

도면50a


A
ADAR
dCas13b
가이드 RNA
5'
3'
가이드 RNA 상의
C 미스매칭으로
특정된 표적 A
표적 A의
I로의 전환

도면50b

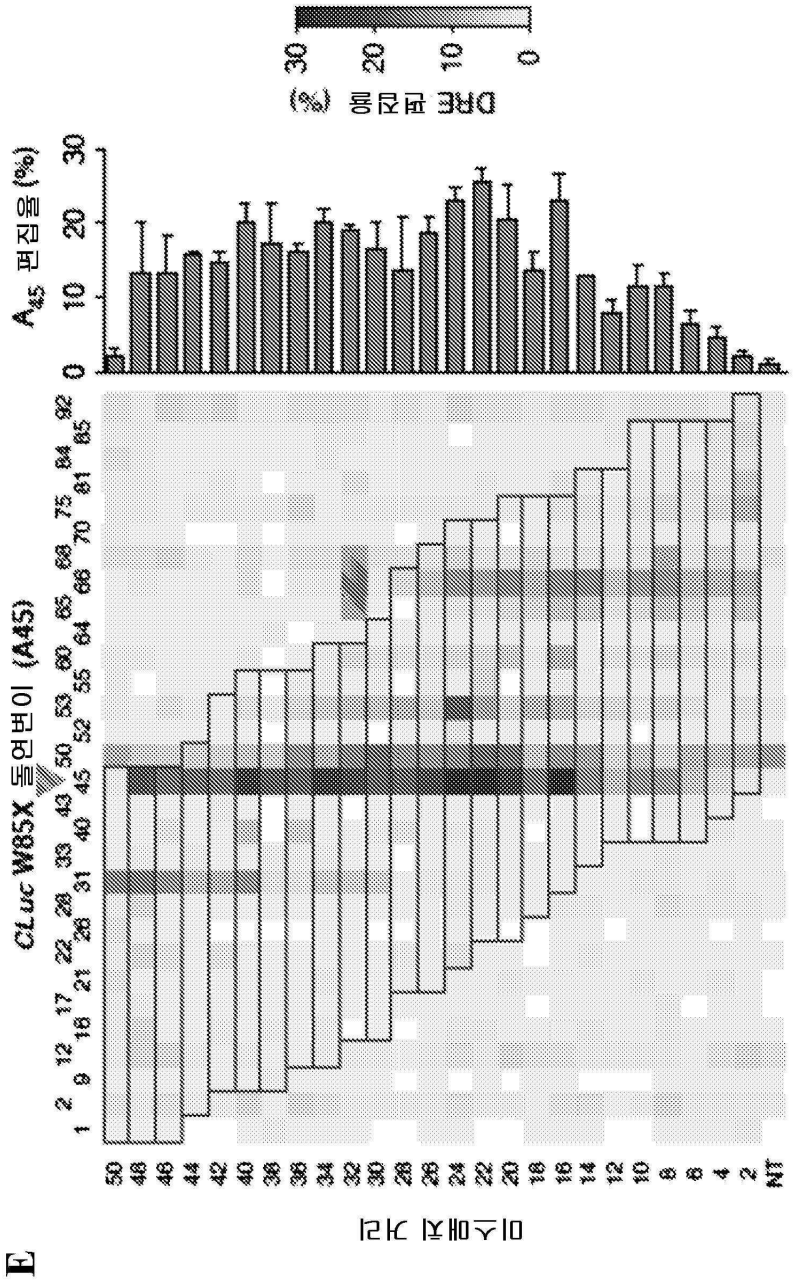
도면50c


C
스페이서 길이
30 nt
50 nt
70 nt
84 nt
배경치 차감 RLU
dCas13b-dADAR1
dCas13b-dADAR2
10
5
0
10 20 30 40 50 60 70 80
미스매치 거리

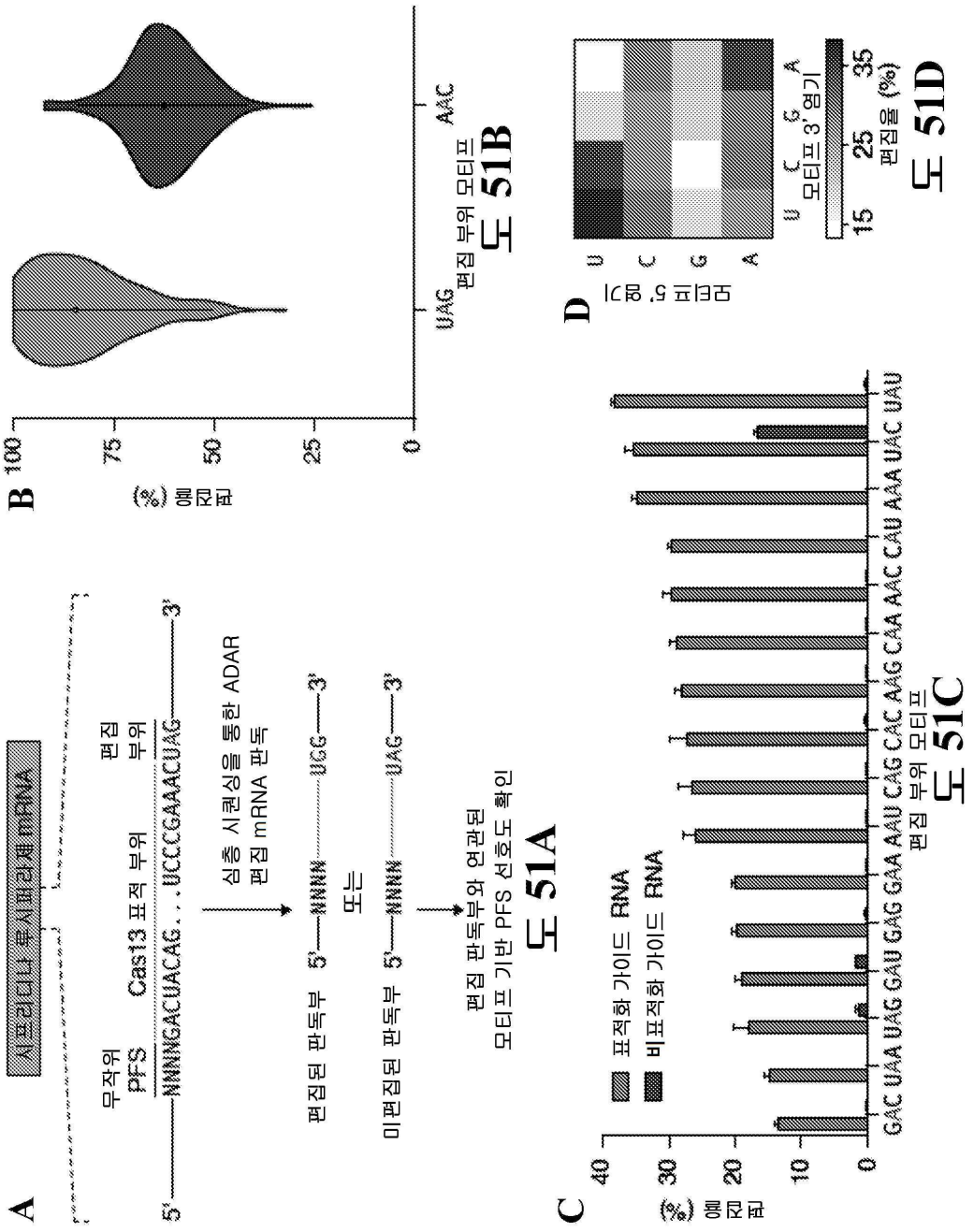
도면50d

D
CLuc W85X mRNA
1 2 9 12 1617 21 22 28 28 31 33 40 43 45 50 52 53 55 60
5′— . . TGC CGA AUG UGU CAG UAU GUA AUU GAA UGC AGA GUA GAG GCC GCA GGA UAG UUU AGA ACA UUC UAU
64 65 66 68 70 75 81 84 85 92
GGA AAG AGA UUC CAG UUC CAG GAA CCU GGU ACA UAC GUG . . —3′

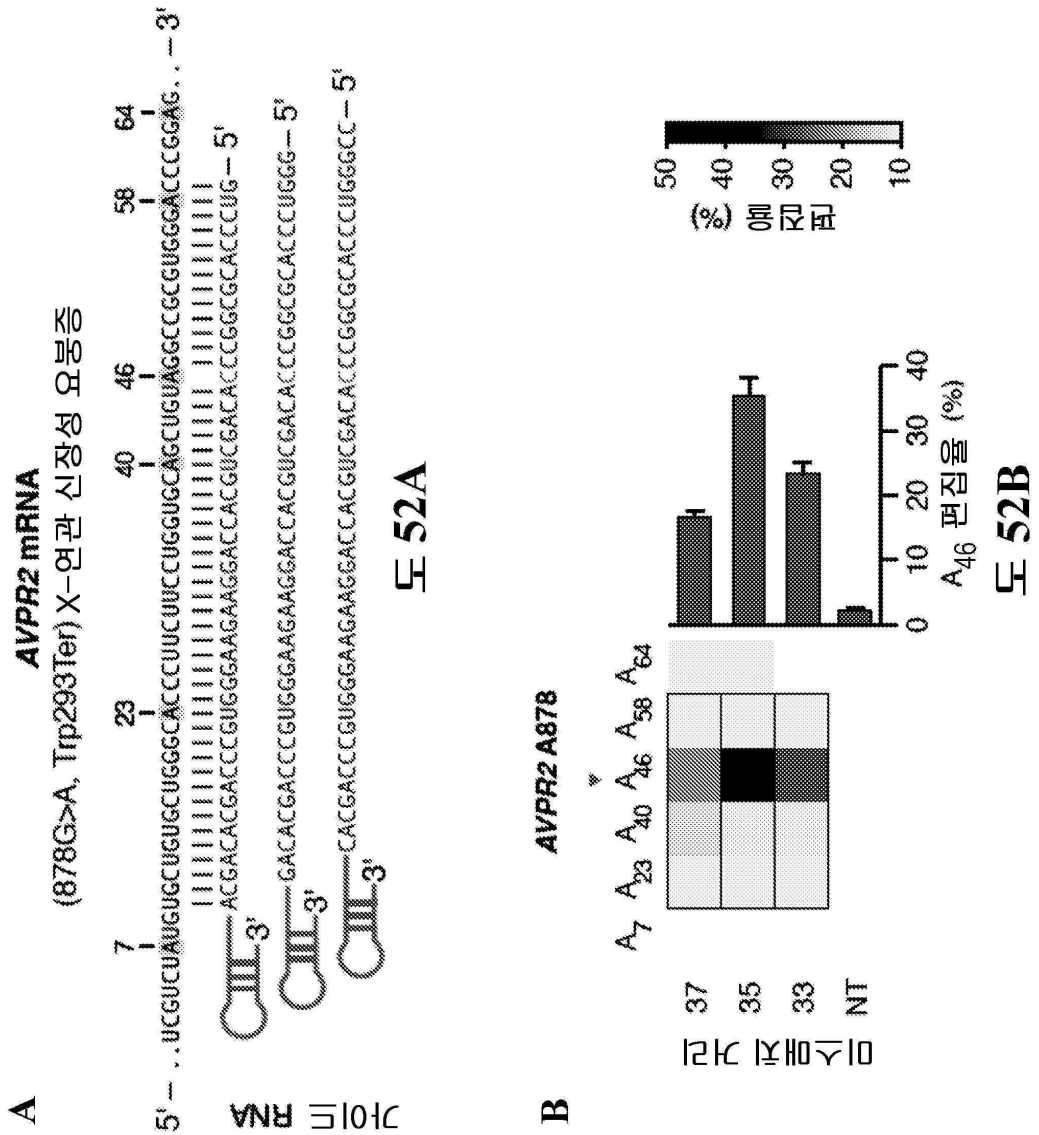
도면50e



도면51


A
시프리디나 루시퍼라제 mRNA
무작위
PFS
Cas13 표적 부위
편집
부위
5'
NNNNGACUACAG...UCCCGAAACUAG
3'
심층 시퀀싱을 통한 ADAR
편집 mRNA 판독
편집된 판독부 5' NNNN UGG 3'
또는
미편집된 판독부 5' NNNN UAG 3'
편집 판독부와 연관된
모티프 기반 PFS 선호도 확인
도 51A
B
100
75
50
25
0
편집율 (%)
UAG
AAC
편집 부위 모티프
도 51B
C
40
30
20
10
0
편집율 (%)
표적화 가이드 RNA
비표적화 가이드 RNA
GAC UAA UAG GAU GAG GAA AAU CAG CAC AAG CAA AAC CAU AAA UAC UAU
편집 부위 모티프
도 51C
D
U
C
G
A
모티프 5' 염기
U C G A
모티프 3' 염기
15
25
35
편집율 (%)
도 51D

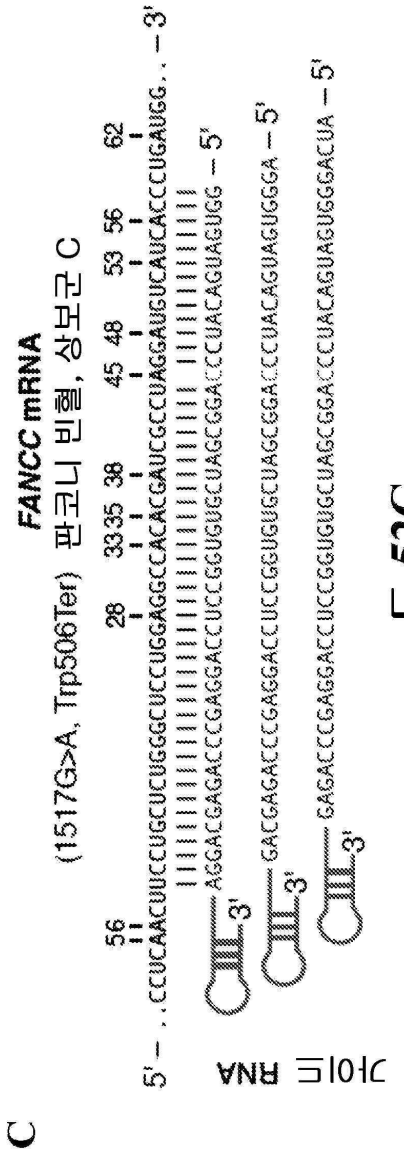
도면52ab



도 52A

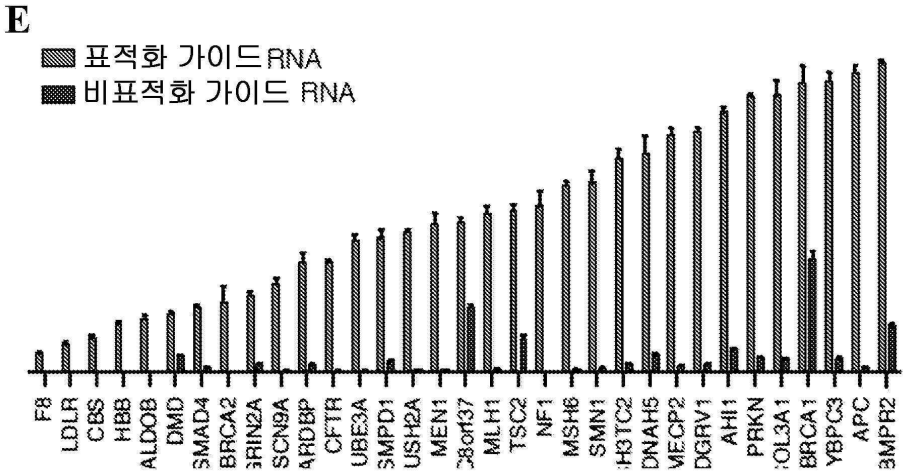
도 52B

도면52cd

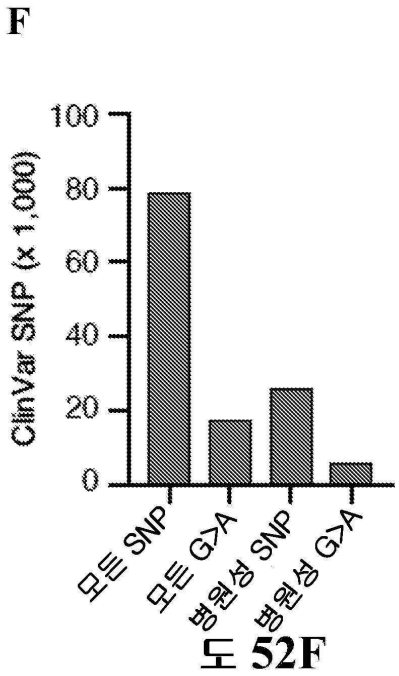

C
FANCC mRNA
(1517G>A, Trp506Ter) 판코니 빈혈, 상보군 C
5' – . . CCUCAACUUCCUGCUCUGGGCUCCUGGAGGCCACACGAUCGCCUAGGAUGUCAUCACCCUGAUGG. . – 3'
AGGACGAGACCCGAGGACCUCCGGUGUGCUAGCGGACCCUACAGUAGUGG – 5'
GACGAGACCCGAGGACCUCCGGUGUGCUAGCGGACCCUACAGUAGUGGGA – 5'
GAGACCCGAGGACCUCCGGUGUGCUAGCGGACCCUACAGUAGUGGGACUA – 5'
RNA 가이드
도 52C

D
A46 편집율 (%)
FANCC A1517
A5 A6 A28 A33 A35 A38 A45 A48 A53 A56 A62
0 10 20 30 40
A45 편집율 (%)
NT 32 35 37
미스매치 거리
도 52D

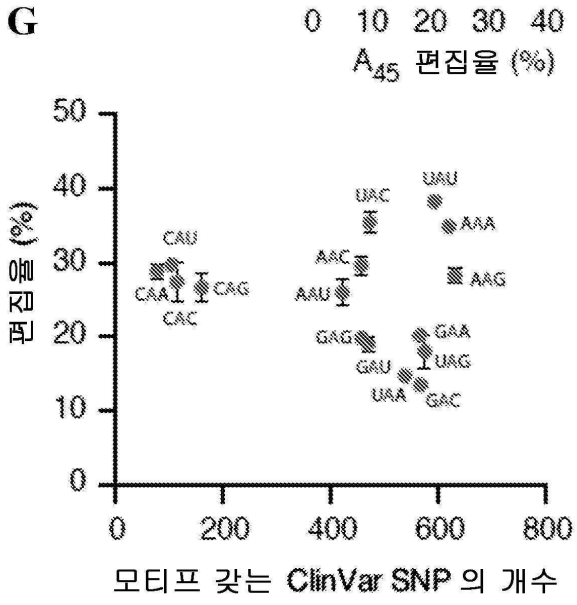
도면52efg



도 52E

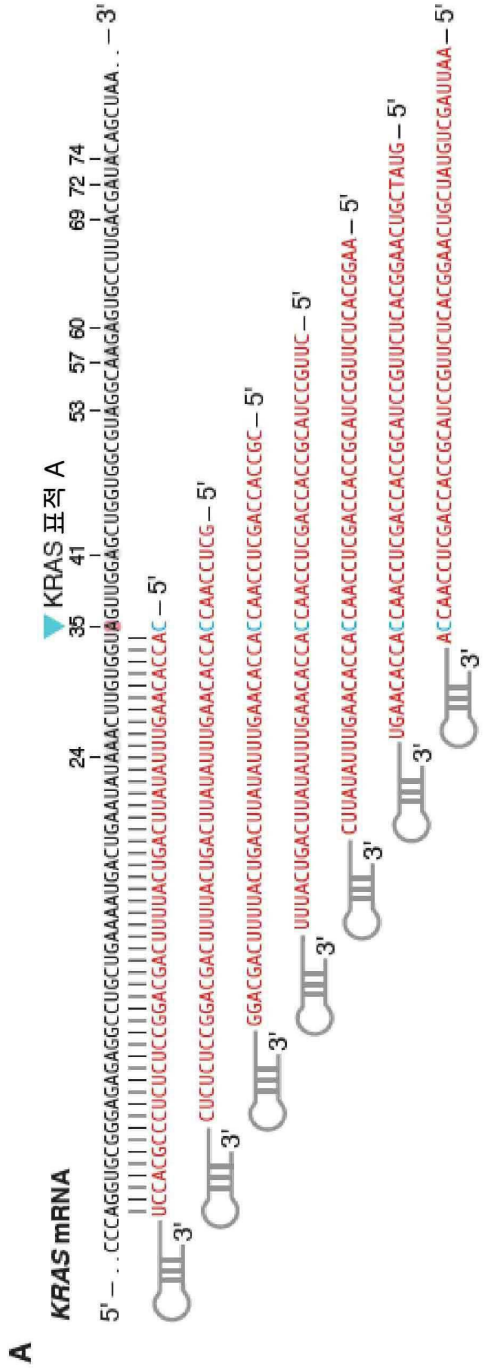


도 52F

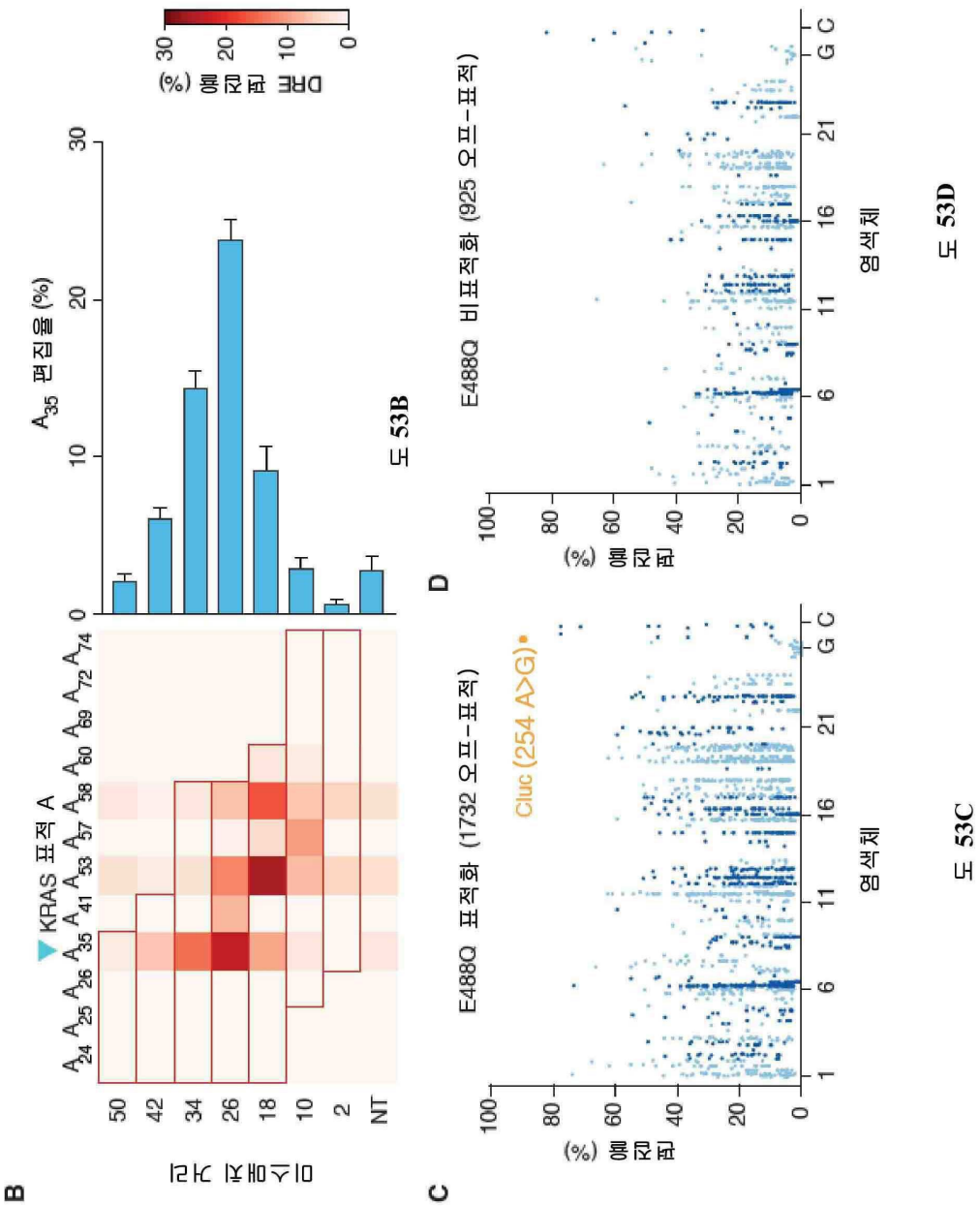


도 52G

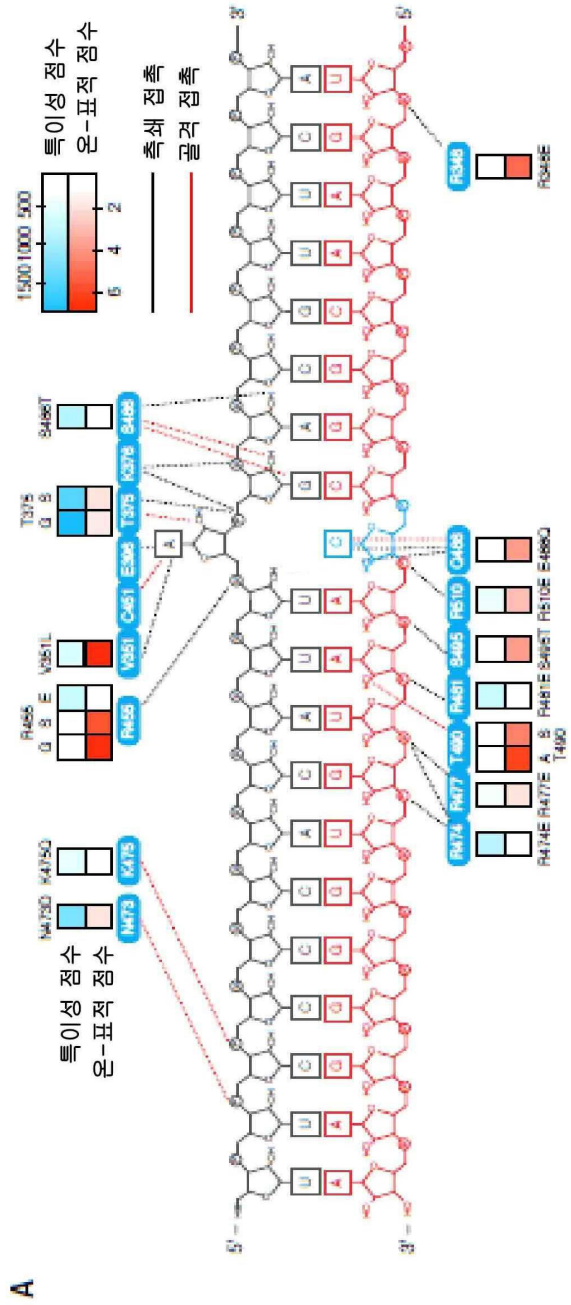
도면53a



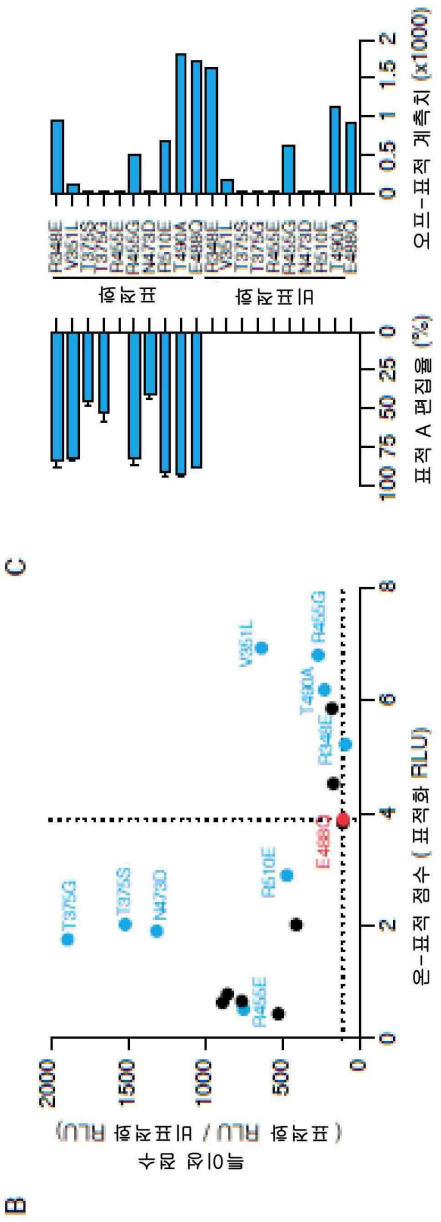
도면53bcd


B
KRAS 표적 A
A24 A25 A26 A35 A41 A53 A57 A58 A60 A69 A72 A74
미스매치 거리
50 42 34 26 18 10 2 NT
A35 편집율 (%)
0 10 20 30
DRE 편집율 (%)
0 10 20 30
도 53B
C
E488Q 표적화 (1732 오프-표적)
Cluc (254 A>G)
편집율 (%)
0 20 40 60 80 100
염색체
1 6 11 16 21 G C
도 53C
D
E488Q 비표적화 (925 오프-표적)
편집율 (%)
0 20 40 60 80 100
염색체
1 6 11 16 21 G C
도 53D

도면54a


A
특이성 점수
온-표적 점수
N473D
K478Q
N473
K478
R455
Q
S
E
V351L
V351
C451
E396
T375
Q
S
T375
K378
S486
S486T
특이성 점수
온-표적 점수
15001000500
6
4
2
측쇄 접촉
골격 접촉
5'
3'
3'
5'
R474
R477
T480
R481
S485
R510
C483
R348
R474E
R477E
A, S
T490
R481E
S485T
R510E
E488Q
R348E

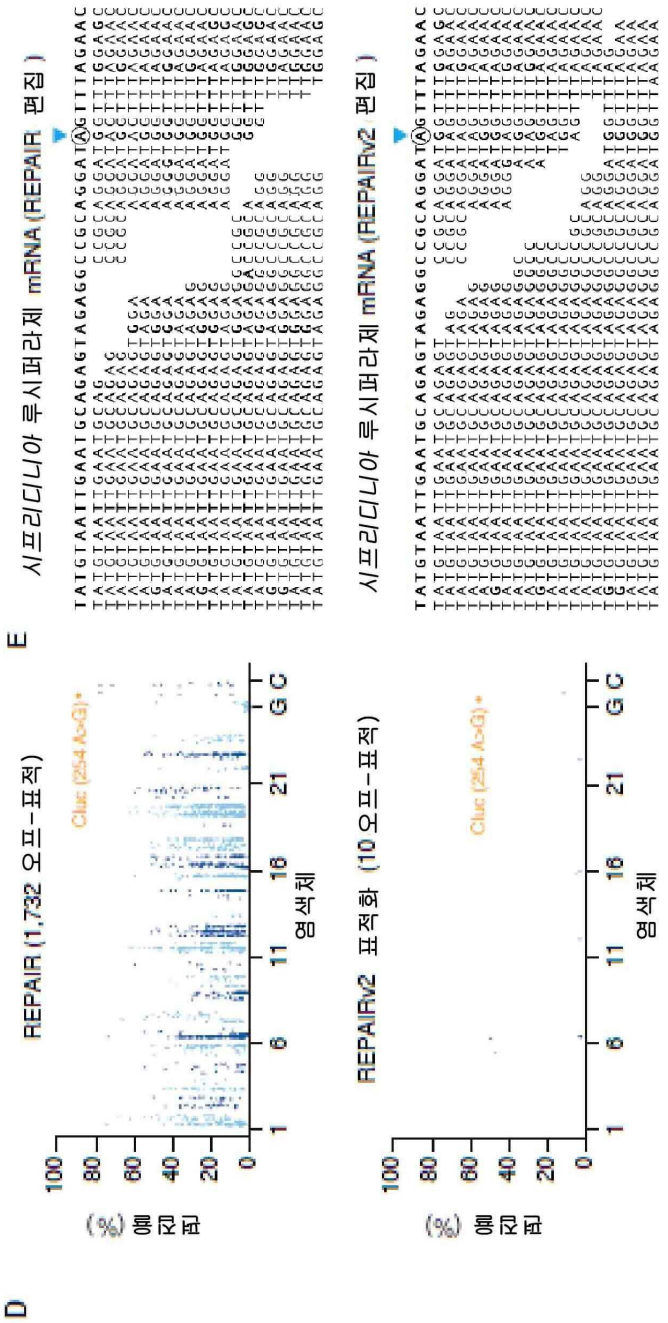
도면54bc



도 54B

도 54C

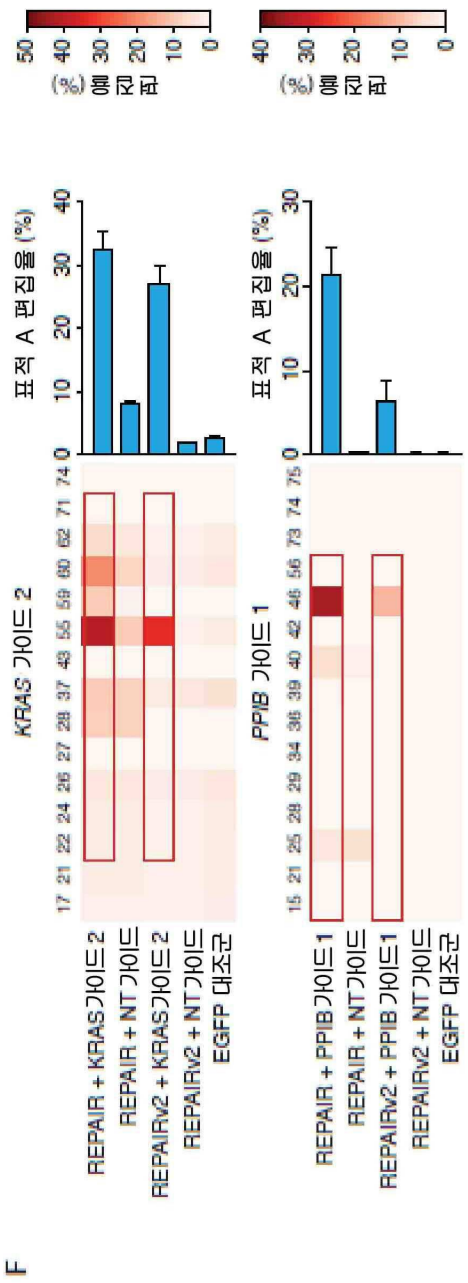
도면54de



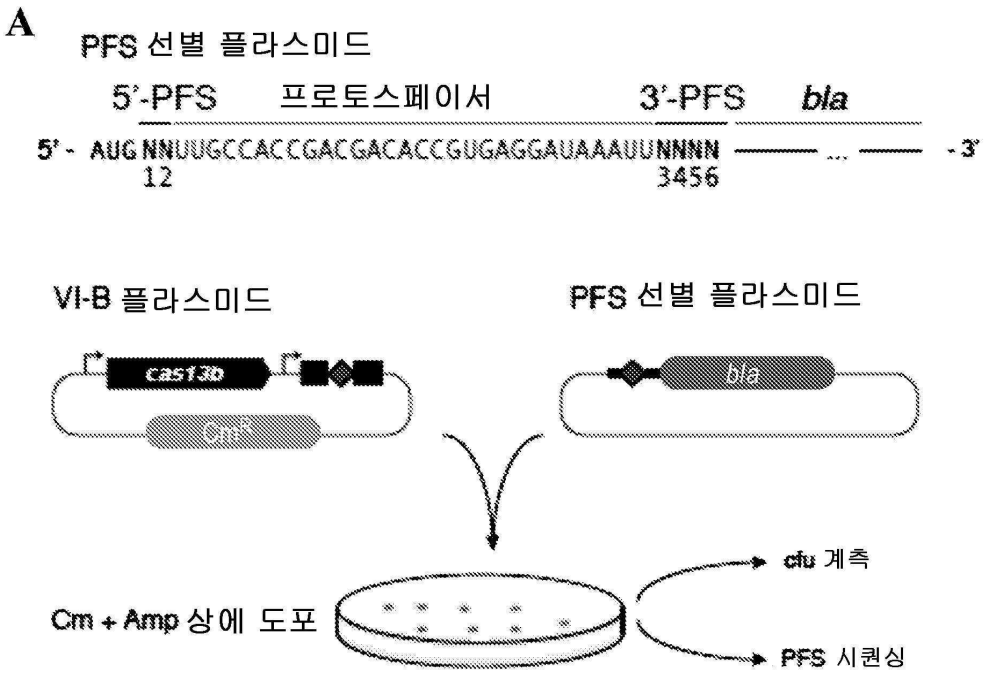
도 54D

도 54E

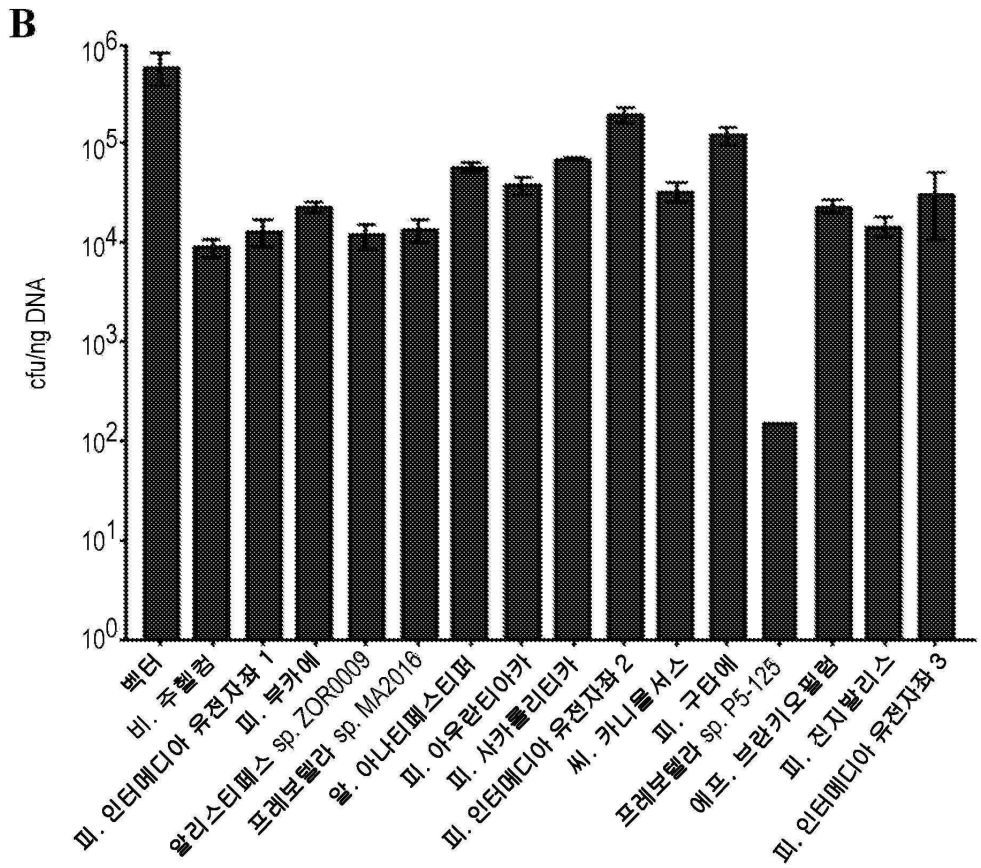
도면54f


F
KRAS 가이드 2
표적 A 편집율 (%)
0
10
20
30
40
17 21 22 24 26 27 28 37 43 55 59 60 62 71 74
REPAIR + KRAS가이드 2
REPAIR + NT 가이드
REPAIRv2 + KRAS가이드 2
REPAIRv2 + NT가이드
EGFP 대조군
편집율 (%)
0
10
20
30
40
50
PPIB 가이드 1
표적 A 편집율 (%)
0
10
20
30
15 21 25 28 29 34 36 39 40 42 46 56 73 74 75
REPAIR + PPIB가이드 1
REPAIR + NT가이드
REPAIRv2 + PPIB 가이드1
REPAIRv2 + NT 가이드
EGFP 대조군
편집율 (%)
0
10
20
30
40

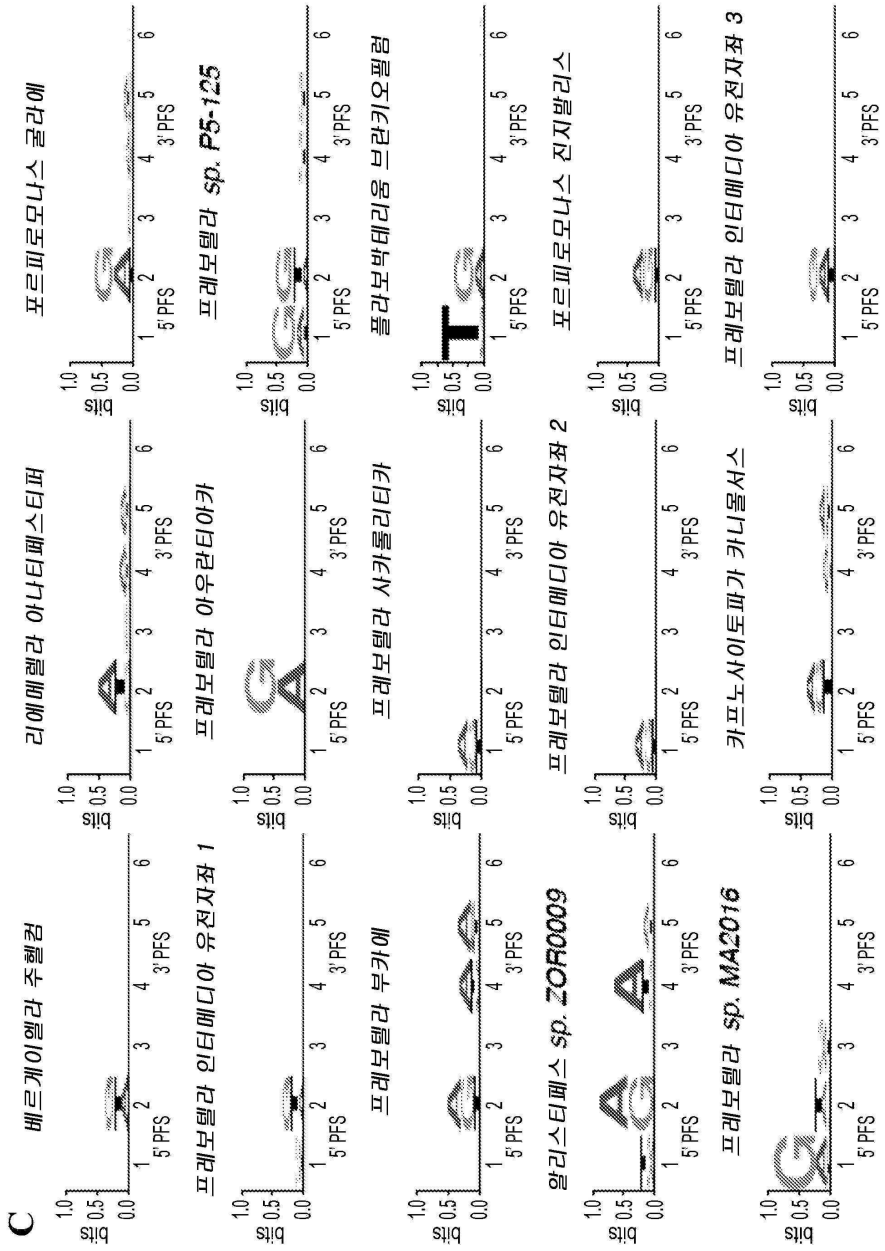
도면55a


A
PFS 선별 플라스미드
5'-PFS
프로토스페이서
3'-PFS
bla
5' - AUG NNUUGCCACCGACGACACCGUGAGGAUAAAAUUNNNN ... - 3'
12
3456
VI-B 플라스미드
PFS 선별 플라스미드
cas1 3b
Cm^R
bla
cfu 계측
Cm + Amp 상에 도포
PFS 시퀀싱

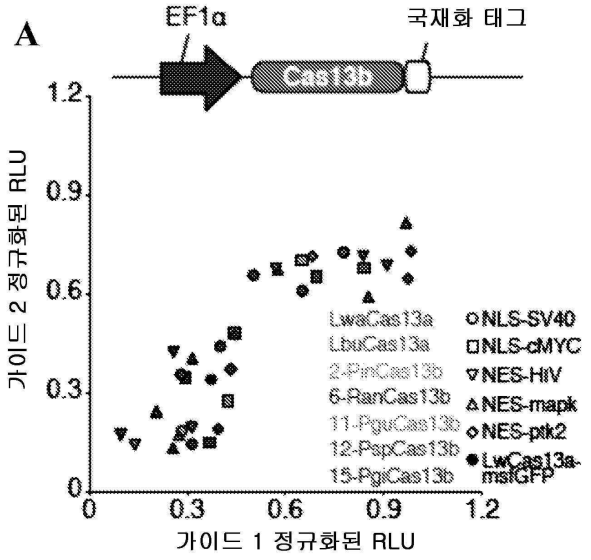
도면55b


B
cfu/ng DNA
10^6
10^5
10^4
10^3
10^2
10^1
10^0
벡터
비. 주헬럼
피. 인터메디아 유전자좌 1
피. 루카에
알리스티페스 sp. ZOR0009
프레보텔라 sp. MA2016
알. 아너티페스티퍼
피. 아우란티아카
피. 사카롤리티카
피. 인터메디아 유전자좌 2
씨. 카니룰서스
피. 구타에
프레보텔라 sp. P5-125
에프. 브라키오팔렘
피. 진지발리스
피. 인터메디아 유전자좌 3

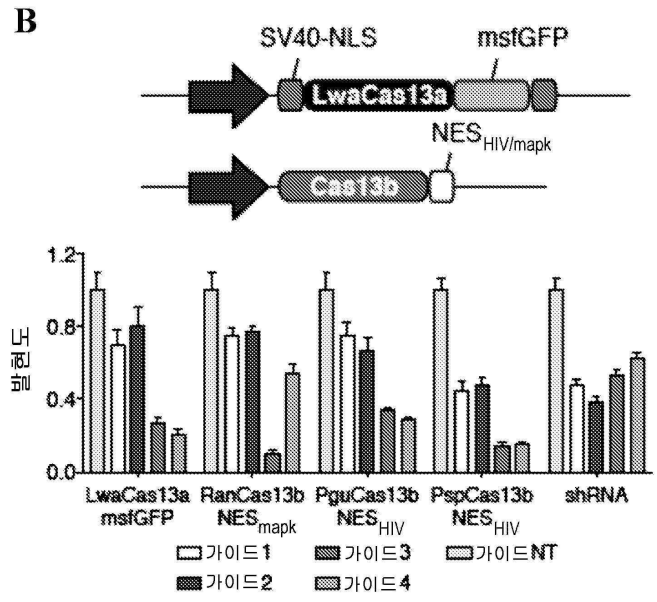
도면55c


C
베르케이엘라 주헬컴
리에메렐라 아나티페스티퍼
포르피로모나스 글라에
프레보텔라 인터메디아 유전자좌 1
프레보텔라 아우라티아카
프레보텔라 sp. P5-125
프레보텔라 부카에
프레보텔라 사카롤리티카
플라브박테리움 브랑키오필럼
알리스티페스 sp. ZOR0009
프레보텔라 인터메디아 유전자좌 2
포르피로모나스 진지발리스
프레보텔라 sp. MA2016
카프노사이토파가 카니몰서스
프레보텔라 인터메디아 유전자좌 3
bits
5' PFS
3' PFS

도면56ab



도 56A



도 56B

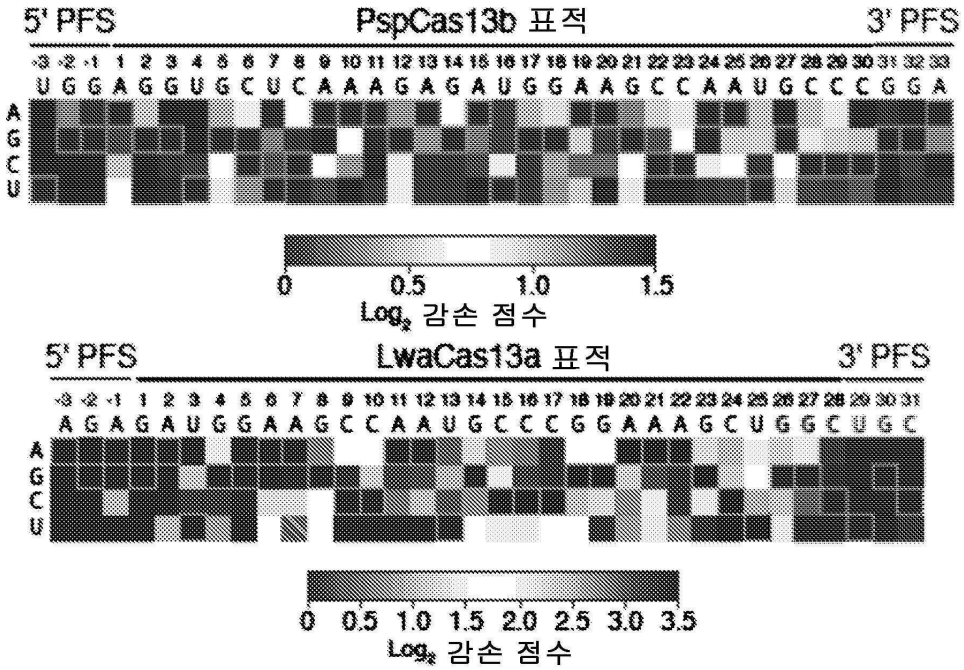
도면56c

C

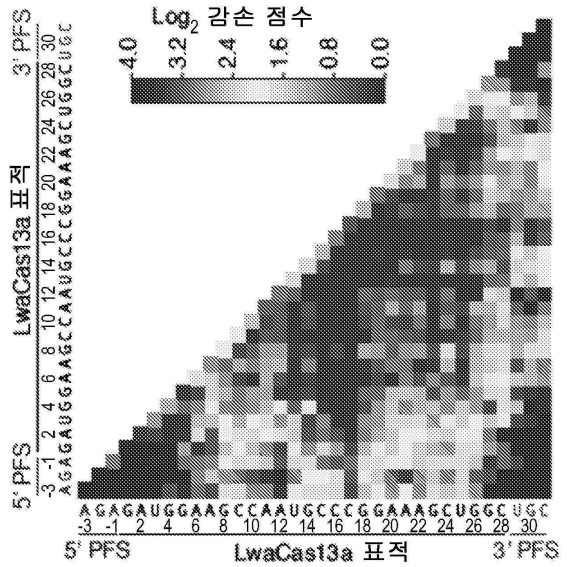


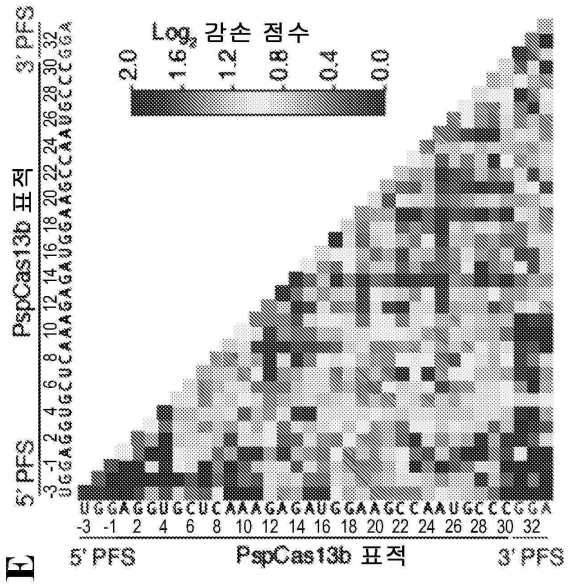
도면56d

D



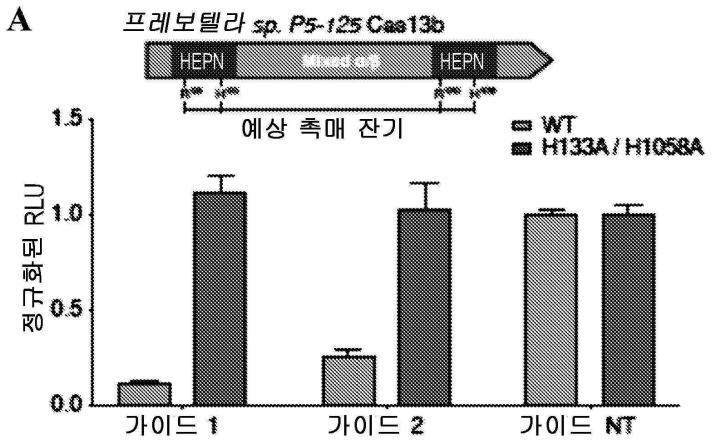
도면56e


Log2 감손 점수
4.0
3.2
2.4
1.6
0.8
0.0
LwaCas13a 표적
5' PFS
3' PFS

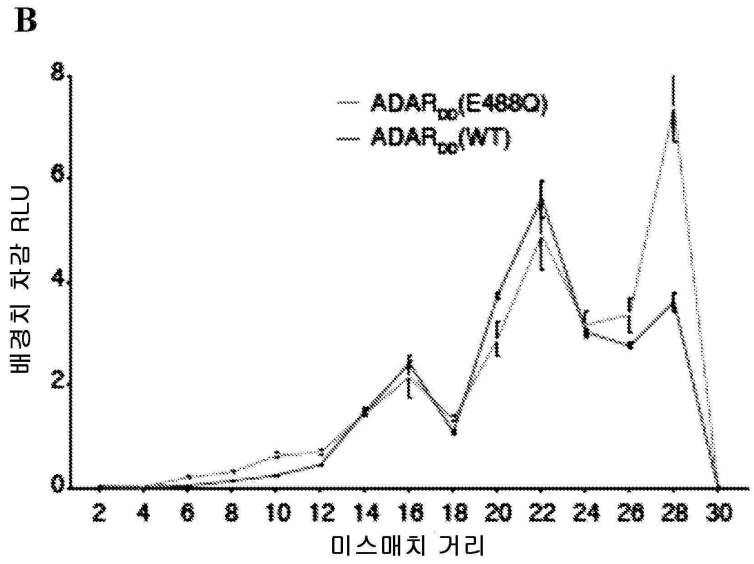

Log2 감손 점수
2.0
1.6
1.2
0.8
0.4
0.0
PspCas13b 표적
5' PFS
3' PFS

E

도면57ab

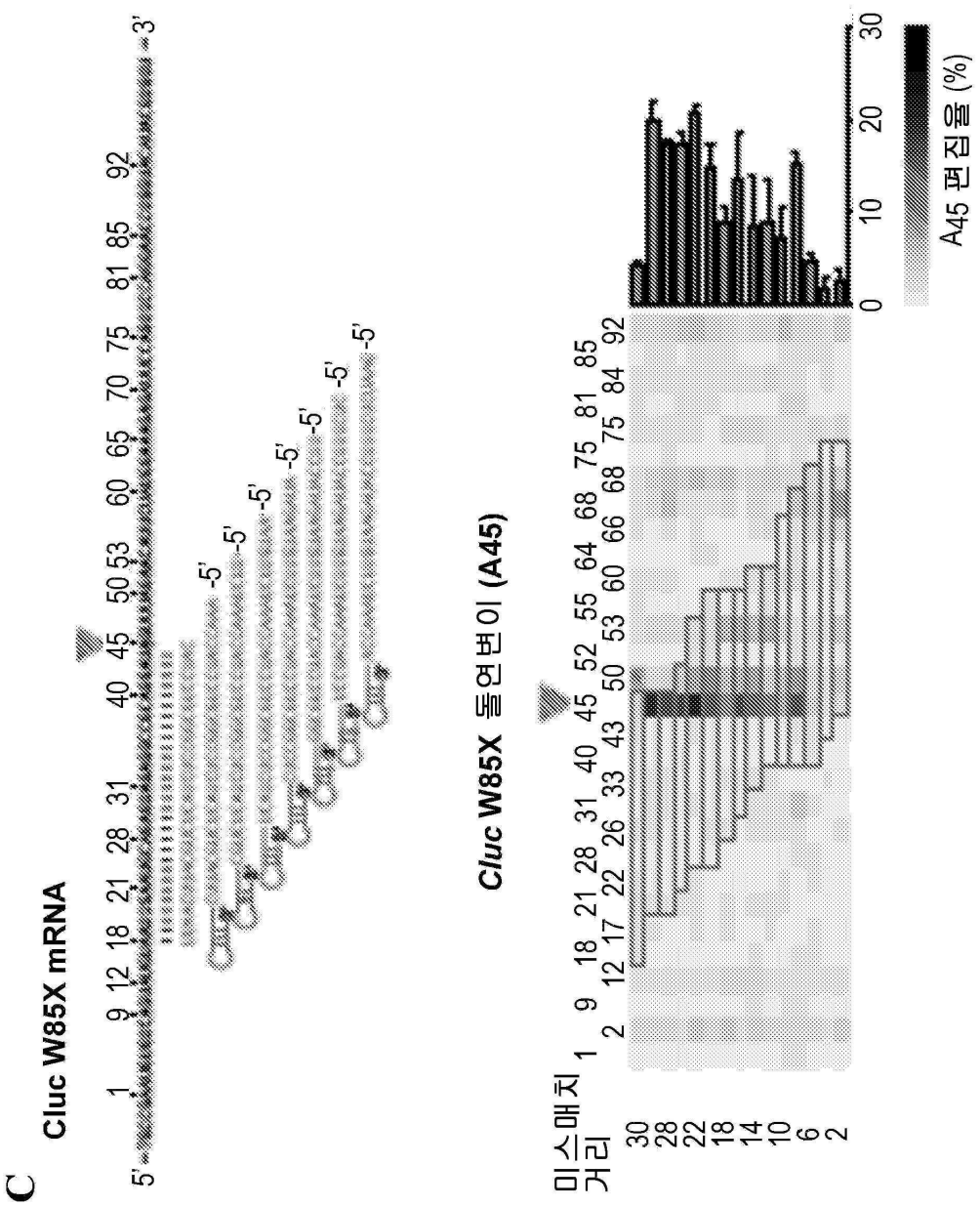


도 57A

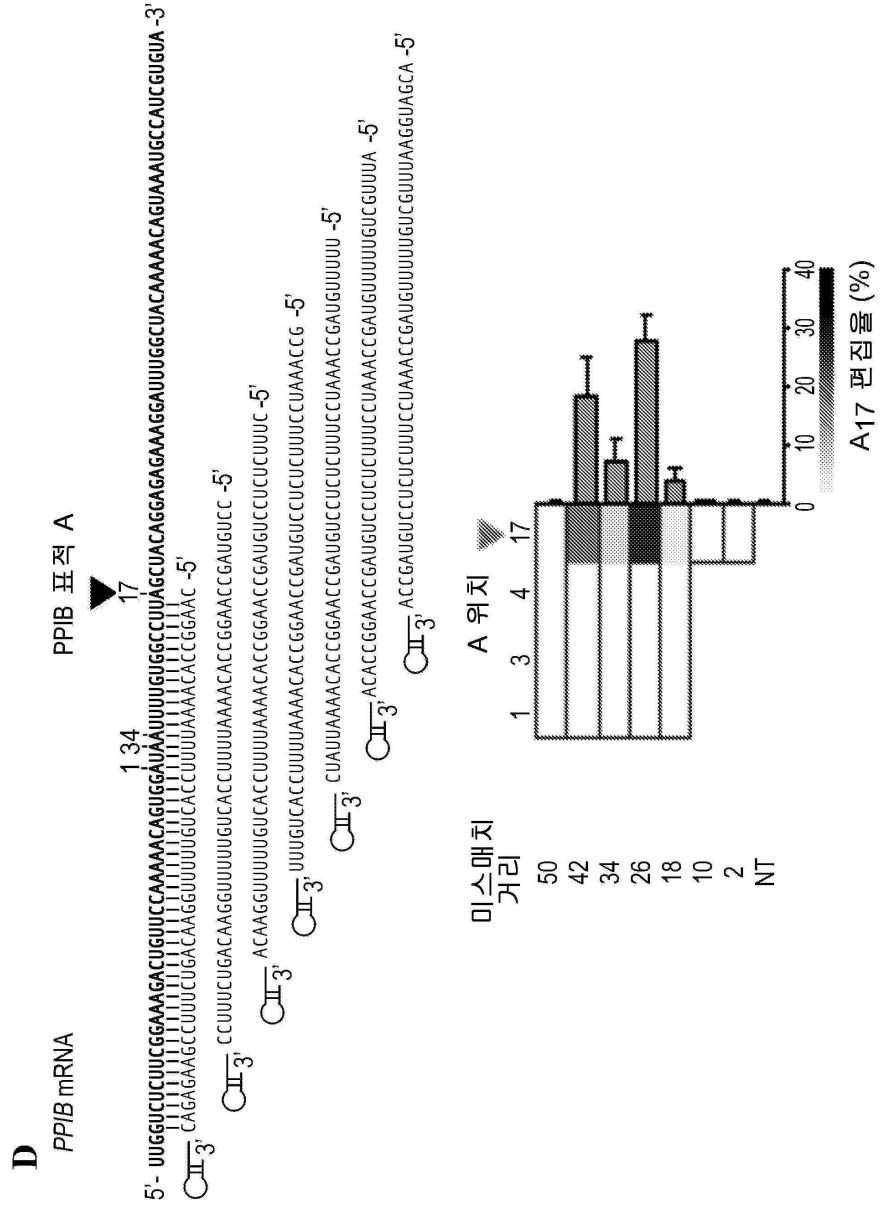


도 57B

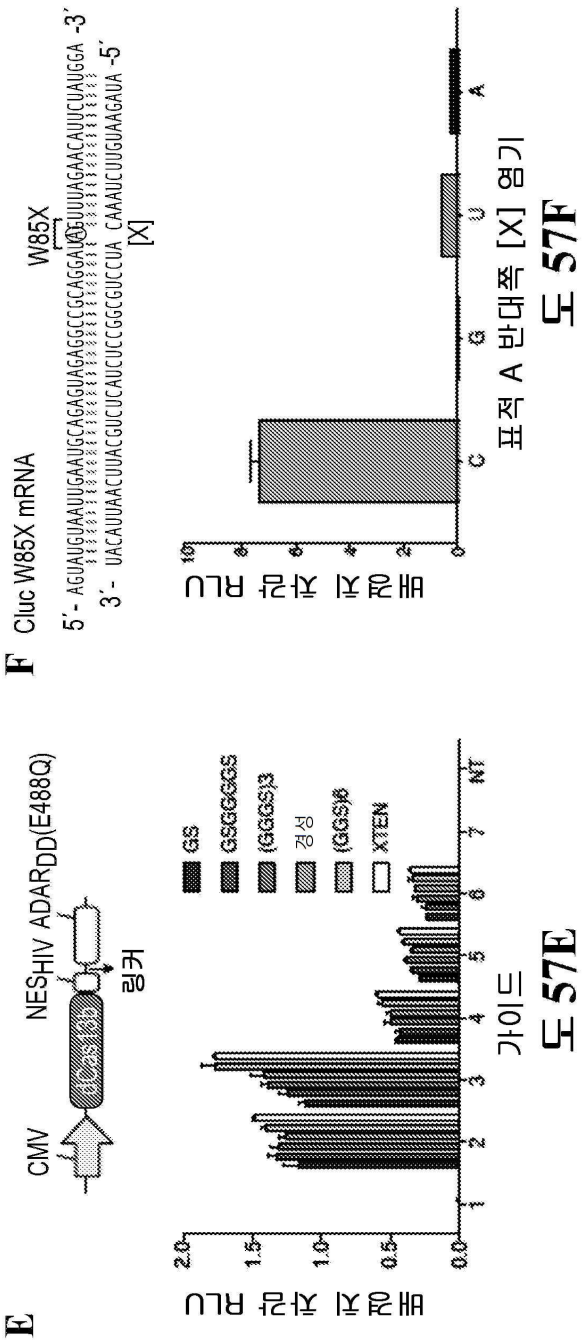
도면57c



도면57d


D
PPIB mRNA
PPIB 표적 A
1 34
17
5'- UUGGUCUCUUCGGAAAGACUGUUCCAAAAACAGUGGAUAAUUUUGUGGCCUUAGCUACAGGAGAAAGGAUUUGGCUACAAAAACAGUAAAUGCCAUCGUGUA -3'
3' CAGAGAAGCCUUUCUGACAAGGUUUUUGUCACCUUUUAAAACACCGGAAC -5'
3' CCUUUCUGACAAGGUUUUUGUCACCUUUUAAAACACCGGAACCGAUGUCC -5'
3' ACAAGGUUUUUGUCACCUUUUAAAACACCGGAACCGAUGUCCUCUCUUUC -5'
3' UUUGUCACCUUUUAAAACACCGGAACCGAUGUCCUCUCUUUCCUAAACCG -5'
3' CUAUUAAAACACCGGAACCGAUGUCCUCUCUUUCCUAAACCGAUGUUUUU -5'
3' ACACCGGAACCGAUGUCCUCUCUUUCCUAAACCGAUGUUUUUGUCGUUUA -5'
3' ACCGAUGUCCUCUCUUUCCUAAACCGAUGUUUUUGUCGUUUUAAGGUAGCA -5'
A 위치
1 3 4 17
미스매치 거리
50
42
34
26
18
10
2
NT
0 10 20 30 40
A17 편집율 (%)

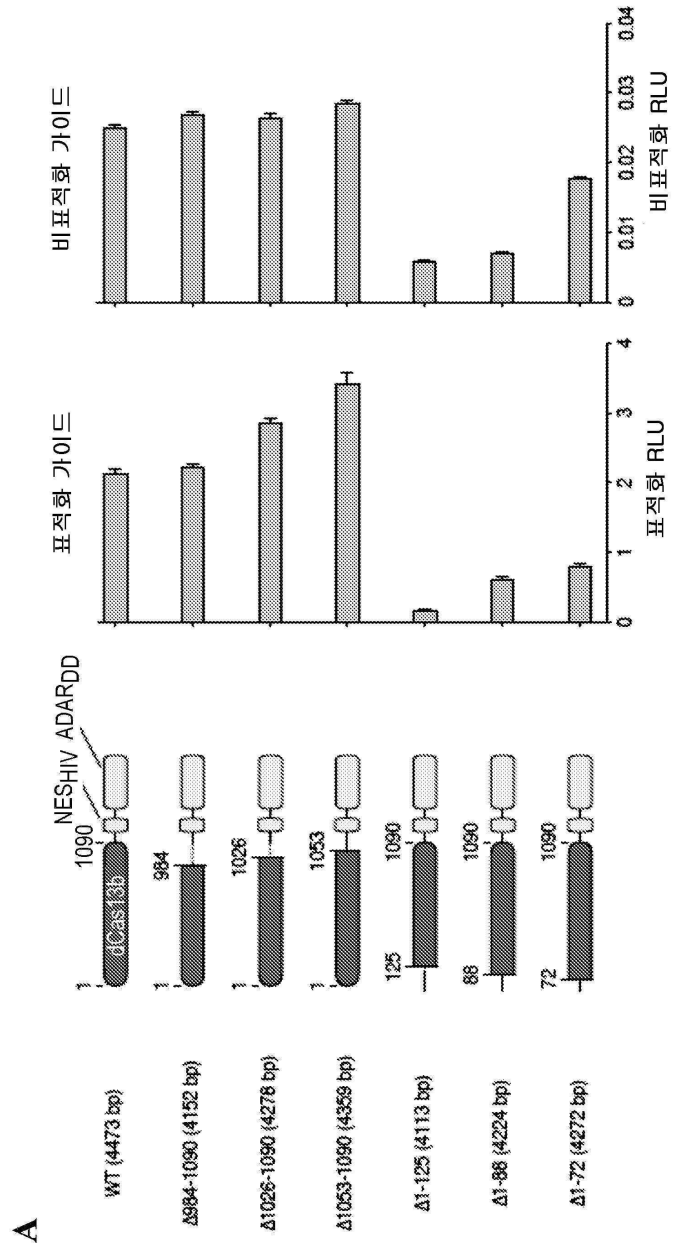
도면57ef


F
Cluc W85X mRNA
W85X
5′- AGUAUGUAAUUGAAUGCAGAGUAGAGGCCGCAGGAUGUUUAGAACAUUCUAUGGA -3′
3′- UACAUUAACUUACGUCUCAUCUCCGGCGUCCUA CAAAUCUUGUAAGAUA -5′
[X]
발현치 차감 RLU
C
G
U
A
표적 A 반대쪽 [X] 염기
도 57F
E
CMV
NESHIV ADARDD(E488Q)
링커
dCas13b
GS
GSGGGGS
(GGGS)3
검성
(GGS)6
XTEN
발현치 차감 RLU
가이드
NT
도 57E

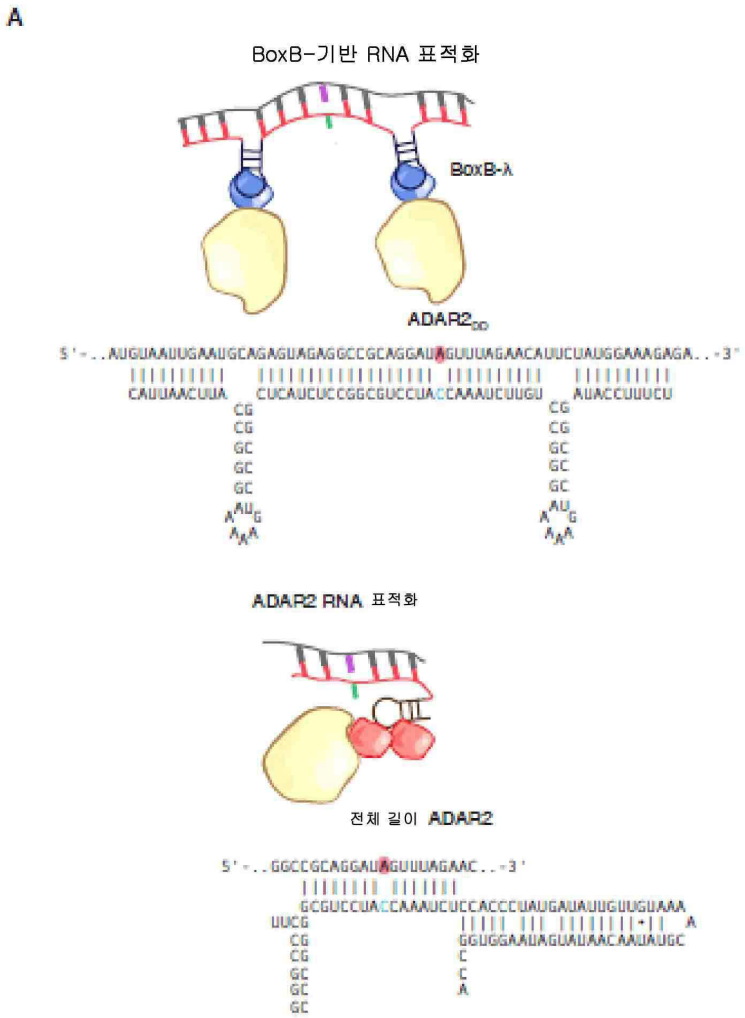
도면58


A
G->A 변화에 대한 ClinVar 모티프 분포도
모티프
0
100
200
300
400
500
600
GAU
CAG
UAU
CAA
CAC
AAC
AAA
UAC
GAG
UAA
GAA
UAG
GAC
AAG
AAU
CAU

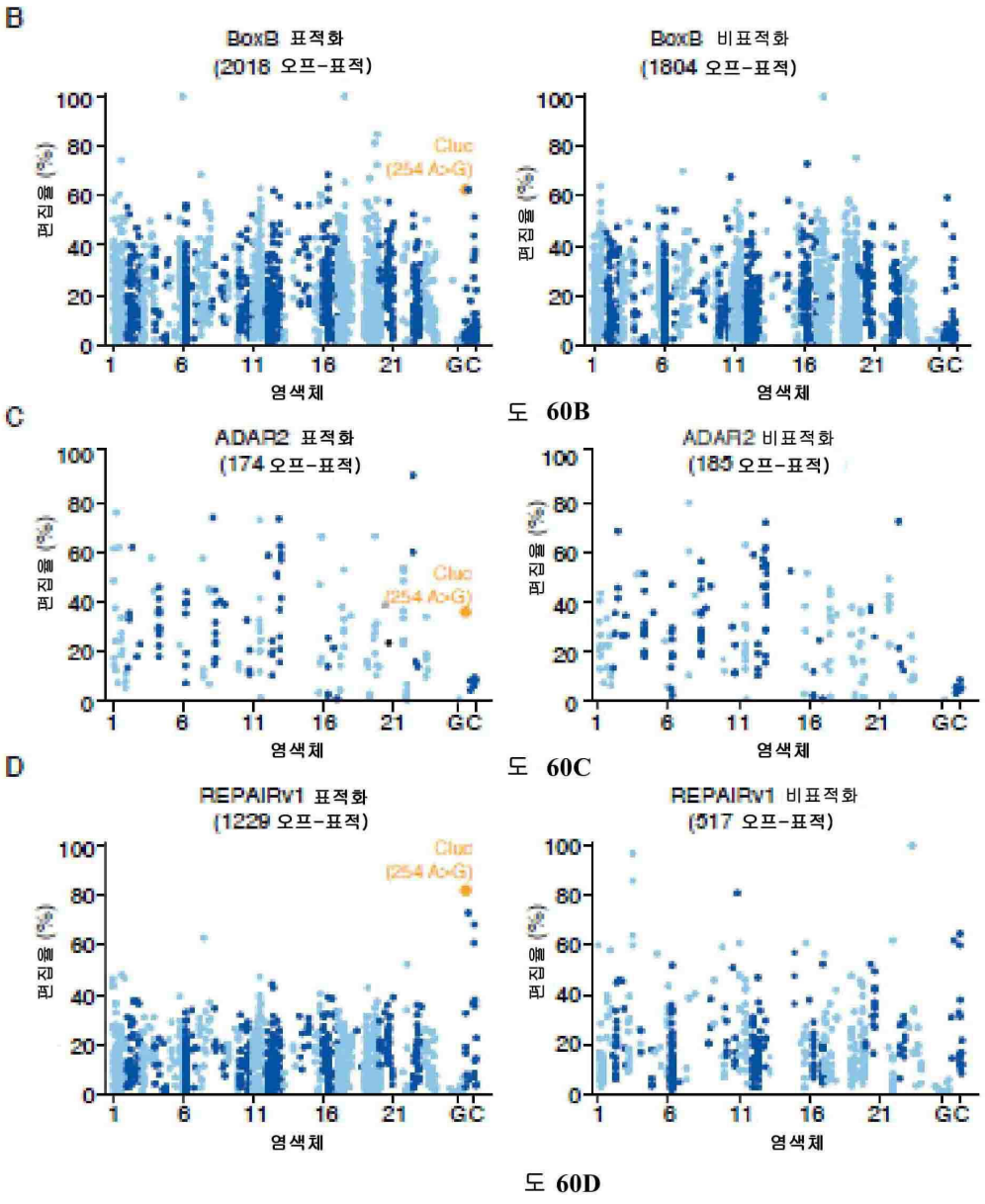
도면59


A
WT (4473 bp)
Δ984-1090 (4152 bp)
Δ1026-1090 (4278 bp)
Δ1053-1090 (4359 bp)
Δ1-125 (4113 bp)
Δ1-88 (4224 bp)
Δ1-72 (4272 bp)
NES
HIV
ADARDD
dCas13b
1
1090
984
1026
1053
125
88
72
표적화 가이드
표적화 RLU
비표적화 가이드
비표적화 RLU
0
1
2
3
4
0.01
0.02
0.03
0.04

도면60a



도면60bcd

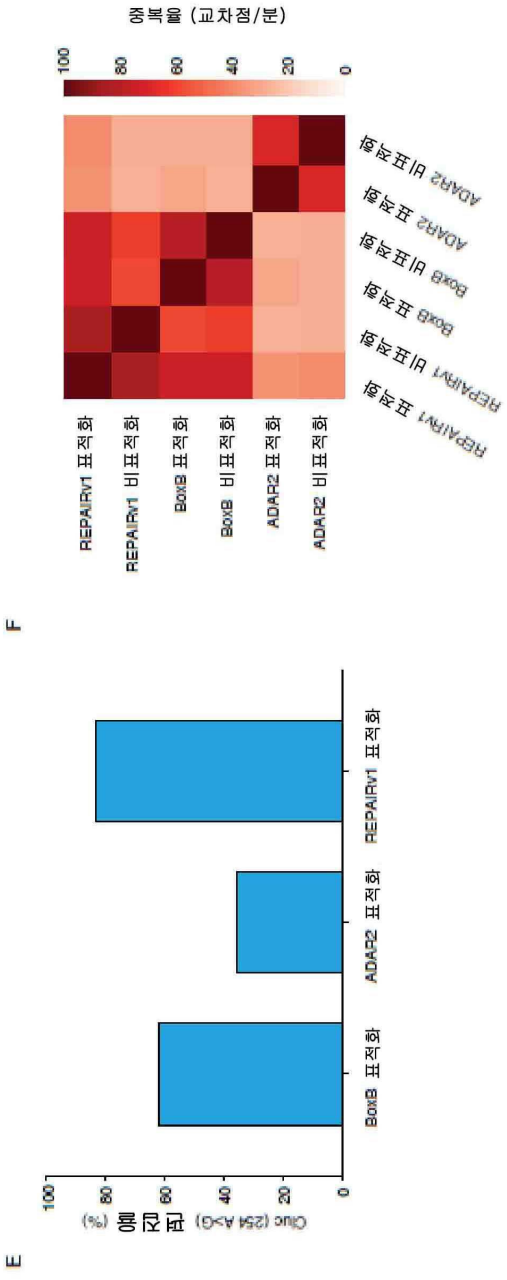


도 60B

도 60C

도 60D

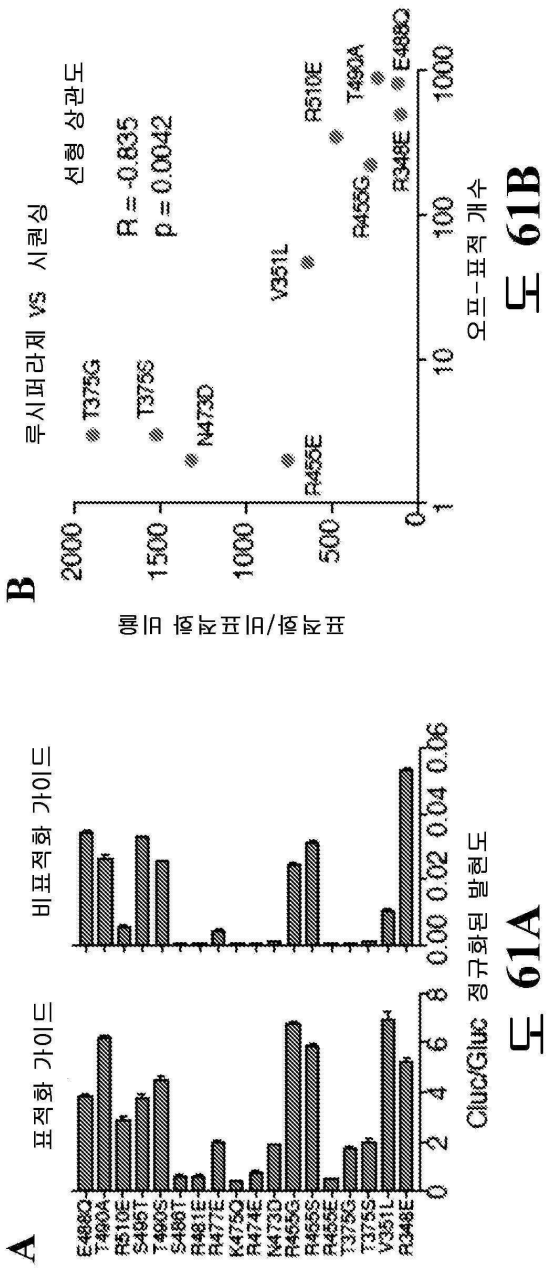
도면60ef



도 60F

도 60E

도면61abc

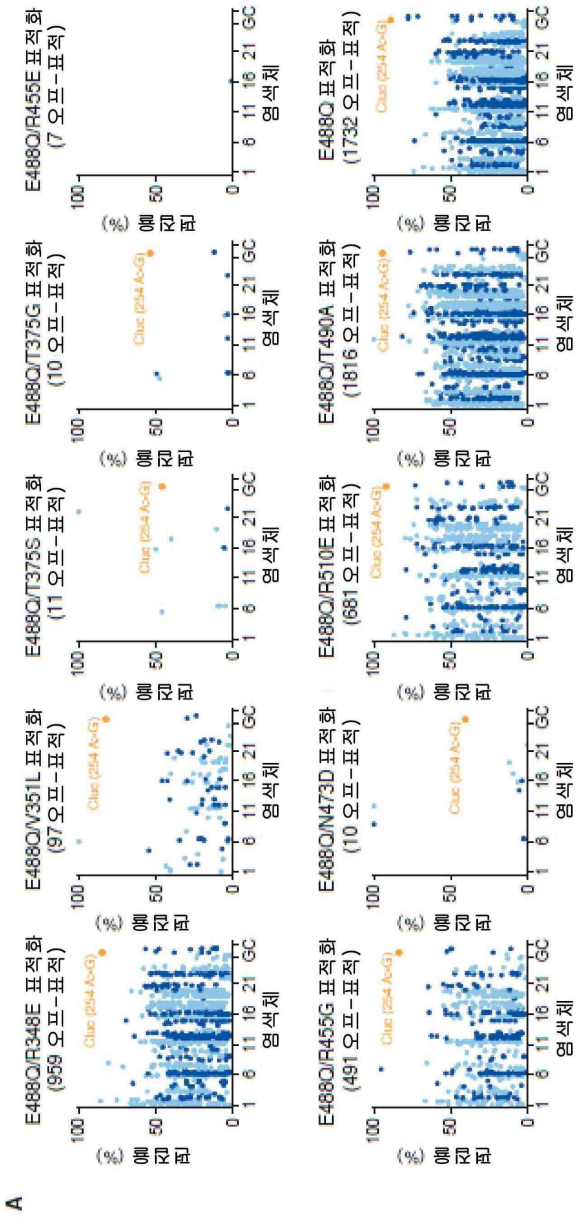


도 61A

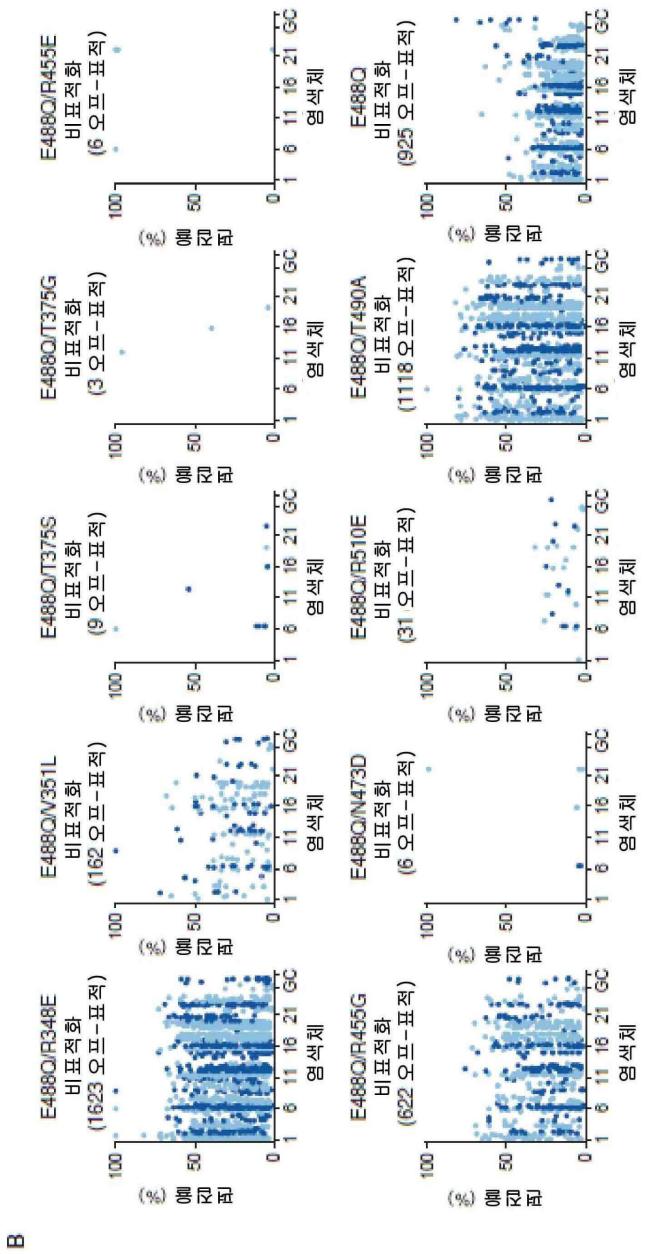
도 61B

도 61C

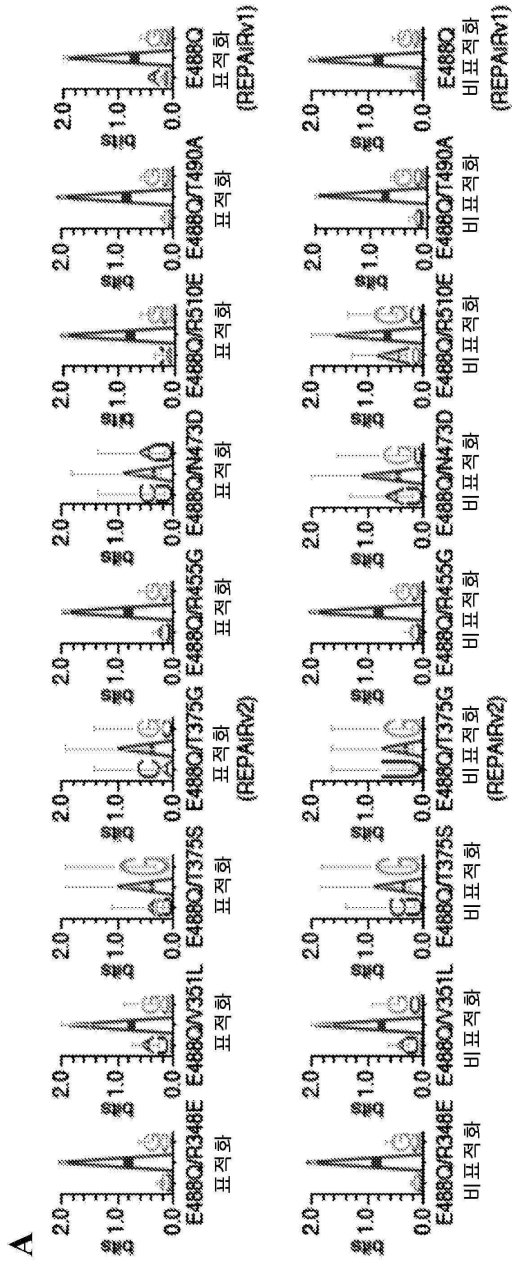
도면62a


A
E488Q/R348E 표적화 (959 오프-표적)
E488Q/V351L 표적화 (97 오프-표적)
E488Q/T375S 표적화 (11 오프-표적)
E488Q/T375G 표적화 (10 오프-표적)
E488Q/R455E 표적화 (7 오프-표적)
E488Q/R455G 표적화 (491 오프-표적)
E488Q/N473D 표적화 (10 오프-표적)
E488Q/R510E 표적화 (681 오프-표적)
E488Q/T490A 표적화 (1816 오프-표적)
E488Q 표적화 (1732 오프-표적)
편집 율 (%)
염색체
Cluc (254 A>G)

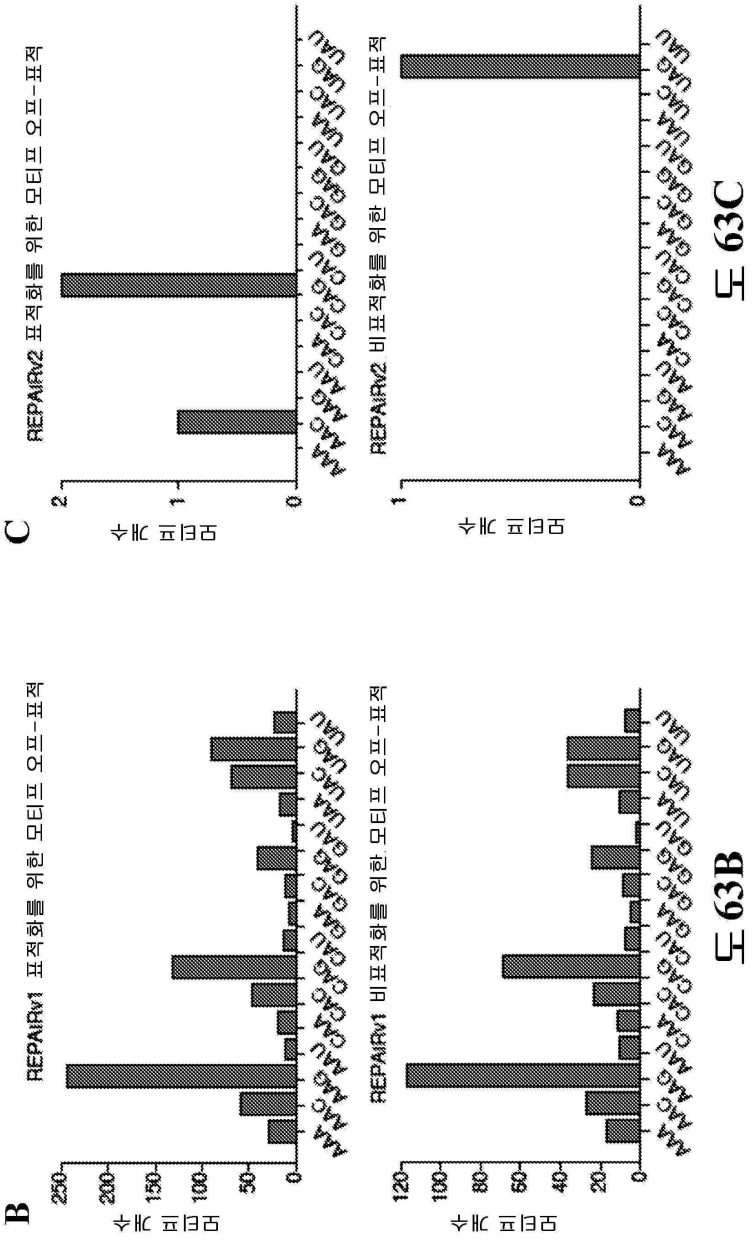
도면62b


B
E488Q/R348E
비표적화
(1623 오프-표적)
E488Q/V351L
비표적화
(162 오프-표적)
E488Q/T375S
비표적화
(9 오프-표적)
E488Q/T375G
비표적화
(3 오프-표적)
E488Q/R455E
비표적화
(6 오프-표적)
E488Q/R455G
비표적화
(622 오프-표적)
E488Q/N473D
비표적화
(6 오프-표적)
E488Q/R510E
비표적화
(31 오프-표적)
E488Q/T490A
비표적화
(1118 오프-표적)
E488Q
비표적화
(925 오프-표적)
편집율 (%)
0
50
100
1
6
11
16
21
GC
염색체

도면63a


A
E488Q/R348E 비표적화
E488Q/V351L 비표적화
E488Q/T375S 비표적화
E488Q/T375G 비표적화
E488Q/R455G 비표적화
E488Q/N473D 비표적화
E488Q/R510E 비표적화
E488Q/T490A 비표적화
E488Q 비표적화
(REPAIRv2)
E488Q/R348E 표적화
E488Q/V351L 표적화
E488Q/T375S 표적화
E488Q/T375G 표적화
E488Q/R455G 표적화
E488Q/N473D 표적화
E488Q/R510E 표적화
E488Q/T490A 표적화
E488Q 표적화
(REPAIRv1)

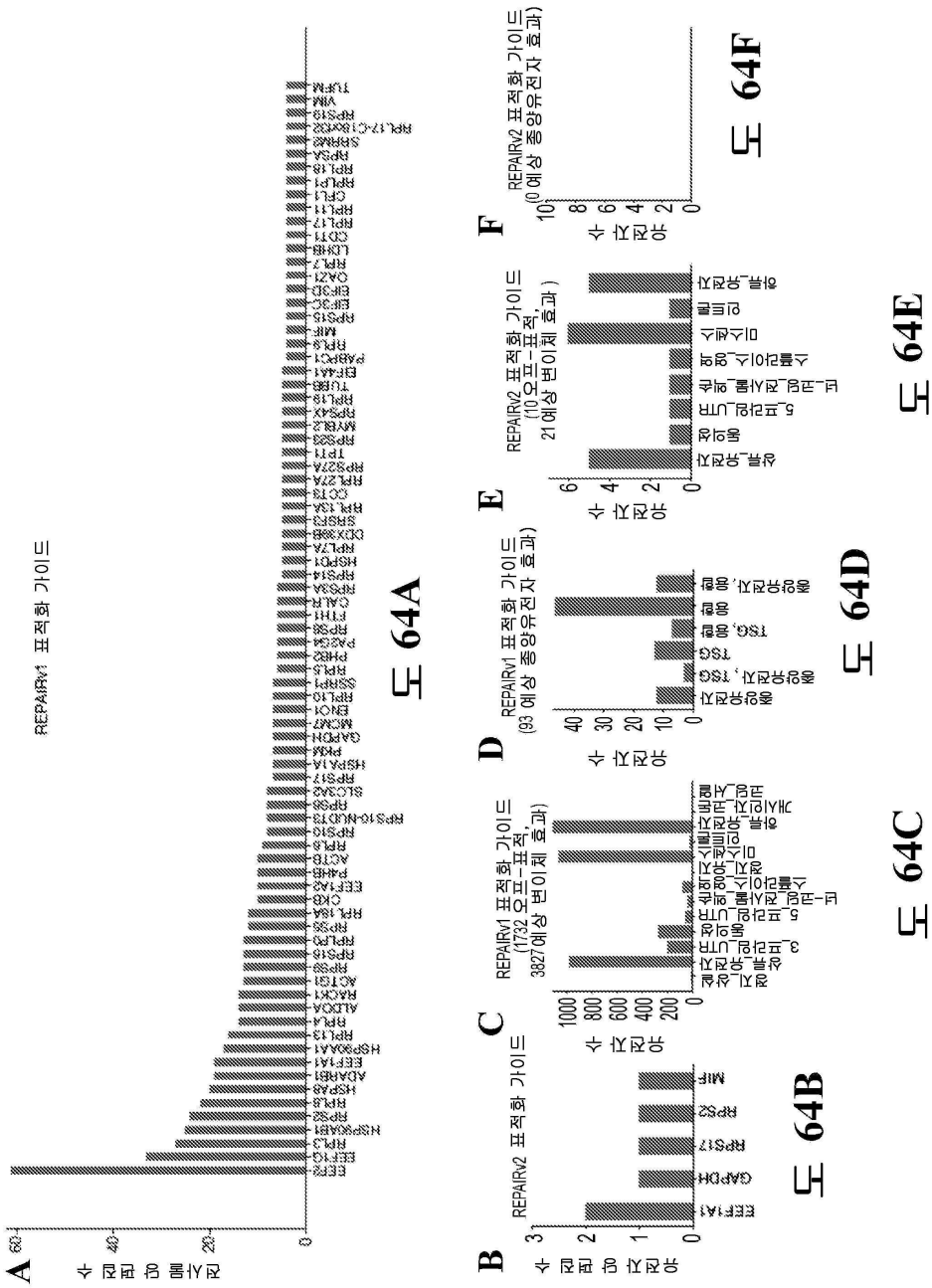
도면63bc



도 63B

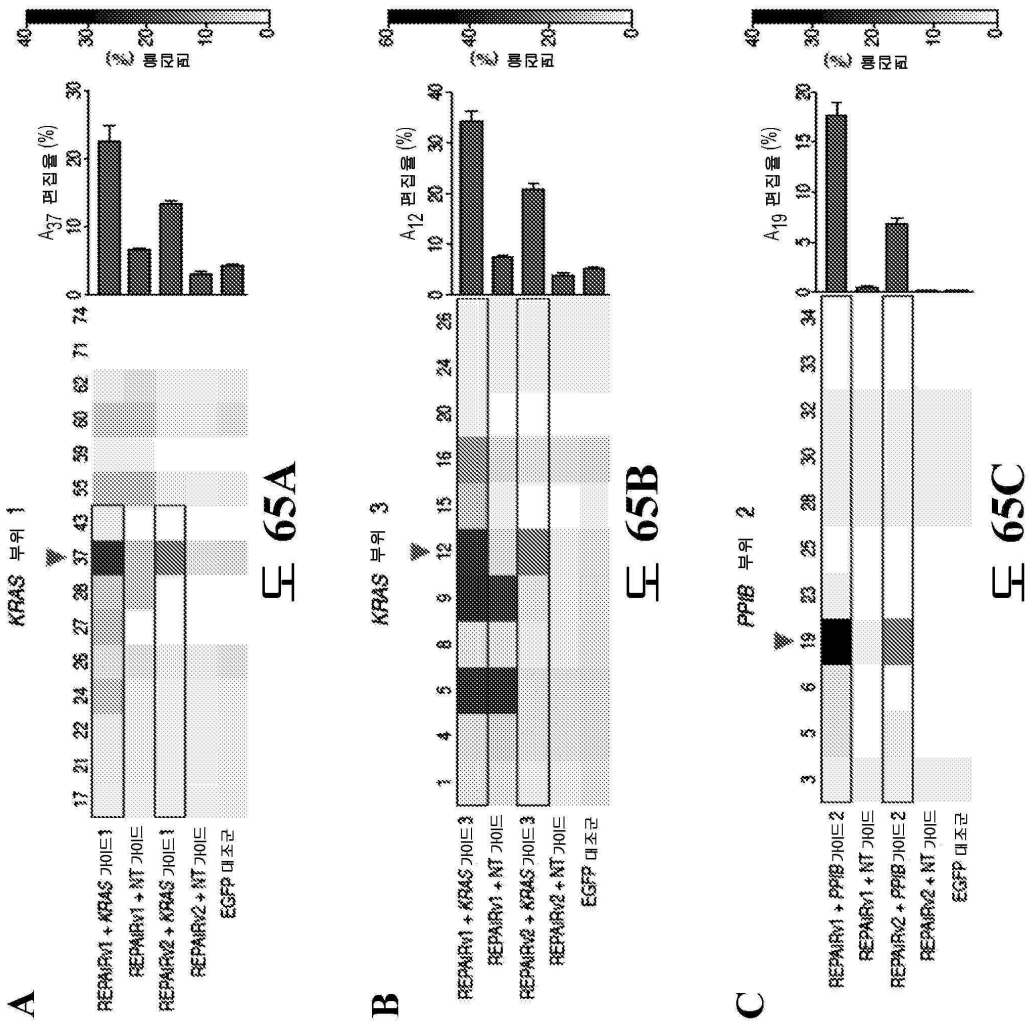
도 63C

도면64



도 64A 도 64B 도 64C 도 64D 도 64E 도 64F

도면65

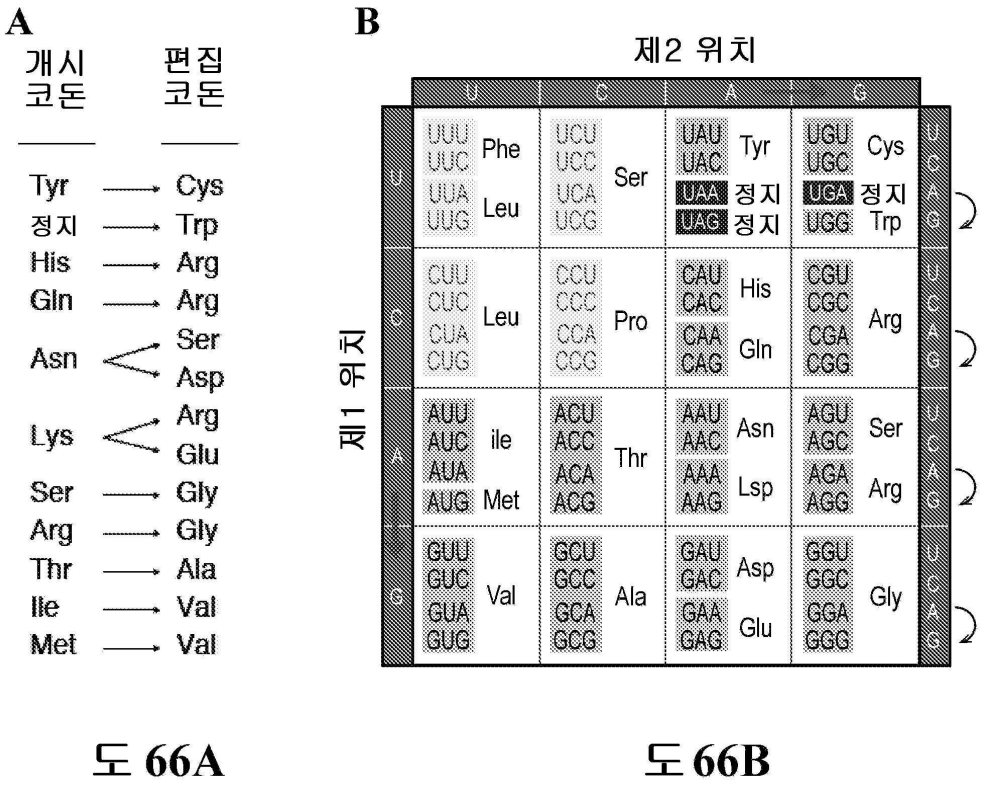


도 65A

도 65B

도 65C

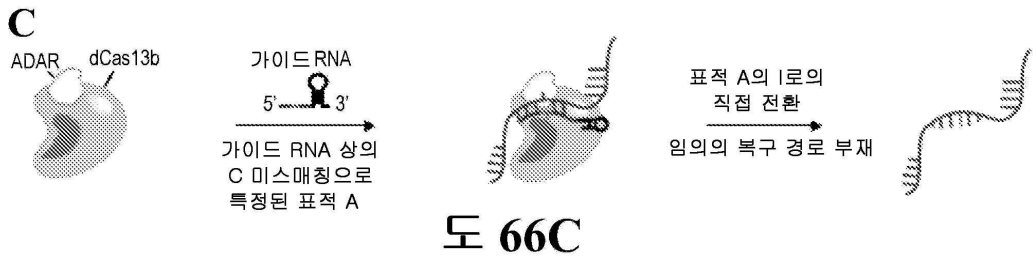
도면66ab



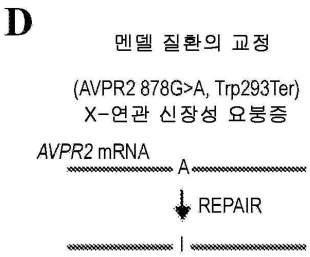
도 66A

도 66B

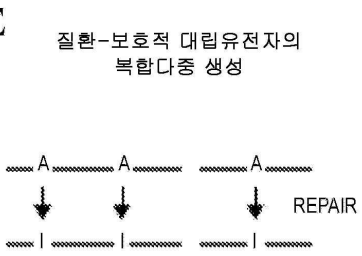
도면66cfg



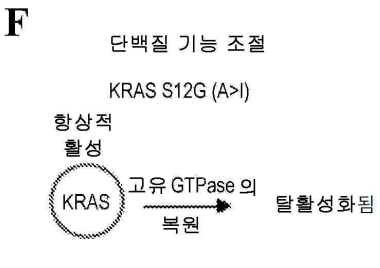
도 66C



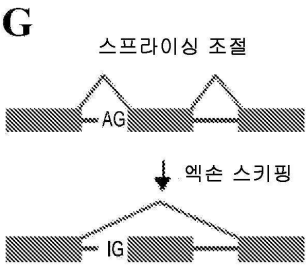
도 66D



도 66E

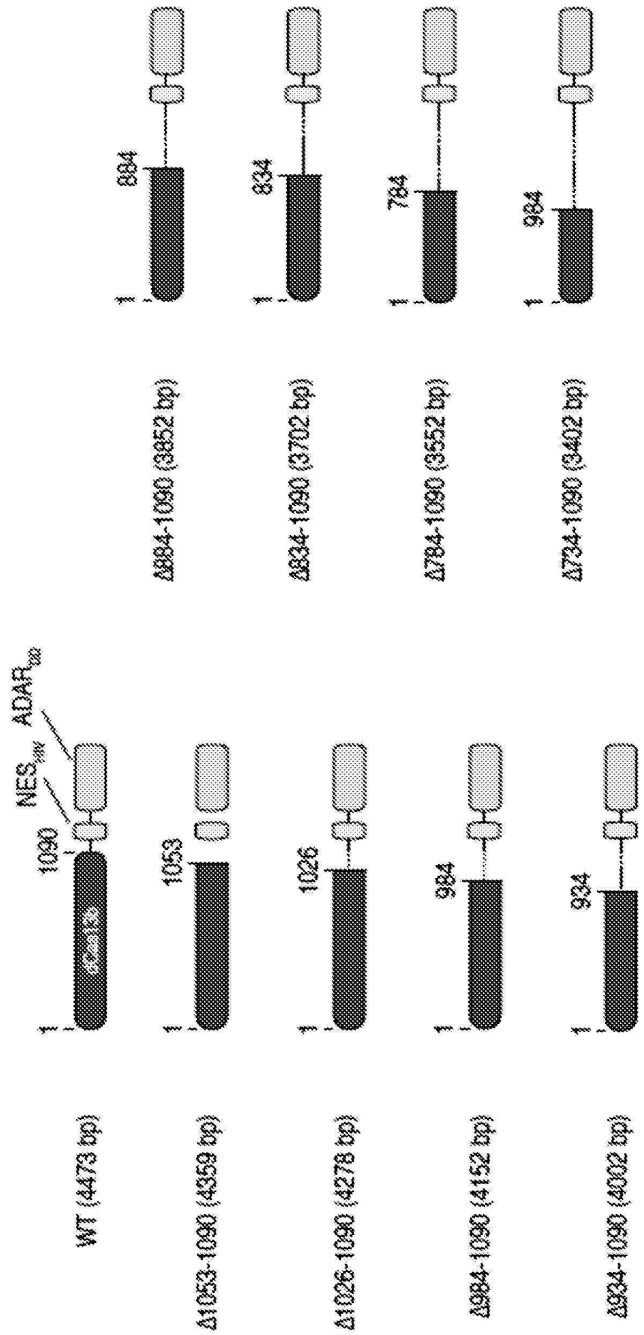


도 66F

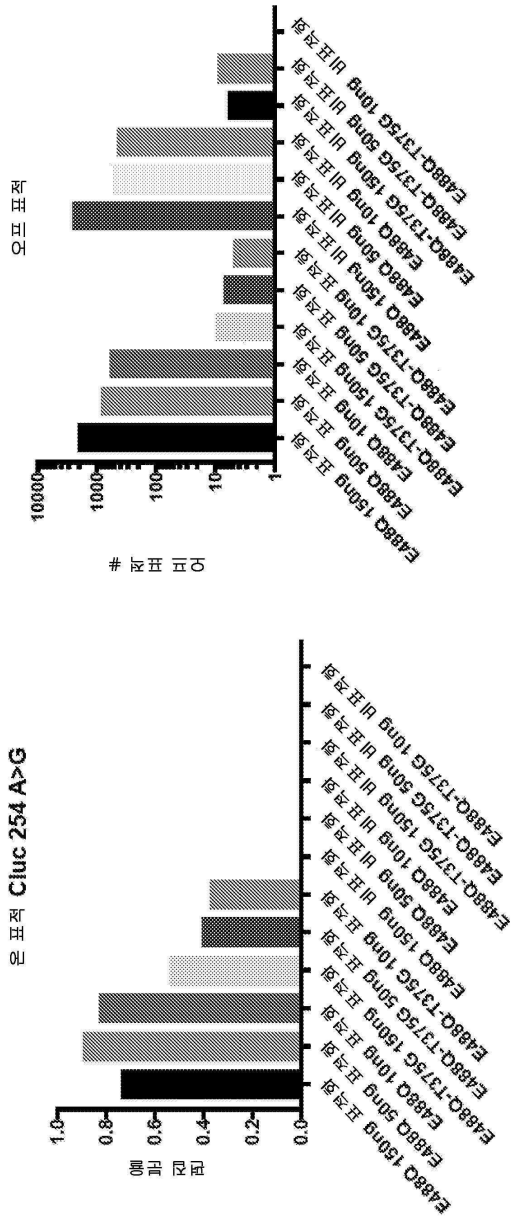


도 66G

도면67

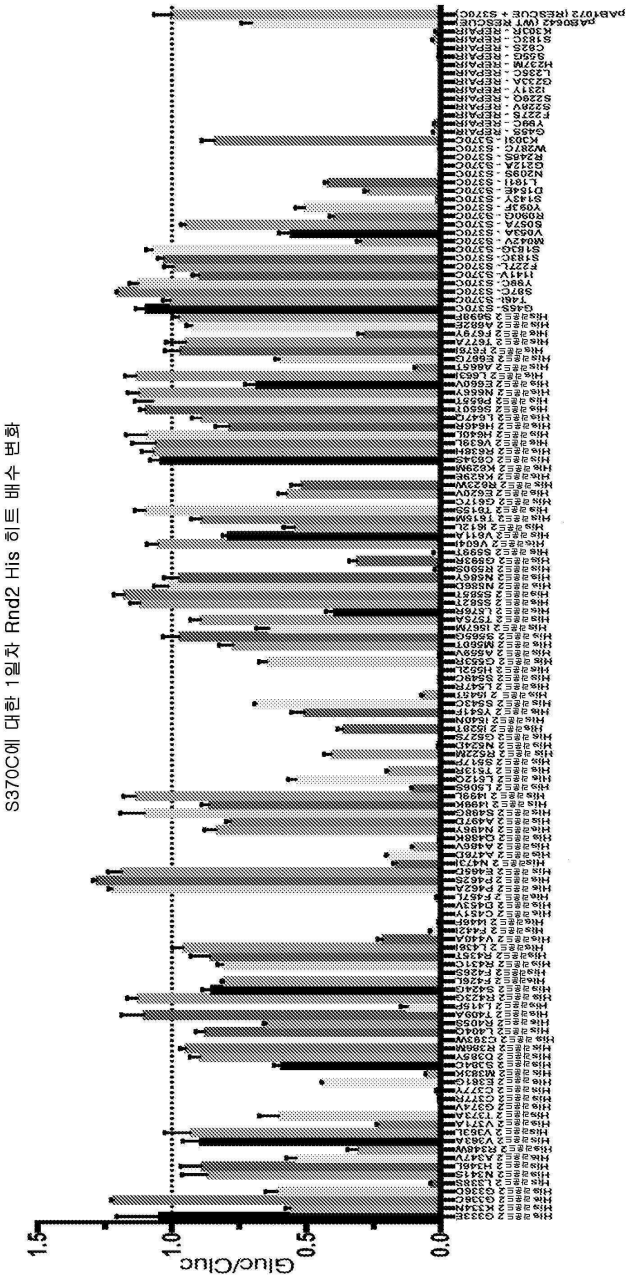

ADAR$_{DD}$
NES$_{mv}$
1090
dCas13
1
WT (4473 bp)
1053
Δ1053-1090 (4359 bp)
1026
Δ1026-1090 (4278 bp)
984
Δ984-1090 (4152 bp)
934
Δ934-1090 (4002 bp)
884
Δ884-1090 (3852 bp)
834
Δ834-1090 (3702 bp)
784
Δ784-1090 (3552 bp)
984
Δ734-1090 (3402 bp)

도면68

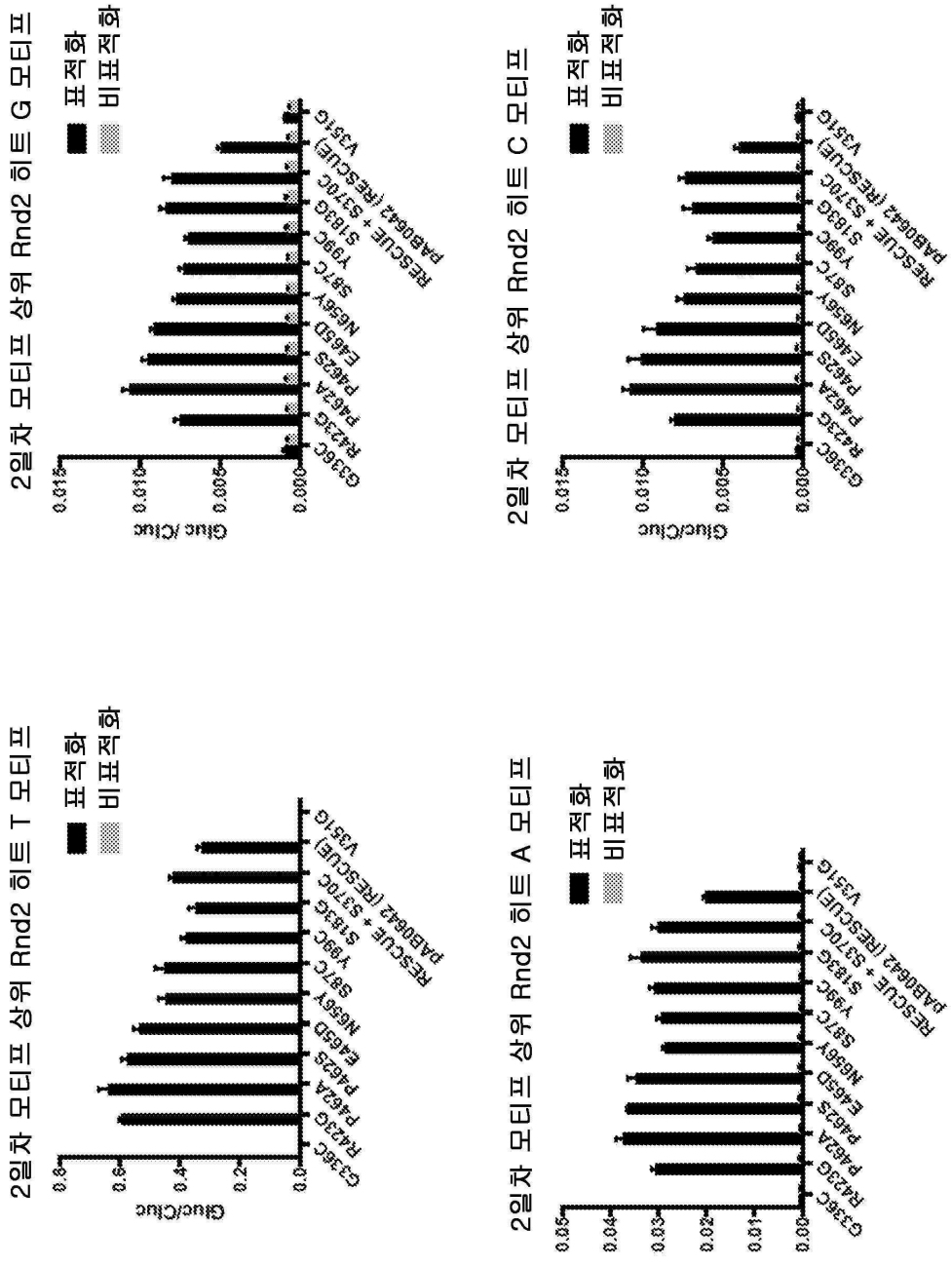

오프 표적
오프 표적 #
10000
1000
100
10
1
E488Q 15ng 표적
E488Q 5ng 표적
E488Q 10ng 표적
E488Q-T375G 15ng 표적
E488Q-T375G 5ng 표적
E488Q 15ng 표적
E488Q 5ng 표적
E488Q 10ng 표적
E488Q-T375G 15ng 표적
E488Q-T375G 5ng 표적
E488Q 15ng 표적
E488Q 5ng 표적
E488Q 10ng 표적
E488Q-T375G 15ng 표적
E488Q-T375G 5ng 표적
E488Q-T375G 10ng 비표적

온 표적 Cluc 254 A>G
편집 분율
1.0
0.8
0.6
0.4
0.2
0.0
E488Q 15ng 표적
E488Q 5ng 표적
E488Q 10ng 표적
E488Q-T375G 15ng 표적
E488Q-T375G 5ng 표적
E488Q 15ng 표적
E488Q 5ng 표적
E488Q 10ng 표적
E488Q-T375G 15ng 표적
E488Q-T375G 5ng 표적
E488Q 15ng 표적
E488Q 5ng 표적
E488Q 10ng 표적
E488Q-T375G 15ng 표적
E488Q-T375G 5ng 표적
E488Q-T375G 10ng 비표적

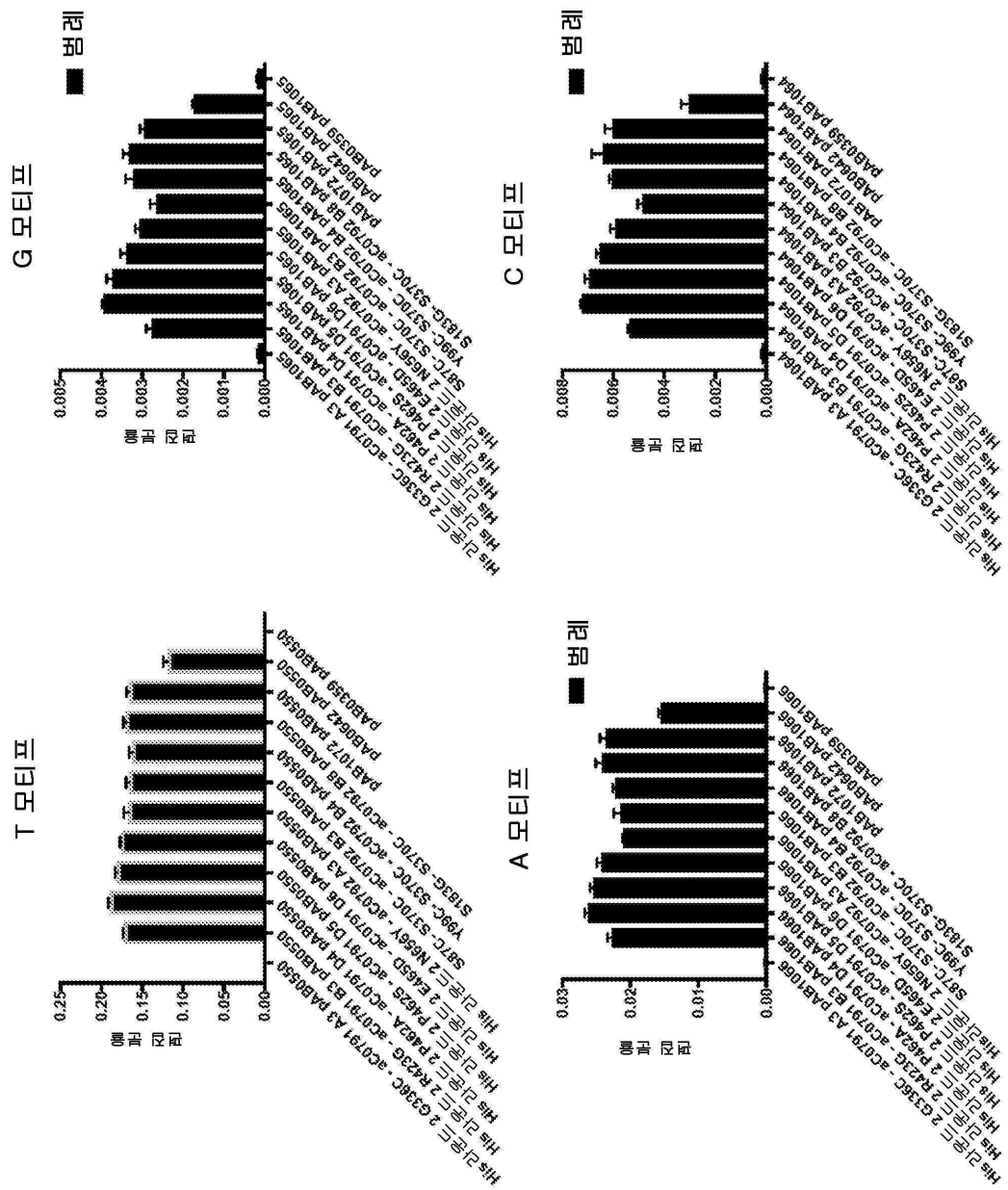
도면69



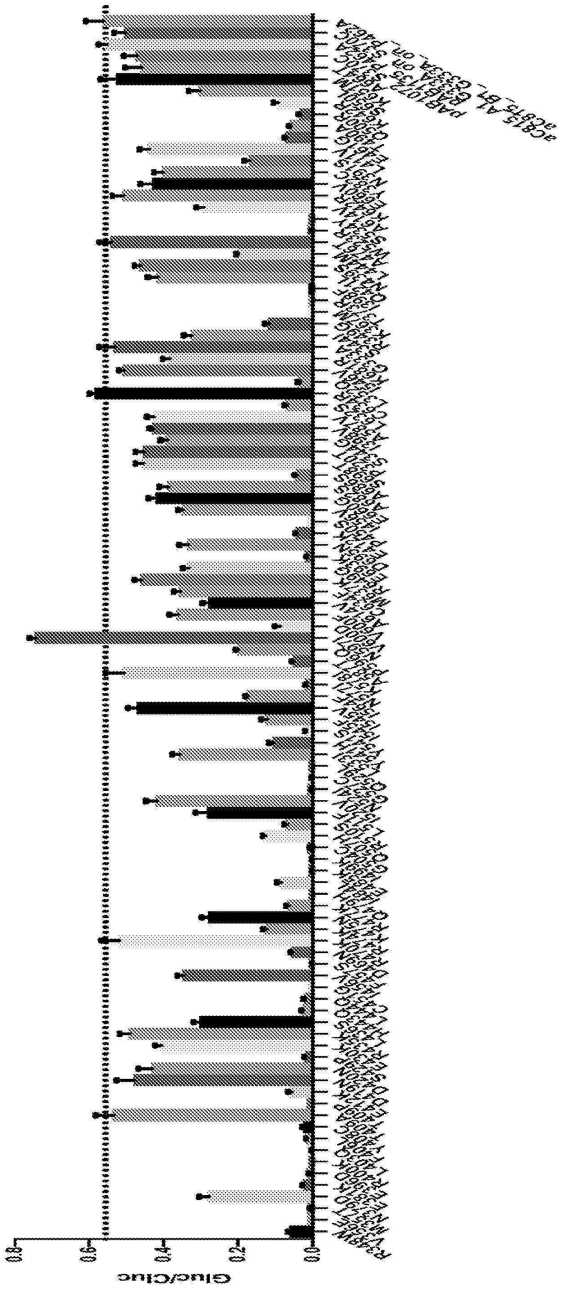
도면70


2일차 모티프 상위 Rnd2 히트 G 모티프
표적화
비표적화
Gluc/Cluc
0.015
0.010
0.005
0.000
G336C
R423G
P462A
P462S
E465D
N656Y
S97C
Y99C
S183G
RESCUE + S370C
PAB0642 (RESCUE)
V351G
2일차 모티프 상위 Rnd2 히트 C 모티프
표적화
비표적화
Gluc/Cluc
0.015
0.010
0.005
0.000
G336C
R423G
P462A
P462S
E465D
N656Y
S97C
Y99C
S183G
RESCUE + S370C
PAB0642 (RESCUE)
V351G
2일차 모티프 상위 Rnd2 히트 T 모티프
표적화
비표적화
Gluc/Cluc
0.8
0.6
0.4
0.2
0.0
G336C
R423G
P462A
P462S
E465D
N656Y
S97C
Y99C
S183G
RESCUE + S370C
PAB0642 (RESCUE)
V351G
2일차 모티프 상위 Rnd2 히트 A 모티프
표적화
비표적화
0.05
0.04
0.03
0.02
0.01
0.00
G336C
R423G
P462A
P462S
E465D
N656Y
S97C
Y99C
S183G
RESCUE + S370C
PAB0642 (RESCUE)
V351G

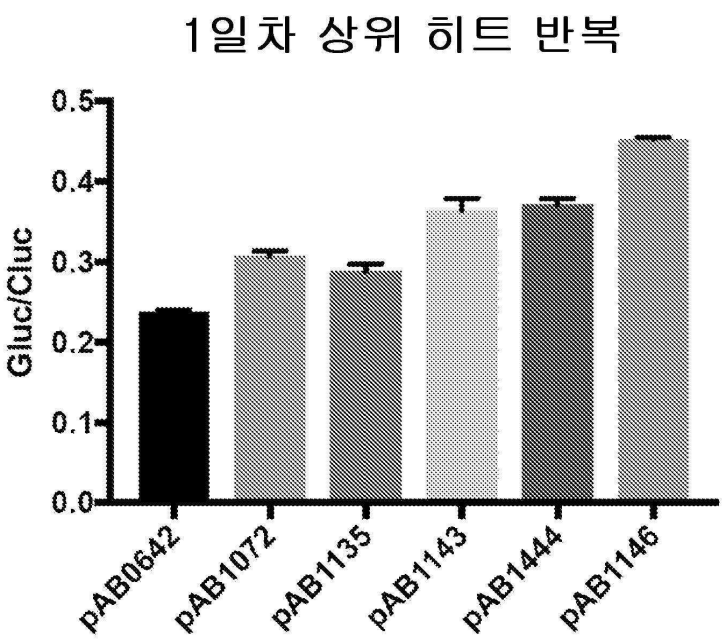
도면71



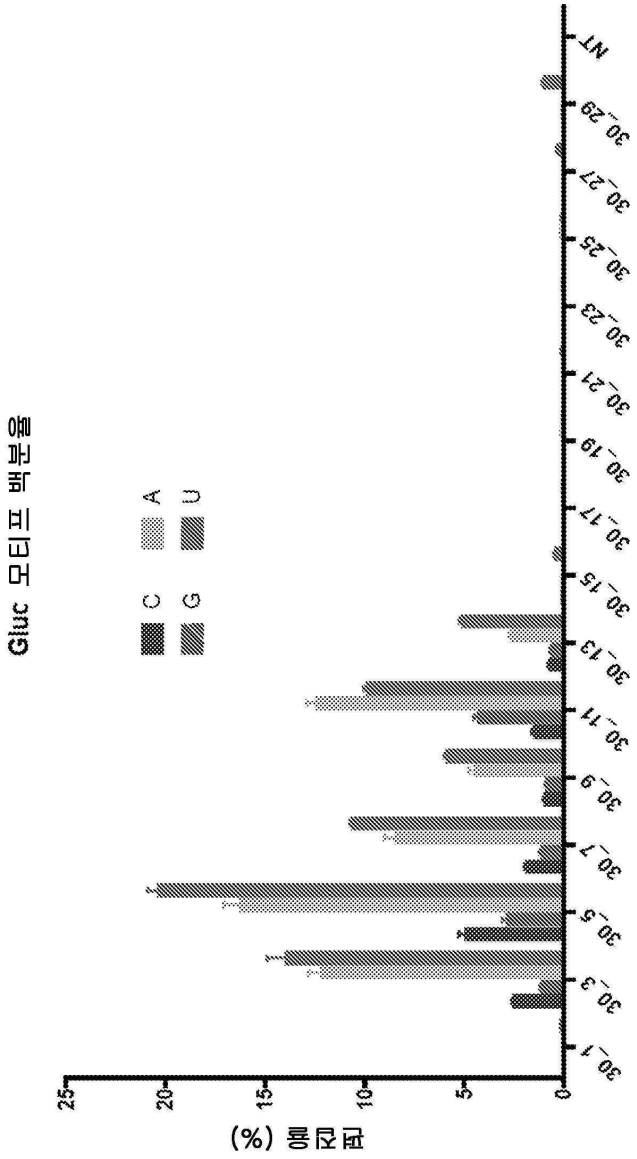
도면72



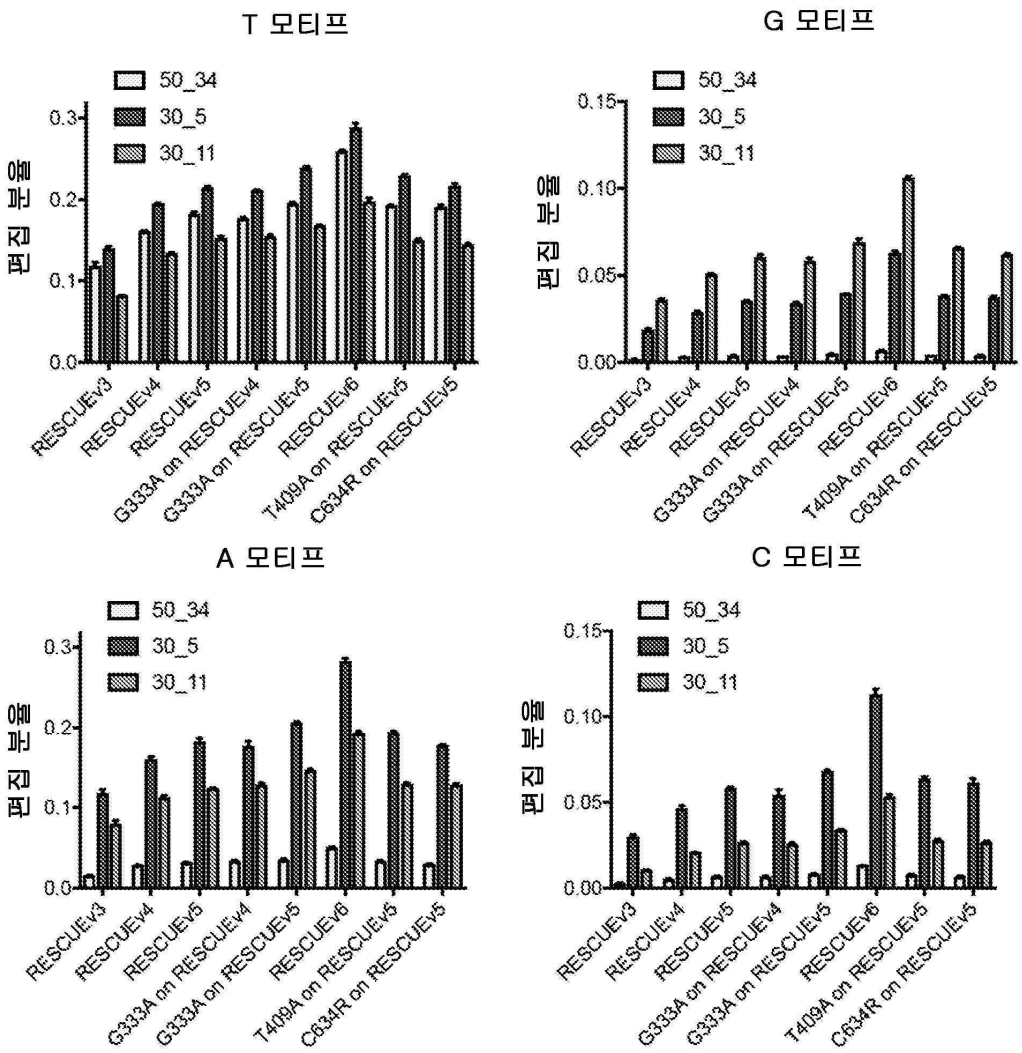
도면73


1일차 상위 히트 반복
Gluc/Cluc
0.5
0.4
0.3
0.2
0.1
0.0
pAB0642
pAB1072
pAB1135
pAB1143
pAB1444
pAB1146

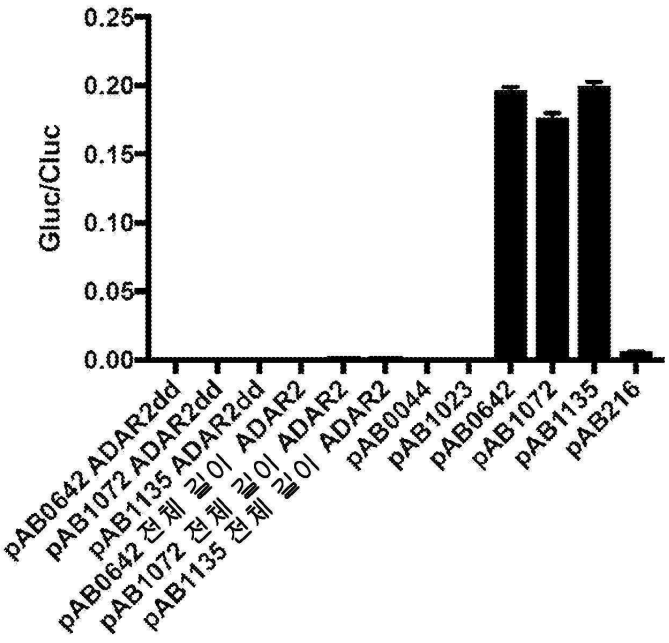
도면74


Gluc 모티프 백분율
A
U
C
G
평균율 (%)
25
20
15
10
5
0
30_1
30_3
30_5
30_7
30_9
30_11
30_13
30_15
30_17
30_19
30_21
30_23
30_25
30_27
30_29
NT

도면75


T 모티프
G 모티프
A 모티프
C 모티프
50_34
30_5
30_11
편집 분율
RESCUEv3
RESCUEv4
RESCUEv5
G333A on RESCUEv4
G333A on RESCUEv5
RESCUEv6
T409A on RESCUEv5
C634R on RESCUEv5

도면76


Gluc/Cluc
pAB0642 ADAR2dd
pAB1072 ADAR2dd
pAB1135 ADAR2dd
pAB0642 전체 길이 ADAR2
pAB1072 전체 길이 ADAR2
pAB1135 전체 길이 ADAR2
pAB0044
pAB1023
pAB0642
pAB1072
pAB1135
pAB216

서 열 목 록

SEQUENCE LISTING

<110> Massachusetts Institute of Technology

The Broad Institute, Inc.

The President and Fellows of Harvard College

Gootenberg, Jonathan

Abudayyeh, Omar

Zhang, Feng

<120> SYSTEMS, METHODS, AND COMPOSITIONS FOR TARGETED NUCLEIC ACID EDITING

<130> BROD-2345WP

<150> 62/568,313

<151> 2017-10-04

<160> 805

<170> PatentIn version 3.5

<210> 1

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 1

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 2

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 2

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser

20

<210> 3

<211> 25

<212> PRT

<213> Artificial Sequence

<220

><223> Synthetic Peptide

<400> 3

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser

20 25

<210> 4

<211> 30

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 4

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

20 25 30

<210> 5

<211> 35

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 5

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly

20 25 30

Gly Gly Ser

35

<210> 6

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 6

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly

20 25 30

Gly Gly Ser Gly Gly Gly Gly Ser

35 40

<210> 7

<211> 45

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 7

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly

20 25 30

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

35 40 45

<210> 8

<211> 50

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 8

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly

20 25 30

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly

35 40 45

Gly Ser

50

<210> 9

<211> 55

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 9

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly

20 25 30

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly

35 40 45

Gly Ser Gly Gly Gly Gly Ser

50 55

<210> 10

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 10

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly

1 5 10 15

<210> 11

<211> 32

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 11

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Gln Ser Gly Ala Leu Thr Arg His Gln Arg Thr His Thr Arg

20 25 30

<210> 12

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 12

Gly Gly Ser Gly Gly Ser Gly Gly Ser

1 5

<210> 13

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 13

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 14

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 14

Gly Gly Gly Gly Ser

1 5

<210> 15

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 15

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10

<210> 16

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 16

Pro Lys Lys Lys Arg Lys Val Glu Ala Ser Ser Pro Lys Lys Arg Lys

1 5 10 15

Val Glu Ala Ser

20

<210> 17

<211> 7

<212> PRT

<213> SV40 virus

<400> 17

Pro Lys Lys Lys Arg Lys Val

1 5

<210> 18

<211> 9

<212> PRT

<213> SV40 virus

<400> 18

Pro Lys Lys Lys Arg Lys Val Glu Ala

1 5

<210> 19

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 19

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

1 5 10 15

<210> 20

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 20

Pro Ala Ala Lys Arg Val Lys Leu Asp

1 5

<210> 21

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 21

Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro

1 5 10

<210> 22

<211> 38

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 22

Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly

1 5 10 15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro

20 25 30

Arg Asn Gln Gly Gly Tyr

35

<210> 23

<211> 42

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 23

Arg Met Arg Ile Glx Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu
1 5 10 15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys
20 25 30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val

35 40

<210> 24

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 24

Val Ser Arg Lys Arg Pro Arg Pro
1 5

<210> 25

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 25

Pro Pro Lys Lys Ala Arg Glu Asp
1 5

<210> 26

<211> 8

<212> PRT

<213> Homo sapiens

<400> 26

Pro Gln Pro Lys Lys Lys Pro Leu
1 5

<210> 27

<211> 12

<212> PRT

<213> Mus musculus

<400> 27

Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro

1 5 10

<210> 28

<211> 5

<212> PRT

<213> Influenza virus

<400> 28

Asp Arg Leu Arg Arg

1 5

<210> 29

<211> 7

<212> PRT

<213> Influenza virus

<400> 29

Pro Lys Gln Lys Lys Arg Lys

1 5

<210> 30

<211> 10

<212> PRT

<213> Hepatitis delta virus

<400> 30

Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu

1 5 10

<210> 31

<211> 10

<212> PRT

<213> Mus musculus

<400> 31

Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg

1 5 10

<210> 32

<211> 20

<212> PRT

<213> Homo sapiens

<400> 32

Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys

1 5 10 15

Lys Ser Lys Lys

20

<210> 33

<211> 17

<212> PRT

<213> Homo sapiens

<400> 33

Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys

1 5 10 15

Lys

<210> 34

<211> 1340

<212> PRT

<213> Lachnospiraceae bacterium

<400> 34

Met Gln Ile Ser Lys Val Asn His Lys His Val Ala Val Gly Gln Lys

1 5 10 15

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

20 25 30

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

35 40 45

Val Leu Phe Asn Val Phe Asn Thr Lys Asp Leu Tyr Asp Ser Gln Glu

50 55 60

Ser Asp Lys Ser Glu Lys Asp Lys Glu Ile Ile Ser Lys Gly Ala Lys

65 70 75 80

Phe Val Ala Lys Ser Phe Asn Ser Ala Ile Thr Ile Leu Lys Lys Gln

85 90 95

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

100 105 110

Lys Asp Lys Phe Gly Gly Ala Arg Ile Tyr Asp Asp Asp Ile Glu Glu

115 120 125

Ala Leu Thr Glu Thr Leu Lys Lys Ser Phe Arg Lys Glu Asn Val Arg

130 135 140

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

145 150 155 160

Leu Ser Lys Asp Glu Glu Glu Leu Ile Gln Glu Tyr Phe Val Lys Gln

165 170 175

Leu Val Glu Glu Tyr Thr Lys Thr Lys Leu Gln Lys Asn Val Val Lys

180 185 190

Ser Ile Lys Asn Gln Asn Met Val Ile Gln Pro Asp Ser Asp Ser Gln

195 200 205

Val Leu Ser Leu Ser Glu Ser Arg Arg Glu Lys Gln Ser Ser Ala Val

210 215 220

Ser Ser Asp Thr Leu Val Asn Cys Lys Glu Lys Asp Val Leu Lys Ala

225 230 235 240

Phe Leu Thr Asp Tyr Ala Val Leu Asp Glu Asp Glu Arg Asn Ser Leu

245 250 255

Leu Trp Lys Leu Arg Asn Leu Val Asn Leu Tyr Phe Tyr Gly Ser Glu

260 265 270

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

275 280 285

His Asp Glu Gln Lys Ala Asn Lys Thr Leu Phe Ile Asp Glu Ile Cys

290 295 300

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

305 310 315 320

Tyr Glu Glu Asn Arg Ser Arg Cys Arg Lys Gln Asn Ile Asn Tyr Tyr

325 330 335

Arg Ser Ala Leu Asn Tyr Ala Lys Asn Asn Thr Ser Gly Ile Phe Glu
340 345 350
Asn Glu Asp Ser Asn His Phe Trp Ile His Leu Ile Glu Asn Glu Val
355 360 365

Glu Arg Leu Tyr Asn Gly Ile Glu Asn Gly Glu Glu Phe Lys Phe Glu
370 375 380
Thr Gly Tyr Ile Ser Glu Lys Val Trp Lys Ala Val Ile Asn His Leu
385 390 395 400
Ser Ile Lys Tyr Ile Ala Leu Gly Lys Ala Val Tyr Asn Tyr Ala Met
405 410 415
Lys Glu Leu Ser Ser Pro Gly Asp Ile Glu Pro Gly Lys Ile Asp Asp
420 425 430

Ser Tyr Ile Asn Gly Ile Thr Ser Phe Asp Tyr Glu Ile Ile Lys Ala
435 440 445
Glu Glu Ser Leu Gln Arg Asp Ile Ser Met Asn Val Val Phe Ala Thr
450 455 460
Asn Tyr Leu Ala Cys Ala Thr Val Asp Thr Asp Lys Asp Phe Leu Leu
465 470 475 480
Phe Ser Lys Glu Asp Ile Arg Ser Cys Thr Lys Lys Asp Gly Asn Leu
485 490 495

Cys Lys Asn Ile Met Gln Phe Trp Gly Gly Tyr Ser Thr Trp Lys Asn
500 505 510
Phe Cys Glu Glu Tyr Leu Lys Asp Asp Lys Asp Ala Leu Glu Leu Leu
515 520 525
Tyr Ser Leu Lys Ser Met Leu Tyr Ser Met Arg Asn Ser Ser Phe His
530 535 540
Phe Ser Thr Glu Asn Val Asp Asn Gly Ser Trp Asp Thr Glu Leu Ile
545 550 555 560

Gly Lys Leu Phe Glu Glu Asp Cys Asn Arg Ala Ala Arg Ile Glu Lys
565 570 575
Glu Lys Phe Tyr Asn Asn Asn Leu His Met Phe Tyr Ser Ser Ser Leu

580 585 590

Leu Glu Lys Val Leu Glu Arg Leu Tyr Ser Ser His His Glu Arg Ala

595 600 605

Ser Gln Val Pro Ser Phe Asn Arg Val Phe Val Arg Lys Asn Phe Pro

610 615 620

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

625 630 635 640

Asp Glu Gln Ile Trp Gln Ser Ala Val Tyr Tyr Leu Cys Lys Glu Ile

645 650 655

Tyr Tyr Asn Asp Phe Leu Gln Ser Lys Glu Ala Tyr Lys Leu Phe Arg

660 665 670

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

675 680 685

Ala Asp Ser Phe Lys Gln Ala Val Val Tyr Tyr Gly Lys Ala Ile Gly

690 695 700

Asn Ala Thr Leu Ser Gln Val Cys Gln Ala Ile Met Thr Glu Tyr Asn

705 710 715 720

Arg Gln Asn Asn Asp Gly Leu Lys Lys Lys Ser Ala Tyr Ala Glu Lys

725 730 735

Gln Asn Ser Asn Lys Tyr Lys His Tyr Pro Leu Phe Leu Lys Gln Val

740 745 750

Leu Gln Ser Ala Phe Trp Glu Tyr Leu Asp Glu Asn Lys Glu Ile Tyr

755 760 765

Gly Phe Ile Ser Ala Gln Ile His Lys Ser Asn Val Glu Ile Lys Ala

770 775 780

Glu Asp Phe Ile Ala Asn Tyr Ser Ser Gln Gln Tyr Lys Lys Leu Val

785 790 795 800

Asp Lys Val Lys Lys Thr Pro Glu Leu Gln Lys Trp Tyr Thr Leu Gly

805 810 815

Arg Leu Ile Asn Pro Arg Gln Ala Asn Gln Phe Leu Gly Ser Ile Arg

820 825 830

Asn Tyr Val Gln Phe Val Lys Asp Ile Gln Arg Arg Ala Lys Glu Asn
835 840 845
Gly Asn Pro Ile Arg Asn Tyr Tyr Glu Val Leu Glu Ser Asp Ser Ile
850 855 860
Ile Lys Ile Leu Glu Met Cys Thr Lys Leu Asn Gly Thr Thr Ser Asn
865 870 875 880

Asp Ile His Asp Tyr Phe Arg Asp Glu Asp Glu Tyr Ala Glu Tyr Ile
885 890 895
Ser Gln Phe Val Asn Phe Gly Asp Val His Ser Gly Ala Ala Leu Asn
900 905 910
Ala Phe Cys Asn Ser Glu Ser Glu Gly Lys Lys Asn Gly Ile Tyr Tyr
915 920 925
Asp Gly Ile Asn Pro Ile Val Asn Arg Asn Trp Val Leu Cys Lys Leu
930 935 940

Tyr Gly Ser Pro Asp Leu Ile Ser Lys Ile Ile Ser Arg Val Asn Glu
945 950 955 960
Asn Met Ile His Asp Phe His Lys Gln Glu Asp Leu Ile Arg Glu Tyr
965 970 975
Gln Ile Lys Gly Ile Cys Ser Asn Lys Lys Glu Gln Gln Asp Leu Arg
980 985 990
Thr Phe Gln Val Leu Lys Asn Arg Val Glu Leu Arg Asp Ile Val Glu
995 1000 1005

Tyr Ser Glu Ile Ile Asn Glu Leu Tyr Gly Gln Leu Ile Lys Trp
1010 1015 1020
Cys Tyr Leu Arg Glu Arg Asp Leu Met Tyr Phe Gln Leu Gly Phe
1025 1030 1035
His Tyr Leu Cys Leu Asn Asn Ala Ser Ser Lys Glu Ala Asp Tyr
1040 1045 1050
Ile Lys Ile Asn Val Asp Asp Arg Asn Ile Ser Gly Ala Ile Leu
1055 1060 1065

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

1070 1075 1080

Lys Lys Asp Asp Met Tyr Val Ala Leu Lys Ser Gly Lys Lys Ala

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

Gln Leu Tyr Leu Ala Gly Leu Glu Leu Phe Glu Asn Val Ala Glu

1130 1135 1140

His Glu Asn Ile Ile Ile Phe Arg Asn Glu Ile Asp His Phe His

1145 1150 1155

Tyr Phe Tyr Asp Arg Asp Arg Ser Met Leu Asp Leu Tyr Ser Glu

1160 1165 1170

Val Phe Asp Arg Phe Phe Thr Tyr Asp Met Lys Leu Arg Lys Asn

1175 1180 1185

Val Val Asn Met Leu Tyr Asn Ile Leu Leu Asp His Asn Ile Val

1190 1195 1200

Ser Ser Phe Val Phe Glu Thr Gly Glu Lys Lys Val Gly Arg Gly

1205 1210 1215

Asp Ser Glu Val Ile Lys Pro Ser Ala Lys Ile Arg Leu Arg Ala

1220 1225 1230

Asn Asn Gly Val Ser Ser Asp Val Phe Thr Tyr Lys Val Gly Ser

1235 1240 1245

Lys Asp Glu Leu Lys Ile Ala Thr Leu Pro Ala Lys Asn Glu Glu

1250 1255 1260

Phe Leu Leu Asn Val Ala Arg Leu Ile Tyr Tyr Pro Asp Met Glu

1265 1270 1275

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

1280 1285 1290

Lys Ser Asn Asp Lys Lys Gly Lys Ile Ser Arg Gly Ser Asn Thr

1295 1300 1305

Arg Ser Ser Asn Gln Ser Lys Tyr Asn Asn Lys Ser Lys Asn Arg

1310 1315 1320

Met Asn Tyr Ser Met Gly Ser Ile Phe Glu Lys Met Asp Leu Lys

1325 1330 1335

Phe Asp

1340

<210> 35

<211> 1437

<212> PRT

<213> Lachnospiraceae bacterium

<400> 35

Met Lys Ile Ser Lys Val Arg Glu Glu Asn Arg Gly Ala Lys Leu Thr

1 5 10 15

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

20 25 30

Ile Leu Tyr Asn Asp Pro Ser Arg Tyr Gly Lys Ser Arg Lys Asn Asp

35 40 45

Glu Asp Arg Asp Arg Tyr Ile Glu Ser Arg Leu Lys Ser Ser Gly Lys

50 55 60

Leu Tyr Arg Ile Phe Asn Glu Asp Lys Asn Lys Arg Glu Thr Asp Glu

65 70 75 80

Leu Gln Trp Phe Leu Ser Glu Ile Val Lys Lys Ile Asn Arg Arg Asn

85 90 95

Gly Leu Val Leu Ser Asp Met Leu Ser Val Asp Asp Arg Ala Phe Glu

100 105 110

Lys Ala Phe Glu Lys Tyr Ala Glu Leu Ser Tyr Thr Asn Arg Arg Asn

115 120 125

Lys Val Ser Gly Ser Pro Ala Phe Glu Thr Cys Gly Val Asp Ala Ala

130 135 140

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

145 150 155 160

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

165 170 175

Arg Ile Ile Ala Lys Tyr Leu Lys Lys Ser Leu Cys Arg Glu Arg Val

180 185 190

Lys Arg Gly Leu Lys Lys Leu Leu Met Asn Ala Phe Asp Leu Pro Tyr

195 200 205

Ser Asp Pro Asp Ile Asp Val Gln Arg Asp Phe Ile Asp Tyr Val Leu

210 215 220

Glu Asp Phe Tyr His Val Arg Ala Lys Ser Gln Val Ser Arg Ser Ile

225 230 235 240

Lys Asn Met Asn Met Pro Val Gln Pro Glu Gly Asp Gly Lys Phe Ala

245 250 255

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

260 265 270

Glu Lys Glu Ala Phe Lys Lys Phe Leu Ser Asp Tyr Ala Ser Leu Asp

275 280 285

Glu Arg Val Arg Asp Asp Met Leu Arg Arg Met Arg Arg Leu Val Val

290 295 300

Leu Tyr Phe Tyr Gly Ser Asp Asp Ser Lys Leu Ser Asp Val Asn Glu

305 310 315 320

Lys Phe Asp Val Trp Glu Asp His Ala Ala Arg Arg Val Asp Asn Arg

325 330 335

Glu Phe Ile Lys Leu Pro Leu Glu Asn Lys Leu Ala Asn Gly Lys Thr

340 345 350

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

355 360 365

Arg Asn Gln Asn Ile Gly Cys Tyr Arg Gln Ala Val Lys Ala Val Glu

370 375 380

Glu Asp Asn Asn Gly Arg Tyr Phe Asp Asp Lys Met Leu Asn Met Phe

385 390 395 400

Phe Ile His Arg Ile Glu Tyr Gly Val Glu Lys Ile Tyr Ala Asn Leu

405 410 415

Lys Gln Val Thr Glu Phe Lys Ala Arg Thr Gly Tyr Leu Ser Glu Lys
420 425 430
Ile Trp Lys Asp Leu Ile Asn Tyr Ile Ser Ile Lys Tyr Ile Ala Met
435 440 445
Gly Lys Ala Val Tyr Asn Tyr Ala Met Asp Glu Leu Asn Ala Ser Asp
450 455 460
Lys Lys Glu Ile Glu Leu Gly Lys Ile Ser Glu Glu Tyr Leu Ser Gly

465 470 475 480
Ile Ser Ser Phe Asp Tyr Glu Leu Ile Lys Ala Glu Glu Met Leu Gln
485 490 495
Arg Glu Thr Ala Val Tyr Val Ala Phe Ala Ala Arg His Leu Ser Ser
500 505 510
Gln Thr Val Glu Leu Asp Ser Glu Asn Ser Asp Phe Leu Leu Leu Lys
515 520 525
Pro Lys Gly Thr Met Asp Lys Asn Asp Lys Asn Lys Leu Ala Ser Asn

530 535 540
Asn Ile Leu Asn Phe Leu Lys Asp Lys Glu Thr Leu Arg Asp Thr Ile
545 550 555 560
Leu Gln Tyr Phe Gly Gly His Ser Leu Trp Thr Asp Phe Pro Phe Asp
565 570 575
Lys Tyr Leu Ala Gly Gly Lys Asp Asp Val Asp Phe Leu Thr Asp Leu
580 585 590
Lys Asp Val Ile Tyr Ser Met Arg Asn Asp Ser Phe His Tyr Ala Thr

595 600 605
Glu Asn His Asn Asn Gly Lys Trp Asn Lys Glu Leu Ile Ser Ala Met
610 615 620
Phe Glu His Glu Thr Glu Arg Met Thr Val Val Met Lys Asp Lys Phe
625 630 635 640
Tyr Ser Asn Asn Leu Pro Met Phe Tyr Lys Asn Asp Asp Leu Lys Lys
645 650 655
Leu Leu Ile Asp Leu Tyr Lys Asp Asn Val Glu Arg Ala Ser Gln Val

660 665 670

Pro Ser Phe Asn Lys Val Phe Val Arg Lys Asn Phe Pro Ala Leu Val

675 680 685

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

690 695 700

Asp Lys Gly Glu Asn Glu Leu Lys Phe Tyr Asn Ala Leu Tyr Tyr Met

705 710 715 720

Phe Lys Glu Ile Tyr Tyr Asn Ala Phe Leu Asn Asp Lys Asn Val Arg

725 730 735

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

740 745 750

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

755 760 765

Glu Lys Lys Lys Leu Arg Glu Gln Leu Gln Asn Tyr Ile Ala Glu Asn

770 775 780

Asp Phe Gly Gln Arg Ile Lys Asn Ile Val Gln Val Asn Pro Asp Tyr

785 790 795 800

Thr Leu Ala Gln Ile Cys Gln Leu Ile Met Thr Glu Tyr Asn Gln Gln

805 810 815

Asn Asn Gly Cys Met Gln Lys Lys Ser Ala Ala Arg Lys Asp Ile Asn

820 825 830

Lys Asp Ser Tyr Gln His Tyr Lys Met Leu Leu Leu Val Asn Leu Arg

835 840 845

Lys Ala Phe Leu Glu Phe Ile Lys Glu Asn Tyr Ala Phe Val Leu Lys

850 855 860

Pro Tyr Lys His Asp Leu Cys Asp Lys Ala Asp Phe Val Pro Asp Phe

865 870 875 880

Ala Lys Tyr Val Lys Pro Tyr Ala Gly Leu Ile Ser Arg Val Ala Gly

885 890 895

Ser Ser Glu Leu Gln Lys Trp Tyr Ile Val Ser Arg Phe Leu Ser Pro

900 905 910

Ala Gln Ala Asn His Met Leu Gly Phe Leu His Ser Tyr Lys Gln Tyr

915 920 925

Val Trp Asp Ile Tyr Arg Arg Ala Ser Glu Thr Gly Thr Glu Ile Asn

930 935 940

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

945 950 955 960

Asp Ala Val Ile Asp Leu Ser Val Lys Leu Cys Gly Thr Ile Ser Ser

965 970 975

Glu Ile Ser Asp Tyr Phe Lys Asp Asp Glu Val Tyr Ala Glu Tyr Ile

980 985 990

Ser Ser Tyr Leu Asp Phe Glu Tyr Asp Gly Gly Asn Tyr Lys Asp Ser

995 1000 1005

Leu Asn Arg Phe Cys Asn Ser Asp Ala Val Asn Asp Gln Lys Val

1010 1015 1020

Ala Leu Tyr Tyr Asp Gly Glu His Pro Lys Leu Asn Arg Asn Ile

1025 1030 1035

Ile Leu Ser Lys Leu Tyr Gly Glu Arg Arg Phe Leu Glu Lys Ile

1040 1045 1050

Thr Asp Arg Val Ser Arg Ser Asp Ile Val Glu Tyr Tyr Lys Leu

1055 1060 1065

Lys Lys Glu Thr Ser Gln Tyr Gln Thr Lys Gly Ile Phe Asp Ser

1070 1075 1080

Glu Asp Glu Gln Lys Asn Ile Lys Lys Phe Gln Glu Met Lys Asn

1085 1090 1095

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

1100 1105 1110

Glu Leu Gln Gly Gln Leu Ile Asn Trp Ile Tyr Leu Arg Glu Arg

1115 1120 1125

Asp Leu Met Asn Phe Gln Leu Gly Tyr His Tyr Ala Cys Leu Asn

1130 1135 1140

Asn Asp Ser Asn Lys Gln Ala Thr Tyr Val Thr Leu Asp Tyr Gln

1145 1150 1155

Gly Lys Lys Asn Arg Lys Ile Asn Gly Ala Ile Leu Tyr Gln Ile

1160 1165 1170

Cys Ala Met Tyr Ile Asn Gly Leu Pro Leu Tyr Tyr Val Asp Lys

1175 1180 1185

Asp Ser Ser Glu Trp Thr Val Ser Asp Gly Lys Glu Ser Thr Gly

1190 1195 1200

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

1205 1210 1215

Thr Ser Asp Cys Tyr Ala Ser Gly Leu Glu Ile Phe Glu Asn Ile

1220 1225 1230

Ser Glu His Asp Asn Ile Thr Glu Leu Arg Asn Tyr Ile Glu His

1235 1240 1245

Phe Arg Tyr Tyr Ser Ser Phe Asp Arg Ser Phe Leu Gly Ile Tyr

1250 1255 1260

Ser Glu Val Phe Asp Arg Phe Phe Thr Tyr Asp Leu Lys Tyr Arg

1265 1270 1275

Lys Asn Val Pro Thr Ile Leu Tyr Asn Ile Leu Leu Gln His Phe

1280 1285 1290

Val Asn Val Arg Phe Glu Phe Val Ser Gly Lys Lys Met Ile Gly

1295 1300 1305

Ile Asp Lys Lys Asp Arg Lys Ile Ala Lys Glu Lys Glu Cys Ala

1310 1315 1320

Arg Ile Thr Ile Arg Glu Lys Asn Gly Val Tyr Ser Glu Gln Phe

1325 1330 1335

Thr Tyr Lys Leu Lys Asn Gly Thr Val Tyr Val Asp Ala Arg Asp

1340 1345 1350

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

1355 1360 1365

Lys Val Asn Met Asp Glu Met Ile Glu Val Lys Glu Lys Lys Lys

1370 1375 1380

Pro Ser Asp Asn Asn Thr Gly Lys Gly Tyr Ser Lys Arg Asp Arg

1385 1390 1395

Gln Gln Asp Arg Lys Glu Tyr Asp Lys Tyr Lys Glu Lys Lys Lys

1400 1405 1410

Lys Glu Gly Asn Phe Leu Ser Gly Met Gly Gly Asn Ile Asn Trp

1415 1420 1425

Asp Glu Ile Asn Ala Gln Leu Lys Asn

1430 1435

<210> 36

<211> 1385

<212> PRT

<213> Clostridium aminophilum

<400> 36

Met Lys Phe Ser Lys Val Asp His Thr Arg Ser Ala Val Gly Ile Gln

1 5 10 15

Lys Ala Thr Asp Ser Val His Gly Met Leu Tyr Thr Asp Pro Lys Lys

20 25 30

Gln Glu Val Asn Asp Leu Asp Lys Arg Phe Asp Gln Leu Asn Val Lys

35 40 45

Ala Lys Arg Leu Tyr Asn Val Phe Asn Gln Ser Lys Ala Glu Glu Asp

50 55 60

Asp Asp Glu Lys Arg Phe Gly Lys Val Val Lys Lys Leu Asn Arg Glu

65 70 75 80

Leu Lys Asp Leu Leu Phe His Arg Glu Val Ser Arg Tyr Asn Ser Ile

85 90 95

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

100 105 110

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

115 120 125

Pro Gln Lys Val Ile Ser Lys Leu Leu Leu Tyr Tyr Leu Arg Lys Gly

130 135 140

Leu Lys Pro Gly Thr Asp Gly Leu Arg Met Ile Leu Glu Ala Ser Cys

145 150 155 160

Gly Leu Arg Lys Leu Ser Gly Asp Glu Lys Glu Leu Lys Val Phe Leu

165 170 175

Gln Thr Leu Asp Glu Asp Phe Glu Lys Lys Thr Phe Lys Lys Asn Leu

180 185 190

Ile Arg Ser Ile Glu Asn Gln Asn Met Ala Val Gln Pro Ser Asn Glu

195 200 205

Gly Asp Pro Ile Ile Gly Ile Thr Gln Gly Arg Phe Asn Ser Gln Lys

210 215 220

Asn Glu Glu Lys Ser Ala Ile Glu Arg Met Met Ser Met Tyr Ala Asp

225 230 235 240

Leu Asn Glu Asp His Arg Glu Asp Val Leu Arg Lys Leu Arg Arg Leu

245 250 255

Asn Val Leu Tyr Phe Asn Val Asp Thr Glu Lys Thr Glu Glu Pro Thr

260 265 270

Leu Pro Gly Glu Val Asp Thr Asn Pro Val Phe Glu Val Trp His Asp

275 280 285

His Glu Lys Gly Lys Glu Asn Asp Arg Gln Phe Ala Thr Phe Ala Lys

290 295 300

Ile Leu Thr Glu Asp Arg Glu Thr Arg Lys Lys Glu Lys Leu Ala Val

305 310 315 320

Lys Glu Ala Leu Asn Asp Leu Lys Ser Ala Ile Arg Asp His Asn Ile

325 330 335

Met Ala Tyr Arg Cys Ser Ile Lys Val Thr Glu Gln Asp Lys Asp Gly

340 345 350

Leu Phe Phe Glu Asp Gln Arg Ile Asn Arg Phe Trp Ile His His Ile

355 360 365

Glu Ser Ala Val Glu Arg Ile Leu Ala Ser Ile Asn Pro Glu Lys Leu

370 375 380

Tyr Lys Leu Arg Ile Gly Tyr Leu Gly Glu Lys Val Trp Lys Asp Leu

385 390 395 400

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

405 410 415

His Phe Ala Met Glu Asp Leu Gly Lys Thr Gly Gln Asp Ile Glu Leu

420 425 430

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

435 440 445

Glu Gln Ile Arg Ala Asp Glu Thr Leu Gln Arg Gln Leu Ser Val Glu

450 455 460

Val Ala Phe Ala Ala Asn Asn Leu Phe Arg Ala Val Val Gly Gln Thr

465 470 475 480

Gly Lys Lys Ile Glu Gln Ser Lys Ser Glu Glu Asn Glu Glu Asp Phe

485 490 495

Leu Leu Trp Lys Ala Glu Lys Ile Ala Glu Ser Ile Lys Lys Glu Gly

500 505 510

Glu Gly Asn Thr Leu Lys Ser Ile Leu Gln Phe Phe Gly Gly Ala Ser

515 520 525

Ser Trp Asp Leu Asn His Phe Cys Ala Ala Tyr Gly Asn Glu Ser Ser

530 535 540

Ala Leu Gly Tyr Glu Thr Lys Phe Ala Asp Asp Leu Arg Lys Ala Ile

545 550 555 560

Tyr Ser Leu Arg Asn Glu Thr Phe His Phe Thr Thr Leu Asn Lys Gly

565 570 575

Ser Phe Asp Trp Asn Ala Lys Leu Ile Gly Asp Met Phe Ser His Glu

580 585 590

Ala Ala Thr Gly Ile Ala Val Glu Arg Thr Arg Phe Tyr Ser Asn Asn

595 600 605

Leu Pro Met Phe Tyr Arg Glu Ser Asp Leu Lys Arg Ile Met Asp His

610 615 620

Leu Tyr Asn Thr Tyr His Pro Arg Ala Ser Gln Val Pro Ser Phe Asn

625 630 635 640

Ser Val Phe Val Arg Lys Asn Phe Arg Leu Phe Leu Ser Asn Thr Leu
645 650 655
Asn Thr Asn Thr Ser Phe Asp Thr Glu Val Tyr Gln Lys Trp Glu Ser

660 665 670
Gly Val Tyr Tyr Leu Phe Lys Glu Ile Tyr Tyr Asn Ser Phe Leu Pro
675 680 685
Ser Gly Asp Ala His His Leu Phe Phe Glu Gly Leu Arg Arg Ile Arg
690 695 700
Lys Glu Ala Asp Asn Leu Pro Ile Val Gly Lys Glu Ala Lys Lys Arg
705 710 715 720
Asn Ala Val Gln Asp Phe Gly Arg Arg Cys Asp Glu Leu Lys Asn Leu

725 730 735
Ser Leu Ser Ala Ile Cys Gln Met Ile Met Thr Glu Tyr Asn Glu Gln
740 745 750
Asn Asn Gly Asn Arg Lys Val Lys Ser Thr Arg Glu Asp Lys Arg Lys
755 760 765
Pro Asp Ile Phe Gln His Tyr Lys Met Leu Leu Leu Arg Thr Leu Gln
770 775 780
Glu Ala Phe Ala Ile Tyr Ile Arg Arg Glu Glu Phe Lys Phe Ile Phe

785 790 795 800
Asp Leu Pro Lys Thr Leu Tyr Val Met Lys Pro Val Glu Glu Phe Leu
805 810 815
Pro Asn Trp Lys Ser Gly Met Phe Asp Ser Leu Val Glu Arg Val Lys
820 825 830
Gln Ser Pro Asp Leu Gln Arg Trp Tyr Val Leu Cys Lys Phe Leu Asn
835 840 845
Gly Arg Leu Leu Asn Gln Leu Ser Gly Val Ile Arg Ser Tyr Ile Gln

850 855 860
Phe Ala Gly Asp Ile Gln Arg Arg Ala Lys Ala Asn His Asn Arg Leu
865 870 875 880
Tyr Met Asp Asn Thr Gln Arg Val Glu Tyr Tyr Ser Asn Val Leu Glu

885 890 895

Val Val Asp Phe Cys Ile Lys Gly Thr Ser Arg Phe Ser Asn Val Phe

900 905 910

Ser Asp Tyr Phe Arg Asp Glu Asp Ala Tyr Ala Asp Tyr Leu Asp Asn

915 920 925

Tyr Leu Gln Phe Lys Asp Glu Lys Ile Ala Glu Val Ser Ser Phe Ala

930 935 940

Ala Leu Lys Thr Phe Cys Asn Glu Glu Glu Val Lys Ala Gly Ile Tyr

945 950 955 960

Met Asp Gly Glu Asn Pro Val Met Gln Arg Asn Ile Val Met Ala Lys

965 970 975

Leu Phe Gly Pro Asp Glu Val Leu Lys Asn Val Val Pro Lys Val Thr

980 985 990

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

995 1000 1005

Tyr Arg Gln Asn Gly Tyr Cys Lys Ser Glu Glu Asp Gln Lys Lys

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Leu Gly Phe His Tyr Leu Cys Leu His Asn Asp Thr Glu Lys Pro

1070 1075 1080

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

1085 1090 1095

Asn Ala Ile Leu His Gln Val Ala Ala Met Tyr Val Gly Gly Leu

1100 1105 1110

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

1115 1120 1125

Gly Glu Ala Asp Cys Lys Leu Ser Ile Ser Lys Asp Thr Ala Gly

1130 1135 1140

Ala Gly Lys Lys Ile Lys Asp Phe Phe Arg Tyr Ser Lys Tyr Val

1145 1150 1155

Leu Ile Lys Asp Arg Met Leu Thr Asp Gln Asn Gln Lys Tyr Thr

1160 1165 1170

Ile Tyr Leu Ala Gly Leu Glu Leu Phe Glu Asn Thr Asp Glu His

1175 1180 1185

Asp Asn Ile Thr Asp Val Arg Lys Tyr Val Asp His Phe Lys Tyr

1190 1195 1200

Tyr Ala Thr Ser Asp Glu Asn Ala Met Ser Ile Leu Asp Leu Tyr

1205 1210 1215

Ser Glu Ile His Asp Arg Phe Phe Thr Tyr Asp Met Lys Tyr Gln

1220 1225 1230

Lys Asn Val Ala Asn Met Leu Glu Asn Ile Leu Leu Arg His Phe

1235 1240 1245

Val Leu Ile Arg Pro Glu Phe Phe Thr Gly Ser Lys Lys Val Gly

1250 1255 1260

Glu Gly Lys Lys Ile Thr Cys Lys Ala Arg Ala Gln Ile Glu Ile

1265 1270 1275

Ala Glu Asn Gly Met Arg Ser Glu Asp Phe Thr Tyr Lys Leu Ser

1280 1285 1290

Asp Gly Lys Lys Asn Ile Ser Thr Cys Met Ile Ala Ala Arg Asp

1295 1300 1305

Gln Lys Tyr Leu Asn Thr Val Ala Arg Leu Leu Tyr Tyr Pro His

1310 1315 1320

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

1325 1330 1335

Lys Lys Thr Asn Arg Gly Asp Gly Thr Phe Asn Lys Gln Lys Gly

1340 1345 1350

Thr Ala Arg Lys Glu Lys Asp Asn Gly Pro Arg Glu Phe Asn Asp

1355 1360 1365

Thr Gly Phe Ser Asn Thr Pro Phe Ala Gly Phe Asp Pro Phe Arg

1370 1375 1380

Asn Ser

1385

<210> 37

<211> 1175

<212> PRT

<213> Carnobacterium gallinarum

<400> 37

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

1 5 10 15

Val Thr Met Gln Lys Glu Glu Lys Tyr Gly Thr Leu Lys Leu Asn Glu

20 25 30

Glu Ser Arg Lys Ser Thr Ala Glu Ile Leu Arg Leu Lys Lys Ala Ser

35 40 45

Phe Asn Lys Ser Phe His Ser Lys Thr Ile Asn Ser Gln Lys Glu Asn

50 55 60

Lys Asn Ala Thr Ile Lys Lys Asn Gly Asp Tyr Ile Ser Gln Ile Phe

65 70 75 80

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

85 90 95

Met Ser Leu Thr Asp Leu Lys Asp Leu Pro Lys Lys Asp Leu Ala Leu

100 105 110

Phe Ile Lys Arg Lys Phe Lys Asn Asp Asp Ile Val Glu Ile Lys Asn

115 120 125

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

130 135 140

Glu His Phe Thr Asp Glu Ser Trp Ala Asp Phe Cys Gln Glu Met Met

145 150 155 160

Pro Tyr Arg Glu Tyr Lys Asn Lys Phe Ile Glu Arg Lys Ile Ile Leu

165 170 175

Leu Ala Asn Ser Ile Glu Gln Asn Lys Gly Phe Ser Ile Asn Pro Glu
180 185 190
Thr Phe Ser Lys Arg Lys Arg Val Leu His Gln Trp Ala Ile Glu Val
195 200 205

Gln Glu Arg Gly Asp Phe Ser Ile Leu Asp Glu Lys Leu Ser Lys Leu
210 215 220
Ala Glu Ile Tyr Asn Phe Lys Lys Met Cys Lys Arg Val Gln Asp Glu
225 230 235 240
Leu Asn Asp Leu Glu Lys Ser Met Lys Lys Gly Lys Asn Pro Glu Lys
245 250 255
Glu Lys Glu Ala Tyr Lys Lys Gln Lys Asn Phe Lys Ile Lys Thr Ile
260 265 270

Trp Lys Asp Tyr Pro Tyr Lys Thr His Ile Gly Leu Ile Glu Lys Ile
275 280 285
Lys Glu Asn Glu Glu Leu Asn Gln Phe Asn Ile Glu Ile Gly Lys Tyr
290 295 300
Phe Glu His Tyr Phe Pro Ile Lys Lys Glu Arg Cys Thr Glu Asp Glu
305 310 315 320
Pro Tyr Tyr Leu Asn Ser Glu Thr Ile Ala Thr Thr Val Asn Tyr Gln
325 330 335

Leu Lys Asn Ala Leu Ile Ser Tyr Leu Met Gln Ile Gly Lys Tyr Lys
340 345 350
Gln Phe Gly Leu Glu Asn Gln Val Leu Asp Ser Lys Lys Leu Gln Glu
355 360 365
Ile Gly Ile Tyr Glu Gly Phe Gln Thr Lys Phe Met Asp Ala Cys Val
370 375 380
Phe Ala Thr Ser Ser Leu Lys Asn Ile Ile Glu Pro Met Arg Ser Gly
385 390 395 400

Asp Ile Leu Gly Lys Arg Glu Phe Lys Glu Ala Ile Ala Thr Ser Ser
405 410 415
Phe Val Asn Tyr His His Phe Phe Pro Tyr Phe Pro Phe Glu Leu Lys

420 425 430

Gly Met Lys Asp Arg Glu Ser Glu Leu Ile Pro Phe Gly Glu Gln Thr

435 440 445

Glu Ala Lys Gln Met Gln Asn Ile Trp Ala Leu Arg Gly Ser Val Gln

450 455 460

Gln Ile Arg Asn Glu Ile Phe His Ser Phe Asp Lys Asn Gln Lys Phe

465 470 475 480

Asn Leu Pro Gln Leu Asp Lys Ser Asn Phe Glu Phe Asp Ala Ser Glu

485 490 495

Asn Ser Thr Gly Lys Ser Gln Ser Tyr Ile Glu Thr Asp Tyr Lys Phe

500 505 510

Leu Phe Glu Ala Glu Lys Asn Gln Leu Glu Gln Phe Phe Ile Glu Arg

515 520 525

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

530 535 540

Lys Leu Phe Ala Lys Lys Glu Met Lys Phe Ser Leu Gly Ser Gln Val

545 550 555 560

Val Ala Phe Ala Pro Ser Tyr Lys Lys Leu Val Lys Lys Gly His Ser

565 570 575

Tyr Gln Thr Ala Thr Glu Gly Thr Ala Asn Tyr Leu Gly Leu Ser Tyr

580 585 590

Tyr Asn Arg Tyr Glu Leu Lys Glu Glu Ser Phe Gln Ala Gln Tyr Tyr

595 600 605

Leu Leu Lys Leu Ile Tyr Gln Tyr Val Phe Leu Pro Asn Phe Ser Gln

610 615 620

Gly Asn Ser Pro Ala Phe Arg Glu Thr Val Lys Ala Ile Leu Arg Ile

625 630 635 640

Asn Lys Asp Glu Ala Arg Lys Lys Met Lys Lys Asn Lys Lys Phe Leu

645 650 655

Arg Lys Tyr Ala Phe Glu Gln Val Arg Glu Met Glu Phe Lys Glu Thr

660 665 670

Pro Asp Gln Tyr Met Ser Tyr Leu Gln Ser Glu Met Arg Glu Glu Lys
675 680 685

Val Arg Lys Ala Glu Lys Asn Asp Lys Gly Phe Glu Lys Asn Ile Thr
690 695 700

Met Asn Phe Glu Lys Leu Leu Met Gln Ile Phe Val Lys Gly Phe Asp
705 710 715 720

Val Phe Leu Thr Thr Phe Ala Gly Lys Glu Leu Leu Leu Ser Ser Glu
725 730 735

Glu Lys Val Ile Lys Glu Thr Glu Ile Ser Leu Ser Lys Lys Ile Asn
740 745 750

Glu Arg Glu Lys Thr Leu Lys Ala Ser Ile Gln Val Glu His Gln Leu
755 760 765

Val Ala Thr Asn Ser Ala Ile Ser Tyr Trp Leu Phe Cys Lys Leu Leu
770 775 780

Asp Ser Arg His Leu Asn Glu Leu Arg Asn Glu Met Ile Lys Phe Lys
785 790 795 800

Gln Ser Arg Ile Lys Phe Asn His Thr Gln His Ala Glu Leu Ile Gln
805 810 815

Asn Leu Leu Pro Ile Val Glu Leu Thr Ile Leu Ser Asn Asp Tyr Asp
820 825 830

Glu Lys Asn Asp Ser Gln Asn Val Asp Val Ser Ala Tyr Phe Glu Asp
835 840 845

Lys Ser Leu Tyr Glu Thr Ala Pro Tyr Val Gln Thr Asp Asp Arg Thr
850 855 860

Arg Val Ser Phe Arg Pro Ile Leu Lys Leu Glu Lys Tyr His Thr Lys
865 870 875 880

Ser Leu Ile Glu Ala Leu Leu Lys Asp Asn Pro Gln Phe Arg Val Ala
885 890 895

Ala Thr Asp Ile Gln Glu Trp Met His Lys Arg Glu Glu Ile Gly Glu
900 905 910

Leu Val Glu Lys Arg Lys Asn Leu His Thr Glu Trp Ala Glu Gly Gln

915 920 925

Gln Thr Leu Gly Ala Glu Lys Arg Glu Glu Tyr Arg Asp Tyr Cys Lys

930 935 940

Lys Ile Asp Arg Phe Asn Trp Lys Ala Asn Lys Val Thr Leu Thr Tyr

945 950 955 960

Leu Ser Gln Leu His Tyr Leu Ile Thr Asp Leu Leu Gly Arg Met Val

965 970 975

Gly Phe Ser Ala Leu Phe Glu Arg Asp Leu Val Tyr Phe Ser Arg Ser

980 985 990

Phe Ser Glu Leu Gly Gly Glu Thr Tyr His Ile Ser Asp Tyr Lys Asn

995 1000 1005

Leu Ser Gly Val Leu Arg Leu Asn Ala Glu Val Lys Pro Ile Lys

1010 1015 1020

Ile Lys Asn Ile Lys Val Ile Asp Asn Glu Glu Asn Pro Tyr Lys

1025 1030 1035

Gly Asn Glu Pro Glu Val Lys Pro Phe Leu Asp Arg Leu His Ala

1040 1045 1050

Tyr Leu Glu Asn Val Ile Gly Ile Lys Ala Val His Gly Lys Ile

1055 1060 1065

Arg Asn Gln Thr Ala His Leu Ser Val Leu Gln Leu Glu Leu Ser

1070 1075 1080

Met Ile Glu Ser Met Asn Asn Leu Arg Asp Leu Met Ala Tyr Asp

1085 1090 1095

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

1100 1105 1110

Asp Lys His Gly Met Ile Leu Lys Leu Lys Ile Asp Glu Asn His

1115 1120 1125

Lys Asn Phe Glu Ile Glu Ser Leu Ile Pro Lys Glu Ile Ile His

1130 1135 1140

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

1145 1150 1155

Cys Gln Leu Val Leu Ala Leu Leu Thr Thr Asn Pro Gly Asn Gln

1160 1165 1170

Leu Asn

1175

<210> 38

<211> 1164

<212> PRT

<213> Carnobacterium gallinarum

<400> 38

Met Arg Met Thr Lys Val Lys Ile Asn Gly Ser Pro Val Ser Met Asn

1 5 10 15

Arg Ser Lys Leu Asn Gly His Leu Val Trp Asn Gly Thr Thr Asn Thr

20 25 30

Val Asn Ile Leu Thr Lys Lys Glu Gln Ser Phe Ala Ala Ser Phe Leu

35 40 45

Asn Lys Thr Leu Val Lys Ala Asp Gln Val Lys Gly Tyr Lys Val Leu

50 55 60

Ala Glu Asn Ile Phe Ile Ile Phe Glu Gln Leu Glu Lys Ser Asn Ser

65 70 75 80

Glu Lys Pro Ser Val Tyr Leu Asn Asn Ile Arg Arg Leu Lys Glu Ala

85 90 95

Gly Leu Lys Arg Phe Phe Lys Ser Lys Tyr His Glu Glu Ile Lys Tyr

100 105 110

Thr Ser Glu Lys Asn Gln Ser Val Pro Thr Lys Leu Asn Leu Ile Pro

115 120 125

Leu Phe Phe Asn Ala Val Asp Arg Ile Gln Glu Asp Lys Phe Asp Glu

130 135 140

Lys Asn Trp Ser Tyr Phe Cys Lys Glu Met Ser Pro Tyr Leu Asp Tyr

145 150 155 160

Lys Lys Ser Tyr Leu Asn Arg Lys Lys Glu Ile Leu Ala Asn Ser Ile

165 170 175

Gln Gln Asn Arg Gly Phe Ser Met Pro Thr Ala Glu Glu Pro Asn Leu

180 185 190

Leu Ser Lys Arg Lys Gln Leu Phe Gln Gln Trp Ala Met Lys Phe Gln

195 200 205

Glu Ser Pro Leu Ile Gln Gln Asn Asn Phe Ala Val Glu Gln Phe Asn

210 215 220

Lys Glu Phe Ala Asn Lys Ile Asn Glu Leu Ala Ala Val Tyr Asn Val

225 230 235 240

Asp Glu Leu Cys Thr Ala Ile Thr Glu Lys Leu Met Asn Phe Asp Lys

245 250 255

Asp Lys Ser Asn Lys Thr Arg Asn Phe Glu Ile Lys Lys Leu Trp Lys

260 265 270

Gln His Pro His Asn Lys Asp Lys Ala Leu Ile Lys Leu Phe Asn Gln

275 280 285

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

290 295 300

Phe Glu His Tyr Phe Pro Lys Thr Gly Lys Lys Glu Ser Ala Glu Ser

305 310 315 320

Tyr Tyr Leu Asn Pro Gln Thr Ile Ile Lys Thr Val Gly Tyr Gln Leu

325 330 335

Arg Asn Ala Phe Val Gln Tyr Leu Leu Gln Val Gly Lys Leu His Gln

340 345 350

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

355 360 365

Tyr Glu Gly Phe Gln Thr Lys Phe Met Asp Ala Cys Val Phe Ala Ser

370 375 380

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

385 390 395 400

Thr Arg Glu Lys Phe Lys Lys Glu Leu Glu Lys Asn Val Glu Leu Lys

405 410 415

His Asp Leu Phe Phe Lys Thr Glu Ile Val Glu Glu Arg Asp Glu Asn

420 425 430

Pro Ala Lys Lys Ile Ala Met Thr Pro Asn Glu Leu Asp Leu Trp Ala
435 440 445
Ile Arg Gly Ala Val Gln Arg Val Arg Asn Gln Ile Phe His Gln Gln
450 455 460
Ile Asn Lys Arg His Glu Pro Asn Gln Leu Lys Val Gly Ser Phe Glu
465 470 475 480
Asn Gly Asp Leu Gly Asn Val Ser Tyr Gln Lys Thr Ile Tyr Gln Lys

485 490 495
Leu Phe Asp Ala Glu Ile Lys Asp Ile Glu Ile Tyr Phe Ala Glu Lys
500 505 510
Ile Lys Ser Ser Gly Ala Leu Glu Gln Tyr Ser Met Lys Asp Leu Glu
515 520 525
Lys Leu Phe Ser Asn Lys Glu Leu Thr Leu Ser Leu Gly Gly Gln Val
530 535 540
Val Ala Phe Ala Pro Ser Tyr Lys Lys Leu Tyr Lys Gln Gly Tyr Phe

545 550 555 560
Tyr Gln Asn Glu Lys Thr Ile Glu Leu Glu Gln Phe Thr Asp Tyr Asp
565 570 575
Phe Ser Asn Asp Val Phe Lys Ala Asn Tyr Tyr Leu Ile Lys Leu Ile
580 585 590
Tyr His Tyr Val Phe Leu Pro Gln Phe Ser Gln Ala Asn Asn Lys Leu
595 600 605
Phe Lys Asp Thr Val His Tyr Val Ile Gln Gln Asn Lys Glu Leu Asn

610 615 620
Thr Thr Glu Lys Asp Lys Lys Asn Asn Lys Lys Ile Arg Lys Tyr Ala
625 630 635 640
Phe Glu Gln Val Lys Leu Met Lys Asn Glu Ser Pro Glu Lys Tyr Met
645 650 655
Gln Tyr Leu Gln Arg Glu Met Gln Glu Glu Arg Thr Ile Lys Glu Ala
660 665 670
Lys Lys Thr Asn Glu Glu Lys Pro Asn Tyr Asn Phe Glu Lys Leu Leu

675 680 685

Ile Gln Ile Phe Ile Lys Gly Phe Asp Thr Phe Leu Arg Asn Phe Asp

690 695 700

Leu Asn Leu Asn Pro Ala Glu Glu Leu Val Gly Thr Val Lys Glu Lys

705 710 715 720

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

725 730 735

Val Asp Glu Gln Ile Lys Thr Gly Asp Glu Glu Ile Ala Phe Trp Ile

740 745 750

Phe Ala Lys Leu Leu Asp Ala Arg His Leu Ser Glu Leu Arg Asn Glu

755 760 765

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

770 775 780

Gly Asp Leu Ile Glu Gln Met Gln Pro Ile Leu Glu Leu Cys Ile Leu

785 790 795 800

Ser Asn Asp Ser Glu Ser Met Glu Lys Glu Ser Phe Asp Lys Ile Glu

805 810 815

Val Phe Leu Glu Lys Val Glu Leu Ala Lys Asn Glu Pro Tyr Met Gln

820 825 830

Glu Asp Lys Leu Thr Pro Val Lys Phe Arg Phe Met Lys Gln Leu Glu

835 840 845

Lys Tyr Gln Thr Arg Asn Phe Ile Glu Asn Leu Val Ile Glu Asn Pro

850 855 860

Glu Phe Lys Val Ser Glu Lys Ile Val Leu Asn Trp His Glu Glu Lys

865 870 875 880

Glu Lys Ile Ala Asp Leu Val Asp Lys Arg Thr Lys Leu His Glu Glu

885 890 895

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

900 905 910

Lys Asn Lys Ser Lys Lys Leu Asp Lys Pro Ala Glu Phe Ala Lys Phe

915 920 925

Ala Glu Tyr Lys Ile Ile Cys Glu Ala Ile Glu Asn Phe Asn Arg Leu

930 935 940

Asp His Lys Val Arg Leu Thr Tyr Leu Lys Asn Leu His Tyr Leu Met

945 950 955 960

Ile Asp Leu Met Gly Arg Met Val Gly Phe Ser Val Leu Phe Glu Arg

965 970 975

Asp Phe Val Tyr Met Gly Arg Ser Tyr Ser Ala Leu Lys Lys Gln Ser

980 985 990

Ile Tyr Leu Asn Asp Tyr Asp Thr Phe Ala Asn Ile Arg Asp Trp Glu

995 1000 1005

Val Asn Glu Asn Lys His Leu Phe Gly Thr Ser Ser Ser Asp Leu

1010 1015 1020

Thr Phe Gln Glu Thr Ala Glu Phe Lys Asn Leu Lys Lys Pro Met

1025 1030 1035

Glu Asn Gln Leu Lys Ala Leu Leu Gly Val Thr Asn His Ser Phe

1040 1045 1050

Glu Ile Arg Asn Asn Ile Ala His Leu His Val Leu Arg Asn Asp

1055 1060 1065

Gly Lys Gly Glu Gly Val Ser Leu Leu Ser Cys Met Asn Asp Leu

1070 1075 1080

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

1085 1090 1095

Lys Ala Ile Ile Lys Ile Leu Asp Lys His Gly Met Ile Leu Lys

1100 1105 1110

Leu Thr Asn Asn Asp His Thr Lys Pro Phe Glu Ile Glu Ser Leu

1115 1120 1125

Lys Pro Lys Lys Ile Ile His Leu Glu Lys Ser Asn His Ser Phe

1130 1135 1140

Pro Met Asp Gln Val Ser Gln Glu Tyr Cys Asp Leu Val Lys Lys

1145 1150 1155

Met Leu Val Phe Thr Asn

1160

<210> 39

<211> 1154

<212> PRT

<213> Paludibacter propionicigenes

<400> 39

Met Arg Val Ser Lys Val Lys Val Lys Asp Gly Gly Lys Asp Lys Met

1 5 10 15

Val Leu Val His Arg Lys Thr Thr Gly Ala Gln Leu Val Tyr Ser Gly

20 25 30

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

35 40 45

Gln Ser Phe Asp Leu Ser Thr Leu Asn Lys Thr Ile Ile Lys Phe Asp

50 55 60

Thr Ala Lys Lys Gln Lys Leu Asn Val Asp Gln Tyr Lys Ile Val Glu

65 70 75 80

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

85 90 95

Glu Glu Ile Leu Pro Phe Leu Asn His Lys Phe Gln Glu Pro Val Lys

100 105 110

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

115 120 125

Val Glu Ala Val Gln Ala Gln Asp Lys Arg Lys Leu Gln Pro Tyr Tyr

130 135 140

Asp Trp Lys Thr Trp Tyr Ile Gln Thr Lys Ser Asp Leu Leu Lys Lys

145 150 155 160

Ser Ile Glu Asn Asn Arg Ile Asp Leu Thr Glu Asn Leu Ser Lys Arg

165 170 175

Lys Lys Ala Leu Leu Ala Trp Glu Thr Glu Phe Thr Ala Ser Gly Ser

180 185 190

Ile Asp Leu Thr His Tyr His Lys Val Tyr Met Thr Asp Val Leu Cys

195 200 205

Lys Met Leu Gln Asp Val Lys Pro Leu Thr Asp Asp Lys Gly Lys Ile

210 215 220

Asn Thr Asn Ala Tyr His Arg Gly Leu Lys Lys Ala Leu Gln Asn His

225 230 235 240

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

245 250 255

Ala Asp Asn Gln Leu Ser Ile Tyr His Leu Glu Val Val Lys Tyr Leu

260 265 270

Glu His Tyr Phe Pro Ile Lys Thr Ser Lys Arg Arg Asn Thr Ala Asp

275 280 285

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

290 295 300

Lys Gln Leu Val Asn Ala Ile Arg Ala Asn Ile Ile Gln Gln Gly Lys

305 310 315 320

Thr Asn His His Glu Leu Lys Ala Asp Thr Thr Ser Asn Asp Leu Ile

325 330 335

Arg Ile Lys Thr Asn Glu Ala Phe Val Leu Asn Leu Thr Gly Thr Cys

340 345 350

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

355 360 365

Asn Asp Ile Leu Gly Lys Gly Asp Phe Ile Lys Ser Leu Leu Lys Asp

370 375 380

Asn Thr Asn Ser Gln Leu Tyr Ser Phe Phe Phe Gly Glu Gly Leu Ser

385 390 395 400

Thr Asn Lys Ala Glu Lys Glu Thr Gln Leu Trp Gly Ile Arg Gly Ala

405 410 415

Val Gln Gln Ile Arg Asn Asn Val Asn His Tyr Lys Lys Asp Ala Leu

420 425 430

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

435 440 445

Pro Lys Gln Gln Thr Asn Tyr Ala Asp Thr Ile Tyr Lys Ala Arg Phe

450 455 460

Ile Asn Glu Leu Glu Lys Ile Pro Glu Ala Phe Ala Gln Gln Leu Lys
465 470 475 480
Thr Gly Gly Ala Val Ser Tyr Tyr Thr Ile Glu Asn Leu Lys Ser Leu
485 490 495
Leu Thr Thr Phe Gln Phe Ser Leu Cys Arg Ser Thr Ile Pro Phe Ala
500 505 510
Pro Gly Phe Lys Lys Val Phe Asn Gly Gly Ile Asn Tyr Gln Asn Ala

515 520 525

Lys Gln Asp Glu Ser Phe Tyr Glu Leu Met Leu Glu Gln Tyr Leu Arg
530 535 540
Lys Glu Asn Phe Ala Glu Glu Ser Tyr Asn Ala Arg Tyr Phe Met Leu
545 550 555 560
Lys Leu Ile Tyr Asn Asn Leu Phe Leu Pro Gly Phe Thr Thr Asp Arg
565 570 575
Lys Ala Phe Ala Asp Ser Val Gly Phe Val Gln Met Gln Asn Lys Lys

580 585 590

Gln Ala Glu Lys Val Asn Pro Arg Lys Lys Glu Ala Tyr Ala Phe Glu
595 600 605
Ala Val Arg Pro Met Thr Ala Ala Asp Ser Ile Ala Asp Tyr Met Ala
610 615 620
Tyr Val Gln Ser Glu Leu Met Gln Glu Gln Asn Lys Lys Glu Glu Lys
625 630 635 640
Val Ala Glu Glu Thr Arg Ile Asn Phe Glu Lys Phe Val Leu Gln Val

645 650 655

Phe Ile Lys Gly Phe Asp Ser Phe Leu Arg Ala Lys Glu Phe Asp Phe
660 665 670
Val Gln Met Pro Gln Pro Gln Leu Thr Ala Thr Ala Ser Asn Gln Gln
675 680 685
Lys Ala Asp Lys Leu Asn Gln Leu Glu Ala Ser Ile Thr Ala Asp Cys

690 695 700

Lys Leu Thr Pro Gln Tyr Ala Lys Ala Asp Asp Ala Thr His Ile Ala

705 710 715 720

Phe Tyr Val Phe Cys Lys Leu Leu Asp Ala Ala His Leu Ser Asn Leu

725 730 735

Arg Asn Glu Leu Ile Lys Phe Arg Glu Ser Val Asn Glu Phe Lys Phe

740 745 750

His His Leu Leu Glu Ile Ile Glu Ile Cys Leu Leu Ser Ala Asp Val

755 760 765

Val Pro Thr Asp Tyr Arg Asp Leu Tyr Ser Ser Glu Ala Asp Cys Leu

770 775 780

Ala Arg Leu Arg Pro Phe Ile Glu Gln Gly Ala Asp Ile Thr Asn Trp

785 790 795 800

Ser Asp Leu Phe Val Gln Ser Asp Lys His Ser Pro Val Ile His Ala

805 810 815

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

820 825 830

Ile Ile Asn Lys Asp Thr Gln Phe Lys Thr Thr Glu Ala Asn Phe Thr

835 840 845

Ala Trp Asn Thr Ala Gln Lys Ser Ile Glu Gln Leu Ile Lys Gln Arg

850 855 860

Glu Asp His His Glu Gln Trp Val Lys Ala Lys Asn Ala Asp Asp Lys

865 870 875 880

Glu Lys Gln Glu Arg Lys Arg Glu Lys Ser Asn Phe Ala Gln Lys Phe

885 890 895

Ile Glu Lys His Gly Asp Asp Tyr Leu Asp Ile Cys Asp Tyr Ile Asn

900 905 910

Thr Tyr Asn Trp Leu Asp Asn Lys Met His Phe Val His Leu Asn Arg

915 920 925

Leu His Gly Leu Thr Ile Glu Leu Leu Gly Arg Met Ala Gly Phe Val

930 935 940

Ala Leu Phe Asp Arg Asp Phe Gln Phe Phe Asp Glu Gln Gln Ile Ala
945 950 955 960
Asp Glu Phe Lys Leu His Gly Phe Val Asn Leu His Ser Ile Asp Lys

965 970 975
Lys Leu Asn Glu Val Pro Thr Lys Lys Ile Lys Glu Ile Tyr Asp Ile
980 985 990
Arg Asn Lys Ile Ile Gln Ile Asn Gly Asn Lys Ile Asn Glu Ser Val
995 1000 1005
Arg Ala Asn Leu Ile Gln Phe Ile Ser Ser Lys Arg Asn Tyr Tyr
1010 1015 1020
Asn Asn Ala Phe Leu His Val Ser Asn Asp Glu Ile Lys Glu Lys

1025 1030 1035
Gln Met Tyr Asp Ile Arg Asn His Ile Ala His Phe Asn Tyr Leu
1040 1045 1050
Thr Lys Asp Ala Ala Asp Phe Ser Leu Ile Asp Leu Ile Asn Glu
1055 1060 1065
Leu Arg Glu Leu Leu His Tyr Asp Arg Lys Leu Lys Asn Ala Val
1070 1075 1080
Ser Lys Ala Phe Ile Asp Leu Phe Asp Lys His Gly Met Ile Leu

1085 1090 1095
Lys Leu Lys Leu Asn Ala Asp His Lys Leu Lys Val Glu Ser Leu
1100 1105 1110
Glu Pro Lys Lys Ile Tyr His Leu Gly Ser Ser Ala Lys Asp Lys
1115 1120 1125
Pro Glu Tyr Gln Tyr Cys Thr Asn Gln Val Met Met Ala Tyr Cys
1130 1135 1140
Asn Met Cys Arg Ser Leu Leu Glu Met Lys Lys
1145 1150

<210> 40
<211> 970
<212> PRT

<213> Listeria weihenstephanensis

<400> 40

Met Leu Ala Leu Leu His Gln Glu Val Pro Ser Gln Lys Leu His Asn
1 5 10 15

Leu Lys Ser Leu Asn Thr Glu Ser Leu Thr Lys Leu Phe Lys Pro Lys
20 25 30

Phe Gln Asn Met Ile Ser Tyr Pro Pro Ser Lys Gly Ala Glu His Val
35 40 45

Gln Phe Cys Leu Thr Asp Ile Ala Val Pro Ala Ile Arg Asp Leu Asp
50 55 60

Glu Ile Lys Pro Asp Trp Gly Ile Phe Phe Glu Lys Leu Lys Pro Tyr
65 70 75 80

Thr Asp Trp Ala Glu Ser Tyr Ile His Tyr Lys Gln Thr Thr Ile Gln
85 90 95

Lys Ser Ile Glu Gln Asn Lys Ile Gln Ser Pro Asp Ser Pro Arg Lys
100 105 110

Leu Val Leu Gln Lys Tyr Val Thr Ala Phe Leu Asn Gly Glu Pro Leu
115 120 125

Gly Leu Asp Leu Val Ala Lys Lys Tyr Lys Leu Ala Asp Leu Ala Glu
130 135 140

Ser Phe Lys Val Val Asp Leu Asn Glu Asp Lys Ser Ala Asn Tyr Lys
145 150 155 160

Ile Lys Ala Cys Leu Gln Gln His Gln Arg Asn Ile Leu Asp Glu Leu
165 170 175

Lys Glu Asp Pro Glu Leu Asn Gln Tyr Gly Ile Glu Val Lys Lys Tyr
180 185 190

Ile Gln Arg Tyr Phe Pro Ile Lys Arg Ala Pro Asn Arg Ser Lys His
195 200 205

Ala Arg Ala Asp Phe Leu Lys Lys Glu Leu Ile Glu Ser Thr Val Glu
210 215 220

Gln Gln Phe Lys Asn Ala Val Tyr His Tyr Val Leu Glu Gln Gly Lys
225 230 235 240

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

245 250 255

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

260 265 270

Phe Ala Ser Asn Asn Leu Lys Met Ile Leu Asn Pro Glu Cys Glu Lys

275 280 285

Asp Ile Leu Gly Lys Gly Asp Phe Lys Lys Asn Leu Pro Asn Ser Thr

290 295 300

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

305 310 315 320

Ile Gln Asn Val Asn Phe Asp Glu Ala Ile Trp Ala Ile Arg Gly Ser

325 330 335

Ile Gln Gln Ile Arg Asn Glu Val Tyr His Cys Lys Lys His Ser Trp

340 345 350

Lys Ser Ile Leu Lys Ile Lys Gly Phe Glu Phe Glu Pro Asn Asn Met

355 360 365

Lys Tyr Thr Asp Ser Asp Met Gln Lys Leu Met Asp Lys Asp Ile Ala

370 375 380

Lys Ile Pro Asp Phe Ile Glu Glu Lys Leu Lys Ser Ser Gly Ile Ile

385 390 395 400

Arg Phe Tyr Ser His Asp Lys Leu Gln Ser Ile Trp Glu Met Lys Gln

405 410 415

Gly Phe Ser Leu Leu Thr Thr Asn Ala Pro Phe Val Pro Ser Phe Lys

420 425 430

Arg Val Tyr Ala Lys Gly His Asp Tyr Gln Thr Ser Lys Asn Arg Tyr

435 440 445

Tyr Asp Leu Gly Leu Thr Thr Phe Asp Ile Leu Glu Tyr Gly Glu Glu

450 455 460

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

465 470 475 480

Phe Met Pro Trp Phe Thr Ala Asp Asn Asn Ala Phe Arg Asp Ala Ala

485 490 495

Asn Phe Val Leu Arg Leu Asn Lys Asn Arg Gln Gln Asp Ala Lys Ala

500 505 510

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

515 520 525

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

530 535 540

Asp Thr Pro Asn His Phe Glu Lys Phe Ile Ser Gln Val Phe Ile Lys

545 550 555 560

Gly Phe Asp Ser His Met Arg Ser Ala Asp Leu Lys Phe Ile Lys Asn

565 570 575

Pro Arg Asn Gln Gly Leu Glu Gln Ser Glu Ile Glu Glu Met Ser Phe

580 585 590

Asp Ile Lys Val Glu Pro Ser Phe Leu Lys Asn Lys Asp Asp Tyr Ile

595 600 605

Ala Phe Trp Thr Phe Cys Lys Met Leu Asp Ala Arg His Leu Ser Glu

610 615 620

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

625 630 635 640

Glu Ile Ile Gly Leu Ala Leu Leu Gly Val Asp Ser Arg Glu Asn Asp

645 650 655

Trp Lys Gln Phe Phe Ser Ser Glu Arg Glu Tyr Glu Lys Ile Met Lys

660 665 670

Gly Tyr Val Gly Glu Glu Leu Tyr Gln Arg Glu Pro Tyr Arg Gln Ser

675 680 685

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

690 695 700

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

705 710 715 720

Phe Lys Val Ser Lys Cys Asn Ile Thr Glu Trp Glu Arg Gln Lys Glu

725 730 735

Thr Ile Glu Glu Thr Ile Glu Arg Arg Lys Glu Leu His Asn Glu Trp
740 745 750
Glu Lys Asn Pro Lys Lys Pro Gln Asn Asn Ala Phe Phe Lys Glu Tyr

755 760 765
Lys Glu Cys Cys Asp Ala Ile Asp Ala Tyr Asn Trp His Lys Asn Lys
770 775 780
Thr Thr Leu Val Tyr Val Asn Glu Leu His His Leu Leu Ile Glu Ile
785 790 795 800
Leu Gly Arg Tyr Val Gly Tyr Val Ala Ile Ala Asp Arg Asp Phe Gln
805 810 815
Cys Met Ala Asn Gln Tyr Phe Lys His Ser Gly Ile Thr Glu Arg Val

820 825 830
Glu Tyr Trp Gly Asp Asn Arg Leu Lys Ser Ile Lys Lys Leu Asp Thr
835 840 845
Phe Leu Lys Lys Glu Gly Leu Phe Val Ser Glu Lys Asn Ala Arg Asn
850 855 860
His Ile Ala His Leu Asn Tyr Leu Ser Leu Lys Ser Glu Cys Thr Leu
865 870 875 880
Leu Tyr Leu Ser Glu Arg Leu Arg Glu Ile Phe Lys Tyr Asp Arg Lys

885 890 895
Leu Lys Asn Ala Val Ser Lys Ser Leu Ile Asp Ile Leu Asp Arg His
900 905 910
Gly Met Ser Val Val Phe Ala Asn Leu Lys Glu Asn Lys His Arg Leu
915 920 925
Val Ile Lys Ser Leu Glu Pro Lys Lys Leu Arg His Leu Gly Glu Lys
930 935 940
Lys Ile Asp Asn Gly Tyr Ile Glu Thr Asn Gln Val Ser Glu Glu Tyr

945 950 955 960
Cys Gly Ile Val Lys Arg Leu Leu Glu Ile
965 970

<210> 41

<211> 1051

<212> PRT

<213> Listeria newyorkensis

<400> 41

Met Lys Ile Thr Lys Met Arg Val Asp Gly Arg Thr Ile Val Met Glu
1 5 10 15

Arg Thr Ser Lys Glu Gly Gln Leu Gly Tyr Glu Gly Ile Asp Gly Asn
20 25 30

Lys Thr Thr Glu Ile Ile Phe Asp Lys Lys Lys Glu Ser Phe Tyr Lys
35 40 45

Ser Ile Leu Asn Lys Thr Val Arg Lys Pro Asp Glu Lys Glu Lys Asn
50 55 60

Arg Arg Lys Gln Ala Ile Asn Lys Ala Ile Asn Lys Glu Ile Thr Glu
65 70 75 80

Leu Met Leu Ala Val Leu His Gln Glu Val Pro Ser Gln Lys Leu His
85 90 95

Asn Leu Lys Ser Leu Asn Thr Glu Ser Leu Thr Lys Leu Phe Lys Pro
100 105 110

Lys Phe Gln Asn Met Ile Ser Tyr Pro Pro Ser Lys Gly Ala Glu His
115 120 125

Val Gln Phe Cys Leu Thr Asp Ile Ala Val Pro Ala Ile Arg Asp Leu
130 135 140

Asp Glu Ile Lys Pro Asp Trp Gly Ile Phe Phe Glu Lys Leu Lys Pro
145 150 155 160

Tyr Thr Asp Trp Ala Glu Ser Tyr Ile His Tyr Lys Gln Thr Thr Ile
165 170 175

Gln Lys Ser Ile Glu Gln Asn Lys Ile Gln Ser Pro Asp Ser Pro Arg
180 185 190

Lys Leu Val Leu Gln Lys Tyr Val Thr Ala Phe Leu Asn Gly Glu Pro
195 200 205

Leu Gly Leu Asp Leu Val Ala Lys Lys Tyr Lys Leu Ala Asp Leu Ala
210 215 220

Glu Ser Phe Lys Leu Val Asp Leu Asn Glu Asp Lys Ser Ala Asn Tyr

225 230 235 240

Lys Ile Lys Ala Cys Leu Gln Gln His Gln Arg Asn Ile Leu Asp Glu

245 250 255

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

260 265 270

Tyr Ile Gln Arg Tyr Phe Pro Ile Lys Arg Ala Pro Asn Arg Ser Lys

275 280 285

His Ala Arg Ala Asp Phe Leu Lys Lys Glu Leu Ile Glu Ser Thr Val

290 295 300

Glu Gln Gln Phe Lys Asn Ala Val Tyr His Tyr Val Leu Glu Gln Gly

305 310 315 320

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

325 330 335

Asp Ile Arg Ser Gly Glu Ala Phe Ser Phe Lys Phe Ile Asn Ala Cys

340 345 350

Ala Phe Ala Ser Asn Asn Leu Lys Met Ile Leu Asn Pro Glu Cys Glu

355 360 365

Lys Asp Ile Leu Gly Lys Gly Asn Phe Lys Lys Asn Leu Pro Asn Ser

370 375 380

Thr Thr Arg Ser Asp Val Val Lys Lys Met Ile Pro Phe Phe Ser Asp

385 390 395 400

Glu Leu Gln Asn Val Asn Phe Asp Glu Ala Ile Trp Ala Ile Arg Gly

405 410 415

Ser Ile Gln Gln Ile Arg Asn Glu Val Tyr His Cys Lys Lys His Ser

420 425 430

Trp Lys Ser Ile Leu Lys Ile Lys Gly Phe Glu Phe Glu Pro Asn Asn

435 440 445

Met Lys Tyr Ala Asp Ser Asp Met Gln Lys Leu Met Asp Lys Asp Ile

450 455 460

Ala Lys Ile Pro Glu Phe Ile Glu Glu Lys Leu Lys Ser Ser Gly Val

465 470 475 480

Val Arg Phe Tyr Arg His Asp Glu Leu Gln Ser Ile Trp Glu Met Lys

485 490 495

Gln Gly Phe Ser Leu Leu Thr Thr Asn Ala Pro Phe Val Pro Ser Phe

500 505 510

Lys Arg Val Tyr Ala Lys Gly His Asp Tyr Gln Thr Ser Lys Asn Arg

515 520 525

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

530 535 540

Glu Asp Phe Arg Ala Arg Tyr Phe Leu Thr Lys Leu Val Tyr Tyr Gln

545 550 555 560

Gln Phe Met Pro Trp Phe Thr Ala Asp Asn Asn Ala Phe Arg Asp Ala

565 570 575

Ala Asn Phe Val Leu Arg Leu Asn Lys Asn Arg Gln Gln Asp Ala Lys

580 585 590

Ala Phe Ile Asn Ile Arg Glu Val Glu Glu Gly Glu Met Pro Arg Asp

595 600 605

Tyr Met Gly Tyr Val Gln Gly Gln Ile Ala Ile His Glu Asp Ser Ile

610 615 620

Glu Asp Thr Pro Asn His Phe Glu Lys Phe Ile Ser Gln Val Phe Ile

625 630 635 640

Lys Gly Phe Asp Arg His Met Arg Ser Ala Asn Leu Lys Phe Ile Lys

645 650 655

Asn Pro Arg Asn Gln Gly Leu Glu Gln Ser Glu Ile Glu Glu Met Ser

660 665 670

Phe Asp Ile Lys Val Glu Pro Ser Phe Leu Lys Asn Lys Asp Asp Tyr

675 680 685

Ile Ala Phe Trp Ile Phe Cys Lys Met Leu Asp Ala Arg His Leu Ser

690 695 700

Glu Leu Arg Asn Glu Met Ile Lys Tyr Asp Gly His Leu Thr Gly Glu

705 710 715 720

Gln Glu Ile Ile Gly Leu Ala Leu Leu Gly Val Asp Ser Arg Glu Asn
725 730 735
Asp Trp Lys Gln Phe Phe Ser Ser Glu Arg Glu Tyr Glu Lys Ile Met

740 745 750
Lys Gly Tyr Val Val Glu Glu Leu Tyr Gln Arg Glu Pro Tyr Arg Gln
755 760 765
Ser Asp Gly Lys Thr Pro Ile Leu Phe Arg Gly Val Glu Gln Ala Arg
770 775 780
Lys Tyr Gly Thr Glu Thr Val Ile Gln Arg Leu Phe Asp Ala Asn Pro
785 790 795 800
Glu Phe Lys Val Ser Lys Cys Asn Leu Ala Glu Trp Glu Arg Gln Lys

805 810 815
Glu Thr Ile Glu Glu Thr Ile Lys Arg Arg Lys Glu Leu His Asn Glu
820 825 830
Trp Ala Lys Asn Pro Lys Lys Pro Gln Asn Asn Ala Phe Phe Lys Glu
835 840 845
Tyr Lys Glu Cys Cys Asp Ala Ile Asp Ala Tyr Asn Trp His Lys Asn
850 855 860
Lys Thr Thr Leu Ala Tyr Val Asn Glu Leu His His Leu Leu Ile Glu

865 870 875 880
Ile Leu Gly Arg Tyr Val Gly Tyr Val Ala Ile Ala Asp Arg Asp Phe
885 890 895
Gln Cys Met Ala Asn Gln Tyr Phe Lys His Ser Gly Ile Thr Glu Arg
900 905 910
Val Glu Tyr Trp Gly Asp Asn Arg Leu Lys Ser Ile Lys Lys Leu Asp
915 920 925
Thr Phe Leu Lys Lys Glu Gly Leu Phe Val Ser Glu Lys Asn Ala Arg

930 935 940
Asn His Ile Ala His Leu Asn Tyr Leu Ser Leu Lys Ser Glu Cys Thr
945 950 955 960
Leu Leu Tyr Leu Ser Glu Arg Leu Arg Glu Ile Phe Lys Tyr Asp Arg

965 970 975

Lys Leu Lys Asn Ala Val Ser Lys Ser Leu Ile Asp Ile Leu Asp Arg

980 985 990

His Gly Met Ser Val Val Phe Ala Asn Leu Lys Glu Asn Lys His Arg

995 1000 1005

Leu Val Ile Lys Ser Leu Glu Pro Lys Lys Leu Arg His Leu Gly

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

<210> 42

<211> 1152

<212> PRT

<213> Leptotrichia wadei

<400> 42

Met Lys Val Thr Lys Val Asp Gly Ile Ser His Lys Lys Tyr Ile Glu

1 5 10 15

Glu Gly Lys Leu Val Lys Ser Thr Ser Glu Glu Asn Arg Thr Ser Glu

20 25 30

Arg Leu Ser Glu Leu Leu Ser Ile Arg Leu Asp Ile Tyr Ile Lys Asn

35 40 45

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

50 55 60

Lys Lys Phe Phe Ser Asn Lys Val Leu His Leu Lys Asp Ser Val Leu

65 70 75 80

Tyr Leu Lys Asn Arg Lys Glu Lys Asn Ala Val Gln Asp Lys Asn Tyr

85 90 95

Ser Glu Glu Asp Ile Ser Glu Tyr Asp Leu Lys Asn Lys Asn Ser Phe

100 105 110

Ser Val Leu Lys Lys Ile Leu Leu Asn Glu Asp Val Asn Ser Glu Glu

115 120 125

Leu Glu Ile Phe Arg Lys Asp Val Glu Ala Lys Leu Asn Lys Ile Asn
130 135 140
Ser Leu Lys Tyr Ser Phe Glu Glu Asn Lys Ala Asn Tyr Gln Lys Ile
145 150 155 160
Asn Glu Asn Asn Val Glu Lys Val Gly Gly Lys Ser Lys Arg Asn Ile
165 170 175
Ile Tyr Asp Tyr Tyr Arg Glu Ser Ala Lys Arg Asn Asp Tyr Ile Asn
180 185 190

Asn Val Gln Glu Ala Phe Asp Lys Leu Tyr Lys Lys Glu Asp Ile Glu
195 200 205
Lys Leu Phe Phe Leu Ile Glu Asn Ser Lys Lys His Glu Lys Tyr Lys
210 215 220
Ile Arg Glu Tyr Tyr His Lys Ile Ile Gly Arg Lys Asn Asp Lys Glu
225 230 235 240
Asn Phe Ala Lys Ile Ile Tyr Glu Glu Ile Gln Asn Val Asn Asn Ile
245 250 255

Lys Glu Leu Ile Glu Lys Ile Pro Asp Met Ser Glu Leu Lys Lys Ser
260 265 270
Gln Val Phe Tyr Lys Tyr Tyr Leu Asp Lys Glu Glu Leu Asn Asp Lys
275 280 285
Asn Ile Lys Tyr Ala Phe Cys His Phe Val Glu Ile Glu Met Ser Gln
290 295 300
Leu Leu Lys Asn Tyr Val Tyr Lys Arg Leu Ser Asn Ile Ser Asn Asp
305 310 315 320

Lys Ile Lys Arg Ile Phe Glu Tyr Gln Asn Leu Lys Lys Leu Ile Glu
325 330 335
Asn Lys Leu Leu Asn Lys Leu Asp Thr Tyr Val Arg Asn Cys Gly Lys
340 345 350
Tyr Asn Tyr Tyr Leu Gln Val Gly Glu Ile Ala Thr Ser Asp Phe Ile
355 360 365

Ala Arg Asn Arg Gln Asn Glu Ala Phe Leu Arg Asn Ile Ile Gly Val
370 375 380

Ser Ser Val Ala Tyr Phe Ser Leu Arg Asn Ile Leu Glu Thr Glu Asn
385 390 395 400
Glu Asn Asp Ile Thr Gly Arg Met Arg Gly Lys Thr Val Lys Asn Asn
405 410 415
Lys Gly Glu Glu Lys Tyr Val Ser Gly Glu Val Asp Lys Ile Tyr Asn
420 425 430
Glu Asn Lys Gln Asn Glu Val Lys Glu Asn Leu Lys Met Phe Tyr Ser
435 440 445

Tyr Asp Phe Asn Met Asp Asn Lys Asn Glu Ile Glu Asp Phe Phe Ala
450 455 460
Asn Ile Asp Glu Ala Ile Ser Ser Ile Arg His Gly Ile Val His Phe
465 470 475 480
Asn Leu Glu Leu Glu Gly Lys Asp Ile Phe Ala Phe Lys Asn Ile Ala
485 490 495
Pro Ser Glu Ile Ser Lys Lys Met Phe Gln Asn Glu Ile Asn Glu Lys
500 505 510

Lys Leu Lys Leu Lys Ile Phe Lys Gln Leu Asn Ser Ala Asn Val Phe
515 520 525
Asn Tyr Tyr Glu Lys Asp Val Ile Ile Lys Tyr Leu Lys Asn Thr Lys
530 535 540
Phe Asn Phe Val Asn Lys Asn Ile Pro Phe Val Pro Ser Phe Thr Lys
545 550 555 560
Leu Tyr Asn Lys Ile Glu Asp Leu Arg Asn Thr Leu Lys Phe Phe Trp
565 570 575

Ser Val Pro Lys Asp Lys Glu Glu Lys Asp Ala Gln Ile Tyr Leu Leu
580 585 590
Lys Asn Ile Tyr Tyr Gly Glu Phe Leu Asn Lys Phe Val Lys Asn Ser
595 600 605
Lys Val Phe Phe Lys Ile Thr Asn Glu Val Ile Lys Ile Asn Lys Gln

610 615 620

Arg Asn Gln Lys Thr Gly His Tyr Lys Tyr Gln Lys Phe Glu Asn Ile
625 630 635 640

Glu Lys Thr Val Pro Val Glu Tyr Leu Ala Ile Ile Gln Ser Arg Glu
645 650 655

Met Ile Asn Asn Gln Asp Lys Glu Glu Lys Asn Thr Tyr Ile Asp Phe
660 665 670

Ile Gln Gln Ile Phe Leu Lys Gly Phe Ile Asp Tyr Leu Asn Lys Asn
675 680 685

Asn Leu Lys Tyr Ile Glu Ser Asn Asn Asn Asn Asp Asn Asn Asp Ile
690 695 700

Phe Ser Lys Ile Lys Ile Lys Lys Asp Asn Lys Glu Lys Tyr Asp Lys
705 710 715 720

Ile Leu Lys Asn Tyr Glu Lys His Asn Arg Asn Lys Glu Ile Pro His
725 730 735

Glu Ile Asn Glu Phe Val Arg Glu Ile Lys Leu Gly Lys Ile Leu Lys
740 745 750

Tyr Thr Glu Asn Leu Asn Met Phe Tyr Leu Ile Leu Lys Leu Leu Asn
755 760 765

His Lys Glu Leu Thr Asn Leu Lys Gly Ser Leu Glu Lys Tyr Gln Ser
770 775 780

Ala Asn Lys Glu Glu Thr Phe Ser Asp Glu Leu Glu Leu Ile Asn Leu
785 790 795 800

Leu Asn Leu Asp Asn Asn Arg Val Thr Glu Asp Phe Glu Leu Glu Ala
805 810 815

Asn Glu Ile Gly Lys Phe Leu Asp Phe Asn Glu Asn Lys Ile Lys Asp
820 825 830

Arg Lys Glu Leu Lys Lys Phe Asp Thr Asn Lys Ile Tyr Phe Asp Gly
835 840 845

Glu Asn Ile Ile Lys His Arg Ala Phe Tyr Asn Ile Lys Lys Tyr Gly
850 855 860

Met Leu Asn Leu Leu Glu Lys Ile Ala Asp Lys Ala Lys Tyr Lys Ile
865 870 875 880
Ser Leu Lys Glu Leu Lys Glu Tyr Ser Asn Lys Lys Asn Glu Ile Glu
885 890 895

Lys Asn Tyr Thr Met Gln Gln Asn Leu His Arg Lys Tyr Ala Arg Pro
900 905 910
Lys Lys Asp Glu Lys Phe Asn Asp Glu Asp Tyr Lys Glu Tyr Glu Lys
915 920 925
Ala Ile Gly Asn Ile Gln Lys Tyr Thr His Leu Lys Asn Lys Val Glu
930 935 940
Phe Asn Glu Leu Asn Leu Leu Gln Gly Leu Leu Leu Lys Ile Leu His
945 950 955 960

Arg Leu Val Gly Tyr Thr Ser Ile Trp Glu Arg Asp Leu Arg Phe Arg
965 970 975
Leu Lys Gly Glu Phe Pro Glu Asn His Tyr Ile Glu Glu Ile Phe Asn
980 985 990
Phe Asp Asn Ser Lys Asn Val Lys Tyr Lys Ser Gly Gln Ile Val Glu
995 1000 1005
Lys Tyr Ile Asn Phe Tyr Lys Glu Leu Tyr Lys Asp Asn Val Glu
1010 1015 1020

Lys Arg Ser Ile Tyr Ser Asp Lys Lys Val Lys Lys Leu Lys Gln
1025 1030 1035
Glu Lys Lys Asp Leu Tyr Ile Arg Asn Tyr Ile Ala His Phe Asn
1040 1045 1050
Tyr Ile Pro His Ala Glu Ile Ser Leu Leu Glu Val Leu Glu Asn
1055 1060 1065
Leu Arg Lys Leu Leu Ser Tyr Asp Arg Lys Leu Lys Asn Ala Ile
1070 1075 1080

Met Lys Ser Ile Val Asp Ile Leu Lys Glu Tyr Gly Phe Val Ala
1085 1090 1095
Thr Phe Lys Ile Gly Ala Asp Lys Lys Ile Glu Ile Gln Thr Leu

1100 1105 1110

Glu Ser Glu Lys Ile Val His Leu Lys Asn Leu Lys Lys Lys Lys

1115 1120 1125

Leu Met Thr Asp Arg Asn Ser Glu Glu Leu Cys Glu Leu Val Lys

1130 1135 1140

Val Met Phe Glu Tyr Lys Ala Leu Glu

1145 1150

<210> 43

<211> 1285

<212> PRT

<213> Rhodobacter capsulatus

<400> 43

Met Gln Ile Gly Lys Val Gln Gly Arg Thr Ile Ser Glu Phe Gly Asp

1 5 10 15

Pro Ala Gly Gly Leu Lys Arg Lys Ile Ser Thr Asp Gly Lys Asn Arg

20 25 30

Lys Glu Leu Pro Ala His Leu Ser Ser Asp Pro Lys Ala Leu Ile Gly

35 40 45

Gln Trp Ile Ser Gly Ile Asp Lys Ile Tyr Arg Lys Pro Asp Ser Arg

50 55 60

Lys Ser Asp Gly Lys Ala Ile His Ser Pro Thr Pro Ser Lys Met Gln

65 70 75 80

Phe Asp Ala Arg Asp Asp Leu Gly Glu Ala Phe Trp Lys Leu Val Ser

85 90 95

Glu Ala Gly Leu Ala Gln Asp Ser Asp Tyr Asp Gln Phe Lys Arg Arg

100 105 110

Leu His Pro Tyr Gly Asp Lys Phe Gln Pro Ala Asp Ser Gly Ala Lys

115 120 125

Leu Lys Phe Glu Ala Asp Pro Pro Glu Pro Gln Ala Phe His Gly Arg

130 135 140

Trp Tyr Gly Ala Met Ser Lys Arg Gly Asn Asp Ala Lys Glu Leu Ala

145 150 155 160

Ala Ala Leu Tyr Glu His Leu His Val Asp Glu Lys Arg Ile Asp Gly
165 170 175

Gln Pro Lys Arg Asn Pro Lys Thr Asp Lys Phe Ala Pro Gly Leu Val
180 185 190
Val Ala Arg Ala Leu Gly Ile Glu Ser Ser Val Leu Pro Arg Gly Met
195 200 205
Ala Arg Leu Ala Arg Asn Trp Gly Glu Glu Glu Ile Gln Thr Tyr Phe
210 215 220
Val Val Asp Val Ala Ala Ser Val Lys Glu Val Ala Lys Ala Ala Val
225 230 235 240

Ser Ala Ala Gln Ala Phe Asp Pro Pro Arg Gln Val Ser Gly Arg Ser
245 250 255
Leu Ser Pro Lys Val Gly Phe Ala Leu Ala Glu His Leu Glu Arg Val
260 265 270
Thr Gly Ser Lys Arg Cys Ser Phe Asp Pro Ala Ala Gly Pro Ser Val
275 280 285
Leu Ala Leu His Asp Glu Val Lys Lys Thr Tyr Lys Arg Leu Cys Ala
290 295 300

Arg Gly Lys Asn Ala Ala Arg Ala Phe Pro Ala Asp Lys Thr Glu Leu
305 310 315 320
Leu Ala Leu Met Arg His Thr His Glu Asn Arg Val Arg Asn Gln Met
325 330 335
Val Arg Met Gly Arg Val Ser Glu Tyr Arg Gly Gln Gln Ala Gly Asp
340 345 350
Leu Ala Gln Ser His Tyr Trp Thr Ser Ala Gly Gln Thr Glu Ile Lys
355 360 365

Glu Ser Glu Ile Phe Val Arg Leu Trp Val Gly Ala Phe Ala Leu Ala
370 375 380
Gly Arg Ser Met Lys Ala Trp Ile Asp Pro Met Gly Lys Ile Val Asn
385 390 395 400
Thr Glu Lys Asn Asp Arg Asp Leu Thr Ala Ala Val Asn Ile Arg Gln

405 410 415

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

420 425 430

Ile Tyr Phe Gly Glu Thr Pro Glu Leu Asp Arg Leu Gly Ala Glu Gly

435 440 445

Asn Glu Gly Phe Val Phe Ala Leu Leu Arg Tyr Leu Arg Gly Cys Arg

450 455 460

Asn Gln Thr Phe His Leu Gly Ala Arg Ala Gly Phe Leu Lys Glu Ile

465 470 475 480

Arg Lys Glu Leu Glu Lys Thr Arg Trp Gly Lys Ala Lys Glu Ala Glu

485 490 495

His Val Val Leu Thr Asp Lys Thr Val Ala Ala Ile Arg Ala Ile Ile

500 505 510

Asp Asn Asp Ala Lys Ala Leu Gly Ala Arg Leu Leu Ala Asp Leu Ser

515 520 525

Gly Ala Phe Val Ala His Tyr Ala Ser Lys Glu His Phe Ser Thr Leu

530 535 540

Tyr Ser Glu Ile Val Lys Ala Val Lys Asp Ala Pro Glu Val Ser Ser

545 550 555 560

Gly Leu Pro Arg Leu Lys Leu Leu Leu Lys Arg Ala Asp Gly Val Arg

565 570 575

Gly Tyr Val His Gly Leu Arg Asp Thr Arg Lys His Ala Phe Ala Thr

580 585 590

Lys Leu Pro Pro Pro Pro Ala Pro Arg Glu Leu Asp Asp Pro Ala Thr

595 600 605

Lys Ala Arg Tyr Ile Ala Leu Leu Arg Leu Tyr Asp Gly Pro Phe Arg

610 615 620

Ala Tyr Ala Ser Gly Ile Thr Gly Thr Ala Leu Ala Gly Pro Ala Ala

625 630 635 640

Arg Ala Lys Glu Ala Ala Thr Ala Leu Ala Gln Ser Val Asn Val Thr

645 650 655

Lys Ala Tyr Ser Asp Val Met Glu Gly Arg Thr Ser Arg Leu Arg Pro
660 665 670

Pro Asn Asp Gly Glu Thr Leu Arg Glu Tyr Leu Ser Ala Leu Thr Gly
675 680 685

Glu Thr Ala Thr Glu Phe Arg Val Gln Ile Gly Tyr Glu Ser Asp Ser
690 695 700

Glu Asn Ala Arg Lys Gln Ala Glu Phe Ile Glu Asn Tyr Arg Arg Asp
705 710 715 720

Met Leu Ala Phe Met Phe Glu Asp Tyr Ile Arg Ala Lys Gly Phe Asp
725 730 735

Trp Ile Leu Lys Ile Glu Pro Gly Ala Thr Ala Met Thr Arg Ala Pro
740 745 750

Val Leu Pro Glu Pro Ile Asp Thr Arg Gly Gln Tyr Glu His Trp Gln
755 760 765

Ala Ala Leu Tyr Leu Val Met His Phe Val Pro Ala Ser Asp Val Ser
770 775 780

Asn Leu Leu His Gln Leu Arg Lys Trp Glu Ala Leu Gln Gly Lys Tyr
785 790 795 800

Glu Leu Val Gln Asp Gly Asp Ala Thr Asp Gln Ala Asp Ala Arg Arg
805 810 815

Glu Ala Leu Asp Leu Val Lys Arg Phe Arg Asp Val Leu Val Leu Phe
820 825 830

Leu Lys Thr Gly Glu Ala Arg Phe Glu Gly Arg Ala Ala Pro Phe Asp
835 840 845

Leu Lys Pro Phe Arg Ala Leu Phe Ala Asn Pro Ala Thr Phe Asp Arg
850 855 860

Leu Phe Met Ala Thr Pro Thr Thr Ala Arg Pro Ala Glu Asp Asp Pro
865 870 875 880

Glu Gly Asp Gly Ala Ser Glu Pro Glu Leu Arg Val Ala Arg Thr Leu
885 890 895

Arg Gly Leu Arg Gln Ile Ala Arg Tyr Asn His Met Ala Val Leu Ser

900 905 910

Asp Leu Phe Ala Lys His Lys Val Arg Asp Glu Glu Val Ala Arg Leu

915 920 925

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

930 935 940

Gln Glu Leu Arg Thr Asp Leu His Asp Lys Val Met Lys Cys His Pro

945 950 955 960

Lys Thr Ile Ser Pro Glu Glu Arg Gln Ser Tyr Ala Ala Ala Ile Lys

965 970 975

Thr Ile Glu Glu His Arg Phe Leu Val Gly Arg Val Tyr Leu Gly Asp

980 985 990

His Leu Arg Leu His Arg Leu Met Met Asp Val Ile Gly Arg Leu Ile

995 1000 1005

Asp Tyr Ala Gly Ala Tyr Glu Arg Asp Thr Gly Thr Phe Leu Ile

1010 1015 1020

Asn Ala Ser Lys Gln Leu Gly Ala Gly Ala Asp Trp Ala Val Thr

1025 1030 1035

Ile Ala Gly Ala Ala Asn Thr Asp Ala Arg Thr Gln Thr Arg Lys

1040 1045 1050

Asp Leu Ala His Phe Asn Val Leu Asp Arg Ala Asp Gly Thr Pro

1055 1060 1065

Asp Leu Thr Ala Leu Val Asn Arg Ala Arg Glu Met Met Ala Tyr

1070 1075 1080

Asp Arg Lys Arg Lys Asn Ala Val Pro Arg Ser Ile Leu Asp Met

1085 1090 1095

Leu Ala Arg Leu Gly Leu Thr Leu Lys Trp Gln Met Lys Asp His

1100 1105 1110

Leu Leu Gln Asp Ala Thr Ile Thr Gln Ala Ala Ile Lys His Leu

1115 1120 1125

Asp Lys Val Arg Leu Thr Val Gly Gly Pro Ala Ala Val Thr Glu

1130 1135 1140

Ala Arg Phe Ser Gln Asp Tyr Leu Gln Met Val Ala Ala Val Phe
1145 1150 1155
Asn Gly Ser Val Gln Asn Pro Lys Pro Arg Arg Arg Asp Asp Gly
1160 1165 1170
Asp Ala Trp His Lys Pro Pro Lys Pro Ala Thr Ala Gln Ser Gln
1175 1180 1185

Pro Asp Gln Lys Pro Pro Asn Lys Ala Pro Ser Ala Gly Ser Arg
1190 1195 1200
Leu Pro Pro Pro Gln Val Gly Glu Val Tyr Glu Gly Val Val Val
1205 1210 1215
Lys Val Ile Asp Thr Gly Ser Leu Gly Phe Leu Ala Val Glu Gly
1220 1225 1230
Val Ala Gly Asn Ile Gly Leu His Ile Ser Arg Leu Arg Arg Ile
1235 1240 1245

Arg Glu Asp Ala Ile Ile Val Gly Arg Arg Tyr Arg Phe Arg Val
1250 1255 1260
Glu Ile Tyr Val Pro Pro Lys Ser Asn Thr Ser Lys Leu Asn Ala
1265 1270 1275
Ala Asp Leu Val Arg Ile Asp
1280 1285

<210> 44
<211> 1285
<212> PRT
<213> Rhodobacter capsulatus
<400> 44

Met Gln Ile Gly Lys Val Gln Gly Arg Thr Ile Ser Glu Phe Gly Asp
1 5 10 15

Pro Ala Gly Gly Leu Lys Arg Lys Ile Ser Thr Asp Gly Lys Asn Arg
20 25 30
Lys Glu Leu Pro Ala His Leu Ser Ser Asp Pro Lys Ala Leu Ile Gly
35 40 45
Gln Trp Ile Ser Gly Ile Asp Lys Ile Tyr Arg Lys Pro Asp Ser Arg

50 55 60

Lys Ser Asp Gly Lys Ala Ile His Ser Pro Thr Pro Ser Lys Met Gln
65 70 75 80

Phe Asp Ala Arg Asp Asp Leu Gly Glu Ala Phe Trp Lys Leu Val Ser
85 90 95

Glu Ala Gly Leu Ala Gln Asp Ser Asp Tyr Asp Gln Phe Lys Arg Arg
100 105 110

Leu His Pro Tyr Gly Asp Lys Phe Gln Pro Ala Asp Ser Gly Ala Lys
115 120 125

Leu Lys Phe Glu Ala Asp Pro Pro Glu Pro Gln Ala Phe His Gly Arg
130 135 140

Trp Tyr Gly Ala Met Ser Lys Arg Gly Asn Asp Ala Lys Glu Leu Ala
145 150 155 160

Ala Ala Leu Tyr Glu His Leu His Val Asp Glu Lys Arg Ile Asp Gly
165 170 175

Gln Pro Lys Arg Asn Pro Lys Thr Asp Lys Phe Ala Pro Gly Leu Val
180 185 190

Val Ala Arg Ala Leu Gly Ile Glu Ser Ser Val Leu Pro Arg Gly Met
195 200 205

Ala Arg Leu Ala Arg Asn Trp Gly Glu Glu Glu Ile Gln Thr Tyr Phe
210 215 220

Val Val Asp Val Ala Ala Ser Val Lys Glu Val Ala Lys Ala Ala Val
225 230 235 240

Ser Ala Ala Gln Ala Phe Asp Pro Pro Arg Gln Val Ser Gly Arg Ser
245 250 255

Leu Ser Pro Lys Val Gly Phe Ala Leu Ala Glu His Leu Glu Arg Val
260 265 270

Thr Gly Ser Lys Arg Cys Ser Phe Asp Pro Ala Ala Gly Pro Ser Val
275 280 285

Leu Ala Leu His Asp Glu Val Lys Lys Thr Tyr Lys Arg Leu Cys Ala
290 295 300

Arg Gly Lys Asn Ala Ala Arg Ala Phe Pro Ala Asp Lys Thr Glu Leu
305 310 315 320
Leu Ala Leu Met Arg His Thr His Glu Asn Arg Val Arg Asn Gln Met
325 330 335

Val Arg Met Gly Arg Val Ser Glu Tyr Arg Gly Gln Gln Ala Gly Asp
340 345 350
Leu Ala Gln Ser His Tyr Trp Thr Ser Ala Gly Gln Thr Glu Ile Lys
355 360 365
Glu Ser Glu Ile Phe Val Arg Leu Trp Val Gly Ala Phe Ala Leu Ala
370 375 380
Gly Arg Ser Met Lys Ala Trp Ile Asp Pro Met Gly Lys Ile Val Asn
385 390 395 400

Thr Glu Lys Asn Asp Arg Asp Leu Thr Ala Ala Val Asn Ile Arg Gln
405 410 415
Val Ile Ser Asn Lys Glu Met Val Ala Glu Ala Met Ala Arg Arg Gly
420 425 430
Ile Tyr Phe Gly Glu Thr Pro Glu Leu Asp Arg Leu Gly Ala Glu Gly
435 440 445
Asn Glu Gly Phe Val Phe Ala Leu Leu Arg Tyr Leu Arg Gly Cys Arg
450 455 460

Asn Gln Thr Phe His Leu Gly Ala Arg Ala Gly Phe Leu Lys Glu Ile
465 470 475 480
Arg Lys Glu Leu Glu Lys Thr Arg Trp Gly Lys Ala Lys Glu Ala Glu
485 490 495
His Val Val Leu Thr Asp Lys Thr Val Ala Ala Ile Arg Ala Ile Ile
500 505 510
Asp Asn Asp Ala Lys Ala Leu Gly Ala Arg Leu Leu Ala Asp Leu Ser
515 520 525

Gly Ala Phe Val Ala His Tyr Ala Ser Lys Glu His Phe Ser Thr Leu
530 535 540
Tyr Ser Glu Ile Val Lys Ala Val Lys Asp Ala Pro Glu Val Ser Ser

545 550 555 560

Gly Leu Pro Arg Leu Lys Leu Leu Leu Lys Arg Ala Asp Gly Val Arg

565 570 575

Gly Tyr Val His Gly Leu Arg Asp Thr Arg Lys His Ala Phe Ala Thr

580 585 590

Lys Leu Pro Pro Pro Pro Ala Pro Arg Glu Leu Asp Asp Pro Ala Thr

595 600 605

Lys Ala Arg Tyr Ile Ala Leu Leu Arg Leu Tyr Asp Gly Pro Phe Arg

610 615 620

Ala Tyr Ala Ser Gly Ile Thr Gly Thr Ala Leu Ala Gly Pro Ala Ala

625 630 635 640

Arg Ala Lys Glu Ala Ala Thr Ala Leu Ala Gln Ser Val Asn Val Thr

645 650 655

Lys Ala Tyr Ser Asp Val Met Glu Gly Arg Ser Ser Arg Leu Arg Pro

660 665 670

Pro Asn Asp Gly Glu Thr Leu Arg Glu Tyr Leu Ser Ala Leu Thr Gly

675 680 685

Glu Thr Ala Thr Glu Phe Arg Val Gln Ile Gly Tyr Glu Ser Asp Ser

690 695 700

Glu Asn Ala Arg Lys Gln Ala Glu Phe Ile Glu Asn Tyr Arg Arg Asp

705 710 715 720

Met Leu Ala Phe Met Phe Glu Asp Tyr Ile Arg Ala Lys Gly Phe Asp

725 730 735

Trp Ile Leu Lys Ile Glu Pro Gly Ala Thr Ala Met Thr Arg Ala Pro

740 745 750

Val Leu Pro Glu Pro Ile Asp Thr Arg Gly Gln Tyr Glu His Trp Gln

755 760 765

Ala Ala Leu Tyr Leu Val Met His Phe Val Pro Ala Ser Asp Val Ser

770 775 780

Asn Leu Leu His Gln Leu Arg Lys Trp Glu Ala Leu Gln Gly Lys Tyr

785 790 795 800

Glu Leu Val Gln Asp Gly Asp Ala Thr Asp Gln Ala Asp Ala Arg Arg
805 810 815
Glu Ala Leu Asp Leu Val Lys Arg Phe Arg Asp Val Leu Val Leu Phe
820 825 830
Leu Lys Thr Gly Glu Ala Arg Phe Glu Gly Arg Ala Ala Pro Phe Asp
835 840 845

Leu Lys Pro Phe Arg Ala Leu Phe Ala Asn Pro Ala Thr Phe Asp Arg
850 855 860
Leu Phe Met Ala Thr Pro Thr Thr Ala Arg Pro Ala Glu Asp Asp Pro
865 870 875 880
Glu Gly Asp Gly Ala Ser Glu Pro Glu Leu Arg Val Ala Arg Thr Leu
885 890 895
Arg Gly Leu Arg Gln Ile Ala Arg Tyr Asn His Met Ala Val Leu Ser
900 905 910

Asp Leu Phe Ala Lys His Lys Val Arg Asp Glu Glu Val Ala Arg Leu
915 920 925
Ala Glu Ile Glu Asp Glu Thr Gln Glu Lys Ser Gln Ile Val Ala Ala
930 935 940
Gln Glu Leu Arg Thr Asp Leu His Asp Lys Val Met Lys Cys His Pro
945 950 955 960
Lys Thr Ile Ser Pro Glu Glu Arg Gln Ser Tyr Ala Ala Ala Ile Lys
965 970 975

Thr Ile Glu Glu His Arg Phe Leu Val Gly Arg Val Tyr Leu Gly Asp
980 985 990
His Leu Arg Leu His Arg Leu Met Met Asp Val Ile Gly Arg Leu Ile
995 1000 1005
Asp Tyr Ala Gly Ala Tyr Glu Arg Asp Thr Gly Thr Phe Leu Ile
1010 1015 1020
Asn Ala Ser Lys Gln Leu Gly Ala Gly Ala Asp Trp Ala Val Thr
1025 1030 1035

Ile Ala Gly Ala Ala Asn Thr Asp Ala Arg Thr Gln Thr Arg Lys

1040 1045 1050

Asp Leu Ala His Phe Asn Val Leu Asp Arg Ala Asp Gly Thr Pro

1055 1060 1065

Asp Leu Thr Ala Leu Val Asn Arg Ala Arg Glu Met Met Ala Tyr

1070 1075 1080

Asp Arg Lys Arg Lys Asn Ala Val Pro Arg Ser Ile Leu Asp Met

1085 1090 1095

Leu Ala Arg Leu Gly Leu Thr Leu Lys Trp Gln Met Lys Asp His

1100 1105 1110

Leu Leu Gln Asp Ala Thr Ile Thr Gln Ala Ala Ile Lys His Leu

1115 1120 1125

Asp Lys Val Arg Leu Thr Val Gly Gly Pro Ala Ala Val Thr Glu

1130 1135 1140

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

1145 1150 1155

Asn Gly Ser Val Gln Asn Pro Lys Pro Arg Arg Arg Asp Asp Gly

1160 1165 1170

Asp Ala Trp His Lys Pro Pro Lys Pro Ala Thr Ala Gln Ser Gln

1175 1180 1185

Pro Asp Gln Lys Pro Pro Asn Lys Ala Pro Ser Ala Gly Ser Arg

1190 1195 1200

Leu Pro Pro Pro Gln Val Gly Glu Val Tyr Glu Gly Val Val Val

1205 1210 1215

Lys Val Ile Asp Thr Gly Ser Leu Gly Phe Leu Ala Val Glu Gly

1220 1225 1230

Val Ala Gly Asn Ile Gly Leu His Ile Ser Arg Leu Arg Arg Ile

1235 1240 1245

Arg Glu Asp Ala Ile Ile Val Gly Arg Arg Tyr Arg Phe Arg Val

1250 1255 1260

Glu Ile Tyr Val Pro Pro Lys Ser Asn Thr Ser Lys Leu Asn Ala

1265 1270 1275

Ala Asp Leu Val Arg Ile Asp
1280 1285

<210> 45
<211> 1285
<212> PRT
<213> Rhodobacter capsulatus
<400> 45

Met Gln Ile Gly Lys Val Gln Gly Arg Thr Ile Ser Glu Phe Gly Asp
1 5 10 15
Pro Ala Gly Gly Leu Lys Arg Lys Ile Ser Thr Asp Gly Lys Asn Arg
20 25 30
Lys Glu Leu Pro Ala His Leu Ser Ser Asp Pro Lys Ala Leu Ile Gly
35 40 45

Gln Trp Ile Ser Gly Ile Asp Lys Ile Tyr Arg Lys Pro Asp Ser Arg
50 55 60
Lys Ser Asp Gly Lys Ala Ile His Ser Pro Thr Pro Ser Lys Met Gln
65 70 75 80
Phe Asp Ala Arg Asp Asp Leu Gly Glu Ala Phe Trp Lys Leu Val Ser
85 90 95
Glu Ala Gly Leu Ala Gln Asp Ser Asp Tyr Asp Gln Phe Lys Arg Arg
100 105 110

Leu His Pro Tyr Gly Asp Lys Phe Gln Pro Ala Asp Ser Gly Ala Lys
115 120 125
Leu Lys Phe Glu Ala Asp Pro Pro Glu Pro Gln Ala Phe His Gly Arg
130 135 140
Trp Tyr Gly Ala Met Ser Lys Arg Gly Asn Asp Ala Lys Glu Leu Ala
145 150 155 160
Ala Ala Leu Tyr Glu His Leu His Val Asp Glu Lys Arg Ile Asp Gly
165 170 175

Gln Pro Lys Arg Asn Pro Lys Thr Asp Lys Phe Ala Pro Gly Leu Val
180 185 190
Val Ala Arg Ala Leu Gly Ile Glu Ser Ser Val Leu Pro Arg Gly Met

195 200 205

Ala Arg Leu Ala Arg Asn Trp Gly Glu Glu Glu Ile Gln Thr Tyr Phe

210 215 220

Val Val Asp Val Ala Ala Ser Val Lys Glu Val Ala Lys Ala Ala Val

225 230 235 240

Ser Ala Ala Gln Ala Phe Asp Pro Pro Arg Gln Val Ser Gly Arg Ser

245 250 255

Leu Ser Pro Lys Val Gly Phe Ala Leu Ala Glu His Leu Glu Arg Val

260 265 270

Thr Gly Ser Lys Arg Cys Ser Phe Asp Pro Ala Ala Gly Pro Ser Val

275 280 285

Leu Ala Leu His Asp Glu Val Lys Lys Thr Tyr Lys Arg Leu Cys Ala

290 295 300

Arg Gly Lys Asn Ala Ala Arg Ala Phe Pro Ala Asp Lys Thr Glu Leu

305 310 315 320

Leu Ala Leu Met Arg His Thr His Glu Asn Arg Val Arg Asn Gln Met

325 330 335

Val Arg Met Gly Arg Val Ser Glu Tyr Arg Gly Gln Gln Ala Gly Asp

340 345 350

Leu Ala Gln Ser His Tyr Trp Thr Ser Ala Gly Gln Thr Glu Ile Lys

355 360 365

Glu Ser Glu Ile Phe Val Arg Leu Trp Val Gly Ala Phe Ala Leu Ala

370 375 380

Gly Arg Ser Met Lys Ala Trp Ile Asp Pro Met Gly Lys Ile Val Asn

385 390 395 400

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

405 410 415

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

420 425 430

Ile Tyr Phe Gly Glu Thr Pro Glu Leu Asp Arg Leu Gly Ala Glu Gly

435 440 445

Asn Glu Gly Phe Val Phe Ala Leu Leu Arg Tyr Leu Arg Gly Cys Arg
450 455 460
Asn Gln Thr Phe His Leu Gly Ala Arg Ala Gly Phe Leu Lys Glu Ile
465 470 475 480
Arg Lys Glu Leu Glu Lys Thr Arg Trp Gly Lys Ala Lys Glu Ala Glu
485 490 495

His Val Val Leu Thr Asp Lys Thr Val Ala Ala Ile Arg Ala Ile Ile
500 505 510
Asp Asn Asp Ala Lys Ala Leu Gly Ala Arg Leu Leu Ala Asp Leu Ser
515 520 525
Gly Ala Phe Val Ala His Tyr Ala Ser Lys Glu His Phe Ser Thr Leu
530 535 540
Tyr Ser Glu Ile Val Lys Ala Val Lys Asp Ala Pro Glu Val Ser Ser
545 550 555 560

Gly Leu Pro Arg Leu Lys Leu Leu Leu Lys Arg Ala Asp Gly Val Arg
565 570 575
Gly Tyr Val His Gly Leu Arg Asp Thr Arg Lys His Ala Phe Ala Thr
580 585 590
Lys Leu Pro Pro Pro Pro Ala Pro Arg Glu Leu Asp Asp Pro Ala Thr
595 600 605
Lys Ala Arg Tyr Ile Ala Leu Leu Arg Leu Tyr Asp Gly Pro Phe Arg
610 615 620

Ala Tyr Ala Ser Gly Ile Thr Gly Thr Ala Leu Ala Gly Pro Ala Ala
625 630 635 640
Arg Ala Lys Glu Ala Ala Thr Ala Leu Ala Gln Ser Val Asn Val Thr
645 650 655
Lys Ala Tyr Ser Asp Val Met Glu Gly Arg Ser Ser Arg Leu Arg Pro
660 665 670
Pro Asn Asp Gly Glu Thr Leu Arg Glu Tyr Leu Ser Ala Leu Thr Gly
675 680 685

Glu Thr Ala Thr Glu Phe Arg Val Gln Ile Gly Tyr Glu Ser Asp Ser

690 695 700

Glu Asn Ala Arg Lys Gln Ala Glu Phe Ile Glu Asn Tyr Arg Arg Asp

705 710 715 720

Met Leu Ala Phe Met Phe Glu Asp Tyr Ile Arg Ala Lys Gly Phe Asp

725 730 735

Trp Ile Leu Lys Ile Glu Pro Gly Ala Thr Ala Met Thr Arg Ala Pro

740 745 750

Val Leu Pro Glu Pro Ile Asp Thr Arg Gly Gln Tyr Glu His Trp Gln

755 760 765

Ala Ala Leu Tyr Leu Val Met His Phe Val Pro Ala Ser Asp Val Ser

770 775 780

Asn Leu Leu His Gln Leu Arg Lys Trp Glu Ala Leu Gln Gly Lys Tyr

785 790 795 800

Glu Leu Val Gln Asp Gly Asp Ala Thr Asp Gln Ala Asp Ala Arg Arg

805 810 815

Glu Ala Leu Asp Leu Val Lys Arg Phe Arg Asp Val Leu Val Leu Phe

820 825 830

Leu Lys Thr Gly Glu Ala Arg Phe Glu Gly Arg Ala Ala Pro Phe Asp

835 840 845

Leu Lys Pro Phe Arg Ala Leu Phe Ala Asn Pro Ala Thr Phe Asp Arg

850 855 860

Leu Phe Met Ala Thr Pro Thr Thr Ala Arg Pro Ala Glu Asp Asp Pro

865 870 875 880

Glu Gly Asp Gly Ala Ser Glu Pro Glu Leu Arg Val Ala Arg Thr Leu

885 890 895

Arg Gly Leu Arg Gln Ile Ala Arg Tyr Asn His Met Ala Val Leu Ser

900 905 910

Asp Leu Phe Ala Lys His Lys Val Arg Asp Glu Glu Val Ala Arg Leu

915 920 925

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

930 935 940

Gln Glu Leu Arg Thr Asp Leu His Asp Lys Val Met Lys Cys His Pro
945 950 955 960
Lys Thr Ile Ser Pro Glu Glu Arg Gln Ser Tyr Ala Ala Ala Ile Lys
965 970 975
Thr Ile Glu Glu His Arg Phe Leu Val Gly Arg Val Tyr Leu Gly Asp
980 985 990
His Leu Arg Leu His Arg Leu Met Met Asp Val Ile Gly Arg Leu Ile
995 1000 1005

Asp Tyr Ala Gly Ala Tyr Glu Arg Asp Thr Gly Thr Phe Leu Ile
1010 1015 1020
Asn Ala Ser Lys Gln Leu Gly Ala Gly Ala Asp Trp Ala Val Thr
1025 1030 1035
Ile Ala Gly Ala Ala Asn Thr Asp Ala Arg Thr Gln Thr Arg Lys
1040 1045 1050
Asp Leu Ala His Phe Asn Val Leu Asp Arg Ala Asp Gly Thr Pro
1055 1060 1065

Asp Leu Thr Ala Leu Val Asn Arg Ala Arg Glu Met Met Ala Tyr
1070 1075 1080
Asp Arg Lys Arg Lys Asn Ala Val Pro Arg Ser Ile Leu Asp Met
1085 1090 1095
Leu Ala Arg Leu Gly Leu Thr Leu Lys Trp Gln Met Lys Asp His
1100 1105 1110
Leu Leu Gln Asp Ala Thr Ile Thr Gln Ala Ala Ile Lys His Leu
1115 1120 1125

Asp Lys Val Arg Leu Thr Val Gly Gly Pro Ala Ala Val Thr Glu
1130 1135 1140
Ala Arg Phe Ser Gln Asp Tyr Leu Gln Met Val Ala Ala Val Phe
1145 1150 1155
Asn Gly Ser Val Gln Asn Pro Lys Pro Arg Arg Arg Asp Asp Gly
1160 1165 1170
Asp Ala Trp His Lys Pro Pro Lys Pro Ala Thr Ala Gln Ser Gln

1175 1180 1185

Pro Asp Gln Lys Pro Pro Asn Lys Ala Pro Ser Ala Gly Ser Arg
1190 1195 1200
Leu Pro Pro Pro Gln Val Gly Glu Val Tyr Glu Gly Val Val Val
1205 1210 1215
Lys Val Ile Asp Thr Gly Ser Leu Gly Phe Leu Ala Val Glu Gly
1220 1225 1230
Val Ala Gly Asn Ile Gly Leu His Ile Ser Arg Leu Arg Arg Ile
1235 1240 1245

Arg Glu Asp Ala Ile Ile Val Gly Arg Arg Tyr Arg Phe Arg Val
1250 1255 1260
Glu Ile Tyr Val Pro Pro Lys Ser Asn Thr Ser Lys Leu Asn Ala
1265 1270 1275
Ala Asp Leu Val Arg Ile Asp
1280 1285

<210> 46
<211> 1389
<212> PRT
<213> Leptotrichia shahii
<400> 46
Met Gly Asn Leu Phe Gly His Lys Arg Trp Tyr Glu Val Arg Asp Lys
1 5 10 15

Lys Asp Phe Lys Ile Lys Arg Lys Val Lys Val Lys Arg Asn Tyr Asp
20 25 30
Gly Asn Lys Tyr Ile Leu Asn Ile Asn Glu Asn Asn Asn Lys Glu Lys
35 40 45
Ile Asp Asn Asn Lys Phe Ile Arg Lys Tyr Ile Asn Tyr Lys Lys Asn
50 55 60
Asp Asn Ile Leu Lys Glu Phe Thr Arg Lys Phe His Ala Gly Asn Ile
65 70 75 80

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

85 90 95

Asp Asp Phe Leu Glu Thr Glu Glu Val Val Leu Tyr Ile Glu Ala Tyr

100 105 110

Gly Lys Ser Glu Lys Leu Lys Ala Leu Gly Ile Thr Lys Lys Lys Ile

115 120 125

Ile Asp Glu Ala Ile Arg Gln Gly Ile Thr Lys Asp Asp Lys Lys Ile

130 135 140

Glu Ile Lys Arg Gln Glu Asn Glu Glu Glu Ile Glu Ile Asp Ile Arg

145 150 155 160

Asp Glu Tyr Thr Asn Lys Thr Leu Asn Asp Cys Ser Ile Ile Leu Arg

165 170 175

Ile Ile Glu Asn Asp Glu Leu Glu Thr Lys Lys Ser Ile Tyr Glu Ile

180 185 190

Phe Lys Asn Ile Asn Met Ser Leu Tyr Lys Ile Ile Glu Lys Ile Ile

195 200 205

Glu Asn Glu Thr Glu Lys Val Phe Glu Asn Arg Tyr Tyr Glu Glu His

210 215 220

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

225 230 235 240

Asn Phe Met Glu Ile Arg Glu Lys Ile Lys Ser Asn Leu Glu Ile Leu

245 250 255

Gly Phe Val Lys Phe Tyr Leu Asn Val Gly Gly Asp Lys Lys Lys Ser

260 265 270

Lys Asn Lys Lys Met Leu Val Glu Lys Ile Leu Asn Ile Asn Val Asp

275 280 285

Leu Thr Val Glu Asp Ile Ala Asp Phe Val Ile Lys Glu Leu Glu Phe

290 295 300

Trp Asn Ile Thr Lys Arg Ile Glu Lys Val Lys Lys Val Asn Asn Glu

305 310 315 320

Phe Leu Glu Lys Arg Arg Asn Arg Thr Tyr Ile Lys Ser Tyr Val Leu

325 330 335

Leu Asp Lys His Glu Lys Phe Lys Ile Glu Arg Glu Asn Lys Lys Asp
340 345 350

Lys Ile Val Lys Phe Phe Val Glu Asn Ile Lys Asn Asn Ser Ile Lys
355 360 365

Glu Lys Ile Glu Lys Ile Leu Ala Glu Phe Lys Ile Asp Glu Leu Ile
370 375 380

Lys Lys Leu Glu Lys Glu Leu Lys Lys Gly Asn Cys Asp Thr Glu Ile
385 390 395 400

Phe Gly Ile Phe Lys Lys His Tyr Lys Val Asn Phe Asp Ser Lys Lys
405 410 415

Phe Ser Lys Lys Ser Asp Glu Glu Lys Glu Leu Tyr Lys Ile Ile Tyr
420 425 430

Arg Tyr Leu Lys Gly Arg Ile Glu Lys Ile Leu Val Asn Glu Gln Lys
435 440 445

Val Arg Leu Lys Lys Met Glu Lys Ile Glu Ile Glu Lys Ile Leu Asn
450 455 460

Glu Ser Ile Leu Ser Glu Lys Ile Leu Lys Arg Val Lys Gln Tyr Thr
465 470 475 480

Leu Glu His Ile Met Tyr Leu Gly Lys Leu Arg His Asn Asp Ile Asp
485 490 495

Met Thr Thr Val Asn Thr Asp Asp Phe Ser Arg Leu His Ala Lys Glu
500 505 510

Glu Leu Asp Leu Glu Leu Ile Thr Phe Phe Ala Ser Thr Asn Met Glu
515 520 525

Leu Asn Lys Ile Phe Ser Arg Glu Asn Ile Asn Asn Asp Glu Asn Ile
530 535 540

Asp Phe Phe Gly Gly Asp Arg Glu Lys Asn Tyr Val Leu Asp Lys Lys
545 550 555 560

Ile Leu Asn Ser Lys Ile Lys Ile Ile Arg Asp Leu Asp Phe Ile Asp
565 570 575

Asn Lys Asn Asn Ile Thr Asn Asn Phe Ile Arg Lys Phe Thr Lys Ile

580 585 590

Gly Thr Asn Glu Arg Asn Arg Ile Leu His Ala Ile Ser Lys Glu Arg
595 600 605

Asp Leu Gln Gly Thr Gln Asp Asp Tyr Asn Lys Val Ile Asn Ile Ile
610 615 620

Gln Asn Leu Lys Ile Ser Asp Glu Glu Val Ser Lys Ala Leu Asn Leu
625 630 635 640

Asp Val Val Phe Lys Asp Lys Lys Asn Ile Ile Thr Lys Ile Asn Asp
645 650 655

Ile Lys Ile Ser Glu Glu Asn Asn Asn Asp Ile Lys Tyr Leu Pro Ser
660 665 670

Phe Ser Lys Val Leu Pro Glu Ile Leu Asn Leu Tyr Arg Asn Asn Pro
675 680 685

Lys Asn Glu Pro Phe Asp Thr Ile Glu Thr Glu Lys Ile Val Leu Asn
690 695 700

Ala Leu Ile Tyr Val Asn Lys Glu Leu Tyr Lys Lys Leu Ile Leu Glu
705 710 715 720

Asp Asp Leu Glu Glu Asn Glu Ser Lys Asn Ile Phe Leu Gln Glu Leu
725 730 735

Lys Lys Thr Leu Gly Asn Ile Asp Glu Ile Asp Glu Asn Ile Ile Glu
740 745 750

Asn Tyr Tyr Lys Asn Ala Gln Ile Ser Ala Ser Lys Gly Asn Asn Lys
755 760 765

Ala Ile Lys Lys Tyr Gln Lys Lys Val Ile Glu Cys Tyr Ile Gly Tyr
770 775 780

Leu Arg Lys Asn Tyr Glu Glu Leu Phe Asp Phe Ser Asp Phe Lys Met
785 790 795 800

Asn Ile Gln Glu Ile Lys Lys Gln Ile Lys Asp Ile Asn Asp Asn Lys
805 810 815

Thr Tyr Glu Arg Ile Thr Val Lys Thr Ser Asp Lys Thr Ile Val Ile
820 825 830

Asn Asp Asp Phe Glu Tyr Ile Ile Ser Ile Phe Ala Leu Leu Asn Ser
835 840 845

Asn Ala Val Ile Asn Lys Ile Arg Asn Arg Phe Phe Ala Thr Ser Val
850 855 860
Trp Leu Asn Thr Ser Glu Tyr Gln Asn Ile Ile Asp Ile Leu Asp Glu
865 870 875 880
Ile Met Gln Leu Asn Thr Leu Arg Asn Glu Cys Ile Thr Glu Asn Trp
885 890 895
Asn Leu Asn Leu Glu Glu Phe Ile Gln Lys Met Lys Glu Ile Glu Lys
900 905 910

Asp Phe Asp Asp Phe Lys Ile Gln Thr Lys Lys Glu Ile Phe Asn Asn
915 920 925
Tyr Tyr Glu Asp Ile Lys Asn Asn Ile Leu Thr Glu Phe Lys Asp Asp
930 935 940
Ile Asn Gly Cys Asp Val Leu Glu Lys Lys Leu Glu Lys Ile Val Ile
945 950 955 960
Phe Asp Asp Glu Thr Lys Phe Glu Ile Asp Lys Lys Ser Asn Ile Leu
965 970 975

Gln Asp Glu Gln Arg Lys Leu Ser Asn Ile Asn Lys Lys Asp Leu Lys
980 985 990
Lys Lys Val Asp Gln Tyr Ile Lys Asp Lys Asp Gln Glu Ile Lys Ser
995 1000 1005
Lys Ile Leu Cys Arg Ile Ile Phe Asn Ser Asp Phe Leu Lys Lys
1010 1015 1020
Tyr Lys Lys Glu Ile Asp Asn Leu Ile Glu Asp Met Glu Ser Glu
1025 1030 1035

Asn Glu Asn Lys Phe Gln Glu Ile Tyr Tyr Pro Lys Glu Arg Lys
1040 1045 1050
Asn Glu Leu Tyr Ile Tyr Lys Lys Asn Leu Phe Leu Asn Ile Gly
1055 1060 1065
Asn Pro Asn Phe Asp Lys Ile Tyr Gly Leu Ile Ser Asn Asp Ile

1070 1075 1080

Lys Met Ala Asp Ala Lys Phe Leu Phe Asn Ile Asp Gly Lys Asn

1085 1090 1095

Ile Arg Lys Asn Lys Ile Ser Glu Ile Asp Ala Ile Leu Lys Asn

1100 1105 1110

Leu Asn Asp Lys Leu Asn Gly Tyr Ser Lys Glu Tyr Lys Glu Lys

1115 1120 1125

Tyr Ile Lys Lys Leu Lys Glu Asn Asp Asp Phe Phe Ala Lys Asn

1130 1135 1140

Ile Gln Asn Lys Asn Tyr Lys Ser Phe Glu Lys Asp Tyr Asn Arg

1145 1150 1155

Val Ser Glu Tyr Lys Lys Ile Arg Asp Leu Val Glu Phe Asn Tyr

1160 1165 1170

Leu Asn Lys Ile Glu Ser Tyr Leu Ile Asp Ile Asn Trp Lys Leu

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

Thr Gly Ile Ser Arg Ala Tyr Pro Lys Arg Asn Gly Ser Asp Gly

1220 1225 1230

Phe Tyr Thr Thr Thr Ala Tyr Tyr Lys Phe Phe Asp Glu Glu Ser

1235 1240 1245

Tyr Lys Lys Phe Glu Lys Ile Cys Tyr Gly Phe Gly Ile Asp Leu

1250 1255 1260

Ser Glu Asn Ser Glu Ile Asn Lys Pro Glu Asn Glu Ser Ile Arg

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

Tyr Ser Thr Arg Tyr Asn Asn Ser Thr Tyr Ala Ser Val Phe Glu
1310 1315 1320
Val Phe Lys Lys Asp Val Asn Leu Asp Tyr Asp Glu Leu Lys Lys
1325 1330 1335

Lys Phe Lys Leu Ile Gly Asn Asn Asp Ile Leu Glu Arg Leu Met
1340 1345 1350
Lys Pro Lys Lys Val Ser Val Leu Glu Leu Glu Ser Tyr Asn Ser
1355 1360 1365
Asp Tyr Ile Lys Asn Leu Ile Ile Glu Leu Leu Thr Lys Ile Glu
1370 1375 1380
Asn Thr Asn Asp Thr Leu
1385

<210> 47
<211> 1197
<212> PRT
<213> Leptotrichia wadei

<400> 47
Met Lys Val Thr Lys Val Asp Gly Ile Ser His Lys Lys Tyr Ile Glu
1 5 10 15
Glu Gly Lys Leu Val Lys Ser Thr Ser Glu Glu Asn Arg Thr Ser Glu
20 25 30
Arg Leu Ser Glu Leu Leu Ser Ile Arg Leu Asp Ile Tyr Ile Lys Asn
35 40 45
Pro Asp Asn Ala Ser Glu Glu Glu Asn Arg Ile Arg Arg Glu Asn Leu
50 55 60

Lys Lys Phe Phe Ser Asn Lys Val Leu His Leu Lys Asp Ser Val Leu
65 70 75 80
Tyr Leu Lys Asn Arg Lys Glu Lys Asn Ala Val Gln Asp Lys Asn Tyr
85 90 95
Ser Glu Glu Asp Ile Ser Glu Tyr Asp Leu Lys Asn Lys Asn Ser Phe
100 105 110
Ser Val Leu Lys Lys Ile Leu Leu Asn Glu Asp Val Asn Ser Glu Glu

115 120 125

Leu Glu Ile Phe Arg Lys Asp Val Glu Ala Lys Leu Asn Lys Ile Asn
130 135 140
Ser Leu Lys Tyr Ser Phe Glu Glu Asn Lys Ala Asn Tyr Gln Lys Ile
145 150 155 160
Asn Glu Asn Asn Val Glu Lys Val Gly Gly Lys Ser Lys Arg Asn Ile
165 170 175
Ile Tyr Asp Tyr Tyr Arg Glu Ser Ala Lys Arg Asn Asp Tyr Ile Asn
180 185 190

Asn Val Gln Glu Ala Phe Asp Lys Leu Tyr Lys Lys Glu Asp Ile Glu
195 200 205
Lys Leu Phe Phe Leu Ile Glu Asn Ser Lys Lys His Glu Lys Tyr Lys
210 215 220
Ile Arg Glu Tyr Tyr His Lys Ile Ile Gly Arg Lys Asn Asp Lys Glu
225 230 235 240
Asn Phe Ala Lys Ile Ile Tyr Glu Glu Ile Gln Asn Val Asn Asn Ile
245 250 255

Lys Glu Leu Ile Glu Lys Ile Pro Asp Met Ser Glu Leu Lys Lys Ser
260 265 270
Gln Val Phe Tyr Lys Tyr Tyr Leu Asp Lys Glu Glu Leu Asn Asp Lys
275 280 285
Asn Ile Lys Tyr Ala Phe Cys His Phe Val Glu Ile Glu Met Ser Gln
290 295 300
Leu Leu Lys Asn Tyr Val Tyr Lys Arg Leu Ser Asn Ile Ser Asn Asp
305 310 315 320

Lys Ile Lys Arg Ile Phe Glu Tyr Gln Asn Leu Lys Lys Leu Ile Glu
325 330 335
Asn Lys Leu Leu Asn Lys Leu Asp Thr Tyr Val Arg Asn Cys Gly Lys
340 345 350
Tyr Asn Tyr Tyr Leu Gln Val Gly Glu Ile Ala Thr Ser Asp Phe Ile
355 360 365

Ala Arg Asn Arg Gln Asn Glu Ala Phe Leu Arg Asn Ile Ile Gly Val
370 375 380

Ser Ser Val Ala Tyr Phe Ser Leu Arg Asn Ile Leu Glu Thr Glu Asn
385 390 395 400
Glu Asn Asp Ile Thr Gly Arg Met Arg Gly Lys Thr Val Lys Asn Asn
405 410 415
Lys Gly Glu Glu Lys Tyr Val Ser Gly Glu Val Asp Lys Ile Tyr Asn
420 425 430
Glu Asn Lys Gln Asn Glu Val Lys Glu Asn Leu Lys Met Phe Tyr Ser
435 440 445

Tyr Asp Phe Asn Met Asp Asn Lys Asn Glu Ile Glu Asp Phe Phe Ala
450 455 460
Asn Ile Asp Glu Ala Ile Ser Ser Ile Arg His Gly Ile Val His Phe
465 470 475 480
Asn Leu Glu Leu Glu Gly Lys Asp Ile Phe Ala Phe Lys Asn Ile Ala
485 490 495
Pro Ser Glu Ile Ser Lys Lys Met Phe Gln Asn Glu Ile Asn Glu Lys
500 505 510

Lys Leu Lys Leu Lys Ile Phe Lys Gln Leu Asn Ser Ala Asn Val Phe
515 520 525
Asn Tyr Tyr Glu Lys Asp Val Ile Ile Lys Tyr Leu Lys Asn Thr Lys
530 535 540
Phe Asn Phe Val Asn Lys Asn Ile Pro Phe Val Pro Ser Phe Thr Lys
545 550 555 560
Leu Tyr Asn Lys Ile Glu Asp Leu Arg Asn Thr Leu Lys Phe Phe Trp
565 570 575

Ser Val Pro Lys Asp Lys Glu Glu Lys Asp Ala Gln Ile Tyr Leu Leu
580 585 590
Lys Asn Ile Tyr Tyr Gly Glu Phe Leu Asn Lys Phe Val Lys Asn Ser
595 600 605
Lys Val Phe Phe Lys Ile Thr Asn Glu Val Ile Lys Ile Asn Lys Gln

610 615 620

Arg Asn Gln Lys Thr Gly His Tyr Lys Tyr Gln Lys Phe Glu Asn Ile
625 630 635 640

Glu Lys Thr Val Pro Val Glu Tyr Leu Ala Ile Ile Gln Ser Arg Glu
645 650 655

Met Ile Asn Asn Gln Asp Lys Glu Glu Lys Asn Thr Tyr Ile Asp Phe
660 665 670

Ile Gln Gln Ile Phe Leu Lys Gly Phe Ile Asp Tyr Leu Asn Lys Asn
675 680 685

Asn Leu Lys Tyr Ile Glu Ser Asn Asn Asn Asn Asp Asn Asn Asp Ile
690 695 700

Phe Ser Lys Ile Lys Ile Lys Lys Asp Asn Lys Glu Lys Tyr Asp Lys
705 710 715 720

Ile Leu Lys Asn Tyr Glu Lys His Asn Arg Asn Lys Glu Ile Pro His
725 730 735

Glu Ile Asn Glu Phe Val Arg Glu Ile Lys Leu Gly Lys Ile Leu Lys
740 745 750

Tyr Thr Glu Asn Leu Asn Met Phe Tyr Leu Ile Leu Lys Leu Leu Asn
755 760 765

His Lys Glu Leu Thr Asn Leu Lys Gly Ser Leu Glu Lys Tyr Gln Ser
770 775 780

Ala Asn Lys Glu Glu Thr Phe Ser Asp Glu Leu Glu Leu Ile Asn Leu
785 790 795 800

Leu Asn Leu Asp Asn Asn Arg Val Thr Glu Asp Phe Glu Leu Glu Ala
805 810 815

Asn Glu Ile Gly Lys Phe Leu Asp Phe Asn Glu Asn Lys Ile Lys Asp
820 825 830

Arg Lys Glu Leu Lys Lys Phe Asp Thr Asn Lys Ile Tyr Phe Asp Gly
835 840 845

Glu Asn Ile Ile Lys His Arg Ala Phe Tyr Asn Ile Lys Lys Tyr Gly
850 855 860

Met Leu Asn Leu Leu Glu Lys Ile Ala Asp Lys Ala Lys Tyr Lys Ile
865 870 875 880
Ser Leu Lys Glu Leu Lys Glu Tyr Ser Asn Lys Lys Asn Glu Ile Glu
885 890 895

Lys Asn Tyr Thr Met Gln Gln Asn Leu His Arg Lys Tyr Ala Arg Pro
900 905 910
Lys Lys Asp Glu Lys Phe Asn Asp Glu Asp Tyr Lys Glu Tyr Glu Lys
915 920 925
Ala Ile Gly Asn Ile Gln Lys Tyr Thr His Leu Lys Asn Lys Val Glu
930 935 940
Phe Asn Glu Leu Asn Leu Leu Gln Gly Leu Leu Leu Lys Ile Leu His
945 950 955 960

Arg Leu Val Gly Tyr Thr Ser Ile Trp Glu Arg Asp Leu Arg Phe Arg
965 970 975
Leu Lys Gly Glu Phe Pro Glu Asn His Tyr Ile Glu Glu Ile Phe Asn
980 985 990
Phe Asp Asn Ser Lys Asn Val Lys Tyr Lys Ser Gly Gln Ile Val Glu
995 1000 1005
Lys Tyr Ile Asn Phe Tyr Lys Glu Leu Tyr Lys Asp Asn Val Glu
1010 1015 1020

Lys Arg Ser Ile Tyr Ser Asp Lys Lys Val Lys Lys Leu Lys Gln
1025 1030 1035
Glu Lys Lys Asp Leu Tyr Ile Arg Asn Tyr Ile Ala His Phe Asn
1040 1045 1050
Tyr Ile Pro His Ala Glu Ile Ser Leu Leu Glu Val Leu Glu Asn
1055 1060 1065
Leu Arg Lys Leu Leu Ser Tyr Asp Arg Lys Leu Lys Asn Ala Ile
1070 1075 1080

Met Lys Ser Ile Val Asp Ile Leu Lys Glu Tyr Gly Phe Val Ala
1085 1090 1095
Thr Phe Lys Ile Gly Ala Asp Lys Lys Ile Glu Ile Gln Thr Leu

1100 1105 1110

Glu Ser Glu Lys Ile Val His Leu Lys Asn Leu Lys Lys Lys Lys

1115 1120 1125

Leu Met Thr Asp Arg Asn Ser Glu Glu Leu Cys Glu Leu Val Lys

1130 1135 1140

Val Met Phe Glu Tyr Lys Ala Leu Glu Lys Arg Pro Ala Ala Thr

1145 1150 1155

Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser Tyr Pro Tyr

1160 1165 1170

Asp Val Pro Asp Tyr Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

1175 1180 1185

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

1190 1195

<210> 48

<211> 1120

<212> PRT

<213> Listeria seeligeri

<400> 48

Met Trp Ile Ser Ile Lys Thr Leu Ile His His Leu Gly Val Leu Phe

1 5 10 15

Phe Cys Asp Tyr Met Tyr Asn Arg Arg Glu Lys Lys Ile Ile Glu Val

20 25 30

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

35 40 45

Ile Ser Arg Asp Lys Asn Gly Gly Lys Leu Val Tyr Glu Asn Glu Met

50 55 60

Gln Asp Asn Thr Glu Gln Ile Met His His Lys Lys Ser Ser Phe Tyr

65 70 75 80

Lys Ser Val Val Asn Lys Thr Ile Cys Arg Pro Glu Gln Lys Gln Met

85 90 95

Lys Lys Leu Val His Gly Leu Leu Gln Glu Asn Ser Gln Glu Lys Ile

100 105 110

Lys Val Ser Asp Val Thr Lys Leu Asn Ile Ser Asn Phe Leu Asn His
115 120 125

Arg Phe Lys Lys Ser Leu Tyr Tyr Phe Pro Glu Asn Ser Pro Asp Lys
130 135 140
Ser Glu Glu Tyr Arg Ile Glu Ile Asn Leu Ser Gln Leu Leu Glu Asp
145 150 155 160
Ser Leu Lys Lys Gln Gln Gly Thr Phe Ile Cys Trp Glu Ser Phe Ser
165 170 175
Lys Asp Met Glu Leu Tyr Ile Asn Trp Ala Glu Asn Tyr Ile Ser Ser
180 185 190

Lys Thr Lys Leu Ile Lys Lys Ser Ile Arg Asn Asn Arg Ile Gln Ser
195 200 205
Thr Glu Ser Arg Ser Gly Gln Leu Met Asp Arg Tyr Met Lys Asp Ile
210 215 220
Leu Asn Lys Asn Lys Pro Phe Asp Ile Gln Ser Val Ser Glu Lys Tyr
225 230 235 240
Gln Leu Glu Lys Leu Thr Ser Ala Leu Lys Ala Thr Phe Lys Glu Ala
245 250 255

Lys Lys Asn Asp Lys Glu Ile Asn Tyr Lys Leu Lys Ser Thr Leu Gln
260 265 270
Asn His Glu Arg Gln Ile Ile Glu Glu Leu Lys Glu Asn Ser Glu Leu
275 280 285
Asn Gln Phe Asn Ile Glu Ile Arg Lys His Leu Glu Thr Tyr Phe Pro
290 295 300
Ile Lys Lys Thr Asn Arg Lys Val Gly Asp Ile Arg Asn Leu Glu Ile
305 310 315 320

Gly Glu Ile Gln Lys Ile Val Asn His Arg Leu Lys Asn Lys Ile Val
325 330 335
Gln Arg Ile Leu Gln Glu Gly Lys Leu Ala Ser Tyr Glu Ile Glu Ser
340 345 350
Thr Val Asn Ser Asn Ser Leu Gln Lys Ile Lys Ile Glu Glu Ala Phe

355 360 365

Ala Leu Lys Phe Ile Asn Ala Cys Leu Phe Ala Ser Asn Asn Leu Arg

370 375 380

Asn Met Val Tyr Pro Val Cys Lys Lys Asp Ile Leu Met Ile Gly Glu

385 390 395 400

Phe Lys Asn Ser Phe Lys Glu Ile Lys His Lys Lys Phe Ile Arg Gln

405 410 415

Trp Ser Gln Phe Phe Ser Gln Glu Ile Thr Val Asp Asp Ile Glu Leu

420 425 430

Ala Ser Trp Gly Leu Arg Gly Ala Ile Ala Pro Ile Arg Asn Glu Ile

435 440 445

Ile His Leu Lys Lys His Ser Trp Lys Lys Phe Phe Asn Asn Pro Thr

450 455 460

Phe Lys Val Lys Lys Ser Lys Ile Ile Asn Gly Lys Thr Lys Asp Val

465 470 475 480

Thr Ser Glu Phe Leu Tyr Lys Glu Thr Leu Phe Lys Asp Tyr Phe Tyr

485 490 495

Ser Glu Leu Asp Ser Val Pro Glu Leu Ile Ile Asn Lys Met Glu Ser

500 505 510

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

515 520 525

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

530 535 540

Ala Pro Ser Phe Lys Arg Val Tyr Leu Lys Gly Phe Asp Tyr Gln Asn

545 550 555 560

Gln Asp Glu Ala Gln Pro Asp Tyr Asn Leu Lys Leu Asn Ile Tyr Asn

565 570 575

Glu Lys Ala Phe Asn Ser Glu Ala Phe Gln Ala Gln Tyr Ser Leu Phe

580 585 590

Lys Met Val Tyr Tyr Gln Val Phe Leu Pro Gln Phe Thr Thr Asn Asn

595 600 605

Asp Leu Phe Lys Ser Ser Val Asp Phe Ile Leu Thr Leu Asn Lys Glu
610 615 620
Arg Lys Gly Tyr Ala Lys Ala Phe Gln Asp Ile Arg Lys Met Asn Lys
625 630 635 640

Asp Glu Lys Pro Ser Glu Tyr Met Ser Tyr Ile Gln Ser Gln Leu Met
645 650 655
Leu Tyr Gln Lys Lys Gln Glu Glu Lys Glu Lys Ile Asn His Phe Glu
660 665 670
Lys Phe Ile Asn Gln Val Phe Ile Lys Gly Phe Asn Ser Phe Ile Glu
675 680 685
Lys Asn Arg Leu Thr Tyr Ile Cys His Pro Thr Lys Asn Thr Val Pro
690 695 700

Glu Asn Asp Asn Ile Glu Ile Pro Phe His Thr Asp Met Asp Asp Ser
705 710 715 720
Asn Ile Ala Phe Trp Leu Met Cys Lys Leu Leu Asp Ala Lys Gln Leu
725 730 735
Ser Glu Leu Arg Asn Glu Met Ile Lys Phe Ser Cys Ser Leu Gln Ser
740 745 750
Thr Glu Glu Ile Ser Thr Phe Thr Lys Ala Arg Glu Val Ile Gly Leu
755 760 765

Ala Leu Leu Asn Gly Glu Lys Gly Cys Asn Asp Trp Lys Glu Leu Phe
770 775 780
Asp Asp Lys Glu Ala Trp Lys Lys Asn Met Ser Leu Tyr Val Ser Glu
785 790 795 800
Glu Leu Leu Gln Ser Leu Pro Tyr Thr Gln Glu Asp Gly Gln Thr Pro
805 810 815
Val Ile Asn Arg Ser Ile Asp Leu Val Lys Lys Tyr Gly Thr Glu Thr
820 825 830

Ile Leu Glu Lys Leu Phe Ser Ser Ser Asp Asp Tyr Lys Val Ser Ala
835 840 845
Lys Asp Ile Ala Lys Leu His Glu Tyr Asp Val Thr Glu Lys Ile Ala

850 855 860

Gln Gln Glu Ser Leu His Lys Gln Trp Ile Glu Lys Pro Gly Leu Ala
865 870 875 880

Arg Asp Ser Ala Trp Thr Lys Lys Tyr Gln Asn Val Ile Asn Asp Ile
885 890 895

Ser Asn Tyr Gln Trp Ala Lys Thr Lys Val Glu Leu Thr Gln Val Arg
900 905 910

His Leu His Gln Leu Thr Ile Asp Leu Leu Ser Arg Leu Ala Gly Tyr
915 920 925

Met Ser Ile Ala Asp Arg Asp Phe Gln Phe Ser Ser Asn Tyr Ile Leu
930 935 940

Glu Arg Glu Asn Ser Glu Tyr Arg Val Thr Ser Trp Ile Leu Leu Ser
945 950 955 960

Glu Asn Lys Asn Lys Asn Lys Tyr Asn Asp Tyr Glu Leu Tyr Asn Leu
965 970 975

Lys Asn Ala Ser Ile Lys Val Ser Ser Lys Asn Asp Pro Gln Leu Lys
980 985 990

Val Asp Leu Lys Gln Leu Arg Leu Thr Leu Glu Tyr Leu Glu Leu Phe
995 1000 1005

Asp Asn Arg Leu Lys Glu Lys Arg Asn Asn Ile Ser His Phe Asn
1010 1015 1020

Tyr Leu Asn Gly Gln Leu Gly Asn Ser Ile Leu Glu Leu Phe Asp
1025 1030 1035

Asp Ala Arg Asp Val Leu Ser Tyr Asp Arg Lys Leu Lys Asn Ala
1040 1045 1050

Val Ser Lys Ser Leu Lys Glu Ile Leu Ser Ser His Gly Met Glu
1055 1060 1065

Val Thr Phe Lys Pro Leu Tyr Gln Thr Asn His His Leu Lys Ile
1070 1075 1080

Asp Lys Leu Gln Pro Lys Lys Ile His His Leu Gly Glu Lys Ser
1085 1090 1095

Thr Val Ser Ser Asn Gln Val Ser Asn Glu Tyr Cys Gln Leu Val
1100 1105 1110

Arg Thr Leu Leu Thr Met Lys
1115 1120

<210> 49

<211> 1159

<212> PRT

<213> Leptotrichia buccalis

<400> 49

Met Lys Val Thr Lys Val Gly Gly Ile Ser His Lys Lys Tyr Thr Ser
1 5 10 15

Glu Gly Arg Leu Val Lys Ser Glu Ser Glu Glu Asn Arg Thr Asp Glu
20 25 30

Arg Leu Ser Ala Leu Leu Asn Met Arg Leu Asp Met Tyr Ile Lys Asn
35 40 45

Pro Ser Ser Thr Glu Thr Lys Glu Asn Gln Lys Arg Ile Gly Lys Leu
50 55 60

Lys Lys Phe Phe Ser Asn Lys Met Val Tyr Leu Lys Asp Asn Thr Leu
65 70 75 80

Ser Leu Lys Asn Gly Lys Lys Glu Asn Ile Asp Arg Glu Tyr Ser Glu
85 90 95

Thr Asp Ile Leu Glu Ser Asp Val Arg Asp Lys Lys Asn Phe Ala Val
100 105 110

Leu Lys Lys Ile Tyr Leu Asn Glu Asn Val Asn Ser Glu Glu Leu Glu
115 120 125

Val Phe Arg Asn Asp Ile Lys Lys Lys Leu Asn Lys Ile Asn Ser Leu
130 135 140

Lys Tyr Ser Phe Glu Lys Asn Lys Ala Asn Tyr Gln Lys Ile Asn Glu
145 150 155 160

Asn Asn Ile Glu Lys Val Glu Gly Lys Ser Lys Arg Asn Ile Ile Tyr
165 170 175

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

180 185 190

Lys Glu Ala Phe Asp Lys Leu Tyr Lys Glu Glu Asp Ile Ala Lys Leu

195 200 205

Val Leu Glu Ile Glu Asn Leu Thr Lys Leu Glu Lys Tyr Lys Ile Arg

210 215 220

Glu Phe Tyr His Glu Ile Ile Gly Arg Lys Asn Asp Lys Glu Asn Phe

225 230 235 240

Ala Lys Ile Ile Tyr Glu Glu Ile Gln Asn Val Asn Asn Met Lys Glu

245 250 255

Leu Ile Glu Lys Val Pro Asp Met Ser Glu Leu Lys Lys Ser Gln Val

260 265 270

Phe Tyr Lys Tyr Tyr Leu Asp Lys Glu Glu Leu Asn Asp Lys Asn Ile

275 280 285

Lys Tyr Ala Phe Cys His Phe Val Glu Ile Glu Met Ser Gln Leu Leu

290 295 300

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

305 310 315 320

Lys Arg Ile Phe Glu Tyr Gln Asn Leu Lys Lys Leu Ile Glu Asn Lys

325 330 335

Leu Leu Asn Lys Leu Asp Thr Tyr Val Arg Asn Cys Gly Lys Tyr Asn

340 345 350

Tyr Tyr Leu Gln Asp Gly Glu Ile Ala Thr Ser Asp Phe Ile Ala Arg

355 360 365

Asn Arg Gln Asn Glu Ala Phe Leu Arg Asn Ile Ile Gly Val Ser Ser

370 375 380

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

385 390 395 400

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

405 410 415

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

420 425 430

Lys Lys Asn Glu Val Lys Glu Asn Leu Lys Met Phe Tyr Ser Tyr Asp
435 440 445
Phe Asn Met Asp Asn Lys Asn Glu Ile Glu Asp Phe Phe Ala Asn Ile
450 455 460

Asp Glu Ala Ile Ser Ser Ile Arg His Gly Ile Val His Phe Asn Leu
465 470 475 480
Glu Leu Glu Gly Lys Asp Ile Phe Ala Phe Lys Asn Ile Ala Pro Ser
485 490 495
Glu Ile Ser Lys Lys Met Phe Gln Asn Glu Ile Asn Glu Lys Lys Leu
500 505 510
Lys Leu Lys Ile Phe Arg Gln Leu Asn Ser Ala Asn Val Phe Arg Tyr
515 520 525

Leu Glu Lys Tyr Lys Ile Leu Asn Tyr Leu Lys Arg Thr Arg Phe Glu
530 535 540
Phe Val Asn Lys Asn Ile Pro Phe Val Pro Ser Phe Thr Lys Leu Tyr
545 550 555 560
Ser Arg Ile Asp Asp Leu Lys Asn Ser Leu Gly Ile Tyr Trp Lys Thr
565 570 575
Pro Lys Thr Asn Asp Asp Asn Lys Thr Lys Glu Ile Ile Asp Ala Gln
580 585 590

Ile Tyr Leu Leu Lys Asn Ile Tyr Tyr Gly Glu Phe Leu Asn Tyr Phe
595 600 605
Met Ser Asn Asn Gly Asn Phe Phe Glu Ile Ser Lys Glu Ile Ile Glu
610 615 620
Leu Asn Lys Asn Asp Lys Arg Asn Leu Lys Thr Gly Phe Tyr Lys Leu
625 630 635 640
Gln Lys Phe Glu Asp Ile Gln Glu Lys Ile Pro Lys Glu Tyr Leu Ala
645 650 655

Asn Ile Gln Ser Leu Tyr Met Ile Asn Ala Gly Asn Gln Asp Glu Glu
660 665 670
Glu Lys Asp Thr Tyr Ile Asp Phe Ile Gln Lys Ile Phe Leu Lys Gly

675 680 685

Phe Met Thr Tyr Leu Ala Asn Asn Gly Arg Leu Ser Leu Ile Tyr Ile

690 695 700

Gly Ser Asp Glu Glu Thr Asn Thr Ser Leu Ala Glu Lys Lys Gln Glu

705 710 715 720

Phe Asp Lys Phe Leu Lys Lys Tyr Glu Gln Asn Asn Asn Ile Lys Ile

725 730 735

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

740 745 750

Leu Lys Tyr Thr Glu Arg Leu Asn Met Phe Tyr Leu Ile Leu Lys Leu

755 760 765

Leu Asn His Lys Glu Leu Thr Asn Leu Lys Gly Ser Leu Glu Lys Tyr

770 775 780

Gln Ser Ala Asn Lys Glu Glu Ala Phe Ser Asp Gln Leu Glu Leu Ile

785 790 795 800

Asn Leu Leu Asn Leu Asp Asn Asn Arg Val Thr Glu Asp Phe Glu Leu

805 810 815

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

820 825 830

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

835 840 845

Asp Gly Glu Asn Ile Ile Lys His Arg Ala Phe Tyr Asn Ile Lys Lys

850 855 860

Tyr Gly Met Leu Asn Leu Leu Glu Lys Ile Ala Asp Lys Ala Gly Tyr

865 870 875 880

Lys Ile Ser Ile Glu Glu Leu Lys Lys Tyr Ser Asn Lys Lys Asn Glu

885 890 895

Ile Glu Lys Asn His Lys Met Gln Glu Asn Leu His Arg Lys Tyr Ala

900 905 910

Arg Pro Arg Lys Asp Glu Lys Phe Thr Asp Glu Asp Tyr Glu Ser Tyr

915 920 925

Lys Gln Ala Ile Glu Asn Ile Glu Glu Tyr Thr His Leu Lys Asn Lys
930 935 940
Val Glu Phe Asn Glu Leu Asn Leu Leu Gln Gly Leu Leu Leu Arg Ile
945 950 955 960
Leu His Arg Leu Val Gly Tyr Thr Ser Ile Trp Glu Arg Asp Leu Arg
965 970 975

Phe Arg Leu Lys Gly Glu Phe Pro Glu Asn Gln Tyr Ile Glu Glu Ile
980 985 990
Phe Asn Phe Glu Asn Lys Lys Asn Val Lys Tyr Lys Gly Gly Gln Ile
995 1000 1005
Val Glu Lys Tyr Ile Lys Phe Tyr Lys Glu Leu His Gln Asn Asp
1010 1015 1020
Glu Val Lys Ile Asn Lys Tyr Ser Ser Ala Asn Ile Lys Val Leu
1025 1030 1035

Lys Gln Glu Lys Lys Asp Leu Tyr Ile Arg Asn Tyr Ile Ala His
1040 1045 1050
Phe Asn Tyr Ile Pro His Ala Glu Ile Ser Leu Leu Glu Val Leu
1055 1060 1065
Glu Asn Leu Arg Lys Leu Leu Ser Tyr Asp Arg Lys Leu Lys Asn
1070 1075 1080
Ala Val Met Lys Ser Val Val Asp Ile Leu Lys Glu Tyr Gly Phe
1085 1090 1095

Val Ala Thr Phe Lys Ile Gly Ala Asp Lys Lys Ile Gly Ile Gln
1100 1105 1110
Thr Leu Glu Ser Glu Lys Ile Val His Leu Lys Asn Leu Lys Lys
1115 1120 1125
Lys Lys Leu Met Thr Asp Arg Asn Ser Glu Glu Leu Cys Lys Leu
1130 1135 1140
Val Lys Ile Met Phe Glu Tyr Lys Met Glu Glu Lys Lys Ser Glu
1145 1150 1155

Asn

<210> 50

<211> 1285

<212> PRT

<213> Herbinix hemicellulosilytica

<400> 50

Met Lys Leu Thr Arg Arg Arg Ile Ser Gly Asn Ser Val Asp Gln Lys
1 5 10 15

Ile Thr Ala Ala Phe Tyr Arg Asp Met Ser Gln Gly Leu Leu Tyr Tyr
20 25 30

Asp Ser Glu Asp Asn Asp Cys Thr Asp Lys Val Ile Glu Ser Met Asp
35 40 45

Phe Glu Arg Ser Trp Arg Gly Arg Ile Leu Lys Asn Gly Glu Asp Asp

50 55 60

Lys Asn Pro Phe Tyr Met Phe Val Lys Gly Leu Val Gly Ser Asn Asp
65 70 75 80

Lys Ile Val Cys Glu Pro Ile Asp Val Asp Ser Asp Pro Asp Asn Leu
85 90 95

Asp Ile Leu Ile Asn Lys Asn Leu Thr Gly Phe Gly Arg Asn Leu Lys
100 105 110

Ala Pro Asp Ser Asn Asp Thr Leu Glu Asn Leu Ile Arg Lys Ile Gln

115 120 125

Ala Gly Ile Pro Glu Glu Glu Val Leu Pro Glu Leu Lys Lys Ile Lys
130 135 140

Glu Met Ile Gln Lys Asp Ile Val Asn Arg Lys Glu Gln Leu Leu Lys
145 150 155 160

Ser Ile Lys Asn Asn Arg Ile Pro Phe Ser Leu Glu Gly Ser Lys Leu
165 170 175

Val Pro Ser Thr Lys Lys Met Lys Trp Leu Phe Lys Leu Ile Asp Val

180 185 190

Pro Asn Lys Thr Phe Asn Glu Lys Met Leu Glu Lys Tyr Trp Glu Ile
195 200 205

Tyr Asp Tyr Asp Lys Leu Lys Ala Asn Ile Thr Asn Arg Leu Asp Lys

210 215 220

Thr Asp Lys Lys Ala Arg Ser Ile Ser Arg Ala Val Ser Glu Glu Leu

225 230 235 240

Arg Glu Tyr His Lys Asn Leu Arg Thr Asn Tyr Asn Arg Phe Val Ser

245 250 255

Gly Asp Arg Pro Ala Ala Gly Leu Asp Asn Gly Gly Ser Ala Lys Tyr

260 265 270

Asn Pro Asp Lys Glu Glu Phe Leu Leu Phe Leu Lys Glu Val Glu Gln

275 280 285

Tyr Phe Lys Lys Tyr Phe Pro Val Lys Ser Lys His Ser Asn Lys Ser

290 295 300

Lys Asp Lys Ser Leu Val Asp Lys Tyr Lys Asn Tyr Cys Ser Tyr Lys

305 310 315 320

Val Val Lys Lys Glu Val Asn Arg Ser Ile Ile Asn Gln Leu Val Ala

325 330 335

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

340 345 350

Thr Trp Gln Glu Asp Phe Leu Asn Ser Tyr Gly Leu Ser Tyr Ile Gln

355 360 365

Val Glu Glu Ala Phe Lys Lys Ser Val Met Thr Ser Leu Ser Trp Gly

370 375 380

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

385 390 395 400

Phe Asp Asp Ile Thr Thr Lys Lys Ala Lys Glu Ala Ile Glu Ser Asn

405 410 415

Tyr Phe Asn Lys Leu Arg Thr Cys Ser Arg Met Gln Asp His Phe Lys

420 425 430

Glu Lys Leu Ala Phe Phe Tyr Pro Val Tyr Val Lys Asp Lys Lys Asp

435 440 445

Arg Pro Asp Asp Asp Ile Glu Asn Leu Ile Val Leu Val Lys Asn Ala

450 455 460

Ile Glu Ser Val Ser Tyr Leu Arg Asn Arg Thr Phe His Phe Lys Glu

465 470 475 480

Ser Ser Leu Leu Glu Leu Leu Lys Glu Leu Asp Asp Lys Asn Ser Gly

485 490 495

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

500 505 510

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

515 520 525

Ile Ala Glu Tyr Tyr Lys Ala Asp Met Ile Ser Asp Cys Phe Lys Thr

530 535 540

Cys Gly Leu Glu Phe Ala Leu Tyr Ser Pro Lys Asn Ser Leu Met Pro

545 550 555 560

Ala Phe Lys Asn Val Tyr Lys Arg Gly Ala Asn Leu Asn Lys Ala Tyr

565 570 575

Ile Arg Asp Lys Gly Pro Lys Glu Thr Gly Asp Gln Gly Gln Asn Ser

580 585 590

Tyr Lys Ala Leu Glu Glu Tyr Arg Glu Leu Thr Trp Tyr Ile Glu Val

595 600 605

Lys Asn Asn Asp Gln Ser Tyr Asn Ala Tyr Lys Asn Leu Leu Gln Leu

610 615 620

Ile Tyr Tyr His Ala Phe Leu Pro Glu Val Arg Glu Asn Glu Ala Leu

625 630 635 640

Ile Thr Asp Phe Ile Asn Arg Thr Lys Glu Trp Asn Arg Lys Glu Thr

645 650 655

Glu Glu Arg Leu Asn Thr Lys Asn Asn Lys Lys His Lys Asn Phe Asp

660 665 670

Glu Asn Asp Asp Ile Thr Val Asn Thr Tyr Arg Tyr Glu Ser Ile Pro

675 680 685

Asp Tyr Gln Gly Glu Ser Leu Asp Asp Tyr Leu Lys Val Leu Gln Arg

690 695 700

Lys Gln Met Ala Arg Ala Lys Glu Val Asn Glu Lys Glu Glu Gly Asn
705 710 715 720
Asn Asn Tyr Ile Gln Phe Ile Arg Asp Val Val Val Trp Ala Phe Gly
725 730 735
Ala Tyr Leu Glu Asn Lys Leu Lys Asn Tyr Lys Asn Glu Leu Gln Pro
740 745 750
Pro Leu Ser Lys Glu Asn Ile Gly Leu Asn Asp Thr Leu Lys Glu Leu

755 760 765
Phe Pro Glu Glu Lys Val Lys Ser Pro Phe Asn Ile Lys Cys Arg Phe
770 775 780
Ser Ile Ser Thr Phe Ile Asp Asn Lys Gly Lys Ser Thr Asp Asn Thr
785 790 795 800
Ser Ala Glu Ala Val Lys Thr Asp Gly Lys Glu Asp Glu Lys Asp Lys
805 810 815
Lys Asn Ile Lys Arg Lys Asp Leu Leu Cys Phe Tyr Leu Phe Leu Arg

820 825 830
Leu Leu Asp Glu Asn Glu Ile Cys Lys Leu Gln His Gln Phe Ile Lys
835 840 845
Tyr Arg Cys Ser Leu Lys Glu Arg Arg Phe Pro Gly Asn Arg Thr Lys
850 855 860
Leu Glu Lys Glu Thr Glu Leu Leu Ala Glu Leu Glu Glu Leu Met Glu
865 870 875 880
Leu Val Arg Phe Thr Met Pro Ser Ile Pro Glu Ile Ser Ala Lys Ala

885 890 895
Glu Ser Gly Tyr Asp Thr Met Ile Lys Lys Tyr Phe Lys Asp Phe Ile
900 905 910
Glu Lys Lys Val Phe Lys Asn Pro Lys Thr Ser Asn Leu Tyr Tyr His
915 920 925
Ser Asp Ser Lys Thr Pro Val Thr Arg Lys Tyr Met Ala Leu Leu Met
930 935 940
Arg Ser Ala Pro Leu His Leu Tyr Lys Asp Ile Phe Lys Gly Tyr Tyr

945 950 955 960

Leu Ile Thr Lys Lys Glu Cys Leu Glu Tyr Ile Lys Leu Ser Asn Ile

965 970 975

Ile Lys Asp Tyr Gln Asn Ser Leu Asn Glu Leu His Glu Gln Leu Glu

980 985 990

Arg Ile Lys Leu Lys Ser Glu Lys Gln Asn Gly Lys Asp Ser Leu Tyr

995 1000 1005

Leu Asp Lys Lys Asp Phe Tyr Lys Val Lys Glu Tyr Val Glu Asn

1010 1015 1020

Leu Glu Gln Val Ala Arg Tyr Lys His Leu Gln His Lys Ile Asn

1025 1030 1035

Phe Glu Ser Leu Tyr Arg Ile Phe Arg Ile His Val Asp Ile Ala

1040 1045 1050

Ala Arg Met Val Gly Tyr Thr Gln Asp Trp Glu Arg Asp Met His

1055 1060 1065

Phe Leu Phe Lys Ala Leu Val Tyr Asn Gly Val Leu Glu Glu Arg

1070 1075 1080

Arg Phe Glu Ala Ile Phe Asn Asn Asn Asp Asp Asn Asn Asp Gly

1085 1090 1095

Arg Ile Val Lys Lys Ile Gln Asn Asn Leu Asn Asn Lys Asn Arg

1100 1105 1110

Glu Leu Val Ser Met Leu Cys Trp Asn Lys Lys Leu Asn Lys Asn

1115 1120 1125

Glu Phe Gly Ala Ile Ile Trp Lys Arg Asn Pro Ile Ala His Leu

1130 1135 1140

Asn His Phe Thr Gln Thr Glu Gln Asn Ser Lys Ser Ser Leu Glu

1145 1150 1155

Ser Leu Ile Asn Ser Leu Arg Ile Leu Leu Ala Tyr Asp Arg Lys

1160 1165 1170

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

1175 1180 1185

Asp Tyr His Ile Arg Ile Lys Trp Glu Gly Arg Val Asp Glu Gly

1190 1195 1200

Gln Ile Tyr Phe Asn Ile Lys Glu Lys Glu Asp Ile Glu Asn Glu

1205 1210 1215

Pro Ile Ile His Leu Lys His Leu His Lys Lys Asp Cys Tyr Ile

1220 1225 1230

Tyr Lys Asn Ser Tyr Met Phe Asp Lys Gln Lys Glu Trp Ile Cys

1235 1240 1245

Asn Gly Ile Lys Glu Glu Val Tyr Asp Lys Ser Ile Leu Lys Cys

1250 1255 1260

Ile Gly Asn Leu Phe Lys Phe Asp Tyr Glu Asp Lys Asn Lys Ser

1265 1270 1275

Ser Ala Asn Pro Lys His Thr

1280 1285

<210> 51

<211> 1344

<212> PRT

<213> Eubacterium rectale

<400> 51

Met Leu Arg Arg Asp Lys Glu Val Lys Lys Leu Tyr Asn Val Phe Asn

1 5 10 15

Gln Ile Gln Val Gly Thr Lys Pro Lys Lys Trp Asn Asn Asp Glu Lys

20 25 30

Leu Ser Pro Glu Glu Asn Glu Arg Arg Ala Gln Gln Lys Asn Ile Lys

35 40 45

Met Lys Asn Tyr Lys Trp Arg Glu Ala Cys Ser Lys Tyr Val Glu Ser

50 55 60

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

65 70 75 80

Lys Asn Lys Leu Arg Tyr Met Arg Lys Asn Glu Asp Ile Leu Lys Lys

85 90 95

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

145 150 155 160
Lys Leu Leu Glu Leu Ile Arg Lys Asp Phe Ser Lys Leu Asp Pro Asn
165 170 175
Val Lys Gly Ser Gln Gly Ala Asn Ile Val Arg Ser Val Arg Asn Gln
180 185 190
Asn Met Ile Val Gln Pro Gln Gly Asp Arg Phe Leu Phe Pro Gln Val
195 200 205
Tyr Ala Lys Glu Asn Glu Thr Val Thr Asn Lys Asn Val Glu Lys Glu

210 215 220
Gly Leu Asn Glu Phe Leu Leu Asn Tyr Ala Asn Leu Asp Asp Glu Lys
225 230 235 240
Arg Ala Glu Ser Leu Arg Lys Leu Arg Arg Ile Leu Asp Val Tyr Phe
245 250 255
Ser Ala Pro Asn His Tyr Glu Lys Asp Met Asp Ile Thr Leu Ser Asp
260 265 270
Asn Ile Glu Lys Glu Lys Phe Asn Val Trp Glu Lys His Glu Cys Gly

275 280 285
Lys Lys Glu Thr Gly Leu Phe Val Asp Ile Pro Asp Val Leu Met Glu
290 295 300
Ala Glu Ala Glu Asn Ile Lys Leu Asp Ala Val Val Glu Lys Arg Glu
305 310 315 320
Arg Lys Val Leu Asn Asp Arg Val Arg Lys Gln Asn Ile Ile Cys Tyr
325 330 335
Arg Tyr Thr Arg Ala Val Val Glu Lys Tyr Asn Ser Asn Glu Pro Leu

340 345 350

Phe Phe Glu Asn Asn Ala Ile Asn Gln Tyr Trp Ile His His Ile Glu

355 360 365

Asn Ala Val Glu Arg Ile Leu Lys Asn Cys Lys Ala Gly Lys Leu Phe

370 375 380

Lys Leu Arg Lys Gly Tyr Leu Ala Glu Lys Val Trp Lys Asp Ala Ile

385 390 395 400

Asn Leu Ile Ser Ile Lys Tyr Ile Ala Leu Gly Lys Ala Val Tyr Asn

405 410 415

Phe Ala Leu Asp Asp Ile Trp Lys Asp Lys Lys Asn Lys Glu Leu Gly

420 425 430

Ile Val Asp Glu Arg Ile Arg Asn Gly Ile Thr Ser Phe Asp Tyr Glu

435 440 445

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

450 455 460

Ala Phe Ser Val Asn Asn Leu Ala Arg Ala Val Cys Asp Met Ser Asn

465 470 475 480

Leu Gly Asn Lys Glu Ser Asp Phe Leu Leu Trp Lys Arg Asn Asp Ile

485 490 495

Ala Asp Lys Leu Lys Asn Lys Asp Asp Met Ala Ser Val Ser Ala Val

500 505 510

Leu Gln Phe Phe Gly Gly Lys Ser Ser Trp Asp Ile Asn Ile Phe Lys

515 520 525

Asp Ala Tyr Lys Gly Lys Lys Lys Tyr Asn Tyr Glu Val Arg Phe Ile

530 535 540

Asp Asp Leu Arg Lys Ala Ile Tyr Cys Ala Arg Asn Glu Asn Phe His

545 550 555 560

Phe Lys Thr Ala Leu Val Asn Asp Glu Lys Trp Asn Thr Glu Leu Phe

565 570 575

Gly Lys Ile Phe Glu Arg Glu Thr Glu Phe Cys Leu Asn Val Glu Lys

580 585 590

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

595 600 605

Leu Arg Asn Met Leu Asp His Leu Tyr Ser Arg Ser Val Ser Arg Ala

610 615 620

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

625 630 635 640

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

645 650 655

Asp Thr Leu Gly Lys Trp Tyr Ser Ala Cys Tyr Tyr Leu Leu Lys Glu

660 665 670

Ile Tyr Tyr Asn Ser Phe Leu Gln Ser Asp Arg Ala Leu Gln Leu Phe

675 680 685

Glu Lys Ser Val Lys Thr Leu Ser Trp Asp Asp Lys Lys Gln Gln Arg

690 695 700

Ala Val Asp Asn Phe Lys Asp His Phe Ser Asp Ile Lys Ser Ala Cys

705 710 715 720

Thr Ser Leu Ala Gln Val Cys Gln Ile Tyr Met Thr Glu Tyr Asn Gln

725 730 735

Gln Asn Asn Gln Ile Lys Lys Val Arg Ser Ser Asn Asp Ser Ile Phe

740 745 750

Asp Gln Pro Val Tyr Gln His Tyr Lys Val Leu Leu Lys Lys Ala Ile

755 760 765

Ala Asn Ala Phe Ala Asp Tyr Leu Lys Asn Asn Lys Asp Leu Phe Gly

770 775 780

Phe Ile Gly Lys Pro Phe Lys Ala Asn Glu Ile Arg Glu Ile Asp Lys

785 790 795 800

Glu Gln Phe Leu Pro Asp Trp Thr Ser Arg Lys Tyr Glu Ala Leu Cys

805 810 815

Ile Glu Val Ser Gly Ser Gln Glu Leu Gln Lys Trp Tyr Ile Val Gly

820 825 830

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

835 840 845

Ser Tyr Ile Gln Tyr Val Thr Asp Ile Lys Arg Arg Ala Ala Ser Ile

850 855 860

Gly Asn Glu Leu His Val Ser Val His Asp Val Glu Lys Val Glu Lys

865 870 875 880

Trp Val Gln Val Ile Glu Val Cys Ser Leu Leu Ala Ser Arg Thr Ser

885 890 895

Asn Gln Phe Glu Asp Tyr Phe Asn Asp Lys Asp Asp Tyr Ala Arg Tyr

900 905 910

Leu Lys Ser Tyr Val Asp Phe Ser Asn Val Asp Met Pro Ser Glu Tyr

915 920 925

Ser Ala Leu Val Asp Phe Ser Asn Glu Glu Gln Ser Asp Leu Tyr Val

930 935 940

Asp Pro Lys Asn Pro Lys Val Asn Arg Asn Ile Val His Ser Lys Leu

945 950 955 960

Phe Ala Ala Asp His Ile Leu Arg Asp Ile Val Glu Pro Val Ser Lys

965 970 975

Asp Asn Ile Glu Glu Phe Tyr Ser Gln Lys Ala Glu Ile Ala Tyr Cys

980 985 990

Lys Ile Lys Gly Lys Glu Ile Thr Ala Glu Glu Gln Lys Ala Val Leu

995 1000 1005

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

1010 1015 1020

Glu Tyr Gly Glu Ile Ile Asn Glu Leu Leu Gly Gln Leu Ile Asn

1025 1030 1035

Trp Ser Phe Met Arg Glu Arg Asp Leu Leu Tyr Phe Gln Leu Gly

1040 1045 1050

Phe His Tyr Asp Cys Leu Arg Asn Asp Ser Lys Lys Pro Glu Gly

1055 1060 1065

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

1070 1075 1080

Leu Tyr Gln Ile Ile Gly Met Tyr Val Asn Gly Val Thr Val Tyr
1085 1090 1095
Ala Pro Glu Lys Asp Gly Asp Lys Leu Lys Glu Gln Cys Val Lys

1100 1105 1110
Gly Gly Val Gly Val Lys Val Ser Ala Phe His Arg Tyr Ser Lys
1115 1120 1125
Tyr Leu Gly Leu Asn Glu Lys Thr Leu Tyr Asn Ala Gly Leu Glu
1130 1135 1140
Ile Phe Glu Val Val Ala Glu His Glu Asp Ile Ile Asn Leu Arg
1145 1150 1155
Asn Gly Ile Asp His Phe Lys Tyr Tyr Leu Gly Asp Tyr Arg Ser

1160 1165 1170
Met Leu Ser Ile Tyr Ser Glu Val Phe Asp Arg Phe Phe Thr Tyr
1175 1180 1185
Asp Ile Lys Tyr Gln Lys Asn Val Leu Asn Leu Leu Gln Asn Ile
1190 1195 1200
Leu Leu Arg His Asn Val Ile Val Glu Pro Ile Leu Glu Ser Gly
1205 1210 1215
Phe Lys Thr Ile Gly Glu Gln Thr Lys Pro Gly Ala Lys Leu Ser

1220 1225 1230
Ile Arg Ser Ile Lys Ser Asp Thr Phe Gln Tyr Lys Val Lys Gly
1235 1240 1245
Gly Thr Leu Ile Thr Asp Ala Lys Asp Glu Arg Tyr Leu Glu Thr
1250 1255 1260
Ile Arg Lys Ile Leu Tyr Tyr Ala Glu Asn Glu Glu Asp Asn Leu
1265 1270 1275
Lys Lys Ser Val Val Val Thr Asn Ala Asp Lys Tyr Glu Lys Asn

1280 1285 1290
Lys Glu Ser Asp Asp Gln Asn Lys Gln Lys Glu Lys Lys Asn Lys
1295 1300 1305
Asp Asn Lys Gly Lys Lys Asn Glu Glu Thr Lys Ser Asp Ala Glu

1310 1315 1320

Lys Asn Asn Asn Glu Arg Leu Ser Tyr Asn Pro Phe Ala Asn Leu

1325 1330 1335

Asn Phe Lys Leu Ser Asn

1340

<210> 52

<211> 1373

<212> PRT

<213> Eubacteriaceae bacterium

<400> 52

Met Lys Ile Ser Lys Glu Ser His Lys Arg Thr Ala Val Ala Val Met

1 5 10 15

Glu Asp Arg Val Gly Gly Val Val Tyr Val Pro Gly Gly Ser Gly Ile

20 25 30

Asp Leu Ser Asn Asn Leu Lys Lys Arg Ser Met Asp Thr Lys Ser Leu

35 40 45

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

50 55 60

Glu Trp Lys Asp Tyr Leu Ser Glu Ala Glu Asn Lys Lys Arg Glu Ala

65 70 75 80

Gln Lys Met Ile Gln Lys Ala Asn Tyr Glu Leu Arg Arg Glu Cys Glu

85 90 95

Asp Tyr Ala Lys Lys Ala Asn Leu Ala Val Ser Arg Ile Ile Phe Ser

100 105 110

Lys Lys Pro Lys Lys Ile Phe Ser Asp Asp Asp Ile Ile Ser His Met

115 120 125

Lys Lys Gln Arg Leu Ser Lys Phe Lys Gly Arg Met Glu Asp Phe Val

130 135 140

Leu Ile Ala Leu Arg Lys Ser Leu Val Val Ser Thr Tyr Asn Gln Glu

145 150 155 160

Val Phe Asp Ser Arg Lys Ala Ala Thr Val Phe Leu Lys Asn Ile Gly

165 170 175

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

180 185 190

Ala Leu Ile Arg Glu Asp Tyr Asp Lys Trp Asn Pro Asp Lys Asp Ser

195 200 205

Ser Asp Lys Lys Glu Ser Ser Gly Thr Lys Val Ile Arg Ser Ile Glu

210 215 220

His Gln Asn Met Val Ile Gln Pro Glu Lys Asn Lys Leu Ser Leu Ser

225 230 235 240

Lys Ile Ser Asn Val Gly Lys Lys Thr Lys Thr Lys Gln Lys Glu Lys

245 250 255

Ala Gly Leu Asp Ala Phe Leu Lys Glu Tyr Ala Gln Ile Asp Glu Asn

260 265 270

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

275 280 285

Phe Ala Ala Pro Ser Ser Tyr Ile Lys Gly Ala Ala Val Ser Leu Pro

290 295 300

Glu Asn Ile Asn Phe Ser Ser Glu Leu Asn Val Trp Glu Arg His Glu

305 310 315 320

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

325 330 335

Leu Asn Ala Glu Gln Asn Asn Asn Lys Ile Asn Lys Val Glu Gln Glu

340 345 350

His Ser Leu Glu Gln Leu Arg Thr Asp Ile Arg Arg Arg Asn Ile Thr

355 360 365

Cys Tyr His Phe Ala Asn Ala Leu Ala Ala Asp Glu Arg Tyr His Thr

370 375 380

Leu Phe Phe Glu Asn Met Ala Met Asn Gln Phe Trp Ile His His Met

385 390 395 400

Glu Asn Ala Val Glu Arg Ile Leu Lys Lys Cys Asn Val Gly Thr Leu

405 410 415

Phe Lys Leu Arg Ile Gly Tyr Leu Ser Glu Lys Val Trp Lys Asp Met

420 425 430

Leu Asn Leu Leu Ser Ile Lys Tyr Ile Ala Leu Gly Lys Ala Val Tyr

435 440 445

His Phe Ala Leu Asp Asp Ile Trp Lys Ala Asp Ile Trp Lys Asp Ala

450 455 460

Ser Asp Lys Asn Ser Gly Lys Ile Asn Asp Leu Thr Leu Lys Gly Ile

465 470 475 480

Ser Ser Phe Asp Tyr Glu Met Val Lys Ala Gln Glu Asp Leu Gln Arg

485 490 495

Glu Met Ala Val Gly Val Ala Phe Ser Thr Asn Asn Leu Ala Arg Val

500 505 510

Thr Cys Lys Met Asp Asp Leu Ser Asp Ala Glu Ser Asp Phe Leu Leu

515 520 525

Trp Asn Lys Glu Ala Ile Arg Arg His Val Lys Tyr Thr Glu Lys Gly

530 535 540

Glu Ile Leu Ser Ala Ile Leu Gln Phe Phe Gly Gly Arg Ser Leu Trp

545 550 555 560

Asp Glu Ser Leu Phe Glu Lys Ala Tyr Ser Asp Ser Asn Tyr Glu Leu

565 570 575

Lys Phe Leu Asp Asp Leu Lys Arg Ala Ile Tyr Ala Ala Arg Asn Glu

580 585 590

Thr Phe His Phe Lys Thr Ala Ala Ile Asp Gly Gly Ser Trp Asn Thr

595 600 605

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

610 615 620

Val Glu Lys Asn Lys Phe Tyr Ser Asn Asn Leu Val Leu Phe Tyr Lys

625 630 635 640

Gln Glu Asp Leu Arg Val Phe Leu Asp Lys Leu Tyr Gly Lys Glu Cys

645 650 655

Ser Arg Ala Ala Gln Ile Pro Ser Tyr Asn Thr Ile Leu Pro Arg Lys

660 665 670

Ser Phe Ser Asp Phe Met Lys Gln Leu Leu Gly Leu Lys Glu Pro Val
675 680 685
Tyr Gly Ser Ala Ile Leu Asp Gln Trp Tyr Ser Ala Cys Tyr Tyr Leu

690 695 700
Phe Lys Glu Val Tyr Tyr Asn Leu Phe Leu Gln Asp Ser Ser Ala Lys
705 710 715 720
Ala Leu Phe Glu Lys Ala Val Lys Ala Leu Lys Gly Ala Asp Lys Lys
725 730 735
Gln Glu Lys Ala Val Glu Ser Phe Arg Lys Arg Tyr Trp Glu Ile Ser
740 745 750
Lys Asn Ala Ser Leu Ala Glu Ile Cys Gln Ser Tyr Ile Thr Glu Tyr

755 760 765
Asn Gln Gln Asn Asn Lys Glu Arg Lys Val Arg Ser Ala Asn Asp Gly
770 775 780
Met Phe Asn Glu Pro Ile Tyr Gln His Tyr Lys Met Leu Leu Lys Glu
785 790 795 800
Ala Leu Lys Met Ala Phe Ala Ser Tyr Ile Lys Asn Asp Lys Glu Leu
805 810 815
Lys Phe Val Tyr Lys Pro Thr Glu Lys Leu Phe Glu Val Ser Gln Asp

820 825 830
Asn Phe Leu Pro Asn Trp Asn Ser Glu Lys Tyr Asn Thr Leu Ile Ser
835 840 845
Glu Val Lys Asn Ser Pro Asp Leu Gln Lys Trp Tyr Ile Val Gly Lys
850 855 860
Phe Met Asn Ala Arg Met Leu Asn Leu Leu Leu Gly Ser Met Arg Ser
865 870 875 880
Tyr Leu Gln Tyr Val Ser Asp Ile Gln Lys Arg Ala Ala Gly Leu Gly

885 890 895
Glu Asn Gln Leu His Leu Ser Ala Glu Asn Val Gly Gln Val Lys Lys
900 905 910
Trp Ile Gln Val Leu Glu Val Cys Leu Leu Leu Ser Val Arg Ile Ser

915 920 925

Asp Lys Phe Thr Asp Tyr Phe Lys Asp Glu Glu Glu Tyr Ala Ser Tyr

930 935 940

Leu Lys Glu Tyr Val Asp Phe Glu Asp Ser Ala Met Pro Ser Asp Tyr

945 950 955 960

Ser Ala Leu Leu Ala Phe Ser Asn Glu Gly Lys Ile Asp Leu Tyr Val

965 970 975

Asp Ala Ser Asn Pro Lys Val Asn Arg Asn Ile Ile Gln Ala Lys Leu

980 985 990

Tyr Ala Pro Asp Met Val Leu Lys Lys Val Val Lys Lys Ile Ser Gln

995 1000 1005

Asp Glu Cys Lys Glu Phe Asn Glu Lys Lys Glu Gln Ile Met Gln

1010 1015 1020

Phe Lys Asn Lys Gly Asp Glu Val Ser Trp Glu Glu Gln Gln Lys

1025 1030 1035

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

1040 1045 1050

Leu Ser Glu Tyr Gly Glu Leu Ile Asn Glu Leu Leu Gly Gln Leu

1055 1060 1065

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

1070 1075 1080

Leu Gly Phe His Tyr Ser Cys Leu Met Asn Glu Ser Lys Lys Pro

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

Val Tyr Ala Pro Glu Lys Gly Glu Leu Lys Pro Gln Cys Lys Thr

1130 1135 1140

Gly Ser Ala Gly Gln Lys Ile Arg Ala Phe Cys Gln Trp Ala Ser

1145 1150 1155

Met Val Glu Lys Lys Lys Tyr Glu Leu Tyr Asn Ala Gly Leu Glu
1160 1165 1170
Leu Phe Glu Val Val Lys Glu His Asp Asn Ile Ile Asp Leu Arg
1175 1180 1185
Asn Lys Ile Asp His Phe Lys Tyr Tyr Gln Gly Asn Asp Ser Ile

1190 1195 1200
Leu Ala Leu Tyr Gly Glu Ile Phe Asp Arg Phe Phe Thr Tyr Asp
1205 1210 1215
Met Lys Tyr Arg Asn Asn Val Leu Asn His Leu Gln Asn Ile Leu
1220 1225 1230
Leu Arg His Asn Val Ile Ile Lys Pro Ile Ile Ser Lys Asp Lys
1235 1240 1245
Lys Glu Val Gly Arg Gly Lys Met Lys Asp Arg Ala Ala Phe Leu

1250 1255 1260
Leu Glu Glu Val Ser Ser Asp Arg Phe Thr Tyr Lys Val Lys Glu
1265 1270 1275
Gly Glu Arg Lys Ile Asp Ala Lys Asn Arg Leu Tyr Leu Glu Thr
1280 1285 1290
Val Arg Asp Ile Leu Tyr Phe Pro Asn Arg Ala Val Asn Asp Lys
1295 1300 1305
Gly Glu Asp Val Ile Ile Cys Ser Lys Lys Ala Gln Asp Leu Asn

1310 1315 1320
Glu Lys Lys Ala Asp Arg Asp Lys Asn His Asp Lys Ser Lys Asp
1325 1330 1335
Thr Asn Gln Lys Lys Glu Gly Lys Asn Gln Glu Glu Lys Ser Glu
1340 1345 1350
Asn Lys Glu Pro Tyr Ser Asp Arg Met Thr Trp Lys Pro Phe Ala
1355 1360 1365
Gly Ile Lys Leu Glu
1370
<210> 53

<211> 1276

<212> PRT

<213> Blautia species

<400> 53

Met Lys Ile Ser Lys Val Asp His Val Lys Ser Gly Ile Asp Gln Lys
1 5 10 15

Leu Ser Ser Gln Arg Gly Met Leu Tyr Lys Gln Pro Gln Lys Lys Tyr
20 25 30

Glu Gly Lys Gln Leu Glu Glu His Val Arg Asn Leu Ser Arg Lys Ala
35 40 45

Lys Ala Leu Tyr Gln Val Phe Pro Val Ser Gly Asn Ser Lys Met Glu
50 55 60

Lys Glu Leu Gln Ile Ile Asn Ser Phe Ile Lys Asn Ile Leu Leu Arg
65 70 75 80

Leu Asp Ser Gly Lys Thr Ser Glu Glu Ile Val Gly Tyr Ile Asn Thr
85 90 95

Tyr Ser Val Ala Ser Gln Ile Ser Gly Asp His Ile Gln Glu Leu Val
100 105 110

Asp Gln His Leu Lys Glu Ser Leu Arg Lys Tyr Thr Cys Val Gly Asp
115 120 125

Lys Arg Ile Tyr Val Pro Asp Ile Ile Val Ala Leu Leu Lys Ser Lys
130 135 140

Phe Asn Ser Glu Thr Leu Gln Tyr Asp Asn Ser Glu Leu Lys Ile Leu
145 150 155 160

Ile Asp Phe Ile Arg Glu Asp Tyr Leu Lys Glu Lys Gln Ile Lys Gln
165 170 175

Ile Val His Ser Ile Glu Asn Asn Ser Thr Pro Leu Arg Ile Ala Glu
180 185 190

Ile Asn Gly Gln Lys Arg Leu Ile Pro Ala Asn Val Asp Asn Pro Lys
195 200 205

Lys Ser Tyr Ile Phe Glu Phe Leu Lys Glu Tyr Ala Gln Ser Asp Pro
210 215 220

Lys Gly Gln Glu Ser Leu Leu Gln His Met Arg Tyr Leu Ile Leu Leu
225 230 235 240
Tyr Leu Tyr Gly Pro Asp Lys Ile Thr Asp Asp Tyr Cys Glu Glu Ile

245 250 255
Glu Ala Trp Asn Phe Gly Ser Ile Val Met Asp Asn Glu Gln Leu Phe
260 265 270
Ser Glu Glu Ala Ser Met Leu Ile Gln Asp Arg Ile Tyr Val Asn Gln
275 280 285
Gln Ile Glu Glu Gly Arg Gln Ser Lys Asp Thr Ala Lys Val Lys Lys
290 295 300
Asn Lys Ser Lys Tyr Arg Met Leu Gly Asp Lys Ile Glu His Ser Ile

305 310 315 320
Asn Glu Ser Val Val Lys His Tyr Gln Glu Ala Cys Lys Ala Val Glu
325 330 335
Glu Lys Asp Ile Pro Trp Ile Lys Tyr Ile Ser Asp His Val Met Ser
340 345 350
Val Tyr Ser Ser Lys Asn Arg Val Asp Leu Asp Lys Leu Ser Leu Pro
355 360 365
Tyr Leu Ala Lys Asn Thr Trp Asn Thr Trp Ile Ser Phe Ile Ala Met

370 375 380
Lys Tyr Val Asp Met Gly Lys Gly Val Tyr His Phe Ala Met Ser Asp
385 390 395 400
Val Asp Lys Val Gly Lys Gln Asp Asn Leu Ile Ile Gly Gln Ile Asp
405 410 415
Pro Lys Phe Ser Asp Gly Ile Ser Ser Phe Asp Tyr Glu Arg Ile Lys
420 425 430
Ala Glu Asp Asp Leu His Arg Ser Met Ser Gly Tyr Ile Ala Phe Ala

435 440 445
Val Asn Asn Phe Ala Arg Ala Ile Cys Ser Asp Glu Phe Arg Lys Lys
450 455 460
Asn Arg Lys Glu Asp Val Leu Thr Val Gly Leu Asp Glu Ile Pro Leu

465 470 475 480

Tyr Asp Asn Val Lys Arg Lys Leu Leu Gln Tyr Phe Gly Gly Ala Ser

485 490 495

Asn Trp Asp Asp Ser Ile Ile Asp Ile Ile Asp Asp Lys Asp Leu Val

500 505 510

Ala Cys Ile Lys Glu Asn Leu Tyr Val Ala Arg Asn Val Asn Phe His

515 520 525

Phe Ala Gly Ser Glu Lys Val Gln Lys Lys Gln Asp Asp Ile Leu Glu

530 535 540

Glu Ile Val Arg Lys Glu Thr Arg Asp Ile Gly Lys His Tyr Arg Lys

545 550 555 560

Val Phe Tyr Ser Asn Asn Val Ala Val Phe Tyr Cys Asp Glu Asp Ile

565 570 575

Ile Lys Leu Met Asn His Leu Tyr Gln Arg Glu Lys Pro Tyr Gln Ala

580 585 590

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

595 600 605

Leu Ile Phe Met Leu Leu Lys Gly Lys Asn Arg Thr Lys Ile Ser Asp

610 615 620

Pro Ser Ile Met Asn Met Phe Arg Gly Thr Phe Tyr Phe Leu Leu Lys

625 630 635 640

Glu Ile Tyr Tyr Asn Asp Phe Leu Gln Ala Ser Asn Leu Lys Glu Met

645 650 655

Phe Cys Glu Gly Leu Lys Asn Asn Val Lys Asn Lys Lys Ser Glu Lys

660 665 670

Pro Tyr Gln Asn Phe Met Arg Arg Phe Glu Glu Leu Glu Asn Met Gly

675 680 685

Met Asp Phe Gly Glu Ile Cys Gln Gln Ile Met Thr Asp Tyr Glu Gln

690 695 700

Gln Asn Lys Gln Lys Lys Lys Thr Ala Thr Ala Val Met Ser Glu Lys

705 710 715 720

Asp Lys Lys Ile Arg Thr Leu Asp Asn Asp Thr Gln Lys Tyr Lys His
725 730 735
Phe Arg Thr Leu Leu Tyr Ile Gly Leu Arg Glu Ala Phe Ile Ile Tyr
740 745 750
Leu Lys Asp Glu Lys Asn Lys Glu Trp Tyr Glu Phe Leu Arg Glu Pro

755 760 765
Val Lys Arg Glu Gln Pro Glu Glu Lys Glu Phe Val Asn Lys Trp Lys
770 775 780
Leu Asn Gln Tyr Ser Asp Cys Ser Glu Leu Ile Leu Lys Asp Ser Leu
785 790 795 800
Ala Ala Ala Trp Tyr Val Val Ala His Phe Ile Asn Gln Ala Gln Leu
805 810 815
Asn His Leu Ile Gly Asp Ile Lys Asn Tyr Ile Gln Phe Ile Ser Asp

820 825 830
Ile Asp Arg Arg Ala Lys Ser Thr Gly Asn Pro Val Ser Glu Ser Thr
835 840 845
Glu Ile Gln Ile Glu Arg Tyr Arg Lys Ile Leu Arg Val Leu Glu Phe
850 855 860
Ala Lys Phe Phe Cys Gly Gln Ile Thr Asn Val Leu Thr Asp Tyr Tyr
865 870 875 880
Gln Asp Glu Asn Asp Phe Ser Thr His Val Gly His Tyr Val Lys Phe

885 890 895
Glu Lys Lys Asn Met Glu Pro Ala His Ala Leu Gln Ala Phe Ser Asn
900 905 910
Ser Leu Tyr Ala Cys Gly Lys Glu Lys Lys Lys Ala Gly Phe Tyr Tyr
915 920 925
Asp Gly Met Asn Pro Ile Val Asn Arg Asn Ile Thr Leu Ala Ser Met
930 935 940
Tyr Gly Asn Lys Lys Leu Leu Glu Asn Ala Met Asn Pro Val Thr Glu

945 950 955 960
Gln Asp Ile Arg Lys Tyr Tyr Ser Leu Met Ala Glu Leu Asp Ser Val

965 970 975

Leu Lys Asn Gly Ala Val Cys Lys Ser Glu Asp Glu Gln Lys Asn Leu

980 985 990

Arg His Phe Gln Asn Leu Lys Asn Arg Ile Glu Leu Val Asp Val Leu

995 1000 1005

Thr Leu Ser Glu Leu Val Asn Asp Leu Val Ala Gln Leu Ile Gly

1010 1015 1020

Trp Val Tyr Ile Arg Glu Arg Asp Met Met Tyr Leu Gln Leu Gly

1025 1030 1035

Leu His Tyr Ile Lys Leu Tyr Phe Thr Asp Ser Val Ala Glu Asp

1040 1045 1050

Ser Tyr Leu Arg Thr Leu Asp Leu Glu Glu Gly Ser Ile Ala Asp

1055 1060 1065

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

1070 1075 1080

Pro Met Tyr Val Lys Pro Asn Lys Ser Ser Val Tyr Cys Lys Lys

1085 1090 1095

His Val Asn Ser Val Ala Thr Lys Phe Asp Ile Phe Glu Lys Glu

1100 1105 1110

Tyr Cys Asn Gly Asp Glu Thr Val Ile Glu Asn Gly Leu Arg Leu

1115 1120 1125

Phe Glu Asn Ile Asn Leu His Lys Asp Met Val Lys Phe Arg Asp

1130 1135 1140

Tyr Leu Ala His Phe Lys Tyr Phe Ala Lys Leu Asp Glu Ser Ile

1145 1150 1155

Leu Glu Leu Tyr Ser Lys Ala Tyr Asp Phe Phe Phe Ser Tyr Asn

1160 1165 1170

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

1175 1180 1185

Leu Ser Tyr Phe Ile Asn Ala Lys Leu Ser Phe Ser Thr Tyr Lys

1190 1195 1200

Ser Ser Gly Asn Lys Thr Val Gln His Arg Thr Thr Lys Ile Ser
1205 1210 1215
Val Val Ala Gln Thr Asp Tyr Phe Thr Tyr Lys Leu Arg Ser Ile
1220 1225 1230
Val Lys Asn Lys Asn Gly Val Glu Ser Ile Glu Asn Asp Asp Arg
1235 1240 1245
Arg Cys Glu Val Val Asn Ile Ala Ala Arg Asp Lys Glu Phe Val

1250 1255 1260
Asp Glu Val Cys Asn Val Ile Asn Tyr Asn Ser Asp Lys
1265 1270 1275
<210> 54
<211> 1385
<212> PRT
<213> Leptotrichia species
<400> 54
Met Gly Asn Leu Phe Gly His Lys Arg Trp Tyr Glu Val Arg Asp Lys
1 5 10 15
Lys Asp Phe Lys Ile Lys Arg Lys Val Lys Val Lys Arg Asn Tyr Asp
20 25 30

Gly Asn Lys Tyr Ile Leu Asn Ile Asn Glu Asn Asn Asn Lys Glu Lys
35 40 45
Ile Asp Asn Asn Lys Phe Ile Gly Glu Phe Val Asn Tyr Lys Lys Asn
50 55 60
Asn Asn Val Leu Lys Glu Phe Lys Arg Lys Phe His Ala Gly Asn Ile
65 70 75 80
Leu Phe Lys Leu Lys Gly Lys Glu Glu Ile Ile Arg Ile Glu Asn Asn
85 90 95

Asp Asp Phe Leu Glu Thr Glu Glu Val Val Leu Tyr Ile Glu Val Tyr
100 105 110
Gly Lys Ser Glu Lys Leu Lys Ala Leu Glu Ile Thr Lys Lys Lys Ile
115 120 125
Ile Asp Glu Ala Ile Arg Gln Gly Ile Thr Lys Asp Asp Lys Lys Ile

130 135 140

Glu Ile Lys Arg Gln Glu Asn Glu Glu Glu Ile Glu Ile Asp Ile Arg
145 150 155 160

Asp Glu Tyr Thr Asn Lys Thr Leu Asn Asp Cys Ser Ile Ile Leu Arg
165 170 175

Ile Ile Glu Asn Asp Glu Leu Glu Thr Lys Lys Ser Ile Tyr Glu Ile
180 185 190

Phe Lys Asn Ile Asn Met Ser Leu Tyr Lys Ile Ile Glu Lys Ile Ile
195 200 205

Glu Asn Glu Thr Glu Lys Val Phe Glu Asn Arg Tyr Tyr Glu Glu His
210 215 220

Leu Arg Glu Lys Leu Leu Lys Asp Asn Lys Ile Asp Val Ile Leu Thr
225 230 235 240

Asn Phe Met Glu Ile Arg Glu Lys Ile Lys Ser Asn Leu Glu Ile Met
245 250 255

Gly Phe Val Lys Phe Tyr Leu Asn Val Ser Gly Asp Lys Lys Lys Ser
260 265 270

Glu Asn Lys Lys Met Phe Val Glu Lys Ile Leu Asn Thr Asn Val Asp
275 280 285

Leu Thr Val Glu Asp Ile Val Asp Phe Ile Val Lys Glu Leu Lys Phe
290 295 300

Trp Asn Ile Thr Lys Arg Ile Glu Lys Val Lys Lys Phe Asn Asn Glu
305 310 315 320

Phe Leu Glu Asn Arg Arg Asn Arg Thr Tyr Ile Lys Ser Tyr Val Leu
325 330 335

Leu Asp Lys His Glu Lys Phe Lys Ile Glu Arg Glu Asn Lys Lys Asp
340 345 350

Lys Ile Val Lys Phe Phe Val Glu Asn Ile Lys Asn Asn Ser Ile Lys
355 360 365

Glu Lys Ile Glu Lys Ile Leu Ala Glu Phe Lys Ile Asn Glu Leu Ile
370 375 380

Lys Lys Leu Glu Lys Glu Leu Lys Lys Gly Asn Cys Asp Thr Glu Ile
385 390 395 400
Phe Gly Ile Phe Lys Lys His Tyr Lys Val Asn Phe Asp Ser Lys Lys
405 410 415

Phe Ser Asn Lys Ser Asp Glu Glu Lys Glu Leu Tyr Lys Ile Ile Tyr
420 425 430
Arg Tyr Leu Lys Gly Arg Ile Glu Lys Ile Leu Val Asn Glu Gln Lys
435 440 445
Val Arg Leu Lys Lys Met Glu Lys Ile Glu Ile Glu Lys Ile Leu Asn
450 455 460
Glu Ser Ile Leu Ser Glu Lys Ile Leu Lys Arg Val Lys Gln Tyr Thr
465 470 475 480

Leu Glu His Ile Met Tyr Leu Gly Lys Leu Arg His Asn Asp Ile Val
485 490 495
Lys Met Thr Val Asn Thr Asp Asp Phe Ser Arg Leu His Ala Lys Glu
500 505 510
Glu Leu Asp Leu Glu Leu Ile Thr Phe Phe Ala Ser Thr Asn Met Glu
515 520 525
Leu Asn Lys Ile Phe Asn Gly Lys Glu Lys Val Thr Asp Phe Phe Gly
530 535 540

Phe Asn Leu Asn Gly Gln Lys Ile Thr Leu Lys Glu Lys Val Pro Ser
545 550 555 560
Phe Lys Leu Asn Ile Leu Lys Lys Leu Asn Phe Ile Asn Asn Glu Asn
565 570 575
Asn Ile Asp Glu Lys Leu Ser His Phe Tyr Ser Phe Gln Lys Glu Gly
580 585 590
Tyr Leu Leu Arg Asn Lys Ile Leu His Asn Ser Tyr Gly Asn Ile Gln
595 600 605

Glu Thr Lys Asn Leu Lys Gly Glu Tyr Glu Asn Val Glu Lys Leu Ile
610 615 620
Lys Glu Leu Lys Val Ser Asp Glu Glu Ile Ser Lys Ser Leu Ser Leu

625 630 635 640

Asp Val Ile Phe Glu Gly Lys Val Asp Ile Ile Asn Lys Ile Asn Ser

645 650 655

Leu Lys Ile Gly Glu Tyr Lys Asp Lys Lys Tyr Leu Pro Ser Phe Ser

660 665 670

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

675 680 685

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

690 695 700

Tyr Val Asn Lys Ile Leu Tyr Glu Lys Ile Thr Ser Asn Glu Glu Asn

705 710 715 720

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

725 730 735

Lys Lys Glu Asn Lys Asn Leu Leu Ser Ile Glu Glu Tyr Tyr Lys Asn

740 745 750

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

755 760 765

Gln Asn Lys Val Thr Asn Ala Tyr Leu Glu Tyr Leu Glu Asn Thr Phe

770 775 780

Thr Glu Ile Ile Asp Phe Ser Lys Phe Asn Leu Asn Tyr Asp Glu Ile

785 790 795 800

Lys Thr Lys Ile Glu Glu Arg Lys Asp Asn Lys Ser Lys Ile Ile Ile

805 810 815

Asp Ser Ile Ser Thr Asn Ile Asn Ile Thr Asn Asp Ile Glu Tyr Ile

820 825 830

Ile Ser Ile Phe Ala Leu Leu Asn Ser Asn Thr Tyr Ile Asn Lys Ile

835 840 845

Arg Asn Arg Phe Phe Ala Thr Ser Val Trp Leu Glu Lys Gln Asn Gly

850 855 860

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

865 870 875 880

Leu Leu Ile Asn Leu Leu Arg Glu Asn Asn Ile Thr Asp Ile Leu Asp
885 890 895
Leu Lys Asn Ala Ile Ile Asp Ala Lys Ile Val Glu Asn Asp Glu Thr
900 905 910
Tyr Ile Lys Asn Tyr Ile Phe Glu Ser Asn Glu Glu Lys Leu Lys Lys
915 920 925

Arg Leu Phe Cys Glu Glu Leu Val Asp Lys Glu Asp Ile Arg Lys Ile
930 935 940
Phe Glu Asp Glu Asn Phe Lys Phe Lys Ser Phe Ile Lys Lys Asn Glu
945 950 955 960
Ile Gly Asn Phe Lys Ile Asn Phe Gly Ile Leu Ser Asn Leu Glu Cys
965 970 975
Asn Ser Glu Val Glu Ala Lys Lys Ile Ile Gly Lys Asn Ser Lys Lys
980 985 990

Leu Glu Ser Phe Ile Gln Asn Ile Ile Asp Glu Tyr Lys Ser Asn Ile
995 1000 1005
Arg Thr Leu Phe Ser Ser Glu Phe Leu Glu Lys Tyr Lys Glu Glu
1010 1015 1020
Ile Asp Asn Leu Val Glu Asp Thr Glu Ser Glu Asn Lys Asn Lys
1025 1030 1035
Phe Glu Lys Ile Tyr Tyr Pro Lys Glu His Lys Asn Glu Leu Tyr
1040 1045 1050

Ile Tyr Lys Lys Asn Leu Phe Leu Asn Ile Gly Asn Pro Asn Phe
1055 1060 1065
Asp Lys Ile Tyr Gly Leu Ile Ser Lys Asp Ile Lys Asn Val Asp
1070 1075 1080
Thr Lys Ile Leu Phe Asp Asp Asp Ile Lys Lys Asn Lys Ile Ser
1085 1090 1095
Glu Ile Asp Ala Ile Leu Lys Asn Leu Asn Asp Lys Leu Asn Gly
1100 1105 1110

Tyr Ser Asn Asp Tyr Lys Ala Lys Tyr Val Asn Lys Leu Lys Glu

1115 1120 1125

Asn Asp Asp Phe Phe Ala Lys Asn Ile Gln Asn Glu Asn Tyr Ser

1130 1135 1140

Ser Phe Gly Glu Phe Glu Lys Asp Tyr Asn Lys Val Ser Glu Tyr

1145 1150 1155

Lys Lys Ile Arg Asp Leu Val Glu Phe Asn Tyr Leu Asn Lys Ile

1160 1165 1170

Glu Ser Tyr Leu Ile Asp Ile Asn Trp Lys Leu Ala Ile Gln Met

1175 1180 1185

Ala Arg Phe Glu Arg Asp Met His Tyr Ile Val Asn Gly Leu Arg

1190 1195 1200

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

1205 1210 1215

Arg Ala Tyr Pro Lys Arg Asn Gly Ser Asp Gly Phe Tyr Thr Thr

1220 1225 1230

Thr Ala Tyr Tyr Lys Phe Phe Asp Glu Glu Ser Tyr Lys Lys Phe

1235 1240 1245

Glu Lys Ile Cys Tyr Gly Phe Gly Ile Asp Leu Ser Glu Asn Ser

1250 1255 1260

Glu Ile Asn Lys Pro Glu Asn Glu Ser Ile Arg Asn Tyr Ile Ser

1265 1270 1275

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

1280 1285 1290

Glu Gln Ile Asp Arg Val Ser Asn Leu Leu Ser Tyr Ser Thr Arg

1295 1300 1305

Tyr Asn Asn Ser Thr Tyr Ala Ser Val Phe Glu Val Phe Lys Lys

1310 1315 1320

Asp Val Asn Leu Asp Tyr Asp Glu Leu Lys Lys Lys Phe Arg Leu

1325 1330 1335

Ile Gly Asn Asn Asp Ile Leu Glu Arg Leu Met Lys Pro Lys Lys

1340 1345 1350

Val Ser Val Leu Glu Leu Glu Ser Tyr Asn Ser Asp Tyr Ile Lys

1355 1360 1365

Asn Leu Ile Ile Glu Leu Leu Thr Lys Ile Glu Asn Thr Asn Asp

1370 1375 1380

Thr Leu

1385

<210> 55

<211> 1334

<212> PRT

<213> Lachnospiraceae bacterium

<400> 55

Met Lys Ile Ser Lys Val Asp His Thr Arg Met Ala Val Ala Lys Gly

1 5 10 15

Asn Gln His Arg Arg Asp Glu Ile Ser Gly Ile Leu Tyr Lys Asp Pro

20 25 30

Thr Lys Thr Gly Ser Ile Asp Phe Asp Glu Arg Phe Lys Lys Leu Asn

35 40 45

Cys Ser Ala Lys Ile Leu Tyr His Val Phe Asn Gly Ile Ala Glu Gly

50 55 60

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

65 70 75 80

Arg Val Leu Phe Thr Gly Lys Ser Tyr Asp Arg Lys Ser Ile Ile Asp

85 90 95

Ile Asp Thr Val Leu Arg Asn Val Glu Lys Ile Asn Ala Phe Asp Arg

100 105 110

Ile Ser Thr Glu Glu Arg Glu Gln Ile Ile Asp Asp Leu Leu Glu Ile

115 120 125

Gln Leu Arg Lys Gly Leu Arg Lys Gly Lys Ala Gly Leu Arg Glu Val

130 135 140

Leu Leu Ile Gly Ala Gly Val Ile Val Arg Thr Asp Lys Lys Gln Glu

145 150 155 160

Ile Ala Asp Phe Leu Glu Ile Leu Asp Glu Asp Phe Asn Lys Thr Asn

165 170 175

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

180 185 190

Ser Pro Val Ser Arg Gly Glu Glu Arg Ile Phe Asp Val Ser Gly Ala

195 200 205

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

210 215 220

Ala Phe Leu Leu Asp Tyr Ala Asp Leu Asp Lys Asn Val Arg Phe Glu

225 230 235 240

Tyr Leu Arg Lys Ile Arg Arg Leu Ile Asn Leu Tyr Phe Tyr Val Lys

245 250 255

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

260 265 270

Glu Lys Asp Phe Asp Ile Trp Arg Asp His Glu Gln Arg Lys Glu Glu

275 280 285

Asn Gly Asp Phe Val Gly Cys Pro Asp Ile Leu Leu Ala Asp Arg Asp

290 295 300

Val Lys Lys Ser Asn Ser Lys Gln Val Lys Ile Ala Glu Arg Gln Leu

305 310 315 320

Arg Glu Ser Ile Arg Glu Lys Asn Ile Lys Arg Tyr Arg Phe Ser Ile

325 330 335

Lys Thr Ile Glu Lys Asp Asp Gly Thr Tyr Phe Phe Ala Asn Lys Gln

340 345 350

Ile Ser Val Phe Trp Ile His Arg Ile Glu Asn Ala Val Glu Arg Ile

355 360 365

Leu Gly Ser Ile Asn Asp Lys Lys Leu Tyr Arg Leu Arg Leu Gly Tyr

370 375 380

Leu Gly Glu Lys Val Trp Lys Asp Ile Leu Asn Phe Leu Ser Ile Lys

385 390 395 400

Tyr Ile Ala Val Gly Lys Ala Val Phe Asn Phe Ala Met Asp Asp Leu

405 410 415

Gln Glu Lys Asp Arg Asp Ile Glu Pro Gly Lys Ile Ser Glu Asn Ala
420 425 430
Val Asn Gly Leu Thr Ser Phe Asp Tyr Glu Gln Ile Lys Ala Asp Glu
435 440 445
Met Leu Gln Arg Glu Val Ala Val Asn Val Ala Phe Ala Ala Asn Asn
450 455 460
Leu Ala Arg Val Thr Val Asp Ile Pro Gln Asn Gly Glu Lys Glu Asp

465 470 475 480
Ile Leu Leu Trp Asn Lys Ser Asp Ile Lys Lys Tyr Lys Lys Asn Ser
485 490 495
Lys Lys Gly Ile Leu Lys Ser Ile Leu Gln Phe Phe Gly Gly Ala Ser
500 505 510
Thr Trp Asn Met Lys Met Phe Glu Ile Ala Tyr His Asp Gln Pro Gly
515 520 525
Asp Tyr Glu Glu Asn Tyr Leu Tyr Asp Ile Ile Gln Ile Ile Tyr Ser

530 535 540
Leu Arg Asn Lys Ser Phe His Phe Lys Thr Tyr Asp His Gly Asp Lys
545 550 555 560
Asn Trp Asn Arg Glu Leu Ile Gly Lys Met Ile Glu His Asp Ala Glu
565 570 575
Arg Val Ile Ser Val Glu Arg Glu Lys Phe His Ser Asn Asn Leu Pro
580 585 590
Met Phe Tyr Lys Asp Ala Asp Leu Lys Lys Ile Leu Asp Leu Leu Tyr

595 600 605
Ser Asp Tyr Ala Gly Arg Ala Ser Gln Val Pro Ala Phe Asn Thr Val
610 615 620
Leu Val Arg Lys Asn Phe Pro Glu Phe Leu Arg Lys Asp Met Gly Tyr
625 630 635 640
Lys Val His Phe Asn Asn Pro Glu Val Glu Asn Gln Trp His Ser Ala
645 650 655
Val Tyr Tyr Leu Tyr Lys Glu Ile Tyr Tyr Asn Leu Phe Leu Arg Asp

660 665 670

Lys Glu Val Lys Asn Leu Phe Tyr Thr Ser Leu Lys Asn Ile Arg Ser

675 680 685

Glu Val Ser Asp Lys Lys Gln Lys Leu Ala Ser Asp Asp Phe Ala Ser

690 695 700

Arg Cys Glu Glu Ile Glu Asp Arg Ser Leu Pro Glu Ile Cys Gln Ile

705 710 715 720

Ile Met Thr Glu Tyr Asn Ala Gln Asn Phe Gly Asn Arg Lys Val Lys

725 730 735

Ser Gln Arg Val Ile Glu Lys Asn Lys Asp Ile Phe Arg His Tyr Lys

740 745 750

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

755 760 765

Gln Glu Arg Phe Ala Phe Ile Gly Lys Ala Thr Pro Ile Pro Tyr Glu

770 775 780

Thr Thr Asp Val Lys Asn Phe Leu Pro Glu Trp Lys Ser Gly Met Tyr

785 790 795 800

Ala Ser Phe Val Glu Glu Ile Lys Asn Asn Leu Asp Leu Gln Glu Trp

805 810 815

Tyr Ile Val Gly Arg Phe Leu Asn Gly Arg Met Leu Asn Gln Leu Ala

820 825 830

Gly Ser Leu Arg Ser Tyr Ile Gln Tyr Ala Glu Asp Ile Glu Arg Arg

835 840 845

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

850 855 860

Glu Ala Cys Lys Lys Ala Val Arg Val Leu Asp Leu Cys Ile Lys Ile

865 870 875 880

Ser Thr Arg Ile Ser Ala Glu Phe Thr Asp Tyr Phe Asp Ser Glu Asp

885 890 895

Asp Tyr Ala Asp Tyr Leu Glu Lys Tyr Leu Lys Tyr Gln Asp Asp Ala

900 905 910

Ile Lys Glu Leu Ser Gly Ser Ser Tyr Ala Ala Leu Asp His Phe Cys

915 920 925

Asn Lys Asp Asp Leu Lys Phe Asp Ile Tyr Val Asn Ala Gly Gln Lys

930 935 940

Pro Ile Leu Gln Arg Asn Ile Val Met Ala Lys Leu Phe Gly Pro Asp

945 950 955 960

Asn Ile Leu Ser Glu Val Met Glu Lys Val Thr Glu Ser Ala Ile Arg

965 970 975

Glu Tyr Tyr Asp Tyr Leu Lys Lys Val Ser Gly Tyr Arg Val Arg Gly

980 985 990

Lys Cys Ser Thr Glu Lys Glu Gln Glu Asp Leu Leu Lys Phe Gln Arg

995 1000 1005

Leu Lys Asn Ala Val Glu Phe Arg Asp Val Thr Glu Tyr Ala Glu

1010 1015 1020

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

1025 1030 1035

Arg Glu Arg Asp Leu Leu Tyr Phe Gln Leu Gly Phe His Tyr Met

1040 1045 1050

Cys Leu Lys Asn Lys Ser Phe Lys Pro Ala Glu Tyr Val Asp Ile

1055 1060 1065

Arg Arg Asn Asn Gly Thr Ile Ile His Asn Ala Ile Leu Tyr Gln

1070 1075 1080

Ile Val Ser Met Tyr Ile Asn Gly Leu Asp Phe Tyr Ser Cys Asp

1085 1090 1095

Lys Glu Gly Lys Thr Leu Lys Pro Ile Glu Thr Gly Lys Gly Val

1100 1105 1110

Gly Ser Lys Ile Gly Gln Phe Ile Lys Tyr Ser Gln Tyr Leu Tyr

1115 1120 1125

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

1130 1135 1140

Val Phe Glu Asn Ile Asp Glu His Asp Asn Ile Thr Asp Leu Arg

1145 1150 1155

Lys Tyr Val Asp His Phe Lys Tyr Tyr Ala Tyr Gly Asn Lys Met

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

Ile Leu Leu Arg His Phe Val Ile Phe Tyr Pro Lys Phe Gly Ser

1205 1210 1215

Gly Lys Lys Asp Val Gly Ile Arg Asp Cys Lys Lys Glu Arg Ala

1220 1225 1230

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

1235 1240 1245

Phe Lys Leu Asp Asp Lys Ala Gly Glu Glu Ala Lys Lys Phe Pro

1250 1255 1260

Ala Arg Asp Glu Arg Tyr Leu Gln Thr Ile Ala Lys Leu Leu Tyr

1265 1270 1275

Tyr Pro Asn Glu Ile Glu Asp Met Asn Arg Phe Met Lys Lys Gly

1280 1285 1290

Glu Thr Ile Asn Lys Lys Val Gln Phe Asn Arg Lys Lys Lys Ile

1295 1300 1305

Thr Arg Lys Gln Lys Asn Asn Ser Ser Asn Glu Val Leu Ser Ser

1310 1315 1320

Thr Met Gly Tyr Leu Phe Lys Asn Ile Lys Leu

1325 1330

<210> 56

<211> 500

<212> PRT

<213> Chloroflexus aggregans

<400> 56

Met Thr Asp Gln Val Arg Arg Glu Glu Val Ala Ala Gly Glu Leu Ala

1 5 10 15

Asp Thr Pro Leu Ala Ala Ala Gln Thr Pro Ala Ala Asp Ala Ala Val

20 25 30

Ala Ala Thr Pro Ala Pro Ala Glu Ala Val Ala Pro Thr Pro Glu Gln

35 40 45

Ala Val Asp Gln Pro Ala Thr Thr Gly Glu Ser Glu Ala Pro Val Thr

50 55 60

Thr Ala Gln Ala Ala Ala His Glu Ala Glu Pro Ala Glu Ala Thr Gly

65 70 75 80

Ala Ser Phe Thr Pro Val Ser Glu Gln Gln Pro Gln Lys Pro Arg Arg

85 90 95

Leu Lys Asp Leu Gln Pro Gly Met Glu Leu Glu Gly Lys Val Thr Ser

100 105 110

Ile Ala Leu Tyr Gly Ile Phe Val Asp Val Gly Val Gly Arg Asp Gly

115 120 125

Leu Val His Ile Ser Glu Met Ser Asp Arg Arg Ile Asp Thr Pro Ser

130 135 140

Glu Leu Val Gln Ile Gly Asp Thr Val Lys Val Trp Val Lys Ser Val

145 150 155 160

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

165 170 175

Gly Glu Lys Pro Arg Arg Ser Arg Gln Ser Gln Pro Ala Gln Pro Gln

180 185 190

Pro Arg Arg Gln Glu Val Asp Arg Glu Lys Leu Ala Ser Leu Lys Val

195 200 205

Gly Glu Ile Val Glu Gly Val Ile Thr Gly Phe Ala Pro Phe Gly Ala

210 215 220

Phe Ala Asp Ile Gly Val Gly Lys Asp Gly Leu Ile His Ile Ser Glu

225 230 235 240

Leu Ser Glu Gly Arg Val Glu Lys Pro Glu Asp Ala Val Lys Val Gly

245 250 255

Glu Arg Tyr Gln Phe Lys Val Leu Glu Ile Asp Gly Glu Gly Thr Arg

260 265 270

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

275 280 285

Leu Glu Pro Gly Gln Ile Ile Glu Gly Thr Val Ser Gly Ile Ala Thr

290 295 300

Phe Gly Ala Phe Val Asp Ile Gly Val Gly Arg Asp Gly Leu Val His

305 310 315 320

Ile Ser Ala Leu Ala Pro His Arg Val Ala Lys Val Glu Asp Val Val

325 330 335

Lys Val Gly Asp Lys Val Lys Val Lys Val Leu Gly Val Asp Pro Gln

340 345 350

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

355 360 365

Thr Thr Ala Gly Asp Glu Ala Ala Glu Pro Ala Glu Glu Val Thr Pro

370 375 380

Thr Arg Arg Gly Asn Leu Glu Arg Phe Ala Ala Ala Ala Gln Thr Ala

385 390 395 400

Arg Glu Arg Ser Glu Arg Gly Glu Arg Ser Glu Arg Gly Glu Arg Arg

405 410 415

Glu Arg Arg Glu Arg Arg Pro Ala Gln Ser Ser Pro Asp Thr Tyr Ile

420 425 430

Val Gly Glu Asp Asp Asp Glu Ser Phe Glu Gly Asn Ala Thr Ile Glu

435 440 445

Asp Leu Leu Thr Lys Phe Gly Gly Ser Ser Ser Arg Arg Asp Arg Asp

450 455 460

Arg Arg Arg Arg His Glu Asp Asp Asp Asp Glu Glu Met Glu Arg Pro

465 470 475 480

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

485 490 495

Gly Tyr Asp Glu

500

<210> 57

<211> 249

<212> PRT

<213> Demequina aurantiaca

<400> 57

Met Asp Leu Thr Trp His Ala Leu Leu Ile Leu Phe Ile Val Ala Leu

1 5 10 15

Leu Ala Gly Phe Leu Asp Thr Leu Ala Gly Gly Gly Gly Leu Leu Thr

20 25 30

Val Pro Ala Leu Leu Leu Thr Gly Ile Pro Pro Leu Gln Ala Leu Gly

35 40 45

Thr Asn Lys Leu Gln Ser Ser Phe Gly Thr Gly Met Ala Thr Tyr Gln

50 55 60

Val Ile Arg Lys Lys Arg Val His Trp Arg Asp Val Arg Trp Pro Met

65 70 75 80

Val Trp Ala Phe Leu Gly Ser Ala Ala Gly Ala Val Ala Val Gln Phe

85 90 95

Ile Asp Thr Asp Ala Leu Leu Ile Ile Ile Pro Val Val Leu Ala Leu

100 105 110

Val Ala Ala Tyr Phe Leu Phe Val Pro Lys Ser His Leu Pro Pro Pro

115 120 125

Glu Pro Arg Met Ser Asp Pro Ala Tyr Glu Ala Thr Leu Val Pro Ile

130 135 140

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

145 150 155 160

Ala Leu Ser Gly Val Ala Leu Arg Ala Lys Thr Leu Val Gln Ser Thr

165 170 175

Ala Ile Ala Lys Thr Leu Asn Phe Ala Thr Asn Phe Ala Ala Leu Leu

180 185 190

Val Phe Ala Phe Ala Gly His Met Leu Trp Thr Val Gly Ala Val Met

195 200 205

Ile Ala Gly Gln Leu Ile Gly Ala Tyr Ala Gly Ser His Met Leu Phe

210 215 220

Arg Val Asn Pro Leu Val Leu Arg Val Leu Ile Val Val Met Ser Leu

225 230 235 240

Gly Met Leu Ile Arg Val Leu Leu Asp

245

<210> 58

<211> 1235

<212> PRT

<213> Thalassospira species

<400> 58

Met Arg Ile Ile Lys Pro Tyr Gly Arg Ser His Val Glu Gly Val Ala

1 5 10 15

Thr Gln Glu Pro Arg Arg Lys Leu Arg Leu Asn Ser Ser Pro Asp Ile

20 25 30

Ser Arg Asp Ile Pro Gly Phe Ala Gln Ser His Asp Ala Leu Ile Ile

35 40 45

Ala Gln Trp Ile Ser Ala Ile Asp Lys Ile Ala Thr Lys Pro Lys Pro

50 55 60

Asp Lys Lys Pro Thr Gln Ala Gln Ile Asn Leu Arg Thr Thr Leu Gly

65 70 75 80

Asp Ala Ala Trp Gln His Val Met Ala Glu Asn Leu Leu Pro Ala Ala

85 90 95

Thr Asp Pro Ala Ile Arg Glu Lys Leu His Leu Ile Trp Gln Ser Lys

100 105 110

Ile Ala Pro Trp Gly Thr Ala Arg Pro Gln Ala Glu Lys Asp Gly Lys

115 120 125

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

130 135 140

Pro Glu Ala Ile Thr Gln Asn Val Ala Arg Gln Ile Ala Lys Asp Ile

145 150 155 160

Tyr Asp His Leu His Val Ala Ala Lys Arg Lys Gly Arg Glu Pro Ala
165 170 175
Lys Gln Gly Glu Ser Ser Asn Lys Pro Gly Lys Phe Lys Pro Asp Arg
180 185 190
Lys Arg Gly Leu Ile Glu Glu Arg Ala Glu Ser Ile Ala Lys Asn Ala

195 200 205
Leu Arg Pro Gly Ser His Ala Pro Cys Pro Trp Gly Pro Asp Asp Gln
210 215 220
Ala Thr Tyr Glu Gln Ala Gly Asp Val Ala Gly Gln Ile Tyr Ala Ala
225 230 235 240
Ala Arg Asp Cys Leu Glu Glu Lys Lys Arg Arg Ser Gly Asn Arg Asn
245 250 255
Thr Ser Ser Val Gln Tyr Leu Pro Arg Asp Leu Ala Ala Lys Ile Leu

260 265 270
Tyr Ala Gln Tyr Gly Arg Val Phe Gly Pro Asp Thr Thr Ile Lys Ala
275 280 285
Ala Leu Asp Glu Gln Pro Ser Leu Phe Ala Leu His Lys Ala Ile Lys
290 295 300
Asp Cys Tyr His Arg Leu Ile Asn Asp Ala Arg Lys Arg Asp Ile Leu
305 310 315 320
Arg Ile Leu Pro Arg Asn Met Ala Ala Leu Phe Arg Leu Val Arg Ala

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

385 390 395 400
Ile Trp Arg His Ile Ile Ala Leu Ala Ser Arg Thr Leu His Asp Trp

405 410 415

Ala Asp Pro His Ser Gln Lys Phe Ser Gly Glu Asn Asp Asp Ile Leu

420 425 430

Leu Leu Ala Lys Asp Ala Ile Glu Asp Asp Val Phe Lys Ala Gly His

435 440 445

Tyr Glu Arg Lys Cys Asp Val Leu Phe Gly Ala Gln Ala Ser Leu Phe

450 455 460

Cys Gly Ala Glu Asp Phe Glu Lys Ala Ile Leu Lys Gln Ala Ile Thr

465 470 475 480

Gly Thr Gly Asn Leu Arg Asn Ala Thr Phe His Phe Lys Gly Lys Val

485 490 495

Arg Phe Glu Lys Glu Leu Gln Glu Leu Thr Lys Asp Val Pro Val Glu

500 505 510

Val Gln Ser Ala Ile Ala Ala Leu Trp Gln Lys Asp Ala Glu Gly Arg

515 520 525

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

530 535 540

Leu Thr Glu Glu Gln Asn Arg His Ile Phe Ala Ala Leu Thr Ala Ala

545 550 555 560

Met Ala Gln Pro Gly Asp Val Pro Leu Pro Arg Leu Arg Arg Val Leu

565 570 575

Ala Arg His Asp Ser Ile Cys Gln Arg Gly Arg Ile Leu Pro Leu Ser

580 585 590

Pro Cys Pro Asp Arg Ala Lys Leu Glu Glu Ser Pro Ala Leu Thr Cys

595 600 605

Gln Tyr Thr Val Leu Lys Met Leu Tyr Asp Gly Pro Phe Arg Ala Trp

610 615 620

Leu Ala Gln Gln Asn Ser Thr Ile Leu Asn His Tyr Ile Asp Ser Thr

625 630 635 640

Ile Ala Arg Thr Asp Lys Ala Ala Arg Asp Met Asn Gly Arg Lys Leu

645 650 655

Ala Gln Ala Glu Lys Asp Leu Ile Thr Ser Arg Ala Ala Asp Leu Pro
660 665 670
Arg Leu Ser Val Asp Glu Lys Met Gly Asp Phe Leu Ala Arg Leu Thr
675 680 685
Ala Ala Thr Ala Thr Glu Met Arg Val Gln Arg Gly Tyr Gln Ser Asp
690 695 700
Gly Glu Asn Ala Gln Lys Gln Ala Ala Phe Ile Gly Gln Phe Glu Cys

705 710 715 720
Asp Val Ile Gly Arg Ala Phe Ala Asp Phe Leu Asn Gln Ser Gly Phe
725 730 735
Asp Phe Val Leu Lys Leu Lys Ala Asp Thr Pro Gln Pro Asp Ala Ala
740 745 750
Gln Cys Asp Val Thr Ala Leu Ile Ala Pro Asp Asp Ile Ser Val Ser
755 760 765
Pro Pro Gln Ala Trp Gln Gln Val Leu Tyr Phe Ile Leu His Leu Val

770 775 780
Pro Val Asp Asp Ala Ser His Leu Leu His Gln Ile Arg Lys Trp Gln
785 790 795 800
Val Leu Glu Gly Lys Glu Lys Pro Ala Gln Ile Ala His Asp Val Gln
805 810 815
Ser Val Leu Met Leu Tyr Leu Asp Met His Asp Ala Lys Phe Thr Gly
820 825 830
Gly Ala Ala Leu His Gly Ile Glu Lys Phe Ala Glu Phe Phe Ala His

835 840 845
Ala Ala Asp Phe Arg Ala Val Phe Pro Pro Gln Ser Leu Gln Asp Gln
850 855 860
Asp Arg Ser Ile Pro Arg Arg Gly Leu Arg Glu Ile Val Arg Phe Gly
865 870 875 880
His Leu Pro Leu Leu Gln His Met Ser Gly Thr Val Gln Ile Thr His
885 890 895
Asp Asn Val Val Ala Trp Gln Ala Ala Arg Thr Ala Gly Ala Thr Gly

900 905 910

Met Ser Pro Ile Ala Arg Arg Gln Lys Gln Arg Glu Glu Leu His Ala

915 920 925

Leu Ala Val Glu Arg Thr Ala Arg Phe Arg Asn Ala Asp Leu Gln Asn

930 935 940

Tyr Met His Ala Leu Val Asp Val Ile Lys His Arg Gln Leu Ser Ala

945 950 955 960

Gln Val Thr Leu Ser Asp Gln Val Arg Leu His Arg Leu Met Met Gly

965 970 975

Val Leu Gly Arg Leu Val Asp Tyr Ala Gly Leu Trp Glu Arg Asp Leu

980 985 990

Tyr Phe Val Val Leu Ala Leu Leu Tyr His His Gly Ala Thr Pro Asp

995 1000 1005

Asp Val Phe Lys Gly Gln Gly Lys Lys Asn Leu Ala Asp Gly Gln

1010 1015 1020

Val Val Ala Ala Leu Lys Pro Lys Asn Arg Lys Ala Ala Ala Pro

1025 1030 1035

Val Gly Val Phe Asp Asp Leu Asp His Tyr Gly Ile Tyr Gln Asp

1040 1045 1050

Asp Arg Gln Ser Ile Arg Asn Gly Leu Ser His Phe Asn Met Leu

1055 1060 1065

Arg Gly Gly Lys Ala Pro Asp Leu Ser His Trp Val Asn Gln Thr

1070 1075 1080

Arg Ser Leu Val Ala His Asp Arg Lys Leu Lys Asn Ala Val Ala

1085 1090 1095

Lys Ser Val Ile Glu Met Leu Ala Arg Glu Gly Phe Asp Leu Asp

1100 1105 1110

Trp Gly Ile Gln Thr Asp Arg Gly Gln His Ile Leu Ser His Gly

1115 1120 1125

Lys Ile Arg Thr Arg Gln Ala Gln His Phe Gln Lys Ser Arg Leu

1130 1135 1140

His Ile Val Lys Lys Ser Ala Lys Pro Asp Lys Asn Asp Thr Val

1145 1150 1155

Lys Ile Arg Glu Asn Leu His Gly Asp Ala Met Val Glu Arg Val

1160 1165 1170

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

1175 1180 1185

Val Glu Lys Arg Leu Asp His Leu Phe Leu Lys Pro Gln Asp Gln

1190 1195 1200

Lys Gly Lys Asn Gly Ile His Thr His Asn Gly Trp Ser Lys Thr

1205 1210 1215

Glu Lys Lys Arg Arg Pro Ser Arg Glu Asn Arg Lys Gly Asn His

1220 1225 1230

Glu Asn

1235

<210> 59

<211> 1312

<212> PRT

<213> Pseudobutyrivibrio species

<400> 59

Met Lys Phe Ser Lys Glu Ser His Arg Lys Thr Ala Val Gly Val Thr

1 5 10 15

Glu Ser Asn Gly Ile Ile Gly Leu Leu Tyr Lys Asp Pro Leu Asn Glu

20 25 30

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

35 40 45

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

50 55 60

Arg Ala Ser Lys Asp Leu Ala Lys Val Val Asn Lys Ala Ile Gly Asn

65 70 75 80

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

85 90 95

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

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

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

Ser Ile Lys Thr Asp Arg Lys Gly Asn Val Lys Glu Val Leu Asp Val
290 295 300
Asp Ala Thr Ile Asp Ala Ile Arg Thr Asn Asn Ile Ala Cys Tyr Arg
305 310 315 320
Arg Ala Leu Ala Tyr Ala Asn Glu Asn Pro Asp Val Phe Phe Ser Asp
325 330 335
Thr Met Leu Asn Lys Phe Trp Ile His His Val Glu Asn Glu Val Glu

340 345 350

Arg Ile Tyr Gly His Ile Asn Asn Asn Thr Gly Asp Tyr Lys Tyr Gln
355 360 365
Leu Gly Tyr Leu Ser Glu Lys Val Trp Lys Gly Ile Ile Asn Tyr Leu
370 375 380
Ser Ile Lys Tyr Ile Ala Glu Gly Lys Ala Val Tyr Asn Tyr Ala Met
385 390 395 400
Asn Ala Leu Ala Lys Asp Asn Asn Ser Asn Ala Phe Gly Lys Leu Asp
405 410 415

Glu Lys Phe Val Asn Gly Ile Thr Ser Phe Glu Tyr Glu Arg Ile Lys
420 425 430
Ala Glu Glu Thr Leu Gln Arg Glu Cys Ala Val Asn Ile Ala Phe Ala
435 440 445
Ala Asn His Leu Ala Asn Ala Thr Val Asp Leu Asn Glu Lys Asp Ser
450 455 460
Asp Phe Leu Leu Leu Lys His Glu Asp Asn Lys Asp Thr Leu Gly Ala
465 470 475 480

Val Ala Arg Pro Asn Ile Leu Arg Asn Ile Leu Gln Phe Phe Gly Gly
485 490 495
Lys Ser Arg Trp Asn Asp Phe Asp Phe Ser Gly Ile Asp Glu Ile Gln
500 505 510
Leu Leu Asp Asp Leu Arg Lys Met Ile Tyr Ser Leu Arg Asn Ser Ser
515 520 525
Phe His Phe Lys Thr Glu Asn Ile Asp Asn Asp Ser Trp Asn Thr Lys
530 535 540

Leu Ile Gly Asp Met Phe Ala Tyr Asp Phe Asn Met Ala Gly Asn Val
545 550 555 560
Gln Lys Asp Lys Met Tyr Ser Asn Asn Val Pro Met Phe Tyr Ser Thr
565 570 575
Ser Asp Ile Glu Lys Met Leu Asp Arg Leu Tyr Ala Glu Val His Glu
580 585 590

Arg Ala Ser Gln Val Pro Ser Phe Asn Ser Val Phe Val Arg Lys Asn
595 600 605

Phe Pro Asp Tyr Leu Lys Asn Asp Leu Lys Ile Thr Ser Ala Phe Gly
610 615 620

Val Asp Asp Ala Leu Lys Trp Gln Ser Ala Val Tyr Tyr Val Cys Lys
625 630 635 640

Glu Ile Tyr Tyr Asn Asp Phe Leu Gln Asn Pro Glu Thr Phe Thr Met
645 650 655

Leu Lys Asp Tyr Val Gln Cys Leu Pro Ile Asp Ile Asp Lys Ser Met
660 665 670

Asp Gln Lys Leu Lys Ser Glu Arg Asn Ala His Lys Asn Phe Lys Glu
675 680 685

Ala Phe Ala Thr Tyr Cys Lys Glu Cys Asp Ser Leu Ser Ala Ile Cys
690 695 700

Gln Met Ile Met Thr Glu Tyr Asn Asn Gln Asn Lys Gly Asn Arg Lys
705 710 715 720

Val Ile Ser Ala Arg Thr Lys Asp Gly Asp Lys Leu Ile Tyr Lys His
725 730 735

Tyr Lys Met Ile Leu Phe Glu Ala Leu Lys Asn Val Phe Thr Ile Tyr
740 745 750

Leu Glu Lys Asn Ile Asn Thr Tyr Gly Phe Leu Lys Lys Pro Lys Leu
755 760 765

Ile Asn Asn Val Pro Ala Ile Glu Glu Phe Leu Pro Asn Tyr Asn Gly
770 775 780

Arg Gln Tyr Glu Thr Leu Val Asn Arg Ile Thr Glu Glu Thr Glu Leu
785 790 795 800

Gln Lys Trp Tyr Ile Val Gly Arg Leu Leu Asn Pro Lys Gln Val Asn
805 810 815

Gln Leu Ile Gly Asn Phe Arg Ser Tyr Val Gln Tyr Val Asn Asp Val
820 825 830

Ala Arg Arg Ala Lys Gln Thr Gly Asn Asn Leu Ser Asn Asp Asn Ile

835 840 845

Ala Trp Asp Val Lys Asn Ile Ile Gln Ile Phe Asp Val Cys Thr Lys

850 855 860

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

865 870 875 880

Asp Asp Tyr Ala Arg Tyr Leu Lys Asn Phe Val Asp Tyr Thr Asn Lys

885 890 895

Asn Asn Asp His Ser Ala Thr Leu Leu Gly Asp Phe Cys Ala Lys Glu

900 905 910

Ile Asp Gly Ile Lys Ile Gly Ile Tyr His Asp Gly Thr Asn Pro Ile

915 920 925

Val Asn Arg Asn Ile Ile Gln Cys Lys Leu Tyr Gly Ala Thr Gly Ile

930 935 940

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

945 950 955 960

Ile Ile Lys Lys Tyr Met Gln Met Gln Lys Glu Ile Lys Val Tyr Gln

965 970 975

Gln Lys Gly Ile Cys Lys Thr Lys Glu Glu Gln Gln Asn Leu Lys Lys

980 985 990

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

995 1000 1005

Ser Glu Ile Leu Asp Glu Leu Gln Gly Gln Leu Ile Asn Trp Gly

1010 1015 1020

Tyr Leu Arg Glu Arg Asp Leu Met Tyr Phe Gln Leu Gly Phe His

1025 1030 1035

Tyr Leu Cys Leu His Asn Glu Ser Lys Lys Pro Val Gly Tyr Asn

1040 1045 1050

Asn Ala Gly Asp Ile Ser Gly Ala Val Leu Tyr Gln Ile Val Ala

1055 1060 1065

Met Tyr Thr Asn Gly Leu Ser Leu Ile Asp Ala Asn Gly Lys Ser

1070 1075 1080

Lys Lys Asn Ala Lys Ala Ser Ala Gly Ala Lys Val Gly Ser Phe
1085 1090 1095
Cys Ser Tyr Ser Lys Glu Ile Arg Gly Val Asp Lys Asp Thr Lys
1100 1105 1110

Glu Asp Asp Asp Pro Ile Tyr Leu Ala Gly Val Glu Leu Phe Glu
1115 1120 1125
Asn Ile Asn Glu His Gln Gln Cys Ile Asn Leu Arg Asn Tyr Ile
1130 1135 1140
Glu His Phe His Tyr Tyr Ala Lys His Asp Arg Ser Met Leu Asp
1145 1150 1155
Leu Tyr Ser Glu Val Phe Asp Arg Phe Phe Thr Tyr Asp Met Lys
1160 1165 1170

Tyr Thr Lys Asn Val Pro Asn Met Met Tyr Asn Ile Leu Leu Gln
1175 1180 1185
His Leu Val Val Pro Ala Phe Glu Phe Gly Ser Ser Glu Lys Arg
1190 1195 1200
Leu Asp Asp Asn Asp Glu Gln Thr Lys Pro Arg Ala Met Phe Thr
1205 1210 1215
Leu Arg Glu Lys Asn Gly Leu Ser Ser Glu Gln Phe Thr Tyr Arg
1220 1225 1230

Leu Gly Asp Gly Asn Ser Thr Val Lys Leu Ser Ala Arg Gly Asp
1235 1240 1245
Asp Tyr Leu Arg Ala Val Ala Ser Leu Leu Tyr Tyr Pro Asp Arg
1250 1255 1260
Ala Pro Glu Gly Leu Ile Arg Asp Ala Glu Ala Glu Asp Lys Phe
1265 1270 1275
Ala Lys Ile Asn His Ser Asn Pro Lys Ser Asp Asn Arg Asn Asn
1280 1285 1290

Arg Gly Asn Phe Lys Asn Pro Lys Val Gln Trp Tyr Asn Asn Lys
1295 1300 1305
Thr Lys Arg Lys

1310

<210> 60

<211> 1216

<212> PRT

<213> Butyrivibrio species

<400> 60

Leu Lys Lys Ser Phe Arg Val Gly Asn Thr Lys Ser Ala Leu Lys Lys
1 5 10 15

Leu Leu Met Ala Ala Tyr Cys Gly Tyr Lys Leu Ser Ile Glu Glu Lys
20 25 30

Glu Glu Ile Gln Asn Tyr Phe Val Asp Lys Leu Val Lys Glu Tyr Asn
35 40 45

Lys Asp Thr Val Leu Lys Tyr Thr Ala Lys Ser Leu Lys His Gln Asn
50 55 60

Met Val Val Gln Pro Asp Thr Asp Asn His Val Phe Leu Pro Ser Arg
65 70 75 80

Ile Ala Gly Ala Thr Gln Asn Lys Met Ser Glu Lys Glu Ala Leu Thr
85 90 95

Glu Phe Leu Lys Ala Tyr Ala Val Leu Asp Glu Glu Lys Arg His Asn
100 105 110

Leu Arg Ile Ile Leu Arg Lys Leu Val Asn Leu Tyr Phe Tyr Glu Ser
115 120 125

Pro Asp Phe Ile Tyr Pro Glu Asn Asn Glu Trp Lys Glu His Asp Asp
130 135 140

Arg Lys Asn Lys Thr Glu Thr Phe Val Ser Pro Val Lys Val Asn Glu
145 150 155 160

Glu Lys Asn Gly Lys Thr Phe Val Lys Ile Asp Val Pro Ala Thr Lys
165 170 175

Asp Leu Ile Arg Leu Lys Asn Ile Glu Cys Tyr Arg Arg Ser Val Ala
180 185 190

Glu Thr Ala Gly Asn Pro Ile Thr Tyr Phe Thr Asp His Asn Ile Ser
195 200 205

Lys Phe Trp Ile His His Ile Glu Asn Glu Val Glu Lys Ile Phe Ala
210 215 220

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

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

Asn Ile Leu Gln Phe Phe Gly Gly Gln Ser Lys Trp Lys Glu Phe Asp
355 360 365
Phe Asp Arg Tyr Ile Asn Pro Val Asn Tyr Thr Asp Ile Asp Phe Leu
370 375 380
Phe Asp Ile Lys Lys Met Val Tyr Ser Leu Arg Asn Glu Ser Phe His
385 390 395 400
Phe Thr Thr Thr Asp Thr Glu Ser Asp Trp Asn Lys Asn Leu Ile Ser
405 410 415

Ala Met Phe Glu Tyr Glu Cys Arg Arg Ile Ser Thr Val Gln Lys Asn
420 425 430
Lys Phe Phe Ser Asn Asn Leu Pro Leu Phe Tyr Gly Glu Asn Ser Leu
435 440 445
Glu Arg Val Leu His Lys Leu Tyr Asp Asp Tyr Val Asp Arg Met Ser

```
    450                 455                 460
Gln Val Pro Ser Phe Gly Asn Val Phe Val Arg Lys Lys Phe Pro Asp
465                 470                 475                 480

Tyr Met Lys Glu Ile Gly Ile Lys His Asn Leu Ser Ser Glu Asp Asn
                485                 490                 495
Leu Lys Leu Gln Gly Ala Leu Tyr Phe Leu Tyr Lys Glu Ile Tyr Tyr
            500                 505                 510
Asn Ala Phe Ile Ser Ser Glu Lys Ala Met Lys Ile Phe Val Asp Leu
        515                 520                 525
Val Asn Lys Leu Asp Thr Asn Ala Arg Asp Asp Lys Gly Arg Ile Thr
    530                 535                 540

His Glu Ala Met Ala His Lys Asn Phe Lys Asp Ala Ile Ser His Tyr
545                 550                 555                 560
Met Thr His Asp Cys Ser Leu Ala Asp Ile Cys Gln Lys Ile Met Thr
                565                 570                 575
Glu Tyr Asn Gln Gln Asn Thr Gly His Arg Lys Lys Gln Thr Thr Tyr
            580                 585                 590
Ser Ser Glu Lys Asn Pro Glu Ile Phe Arg His Tyr Lys Met Ile Leu
        595                 600                 605

Phe Met Leu Leu Gln Lys Ala Met Thr Glu Tyr Ile Ser Ser Glu Glu
    610                 615                 620
Ile Phe Asp Phe Ile Met Lys Pro Asn Ser Pro Lys Thr Asp Ile Lys
625                 630                 635                 640
Glu Glu Glu Phe Leu Pro Gln Tyr Lys Ser Cys Ala Tyr Asp Asn Leu
                645                 650                 655
Ile Lys Leu Ile Ala Asp Asn Val Glu Leu Gln Lys Trp Tyr Ile Thr
            660                 665                 670

Ala Arg Leu Leu Ser Pro Arg Glu Val Asn Gln Leu Ile Gly Ser Phe
        675                 680                 685
Arg Ser Tyr Lys Gln Phe Val Ser Asp Ile Glu Arg Arg Ala Lys Glu
    690                 695                 700
```

Thr Asn Asn Ser Leu Ser Lys Ser Gly Met Thr Val Asp Val Glu Asn
705 710 715 720
Ile Thr Lys Val Leu Asp Leu Cys Thr Lys Leu Asn Gly Arg Phe Ser
725 730 735

Asn Glu Leu Thr Asp Tyr Phe Asp Ser Lys Asp Asp Tyr Ala Val Tyr
740 745 750
Val Ser Lys Phe Leu Asp Phe Gly Phe Lys Ile Asp Glu Lys Phe Pro
755 760 765
Ala Ala Leu Leu Gly Glu Phe Cys Asn Lys Glu Glu Asn Gly Lys Lys
770 775 780
Ile Gly Ile Tyr His Asn Gly Thr Glu Pro Ile Leu Asn Ser Asn Ile
785 790 795 800

Ile Lys Ser Lys Leu Tyr Gly Ile Thr Asp Val Val Ser Arg Ala Val
805 810 815
Lys Pro Val Ser Glu Lys Leu Ile Arg Glu Tyr Leu Gln Gln Glu Val
820 825 830
Lys Ile Lys Pro Tyr Leu Glu Asn Gly Val Cys Lys Asn Lys Glu Glu
835 840 845
Gln Ala Ala Leu Arg Lys Tyr Gln Glu Leu Lys Asn Arg Ile Glu Phe
850 855 860

Arg Asp Ile Val Glu Tyr Ser Glu Ile Ile Asn Glu Leu Met Gly Gln
865 870 875 880
Leu Ile Asn Phe Ser Tyr Leu Arg Glu Arg Asp Leu Met Tyr Phe Gln
885 890 895
Leu Gly Phe His Tyr Leu Cys Leu Asn Asn Tyr Gly Ala Lys Pro Glu
900 905 910
Gly Tyr Tyr Ser Ile Val Asn Asp Lys Arg Thr Ile Lys Gly Ala Ile
915 920 925

Leu Tyr Gln Ile Val Ala Met Tyr Thr Tyr Gly Leu Pro Ile Tyr His
930 935 940
Tyr Val Asp Gly Thr Ile Ser Asp Arg Arg Lys Asn Lys Lys Thr Val

945 950 955 960

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

965 970 975

Phe Ile Tyr Tyr Ser Asp Glu Leu Phe Asn Asp Ser Leu Ile Leu Tyr

980 985 990

Asn Ala Gly Leu Glu Leu Phe Glu Asn Ile Asn Glu His Glu Asn Ile

995 1000 1005

Val Asn Leu Arg Lys Tyr Ile Asp His Phe Lys Tyr Tyr Val Ser

1010 1015 1020

Gln Asp Arg Ser Leu Leu Asp Ile Tyr Ser Glu Val Phe Asp Arg

1025 1030 1035

Tyr Phe Thr Tyr Asp Arg Lys Tyr Lys Lys Asn Val Met Asn Leu

1040 1045 1050

Phe Ser Asn Ile Met Leu Lys His Phe Ile Ile Thr Asp Phe Glu

1055 1060 1065

Phe Ser Thr Gly Glu Lys Thr Ile Gly Glu Lys Asn Thr Ala Lys

1070 1075 1080

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

1085 1090 1095

Asp Lys Phe Thr Tyr Lys Phe Lys Asp Ala Lys Pro Ile Glu Leu

1100 1105 1110

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

1115 1120 1125

Tyr Tyr Pro Glu Asn Val Val Leu Thr Asp Leu Val Arg Asn Ser

1130 1135 1140

Glu Val Glu Asp Glu Lys Arg Ile Glu Lys Tyr Asp Arg Asn His

1145 1150 1155

Asn Ser Ser Pro Thr Arg Lys Asp Lys Thr Tyr Lys Gln Asp Val

1160 1165 1170

Lys Lys Asn Tyr Asn Lys Lys Thr Ser Lys Ala Phe Asp Ser Ser

1175 1180 1185

Lys Leu Asp Thr Lys Ser Val Gly Asn Asn Leu Ser Asp Asn Pro

1190 1195 1200

Val Leu Lys Gln Phe Leu Ser Glu Ser Lys Lys Lys Arg

1205 1210 1215

<210> 61

<211> 1276

<212> PRT

<213> Blautia species

<400> 61

Met Lys Ile Ser Lys Val Asp His Val Lys Ser Gly Ile Asp Gln Lys

1 5 10 15

Leu Ser Ser Gln Arg Gly Met Leu Tyr Lys Gln Pro Gln Lys Lys Tyr

20 25 30

Glu Gly Lys Gln Leu Glu Glu His Val Arg Asn Leu Ser Arg Lys Ala

35 40 45

Lys Ala Leu Tyr Gln Val Phe Pro Val Ser Gly Asn Ser Lys Met Glu

50 55 60

Lys Glu Leu Gln Ile Ile Asn Ser Phe Ile Lys Asn Ile Leu Leu Arg

65 70 75 80

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

85 90 95

Tyr Ser Val Ala Ser Gln Ile Ser Gly Asp His Ile Gln Glu Leu Val

100 105 110

Asp Gln His Leu Lys Glu Ser Leu Arg Lys Tyr Thr Cys Val Gly Asp

115 120 125

Lys Arg Ile Tyr Val Pro Asp Ile Ile Val Ala Leu Leu Lys Ser Lys

130 135 140

Phe Asn Ser Glu Thr Leu Gln Tyr Asp Asn Ser Glu Leu Lys Ile Leu

145 150 155 160

Ile Asp Phe Ile Arg Glu Asp Tyr Leu Lys Glu Lys Gln Ile Lys Gln

165 170 175

Ile Val His Ser Ile Glu Asn Asn Ser Thr Pro Leu Arg Ile Ala Glu

180 185 190

Ile Asn Gly Gln Lys Arg Leu Ile Pro Ala Asn Val Asp Asn Pro Lys

195 200 205

Lys Ser Tyr Ile Phe Glu Phe Leu Lys Glu Tyr Ala Gln Ser Asp Pro

210 215 220

Lys Gly Gln Glu Ser Leu Leu Gln His Met Arg Tyr Leu Ile Leu Leu

225 230 235 240

Tyr Leu Tyr Gly Pro Asp Lys Ile Thr Asp Asp Tyr Cys Glu Glu Ile

245 250 255

Glu Ala Trp Asn Phe Gly Ser Ile Val Met Asp Asn Glu Gln Leu Phe

260 265 270

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

275 280 285

Gln Ile Glu Glu Gly Arg Gln Ser Lys Asp Thr Ala Lys Val Lys Lys

290 295 300

Asn Lys Ser Lys Tyr Arg Met Leu Gly Asp Lys Ile Glu His Ser Ile

305 310 315 320

Asn Glu Ser Val Val Lys His Tyr Gln Glu Ala Cys Lys Ala Val Glu

325 330 335

Glu Lys Asp Ile Pro Trp Ile Lys Tyr Ile Ser Asp His Val Met Ser

340 345 350

Val Tyr Ser Ser Lys Asn Arg Val Asp Leu Asp Lys Leu Ser Leu Pro

355 360 365

Tyr Leu Ala Lys Asn Thr Trp Asn Thr Trp Ile Ser Phe Ile Ala Met

370 375 380

Lys Tyr Val Asp Met Gly Lys Gly Val Tyr His Phe Ala Met Ser Asp

385 390 395 400

Val Asp Lys Val Gly Lys Gln Asp Asn Leu Ile Ile Gly Gln Ile Asp

405 410 415

Pro Lys Phe Ser Asp Gly Ile Ser Ser Phe Asp Tyr Glu Arg Ile Lys

420 425 430

Ala Glu Asp Asp Leu His Arg Ser Met Ser Gly Tyr Ile Ala Phe Ala
435 440 445
Val Asn Asn Phe Ala Arg Ala Ile Cys Ser Asp Glu Phe Arg Lys Lys

450 455 460
Asn Arg Lys Glu Asp Val Leu Thr Val Gly Leu Asp Glu Ile Pro Leu
465 470 475 480
Tyr Asp Asn Val Lys Arg Lys Leu Leu Gln Tyr Phe Gly Gly Ala Ser
485 490 495
Asn Trp Asp Asp Ser Ile Ile Asp Ile Ile Asp Asp Lys Asp Leu Val
500 505 510
Ala Cys Ile Lys Glu Asn Leu Tyr Val Ala Arg Asn Val Asn Phe His

515 520 525
Phe Ala Gly Ser Glu Lys Val Gln Lys Lys Gln Asp Asp Ile Leu Glu
530 535 540
Glu Ile Val Arg Lys Glu Thr Arg Asp Ile Gly Lys His Tyr Arg Lys
545 550 555 560
Val Phe Tyr Ser Asn Asn Val Ala Val Phe Tyr Cys Asp Glu Asp Ile
565 570 575
Ile Lys Leu Met Asn His Leu Tyr Gln Arg Glu Lys Pro Tyr Gln Ala

580 585 590
Gln Ile Pro Ser Tyr Asn Lys Val Ile Ser Lys Thr Tyr Leu Pro Asp
595 600 605
Leu Ile Phe Met Leu Leu Lys Gly Lys Asn Arg Thr Lys Ile Ser Asp
610 615 620
Pro Ser Ile Met Asn Met Phe Arg Gly Thr Phe Tyr Phe Leu Leu Lys
625 630 635 640
Glu Ile Tyr Tyr Asn Asp Phe Leu Gln Ala Ser Asn Leu Lys Glu Met

645 650 655
Phe Cys Glu Gly Leu Lys Asn Asn Val Lys Asn Lys Lys Ser Glu Lys
660 665 670
Pro Tyr Gln Asn Phe Met Arg Arg Phe Glu Glu Leu Glu Asn Met Gly

675 680 685

Met Asp Phe Gly Glu Ile Cys Gln Gln Ile Met Thr Asp Tyr Glu Gln

690 695 700

Gln Asn Lys Gln Lys Lys Lys Thr Ala Thr Ala Val Met Ser Glu Lys

705 710 715 720

Asp Lys Lys Ile Arg Thr Leu Asp Asn Asp Thr Gln Lys Tyr Lys His

725 730 735

Phe Arg Thr Leu Leu Tyr Ile Gly Leu Arg Glu Ala Phe Ile Ile Tyr

740 745 750

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

755 760 765

Val Lys Arg Glu Gln Pro Glu Glu Lys Glu Phe Val Asn Lys Trp Lys

770 775 780

Leu Asn Gln Tyr Ser Asp Cys Ser Glu Leu Ile Leu Lys Asp Ser Leu

785 790 795 800

Ala Ala Ala Trp Tyr Val Val Ala His Phe Ile Asn Gln Ala Gln Leu

805 810 815

Asn His Leu Ile Gly Asp Ile Lys Asn Tyr Ile Gln Phe Ile Ser Asp

820 825 830

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

835 840 845

Glu Ile Gln Ile Glu Arg Tyr Arg Lys Ile Leu Arg Val Leu Glu Phe

850 855 860

Ala Lys Phe Phe Cys Gly Gln Ile Thr Asn Val Leu Thr Asp Tyr Tyr

865 870 875 880

Gln Asp Glu Asn Asp Phe Ser Thr His Val Gly His Tyr Val Lys Phe

885 890 895

Glu Lys Lys Asn Met Glu Pro Ala His Ala Leu Gln Ala Phe Ser Asn

900 905 910

Ser Leu Tyr Ala Cys Gly Lys Glu Lys Lys Lys Ala Gly Phe Tyr Tyr

915 920 925

Asp Gly Met Asn Pro Ile Val Asn Arg Asn Ile Thr Leu Ala Ser Met

930 935 940

Tyr Gly Asn Lys Lys Leu Leu Glu Asn Ala Met Asn Pro Val Thr Glu

945 950 955 960

Gln Asp Ile Arg Lys Tyr Tyr Ser Leu Met Ala Glu Leu Asp Ser Val

965 970 975

Leu Lys Asn Gly Ala Val Cys Lys Ser Glu Asp Glu Gln Lys Asn Leu

980 985 990

Arg His Phe Gln Asn Leu Lys Asn Arg Ile Glu Leu Val Asp Val Leu

995 1000 1005

Thr Leu Ser Glu Leu Val Asn Asp Leu Val Ala Gln Leu Ile Gly

1010 1015 1020

Trp Val Tyr Ile Arg Glu Arg Asp Met Met Tyr Leu Gln Leu Gly

1025 1030 1035

Leu His Tyr Ile Lys Leu Tyr Phe Thr Asp Ser Val Ala Glu Asp

1040 1045 1050

Ser Tyr Leu Arg Thr Leu Asp Leu Glu Glu Gly Ser Ile Ala Asp

1055 1060 1065

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

1070 1075 1080

Pro Met Tyr Val Lys Pro Asn Lys Ser Ser Val Tyr Cys Lys Lys

1085 1090 1095

His Val Asn Ser Val Ala Thr Lys Phe Asp Ile Phe Glu Lys Glu

1100 1105 1110

Tyr Cys Asn Gly Asp Glu Thr Val Ile Glu Asn Gly Leu Arg Leu

1115 1120 1125

Phe Glu Asn Ile Asn Leu His Lys Asp Met Val Lys Phe Arg Asp

1130 1135 1140

Tyr Leu Ala His Phe Lys Tyr Phe Ala Lys Leu Asp Glu Ser Ile

1145 1150 1155

Leu Glu Leu Tyr Ser Lys Ala Tyr Asp Phe Phe Phe Ser Tyr Asn

1160 1165 1170

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

1175 1180 1185

Leu Ser Tyr Phe Ile Asn Ala Lys Leu Ser Phe Ser Thr Tyr Lys

1190 1195 1200

Ser Ser Gly Asn Lys Thr Val Gln His Arg Thr Thr Lys Ile Ser

1205 1210 1215

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

1220 1225 1230

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

1235 1240 1245

Arg Cys Glu Val Val Asn Ile Ala Ala Arg Asp Lys Glu Phe Val

1250 1255 1260

Asp Glu Val Cys Asn Val Ile Asn Tyr Asn Ser Asp Lys

1265 1270 1275

<210> 62

<211> 1168

<212> PRT

<213> Leptotrichia species

<400> 62

Met Lys Ile Thr Lys Ile Asp Gly Ile Ser His Lys Lys Tyr Ile Lys

1 5 10 15

Glu Gly Lys Leu Val Lys Ser Thr Ser Glu Glu Asn Lys Thr Asp Glu

20 25 30

Arg Leu Ser Glu Leu Leu Thr Ile Arg Leu Asp Thr Tyr Ile Lys Asn

35 40 45

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

50 55 60

Lys Glu Phe Phe Ser Asn Lys Val Leu Tyr Leu Lys Asp Gly Ile Leu

65 70 75 80

Tyr Leu Lys Asp Arg Arg Glu Lys Asn Gln Leu Gln Asn Lys Asn Tyr

85 90 95

Ser Glu Glu Asp Ile Ser Glu Tyr Asp Leu Lys Asn Lys Asn Asn Phe
100 105 110

Leu Val Leu Lys Lys Ile Leu Leu Asn Glu Asp Ile Asn Ser Glu Glu
115 120 125
Leu Glu Ile Phe Arg Asn Asp Phe Glu Lys Lys Leu Asp Lys Ile Asn
130 135 140
Ser Leu Lys Tyr Ser Leu Glu Glu Asn Lys Ala Asn Tyr Gln Lys Ile
145 150 155 160
Asn Glu Asn Asn Ile Lys Lys Val Glu Gly Lys Ser Lys Arg Asn Ile
165 170 175

Phe Tyr Asn Tyr Tyr Lys Asp Ser Ala Lys Arg Asn Asp Tyr Ile Asn
180 185 190
Asn Ile Gln Glu Ala Phe Asp Lys Leu Tyr Lys Lys Glu Asp Ile Glu
195 200 205
Asn Leu Phe Phe Leu Ile Glu Asn Ser Lys Lys His Glu Lys Tyr Lys
210 215 220
Ile Arg Glu Cys Tyr His Lys Ile Ile Gly Arg Lys Asn Asp Lys Glu
225 230 235 240

Asn Phe Ala Thr Ile Ile Tyr Glu Glu Ile Gln Asn Val Asn Asn Met
245 250 255
Lys Glu Leu Ile Glu Lys Val Pro Asn Val Ser Glu Leu Lys Lys Ser
260 265 270
Gln Val Phe Tyr Lys Tyr Tyr Leu Asn Lys Glu Lys Leu Asn Asp Glu
275 280 285
Asn Ile Lys Tyr Val Phe Cys His Phe Val Glu Ile Glu Met Ser Lys
290 295 300

Leu Leu Lys Asn Tyr Val Tyr Lys Lys Pro Ser Asn Ile Ser Asn Asp
305 310 315 320
Lys Val Lys Arg Ile Phe Glu Tyr Gln Ser Leu Lys Lys Leu Ile Glu
325 330 335
Asn Lys Leu Leu Asn Lys Leu Asp Thr Tyr Val Arg Asn Cys Gly Lys

340 345 350

Tyr Ser Phe Tyr Leu Gln Asp Gly Glu Ile Ala Thr Ser Asp Phe Ile

355 360 365

Val Gly Asn Arg Gln Asn Glu Ala Phe Leu Arg Asn Ile Ile Gly Val

370 375 380

Ser Ser Thr Ala Tyr Phe Ser Leu Arg Asn Ile Leu Glu Thr Glu Asn

385 390 395 400

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

405 410 415

Lys Gly Glu Glu Lys Tyr Ile Ser Gly Glu Ile Asp Lys Leu Tyr Asp

420 425 430

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

435 440 445

Tyr Asp Phe Asn Met Asn Ser Lys Lys Glu Ile Glu Asp Phe Phe Ser

450 455 460

Asn Ile Asp Glu Ala Ile Ser Ser Ile Arg His Gly Ile Val His Phe

465 470 475 480

Asn Leu Glu Leu Glu Gly Lys Asp Ile Phe Thr Phe Lys Asn Ile Val

485 490 495

Pro Ser Gln Ile Ser Lys Lys Met Phe His Asp Glu Ile Asn Glu Lys

500 505 510

Lys Leu Lys Leu Lys Ile Phe Lys Gln Leu Asn Ser Ala Asn Val Phe

515 520 525

Arg Tyr Leu Glu Lys Tyr Lys Ile Leu Asn Tyr Leu Asn Arg Thr Arg

530 535 540

Phe Glu Phe Val Asn Lys Asn Ile Pro Phe Val Pro Ser Phe Thr Lys

545 550 555 560

Leu Tyr Ser Arg Ile Asp Asp Leu Lys Asn Ser Leu Gly Ile Tyr Trp

565 570 575

Lys Thr Pro Lys Thr Asn Asp Asp Asn Lys Thr Lys Glu Ile Thr Asp

580 585 590

Ala Gln Ile Tyr Leu Leu Lys Asn Ile Tyr Tyr Gly Glu Phe Leu Asn
595 600 605
Tyr Phe Met Ser Asn Asn Gly Asn Phe Phe Glu Ile Thr Lys Glu Ile
610 615 620

Ile Glu Leu Asn Lys Asn Asp Lys Arg Asn Leu Lys Thr Gly Phe Tyr
625 630 635 640
Lys Leu Gln Lys Phe Glu Asn Leu Gln Glu Lys Thr Pro Lys Glu Tyr
645 650 655
Leu Ala Asn Ile Gln Ser Leu Tyr Met Ile Asn Ala Gly Asn Gln Asp
660 665 670
Glu Glu Glu Lys Asp Thr Tyr Ile Asp Phe Ile Gln Lys Ile Phe Leu
675 680 685

Lys Gly Phe Met Thr Tyr Leu Ala Asn Asn Gly Arg Leu Ser Leu Ile
690 695 700
Tyr Ile Gly Ser Asp Glu Glu Thr Asn Thr Ser Leu Ala Glu Lys Lys
705 710 715 720
Gln Glu Phe Asp Lys Phe Leu Lys Lys Tyr Glu Gln Asn Asn Asn Ile
725 730 735
Glu Ile Pro Tyr Glu Ile Asn Glu Phe Val Arg Glu Ile Lys Leu Gly
740 745 750

Lys Ile Leu Lys Tyr Thr Glu Arg Leu Asn Met Phe Tyr Leu Ile Leu
755 760 765
Lys Leu Leu Asn His Lys Glu Leu Thr Asn Leu Lys Gly Ser Leu Glu
770 775 780
Lys Tyr Gln Ser Ala Asn Lys Glu Glu Ala Phe Ser Asp Gln Leu Glu
785 790 795 800
Leu Ile Asn Leu Leu Asn Leu Asp Asn Asn Arg Val Thr Glu Asp Phe
805 810 815

Glu Leu Glu Ala Asp Glu Ile Gly Lys Phe Leu Asp Phe Asn Gly Asn
820 825 830
Lys Val Lys Asp Asn Lys Glu Leu Lys Lys Phe Asp Thr Asn Lys Ile

835 840 845

Tyr Phe Asp Gly Glu Asn Ile Ile Lys His Arg Ala Phe Tyr Asn Ile

850 855 860

Lys Lys Tyr Gly Met Leu Asn Leu Leu Glu Lys Ile Ser Asp Glu Ala

865 870 875 880

Lys Tyr Lys Ile Ser Ile Glu Glu Leu Lys Asn Tyr Ser Lys Lys Lys

885 890 895

Asn Glu Ile Glu Glu Asn His Thr Thr Gln Glu Asn Leu His Arg Lys

900 905 910

Tyr Ala Arg Pro Arg Lys Asp Glu Lys Phe Thr Asp Glu Asp Tyr Lys

915 920 925

Lys Tyr Glu Lys Ala Ile Arg Asn Ile Gln Gln Tyr Thr His Leu Lys

930 935 940

Asn Lys Val Glu Phe Asn Glu Leu Asn Leu Leu Gln Ser Leu Leu Leu

945 950 955 960

Arg Ile Leu His Arg Leu Val Gly Tyr Thr Ser Ile Trp Glu Arg Asp

965 970 975

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

980 985 990

Glu Ile Phe Asn Phe Asp Asn Ser Lys Asn Val Lys Tyr Lys Asn Gly

995 1000 1005

Gln Ile Val Glu Lys Tyr Ile Asn Phe Tyr Lys Glu Leu Tyr Lys

1010 1015 1020

Asp Asp Thr Glu Lys Ile Ser Ile Tyr Ser Asp Lys Lys Val Lys

1025 1030 1035

Glu Leu Lys Lys Glu Lys Lys Asp Leu Tyr Ile Arg Asn Tyr Ile

1040 1045 1050

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

1055 1060 1065

Met Leu Glu Asn Leu Arg Lys Leu Leu Ser Tyr Asp Arg Lys Leu

1070 1075 1080

Lys Asn Ala Ile Met Lys Ser Ile Val Asp Ile Leu Lys Glu Tyr
1085 1090 1095
Gly Phe Val Val Thr Phe Lys Ile Glu Lys Asp Lys Lys Ile Arg
1100 1105 1110
Ile Glu Ser Leu Lys Ser Glu Glu Val Val His Leu Lys Lys Leu
1115 1120 1125

Lys Leu Lys Asp Asn Asp Lys Lys Lys Glu Pro Ile Lys Thr Tyr
1130 1135 1140
Arg Asn Ser Lys Glu Leu Cys Lys Leu Val Lys Val Met Phe Glu
1145 1150 1155
Tyr Lys Met Lys Glu Lys Lys Ser Glu Asn
1160 1165

<210> 63
<211> 1138
<212> PRT
<213> Bacteroides ihuae
<400> 63
Met Arg Ile Thr Lys Val Lys Val Lys Glu Ser Ser Asp Gln Lys Asp
1 5 10 15

Lys Met Val Leu Ile His Arg Lys Val Gly Glu Gly Thr Leu Val Leu
20 25 30
Asp Glu Asn Leu Ala Asp Leu Thr Ala Pro Ile Ile Asp Lys Tyr Lys
35 40 45
Asp Lys Ser Phe Glu Leu Ser Leu Leu Lys Gln Thr Leu Val Ser Glu
50 55 60
Lys Glu Met Asn Ile Pro Lys Cys Asp Lys Cys Thr Ala Lys Glu Arg
65 70 75 80

Cys Leu Ser Cys Lys Gln Arg Glu Lys Arg Leu Lys Glu Val Arg Gly
85 90 95
Ala Ile Glu Lys Thr Ile Gly Ala Val Ile Ala Gly Arg Asp Ile Ile
100 105 110
Pro Arg Leu Asn Ile Phe Asn Glu Asp Glu Ile Cys Trp Leu Ile Lys

115 120 125

Pro Lys Leu Arg Asn Glu Phe Thr Phe Lys Asp Val Asn Lys Gln Val

130 135 140

Val Lys Leu Asn Leu Pro Lys Val Leu Val Glu Tyr Ser Lys Lys Asn

145 150 155 160

Asp Pro Thr Leu Phe Leu Ala Tyr Gln Gln Trp Ile Ala Ala Tyr Leu

165 170 175

Lys Asn Lys Lys Gly His Ile Lys Lys Ser Ile Leu Asn Asn Arg Val

180 185 190

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

195 200 205

Leu Glu Leu Trp Gly Glu Glu Tyr Glu Thr Asn Gln Arg Ile Ala Leu

210 215 220

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

225 230 235 240

Pro Asn Pro Glu Glu Tyr Val Ser Asp Lys Gly Glu Ile Arg Pro Ala

245 250 255

Phe His Tyr Lys Leu Lys Asn Val Leu Gln Met His Gln Ser Thr Val

260 265 270

Phe Gly Thr Asn Glu Ile Leu Cys Ile Asn Pro Ile Phe Asn Glu Asn

275 280 285

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

290 295 300

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

305 310 315 320

Asn Gln Ala Ile Tyr Tyr Leu Lys Val Glu Thr Leu Lys Glu Arg Leu

325 330 335

Ser Leu Gln Leu Glu Asn Ala Leu Arg Met Asn Leu Leu Gln Lys Gly

340 345 350

Lys Ile Lys Lys His Glu Phe Asp Lys Asn Thr Cys Ser Asn Thr Leu

355 360 365

Ser Gln Ile Lys Arg Asp Glu Phe Phe Val Leu Asn Leu Val Glu Met
370 375 380
Cys Ala Phe Ala Ala Asn Asn Ile Arg Asn Ile Val Asp Lys Glu Gln
385 390 395 400

Val Asn Glu Ile Leu Ser Lys Lys Asp Leu Cys Asn Ser Leu Ser Lys
405 410 415
Asn Thr Ile Asp Lys Glu Leu Cys Thr Lys Phe Tyr Gly Ala Asp Phe
420 425 430
Ser Gln Ile Pro Val Ala Ile Trp Ala Met Arg Gly Ser Val Gln Gln
435 440 445
Ile Arg Asn Glu Ile Val His Tyr Lys Ala Glu Ala Ile Asp Lys Ile
450 455 460

Phe Ala Leu Lys Thr Phe Glu Tyr Asp Asp Met Glu Lys Asp Tyr Ser
465 470 475 480
Asp Thr Pro Phe Lys Gln Tyr Leu Glu Leu Ser Ile Glu Lys Ile Asp
485 490 495
Ser Phe Phe Ile Glu Gln Leu Ser Ser Asn Asp Val Leu Asn Tyr Tyr
500 505 510
Cys Thr Glu Asp Val Asn Lys Leu Leu Asn Lys Cys Lys Leu Ser Leu
515 520 525

Arg Arg Thr Ser Ile Pro Phe Ala Pro Gly Phe Lys Thr Ile Tyr Glu
530 535 540
Leu Gly Cys His Leu Gln Asp Ser Ser Asn Thr Tyr Arg Ile Gly His
545 550 555 560
Tyr Leu Met Leu Ile Gly Gly Arg Val Ala Asn Ser Thr Val Thr Lys
565 570 575
Ala Ser Lys Ala Tyr Pro Ala Tyr Arg Phe Met Leu Lys Leu Ile Tyr
580 585 590

Asn His Leu Phe Leu Asn Lys Phe Leu Asp Asn His Asn Lys Arg Phe
595 600 605
Phe Met Lys Ala Val Ala Phe Val Leu Lys Asp Asn Arg Glu Asn Ala

610 615 620

Arg Asn Lys Phe Gln Tyr Ala Phe Lys Glu Ile Arg Met Met Asn Asn

625 630 635 640

Asp Glu Ser Ile Ala Ser Tyr Met Ser Tyr Ile His Ser Leu Ser Val

645 650 655

Gln Glu Gln Glu Lys Lys Gly Asp Lys Asn Asp Lys Val Arg Tyr Asn

660 665 670

Thr Glu Lys Phe Ile Glu Lys Val Phe Val Lys Gly Phe Asp Asp Phe

675 680 685

Leu Ser Trp Leu Gly Val Glu Phe Ile Leu Ser Pro Asn Gln Glu Glu

690 695 700

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

705 710 715 720

Asp Arg Val Glu His Ser Ile Asn Ser Asn Gln Glu Ser His Ile Ala

725 730 735

Phe Phe Thr Phe Cys Lys Leu Leu Asp Ala Asn His Leu Ser Asp Leu

740 745 750

Arg Asn Glu Trp Ile Lys Phe Arg Ser Ser Gly Asp Lys Glu Gly Phe

755 760 765

Ser Tyr Asn Phe Ala Ile Asp Ile Ile Glu Leu Cys Leu Leu Thr Val

770 775 780

Asp Arg Val Glu Gln Arg Arg Asp Gly Tyr Lys Glu Gln Thr Glu Leu

785 790 795 800

Lys Glu Tyr Leu Ser Phe Phe Ile Lys Gly Asn Glu Ser Glu Asn Thr

805 810 815

Val Trp Lys Gly Phe Tyr Phe Gln Gln Asp Asn Tyr Thr Pro Val Leu

820 825 830

Tyr Ser Pro Ile Glu Leu Ile Arg Lys Tyr Gly Thr Leu Glu Leu Leu

835 840 845

Lys Leu Ile Ile Val Asp Glu Asp Lys Ile Thr Gln Gly Glu Phe Glu

850 855 860

Glu Trp Gln Thr Leu Lys Lys Val Val Glu Asp Lys Val Thr Arg Arg
865 870 875 880
Asn Glu Leu His Gln Glu Trp Glu Asp Met Lys Asn Lys Ser Ser Phe
885 890 895
Ser Gln Glu Lys Cys Ser Ile Tyr Gln Lys Leu Cys Arg Asp Ile Asp
900 905 910

Arg Tyr Asn Trp Leu Asp Asn Lys Leu His Leu Val His Leu Arg Lys
915 920 925
Leu His Asn Leu Val Ile Gln Ile Leu Ser Arg Met Ala Arg Phe Ile
930 935 940
Ala Leu Trp Asp Arg Asp Phe Val Leu Leu Asp Ala Ser Arg Ala Asn
945 950 955 960
Asp Asp Tyr Lys Leu Leu Ser Phe Phe Asn Phe Arg Asp Phe Ile Asn
965 970 975

Ala Lys Lys Thr Lys Thr Asp Asp Glu Leu Leu Ala Glu Phe Gly Ser
980 985 990
Lys Ile Glu Lys Lys Asn Ala Pro Phe Ile Lys Ala Glu Asp Val Pro
995 1000 1005
Leu Met Val Glu Cys Ile Glu Ala Lys Arg Ser Phe Tyr Gln Lys
1010 1015 1020
Val Phe Phe Arg Asn Asn Leu Gln Val Leu Ala Asp Arg Asn Phe
1025 1030 1035

Ile Ala His Tyr Asn Tyr Ile Ser Lys Thr Ala Lys Cys Ser Leu
1040 1045 1050
Phe Glu Met Ile Ile Lys Leu Arg Thr Leu Met Tyr Tyr Asp Arg
1055 1060 1065
Lys Leu Arg Asn Ala Val Val Lys Ser Ile Ala Asn Val Phe Asp
1070 1075 1080
Gln Asn Gly Met Val Leu Gln Leu Ser Leu Asp Asp Ser His Glu
1085 1090 1095

Leu Lys Val Asp Lys Val Ile Ser Lys Arg Ile Val His Leu Lys

1100 1105 1110

Asn Asn Asn Ile Met Thr Asp Gln Val Pro Glu Glu Tyr Tyr Lys

1115 1120 1125

Ile Cys Arg Arg Leu Leu Glu Met Lys Lys

1130 1135

<210> 64

<211> 697

<212> PRT

<213> Porphyromonadaceae bacterium

<400> 64

Met Glu Phe Arg Asp Ser Ile Phe Lys Ser Leu Leu Gln Lys Glu Ile

1 5 10 15

Glu Lys Ala Pro Leu Cys Phe Ala Glu Lys Leu Ile Ser Gly Gly Val

20 25 30

Phe Ser Tyr Tyr Pro Ser Glu Arg Leu Lys Glu Phe Val Gly Asn His

35 40 45

Pro Phe Ser Leu Phe Arg Lys Thr Met Pro Phe Ser Pro Gly Phe Lys

50 55 60

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

65 70 75 80

Arg Phe Tyr Asp Leu Asp Ile Gly Val Tyr Leu Pro Lys Asp Gly Phe

85 90 95

Gly Asp Glu Glu Trp Asn Ala Arg Tyr Phe Leu Met Lys Leu Ile Tyr

100 105 110

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

115 120 125

Arg Glu Cys Val Asp Phe Val Lys Arg Val Asn Arg Asp Tyr Asn Cys

130 135 140

Lys Asn Asn Asn Ser Glu Glu Gln Ala Phe Ile Asp Ile Arg Ser Met

145 150 155 160

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

165 170 175

Ile Ile Ile Glu Glu Asn Lys Lys Lys Glu Thr Asn Lys Glu Gly Gln
180 185 190
Ile Asn Phe Asn Lys Phe Leu Leu Gln Val Phe Val Lys Gly Phe Asp
195 200 205

Ser Phe Leu Lys Asp Arg Thr Glu Leu Asn Phe Leu Gln Leu Pro Glu
210 215 220
Leu Gln Gly Asp Gly Thr Arg Gly Asp Asp Leu Glu Ser Leu Asp Lys
225 230 235 240
Leu Gly Ala Val Val Ala Val Asp Leu Lys Leu Asp Ala Thr Gly Ile
245 250 255
Asp Ala Asp Leu Asn Glu Asn Ile Ser Phe Tyr Thr Phe Cys Lys Leu
260 265 270

Leu Asp Ser Asn His Leu Ser Arg Leu Arg Asn Glu Ile Ile Lys Tyr
275 280 285
Gln Ser Ala Asn Ser Asp Phe Ser His Asn Glu Asp Phe Asp Tyr Asp
290 295 300
Arg Ile Ile Ser Ile Ile Glu Leu Cys Met Leu Ser Ala Asp His Val
305 310 315 320
Ser Thr Asn Asp Asn Glu Ser Ile Phe Pro Asn Asn Asp Lys Asp Phe
325 330 335

Ser Gly Ile Arg Pro Tyr Leu Ser Thr Asp Ala Lys Val Glu Thr Phe
340 345 350
Glu Asp Leu Tyr Val His Ser Asp Ala Lys Thr Pro Ile Thr Asn Ala
355 360 365
Thr Met Val Leu Asn Trp Lys Tyr Gly Thr Asp Lys Leu Phe Glu Arg
370 375 380
Leu Met Ile Ser Asp Gln Asp Phe Leu Val Thr Glu Lys Asp Tyr Phe
385 390 395 400

Val Trp Lys Glu Leu Lys Lys Asp Ile Glu Glu Lys Ile Lys Leu Arg
405 410 415
Glu Glu Leu His Ser Leu Trp Val Asn Thr Pro Lys Gly Lys Lys Gly

420 425 430

Ala Lys Lys Lys Asn Gly Arg Glu Thr Thr Gly Glu Phe Ser Glu Glu

435 440 445

Asn Lys Lys Glu Tyr Leu Glu Val Cys Arg Glu Ile Asp Arg Tyr Val

450 455 460

Asn Leu Asp Asn Lys Leu His Phe Val His Leu Lys Arg Met His Ser

465 470 475 480

Leu Leu Ile Glu Leu Leu Gly Arg Phe Val Gly Phe Thr Tyr Leu Phe

485 490 495

Glu Arg Asp Tyr Gln Tyr Tyr His Leu Glu Ile Arg Ser Arg Arg Asn

500 505 510

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

515 520 525

Gln Asn Lys Tyr Asp Lys Asp Asp Phe Phe Ala Cys Thr Phe Leu Tyr

530 535 540

Glu Lys Ala Asn Lys Val Arg Asn Phe Ile Ala His Phe Asn Tyr Leu

545 550 555 560

Thr Met Trp Asn Ser Pro Gln Glu Glu Glu His Asn Ser Asn Leu Ser

565 570 575

Gly Ala Lys Asn Ser Ser Gly Arg Gln Asn Leu Lys Cys Ser Leu Thr

580 585 590

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

595 600 605

Lys Asn Ala Val Thr Lys Ala Val Ile Asp Leu Phe Asp Lys His Gly

610 615 620

Met Val Ile Lys Phe Arg Ile Val Asn Asn Asn Asn Asn Asp Asn Lys

625 630 635 640

Asn Lys His His Leu Glu Leu Asp Asp Ile Val Pro Lys Lys Ile Met

645 650 655

His Leu Arg Gly Ile Lys Leu Lys Arg Gln Asp Gly Lys Pro Ile Pro

660 665 670

Ile Gln Thr Asp Ser Val Asp Pro Leu Tyr Cys Arg Met Trp Lys Lys
675 680 685

Leu Leu Asp Leu Lys Pro Thr Pro Phe
690 695

<210> 65

<211> 204

<212> PRT

<213> Listeria riparia

<400> 65

Met His Asp Ala Trp Ala Glu Asn Pro Lys Lys Pro Gln Ser Asp Ala
1 5 10 15

Phe Leu Lys Glu Tyr Lys Ala Cys Cys Glu Ala Ile Asp Thr Tyr Asn
20 25 30

Trp His Lys Asn Lys Ala Thr Leu Val Tyr Val Asn Glu Leu His His
35 40 45

Leu Leu Ile Asp Ile Leu Gly Arg Leu Val Gly Tyr Val Ala Ile Ala
50 55 60

Asp Arg Asp Phe Gln Cys Met Ala Asn Gln Tyr Leu Lys Ser Ser Gly
65 70 75 80

His Thr Glu Arg Val Asp Ser Trp Ile Asn Thr Ile Arg Lys Asn Arg
85 90 95

Pro Asp Tyr Ile Glu Lys Leu Asp Ile Phe Met Asn Lys Ala Gly Leu
100 105 110

Phe Val Ser Glu Lys Asn Gly Arg Asn Tyr Ile Ala His Leu Asn Tyr
115 120 125

Leu Ser Pro Lys His Lys Tyr Ser Leu Leu Tyr Leu Phe Glu Lys Leu
130 135 140

Arg Glu Met Leu Lys Tyr Asp Arg Lys Leu Lys Asn Ala Val Thr Lys
145 150 155 160

Ser Leu Ile Asp Leu Leu Asp Lys His Gly Met Cys Val Val Phe Ala
165 170 175

Asn Leu Lys Asn Asn Lys His Arg Leu Val Ile Ala Ser Leu Lys Pro

180 185 190

Lys Lys Ile Glu Thr Phe Lys Trp Lys Lys Ile Lys

195 200

<210> 66

<211

> 1082

<212> PRT

<213> Insolitispirillum peregrinum

<400> 66

Met Arg Ile Ile Arg Pro Tyr Gly Ser Ser Thr Val Ala Ser Pro Ser

1 5 10 15

Pro Gln Asp Ala Gln Pro Leu Arg Ser Leu Gln Arg Gln Asn Gly Thr

20 25 30

Phe Asp Val Ala Glu Phe Ser Arg Arg His Pro Glu Leu Val Leu Ala

35 40 45

Gln Trp Val Ala Met Leu Asp Lys Ile Ile Arg Lys Pro Ala Pro Gly

50 55 60

Lys Asn Ser Thr Ala Leu Pro Arg Pro Thr Ala Glu Gln Arg Arg Leu

65 70 75 80

Arg Gln Gln Val Gly Ala Ala Leu Trp Ala Glu Met Gln Arg His Thr

85 90 95

Pro Val Pro Pro Glu Leu Lys Ala Val Trp Asp Ser Lys Val His Pro

100 105 110

Tyr Ser Lys Asp Asn Ala Pro Ala Thr Ala Lys Thr Pro Ser His Arg

115 120 125

Gly Arg Trp Tyr Asp Arg Phe Gly Asp Pro Glu Thr Ser Ala Ala Thr

130 135 140

Val Ala Glu Gly Val Arg Arg His Leu Leu Asp Ser Ala Gln Pro Phe

145 150 155 160

Arg Ala Asn Gly Gly Gln Pro Lys Gly Lys Gly Val Ile Glu His Arg

165 170 175

Ala Leu Thr Ile Gln Asn Gly Thr Leu Leu His His His Gln Ser Glu

180 185 190

Lys Ala Gly Pro Leu Pro Glu Asp Trp Ser Thr Tyr Arg Ala Asp Glu

195 200 205

Leu Val Ser Thr Ile Gly Lys Asp Ala Arg Trp Ile Lys Val Ala Ala

210 215 220

Ser Leu Tyr Gln His Tyr Gly Arg Ile Phe Gly Pro Thr Thr Pro Ile

225 230 235 240

Ser Glu Ala Gln Thr Arg Pro Glu Phe Val Leu His Thr Ala Val Lys

245 250 255

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

260 265 270

Leu Glu Arg Leu Leu Pro Arg Thr Gly Glu Ala Leu Arg His Ala Val

275 280 285

Thr Val Gln His Gly Asn Arg Ser Leu Ala Asp Ala Val Arg Ile Gly

290 295 300

Lys Ile Leu His Tyr Gly Trp Leu Gln Asn Gly Glu Pro Asp Pro Trp

305 310 315 320

Pro Asp Asp Ala Ala Leu Tyr Ser Ser Arg Tyr Trp Gly Ser Asp Gly

325 330 335

Gln Thr Asp Ile Lys His Ser Glu Ala Val Ser Arg Val Trp Arg Arg

340 345 350

Ala Leu Thr Ala Ala Gln Arg Thr Leu Thr Ser Trp Leu Tyr Pro Ala

355 360 365

Gly Thr Asp Ala Gly Asp Ile Leu Leu Ile Gly Gln Lys Pro Asp Ser

370 375 380

Ile Asp Arg Asn Arg Leu Pro Leu Leu Tyr Gly Asp Ser Thr Arg His

385 390 395 400

Trp Thr Arg Ser Pro Gly Asp Val Trp Leu Phe Leu Lys Gln Thr Leu

405 410 415

Glu Asn Leu Arg Asn Ser Ser Phe His Phe Lys Thr Leu Ser Ala Phe

420 425 430

Thr Ser His Leu Asp Gly Thr Cys Glu Ser Glu Pro Ala Glu Gln Gln

435 440 445

Ala Ala Gln Ala Leu Trp Gln Asp Asp Arg Gln Gln Asp His Gln Gln

450 455 460

Val Phe Leu Ser Leu Arg Ala Leu Asp Ala Thr Thr Tyr Leu Pro Thr

465 470 475 480

Gly Pro Leu His Arg Ile Val Asn Ala Val Gln Ser Thr Asp Ala Thr

485 490 495

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

500 505 510

Arg Leu Lys Gly Phe Pro Val Glu Pro Val Asn Arg Arg Thr Met Glu

515 520 525

Asp Asp Pro Leu Leu Arg Cys Arg Tyr Gly Val Leu Lys Leu Leu Tyr

530 535 540

Glu Arg Gly Phe Arg Ala Trp Leu Glu Thr Arg Pro Ser Ile Ala Ser

545 550 555 560

Cys Leu Asp Gln Ser Leu Lys Arg Ser Thr Lys Ala Ala Gln Thr Ile

565 570 575

Asn Gly Lys Asn Ser Pro Gln Gly Val Glu Ile Leu Ser Arg Ala Thr

580 585 590

Lys Leu Leu Gln Ala Glu Gly Gly Gly Gly His Gly Ile His Asp Leu

595 600 605

Phe Asp Arg Leu Tyr Ala Ala Thr Ala Arg Glu Met Arg Val Gln Val

610 615 620

Gly Tyr His His Asp Ala Glu Ala Ala Arg Gln Gln Ala Glu Phe Ile

625 630 635 640

Glu Asp Leu Lys Cys Glu Val Val Ala Arg Ala Phe Cys Ala Tyr Leu

645 650 655

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

660 665 670

Leu Pro Thr Trp Pro Asp Leu Pro Asp Leu Pro Ser Ser Thr Ile Gly

675 680 685

Thr Ala Gln Ala Ala Leu Tyr Ser Val Leu His Leu Met Pro Val Glu

690 695 700

Asp Val Gly Ser Leu Leu His Gln Leu Arg Arg Trp Leu Val Ala Leu

705 710 715 720

Gln Ala Arg Gly Gly Glu Asp Gly Thr Ala Ile Thr Ala Thr Ile Pro

725 730 735

Leu Leu Glu Leu Tyr Leu Asn Arg His Asp Ala Lys Phe Ser Gly Gly

740 745 750

Gly Ala Gly Thr Gly Leu Arg Trp Asp Asp Trp Gln Val Phe Phe Asp

755 760 765

Cys Gln Ala Thr Phe Asp Arg Val Phe Pro Pro Gly Pro Ala Leu Asp

770 775 780

Ser His Arg Leu Pro Leu Arg Gly Leu Arg Glu Val Leu Arg Phe Gly

785 790 795 800

Arg Val Asn Asp Leu Ala Ala Leu Ile Gly Gln Asp Lys Ile Thr Ala

805 810 815

Ala Glu Val Asp Arg Trp His Thr Ala Glu Gln Thr Ile Ala Ala Gln

820 825 830

Gln Gln Arg Arg Glu Ala Leu His Glu Gln Leu Ser Arg Lys Lys Gly

835 840 845

Thr Asp Ala Glu Val Asp Glu Tyr Arg Ala Leu Val Thr Ala Ile Ala

850 855 860

Asp His Arg His Leu Thr Ala His Val Thr Leu Ser Asn Val Val Arg

865 870 875 880

Leu His Arg Leu Met Thr Thr Val Leu Gly Arg Leu Val Asp Tyr Gly

885 890 895

Gly Leu Trp Glu Arg Asp Leu Thr Phe Val Thr Leu Tyr Glu Ala His

900 905 910

Arg Leu Gly Gly Leu Arg Asn Leu Leu Ser Glu Ser Arg Val Asn Lys

915 920 925

Phe Leu Asp Gly Gln Thr Pro Ala Ala Leu Ser Lys Lys Asn Asn Ala

930 935 940

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

945 950 955 960

Gln Ile Arg Asn Asp Phe Ala His Phe Asn Met Leu Gln Gln Gly Lys

965 970 975

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

980 985 990

Ala His Asp Arg Lys Leu Lys Asn Ala Ile Thr Arg Ser Val Thr Thr

995 1000 1005

Leu Leu Gln Gln Asp Gly Leu Asp Ile Val Trp Thr Met Asp Ala

1010 1015 1020

Ser His Arg Leu Thr Asp Ala Lys Ile Asp Ser Arg Asn Ala Ile

1025 1030 1035

His Leu His Lys Thr His Asn Arg Ala Asn Ile Arg Glu Pro Leu

1040 1045 1050

His Gly Lys Ser Tyr Cys Arg Trp Val Ala Ala Leu Phe Gly Ala

1055 1060 1065

Thr Ser Thr Pro Ser Ala Thr Lys Lys Ser Asp Lys Ile Arg

1070 1075 1080

<210> 67

<211> 1224

<212> PRT

<213> Bergeyella zoohelcum

<400> 67

Met Glu Asn Lys Thr Ser Leu Gly Asn Asn Ile Tyr Tyr Asn Pro Phe

1 5 10 15

Lys Pro Gln Asp Lys Ser Tyr Phe Ala Gly Tyr Phe Asn Ala Ala Met

20 25 30

Glu Asn Thr Asp Ser Val Phe Arg Glu Leu Gly Lys Arg Leu Lys Gly

35 40 45

Lys Glu Tyr Thr Ser Glu Asn Phe Phe Asp Ala Ile Phe Lys Glu Asn
50 55 60
Ile Ser Leu Val Glu Tyr Glu Arg Tyr Val Lys Leu Leu Ser Asp Tyr
65 70 75 80
Phe Pro Met Ala Arg Leu Leu Asp Lys Lys Glu Val Pro Ile Lys Glu
85 90 95
Arg Lys Glu Asn Phe Lys Lys Asn Phe Lys Gly Ile Ile Lys Ala Val
100 105 110

Arg Asp Leu Arg Asn Phe Tyr Thr His Lys Glu His Gly Glu Val Glu
115 120 125
Ile Thr Asp Glu Ile Phe Gly Val Leu Asp Glu Met Leu Lys Ser Thr
130 135 140
Val Leu Thr Val Lys Lys Lys Lys Val Lys Thr Asp Lys Thr Lys Glu
145 150 155 160
Ile Leu Lys Lys Ser Ile Glu Lys Gln Leu Asp Ile Leu Cys Gln Lys
165 170 175

Lys Leu Glu Tyr Leu Arg Asp Thr Ala Arg Lys Ile Glu Glu Lys Arg
180 185 190
Arg Asn Gln Arg Glu Arg Gly Glu Lys Glu Leu Val Ala Pro Phe Lys
195 200 205
Tyr Ser Asp Lys Arg Asp Asp Leu Ile Ala Ala Ile Tyr Asn Asp Ala
210 215 220
Phe Asp Val Tyr Ile Asp Lys Lys Lys Asp Ser Leu Lys Glu Ser Ser
225 230 235 240

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

290 295 300

Ser His Gly Lys Asn Ser Ile Cys Phe Met Ala Thr His Glu Ile Phe
305 310 315 320
Ser His Leu Ala Tyr Lys Lys Leu Lys Arg Lys Val Arg Thr Ala Glu
325 330 335
Ile Asn Tyr Gly Glu Ala Glu Asn Ala Glu Gln Leu Ser Val Tyr Ala
340 345 350
Lys Glu Thr Leu Met Met Gln Met Leu Asp Glu Leu Ser Lys Val Pro
355 360 365

Asp Val Val Tyr Gln Asn Leu Ser Glu Asp Val Gln Lys Thr Phe Ile
370 375 380
Glu Asp Trp Asn Glu Tyr Leu Lys Glu Asn Asn Gly Asp Val Gly Thr
385 390 395 400
Met Glu Glu Glu Gln Val Ile His Pro Val Ile Arg Lys Arg Tyr Glu
405 410 415
Asp Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Ala Gln
420 425 430

Phe Pro Thr Leu Arg Phe Gln Val His Leu Gly Asn Tyr Leu His Asp
435 440 445
Ser Arg Pro Lys Glu Asn Leu Ile Ser Asp Arg Arg Ile Lys Glu Lys
450 455 460
Ile Thr Val Phe Gly Arg Leu Ser Glu Leu Glu His Lys Lys Ala Leu
465 470 475 480
Phe Ile Lys Asn Thr Glu Thr Asn Glu Asp Arg Glu His Tyr Trp Glu
485 490 495

Ile Phe Pro Asn Pro Asn Tyr Asp Phe Pro Lys Glu Asn Ile Ser Val
500 505 510
Asn Asp Lys Asp Phe Pro Ile Ala Gly Ser Ile Leu Asp Arg Glu Lys
515 520 525
Gln Pro Val Ala Gly Lys Ile Gly Ile Lys Val Lys Leu Leu Asn Gln
530 535 540

Gln Tyr Val Ser Glu Val Asp Lys Ala Val Lys Ala His Gln Leu Lys
545 550 555 560

Gln Arg Lys Ala Ser Lys Pro Ser Ile Gln Asn Ile Ile Glu Glu Ile
565 570 575
Val Pro Ile Asn Glu Ser Asn Pro Lys Glu Ala Ile Val Phe Gly Gly
580 585 590
Gln Pro Thr Ala Tyr Leu Ser Met Asn Asp Ile His Ser Ile Leu Tyr
595 600 605
Glu Phe Phe Asp Lys Trp Glu Lys Lys Lys Glu Lys Leu Glu Lys Lys
610 615 620

Gly Glu Lys Glu Leu Arg Lys Glu Ile Gly Lys Glu Leu Glu Lys Lys
625 630 635 640
Ile Val Gly Lys Ile Gln Ala Gln Ile Gln Gln Ile Ile Asp Lys Asp
645 650 655
Thr Asn Ala Lys Ile Leu Lys Pro Tyr Gln Asp Gly Asn Ser Thr Ala
660 665 670
Ile Asp Lys Glu Lys Leu Ile Lys Asp Leu Lys Gln Glu Gln Asn Ile
675 680 685

Leu Gln Lys Leu Lys Asp Glu Gln Thr Val Arg Glu Lys Glu Tyr Asn
690 695 700
Asp Phe Ile Ala Tyr Gln Asp Lys Asn Arg Glu Ile Asn Lys Val Arg
705 710 715 720
Asp Arg Asn His Lys Gln Tyr Leu Lys Asp Asn Leu Lys Arg Lys Tyr
725 730 735
Pro Glu Ala Pro Ala Arg Lys Glu Val Leu Tyr Tyr Arg Glu Lys Gly
740 745 750

Lys Val Ala Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met Pro Thr
755 760 765
Asp Phe Lys Asn Glu Trp Lys Gly Glu Gln His Ser Leu Leu Gln Lys
770 775 780
Ser Leu Ala Tyr Tyr Glu Gln Cys Lys Glu Glu Leu Lys Asn Leu Leu

785 790 795 800

Pro Glu Lys Val Phe Gln His Leu Pro Phe Lys Leu Gly Gly Tyr Phe

805 810 815

Gln Gln Lys Tyr Leu Tyr Gln Phe Tyr Thr Cys Tyr Leu Asp Lys Arg

820 825 830

Leu Glu Tyr Ile Ser Gly Leu Val Gln Gln Ala Glu Asn Phe Lys Ser

835 840 845

Glu Asn Lys Val Phe Lys Lys Val Glu Asn Glu Cys Phe Lys Phe Leu

850 855 860

Lys Lys Gln Asn Tyr Thr His Lys Glu Leu Asp Ala Arg Val Gln Ser

865 870 875 880

Ile Leu Gly Tyr Pro Ile Phe Leu Glu Arg Gly Phe Met Asp Glu Lys

885 890 895

Pro Thr Ile Ile Lys Gly Lys Thr Phe Lys Gly Asn Glu Ala Leu Phe

900 905 910

Ala Asp Trp Phe Arg Tyr Tyr Lys Glu Tyr Gln Asn Phe Gln Thr Phe

915 920 925

Tyr Asp Thr Glu Asn Tyr Pro Leu Val Glu Leu Glu Lys Lys Gln Ala

930 935 940

Asp Arg Lys Arg Lys Thr Lys Ile Tyr Gln Gln Lys Lys Asn Asp Val

945 950 955 960

Phe Thr Leu Leu Met Ala Lys His Ile Phe Lys Ser Val Phe Lys Gln

965 970 975

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

980 985 990

Glu Arg Leu Gly Asn Gln Glu Arg Ala Arg Gln Thr Gly Glu Arg Asn

995 1000 1005

Thr Asn Tyr Ile Trp Asn Lys Thr Val Asp Leu Lys Leu Cys Asp

1010 1015 1020

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

1025 1030 1035

Phe Ile Lys Tyr Glu Tyr Asp Gln Arg Val Gln Ala Phe Leu Lys
1040 1045 1050
Tyr Glu Glu Asn Ile Glu Trp Gln Ala Phe Leu Ile Lys Glu Ser
1055 1060 1065

Lys Glu Glu Glu Asn Tyr Pro Tyr Val Val Glu Arg Glu Ile Glu
1070 1075 1080
Gln Tyr Glu Lys Val Arg Arg Glu Glu Leu Leu Lys Glu Val His
1085 1090 1095
Leu Ile Glu Glu Tyr Ile Leu Glu Lys Val Lys Asp Lys Glu Ile
1100 1105 1110
Leu Lys Lys Gly Asp Asn Gln Asn Phe Lys Tyr Tyr Ile Leu Asn
1115 1120 1125

Gly Leu Leu Lys Gln Leu Lys Asn Glu Asp Val Glu Ser Tyr Lys
1130 1135 1140
Val Phe Asn Leu Asn Thr Glu Pro Glu Asp Val Asn Ile Asn Gln
1145 1150 1155
Leu Lys Gln Glu Ala Thr Asp Leu Glu Gln Lys Ala Phe Val Leu
1160 1165 1170
Thr Tyr Ile Arg Asn Lys Phe Ala His Asn Gln Leu Pro Lys Lys
1175 1180 1185

Glu Phe Trp Asp Tyr Cys Gln Glu Lys Tyr Gly Lys Ile Glu Lys
1190 1195 1200
Glu Lys Thr Tyr Ala Glu Tyr Phe Ala Glu Val Phe Lys Lys Glu
1205 1210 1215
Lys Glu Ala Leu Ile Lys
1220

<210> 68
<211> 1126
<212> PRT
<213> Prevotella intermedia
<400> 68
Met Glu Asp Asp Lys Lys Thr Thr Asp Ser Ile Arg Tyr Glu Leu Lys

1 5 10 15

Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val
20 25 30
Tyr Ile Thr Val Asn His Ile Asn Lys Ile Leu Glu Glu Gly Glu Ile
35 40 45
Asn Arg Asp Gly Tyr Glu Thr Thr Leu Lys Asn Thr Trp Asn Glu Ile
50 55 60
Lys Asp Ile Asn Lys Lys Asp Arg Leu Ser Lys Leu Ile Ile Lys His
65 70 75 80

Phe Pro Phe Leu Glu Ala Ala Thr Tyr Arg Leu Asn Pro Thr Asp Thr
85 90 95
Thr Lys Gln Lys Glu Glu Lys Gln Ala Glu Ala Gln Ser Leu Glu Ser
100 105 110
Leu Arg Lys Ser Phe Phe Val Phe Ile Tyr Lys Leu Arg Asp Leu Arg
115 120 125
Asn His Tyr Ser His Tyr Lys His Ser Lys Ser Leu Glu Arg Pro Lys
130 135 140

Phe Glu Glu Gly Leu Leu Glu Lys Met Tyr Asn Ile Phe Asn Ala Ser
145 150 155 160
Ile Arg Leu Val Lys Glu Asp Tyr Gln Tyr Asn Lys Asp Ile Asn Pro
165 170 175
Asp Glu Asp Phe Lys His Leu Asp Arg Thr Glu Glu Glu Phe Asn Tyr
180 185 190
Tyr Phe Thr Lys Asp Asn Glu Gly Asn Ile Thr Glu Ser Gly Leu Leu
195 200 205

Phe Phe Val Ser Leu Phe Leu Glu Lys Lys Asp Ala Ile Trp Met Gln
210 215 220
Gln Lys Leu Arg Gly Phe Lys Asp Asn Arg Glu Asn Lys Lys Lys Met
225 230 235 240
Thr Asn Glu Val Phe Cys Arg Ser Arg Met Leu Leu Pro Lys Leu Arg
245 250 255

Leu Gln Ser Thr Gln Thr Gln Asp Trp Ile Leu Leu Asp Met Leu Asn
260 265 270

Glu Leu Ile Arg Cys Pro Lys Ser Leu Tyr Glu Arg Leu Arg Glu Glu
275 280 285
Asp Arg Glu Lys Phe Arg Val Pro Ile Glu Ile Ala Asp Glu Asp Tyr
290 295 300
Asp Ala Glu Gln Glu Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp
305 310 315 320
Arg Phe Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Tyr Asn Glu Ile Phe
325 330 335

Thr Asn Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ser Ile
340 345 350
Tyr Lys Lys Gln Ile Gly Asp Tyr Lys Glu Ser His His Leu Thr His
355 360 365
Lys Leu Tyr Gly Phe Glu Arg Ile Gln Glu Phe Thr Lys Gln Asn Arg
370 375 380
Pro Asp Glu Trp Arg Lys Phe Val Lys Thr Phe Asn Ser Phe Glu Thr
385 390 395 400

Ser Lys Glu Pro Tyr Ile Pro Glu Thr Thr Pro His Tyr His Leu Glu
405 410 415
Asn Gln Lys Ile Gly Ile Arg Phe Arg Asn Asp Asn Asp Lys Ile Trp
420 425 430
Pro Ser Leu Lys Thr Asn Ser Glu Lys Asn Glu Lys Ser Lys Tyr Lys
435 440 445
Leu Asp Lys Ser Phe Gln Ala Glu Ala Phe Leu Ser Val His Glu Leu
450 455 460

Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu Lys Thr Glu Asn Thr Asp
465 470 475 480
Asn Asp Asn Glu Ile Glu Thr Lys Lys Lys Glu Asn Lys Asn Asp Lys
485 490 495
Gln Glu Lys His Lys Ile Glu Glu Ile Ile Glu Asn Lys Ile Thr Glu

500 505 510

Ile Tyr Ala Leu Tyr Asp Thr Phe Ala Asn Gly Glu Ile Lys Ser Ile

515 520 525

Asp Glu Leu Glu Glu Tyr Cys Lys Gly Lys Asp Ile Glu Ile Gly His

530 535 540

Leu Pro Lys Gln Met Ile Ala Ile Leu Lys Asp Glu His Lys Val Met

545 550 555 560

Ala Thr Glu Ala Glu Arg Lys Gln Glu Glu Met Leu Val Asp Val Gln

565 570 575

Lys Ser Leu Glu Ser Leu Asp Asn Gln Ile Asn Glu Glu Ile Glu Asn

580 585 590

Val Glu Arg Lys Asn Ser Ser Leu Lys Ser Gly Lys Ile Ala Ser Trp

595 600 605

Leu Val Asn Asp Met Met Arg Phe Gln Pro Val Gln Lys Asp Asn Glu

610 615 620

Gly Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Leu

625 630 635 640

Leu Gln Arg Thr Leu Ala Phe Phe Gly Ser Glu His Glu Arg Leu Ala

645 650 655

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

660 665 670

Phe Leu Lys Asp Thr Glu Trp Glu Lys Cys Asn Asn Ile Leu Ser Phe

675 680 685

Tyr Arg Ser Tyr Leu Glu Ala Lys Lys Asn Phe Leu Glu Ser Leu Lys

690 695 700

Pro Glu Asp Trp Glu Lys Asn Gln Tyr Phe Leu Lys Leu Lys Glu Pro

705 710 715 720

Lys Thr Lys Pro Lys Thr Leu Val Gln Gly Trp Lys Asn Gly Phe Asn

725 730 735

Leu Pro Arg Gly Ile Phe Thr Glu Pro Ile Arg Lys Trp Phe Met Lys

740 745 750

```
His Arg Glu Asn Ile Thr Val Ala Glu Leu Lys Arg Val Gly Leu Val
        755                 760                 765
Ala Lys Val Ile Pro Leu Phe Phe Ser Glu Glu Tyr Lys Asp Ser Val
    770                 775                 780

Gln Pro Phe Tyr Asn Tyr His Phe Asn Val Gly Asn Ile Asn Lys Pro
785                 790                 795                 800
Asp Glu Lys Asn Phe Leu Asn Cys Glu Glu Arg Arg Glu Leu Leu Arg
                805                 810                 815
Lys Lys Lys Asp Glu Phe Lys Lys Met Thr Asp Lys Glu Lys Glu Glu
            820                 825                 830
Asn Pro Ser Tyr Leu Glu Phe Lys Ser Trp Asn Lys Phe Glu Arg Glu
        835                 840                 845

Leu Arg Leu Val Arg Asn Gln Asp Ile Val Thr Trp Leu Leu Cys Met
    850                 855                 860
Glu Leu Phe Asn Lys Lys Lys Ile Lys Glu Leu Asn Val Glu Lys Ile
865                 870                 875                 880
Tyr Leu Lys Asn Ile Asn Thr Asn Thr Thr Lys Lys Glu Lys Asn Thr
                885                 890                 895
Glu Glu Lys Asn Gly Glu Glu Lys Asn Ile Lys Glu Lys Asn Asn Ile
            900                 905                 910

Leu Asn Arg Ile Met Pro Met Arg Leu Pro Ile Lys Val Tyr Gly Arg
        915                 920                 925
Glu Asn Phe Ser Lys Asn Lys Lys Lys Lys Ile Arg Arg Asn Thr Phe
    930                 935                 940
Phe Thr Val Tyr Ile Glu Glu Lys Gly Thr Lys Leu Leu Lys Gln Gly
945                 950                 955                 960
Asn Phe Lys Ala Leu Glu Arg Asp Arg Arg Leu Gly Gly Leu Phe Ser
                965                 970                 975

Phe Val Lys Thr Pro Ser Lys Ala Glu Ser Lys Ser Asn Thr Ile Ser
            980                 985                 990
Lys Leu Arg Val Glu Tyr Glu Leu  Gly Glu Tyr Gln Lys  Ala Arg Ile
```

995 1000 1005

Glu Ile Ile Lys Asp Met Leu Ala Leu Glu Lys Thr Leu Ile Asp

1010 1015 1020

Lys Tyr Asn Ser Leu Asp Thr Asp Asn Phe Asn Lys Met Leu Thr

1025 1030 1035

Asp Trp Leu Glu Leu Lys Gly Glu Pro Asp Lys Ala Ser Phe Gln

1040 1045 1050

Asn Asp Val Asp Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His

1055 1060 1065

Asn Gln Tyr Pro Met Arg Asn Arg Ile Ala Phe Ala Asn Ile Asn

1070 1075 1080

Pro Phe Ser Leu Ser Ser Ala Asn Thr Ser Glu Glu Lys Gly Leu

1085 1090 1095

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

1100 1105 1110

Lys Ile Ile Glu Ile Glu Lys Pro Ile Glu Thr Lys Glu

1115 1120 1125

<210> 69

<211> 1127

<212> PRT

<213> Prevotella buccae

<400> 69

Met Gln Lys Gln Asp Lys Leu Phe Val Asp Arg Lys Lys Asn Ala Ile

1 5 10 15

Phe Ala Phe Pro Lys Tyr Ile Thr Ile Met Glu Asn Lys Glu Lys Pro

20 25 30

Glu Pro Ile Tyr Tyr Glu Leu Thr Asp Lys His Phe Trp Ala Ala Phe

35 40 45

Leu Asn Leu Ala Arg His Asn Val Tyr Thr Thr Ile Asn His Ile Asn

50 55 60

Arg Arg Leu Glu Ile Ala Glu Leu Lys Asp Asp Gly Tyr Met Met Gly

65 70 75 80

Ile Lys Gly Ser Trp Asn Glu Gln Ala Lys Lys Leu Asp Lys Lys Val

85 90 95

Arg Leu Arg Asp Leu Ile Met Lys His Phe Pro Phe Leu Glu Ala Ala

100 105 110

Ala Tyr Glu Met Thr Asn Ser Lys Ser Pro Asn Asn Lys Glu Gln Arg

115 120 125

Glu Lys Glu Gln Ser Glu Ala Leu Ser Leu Asn Asn Leu Lys Asn Val

130 135 140

Leu Phe Ile Phe Leu Glu Lys Leu Gln Val Leu Arg Asn Tyr Tyr Ser

145 150 155 160

His Tyr Lys Tyr Ser Glu Glu Ser Pro Lys Pro Ile Phe Glu Thr Ser

165 170 175

Leu Leu Lys Asn Met Tyr Lys Val Phe Asp Ala Asn Val Arg Leu Val

180 185 190

Lys Arg Asp Tyr Met His His Glu Asn Ile Asp Met Gln Arg Asp Phe

195 200 205

Thr His Leu Asn Arg Lys Lys Gln Val Gly Arg Thr Lys Asn Ile Ile

210 215 220

Asp Ser Pro Asn Phe His Tyr His Phe Ala Asp Lys Glu Gly Asn Met

225 230 235 240

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

245 250 255

Asp Ala Ile Trp Met Gln Lys Lys Leu Lys Gly Phe Lys Asp Gly Arg

260 265 270

Asn Leu Arg Glu Gln Met Thr Asn Glu Val Phe Cys Arg Ser Arg Ile

275 280 285

Ser Leu Pro Lys Leu Lys Leu Glu Asn Val Gln Thr Lys Asp Trp Met

290 295 300

Gln Leu Asp Met Leu Asn Glu Leu Val Arg Cys Pro Lys Ser Leu Tyr

305 310 315 320

Glu Arg Leu Arg Glu Lys Asp Arg Glu Ser Phe Lys Val Pro Phe Asp

325 330 335

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

340 345 350

Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Val Leu Arg Tyr Phe

355 360 365

Asp Leu Asn Glu Ile Phe Glu Gln Leu Arg Phe Gln Ile Asp Leu Gly

370 375 380

Thr Tyr His Phe Ser Ile Tyr Asn Lys Arg Ile Gly Asp Glu Asp Glu

385 390 395 400

Val Arg His Leu Thr His His Leu Tyr Gly Phe Ala Arg Ile Gln Asp

405 410 415

Phe Ala Pro Gln Asn Gln Pro Glu Glu Trp Arg Lys Leu Val Lys Asp

420 425 430

Leu Asp His Phe Glu Thr Ser Gln Glu Pro Tyr Ile Ser Lys Thr Ala

435 440 445

Pro His Tyr His Leu Glu Asn Glu Lys Ile Gly Ile Lys Phe Cys Ser

450 455 460

Ala His Asn Asn Leu Phe Pro Ser Leu Gln Thr Asp Lys Thr Cys Asn

465 470 475 480

Gly Arg Ser Lys Phe Asn Leu Gly Thr Gln Phe Thr Ala Glu Ala Phe

485 490 495

Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu

500 505 510

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

515 520 525

Ile Arg Lys Glu Ile Ser Asn Ile Tyr Ala Ile Tyr Asp Ala Phe Ala

530 535 540

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

545 550 555 560

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

565 570 575

Lys Gly Arg Gln Lys Asp Met Gly Lys Glu Ala Glu Arg Lys Ile Gly
580 585 590
Glu Met Ile Asp Asp Thr Gln Arg Arg Leu Asp Leu Leu Cys Lys Gln

595 600 605
Thr Asn Gln Lys Ile Arg Ile Gly Lys Arg Asn Ala Gly Leu Leu Lys
610 615 620
Ser Gly Lys Ile Ala Asp Trp Leu Val Asn Asp Met Met Arg Phe Gln
625 630 635 640
Pro Val Gln Lys Asp Gln Asn Asn Ile Pro Ile Asn Asn Ser Lys Ala
645 650 655
Asn Ser Thr Glu Tyr Arg Met Leu Gln Arg Ala Leu Ala Leu Phe Gly

660 665 670
Ser Glu Asn Phe Arg Leu Lys Ala Tyr Phe Asn Gln Met Asn Leu Val
675 680 685
Gly Asn Asp Asn Pro His Pro Phe Leu Ala Glu Thr Gln Trp Glu His
690 695 700
Gln Thr Asn Ile Leu Ser Phe Tyr Arg Asn Tyr Leu Glu Ala Arg Lys
705 710 715 720
Lys Tyr Leu Lys Gly Leu Lys Pro Gln Asn Trp Lys Gln Tyr Gln His

725 730 735
Phe Leu Ile Leu Lys Val Gln Lys Thr Asn Arg Asn Thr Leu Val Thr
740 745 750
Gly Trp Lys Asn Ser Phe Asn Leu Pro Arg Gly Ile Phe Thr Gln Pro
755 760 765
Ile Arg Glu Trp Phe Glu Lys His Asn Asn Ser Lys Arg Ile Tyr Asp
770 775 780
Gln Ile Leu Ser Phe Asp Arg Val Gly Phe Val Ala Lys Ala Ile Pro

785 790 795 800
Leu Tyr Phe Ala Glu Glu Tyr Lys Asp Asn Val Gln Pro Phe Tyr Asp
805 810 815
Tyr Pro Phe Asn Ile Gly Asn Arg Leu Lys Pro Lys Lys Arg Gln Phe

820 825 830

Leu Asp Lys Lys Glu Arg Val Glu Leu Trp Gln Lys Asn Lys Glu Leu

835 840 845

Phe Lys Asn Tyr Pro Ser Glu Lys Lys Lys Thr Asp Leu Ala Tyr Leu

850 855 860

Asp Phe Leu Ser Trp Lys Lys Phe Glu Arg Glu Leu Arg Leu Ile Lys

865 870 875 880

Asn Gln Asp Ile Val Thr Trp Leu Met Phe Lys Glu Leu Phe Asn Met

885 890 895

Ala Thr Val Glu Gly Leu Lys Ile Gly Glu Ile His Leu Arg Asp Ile

900 905 910

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

915 920 925

Met Pro Met Lys Leu Pro Val Lys Thr Tyr Glu Thr Asp Asn Lys Gly

930 935 940

Asn Ile Leu Lys Glu Arg Pro Leu Ala Thr Phe Tyr Ile Glu Glu Thr

945 950 955 960

Glu Thr Lys Val Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys Asp

965 970 975

Arg Arg Leu Asn Gly Leu Phe Ser Phe Ala Glu Thr Thr Asp Leu Asn

980 985 990

Leu Glu Glu His Pro Ile Ser Lys Leu Ser Val Asp Leu Glu Leu Ile

995 1000 1005

Lys Tyr Gln Thr Thr Arg Ile Ser Ile Phe Glu Met Thr Leu Gly

1010 1015 1020

Leu Glu Lys Lys Leu Ile Asp Lys Tyr Ser Thr Leu Pro Thr Asp

1025 1030 1035

Ser Phe Arg Asn Met Leu Glu Arg Trp Leu Gln Cys Lys Ala Asn

1040 1045 1050

Arg Pro Glu Leu Lys Asn Tyr Val Asn Ser Leu Ile Ala Val Arg

1055 1060 1065

Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp Ala Thr Leu
1070 1075 1080
Phe Ala Glu Val Lys Lys Phe Thr Leu Phe Pro Ser Val Asp Thr
1085 1090 1095
Lys Lys Ile Glu Leu Asn Ile Ala Pro Gln Leu Leu Glu Ile Val

1100 1105 1110
Gly Lys Ala Ile Lys Glu Ile Glu Lys Ser Glu Asn Lys Asn
1115 1120 1125

<210> 70
<211> 1135
<212> PRT
<213> Porphyromonas gingivalis
<400> 70
Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15
Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30

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

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

145 150 155 160

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

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

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

Gly Glu Ile Glu Leu Asp Ser Tyr Ser Lys Lys Val Gly Arg Asn Gly
355 360 365
Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu
370 375 380
Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu
385 390 395 400

Ala Ser Tyr Pro Val Arg Phe Ser Leu Phe Ala Pro Arg Tyr Ala Ile
405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro
420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met
435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu
450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys
465 470 475 480

Ala Asn Arg Ile Leu Asp Glu Thr Ala Glu Gly Lys Leu Gln Phe Ser
485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro
500 505 510

Lys Ser Lys Asp Arg Arg Glu Lys Ala Glu Thr Thr Leu Glu Lys Tyr
515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu
530 535 540

Ser Ala Phe Asp Met Asp Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560

Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr
565 570 575

Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe
580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu
595 600 605

Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620

Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg
625 630 635 640

Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala

645 650 655

Ile Val Ala Glu Leu Arg Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670

Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Gly Phe Tyr

675 680 685

Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Ile Phe Tyr

690 695 700

Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg

725 730 735

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

740 745 750

Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Val Met

755 760 765

Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe

770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830

Lys Thr Val Arg Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro

835 840 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn
900 905 910
Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu
915 920 925

Gly Gly Asp Asn Ser Leu Ser Leu Val Pro Ala Thr Ile Glu Ile Lys
930 935 940
Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg
945 950 955 960
Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu
965 970 975
Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys
980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg
995 1000 1005
Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys
1010 1015 1020
Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys
1025 1030 1035
Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg
1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro
1055 1060 1065
Leu Ile Tyr Arg Asp Val Ser Ala Lys Val Gly Ser Ile Glu Gly
1070 1075 1080
Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser
1085 1090 1095
Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile
1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met
1115 1120 1125
Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 71

<211> 1115

<212> PRT

<213> Bacteroides pyogenes

<400> 71

Met Glu Ser Ile Lys Asn Ser Gln Lys Ser Thr Gly Lys Thr Leu Gln
1 5 10 15

Lys Asp Pro Pro Tyr Phe Gly Leu Tyr Leu Asn Met Ala Leu Leu Asn
20 25 30

Val Arg Lys Val Glu Asn His Ile Arg Lys Trp Leu Gly Asp Val Ala
35 40 45

Leu Leu Pro Glu Lys Ser Gly Phe His Ser Leu Leu Thr Thr Asp Asn
50 55 60

Leu Ser Ser Ala Lys Trp Thr Arg Phe Tyr Tyr Lys Ser Arg Lys Phe
65 70 75 80

Leu Pro Phe Leu Glu Met Phe Asp Ser Asp Lys Lys Ser Tyr Glu Asn
85 90 95

Arg Arg Glu Thr Ala Glu Cys Leu Asp Thr Ile Asp Arg Gln Lys Ile
100 105 110

Ser Ser Leu Leu Lys Glu Val Tyr Gly Lys Leu Gln Asp Ile Arg Asn
115 120 125

Ala Phe Ser His Tyr His Ile Asp Asp Gln Ser Val Lys His Thr Ala
130 135 140

Leu Ile Ile Ser Ser Glu Met His Arg Phe Ile Glu Asn Ala Tyr Ser
145 150 155 160

Phe Ala Leu Gln Lys Thr Arg Ala Arg Phe Thr Gly Val Phe Val Glu
165 170 175

Thr Asp Phe Leu Gln Ala Glu Glu Lys Gly Asp Asn Lys Lys Phe Phe
180 185 190

Ala Ile Gly Gly Asn Glu Gly Ile Lys Leu Lys Asp Asn Ala Leu Ile
195 200 205

Phe Leu Ile Cys Leu Phe Leu Asp Arg Glu Glu Ala Phe Lys Phe Leu
210 215 220

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

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

Pro Phe Ile Arg Phe Arg Ile Asp Leu Gly Cys Leu Glu Leu Ala Ser
355 360 365
Tyr Pro Lys Lys Met Gly Glu Glu Asn Asn Tyr Glu Arg Ser Val Thr
370 375 380
Asp His Ala Met Ala Phe Gly Arg Leu Thr Asp Phe His Asn Glu Asp
385 390 395 400
Ala Val Leu Gln Gln Ile Thr Lys Gly Ile Thr Asp Glu Val Arg Phe
405 410 415

Ser Leu Tyr Ala Pro Arg Tyr Ala Ile Tyr Asn Asn Lys Ile Gly Phe
420 425 430
Val Arg Thr Ser Gly Ser Asp Lys Ile Ser Phe Pro Thr Leu Lys Lys
435 440 445
Lys Gly Gly Glu Gly His Cys Val Ala Tyr Thr Leu Gln Asn Thr Lys

450 455 460

Ser Phe Gly Phe Ile Ser Ile Tyr Asp Leu Arg Lys Ile Leu Leu Leu

465 470 475 480

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

485 490 495

Glu Gln Cys Glu Lys His Trp Lys Asp Leu Ser Glu Asn Leu Phe Asp

500 505 510

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

515 520 525

Tyr Thr Leu Pro Arg Ser Lys Gly Gly Lys Leu Val Ser Ser Lys Leu

530 535 540

Ala Asp Lys Gln Glu Lys Tyr Glu Ser Glu Phe Glu Arg Arg Lys Glu

545 550 555 560

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

565 570 575

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

580 585 590

Glu Lys Lys Leu Lys Gly Tyr Val Glu Thr Leu Lys Leu Asp Cys Arg

595 600 605

Glu Arg Leu Arg Val Phe Glu Lys Arg Glu Lys Gly Glu His Pro Leu

610 615 620

Pro Pro Arg Ile Gly Glu Met Ala Thr Asp Leu Ala Lys Asp Ile Ile

625 630 635 640

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

645 650 655

Tyr Ser Glu Ile Gln Arg Cys Leu Ala Gln Tyr Ala Gly Asp Asp Asn

660 665 670

Arg Arg His Leu Asp Ser Ile Ile Arg Glu Leu Arg Leu Lys Asp Thr

675 680 685

Lys Asn Gly His Pro Phe Leu Gly Lys Val Leu Arg Pro Gly Leu Gly

690 695 700

His Thr Glu Lys Leu Tyr Gln Arg Tyr Phe Glu Glu Lys Lys Glu Trp
705 710 715 720
Leu Glu Ala Thr Phe Tyr Pro Ala Ala Ser Pro Lys Arg Val Pro Arg
725 730 735

Phe Val Asn Pro Pro Thr Gly Lys Gln Lys Glu Leu Pro Leu Ile Ile
740 745 750
Arg Asn Leu Met Lys Glu Arg Pro Glu Trp Arg Asp Trp Lys Gln Arg
755 760 765
Lys Asn Ser His Pro Ile Asp Leu Pro Ser Gln Leu Phe Glu Asn Glu
770 775 780
Ile Cys Arg Leu Leu Lys Asp Lys Ile Gly Lys Glu Pro Ser Gly Lys
785 790 795 800

Leu Lys Trp Asn Glu Met Phe Lys Leu Tyr Trp Asp Lys Glu Phe Pro
805 810 815
Asn Gly Met Gln Arg Phe Tyr Arg Cys Lys Arg Arg Val Glu Val Phe
820 825 830
Asp Lys Val Val Glu Tyr Glu Tyr Ser Glu Glu Gly Gly Asn Tyr Lys
835 840 845
Lys Tyr Tyr Glu Ala Leu Ile Asp Glu Val Val Arg Gln Lys Ile Ser
850 855 860

Ser Ser Lys Glu Lys Ser Lys Leu Gln Val Glu Asp Leu Thr Leu Ser
865 870 875 880
Val Arg Arg Val Phe Lys Arg Ala Ile Asn Glu Lys Glu Tyr Gln Leu
885 890 895
Arg Leu Leu Cys Glu Asp Asp Arg Leu Leu Phe Met Ala Val Arg Asp
900 905 910
Leu Tyr Asp Trp Lys Glu Ala Gln Leu Asp Leu Asp Lys Ile Asp Asn
915 920 925

Met Leu Gly Glu Pro Val Ser Val Ser Gln Val Ile Gln Leu Glu Gly
930 935 940
Gly Gln Pro Asp Ala Val Ile Lys Ala Glu Cys Lys Leu Lys Asp Val

945 950 955 960

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

965 970 975

Pro Tyr Phe Ala Asn His Glu Ala Thr Gln Glu Gln Val Glu Met Glu

980 985 990

Leu Arg His Tyr Glu Asp His Arg Arg Arg Val Phe Asn Trp Val Phe

995 1000 1005

Ala Leu Glu Lys Ser Val Leu Lys Asn Glu Lys Leu Arg Arg Phe

1010 1015 1020

Tyr Glu Glu Ser Gln Gly Gly Cys Glu His Arg Arg Cys Ile Asp

1025 1030 1035

Ala Leu Arg Lys Ala Ser Leu Val Ser Glu Glu Glu Tyr Glu Phe

1040 1045 1050

Leu Val His Ile Arg Asn Lys Ser Ala His Asn Gln Phe Pro Asp

1055 1060 1065

Leu Glu Ile Gly Lys Leu Pro Pro Asn Val Thr Ser Gly Phe Cys

1070 1075 1080

Glu Cys Ile Trp Ser Lys Tyr Lys Ala Ile Ile Cys Arg Ile Ile

1085 1090 1095

Pro Phe Ile Asp Pro Glu Arg Arg Phe Phe Gly Lys Leu Leu Glu

1100 1105 1110

Gln Lys

1115

<210> 72

<211> 953

<212> PRT

<213> Alistipes species

<400> 72

Met Ser Asn Glu Ile Gly Ala Phe Arg Glu His Gln Phe Ala Tyr Ala

1 5 10 15

Pro Gly Asn Glu Lys Gln Glu Glu Ala Thr Phe Ala Thr Tyr Phe Asn

20 25 30

Leu Ala Leu Ser Asn Val Glu Gly Met Met Phe Gly Glu Val Glu Ser

35 40 45

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

50 55 60

Leu Arg Gln Ile Ala Ser Phe Ile Trp Leu Ser Lys Glu Asp His Pro

65 70 75 80

Asp Lys Ala Tyr Ser Thr Glu Glu Val Lys Val Ile Val Thr Asp Leu

85 90 95

Val Arg Arg Leu Cys Phe Tyr Arg Asn Tyr Phe Ser His Cys Phe Tyr

100 105 110

Leu Asp Thr Gln Tyr Phe Tyr Ser Asp Glu Leu Val Asp Thr Thr Ala

115 120 125

Ile Gly Glu Lys Leu Pro Tyr Asn Phe His His Phe Ile Thr Asn Arg

130 135 140

Leu Phe Arg Tyr Ser Leu Pro Glu Ile Thr Leu Phe Arg Trp Asn Glu

145 150 155 160

Gly Glu Arg Lys Tyr Glu Ile Leu Arg Asp Gly Leu Ile Phe Phe Cys

165 170 175

Cys Leu Phe Leu Lys Arg Gly Gln Ala Glu Arg Phe Leu Asn Glu Leu

180 185 190

Arg Phe Phe Lys Arg Thr Asp Glu Glu Gly Arg Ile Lys Arg Thr Ile

195 200 205

Phe Thr Lys Tyr Cys Thr Arg Glu Ser His Lys His Ile Gly Ile Glu

210 215 220

Glu Gln Asp Phe Leu Ile Phe Gln Asp Ile Ile Gly Asp Leu Asn Arg

225 230 235 240

Val Pro Lys Val Cys Asp Gly Val Val Asp Leu Ser Lys Glu Asn Glu

245 250 255

Arg Tyr Ile Lys Asn Arg Glu Thr Ser Asn Glu Ser Asp Glu Asn Lys

260 265 270

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

275 280 285

Leu Met Arg Tyr Ile Val Asp Phe Gly Val Leu Pro Cys Ile Thr Phe

290 295 300

Lys Gln Asn Asp Tyr Ser Thr Lys Glu Gly Arg Gly Gln Phe His Tyr

305 310 315 320

Gln Asp Ala Ala Val Ala Gln Glu Glu Arg Cys Tyr Asn Phe Val Val

325 330 335

Arg Asn Gly Asn Val Tyr Tyr Ser Tyr Met Pro Gln Ala Gln Asn Val

340 345 350

Val Arg Ile Ser Glu Leu Gln Gly Thr Ile Ser Val Glu Glu Leu Arg

355 360 365

Asn Met Val Tyr Ala Ser Ile Asn Gly Lys Asp Val Asn Lys Ser Val

370 375 380

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

385 390 395 400

Ile Ser Gly Gln Thr Ile Lys Glu Gly Arg Val Asp Val Glu Asp Tyr

405 410 415

Arg Pro Leu Leu Asp Lys Leu Leu Leu Arg Pro Ala Ser Asn Gly Glu

420 425 430

Glu Leu Arg Arg Glu Leu Arg Lys Leu Leu Pro Lys Arg Val Cys Asp

435 440 445

Leu Leu Ser Asn Arg Phe Asp Cys Ser Glu Gly Val Ser Ala Val Glu

450 455 460

Lys Arg Leu Lys Ala Ile Leu Leu Arg His Glu Gln Leu Leu Leu Ser

465 470 475 480

Gln Asn Pro Ala Leu His Ile Asp Lys Ile Lys Ser Val Ile Asp Tyr

485 490 495

Leu Tyr Leu Phe Phe Ser Asp Asp Glu Lys Phe Arg Gln Gln Pro Thr

500 505 510

Glu Lys Ala His Arg Gly Leu Lys Asp Glu Glu Phe Gln Met Tyr His

515 520 525

Tyr Leu Val Gly Asp Tyr Asp Ser His Pro Leu Ala Leu Trp Lys Glu
530 535 540
Leu Glu Ala Ser Gly Arg Leu Lys Pro Glu Met Arg Lys Leu Thr Ser
545 550 555 560
Ala Thr Ser Leu His Gly Leu Tyr Met Leu Cys Leu Lys Gly Thr Val

565 570 575
Glu Trp Cys Arg Lys Gln Leu Met Ser Ile Gly Lys Gly Thr Ala Lys
580 585 590
Val Glu Ala Ile Ala Asp Arg Val Gly Leu Lys Leu Tyr Asp Lys Leu
595 600 605
Lys Glu Tyr Thr Pro Glu Gln Leu Glu Arg Glu Val Lys Leu Val Val
610 615 620
Met His Gly Tyr Ala Ala Ala Ala Thr Pro Lys Pro Lys Ala Gln Ala

625 630 635 640
Ala Ile Pro Ser Lys Leu Thr Glu Leu Arg Phe Tyr Ser Phe Leu Gly
645 650 655
Lys Arg Glu Met Ser Phe Ala Ala Phe Ile Arg Gln Asp Lys Lys Ala
660 665 670
Gln Lys Leu Trp Leu Arg Asn Phe Tyr Thr Val Glu Asn Ile Lys Thr
675 680 685
Leu Gln Lys Arg Gln Ala Ala Ala Asp Ala Ala Cys Lys Lys Leu Tyr

690 695 700
Asn Leu Val Gly Glu Val Glu Arg Val His Thr Asn Asp Lys Val Leu
705 710 715 720
Val Leu Val Ala Gln Arg Tyr Arg Glu Arg Leu Leu Asn Val Gly Ser
725 730 735
Lys Cys Ala Val Thr Leu Asp Asn Pro Glu Arg Gln Gln Lys Leu Ala
740 745 750
Asp Val Tyr Glu Val Gln Asn Ala Trp Leu Ser Ile Arg Phe Asp Asp

755 760 765
Leu Asp Phe Thr Leu Thr His Val Asn Leu Ser Asn Leu Arg Lys Ala

770 775 780

Tyr Asn Leu Ile Pro Arg Lys His Ile Leu Ala Phe Lys Glu Tyr Leu

785 790 795 800

Asp Asn Arg Val Lys Gln Lys Leu Cys Glu Glu Cys Arg Asn Val Arg

805 810 815

Arg Lys Glu Asp Leu Cys Thr Cys Cys Ser Pro Arg Tyr Ser Asn Leu

820 825 830

Thr Ser Trp Leu Lys Glu Asn His Ser Glu Ser Ser Ile Glu Arg Glu

835 840 845

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

850 855 860

Leu Leu Asp Glu Arg Arg Lys Ala Ile Ile Glu Tyr Gly Lys Phe Ile

865 870 875 880

Pro Phe Ser Ala Leu Val Lys Glu Cys Arg Leu Ala Asp Ala Gly Leu

885 890 895

Cys Gly Ile Arg Asn Asp Val Leu His Asp Asn Val Ile Ser Tyr Ala

900 905 910

Asp Ala Ile Gly Lys Leu Ser Ala Tyr Phe Pro Lys Glu Ala Ser Glu

915 920 925

Ala Val Glu Tyr Ile Arg Arg Thr Lys Glu Val Arg Glu Gln Arg Arg

930 935 940

Glu Glu Leu Met Ala Asn Ser Ser Gln

945 950

<210> 73

<211> 1322

<212> PRT

<213> Prevotella species

<400> 73

Met Ser Lys Glu Cys Lys Lys Gln Arg Gln Glu Lys Lys Arg Arg Leu

1 5 10 15

Gln Lys Ala Asn Phe Ser Ile Ser Leu Thr Gly Lys His Val Phe Gly

20 25 30

Ala Tyr Phe Asn Met Ala Arg Thr Asn Phe Val Lys Thr Ile Asn Tyr

35 40 45

Ile Leu Pro Ile Ala Gly Val Arg Gly Asn Tyr Ser Glu Asn Gln Ile

50 55 60

Asn Lys Met Leu His Ala Leu Phe Leu Ile Gln Ala Gly Arg Asn Glu

65 70 75 80

Glu Leu Thr Thr Glu Gln Lys Gln Trp Glu Lys Lys Leu Arg Leu Asn

85 90 95

Pro Glu Gln Gln Thr Lys Phe Gln Lys Leu Leu Phe Lys His Phe Pro

100 105 110

Val Leu Gly Pro Met Met Ala Asp Val Ala Asp His Lys Ala Tyr Leu

115 120 125

Asn Lys Lys Lys Ser Thr Val Gln Thr Glu Asp Glu Thr Phe Ala Met

130 135 140

Leu Lys Gly Val Ser Leu Ala Asp Cys Leu Asp Ile Ile Cys Leu Met

145 150 155 160

Ala Asp Thr Leu Thr Glu Cys Arg Asn Phe Tyr Thr His Lys Asp Pro

165 170 175

Tyr Asn Lys Pro Ser Gln Leu Ala Asp Gln Tyr Leu His Gln Glu Met

180 185 190

Ile Ala Lys Lys Leu Asp Lys Val Val Val Ala Ser Arg Arg Ile Leu

195 200 205

Lys Asp Arg Glu Gly Leu Ser Val Asn Glu Val Glu Phe Leu Thr Gly

210 215 220

Ile Asp His Leu His Gln Glu Val Leu Lys Asp Glu Phe Gly Asn Ala

225 230 235 240

Lys Val Lys Asp Gly Lys Val Met Lys Thr Phe Val Glu Tyr Asp Asp

245 250 255

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

260 265 270

Thr Thr Lys Asp Asp Lys Pro Val Asn Val Asn Thr Met Leu Pro Ala

275 280 285

Leu Ser Asp Phe Gly Leu Leu Tyr Phe Cys Val Leu Phe Leu Ser Lys

290 295 300

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

305 310 315 320

Pro Phe Asp Asp Lys Glu Asn Met Ile Met Ser Glu Met Leu Ser Ile

325 330 335

Tyr Arg Ile Arg Thr Pro Arg Leu His Lys Ile Asp Ser His Asp Ser

340 345 350

Lys Ala Thr Leu Ala Met Asp Ile Phe Gly Glu Leu Arg Arg Cys Pro

355 360 365

Met Glu Leu Tyr Asn Leu Leu Asp Lys Asn Ala Gly Gln Pro Phe Phe

370 375 380

His Asp Glu Val Lys His Pro Asn Ser His Thr Pro Asp Val Ser Lys

385 390 395 400

Arg Leu Arg Tyr Asp Asp Arg Phe Pro Thr Leu Ala Leu Arg Tyr Ile

405 410 415

Asp Glu Thr Glu Leu Phe Lys Arg Ile Arg Phe Gln Leu Gln Leu Gly

420 425 430

Ser Phe Arg Tyr Lys Phe Tyr Asp Lys Glu Asn Cys Ile Asp Gly Arg

435 440 445

Val Arg Val Arg Arg Ile Gln Lys Glu Ile Asn Gly Tyr Gly Arg Met

450 455 460

Gln Glu Val Ala Asp Lys Arg Met Asp Lys Trp Gly Asp Leu Ile Gln

465 470 475 480

Lys Arg Glu Glu Arg Ser Val Lys Leu Glu His Glu Glu Leu Tyr Ile

485 490 495

Asn Leu Asp Gln Phe Leu Glu Asp Thr Ala Asp Ser Thr Pro Tyr Val

500 505 510

Thr Asp Arg Arg Pro Ala Tyr Asn Ile His Ala Asn Arg Ile Gly Leu

515 520 525

Tyr Trp Glu Asp Ser Gln Asn Pro Lys Gln Tyr Lys Val Phe Asp Glu
530 535 540
Asn Gly Met Tyr Ile Pro Glu Leu Val Val Thr Glu Asp Lys Lys Ala
545 550 555 560
Pro Ile Lys Met Pro Ala Pro Arg Cys Ala Leu Ser Val Tyr Asp Leu

565 570 575
Pro Ala Met Leu Phe Tyr Glu Tyr Leu Arg Glu Gln Gln Asp Asn Glu
580 585 590
Phe Pro Ser Ala Glu Gln Val Ile Ile Glu Tyr Glu Asp Asp Tyr Arg
595 600 605
Lys Phe Phe Lys Ala Val Ala Glu Gly Lys Leu Lys Pro Phe Lys Arg
610 615 620
Pro Lys Glu Phe Arg Asp Phe Leu Lys Lys Glu Tyr Pro Lys Leu Arg

625 630 635 640
Met Ala Asp Ile Pro Lys Lys Leu Gln Leu Phe Leu Cys Ser His Gly
645 650 655
Leu Cys Tyr Asn Asn Lys Pro Glu Thr Val Tyr Glu Arg Leu Asp Arg
660 665 670
Leu Thr Leu Gln His Leu Glu Glu Arg Glu Leu His Ile Gln Asn Arg
675 680 685
Leu Glu His Tyr Gln Lys Asp Arg Asp Met Ile Gly Asn Lys Asp Asn

690 695 700
Gln Tyr Gly Lys Lys Ser Phe Ser Asp Val Arg His Gly Ala Leu Ala
705 710 715 720
Arg Tyr Leu Ala Gln Ser Met Met Glu Trp Gln Pro Thr Lys Leu Lys
725 730 735
Asp Lys Glu Lys Gly His Asp Lys Leu Thr Gly Leu Asn Tyr Asn Val
740 745 750
Leu Thr Ala Tyr Leu Ala Thr Tyr Gly His Pro Gln Val Pro Glu Glu

755 760 765
Gly Phe Thr Pro Arg Thr Leu Glu Gln Val Leu Ile Asn Ala His Leu

770 775 780

Ile Gly Gly Ser Asn Pro His Pro Phe Ile Asn Lys Val Leu Ala Leu

785 790 795 800

Gly Asn Arg Asn Ile Glu Glu Leu Tyr Leu His Tyr Leu Glu Glu Glu

805 810 815

Leu Lys His Ile Arg Ser Arg Ile Gln Ser Leu Ser Ser Asn Pro Ser

820 825 830

Asp Lys Ala Leu Ser Ala Leu Pro Phe Ile His His Asp Arg Met Arg

835 840 845

Tyr His Glu Arg Thr Ser Glu Glu Met Met Ala Leu Ala Ala Arg Tyr

850 855 860

Thr Thr Ile Gln Leu Pro Asp Gly Leu Phe Thr Pro Tyr Ile Leu Glu

865 870 875 880

Ile Leu Gln Lys His Tyr Thr Glu Asn Ser Asp Leu Gln Asn Ala Leu

885 890 895

Ser Gln Asp Val Pro Val Lys Leu Asn Pro Thr Cys Asn Ala Ala Tyr

900 905 910

Leu Ile Thr Leu Phe Tyr Gln Thr Val Leu Lys Asp Asn Ala Gln Pro

915 920 925

Phe Tyr Leu Ser Asp Lys Thr Tyr Thr Arg Asn Lys Asp Gly Glu Lys

930 935 940

Ala Glu Ser Phe Ser Phe Lys Arg Ala Tyr Glu Leu Phe Ser Val Leu

945 950 955 960

Asn Asn Asn Lys Lys Asp Thr Phe Pro Phe Glu Met Ile Pro Leu Phe

965 970 975

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

980 985 990

Gly Asp Gly Asn Pro Val Pro Glu Val Gly Glu Lys Gly Lys Pro Ala

995 1000 1005

Thr Asp Ser Gln Gly Asn Thr Ile Trp Lys Arg Arg Ile Tyr Ser

1010 1015 1020

Glu Val Asp Asp Tyr Ala Glu Lys Leu Thr Asp Arg Asp Met Lys
1025 1030 1035
Ile Ser Phe Lys Gly Glu Trp Glu Lys Leu Pro Arg Trp Lys Gln
1040 1045 1050
Asp Lys Ile Ile Lys Arg Arg Asp Glu Thr Arg Arg Gln Met Arg
1055 1060 1065
Asp Glu Leu Leu Gln Arg Met Pro Arg Tyr Ile Arg Asp Ile Lys

1070 1075 1080
Asp Asn Glu Arg Thr Leu Arg Arg Tyr Lys Thr Gln Asp Met Val
1085 1090 1095
Leu Phe Leu Leu Ala Glu Lys Met Phe Thr Asn Ile Ile Ser Glu
1100 1105 1110
Gln Ser Ser Glu Phe Asn Trp Lys Gln Met Arg Leu Ser Lys Val
1115 1120 1125
Cys Asn Glu Ala Phe Leu Arg Gln Thr Leu Thr Phe Arg Val Pro

1130 1135 1140
Val Thr Val Gly Glu Thr Thr Ile Tyr Val Glu Gln Glu Asn Met
1145 1150 1155
Ser Leu Lys Asn Tyr Gly Glu Phe Tyr Arg Phe Leu Thr Asp Asp
1160 1165 1170
Arg Leu Met Ser Leu Leu Asn Asn Ile Val Glu Thr Leu Lys Pro
1175 1180 1185
Asn Glu Asn Gly Asp Leu Val Ile Arg His Thr Asp Leu Met Ser

1190 1195 1200
Glu Leu Ala Ala Tyr Asp Gln Tyr Arg Ser Thr Ile Phe Met Leu
1205 1210 1215
Ile Gln Ser Ile Glu Asn Leu Ile Ile Thr Asn Asn Ala Val Leu
1220 1225 1230
Asp Asp Pro Asp Ala Asp Gly Phe Trp Val Arg Glu Asp Leu Pro
1235 1240 1245
Lys Arg Asn Asn Phe Ala Ser Leu Leu Glu Leu Ile Asn Gln Leu

1250 1255 1260

Asn Asn Val Glu Leu Thr Asp Asp Glu Arg Lys Leu Leu Val Ala

1265 1270 1275

Ile Arg Asn Ala Phe Ser His Asn Ser Tyr Asn Ile Asp Phe Ser

1280 1285 1290

Leu Ile Lys Asp Val Lys His Leu Pro Glu Val Ala Lys Gly Ile

1295 1300 1305

Leu Gln His Leu Gln Ser Met Leu Gly Val Glu Ile Thr Lys

1310 1315 1320

<210> 74

<211> 1322

<212> PRT

<213> Prevotella species

<400> 74

Met Ser Lys Glu Cys Lys Lys Gln Arg Gln Glu Lys Lys Arg Arg Leu

1 5 10 15

Gln Lys Ala Asn Phe Ser Ile Ser Leu Thr Gly Lys His Val Phe Gly

20 25 30

Ala Tyr Phe Asn Met Ala Arg Thr Asn Phe Val Lys Thr Ile Asn Tyr

35 40 45

Ile Leu Pro Ile Ala Gly Val Arg Gly Asn Tyr Ser Glu Asn Gln Ile

50 55 60

Asn Lys Met Leu His Ala Leu Phe Leu Ile Gln Ala Gly Arg Asn Glu

65 70 75 80

Glu Leu Thr Thr Glu Gln Lys Gln Trp Glu Lys Lys Leu Arg Leu Asn

85 90 95

Pro Glu Gln Gln Thr Lys Phe Gln Lys Leu Leu Phe Lys His Phe Pro

100 105 110

Val Leu Gly Pro Met Met Ala Asp Val Ala Asp His Lys Ala Tyr Leu

115 120 125

Asn Lys Lys Lys Ser Thr Val Gln Thr Glu Asp Glu Thr Phe Ala Met

130 135 140

Leu Lys Gly Val Ser Leu Ala Asp Cys Leu Asp Ile Ile Cys Leu Met
145 150 155 160
Ala Asp Thr Leu Thr Glu Cys Arg Asn Phe Tyr Thr His Lys Asp Pro
165 170 175

Tyr Asn Lys Pro Ser Gln Leu Ala Asp Gln Tyr Leu His Gln Glu Met
180 185 190
Ile Ala Lys Lys Leu Asp Lys Val Val Val Ala Ser Arg Arg Ile Leu
195 200 205
Lys Asp Arg Glu Gly Leu Ser Val Asn Glu Val Glu Phe Leu Thr Gly
210 215 220
Ile Asp His Leu His Gln Glu Val Leu Lys Asp Glu Phe Gly Asn Ala
225 230 235 240

Lys Val Lys Asp Gly Lys Val Met Lys Thr Phe Val Glu Tyr Asp Asp
245 250 255
Phe Tyr Phe Lys Ile Ser Gly Lys Arg Leu Val Asn Gly Tyr Thr Val
260 265 270
Thr Thr Lys Asp Asp Lys Pro Val Asn Val Asn Thr Met Leu Pro Ala
275 280 285
Leu Ser Asp Phe Gly Leu Leu Tyr Phe Cys Val Leu Phe Leu Ser Lys
290 295 300

Pro Tyr Ala Lys Leu Phe Ile Asp Glu Val Arg Leu Phe Glu Tyr Ser
305 310 315 320
Pro Phe Asp Asp Lys Glu Asn Met Ile Met Ser Glu Met Leu Ser Ile
325 330 335
Tyr Arg Ile Arg Thr Pro Arg Leu His Lys Ile Asp Ser His Asp Ser
340 345 350
Lys Ala Thr Leu Ala Met Asp Ile Phe Gly Glu Leu Arg Arg Cys Pro
355 360 365

Met Glu Leu Tyr Asn Leu Leu Asp Lys Asn Ala Gly Gln Pro Phe Phe
370 375 380
His Asp Glu Val Lys His Pro Asn Ser His Thr Pro Asp Val Ser Lys

385 390 395 400

Arg Leu Arg Tyr Asp Asp Arg Phe Pro Thr Leu Ala Leu Arg Tyr Ile

405 410 415

Asp Glu Thr Glu Leu Phe Lys Arg Ile Arg Phe Gln Leu Gln Leu Gly

420 425 430

Ser Phe Arg Tyr Lys Phe Tyr Asp Lys Glu Asn Cys Ile Asp Gly Arg

435 440 445

Val Arg Val Arg Arg Ile Gln Lys Glu Ile Asn Gly Tyr Gly Arg Met

450 455 460

Gln Glu Val Ala Asp Lys Arg Met Asp Lys Trp Gly Asp Leu Ile Gln

465 470 475 480

Lys Arg Glu Glu Arg Ser Val Lys Leu Glu His Glu Glu Leu Tyr Ile

485 490 495

Asn Leu Asp Gln Phe Leu Glu Asp Thr Ala Asp Ser Thr Pro Tyr Val

500 505 510

Thr Asp Arg Arg Pro Ala Tyr Asn Ile His Ala Asn Arg Ile Gly Leu

515 520 525

Tyr Trp Glu Asp Ser Gln Asn Pro Lys Gln Tyr Lys Val Phe Asp Glu

530 535 540

Asn Gly Met Tyr Ile Pro Glu Leu Val Val Thr Glu Asp Lys Lys Ala

545 550 555 560

Pro Ile Lys Met Pro Ala Pro Arg Cys Ala Leu Ser Val Tyr Asp Leu

565 570 575

Pro Ala Met Leu Phe Tyr Glu Tyr Leu Arg Glu Gln Gln Asp Asn Glu

580 585 590

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

595 600 605

Lys Phe Phe Lys Ala Val Ala Glu Gly Lys Leu Lys Pro Phe Lys Arg

610 615 620

Pro Lys Glu Phe Arg Asp Phe Leu Lys Lys Glu Tyr Pro Lys Leu Arg

625 630 635 640

Met Ala Asp Ile Pro Lys Lys Leu Gln Leu Phe Leu Cys Ser His Gly
645 650 655
Leu Cys Tyr Asn Asn Lys Pro Glu Thr Val Tyr Glu Arg Leu Asp Arg
660 665 670
Leu Thr Leu Gln His Leu Glu Glu Arg Glu Leu His Ile Gln Asn Arg
675 680 685

Leu Glu His Tyr Gln Lys Asp Arg Asp Met Ile Gly Asn Lys Asp Asn
690 695 700
Gln Tyr Gly Lys Lys Ser Phe Ser Asp Val Arg His Gly Ala Leu Ala
705 710 715 720
Arg Tyr Leu Ala Gln Ser Met Met Glu Trp Gln Pro Thr Lys Leu Lys
725 730 735
Asp Lys Glu Lys Gly His Asp Lys Leu Thr Gly Leu Asn Tyr Asn Val
740 745 750

Leu Thr Ala Tyr Leu Ala Thr Tyr Gly His Pro Gln Val Pro Glu Glu
755 760 765
Gly Phe Thr Pro Arg Thr Leu Glu Gln Val Leu Ile Asn Ala His Leu
770 775 780
Ile Gly Gly Ser Asn Pro His Pro Phe Ile Asn Lys Val Leu Ala Leu
785 790 795 800
Gly Asn Arg Asn Ile Glu Glu Leu Tyr Leu His Tyr Leu Glu Glu Glu
805 810 815

Leu Lys His Ile Arg Ser Arg Ile Gln Ser Leu Ser Ser Asn Pro Ser
820 825 830
Asp Lys Ala Leu Ser Ala Leu Pro Phe Ile His His Asp Arg Met Arg
835 840 845
Tyr His Glu Arg Thr Ser Glu Glu Met Met Ala Leu Ala Ala Arg Tyr
850 855 860
Thr Thr Ile Gln Leu Pro Asp Gly Leu Phe Thr Pro Tyr Ile Leu Glu
865 870 875 880

Ile Leu Gln Lys His Tyr Thr Glu Asn Ser Asp Leu Gln Asn Ala Leu

885 890 895

Ser Gln Asp Val Pro Val Lys Leu Asn Pro Thr Cys Asn Ala Ala Tyr

900 905 910

Leu Ile Thr Leu Phe Tyr Gln Thr Val Leu Lys Asp Asn Ala Gln Pro

915 920 925

Phe Tyr Leu Ser Asp Lys Thr Tyr Thr Arg Asn Lys Asp Gly Glu Lys

930 935 940

Ala Glu Ser Phe Ser Phe Lys Arg Ala Tyr Glu Leu Phe Ser Val Leu

945 950 955 960

Asn Asn Asn Lys Lys Asp Thr Phe Pro Phe Glu Met Ile Pro Leu Phe

965 970 975

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

980 985 990

Gly Asp Gly Asn Pro Val Pro Glu Val Gly Glu Lys Gly Lys Pro Ala

995 1000 1005

Thr Asp Ser Gln Gly Asn Thr Ile Trp Lys Arg Arg Ile Tyr Ser

1010 1015 1020

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

1025 1030 1035

Ile Ser Phe Lys Gly Glu Trp Glu Lys Leu Pro Arg Trp Lys Gln

1040 1045 1050

Asp Lys Ile Ile Lys Arg Arg Asp Glu Thr Arg Arg Gln Met Arg

1055 1060 1065

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

1070 1075 1080

Asp Asn Glu Arg Thr Leu Arg Arg Tyr Lys Thr Gln Asp Met Val

1085 1090 1095

Leu Phe Leu Leu Ala Glu Lys Met Phe Thr Asn Ile Ile Ser Glu

1100 1105 1110

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

1115 1120 1125

Cys Asn Glu Ala Phe Leu Arg Gln Thr Leu Thr Phe Arg Val Pro
1130 1135 1140
Val Thr Val Gly Glu Thr Thr Ile Tyr Val Glu Gln Glu Asn Met
1145 1150 1155
Ser Leu Lys Asn Tyr Gly Glu Phe Tyr Arg Phe Leu Thr Asp Asp
1160 1165 1170
Arg Leu Met Ser Leu Leu Asn Asn Ile Val Glu Thr Leu Lys Pro
1175 1180 1185

Asn Glu Asn Gly Asp Leu Val Ile Arg His Thr Asp Leu Met Ser
1190 1195 1200
Glu Leu Ala Ala Tyr Asp Gln Tyr Arg Ser Thr Ile Phe Met Leu
1205 1210 1215
Ile Gln Ser Ile Glu Asn Leu Ile Ile Thr Asn Asn Ala Val Leu
1220 1225 1230
Asp Asp Pro Asp Ala Asp Gly Phe Trp Val Arg Glu Asp Leu Pro
1235 1240 1245

Lys Arg Asn Asn Phe Ala Ser Leu Leu Glu Leu Ile Asn Gln Leu
1250 1255 1260
Asn Asn Val Glu Leu Thr Asp Asp Glu Arg Lys Leu Leu Val Ala
1265 1270 1275
Ile Arg Asn Ala Phe Ser His Asn Ser Tyr Asn Ile Asp Phe Ser
1280 1285 1290
Leu Ile Lys Asp Val Lys His Leu Pro Glu Val Ala Lys Gly Ile
1295 1300 1305

Leu Gln His Leu Gln Ser Met Leu Gly Val Glu Ile Thr Lys
1310 1315 1320
<210> 75
<211> 1095
<212> PRT
<213> Riemerella anatipestifer
<400> 75
Met Glu Lys Pro Leu Leu Pro Asn Val Tyr Thr Leu Lys His Lys Phe

1 5 10 15

Phe Trp Gly Ala Phe Leu Asn Ile Ala Arg His Asn Ala Phe Ile Thr

20 25 30

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

35 40 45

Asp Lys Ile Val Asp Val Val Cys Glu Thr Trp Asn Asn Ile Leu Asn

50 55 60

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

65 70 75 80

Lys His Phe Pro Phe Leu Thr Ala Met Cys Tyr His Pro Pro Lys Lys

85 90 95

Glu Gly Lys Lys Lys Gly His Gln Lys Glu Gln Gln Lys Glu Lys Glu

100 105 110

Ser Glu Ala Gln Ser Gln Ala Glu Ala Leu Asn Pro Ser Lys Leu Ile

115 120 125

Glu Ala Leu Glu Ile Leu Val Asn Gln Leu His Ser Leu Arg Asn Tyr

130 135 140

Tyr Ser His Tyr Lys His Lys Lys Pro Asp Ala Glu Lys Asp Ile Phe

145 150 155 160

Lys His Leu Tyr Lys Ala Phe Asp Ala Ser Leu Arg Met Val Lys Glu

165 170 175

Asp Tyr Lys Ala His Phe Thr Val Asn Leu Thr Arg Asp Phe Ala His

180 185 190

Leu Asn Arg Lys Gly Lys Asn Lys Gln Asp Asn Pro Asp Phe Asn Arg

195 200 205

Tyr Arg Phe Glu Lys Asp Gly Phe Phe Thr Glu Ser Gly Leu Leu Phe

210 215 220

Phe Thr Asn Leu Phe Leu Asp Lys Arg Asp Ala Tyr Trp Met Leu Lys

225 230 235 240

Lys Val Ser Gly Phe Lys Ala Ser His Lys Gln Arg Glu Lys Met Thr

245 250 255

Thr Glu Val Phe Cys Arg Ser Arg Ile Leu Leu Pro Lys Leu Arg Leu
260 265 270
Glu Ser Arg Tyr Asp His Asn Gln Met Leu Leu Asp Met Leu Ser Glu
275 280 285
Leu Ser Arg Cys Pro Lys Leu Leu Tyr Glu Lys Leu Ser Glu Glu Asn

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

355 360 365
Asp Lys Lys Ile Gly Asp Glu Gln Glu Lys Arg His Leu Thr Arg Thr
370 375 380
Leu Leu Ser Phe Gly Arg Leu Gln Asp Phe Thr Glu Ile Asn Arg Pro
385 390 395 400
Gln Glu Trp Lys Ala Leu Thr Lys Asp Leu Asp Tyr Lys Glu Thr Ser
405 410 415
Asn Gln Pro Phe Ile Ser Lys Thr Thr Pro His Tyr His Ile Thr Asp

420 425 430
Asn Lys Ile Gly Phe Arg Leu Gly Thr Ser Lys Glu Leu Tyr Pro Ser
435 440 445
Leu Glu Ile Lys Asp Gly Ala Asn Arg Ile Ala Lys Tyr Pro Tyr Asn
450 455 460
Ser Gly Phe Val Ala His Ala Phe Ile Ser Val His Glu Leu Leu Pro
465 470 475 480
Leu Met Phe Tyr Gln His Leu Thr Gly Lys Ser Glu Asp Leu Leu Lys

485 490 495
Glu Thr Val Arg His Ile Gln Arg Ile Tyr Lys Asp Phe Glu Glu Glu

500 505 510

Arg Ile Asn Thr Ile Glu Asp Leu Glu Lys Ala Asn Gln Gly Arg Leu

515 520 525

Pro Leu Gly Ala Phe Pro Lys Gln Met Leu Gly Leu Leu Gln Asn Lys

530 535 540

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

545 550 555 560

Ala Glu Thr Lys Leu Leu Ser His Arg Leu Asn Thr Lys Leu Lys Ser

565 570 575

Ser Pro Lys Leu Gly Lys Arg Arg Glu Lys Leu Ile Lys Thr Gly Val

580 585 590

Leu Ala Asp Trp Leu Val Lys Asp Phe Met Arg Phe Gln Pro Val Ala

595 600 605

Tyr Asp Ala Gln Asn Gln Pro Ile Lys Ser Ser Lys Ala Asn Ser Thr

610 615 620

Glu Phe Trp Phe Ile Arg Arg Ala Leu Ala Leu Tyr Gly Gly Glu Lys

625 630 635 640

Asn Arg Leu Glu Gly Tyr Phe Lys Gln Thr Asn Leu Ile Gly Asn Thr

645 650 655

Asn Pro His Pro Phe Leu Asn Lys Phe Asn Trp Lys Ala Cys Arg Asn

660 665 670

Leu Val Asp Phe Tyr Gln Gln Tyr Leu Glu Gln Arg Glu Lys Phe Leu

675 680 685

Glu Ala Ile Lys Asn Gln Pro Trp Glu Pro Tyr Gln Tyr Cys Leu Leu

690 695 700

Leu Lys Ile Pro Lys Glu Asn Arg Lys Asn Leu Val Lys Gly Trp Glu

705 710 715 720

Gln Gly Gly Ile Ser Leu Pro Arg Gly Leu Phe Thr Glu Ala Ile Arg

725 730 735

Glu Thr Leu Ser Glu Asp Leu Met Leu Ser Lys Pro Ile Arg Lys Glu

740 745 750

Ile Lys Lys His Gly Arg Val Gly Phe Ile Ser Arg Ala Ile Thr Leu
755 760 765

Tyr Phe Lys Glu Lys Tyr Gln Asp Lys His Gln Ser Phe Tyr Asn Leu
770 775 780

Ser Tyr Lys Leu Glu Ala Lys Ala Pro Leu Leu Lys Arg Glu Glu His
785 790 795 800

Tyr Glu Tyr Trp Gln Gln Asn Lys Pro Gln Ser Pro Thr Glu Ser Gln

805 810 815

Arg Leu Glu Leu His Thr Ser Asp Arg Trp Lys Asp Tyr Leu Leu Tyr
820 825 830

Lys Arg Trp Gln His Leu Glu Lys Lys Leu Arg Leu Tyr Arg Asn Gln
835 840 845

Asp Val Met Leu Trp Leu Met Thr Leu Glu Leu Thr Lys Asn His Phe
850 855 860

Lys Glu Leu Asn Leu Asn Tyr His Gln Leu Lys Leu Glu Asn Leu Ala

865 870 875 880

Val Asn Val Gln Glu Ala Asp Ala Lys Leu Asn Pro Leu Asn Gln Thr
885 890 895

Leu Pro Met Val Leu Pro Val Lys Val Tyr Pro Ala Thr Ala Phe Gly
900 905 910

Glu Val Gln Tyr His Lys Thr Pro Ile Arg Thr Val Tyr Ile Arg Glu
915 920 925

Glu His Thr Lys Ala Leu Lys Met Gly Asn Phe Lys Ala Leu Val Lys

930 935 940

Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe Ile Lys Glu Glu Asn Asp
945 950 955 960

Thr Gln Lys His Pro Ile Ser Gln Leu Arg Leu Arg Arg Glu Leu Glu
965 970 975

Ile Tyr Gln Ser Leu Arg Val Asp Ala Phe Lys Glu Thr Leu Ser Leu
980 985 990

Glu Glu Lys Leu Leu Asn Lys His Thr Ser Leu Ser Ser Leu Glu Asn

995 1000 1005

Glu Phe Arg Ala Leu Leu Glu Glu Trp Lys Lys Glu Tyr Ala Ala

1010 1015 1020

Ser Ser Met Val Thr Asp Glu His Ile Ala Phe Ile Ala Ser Val

1025 1030 1035

Arg Asn Ala Phe Cys His Asn Gln Tyr Pro Phe Tyr Lys Glu Ala

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

Leu Arg Glu Tyr Cys Glu Ile Val Lys Ser Gln Ile

1085 1090 1095

<210> 76

<211> 1124

<212> PRT

<213> Prevotella aurantiaca

<400> 76

Met Glu Asp Asp Lys Lys Thr Thr Gly Ser Ile Ser Tyr Glu Leu Lys

1 5 10 15

Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val

20 25 30

Tyr Ile Thr Ile Asn His Ile Asn Lys Leu Leu Glu Ile Arg Glu Ile

35 40 45

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

50 55 60

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

65 70 75 80

His Phe Pro Phe Leu Glu Thr Ala Ile Tyr Thr Lys Asn Lys Glu Asp

85 90 95

Lys Lys Glu Val Lys Gln Glu Lys Gln Ala Glu Ala Gln Ser Leu Glu

100 105 110

Ser Leu Lys Asp Cys Leu Phe Leu Phe Leu Asp Lys Leu Gln Glu Ala
115 120 125
Arg Asn Tyr Tyr Ser His Tyr Lys Tyr Ser Glu Phe Ser Lys Glu Pro
130 135 140

Glu Phe Glu Glu Gly Leu Leu Glu Lys Met Tyr Asn Ile Phe Gly Asn
145 150 155 160
Asn Ile Gln Leu Val Ile Asn Asp Tyr Gln His Asn Lys Asp Ile Asn
165 170 175
Pro Asp Glu Asp Phe Lys His Leu Asp Arg Lys Gly Gln Phe Lys Tyr
180 185 190
Ser Phe Ala Asp Asn Glu Gly Asn Ile Thr Glu Ser Gly Leu Leu Phe
195 200 205

Phe Val Ser Leu Phe Leu Glu Lys Lys Asp Ala Ile Trp Met Gln Gln
210 215 220
Lys Leu Asn Gly Phe Lys Asp Asn Leu Glu Asn Lys Lys Lys Met Thr
225 230 235 240
His Glu Val Phe Cys Arg Ser Arg Ile Leu Met Pro Lys Leu Arg Leu
245 250 255
Glu Ser Thr Gln Thr Gln Asp Trp Ile Leu Leu Asp Met Leu Asn Glu
260 265 270

Leu Ile Arg Cys Pro Lys Ser Leu Tyr Glu Arg Leu Gln Gly Asp Asp
275 280 285
Arg Glu Lys Phe Lys Val Pro Phe Asp Pro Ala Asp Glu Asp Tyr Asn
290 295 300
Ala Glu Gln Glu Pro Phe Lys Asn Thr Leu Ile Arg His Gln Asp Arg
305 310 315 320
Phe Pro Tyr Phe Val Leu Arg Tyr Phe Asp Tyr Asn Glu Ile Phe Lys
325 330 335

Asn Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ser Ile Tyr
340 345 350
Lys Lys Leu Ile Gly Gly Gln Lys Glu Asp Arg His Leu Thr His Lys

355 360 365

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

370 375 380

Asp Glu Trp Lys Ala Ile Val Lys Asp Leu Asp Thr Tyr Glu Thr Ser

385 390 395 400

Asn Lys Arg Tyr Ile Ser Glu Thr Thr Pro His Tyr His Leu Glu Asn

405 410 415

Gln Lys Ile Gly Ile Arg Phe Arg Asn Gly Asn Lys Glu Ile Trp Pro

420 425 430

Ser Leu Lys Thr Asn Asp Glu Asn Asn Glu Lys Ser Lys Tyr Lys Leu

435 440 445

Asp Lys Gln Tyr Gln Ala Glu Ala Phe Leu Ser Val His Glu Leu Leu

450 455 460

Pro Met Met Phe Tyr Tyr Leu Leu Leu Lys Lys Glu Lys Pro Asn Asn

465 470 475 480

Asp Glu Ile Asn Ala Ser Ile Val Glu Gly Phe Ile Lys Arg Glu Ile

485 490 495

Arg Asn Ile Phe Lys Leu Tyr Asp Ala Phe Ala Asn Gly Glu Ile Asn

500 505 510

Asn Ile Asp Asp Leu Glu Lys Tyr Cys Ala Asp Lys Gly Ile Pro Lys

515 520 525

Arg His Leu Pro Lys Gln Met Val Ala Ile Leu Tyr Asp Glu His Lys

530 535 540

Asp Met Val Lys Glu Ala Lys Arg Lys Gln Lys Glu Met Val Lys Asp

545 550 555 560

Thr Lys Lys Leu Leu Ala Thr Leu Glu Lys Gln Thr Gln Lys Glu Lys

565 570 575

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

580 585 590

Arg Trp Leu Val Asn Asp Met Met Arg Phe Gln Pro Val Gln Lys Asp

595 600 605

Asn Glu Gly Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr
610 615 620
Gln Met Leu Gln Arg Ser Leu Ala Leu Tyr Asn Asn Glu Glu Lys Pro
625 630 635 640
Thr Arg Tyr Phe Arg Gln Val Asn Leu Ile Glu Ser Asn Asn Pro His
645 650 655

Pro Phe Leu Lys Trp Thr Lys Trp Glu Glu Cys Asn Asn Ile Leu Thr
660 665 670
Phe Tyr Tyr Ser Tyr Leu Thr Lys Lys Ile Glu Phe Leu Asn Lys Leu
675 680 685
Lys Pro Glu Asp Trp Lys Lys Asn Gln Tyr Phe Leu Lys Leu Lys Glu
690 695 700
Pro Lys Thr Asn Arg Glu Thr Leu Val Gln Gly Trp Lys Asn Gly Phe
705 710 715 720

Asn Leu Pro Arg Gly Ile Phe Thr Glu Pro Ile Arg Glu Trp Phe Lys
725 730 735
Arg His Gln Asn Asn Ser Lys Glu Tyr Glu Lys Val Glu Ala Leu Asp
740 745 750
Arg Val Gly Leu Val Thr Lys Val Ile Pro Leu Phe Phe Lys Glu Glu
755 760 765
Tyr Phe Lys Asp Lys Glu Glu Asn Phe Lys Glu Asp Thr Gln Lys Glu
770 775 780

Ile Asn Asp Cys Val Gln Pro Phe Tyr Asn Phe Pro Tyr Asn Val Gly
785 790 795 800
Asn Ile His Lys Pro Lys Glu Lys Asp Phe Leu His Arg Glu Glu Arg
805 810 815
Ile Glu Leu Trp Asp Lys Lys Lys Asp Lys Phe Lys Gly Tyr Lys Glu
820 825 830
Lys Ile Lys Ser Lys Lys Leu Thr Glu Lys Asp Lys Glu Glu Phe Arg
835 840 845

Ser Tyr Leu Glu Phe Gln Ser Trp Asn Lys Phe Glu Arg Glu Leu Arg

```
    850                 855                 860
Leu Val Arg Asn Gln Asp Ile Val Thr Trp Leu Leu Cys Lys Glu Leu
865                 870                 875                 880
Ile Asp Lys Leu Lys Ile Asp Glu Leu Asn Ile Glu Glu Leu Lys Lys
                885                 890                 895
Leu Arg Leu Asn Asn Ile Asp Thr Asp Thr Ala Lys Lys Glu Lys Asn
            900                 905                 910

Asn Ile Leu Asn Arg Val Met Pro Met Glu Leu Pro Val Thr Val Tyr
        915                 920                 925
Glu Ile Asp Asp Ser His Lys Ile Val Lys Asp Lys Pro Leu His Thr
    930                 935                 940
Ile Tyr Ile Lys Glu Ala Glu Thr Lys Leu Leu Lys Gln Gly Asn Phe
945                 950                 955                 960
Lys Ala Leu Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe Val
                965                 970                 975

Lys Thr Asn Ser Glu Ala Glu Ser Lys Arg Asn Pro Ile Ser Lys Leu
            980                 985                 990
Arg Val Glu Tyr Glu Leu Gly Glu  Tyr Gln Glu Ala Arg  Ile Glu Ile
        995                 1000                1005
Ile Gln  Asp Met Leu Ala Leu  Glu Glu Lys Leu Ile  Asn Lys Tyr
    1010                1015                1020
Lys Asp  Leu Pro Thr Asn Lys  Phe Ser Glu Met Leu  Asn Ser Trp
    1025                1030                1035

Leu Glu  Gly Lys Asp Glu Ala  Asp Lys Ala Arg Phe  Gln Asn Asp
    1040                1045                1050
Val Asp  Phe Leu Ile Ala Val  Arg Asn Ala Phe Ser  His Asn Gln
    1055                1060                1065
Tyr Pro  Met His Asn Lys Ile  Glu Phe Ala Asn Ile  Lys Pro Phe
    1070                1075                1080
Ser Leu  Tyr Thr Ala Asn Asn  Ser Glu Glu Lys Gly  Leu Gly Ile
    1085                1090                1095
```

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

1100 1105 1110

Lys Lys Ile Glu Lys Pro Ile Glu Thr Lys Glu

1115 1120

<210> 77

<211> 1151

<212> PRT

<213> Prevotella saccharolytica

<400> 77

Met Glu Asp Lys Pro Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His

1 5 10 15

Asn Val Tyr Leu Thr Val Asn His Ile Asn Lys Leu Leu Asp Leu Glu

20 25 30

Lys Leu Tyr Asp Glu Gly Lys His Lys Glu Ile Phe Glu Arg Glu Asp

35 40 45

Ile Phe Asn Ile Ser Asp Asp Val Met Asn Asp Ala Asn Ser Asn Gly

50 55 60

Lys Lys Arg Lys Leu Asp Ile Lys Lys Ile Trp Asp Asp Leu Asp Thr

65 70 75 80

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

85 90 95

Pro Phe Ile Gln Pro Ala Ile Ile Gly Ala Gln Thr Lys Glu Arg Thr

100 105 110

Thr Ile Asp Lys Asp Lys Arg Ser Thr Ser Thr Ser Asn Asp Ser Leu

115 120 125

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

130 135 140

Val Lys Ser Met Phe Phe Arg Leu Leu Gln Ile Leu Glu Gln Leu Arg

145 150 155 160

Asn Tyr Tyr Ser His Val Lys His Ser Lys Ser Ala Thr Met Pro Asn

165 170 175

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

180 185 190

Val Asn Lys Val Lys Glu Asp Tyr Ser Ser Asn Ser Val Ile Asp Pro

195 200 205

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

210 215 220

Lys Pro Cys Arg Tyr Pro Phe Thr Ser Lys Asp Gly Ser Ile Asn Ala

225 230 235 240

Phe Gly Leu Leu Phe Phe Val Ser Leu Phe Leu Glu Lys Gln Asp Ser

245 250 255

Ile Trp Met Gln Lys Lys Ile Pro Gly Phe Lys Lys Ala Ser Glu Asn

260 265 270

Tyr Met Lys Met Thr Asn Glu Val Phe Cys Arg Asn His Ile Leu Leu

275 280 285

Pro Lys Ile Arg Leu Glu Thr Val Tyr Asp Lys Asp Trp Met Leu Leu

290 295 300

Asp Met Leu Asn Glu Val Val Arg Cys Pro Leu Ser Leu Tyr Lys Arg

305 310 315 320

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

325 330 335

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

340 345 350

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

355 360 365

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

370 375 380

Tyr His Phe Ala Ile Cys Lys Lys Gln Ile Gly Asp Lys Lys Glu Val

385 390 395 400

His His Leu Ile Arg Thr Leu Tyr Gly Phe Ser Arg Leu Gln Asn Phe

405 410 415

Thr Gln Asn Thr Arg Pro Glu Glu Trp Asn Thr Leu Val Lys Thr Thr

420 425 430

Glu Pro Ser Ser Gly Asn Asp Gly Lys Thr Val Gln Gly Val Pro Leu
435 440 445
Pro Tyr Ile Ser Tyr Thr Ile Pro His Tyr Gln Ile Glu Asn Glu Lys
450 455 460
Ile Gly Ile Lys Ile Phe Asp Gly Asp Thr Ala Val Asp Thr Asp Ile

465 470 475 480
Trp Pro Ser Val Ser Thr Glu Lys Gln Leu Asn Lys Pro Asp Lys Tyr
485 490 495
Thr Leu Thr Pro Gly Phe Lys Ala Asp Val Phe Leu Ser Val His Glu
500 505 510
Leu Leu Pro Met Met Phe Tyr Tyr Gln Leu Leu Leu Cys Glu Gly Met
515 520 525
Leu Lys Thr Asp Ala Gly Asn Ala Val Glu Lys Val Leu Ile Asp Thr

530 535 540
Arg Asn Ala Ile Phe Asn Leu Tyr Asp Ala Phe Val Gln Glu Lys Ile
545 550 555 560
Asn Thr Ile Thr Asp Leu Glu Asn Tyr Leu Gln Asp Lys Pro Ile Leu
565 570 575
Ile Gly His Leu Pro Lys Gln Met Ile Asp Leu Leu Lys Gly His Gln
580 585 590
Arg Asp Met Leu Lys Ala Val Glu Gln Lys Lys Ala Met Leu Ile Lys

595 600 605
Asp Thr Glu Arg Arg Leu Lys Leu Leu Asp Lys Gln Leu Lys Gln Glu
610 615 620
Thr Asp Val Ala Ala Lys Asn Thr Gly Thr Leu Leu Lys Asn Gly Gln
625 630 635 640
Ile Ala Asp Trp Leu Val Asn Asp Met Met Arg Phe Gln Pro Val Lys
645 650 655
Arg Asp Lys Glu Gly Asn Pro Ile Asn Cys Ser Lys Ala Asn Ser Thr

660 665 670
Glu Tyr Gln Met Leu Gln Arg Ala Phe Ala Phe Tyr Ala Thr Asp Ser

675 680 685

Cys Arg Leu Ser Arg Tyr Phe Thr Gln Leu His Leu Ile His Ser Asp

690 695 700

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

705 710 715 720

Leu Ile Ala Phe Tyr Ala Ala Tyr Leu Lys Ala Lys Leu Glu Phe Leu

725 730 735

Asn Glu Leu Gln Pro Gln Asn Trp Ala Ser Asp Asn Tyr Phe Leu Leu

740 745 750

Leu Arg Ala Pro Lys Asn Asp Arg Gln Lys Leu Ala Glu Gly Trp Lys

755 760 765

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

770 775 780

Trp Phe Asn Glu His Lys Thr Ile Val Asp Ile Ser Asp Cys Asp Ile

785 790 795 800

Phe Lys Asn Arg Val Gly Gln Val Ala Arg Leu Ile Pro Val Phe Phe

805 810 815

Asp Lys Lys Phe Lys Asp His Ser Gln Pro Phe Tyr Arg Tyr Asp Phe

820 825 830

Asn Val Gly Asn Val Ser Lys Pro Thr Glu Ala Asn Tyr Leu Ser Lys

835 840 845

Gly Lys Arg Glu Glu Leu Phe Lys Ser Tyr Gln Asn Lys Phe Lys Asn

850 855 860

Asn Ile Pro Ala Glu Lys Thr Lys Glu Tyr Arg Glu Tyr Lys Asn Phe

865 870 875 880

Ser Leu Trp Lys Lys Phe Glu Arg Glu Leu Arg Leu Ile Lys Asn Gln

885 890 895

Asp Ile Leu Ile Trp Leu Met Cys Lys Asn Leu Phe Asp Glu Lys Ile

900 905 910

Lys Pro Lys Lys Asp Ile Leu Glu Pro Arg Ile Ala Val Ser Tyr Ile

915 920 925

Lys Leu Asp Ser Leu Gln Thr Asn Thr Ser Thr Ala Gly Ser Leu Asn

930 935 940

Ala Leu Ala Lys Val Val Pro Met Thr Leu Ala Ile His Ile Asp Ser

945 950 955 960

Pro Lys Pro Lys Gly Lys Ala Gly Asn Asn Glu Lys Glu Asn Lys Glu

965 970 975

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

980 985 990

Asn Phe Lys Thr Leu Leu Ala Asp Arg Arg Ile Lys Gly Leu Phe Ser

995 1000 1005

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

1010 1015 1020

Lys Arg Arg Val Asp Leu Glu Leu Asp Leu Tyr Gln Thr Cys Arg

1025 1030 1035

Ile Asp Ile Phe Gln Gln Thr Leu Gly Leu Glu Ala Gln Leu Leu

1040 1045 1050

Asp Lys Tyr Ser Asp Leu Asn Thr Asp Asn Phe Tyr Gln Met Leu

1055 1060 1065

Ile Gly Trp Arg Lys Lys Glu Gly Ile Pro Arg Asn Ile Lys Glu

1070 1075 1080

Asp Thr Asp Phe Leu Lys Asp Val Arg Asn Ala Phe Ser His Asn

1085 1090 1095

Gln Tyr Pro Asp Ser Lys Lys Ile Ala Phe Arg Arg Ile Arg Lys

1100 1105 1110

Phe Asn Pro Lys Glu Leu Ile Leu Glu Glu Glu Glu Gly Leu Gly

1115 1120 1125

Ile Ala Thr Gln Met Tyr Lys Glu Val Glu Lys Val Val Asn Arg

1130 1135 1140

Ile Lys Arg Ile Glu Leu Phe Asp

1145 1150

<210> 78

<211> 1159

<212> PRT

<213> Myroides odoratimimus

<400> 78

Met Lys Asp Ile Leu Thr Thr Asp Thr Thr Glu Lys Gln Asn Arg Phe

1 5 10 15

Tyr Ser His Lys Ile Ala Asp Lys Tyr Phe Phe Gly Gly Tyr Phe Asn

20 25 30

Leu Ala Ser Asn Asn Ile Tyr Glu Val Phe Glu Glu Val Asn Lys Arg

35 40 45

Asn Thr Phe Gly Lys Leu Ala Lys Arg Asp Asn Gly Asn Leu Lys Asn

50 55 60

Tyr Ile Ile His Val Phe Lys Asp Glu Leu Ser Ile Ser Asp Phe Glu

65 70 75 80

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

85 90 95

Asp Lys Lys Ser Ile Lys Glu Arg Asn Arg Thr Ile Asp Leu Thr Leu

100 105 110

Ser Gln Arg Ile Arg Gln Phe Arg Glu Met Leu Ile Ser Leu Val Thr

115 120 125

Ala Val Asp Gln Leu Arg Asn Phe Tyr Thr His Tyr His His Ser Asp

130 135 140

Ile Val Ile Glu Asn Lys Val Leu Asp Phe Leu Asn Ser Ser Phe Val

145 150 155 160

Ser Thr Ala Leu His Val Lys Asp Lys Tyr Leu Lys Thr Asp Lys Thr

165 170 175

Lys Glu Phe Leu Lys Glu Thr Ile Ala Ala Glu Leu Asp Ile Leu Ile

180 185 190

Glu Ala Tyr Lys Lys Lys Gln Ile Glu Lys Lys Asn Thr Arg Phe Lys

195 200 205

Ala Asn Lys Arg Glu Asp Ile Leu Asn Ala Ile Tyr Asn Glu Ala Phe

210 215 220

Trp Ser Phe Ile Asn Asp Lys Asp Lys Asp Lys Asp Lys Glu Thr Val
225 230 235 240
Val Ala Lys Gly Ala Asp Ala Tyr Phe Glu Lys Asn His His Lys Ser
245 250 255
Asn Asp Pro Asp Phe Ala Leu Asn Ile Ser Glu Lys Gly Ile Val Tyr

260 265 270
Leu Leu Ser Phe Phe Leu Thr Asn Lys Glu Met Asp Ser Leu Lys Ala
275 280 285
Asn Leu Thr Gly Phe Lys Gly Lys Val Asp Arg Glu Ser Gly Asn Ser
290 295 300
Ile Lys Tyr Met Ala Thr Gln Arg Ile Tyr Ser Phe His Thr Tyr Arg
305 310 315 320
Gly Leu Lys Gln Lys Ile Arg Thr Ser Glu Glu Gly Val Lys Glu Thr

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

385 390 395 400
Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Phe Asp Phe Pro Thr
405 410 415
Leu Arg Phe Gln Val His Leu Gly Asp Tyr Val His Asp Arg Arg Thr
420 425 430
Lys Gln Leu Gly Lys Val Glu Ser Asp Arg Ile Ile Lys Glu Lys Val
435 440 445
Thr Val Phe Ala Arg Leu Lys Asp Ile Asn Ser Ala Lys Ala Ser Tyr

450 455 460
Phe His Ser Leu Glu Glu Gln Asp Lys Glu Glu Leu Asp Asn Lys Trp

465 470 475 480

Thr Leu Phe Pro Asn Pro Ser Tyr Asp Phe Pro Lys Glu His Thr Leu

485 490 495

Gln His Gln Gly Glu Gln Lys Asn Ala Gly Lys Ile Gly Ile Tyr Val

500 505 510

Lys Leu Arg Asp Thr Gln Tyr Lys Glu Lys Ala Ala Leu Glu Glu Ala

515 520 525

Arg Lys Ser Leu Asn Pro Lys Glu Arg Ser Ala Thr Lys Ala Ser Lys

530 535 540

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

545 550 555 560

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

565 570 575

Asn Asp Ile His Ser Met Leu Phe Ser Leu Leu Thr Asp Asn Ala Glu

580 585 590

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

595 600 605

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

610 615 620

Leu Lys Lys Tyr Lys Asp Asn Asp Leu Lys Glu Thr Asp Thr Asp Lys

625 630 635 640

Ile Thr Arg Asp Leu Ala Arg Asp Lys Glu Glu Ile Glu Lys Leu Ile

645 650 655

Leu Glu Gln Lys Gln Arg Ala Asp Asp Tyr Asn Tyr Thr Ser Ser Thr

660 665 670

Lys Phe Asn Ile Asp Lys Ser Arg Lys Arg Lys His Leu Leu Phe Asn

675 680 685

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

690 695 700

Phe Met Phe Lys Glu Ser Lys Ser Lys Trp Lys Gly Tyr Gln His Thr

705 710 715 720

Glu Leu Gln Lys Leu Phe Ala Tyr Phe Asp Thr Ser Lys Ser Asp Leu
725 730 735
Glu Leu Ile Leu Ser Asn Met Val Met Val Lys Asp Tyr Pro Ile Glu
740 745 750
Leu Ile Asp Leu Val Lys Lys Ser Arg Thr Leu Val Asp Phe Leu Asn
755 760 765
Lys Tyr Leu Glu Ala Arg Leu Glu Tyr Ile Glu Asn Val Ile Thr Arg

770 775 780
Val Lys Asn Ser Ile Gly Thr Pro Gln Phe Lys Thr Val Arg Lys Glu
785 790 795 800
Cys Phe Thr Phe Leu Lys Lys Ser Asn Tyr Thr Val Val Ser Leu Asp
805 810 815
Lys Gln Val Glu Arg Ile Leu Ser Met Pro Leu Phe Ile Glu Arg Gly
820 825 830
Phe Met Asp Asp Lys Pro Thr Met Leu Glu Gly Lys Ser Tyr Lys Gln

835 840 845
His Lys Glu Lys Phe Ala Asp Trp Phe Val His Tyr Lys Glu Asn Ser
850 855 860
Asn Tyr Gln Asn Phe Tyr Asp Thr Glu Val Tyr Glu Ile Thr Thr Glu
865 870 875 880
Asp Lys Arg Glu Lys Ala Lys Val Thr Lys Lys Ile Lys Gln Gln Gln
885 890 895
Lys Asn Asp Val Phe Thr Leu Met Met Val Asn Tyr Met Leu Glu Glu

900 905 910
Val Leu Lys Leu Ser Ser Asn Asp Arg Leu Ser Leu Asn Glu Leu Tyr
915 920 925
Gln Thr Lys Glu Glu Arg Ile Val Asn Lys Gln Val Ala Lys Asp Thr
930 935 940
Gln Glu Arg Asn Lys Asn Tyr Ile Trp Asn Lys Val Val Asp Leu Gln
945 950 955 960
Leu Cys Asp Gly Leu Val His Ile Asp Asn Val Lys Leu Lys Asp Ile

965 970 975

Gly Asn Phe Arg Lys Tyr Glu Asn Asp Ser Arg Val Lys Glu Phe Leu

980 985 990

Thr Tyr Gln Ser Asp Ile Val Trp Ser Ala Tyr Leu Ser Asn Glu Val

995 1000 1005

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

1010 1015 1020

Glu Ser Ile Arg Ser Lys Glu Leu Leu Lys Glu Val Gln Glu Ile

1025 1030 1035

Glu Cys Ser Val Tyr Asn Gln Val Ala Asn Lys Glu Ser Leu Lys

1040 1045 1050

Gln Ser Gly Asn Glu Asn Phe Lys Gln Tyr Val Leu Gln Gly Leu

1055 1060 1065

Leu Pro Ile Gly Met Asp Val Arg Glu Met Leu Ile Leu Ser Thr

1070 1075 1080

Asp Val Lys Phe Lys Lys Glu Glu Ile Ile Gln Leu Gly Gln Ala

1085 1090 1095

Gly Glu Val Glu Gln Asp Leu Tyr Ser Leu Ile Tyr Ile Arg Asn

1100 1105 1110

Lys Phe Ala His Asn Gln Leu Pro Ile Lys Glu Phe Phe Asp Phe

1115 1120 1125

Cys Glu Asn Asn Tyr Arg Ser Ile Ser Asp Asn Glu Tyr Tyr Ala

1130 1135 1140

Glu Tyr Tyr Met Glu Ile Phe Arg Ser Ile Lys Glu Lys Tyr Ala

1145 1150 1155

Asn

<210> 79

<211> 1120

<212> PRT

<213> Prevotella intermedia

<400> 79

Met Glu Asp Asp Lys Lys Thr Thr Asp Ser Ile Arg Tyr Glu Leu Lys
1 5 10 15
Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val
20 25 30
Tyr Ile Thr Val Asn His Ile Asn Lys Ile Leu Glu Glu Asp Glu Ile
35 40 45

Asn Arg Asp Gly Tyr Glu Asn Thr Leu Glu Asn Ser Trp Asn Glu Ile
50 55 60
Lys Asp Ile Asn Lys Lys Asp Arg Leu Ser Lys Leu Ile Ile Lys His
65 70 75 80
Phe Pro Phe Leu Glu Ala Thr Thr Tyr Arg Gln Asn Pro Thr Asp Thr
85 90 95
Thr Lys Gln Lys Glu Glu Lys Gln Ala Glu Ala Gln Ser Leu Glu Ser
100 105 110

Leu Lys Lys Ser Phe Phe Val Phe Ile Tyr Lys Leu Arg Asp Leu Arg
115 120 125
Asn His Tyr Ser His Tyr Lys His Ser Lys Ser Leu Glu Arg Pro Lys
130 135 140
Phe Glu Glu Asp Leu Gln Asn Lys Met Tyr Asn Ile Phe Asp Val Ser
145 150 155 160
Ile Gln Phe Val Lys Glu Asp Tyr Lys His Asn Thr Asp Ile Asn Pro
165 170 175

Lys Lys Asp Phe Lys His Leu Asp Arg Lys Arg Lys Gly Lys Phe His
180 185 190
Tyr Ser Phe Ala Asp Asn Glu Gly Asn Ile Thr Glu Ser Gly Leu Leu
195 200 205
Phe Phe Val Ser Leu Phe Leu Glu Lys Lys Asp Ala Ile Trp Val Gln
210 215 220
Lys Lys Leu Glu Gly Phe Lys Cys Ser Asn Lys Ser Tyr Gln Lys Met
225 230 235 240

Thr Asn Glu Val Phe Cys Arg Ser Arg Met Leu Leu Pro Lys Leu Arg

245 250 255

Leu Glu Ser Thr Gln Thr Gln Asp Trp Ile Leu Leu Asp Met Leu Asn

260 265 270

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

275 280 285

Asn Arg Lys Lys Phe Tyr Val Ser Phe Asp Pro Ala Asp Glu Asp Tyr

290 295 300

Asp Ala Glu Gln Glu Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp

305 310 315 320

Arg Phe Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Tyr Asn Glu Val Phe

325 330 335

Ala Asn Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ser Ile

340 345 350

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

355 360 365

Lys Leu Tyr Gly Phe Glu Arg Ile Gln Glu Phe Asp Lys Gln Asn Arg

370 375 380

Pro Asp Glu Trp Lys Ala Ile Val Lys Asp Ser Asp Thr Phe Lys Lys

385 390 395 400

Lys Glu Glu Lys Glu Glu Glu Lys Pro Tyr Ile Ser Glu Thr Thr Pro

405 410 415

His Tyr His Leu Glu Asn Lys Lys Ile Gly Ile Ala Phe Lys Asn His

420 425 430

Asn Ile Trp Pro Ser Thr Gln Thr Glu Leu Thr Asn Asn Lys Arg Lys

435 440 445

Lys Tyr Asn Leu Gly Thr Ser Ile Lys Ala Glu Ala Phe Leu Ser Val

450 455 460

His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu Lys Thr Glu

465 470 475 480

Asn Thr Lys Asn Asp Asn Lys Val Gly Gly Lys Lys Glu Thr Lys Lys

485 490 495

Gln Gly Lys His Lys Ile Glu Ala Ile Ile Glu Ser Lys Ile Lys Asp
500 505 510
Ile Tyr Ala Leu Tyr Asp Ala Phe Ala Asn Gly Glu Ile Asn Ser Glu
515 520 525
Asp Glu Leu Lys Glu Tyr Leu Lys Gly Lys Asp Ile Lys Ile Val His
530 535 540
Leu Pro Lys Gln Met Ile Ala Ile Leu Lys Asn Glu His Lys Asp Met
545 550 555 560

Ala Glu Lys Ala Glu Ala Lys Gln Glu Lys Met Lys Leu Ala Thr Glu
565 570 575
Asn Arg Leu Lys Thr Leu Asp Lys Gln Leu Lys Gly Lys Ile Gln Asn
580 585 590
Gly Lys Arg Tyr Asn Ser Ala Pro Lys Ser Gly Glu Ile Ala Ser Trp
595 600 605
Leu Val Asn Asp Met Met Arg Phe Gln Pro Val Gln Lys Asp Glu Asn
610 615 620

Gly Glu Ser Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Leu
625 630 635 640
Leu Gln Arg Thr Leu Ala Phe Phe Gly Ser Glu His Glu Arg Leu Ala
645 650 655
Pro Tyr Phe Lys Gln Thr Lys Leu Ile Glu Ser Ser Asn Pro His Pro
660 665 670
Phe Leu Asn Asp Thr Glu Trp Glu Lys Cys Ser Asn Ile Leu Ser Phe
675 680 685

Tyr Arg Ser Tyr Leu Lys Ala Arg Lys Asn Phe Leu Glu Ser Leu Lys
690 695 700
Pro Glu Asp Trp Glu Lys Asn Gln Tyr Phe Leu Met Leu Lys Glu Pro
705 710 715 720
Lys Thr Asn Arg Glu Thr Leu Val Gln Gly Trp Lys Asn Gly Phe Asn
725 730 735
Leu Pro Arg Gly Phe Phe Thr Glu Pro Ile Arg Lys Trp Phe Met Glu

740 745 750

His Trp Lys Ser Ile Lys Val Asp Asp Leu Lys Arg Val Gly Leu Val
755 760 765

Ala Lys Val Thr Pro Leu Phe Phe Ser Glu Lys Tyr Lys Asp Ser Val
770 775 780

Gln Pro Phe Tyr Asn Tyr Pro Phe Asn Val Gly Asp Val Asn Lys Pro
785 790 795 800

Lys Glu Glu Asp Phe Leu His Arg Glu Glu Arg Ile Glu Leu Trp Asp
805 810 815

Lys Lys Lys Asp Lys Phe Lys Gly Tyr Lys Ala Lys Lys Lys Phe Lys
820 825 830

Glu Met Thr Asp Lys Glu Lys Glu Glu His Arg Ser Tyr Leu Glu Phe
835 840 845

Gln Ser Trp Asn Lys Phe Glu Arg Glu Leu Arg Leu Val Arg Asn Gln
850 855 860

Asp Ile Val Thr Trp Leu Leu Cys Thr Glu Leu Ile Asp Lys Leu Lys
865 870 875 880

Ile Asp Glu Leu Asn Ile Lys Glu Leu Lys Lys Leu Arg Leu Lys Asp
885 890 895

Ile Asn Thr Asp Thr Ala Lys Lys Glu Lys Asn Asn Ile Leu Asn Arg
900 905 910

Val Met Pro Met Glu Leu Pro Val Thr Val Tyr Lys Val Asn Lys Gly
915 920 925

Gly Tyr Ile Ile Lys Asn Lys Pro Leu His Thr Ile Tyr Ile Lys Glu
930 935 940

Ala Glu Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys
945 950 955 960

Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe Val Lys Thr Pro Ser Glu
965 970 975

Ala Glu Ser Glu Ser Asn Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu
980 985 990

Leu Gly Lys Tyr Gln Asn Ala Arg Leu Asp Ile Ile Glu Asp Met Leu
995 1000 1005

Ala Leu Glu Lys Lys Leu Ile Asp Lys Tyr Asn Ser Leu Asp Thr
1010 1015 1020
Asp Asn Phe His Asn Met Leu Thr Gly Trp Leu Glu Leu Lys Gly
1025 1030 1035
Glu Ala Lys Lys Ala Arg Phe Gln Asn Asp Val Lys Leu Leu Thr
1040 1045 1050
Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp
1055 1060 1065

Glu Asn Leu Phe Gly Asn Ile Glu Arg Phe Ser Leu Ser Ser Ser
1070 1075 1080
Asn Ile Ile Glu Ser Lys Gly Leu Asp Ile Ala Ala Lys Leu Lys
1085 1090 1095
Glu Glu Val Ser Lys Ala Ala Lys Lys Ile Gln Asn Glu Glu Asp
1100 1105 1110
Asn Lys Lys Glu Lys Glu Thr
1115 1120

<210> 80
<211> 1199
<212> PRT
<213> Capnocytophaga canimorsus

<400> 80
Met Lys Asn Ile Gln Arg Leu Gly Lys Gly Asn Glu Phe Ser Pro Phe
1 5 10 15
Lys Lys Glu Asp Lys Phe Tyr Phe Gly Gly Phe Leu Asn Leu Ala Asn
20 25 30
Asn Asn Ile Glu Asp Phe Phe Lys Glu Ile Ile Thr Arg Phe Gly Ile
35 40 45
Val Ile Thr Asp Glu Asn Lys Lys Pro Lys Glu Thr Phe Gly Glu Lys
50 55 60

Ile Leu Asn Glu Ile Phe Lys Lys Asp Ile Ser Ile Val Asp Tyr Glu
65 70 75 80
Lys Trp Val Asn Ile Phe Ala Asp Tyr Phe Pro Phe Thr Lys Tyr Leu
85 90 95
Ser Leu Tyr Leu Glu Glu Met Gln Phe Lys Asn Arg Val Ile Cys Phe
100 105 110
Arg Asp Val Met Lys Glu Leu Leu Lys Thr Val Glu Ala Leu Arg Asn
115 120 125

Phe Tyr Thr His Tyr Asp His Glu Pro Ile Lys Ile Glu Asp Arg Val
130 135 140
Phe Tyr Phe Leu Asp Lys Val Leu Leu Asp Val Ser Leu Thr Val Lys
145 150 155 160
Asn Lys Tyr Leu Lys Thr Asp Lys Thr Lys Glu Phe Leu Asn Gln His
165 170 175
Ile Gly Glu Glu Leu Lys Glu Leu Cys Lys Gln Arg Lys Asp Tyr Leu
180 185 190

Val Gly Lys Gly Lys Arg Ile Asp Lys Glu Ser Glu Ile Ile Asn Gly
195 200 205
Ile Tyr Asn Asn Ala Phe Lys Asp Phe Ile Cys Lys Arg Glu Lys Gln
210 215 220
Asp Asp Lys Glu Asn His Asn Ser Val Glu Lys Ile Leu Cys Asn Lys
225 230 235 240
Glu Pro Gln Asn Lys Lys Gln Lys Ser Ser Ala Thr Val Trp Glu Leu
245 250 255

Cys Ser Lys Ser Ser Ser Lys Tyr Thr Glu Lys Ser Phe Pro Asn Arg
260 265 270
Glu Asn Asp Lys His Cys Leu Glu Val Pro Ile Ser Gln Lys Gly Ile
275 280 285
Val Phe Leu Leu Ser Phe Phe Leu Asn Lys Gly Glu Ile Tyr Ala Leu
290 295 300
Thr Ser Asn Ile Lys Gly Phe Lys Ala Lys Ile Thr Lys Glu Glu Pro

305 310 315 320

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

Val Val Tyr Gln Asn Leu Ser Glu Asp Val Gln Lys Thr Phe Ile Glu
385 390 395 400
Asp Trp Asn Glu Tyr Leu Lys Glu Asn Asn Gly Asp Val Gly Thr Met
405 410 415
Glu Glu Glu Gln Val Ile His Pro Val Ile Arg Lys Arg Tyr Glu Asp
420 425 430
Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Ala Gln Phe
435 440 445

Pro Thr Leu Arg Phe Gln Val His Leu Gly Asn Tyr Leu Cys Asp Lys
450 455 460
Arg Thr Lys Gln Ile Cys Asp Thr Thr Thr Glu Arg Glu Val Lys Lys
465 470 475 480
Lys Ile Thr Val Phe Gly Arg Leu Ser Glu Leu Glu Asn Lys Lys Ala
485 490 495
Ile Phe Leu Asn Glu Arg Glu Glu Ile Lys Gly Trp Glu Val Phe Pro
500 505 510

Asn Pro Ser Tyr Asp Phe Pro Lys Glu Asn Ile Ser Val Asn Tyr Lys
515 520 525
Asp Phe Pro Ile Val Gly Ser Ile Leu Asp Arg Glu Lys Gln Pro Val
530 535 540
Ser Asn Lys Ile Gly Ile Arg Val Lys Ile Ala Asp Glu Leu Gln Arg
545 550 555 560

Glu Ile Asp Lys Ala Ile Lys Glu Lys Lys Leu Arg Asn Pro Lys Asn
565 570 575

Arg Lys Ala Asn Gln Asp Glu Lys Gln Lys Glu Arg Leu Val Asn Glu
580 585 590

Ile Val Ser Thr Asn Ser Asn Glu Gln Gly Glu Pro Val Val Phe Ile
595 600 605

Gly Gln Pro Thr Ala Tyr Leu Ser Met Asn Asp Ile His Ser Val Leu
610 615 620

Tyr Glu Phe Leu Ile Asn Lys Ile Ser Gly Glu Ala Leu Glu Thr Lys
625 630 635 640

Ile Val Glu Lys Ile Glu Thr Gln Ile Lys Gln Ile Ile Gly Lys Asp
645 650 655

Ala Thr Thr Lys Ile Leu Lys Pro Tyr Thr Asn Ala Asn Ser Asn Ser
660 665 670

Ile Asn Arg Glu Lys Leu Leu Arg Asp Leu Glu Gln Glu Gln Gln Ile
675 680 685

Leu Lys Thr Leu Leu Glu Glu Gln Gln Gln Arg Glu Lys Asp Lys Lys
690 695 700

Asp Lys Lys Ser Lys Arg Lys His Glu Leu Tyr Pro Ser Glu Lys Gly
705 710 715 720

Lys Val Ala Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met Pro Lys
725 730 735

Ala Phe Lys Glu Gln Trp Arg Gly Tyr His His Ser Leu Leu Gln Lys
740 745 750

Tyr Leu Ala Tyr Tyr Glu Gln Ser Lys Glu Glu Leu Lys Asn Leu Leu
755 760 765

Pro Lys Glu Val Phe Lys His Phe Pro Phe Lys Leu Lys Gly Tyr Phe
770 775 780

Gln Gln Gln Tyr Leu Asn Gln Phe Tyr Thr Asp Tyr Leu Lys Arg Arg
785 790 795 800

Leu Ser Tyr Val Asn Glu Leu Leu Leu Asn Ile Gln Asn Phe Lys Asn

805 810 815

Asp Lys Asp Ala Leu Lys Ala Thr Glu Lys Glu Cys Phe Lys Phe Phe

820 825 830

Arg Lys Gln Asn Tyr Ile Ile Asn Pro Ile Asn Ile Gln Ile Gln Ser

835 840 845

Ile Leu Val Tyr Pro Ile Phe Leu Lys Arg Gly Phe Leu Asp Glu Lys

850 855 860

Pro Thr Met Ile Asp Arg Glu Lys Phe Lys Glu Asn Lys Asp Thr Glu

865 870 875 880

Leu Ala Asp Trp Phe Met His Tyr Lys Asn Tyr Lys Glu Asp Asn Tyr

885 890 895

Gln Lys Phe Tyr Ala Tyr Pro Leu Glu Lys Val Glu Glu Lys Glu Lys

900 905 910

Phe Lys Arg Asn Lys Gln Ile Asn Lys Gln Lys Lys Asn Asp Val Tyr

915 920 925

Thr Leu Met Met Val Glu Tyr Ile Ile Gln Lys Ile Phe Gly Asp Lys

930 935 940

Phe Val Glu Glu Asn Pro Leu Val Leu Lys Gly Ile Phe Gln Ser Lys

945 950 955 960

Ala Glu Arg Gln Gln Asn Asn Thr His Ala Ala Thr Thr Gln Glu Arg

965 970 975

Asn Leu Asn Gly Ile Leu Asn Gln Pro Lys Asp Ile Lys Ile Gln Gly

980 985 990

Lys Ile Thr Val Lys Gly Val Lys Leu Lys Asp Ile Gly Asn Phe Arg

995 1000 1005

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

1010 1015 1020

Pro Arg Lys Glu Trp Met Ala Tyr Leu Pro Asn Asp Trp Lys Glu

1025 1030 1035

Lys Glu Lys Gln Gly Gln Leu Pro Pro Asn Asn Val Ile Asp Arg

1040 1045 1050

Gln Ile Ser Lys Tyr Glu Thr Val Arg Ser Lys Ile Leu Leu Lys
1055 1060 1065
Asp Val Gln Glu Leu Glu Lys Ile Ile Ser Asp Glu Ile Lys Glu
1070 1075 1080

Glu His Arg His Asp Leu Lys Gln Gly Lys Tyr Tyr Asn Phe Lys
1085 1090 1095
Tyr Tyr Ile Leu Asn Gly Leu Leu Arg Gln Leu Lys Asn Glu Asn
1100 1105 1110
Val Glu Asn Tyr Lys Val Phe Lys Leu Asn Thr Asn Pro Glu Lys
1115 1120 1125
Val Asn Ile Thr Gln Leu Lys Gln Glu Ala Thr Asp Leu Glu Gln
1130 1135 1140

Lys Ala Phe Val Leu Thr Tyr Ile Arg Asn Lys Phe Ala His Asn
1145 1150 1155
Gln Leu Pro Lys Lys Glu Phe Trp Asp Tyr Cys Gln Glu Lys Tyr
1160 1165 1170
Gly Lys Ile Glu Lys Glu Lys Thr Tyr Ala Glu Tyr Phe Ala Glu
1175 1180 1185
Val Phe Lys Arg Glu Lys Glu Ala Leu Ile Lys
1190 1195

<210> 81
<211> 1175

<212> PRT
<213> Porphyromonas gulae
<400> 81
Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15
Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn
20 25 30
Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys
35 40 45

Ala Asp Ile Thr Asn Asp Gln Asp Val Leu Ser Phe Lys Ala Leu Trp

50 55 60

Lys Asn Phe Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu
65 70 75 80
Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95
Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu
100 105 110
Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140
Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Gly Ser Ser Glu Leu
145 150 155 160
Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175
Val Ser Val Gln Arg Val Lys Ile Asp His Glu His Asn Asp Glu Val

180 185 190

Asp Pro His Tyr His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205
Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp
210 215 220
Gly Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240
Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe
260 265 270
Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285
Met Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Asp Asp Arg Ala Cys Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly

325 330 335

Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Ser Gln Thr Ser Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Met Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Arg Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

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

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Arg Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asp Met Glu

565 570 575
Glu Lys Ile Arg Lys Lys Leu Gln Glu Met Met Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Ala Ser Gly

625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700
Arg Ser Tyr Leu Arg Ala Arg Lys Ala Phe Leu Glu Arg Ile Gly Arg
705 710 715 720
Ser Asp Arg Val Glu Asn Arg Pro Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu
740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765
Gly His Asp Glu Val Ala Ser Tyr Lys Glu Val Gly Phe Met Ala Lys
770 775 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Glu Arg Ala Glu Glu Trp Glu Arg Gly

820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser Tyr Ser Ala Ala Arg

835 840 845

Arg Ile Glu Asp Ala Phe Ala Gly Ile Glu Tyr Ala Ser Pro Gly Asn

850 855 860

Lys Lys Lys Ile Glu Gln Leu Leu Arg Asp Leu Ser Leu Trp Glu Ala

865 870 875 880

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

885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Gln Glu His Pro Tyr His Asp Phe

900 905 910

Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln

915 920 925

Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys

930 935 940

Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Pro

945 950 955 960

Asn Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn Arg Val Lys Pro

965 970 975

Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val

980 985 990

His Lys Glu Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp

995 1000 1005

Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp

1010 1015 1020

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

1025 1030 1035

Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu
1040 1045 1050
Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Val Phe Glu Leu Thr
1055 1060 1065
Leu Arg Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro

1070 1075 1080
Asp Glu Ser Phe Arg Glu Met Leu Glu Ser Trp Ser Asp Pro Leu
1085 1090 1095
Leu Ala Lys Trp Pro Glu Leu His Gly Lys Val Arg Leu Leu Ile
1100 1105 1110
Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp
1115 1120 1125
Glu Ala Val Phe Ser Ser Ile Arg Lys Tyr Asp Pro Ser Ser Pro

1130 1135 1140
Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu
1145 1150 1155
Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val Glu Arg Ile Ile
1160 1165 1170
Gln Ala
1175
<210> 82
<211> 1090
<212> PRT
<213> Prevotella species
<400> 82
Met Asn Ile Pro Ala Leu Val Glu Asn Gln Lys Lys Tyr Phe Gly Thr
1 5 10 15

Tyr Ser Val Met Ala Met Leu Asn Ala Gln Thr Val Leu Asp His Ile
20 25 30
Gln Lys Val Ala Asp Ile Glu Gly Glu Gln Asn Glu Asn Asn Glu Asn
35 40 45
Leu Trp Phe His Pro Val Met Ser His Leu Tyr Asn Ala Lys Asn Gly

50 55 60

Tyr Asp Lys Gln Pro Glu Lys Thr Met Phe Ile Ile Glu Arg Leu Gln
65 70 75 80

Ser Tyr Phe Pro Phe Leu Lys Ile Met Ala Glu Asn Gln Arg Glu Tyr
85 90 95

Ser Asn Gly Lys Tyr Lys Gln Asn Arg Val Glu Val Asn Ser Asn Asp
100 105 110

Ile Phe Glu Val Leu Lys Arg Ala Phe Gly Val Leu Lys Met Tyr Arg
115 120 125

Asp Leu Thr Asn His Tyr Lys Thr Tyr Glu Glu Lys Leu Asn Asp Gly
130 135 140

Cys Glu Phe Leu Thr Ser Thr Glu Gln Pro Leu Ser Gly Met Ile Asn
145 150 155 160

Asn Tyr Tyr Thr Val Ala Leu Arg Asn Met Asn Glu Arg Tyr Gly Tyr
165 170 175

Lys Thr Glu Asp Leu Ala Phe Ile Gln Asp Lys Arg Phe Lys Phe Val
180 185 190

Lys Asp Ala Tyr Gly Lys Lys Lys Ser Gln Val Asn Thr Gly Phe Phe
195 200 205

Leu Ser Leu Gln Asp Tyr Asn Gly Asp Thr Gln Lys Lys Leu His Leu
210 215 220

Ser Gly Val Gly Ile Ala Leu Leu Ile Cys Leu Phe Leu Asp Lys Gln
225 230 235 240

Tyr Ile Asn Ile Phe Leu Ser Arg Leu Pro Ile Phe Ser Ser Tyr Asn
245 250 255

Ala Gln Ser Glu Glu Arg Arg Ile Ile Ile Arg Ser Phe Gly Ile Asn
260 265 270

Ser Ile Lys Leu Pro Lys Asp Arg Ile His Ser Glu Lys Ser Asn Lys
275 280 285

Ser Val Ala Met Asp Met Leu Asn Glu Val Lys Arg Cys Pro Asp Glu
290 295 300

Leu Phe Thr Thr Leu Ser Ala Glu Lys Gln Ser Arg Phe Arg Ile Ile
305 310 315 320
Ser Asp Asp His Asn Glu Val Leu Met Lys Arg Ser Ser Asp Arg Phe
325 330 335

Val Pro Leu Leu Leu Gln Tyr Ile Asp Tyr Gly Lys Leu Phe Asp His
340 345 350
Ile Arg Phe His Val Asn Met Gly Lys Leu Arg Tyr Leu Leu Lys Ala
355 360 365
Asp Lys Thr Cys Ile Asp Gly Gln Thr Arg Val Arg Val Ile Glu Gln
370 375 380
Pro Leu Asn Gly Phe Gly Arg Leu Glu Glu Ala Glu Thr Met Arg Lys
385 390 395 400

Gln Glu Asn Gly Thr Phe Gly Asn Ser Gly Ile Arg Ile Arg Asp Phe
405 410 415
Glu Asn Met Lys Arg Asp Asp Ala Asn Pro Ala Asn Tyr Pro Tyr Ile
420 425 430
Val Asp Thr Tyr Thr His Tyr Ile Leu Glu Asn Asn Lys Val Glu Met
435 440 445
Phe Ile Asn Asp Lys Glu Asp Ser Ala Pro Leu Leu Pro Val Ile Glu
450 455 460

Asp Asp Arg Tyr Val Val Lys Thr Ile Pro Ser Cys Arg Met Ser Thr
465 470 475 480
Leu Glu Ile Pro Ala Met Ala Phe His Met Phe Leu Phe Gly Ser Lys
485 490 495
Lys Thr Glu Lys Leu Ile Val Asp Val His Asn Arg Tyr Lys Arg Leu
500 505 510
Phe Gln Ala Met Gln Lys Glu Glu Val Thr Ala Glu Asn Ile Ala Ser
515 520 525

Phe Gly Ile Ala Glu Ser Asp Leu Pro Gln Lys Ile Leu Asp Leu Ile
530 535 540
Ser Gly Asn Ala His Gly Lys Asp Val Asp Ala Phe Ile Arg Leu Thr

545 550 555 560

Val Asp Asp Met Leu Thr Asp Thr Glu Arg Arg Ile Lys Arg Phe Lys

565 570 575

Asp Asp Arg Lys Ser Ile Arg Ser Ala Asp Asn Lys Met Gly Lys Arg

580 585 590

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

595 600 605

Asp Ile Val Leu Phe Gln Pro Ser Val Asn Asp Gly Glu Asn Lys Ile

610 615 620

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

625 630 635 640

Ser Gly Asp Asp Tyr Glu Ala Lys Gln Gln Phe Lys Leu Met Phe Glu

645 650 655

Lys Ala Arg Leu Ile Gly Lys Gly Thr Thr Glu Pro His Pro Phe Leu

660 665 670

Tyr Lys Val Phe Ala Arg Ser Ile Pro Ala Asn Ala Val Glu Phe Tyr

675 680 685

Glu Arg Tyr Leu Ile Glu Arg Lys Phe Tyr Leu Thr Gly Leu Ser Asn

690 695 700

Glu Ile Lys Lys Gly Asn Arg Val Asp Val Pro Phe Ile Arg Arg Asp

705 710 715 720

Gln Asn Lys Trp Lys Thr Pro Ala Met Lys Thr Leu Gly Arg Ile Tyr

725 730 735

Ser Glu Asp Leu Pro Val Glu Leu Pro Arg Gln Met Phe Asp Asn Glu

740 745 750

Ile Lys Ser His Leu Lys Ser Leu Pro Gln Met Glu Gly Ile Asp Phe

755 760 765

Asn Asn Ala Asn Val Thr Tyr Leu Ile Ala Glu Tyr Met Lys Arg Val

770 775 780

Leu Asp Asp Asp Phe Gln Thr Phe Tyr Gln Trp Asn Arg Asn Tyr Arg

785 790 795 800

Tyr Met Asp Met Leu Lys Gly Glu Tyr Asp Arg Lys Gly Ser Leu Gln
805 810 815

His Cys Phe Thr Ser Val Glu Glu Arg Glu Gly Leu Trp Lys Glu Arg
820 825 830

Ala Ser Arg Thr Glu Arg Tyr Arg Lys Gln Ala Ser Asn Lys Ile Arg
835 840 845

Ser Asn Arg Gln Met Arg Asn Ala Ser Ser Glu Glu Ile Glu Thr Ile
850 855 860

Leu Asp Lys Arg Leu Ser Asn Ser Arg Asn Glu Tyr Gln Lys Ser Glu
865 870 875 880

Lys Val Ile Arg Arg Tyr Arg Val Gln Asp Ala Leu Leu Phe Leu Leu
885 890 895

Ala Lys Lys Thr Leu Thr Glu Leu Ala Asp Phe Asp Gly Glu Arg Phe
900 905 910

Lys Leu Lys Glu Ile Met Pro Asp Ala Glu Lys Gly Ile Leu Ser Glu
915 920 925

Ile Met Pro Met Ser Phe Thr Phe Glu Lys Gly Gly Lys Lys Tyr Thr
930 935 940

Ile Thr Ser Glu Gly Met Lys Leu Lys Asn Tyr Gly Asp Phe Phe Val
945 950 955 960

Leu Ala Ser Asp Lys Arg Ile Gly Asn Leu Leu Glu Leu Val Gly Ser
965 970 975

Asp Ile Val Ser Lys Glu Asp Ile Met Glu Glu Phe Asn Lys Tyr Asp
980 985 990

Gln Cys Arg Pro Glu Ile Ser Ser Ile Val Phe Asn Leu Glu Lys Trp
995 1000 1005

Ala Phe Asp Thr Tyr Pro Glu Leu Ser Ala Arg Val Asp Arg Glu
1010 1015 1020

Glu Lys Val Asp Phe Lys Ser Ile Leu Lys Ile Leu Leu Asn Asn
1025 1030 1035

Lys Asn Ile Asn Lys Glu Gln Ser Asp Ile Leu Arg Lys Ile Arg

1040 1045 1050

Asn Ala Phe Asp His Asn Asn Tyr Pro Asp Lys Gly Val Val Glu

1055 1060 1065

Ile Lys Ala Leu Pro Glu Ile Ala Met Ser Ile Lys Lys Ala Phe

1070 1075 1080

Gly Glu Tyr Ala Ile Met Lys

1085 1090

<210> 83

<211> 1150

<212> PRT

<213> Flavobacterium branchiophilum

<400> 83

Met Glu Asn Leu Asn Lys Ile Leu Asp Lys Glu Asn Glu Ile Cys Ile

1 5 10 15

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

20 25 30

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

35 40 45

Ala Arg Leu His Tyr Phe Ser Ile Gly His Ser Phe Lys Gln Ile Asp

50 55 60

Thr Lys Lys Val Phe Asp Tyr Val Leu Ile Glu Glu Leu Lys Asp Glu

65 70 75 80

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

85 90 95

Phe Ser Ile Lys Leu Gln Lys Leu Ile Asn Ser Ile Arg Asn Ile Asn

100 105 110

Asn His Tyr Val His Asn Phe Asn Asp Ile Asn Leu Asn Lys Ile Asp

115 120 125

Ser Asn Val Phe His Phe Leu Lys Glu Ser Phe Glu Leu Ala Ile Ile

130 135 140

Glu Lys Tyr Tyr Lys Val Asn Lys Lys Tyr Pro Leu Asp Asn Glu Ile

145 150 155 160

Val Leu Phe Leu Lys Glu Leu Phe Ile Lys Asp Glu Asn Thr Ala Leu
165 170 175
Leu Asn Tyr Phe Thr Asn Leu Ser Lys Asp Glu Ala Ile Glu Tyr Ile
180 185 190

Leu Thr Phe Thr Ile Thr Glu Asn Lys Ile Trp Asn Ile Asn Asn Glu
195 200 205
His Asn Ile Leu Asn Ile Glu Lys Gly Lys Tyr Leu Thr Phe Glu Ala
210 215 220
Met Leu Phe Leu Ile Thr Ile Phe Leu Tyr Lys Asn Glu Ala Asn His
225 230 235 240
Leu Leu Pro Lys Leu Tyr Asp Phe Lys Asn Asn Lys Ser Lys Gln Glu
245 250 255

Leu Phe Thr Phe Phe Ser Lys Lys Phe Thr Ser Gln Asp Ile Asp Ala
260 265 270
Glu Glu Gly His Leu Ile Lys Phe Arg Asp Met Ile Gln Tyr Leu Asn
275 280 285
His Tyr Pro Thr Ala Trp Asn Asn Asp Leu Lys Leu Glu Ser Glu Asn
290 295 300
Lys Asn Lys Ile Met Thr Thr Lys Leu Ile Asp Ser Ile Ile Glu Phe
305 310 315 320

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

Ser Phe Asn Val Lys Glu Asp Lys Phe Lys Ile Phe Val Lys Lys His
385 390 395 400
Val Leu Glu Glu Tyr Phe Pro Asn Ser Ile Gly Tyr Glu Lys Phe Leu

405 410 415

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

420 425 430

Leu Tyr Ser Asn Pro Lys Thr Asn Lys Leu Ile Glu Arg Ile Asp Asn

435 440 445

His Lys Leu Val Lys Ser His Gly Arg Asn Gln Asp Arg Phe Met Asp

450 455 460

Phe Ser Met Arg Phe Leu Ala Glu Asn Asn Tyr Phe Gly Lys Asp Ala

465 470 475 480

Phe Phe Lys Cys Tyr Lys Phe Tyr Asp Thr Gln Glu Gln Asp Glu Phe

485 490 495

Leu Gln Ser Asn Glu Asn Asn Asp Asp Val Lys Phe His Lys Gly Lys

500 505 510

Val Thr Thr Tyr Ile Lys Tyr Glu Glu His Leu Lys Asn Tyr Ser Tyr

515 520 525

Trp Asp Cys Pro Phe Val Glu Glu Asn Asn Ser Met Ser Val Lys Ile

530 535 540

Ser Ile Gly Ser Glu Glu Lys Ile Leu Lys Ile Gln Arg Asn Leu Met

545 550 555 560

Ile Tyr Phe Leu Glu Asn Ala Leu Tyr Asn Glu Asn Val Glu Asn Gln

565 570 575

Gly Tyr Lys Leu Val Asn Asn Tyr Tyr Arg Glu Leu Lys Lys Asp Val

580 585 590

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

595 600 605

Lys Ser Lys Tyr Lys Lys Ile Leu Pro Lys Arg Leu Leu His Asn Tyr

610 615 620

Ala Pro Ala Lys Gln Asp Lys Ala Pro Glu Asn Ala Phe Glu Thr Leu

625 630 635 640

Leu Lys Lys Ala Asp Phe Arg Glu Glu Gln Tyr Lys Lys Leu Leu Lys

645 650 655

Lys Ala Glu His Glu Lys Asn Lys Glu Asp Phe Val Lys Arg Asn Lys
660 665 670
Gly Lys Gln Phe Lys Leu His Phe Ile Arg Lys Ala Cys Gln Met Met
675 680 685
Tyr Phe Lys Glu Lys Tyr Asn Thr Leu Lys Glu Gly Asn Ala Ala Phe
690 695 700

Glu Lys Lys Asp Pro Val Ile Glu Lys Arg Lys Asn Lys Glu His Glu
705 710 715 720
Phe Gly His His Lys Asn Leu Asn Ile Thr Arg Glu Glu Phe Asn Asp
725 730 735
Tyr Cys Lys Trp Met Phe Ala Phe Asn Gly Asn Asp Ser Tyr Lys Lys
740 745 750
Tyr Leu Arg Asp Leu Phe Ser Glu Lys His Phe Phe Asp Asn Gln Glu
755 760 765

Tyr Lys Asn Leu Phe Glu Ser Ser Val Asn Leu Glu Ala Phe Tyr Ala
770 775 780
Lys Thr Lys Glu Leu Phe Lys Lys Trp Ile Glu Thr Asn Lys Pro Thr
785 790 795 800
Asn Asn Glu Asn Arg Tyr Thr Leu Glu Asn Tyr Lys Asn Leu Ile Leu
805 810 815
Gln Lys Gln Val Phe Ile Asn Val Tyr His Phe Ser Lys Tyr Leu Ile
820 825 830

Asp Lys Asn Leu Leu Asn Ser Glu Asn Asn Val Ile Gln Tyr Lys Ser
835 840 845
Leu Glu Asn Val Glu Tyr Leu Ile Ser Asp Phe Tyr Phe Gln Ser Lys
850 855 860
Leu Ser Ile Asp Gln Tyr Lys Thr Cys Gly Lys Leu Phe Asn Lys Leu
865 870 875 880
Lys Ser Asn Lys Leu Glu Asp Cys Leu Leu Tyr Glu Ile Ala Tyr Asn
885 890 895

Tyr Ile Asp Lys Lys Asn Val His Lys Ile Asp Ile Gln Lys Ile Leu

900 905 910

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

915 920 925

Ile Ser Val Pro Phe Asn Lys Leu Glu Arg Tyr Thr Glu Met Ile Ala

930 935 940

Ile Lys Asn Gln Asn Asn Leu Lys Ala Arg Phe Leu Ile Asp Leu Pro

945 950 955 960

Leu Tyr Leu Ser Lys Asn Lys Ile Lys Lys Gly Lys Asp Ser Ala Gly

965 970 975

Tyr Glu Ile Ile Ile Lys Asn Asp Leu Glu Ile Glu Asp Ile Asn Thr

980 985 990

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

995 1000 1005

Met Glu Leu Glu Lys Tyr Phe Ile Leu Lys Asp Lys Cys Ile Leu

1010 1015 1020

Ser Lys Asn Tyr Ile Asp Asn Ser Glu Ile Pro Ser Leu Lys Gln

1025 1030 1035

Phe Ser Lys Val Trp Ile Lys Glu Asn Glu Asn Glu Ile Ile Asn

1040 1045 1050

Tyr Arg Asn Ile Ala Cys His Phe His Leu Pro Leu Leu Glu Thr

1055 1060 1065

Phe Asp Asn Leu Leu Leu Asn Val Glu Gln Lys Phe Ile Lys Glu

1070 1075 1080

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

1085 1090 1095

Glu Tyr Leu Ile Leu Leu Phe Ile Lys Phe Lys His Asn Asn Phe

1100 1105 1110

Tyr Leu Asn Leu Phe Asn Lys Asn Glu Ser Lys Thr Ile Lys Asn

1115 1120 1125

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

1130 1135 1140

Gln Val Ile Leu Lys Lys Lys

1145 1150

<210> 84

<211> 1159

<212> PRT

<213> Myroides odoratimimus

<400> 84

Met Lys Asp Ile Leu Thr Thr Asp Thr Thr Glu Lys Gln Asn Arg Phe
1 5 10 15

Tyr Ser His Lys Ile Ala Asp Lys Tyr Phe Phe Gly Gly Tyr Phe Asn
20 25 30

Leu Ala Ser Asn Asn Ile Tyr Glu Val Phe Glu Glu Val Asn Lys Arg
35 40 45

Asn Thr Phe Gly Lys Leu Ala Lys Arg Asp Asn Gly Asn Leu Lys Asn
50 55 60

Tyr Ile Ile His Val Phe Lys Asp Glu Leu Ser Ile Ser Asp Phe Glu
65 70 75 80

Lys Arg Val Ala Ile Phe Ala Ser Tyr Phe Pro Ile Leu Glu Thr Val
85 90 95

Asp Lys Lys Ser Ile Lys Glu Arg Asn Arg Thr Ile Asp Leu Thr Leu
100 105 110

Ser Gln Arg Ile Arg Gln Phe Arg Glu Met Leu Ile Ser Leu Val Thr
115 120 125

Ala Val Asp Gln Leu Arg Asn Phe Tyr Thr His Tyr His His Ser Asp
130 135 140

Ile Val Ile Glu Asn Lys Val Leu Asp Phe Leu Asn Ser Ser Phe Val
145 150 155 160

Ser Thr Ala Leu His Val Lys Asp Lys Tyr Leu Lys Thr Asp Lys Thr
165 170 175

Lys Glu Phe Leu Lys Glu Thr Ile Ala Ala Glu Leu Asp Ile Leu Ile
180 185 190

Glu Ala Tyr Lys Lys Lys Gln Ile Glu Lys Lys Asn Thr Arg Phe Lys

195 200 205

Ala Asn Lys Arg Glu Asp Ile Leu Asn Ala Ile Tyr Asn Glu Ala Phe

210 215 220

Trp Ser Phe Ile Asn Asp Lys Asp Lys Asp Lys Asp Lys Glu Thr Val

225 230 235 240

Val Ala Lys Gly Ala Asp Ala Tyr Phe Glu Lys Asn His His Lys Ser

245 250 255

Asn Asp Pro Asp Phe Ala Leu Asn Ile Ser Glu Lys Gly Ile Val Tyr

260 265 270

Leu Leu Ser Phe Phe Leu Thr Asn Lys Glu Met Asp Ser Leu Lys Ala

275 280 285

Asn Leu Thr Gly Phe Lys Gly Lys Val Asp Arg Glu Ser Gly Asn Ser

290 295 300

Ile Lys Tyr Met Ala Thr Gln Arg Ile Tyr Ser Phe His Thr Tyr Arg

305 310 315 320

Gly Leu Lys Gln Lys Ile Arg Thr Ser Glu Glu Gly Val Lys Glu Thr

325 330 335

Leu Leu Met Gln Met Ile Asp Glu Leu Ser Lys Val Pro Asn Val Val

340 345 350

Tyr Gln His Leu Ser Thr Thr Gln Gln Asn Ser Phe Ile Glu Asp Trp

355 360 365

Asn Glu Tyr Tyr Lys Asp Tyr Glu Asp Asp Val Glu Thr Asp Asp Leu

370 375 380

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

385 390 395 400

Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Phe Asp Phe Pro Thr

405 410 415

Leu Arg Phe Gln Val His Leu Gly Asp Tyr Val His Asp Arg Arg Thr

420 425 430

Lys Gln Leu Gly Lys Val Glu Ser Asp Arg Ile Ile Lys Glu Lys Val

435 440 445

Thr Val Phe Ala Arg Leu Lys Asp Ile Asn Ser Ala Lys Ala Ser Tyr
450 455 460
Phe His Ser Leu Glu Glu Gln Asp Lys Glu Glu Leu Asp Asn Lys Trp
465 470 475 480
Thr Leu Phe Pro Asn Pro Ser Tyr Asp Phe Pro Lys Glu His Thr Leu
485 490 495

Gln His Gln Gly Glu Gln Lys Asn Ala Gly Lys Ile Gly Ile Tyr Val
500 505 510
Lys Leu Arg Asp Thr Gln Tyr Lys Glu Lys Ala Ala Leu Glu Glu Ala
515 520 525
Arg Lys Ser Leu Asn Pro Lys Glu Arg Ser Ala Thr Lys Ala Ser Lys
530 535 540
Tyr Asp Ile Ile Thr Gln Ile Ile Glu Ala Asn Asp Asn Val Lys Ser
545 550 555 560

Glu Lys Pro Leu Val Phe Thr Gly Gln Pro Ile Ala Tyr Leu Ser Met
565 570 575
Asn Asp Ile His Ser Met Leu Phe Ser Leu Leu Thr Asp Asn Ala Glu
580 585 590
Leu Lys Lys Thr Pro Glu Glu Val Glu Ala Lys Leu Ile Asp Gln Ile
595 600 605
Gly Lys Gln Ile Asn Glu Ile Leu Ser Lys Asp Thr Asp Thr Lys Ile
610 615 620

Leu Lys Lys Tyr Lys Asp Asn Asp Leu Lys Glu Thr Asp Thr Asp Lys
625 630 635 640
Ile Thr Arg Asp Leu Ala Arg Asp Lys Glu Glu Ile Glu Lys Leu Ile
645 650 655
Leu Glu Gln Lys Gln Arg Ala Asp Asp Tyr Asn Tyr Thr Ser Ser Thr
660 665 670
Lys Phe Asn Ile Asp Lys Ser Arg Lys Arg Lys His Leu Leu Phe Asn
675 680 685

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

690 695 700

Phe Met Phe Lys Glu Ser Lys Ser Lys Trp Lys Gly Tyr Gln His Ile

705 710 715 720

Glu Leu Gln Lys Leu Phe Ala Tyr Phe Asp Thr Ser Lys Ser Asp Leu

725 730 735

Glu Leu Ile Leu Ser Asn Met Val Met Val Lys Asp Tyr Pro Ile Glu

740 745 750

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

755 760 765

Lys Tyr Leu Glu Ala Arg Leu Glu Tyr Ile Glu Asn Val Ile Thr Arg

770 775 780

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

785 790 795 800

Cys Phe Thr Phe Leu Lys Lys Ser Asn Tyr Thr Val Val Ser Leu Asp

805 810 815

Lys Gln Val Glu Arg Ile Leu Ser Met Pro Leu Phe Ile Glu Arg Gly

820 825 830

Phe Met Asp Asp Lys Pro Thr Met Leu Glu Gly Lys Ser Tyr Lys Gln

835 840 845

His Lys Glu Lys Phe Ala Asp Trp Phe Val His Tyr Lys Glu Asn Ser

850 855 860

Asn Tyr Gln Asn Phe Tyr Asp Thr Glu Val Tyr Glu Ile Thr Thr Glu

865 870 875 880

Asp Lys Arg Glu Lys Ala Lys Val Thr Lys Lys Ile Lys Gln Gln Gln

885 890 895

Lys Asn Asp Val Phe Thr Leu Met Met Val Asn Tyr Met Leu Glu Glu

900 905 910

Val Leu Lys Leu Ser Ser Asn Asp Arg Leu Ser Leu Asn Glu Leu Tyr

915 920 925

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

930 935 940

Gln Glu Arg Asn Lys Asn Tyr Ile Trp Asn Lys Val Val Asp Leu Gln
945 950 955 960
Leu Cys Asp Gly Leu Val His Ile Asp Asn Val Lys Leu Lys Asp Ile
965 970 975
Gly Asn Phe Arg Lys Tyr Glu Asn Asp Ser Arg Val Lys Glu Phe Leu
980 985 990
Thr Tyr Gln Ser Asp Ile Val Trp Ser Ala Tyr Leu Ser Asn Glu Val
995 1000 1005

Asp Ser Asn Lys Leu Tyr Val Ile Glu Arg Gln Leu Asp Asn Tyr
1010 1015 1020
Glu Ser Ile Arg Ser Lys Glu Leu Leu Lys Glu Val Gln Glu Ile
1025 1030 1035
Glu Cys Ser Val Tyr Asn Gln Val Ala Asn Lys Glu Ser Leu Lys
1040 1045 1050
Gln Ser Gly Asn Glu Asn Phe Lys Gln Tyr Val Leu Gln Gly Leu
1055 1060 1065

Leu Pro Ile Gly Met Asp Val Arg Glu Met Leu Ile Leu Ser Thr
1070 1075 1080
Asp Val Lys Phe Lys Lys Glu Glu Ile Ile Gln Leu Gly Gln Ala
1085 1090 1095
Gly Glu Val Glu Gln Asp Leu Tyr Ser Leu Ile Tyr Ile Arg Asn
1100 1105 1110
Lys Phe Ala His Asn Gln Leu Pro Ile Lys Glu Phe Phe Asp Phe
1115 1120 1125

Cys Glu Asn Asn Tyr Arg Ser Ile Ser Asp Asn Glu Tyr Tyr Ala
1130 1135 1140
Glu Tyr Tyr Met Glu Ile Phe Arg Ser Ile Lys Glu Lys Tyr Ala
1145 1150 1155
Asn

<210> 85
<211> 1214

<212> PRT

<213> Flavobacterium columnare

<400> 85

Met Ser Ser Lys Asn Glu Ser Tyr Asn Lys Gln Lys Thr Phe Asn His
1 5 10 15

Tyr Lys Gln Glu Asp Lys Tyr Phe Phe Gly Gly Phe Leu Asn Asn Ala
20 25 30

Asp Asp Asn Leu Arg Gln Val Gly Lys Glu Phe Lys Thr Arg Ile Asn
35 40 45

Phe Asn His Asn Asn Asn Glu Leu Ala Ser Val Phe Lys Asp Tyr Phe
50 55 60

Asn Lys Glu Lys Ser Val Ala Lys Arg Glu His Ala Leu Asn Leu Leu
65 70 75 80

Ser Asn Tyr Phe Pro Val Leu Glu Arg Ile Gln Lys His Thr Asn His
85 90 95

Asn Phe Glu Gln Thr Arg Glu Ile Phe Glu Leu Leu Leu Asp Thr Ile
100 105 110

Lys Lys Leu Arg Asp Tyr Tyr Thr His His Tyr His Lys Pro Ile Thr
115 120 125

Ile Asn Pro Lys Ile Tyr Asp Phe Leu Asp Asp Thr Leu Leu Asp Val
130 135 140

Leu Ile Thr Ile Lys Lys Lys Lys Val Lys Asn Asp Thr Ser Arg Glu
145 150 155 160

Leu Leu Lys Glu Lys Leu Arg Pro Glu Leu Thr Gln Leu Lys Asn Gln
165 170 175

Lys Arg Glu Glu Leu Ile Lys Lys Gly Lys Lys Leu Leu Glu Glu Asn
180 185 190

Leu Glu Asn Ala Val Phe Asn His Cys Leu Ile Pro Phe Leu Glu Glu
195 200 205

Asn Lys Thr Asp Asp Lys Gln Asn Lys Thr Val Ser Leu Arg Lys Tyr
210 215 220

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

275 280 285
Thr His Trp Ser Tyr Ser Tyr Tyr Asn Phe Lys Gly Leu Lys His Arg
290 295 300
Ile Lys Thr Asp Gln Gly Val Ser Thr Leu Glu Gln Asn Asn Thr Thr
305 310 315 320
His Ser Leu Thr Asn Thr Asn Thr Lys Glu Ala Leu Leu Thr Gln Ile
325 330 335
Val Asp Tyr Leu Ser Lys Val Pro Asn Glu Ile Tyr Glu Thr Leu Ser

340 345 350
Glu Lys Gln Gln Lys Glu Phe Glu Glu Asp Ile Asn Glu Tyr Met Arg
355 360 365
Glu Asn Pro Glu Asn Glu Asp Ser Thr Phe Ser Ser Ile Val Ser His
370 375 380
Lys Val Ile Arg Lys Arg Tyr Glu Asn Lys Phe Asn Tyr Phe Ala Met
385 390 395 400
Arg Phe Leu Asp Glu Tyr Ala Glu Leu Pro Thr Leu Arg Phe Met Val

405 410 415
Asn Phe Gly Asp Tyr Ile Lys Asp Arg Gln Lys Lys Ile Leu Glu Ser
420 425 430
Ile Gln Phe Asp Ser Glu Arg Ile Ile Lys Lys Glu Ile His Leu Phe
435 440 445
Glu Lys Leu Ser Leu Val Thr Glu Tyr Lys Lys Asn Val Tyr Leu Lys
450 455 460
Glu Thr Ser Asn Ile Asp Leu Ser Arg Phe Pro Leu Phe Pro Asn Pro

465 470 475 480

Ser Tyr Val Met Ala Asn Asn Asn Ile Pro Phe Tyr Ile Asp Ser Arg

485 490 495

Ser Asn Asn Leu Asp Glu Tyr Leu Asn Gln Lys Lys Lys Ala Gln Ser

500 505 510

Gln Asn Lys Lys Arg Asn Leu Thr Phe Glu Lys Tyr Asn Lys Glu Gln

515 520 525

Ser Lys Asp Ala Ile Ile Ala Met Leu Gln Lys Glu Ile Gly Val Lys

530 535 540

Asp Leu Gln Gln Arg Ser Thr Ile Gly Leu Leu Ser Cys Asn Glu Leu

545 550 555 560

Pro Ser Met Leu Tyr Glu Val Ile Val Lys Asp Ile Lys Gly Ala Glu

565 570 575

Leu Glu Asn Lys Ile Ala Gln Lys Ile Arg Glu Gln Tyr Gln Ser Ile

580 585 590

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

595 600 605

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

610 615 620

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

625 630 635 640

Leu Thr Gln Glu Lys Leu Leu Asn Val Lys Glu His Glu Ile Glu Val

645 650 655

Asp Asn Tyr Asn Arg Asn Lys Asn Thr Tyr Lys Phe Lys Asn Gln Pro

660 665 670

Lys Asn Lys Val Asp Asp Lys Lys Leu Gln Arg Lys Tyr Val Phe Tyr

675 680 685

Arg Asn Glu Ile Arg Gln Glu Ala Asn Trp Leu Ala Ser Asp Leu Ile

690 695 700

His Phe Met Lys Asn Lys Ser Leu Trp Lys Gly Tyr Met His Asn Glu

705 710 715 720

Leu Gln Ser Phe Leu Ala Phe Phe Glu Asp Lys Lys Asn Asp Cys Ile

725 730 735

Ala Leu Leu Glu Thr Val Phe Asn Leu Lys Glu Asp Cys Ile Leu Thr

740 745 750

Lys Gly Leu Lys Asn Leu Phe Leu Lys His Gly Asn Phe Ile Asp Phe

755 760 765

Tyr Lys Glu Tyr Leu Lys Leu Lys Glu Asp Phe Leu Ser Thr Glu Ser

770 775 780

Thr Phe Leu Glu Asn Gly Phe Ile Gly Leu Pro Pro Lys Ile Leu Lys

785 790 795 800

Lys Glu Leu Ser Lys Arg Leu Lys Tyr Ile Phe Ile Val Phe Gln Lys

805 810 815

Arg Gln Phe Ile Ile Lys Glu Leu Glu Glu Lys Lys Asn Asn Leu Tyr

820 825 830

Ala Asp Ala Ile Asn Leu Ser Arg Gly Ile Phe Asp Glu Lys Pro Thr

835 840 845

Met Ile Pro Phe Lys Lys Pro Asn Pro Asp Glu Phe Ala Ser Trp Phe

850 855 860

Val Ala Ser Tyr Gln Tyr Asn Asn Tyr Gln Ser Phe Tyr Glu Leu Thr

865 870 875 880

Pro Asp Ile Val Glu Arg Asp Lys Lys Lys Lys Tyr Lys Asn Leu Arg

885 890 895

Ala Ile Asn Lys Val Lys Ile Gln Asp Tyr Tyr Leu Lys Leu Met Val

900 905 910

Asp Thr Leu Tyr Gln Asp Leu Phe Asn Gln Pro Leu Asp Lys Ser Leu

915 920 925

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

930 935 940

Ala Lys Ala Tyr Gln Lys Leu Asn Asp Ser Ser Leu Trp Asn Lys Val

945 950 955 960

Ile His Leu Ser Leu Gln Asn Asn Arg Ile Thr Ala Asn Pro Lys Leu

965 970 975

Lys Asp Ile Gly Lys Tyr Lys Arg Ala Leu Gln Asp Glu Lys Ile Ala

980 985 990

Thr Leu Leu Thr Tyr Asp Ala Arg Thr Trp Thr Tyr Ala Leu Gln Lys

995 1000 1005

Pro Glu Lys Glu Asn Glu Asn Asp Tyr Lys Glu Leu His Tyr Thr

1010 1015 1020

Ala Leu Asn Met Glu Leu Gln Glu Tyr Glu Lys Val Arg Ser Lys

1025 1030 1035

Glu Leu Leu Lys Gln Val Gln Glu Leu Glu Lys Lys Ile Leu Asp

1040 1045 1050

Lys Phe Tyr Asp Phe Ser Asn Asn Ala Ser His Pro Glu Asp Leu

1055 1060 1065

Glu Ile Glu Asp Lys Lys Gly Lys Arg His Pro Asn Phe Lys Leu

1070 1075 1080

Tyr Ile Thr Lys Ala Leu Leu Lys Asn Glu Ser Glu Ile Ile Asn

1085 1090 1095

Leu Glu Asn Ile Asp Ile Glu Ile Leu Leu Lys Tyr Tyr Asp Tyr

1100 1105 1110

Asn Thr Glu Glu Leu Lys Glu Lys Ile Lys Asn Met Asp Glu Asp

1115 1120 1125

Glu Lys Ala Lys Ile Ile Asn Thr Lys Glu Asn Tyr Asn Lys Ile

1130 1135 1140

Thr Asn Val Leu Ile Lys Lys Ala Leu Val Leu Ile Ile Ile Arg

1145 1150 1155

Asn Lys Met Ala His Asn Gln Tyr Pro Pro Lys Phe Ile Tyr Asp

1160 1165 1170

Leu Ala Asn Arg Phe Val Pro Lys Lys Glu Glu Glu Tyr Phe Ala

1175 1180 1185

Thr Tyr Phe Asn Arg Val Phe Glu Thr Ile Thr Lys Glu Leu Trp

1190 1195 1200

Glu Asn Lys Glu Lys Lys Asp Lys Thr Gln Val

1205 1210

<210> 86

<211> 1119

<212> PRT

<213> Porphyromonas gingivalis

<400> 86

Met Thr Glu Gln Asn Glu Lys Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Ala His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu

100 105 110

Leu Ser Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu

145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Lys Tyr Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp

210 215 220

Arg Glu Gly Thr Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Ala Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr

325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

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

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

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

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Val Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

```
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
    690                 695                 700
Arg Ser Tyr Leu Glu Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg

705                 710                 715                 720
Ser Asp Arg Val Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys
                725                 730                 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu
            740                 745                 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
        755                 760                 765
Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

    770                 775                 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Ser Lys Asp Arg Val Gln Pro
785                 790                 795                 800
Phe Tyr Asp Tyr Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
                805                 810                 815
Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly
            820                 825                 830
Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala

        835                 840                 845
Lys Glu His Pro Tyr His Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg
    850                 855                 860
Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Met Cys
865                 870                 875                 880
Arg Asp Leu Met Glu Glu Asn Lys Val Glu Gly Leu Asp Thr Gly Thr
                885                 890                 895
Leu Tyr Leu Lys Asp Ile Arg Thr Asp Val Gln Glu Gln Gly Ser Leu

            900                 905                 910
Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr
        915                 920                 925
Arg Ala Asp Ser Arg Gly His Val His Lys Glu Gln Ala Pro Leu Ala
```

930 935 940

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

945 950 955 960

Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe

965 970 975

Val Asp Thr Gly Ala Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu

980 985 990

Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala

995 1000 1005

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

1010 1015 1020

Pro His Leu Pro Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp

1025 1030 1035

Ser Asp Pro Leu Leu Asp Lys Trp Pro Asp Leu His Gly Asn Val

1040 1045 1050

Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr

1055 1060 1065

Pro Met Tyr Asp Glu Thr Leu Phe Ser Ser Ile Arg Lys Tyr Asp

1070 1075 1080

Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile

1085 1090 1095

Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Met Val

1100 1105 1110

Glu Arg Ile Ile Gln Ala

1115

<210> 87

<211> 1175

<212> PRT

<213> Porphyromonas species

<400> 87

Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Gln Asp Val Leu Ser Phe Lys Ala Leu Trp

50 55 60

Lys Asn Phe Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Glu Ser Ser Glu Leu

145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp

210 215 220

Ser Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Asp Asp Arg Ala Cys Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly

325 330 335

Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Ser Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525
Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys

530 535 540
Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Lys Gln Met Ile Gly Ile Leu Ser Gln Glu Arg Lys Asp Met Glu
565 570 575
Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Arg Ala Arg Lys Ala Phe Leu Glu Arg Ile Gly Arg
705 710 715 720
Ser Asp Arg Val Glu Asn Cys Pro Phe Leu Leu Leu Lys Glu Pro Lys

725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu
740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

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

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Asp Arg Ala Glu Glu Trp Glu Arg Gly

820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg

835 840 845

Arg Ile Lys Asp Ala Phe Ala Gly Ile Glu Tyr Ala Ser Pro Gly Asn

850 855 860

Lys Lys Lys Ile Glu Gln Leu Leu Arg Asp Leu Ser Leu Trp Glu Ala

865 870 875 880

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

885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Gln Glu His Pro Tyr His Asp Phe

900 905 910

Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln

915 920 925

Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys

930 935 940

Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Pro

945 950 955 960

Asn Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn Arg Val Lys Pro

965 970 975

Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val

980 985 990

His Lys Glu Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp

995 1000 1005

Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp

1010 1015 1020

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

1025 1030 1035

Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu

1040 1045 1050

Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Val Phe Glu Leu Thr

1055 1060 1065

Leu Arg Leu Glu Glu Ser Leu Leu Ser Arg Tyr Pro His Leu Pro

1070 1075 1080

Asp Glu Ser Phe Arg Glu Met Leu Glu Ser Trp Ser Asp Pro Leu

1085 1090 1095

Leu Ala Lys Trp Pro Glu Leu His Gly Lys Val Arg Leu Leu Ile

1100 1105 1110

Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp

1115 1120 1125

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

1130 1135 1140

Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu

1145 1150 1155

Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val Glu Arg Ile Ile

1160 1165 1170

Gln Ala

1175

<210> 88

<211> 1132

<212> PRT

<213> Prevotella intermedia

<400> 88

Met Glu Asp Asp Lys Lys Thr Lys Glu Ser Thr Asn Met Leu Asp Asn

1 5 10 15

Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val Tyr

20 25 30

Ile Thr Val Asn His Ile Asn Lys Val Leu Glu Leu Lys Asn Lys Lys

35 40 45

Asp Gln Asp Ile Ile Ile Asp Asn Asp Gln Asp Ile Leu Ala Ile Lys

50 55 60

Thr His Trp Glu Lys Val Asn Gly Asp Leu Asn Lys Thr Glu Arg Leu

65 70 75 80

Arg Glu Leu Met Thr Lys His Phe Pro Phe Leu Glu Thr Ala Ile Tyr

85 90 95

Thr Lys Asn Lys Glu Asp Lys Glu Glu Val Lys Gln Glu Lys Gln Ala

100 105 110

Lys Ala Gln Ser Phe Asp Ser Leu Lys His Cys Leu Phe Leu Phe Leu

115 120 125

Glu Lys Leu Gln Glu Ala Arg Asn Tyr Tyr Ser His Tyr Lys Tyr Ser

130 135 140

Glu Ser Thr Lys Glu Pro Met Leu Glu Lys Glu Leu Leu Lys Lys Met

145 150 155 160

Tyr Asn Ile Phe Asp Asp Asn Ile Gln Leu Val Ile Lys Asp Tyr Gln

165 170 175

His Asn Lys Asp Ile Asn Pro Asp Glu Asp Phe Lys His Leu Asp Arg

180 185 190

Thr Glu Glu Glu Phe Asn Tyr Tyr Phe Thr Thr Asn Lys Lys Gly Asn

195 200 205

Ile Thr Ala Ser Gly Leu Leu Phe Phe Val Ser Leu Phe Leu Glu Lys

210 215 220

Lys Asp Ala Ile Trp Met Gln Gln Lys Leu Arg Gly Phe Lys Asp Asn

225 230 235 240

Arg Glu Ser Lys Lys Lys Met Thr His Glu Val Phe Cys Arg Ser Arg

245 250 255

Met Leu Leu Pro Lys Leu Arg Leu Glu Ser Thr Gln Thr Gln Asp Trp

260 265 270

Ile Leu Leu Asp Met Leu Asn Glu Leu Ile Arg Cys Pro Lys Ser Leu
275 280 285
Tyr Glu Arg Leu Gln Gly Glu Tyr Arg Lys Lys Phe Asn Val Pro Phe
290 295 300

Asp Ser Ala Asp Glu Asp Tyr Asp Ala Glu Gln Glu Pro Phe Lys Asn
305 310 315 320
Thr Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Ala Leu Arg Tyr
325 330 335
Phe Asp Tyr Asn Glu Ile Phe Thr Asn Leu Arg Phe Gln Ile Asp Leu
340 345 350
Gly Thr Tyr His Phe Ser Ile Tyr Lys Lys Leu Ile Gly Gly Gln Lys
355 360 365

Glu Asp Arg His Leu Thr His Lys Leu Tyr Gly Phe Glu Arg Ile Gln
370 375 380
Glu Phe Ala Lys Gln Asn Arg Thr Asp Glu Trp Lys Ala Ile Val Lys
385 390 395 400
Asp Phe Asp Thr Tyr Glu Thr Ser Glu Glu Pro Tyr Ile Ser Glu Thr
405 410 415
Ala Pro His Tyr His Leu Glu Asn Gln Lys Ile Gly Ile Arg Phe Arg
420 425 430

Asn Asp Asn Asp Glu Ile Trp Pro Ser Leu Lys Thr Asn Gly Glu Asn
435 440 445
Asn Glu Lys Arg Lys Tyr Lys Leu Asp Lys Gln Tyr Gln Ala Glu Ala
450 455 460
Phe Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu
465 470 475 480
Leu Lys Lys Glu Glu Pro Asn Asn Asp Lys Lys Asn Ala Ser Ile Val
485 490 495

Glu Gly Phe Ile Lys Arg Glu Ile Arg Asp Ile Tyr Lys Leu Tyr Asp
500 505 510
Ala Phe Ala Asn Gly Glu Ile Asn Asn Ile Asp Asp Leu Glu Lys Tyr

515 520 525

Cys Glu Asp Lys Gly Ile Pro Lys Arg His Leu Pro Lys Gln Met Val

530 535 540

Ala Ile Leu Tyr Asp Glu His Lys Asp Met Ala Glu Glu Ala Lys Arg

545 550 555 560

Lys Gln Lys Glu Met Val Lys Asp Thr Lys Lys Leu Leu Ala Thr Leu

565 570 575

Glu Lys Gln Thr Gln Gly Glu Ile Glu Asp Gly Gly Arg Asn Ile Arg

580 585 590

Leu Leu Lys Ser Gly Glu Ile Ala Arg Trp Leu Val Asn Asp Met Met

595 600 605

Arg Phe Gln Pro Val Gln Lys Asp Asn Glu Gly Asn Pro Leu Asn Asn

610 615 620

Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met Leu Gln Arg Ser Leu Ala

625 630 635 640

Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg Tyr Phe Arg Gln Val Asn

645 650 655

Leu Ile Asn Ser Ser Asn Pro His Pro Phe Leu Lys Trp Thr Lys Trp

660 665 670

Glu Glu Cys Asn Asn Ile Leu Ser Phe Tyr Arg Ser Tyr Leu Thr Lys

675 680 685

Lys Ile Glu Phe Leu Asn Lys Leu Lys Pro Glu Asp Trp Glu Lys Asn

690 695 700

Gln Tyr Phe Leu Lys Leu Lys Glu Pro Lys Thr Asn Arg Glu Thr Leu

705 710 715 720

Val Gln Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg Gly Ile Phe Thr

725 730 735

Glu Pro Ile Arg Glu Trp Phe Lys Arg His Gln Asn Asp Ser Glu Glu

740 745 750

Tyr Glu Lys Val Glu Thr Leu Asp Arg Val Gly Leu Val Thr Lys Val

755 760 765

Ile Pro Leu Phe Phe Lys Lys Glu Asp Ser Lys Asp Lys Glu Glu Tyr
770 775 780
Leu Lys Lys Asp Ala Gln Lys Glu Ile Asn Asn Cys Val Gln Pro Phe
785 790 795 800
Tyr Gly Phe Pro Tyr Asn Val Gly Asn Ile His Lys Pro Asp Glu Lys
805 810 815

Asp Phe Leu Pro Ser Glu Glu Arg Lys Lys Leu Trp Gly Asp Lys Lys
820 825 830
Tyr Lys Phe Lys Gly Tyr Lys Ala Lys Val Lys Ser Lys Lys Leu Thr
835 840 845
Asp Lys Glu Lys Glu Glu Tyr Arg Ser Tyr Leu Glu Phe Gln Ser Trp
850 855 860
Asn Lys Phe Glu Arg Glu Leu Arg Leu Val Arg Asn Gln Asp Ile Val
865 870 875 880

Thr Trp Leu Leu Cys Thr Glu Leu Ile Asp Lys Leu Lys Val Glu Gly
885 890 895
Leu Asn Val Glu Glu Leu Lys Lys Leu Arg Leu Lys Asp Ile Asp Thr
900 905 910
Asp Thr Ala Lys Gln Glu Lys Asn Asn Ile Leu Asn Arg Val Met Pro
915 920 925
Met Gln Leu Pro Val Thr Val Tyr Glu Ile Asp Asp Ser His Asn Ile
930 935 940

Val Lys Asp Arg Pro Leu His Thr Val Tyr Ile Glu Glu Thr Lys Thr
945 950 955 960
Lys Leu Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys Asp Arg Arg
965 970 975
Leu Asn Gly Leu Phe Ser Phe Val Asp Thr Ser Ser Glu Thr Glu Leu
980 985 990
Lys Ser Asn Pro Ile Ser Lys Ser Leu Val Glu Tyr Glu Leu Gly Glu
995 1000 1005

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

1010 1015 1020

Glu Glu Thr Leu Ile Glu Lys Tyr Lys Thr Leu Pro Thr Asp Asn

1025 1030 1035

Phe Ser Asp Met Leu Asn Gly Trp Leu Glu Gly Lys Asp Glu Ala

1040 1045 1050

Asp Lys Ala Arg Phe Gln Asn Asp Val Lys Leu Leu Val Ala Val

1055 1060 1065

Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Arg Asn Arg Ile

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

Thr His Lys Ile Ile Lys Arg Ile Ile Glu Ile Glu Lys Pro Ile

1115 1120 1125

Glu Thr Lys Glu

1130

<210> 89

<211> 1134

<212> PRT

<213> Prevotella intermedia

<400> 89

Met Lys Met Glu Asp Asp Lys Lys Thr Lys Glu Ser Thr Asn Met Leu

1 5 10 15

Asp Asn Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Val Tyr Ile Thr Val Asn His Ile Asn Lys Val Leu Glu Leu Lys Asn

35 40 45

Lys Lys Asp Gln Asp Ile Ile Ile Asp Asn Asp Gln Asp Ile Leu Ala

50 55 60

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

65 70 75 80

Arg Leu Arg Glu Leu Met Thr Lys His Phe Pro Phe Leu Glu Thr Ala
85 90 95
Ile Tyr Thr Lys Asn Lys Glu Asp Lys Glu Glu Val Lys Gln Glu Lys
100 105 110

Gln Ala Lys Ala Gln Ser Phe Asp Ser Leu Lys His Cys Leu Phe Leu
115 120 125
Phe Leu Glu Lys Leu Gln Glu Ala Arg Asn Tyr Tyr Ser His Tyr Lys
130 135 140
Tyr Ser Glu Ser Thr Lys Glu Pro Met Leu Glu Lys Glu Leu Leu Lys
145 150 155 160
Lys Met Tyr Asn Ile Phe Asp Asp Asn Ile Gln Leu Val Ile Lys Asp
165 170 175

Tyr Gln His Asn Lys Asp Ile Asn Pro Asp Glu Asp Phe Lys His Leu
180 185 190
Asp Arg Thr Glu Glu Glu Phe Asn Tyr Tyr Phe Thr Thr Asn Lys Lys
195 200 205
Gly Asn Ile Thr Ala Ser Gly Leu Leu Phe Phe Val Ser Leu Phe Leu
210 215 220
Glu Lys Lys Asp Ala Ile Trp Met Gln Gln Lys Leu Arg Gly Phe Lys
225 230 235 240

Asp Asn Arg Glu Ser Lys Lys Lys Met Thr His Glu Val Phe Cys Arg
245 250 255
Ser Arg Met Leu Leu Pro Lys Leu Arg Leu Glu Ser Thr Gln Thr Gln
260 265 270
Asp Trp Ile Leu Leu Asp Met Leu Asn Glu Leu Ile Arg Cys Pro Lys
275 280 285
Ser Leu Tyr Glu Arg Leu Gln Gly Glu Tyr Arg Lys Lys Phe Asn Val
290 295 300

Pro Phe Asp Ser Ala Asp Glu Asp Tyr Asp Ala Glu Gln Glu Pro Phe
305 310 315 320
Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Ala Leu

325 330 335

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

340 345 350

Asp Leu Gly Thr Tyr His Phe Ser Ile Tyr Lys Lys Leu Ile Gly Gly

355 360 365

Gln Lys Glu Asp Arg His Leu Thr His Lys Leu Tyr Gly Phe Glu Arg

370 375 380

Ile Gln Glu Phe Ala Lys Gln Asn Arg Thr Asp Glu Trp Lys Ala Ile

385 390 395 400

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

405 410 415

Glu Thr Ala Pro His Tyr His Leu Glu Asn Gln Lys Ile Gly Ile Arg

420 425 430

Phe Arg Asn Asp Asn Asp Glu Ile Trp Pro Ser Leu Lys Thr Asn Gly

435 440 445

Glu Asn Asn Glu Lys Arg Lys Tyr Lys Leu Asp Lys Gln Tyr Gln Ala

450 455 460

Glu Ala Phe Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr

465 470 475 480

Leu Leu Leu Lys Lys Glu Glu Pro Asn Asn Asp Lys Lys Asn Ala Ser

485 490 495

Ile Val Glu Gly Phe Ile Lys Arg Glu Ile Arg Asp Ile Tyr Lys Leu

500 505 510

Tyr Asp Ala Phe Ala Asn Gly Glu Ile Asn Asn Ile Asp Asp Leu Glu

515 520 525

Lys Tyr Cys Glu Asp Lys Gly Ile Pro Lys Arg His Leu Pro Lys Gln

530 535 540

Met Val Ala Ile Leu Tyr Asp Glu His Lys Asp Met Ala Glu Glu Ala

545 550 555 560

Lys Arg Lys Gln Lys Glu Met Val Lys Asp Thr Lys Lys Leu Leu Ala

565 570 575

Thr Leu Glu Lys Gln Thr Gln Gly Glu Ile Glu Asp Gly Gly Arg Asn
580 585 590
Ile Arg Leu Leu Lys Ser Gly Glu Ile Ala Arg Trp Leu Val Asn Asp
595 600 605
Met Met Arg Phe Gln Pro Val Gln Lys Asp Asn Glu Gly Asn Pro Leu
610 615 620

Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met Leu Gln Arg Ser
625 630 635 640
Leu Ala Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg Tyr Phe Arg Gln
645 650 655
Val Asn Leu Ile Asn Ser Ser Asn Pro His Pro Phe Leu Lys Trp Thr
660 665 670
Lys Trp Glu Glu Cys Asn Asn Ile Leu Ser Phe Tyr Arg Ser Tyr Leu
675 680 685

Thr Lys Lys Ile Glu Phe Leu Asn Lys Leu Lys Pro Glu Asp Trp Glu
690 695 700
Lys Asn Gln Tyr Phe Leu Lys Leu Lys Glu Pro Lys Thr Asn Arg Glu
705 710 715 720
Thr Leu Val Gln Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg Gly Ile
725 730 735
Phe Thr Glu Pro Ile Arg Glu Trp Phe Lys Arg His Gln Asn Asp Ser
740 745 750

Glu Glu Tyr Glu Lys Val Glu Thr Leu Asp Arg Val Gly Leu Val Thr
755 760 765
Lys Val Ile Pro Leu Phe Phe Lys Lys Glu Asp Ser Lys Asp Lys Glu
770 775 780
Glu Tyr Leu Lys Lys Asp Ala Gln Lys Glu Ile Asn Asn Cys Val Gln
785 790 795 800
Pro Phe Tyr Gly Phe Pro Tyr Asn Val Gly Asn Ile His Lys Pro Asp
805 810 815

Glu Lys Asp Phe Leu Pro Ser Glu Glu Arg Lys Lys Leu Trp Gly Asp

820 825 830

Lys Lys Tyr Lys Phe Lys Gly Tyr Lys Ala Lys Val Lys Ser Lys Lys

835 840 845

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

850 855 860

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

865 870 875 880

Ile Val Thr Trp Leu Leu Cys Thr Glu Leu Ile Asp Lys Leu Lys Val

885 890 895

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

900 905 910

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

915 920 925

Met Pro Met Gln Leu Pro Val Thr Val Tyr Glu Ile Asp Asp Ser His

930 935 940

Asn Ile Val Lys Asp Arg Pro Leu His Thr Val Tyr Ile Glu Glu Thr

945 950 955 960

Lys Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys Asp

965 970 975

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

980 985 990

Glu Leu Lys Ser Asn Pro Ile Ser Lys Ser Leu Val Glu Tyr Glu Leu

995 1000 1005

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

1010 1015 1020

Leu Leu Glu Glu Thr Leu Ile Glu Lys Tyr Lys Thr Leu Pro Thr

1025 1030 1035

Asp Asn Phe Ser Asp Met Leu Asn Gly Trp Leu Glu Gly Lys Asp

1040 1045 1050

Glu Ala Asp Lys Ala Arg Phe Gln Asn Asp Val Lys Leu Leu Val

1055 1060 1065

Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Arg Asn
1070 1075 1080
Arg Ile Ala Phe Ala Asn Ile Asn Pro Phe Ser Leu Ser Ser Ala
1085 1090 1095
Asp Thr Ser Glu Glu Lys Lys Leu Asp Ile Ala Asn Gln Leu Lys
1100 1105 1110
Asp Lys Thr His Lys Ile Ile Lys Arg Ile Ile Glu Ile Glu Lys
1115 1120 1125

Pro Ile Glu Thr Lys Glu
1130

<210> 90
<211> 1133
<212> PRT
<213> Prevotella intermedia
<400> 90

Met Glu Asp Asp Lys Lys Thr Thr Asp Ser Ile Ser Tyr Glu Leu Lys
1 5 10 15
Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val
20 25 30
Tyr Ile Thr Val Asn His Ile Asn Lys Val Leu Glu Leu Lys Asn Lys
35 40 45

Lys Asp Gln Asp Ile Ile Ile Asp Asn Asp Gln Asp Ile Leu Ala Ile
50 55 60
Lys Thr His Trp Glu Lys Val Asn Gly Asp Leu Asn Lys Thr Glu Arg
65 70 75 80
Leu Arg Glu Leu Met Thr Lys His Phe Pro Phe Leu Glu Thr Ala Ile
85 90 95
Tyr Ser Lys Asn Lys Glu Asp Lys Glu Glu Val Lys Gln Glu Lys Gln
100 105 110

Ala Lys Ala Gln Ser Phe Asp Ser Leu Lys His Cys Leu Phe Leu Phe
115 120 125
Leu Glu Lys Leu Gln Glu Thr Arg Asn Tyr Tyr Ser His Tyr Lys Tyr

130 135 140

Ser Glu Ser Thr Lys Glu Pro Met Leu Glu Lys Glu Leu Leu Lys Lys

145 150 155 160

Met Tyr Asn Ile Phe Asp Asp Asn Ile Gln Leu Val Ile Lys Asp Tyr

165 170 175

Gln His Asn Lys Asp Ile Asn Pro Asp Glu Asp Phe Lys His Leu Asp

180 185 190

Arg Thr Glu Glu Asp Phe Asn Tyr Tyr Phe Thr Arg Asn Lys Lys Gly

195 200 205

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

210 215 220

Lys Lys Asp Ala Ile Trp Met Gln Gln Lys Leu Arg Gly Phe Lys Asp

225 230 235 240

Asn Arg Glu Ser Lys Lys Lys Met Thr His Glu Val Phe Cys Arg Ser

245 250 255

Arg Met Leu Leu Pro Lys Leu Arg Leu Glu Ser Thr Gln Thr Gln Asp

260 265 270

Trp Ile Leu Leu Asp Met Leu Asn Glu Leu Ile Arg Cys Pro Lys Ser

275 280 285

Leu Tyr Glu Arg Leu Gln Gly Glu Asp Arg Glu Lys Phe Lys Val Pro

290 295 300

Phe Asp Pro Ala Asp Glu Asp Tyr Asp Ala Glu Gln Glu Pro Phe Lys

305 310 315 320

Asn Thr Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Ala Leu Arg

325 330 335

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

340 345 350

Leu Gly Thr Phe His Phe Ser Ile Tyr Lys Lys Leu Ile Gly Gly Gln

355 360 365

Lys Glu Asp Arg His Leu Thr His Lys Leu Tyr Gly Phe Glu Arg Ile

370 375 380

Gln Glu Phe Ala Lys Gln Asn Arg Pro Asp Glu Trp Lys Ala Ile Val
385 390 395 400
Lys Asp Leu Asp Thr Tyr Glu Thr Ser Asn Glu Arg Tyr Ile Ser Glu
405 410 415
Thr Thr Pro His Tyr His Leu Glu Asn Gln Lys Ile Gly Ile Arg Phe
420 425 430

Arg Asn Asp Asn Asp Glu Ile Trp Pro Ser Leu Lys Thr Asn Gly Glu
435 440 445
Asn Asn Glu Lys Ser Lys Tyr Lys Leu Asp Lys Gln Tyr Gln Ala Glu
450 455 460
Ala Phe Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu
465 470 475 480
Leu Leu Lys Lys Glu Glu Pro Asn Asn Asp Lys Lys Asn Ala Ser Ile
485 490 495

Val Glu Gly Phe Ile Lys Arg Glu Ile Arg Asp Met Tyr Lys Leu Tyr
500 505 510
Asp Ala Phe Ala Asn Gly Glu Ile Asn Asn Ile Asp Asp Leu Glu Lys
515 520 525
Tyr Cys Glu Asp Lys Gly Ile Pro Lys Arg His Leu Pro Lys Gln Met
530 535 540
Val Ala Ile Leu Tyr Asp Glu His Lys Asp Met Val Lys Glu Ala Lys
545 550 555 560

Arg Lys Gln Arg Lys Met Val Lys Asp Thr Glu Lys Leu Leu Ala Ala
565 570 575
Leu Glu Lys Gln Thr Gln Glu Lys Thr Glu Asp Gly Gly Arg Asn Ile
580 585 590
Arg Leu Leu Lys Ser Gly Glu Ile Ala Arg Trp Leu Val Asn Asp Met
595 600 605
Met Arg Phe Gln Pro Val Gln Lys Asp Asn Glu Gly Asn Pro Leu Asn
610 615 620

Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met Leu Gln Arg Ser Leu

625 630 635 640

Ala Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg Tyr Phe Arg Gln Val

645 650 655

Asn Leu Ile Asn Ser Ser Asn Pro His Pro Phe Leu Lys Trp Thr Lys

660 665 670

Trp Glu Glu Cys Asn Asn Ile Leu Ser Phe Tyr Arg Ser Tyr Leu Thr

675 680 685

Lys Lys Ile Glu Phe Leu Asn Lys Leu Lys Pro Glu Asp Trp Glu Lys

690 695 700

Asn Gln Tyr Phe Leu Lys Leu Lys Glu Pro Lys Thr Asn Arg Glu Thr

705 710 715 720

Leu Val Gln Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg Gly Ile Phe

725 730 735

Thr Glu Pro Ile Arg Glu Trp Phe Lys Arg His Gln Asn Asp Ser Lys

740 745 750

Glu Tyr Glu Lys Val Glu Ala Leu Asp Arg Val Gly Leu Val Thr Lys

755 760 765

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

770 775 780

Asp Leu Lys Lys Asp Ala Gln Lys Glu Ile Asn Asn Cys Val Gln Pro

785 790 795 800

Phe Tyr Ser Phe Pro Tyr Asn Val Gly Asn Ile His Lys Pro Asp Glu

805 810 815

Lys Asp Phe Leu His Arg Glu Glu Arg Ile Glu Leu Trp Asp Lys Lys

820 825 830

Lys Asp Lys Phe Lys Gly Tyr Lys Ala Lys Val Lys Ser Lys Lys Leu

835 840 845

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

850 855 860

Trp Asn Lys Phe Glu Arg Glu Leu Arg Leu Val Arg Asn Gln Asp Ile

865 870 875 880

Val Thr Trp Leu Leu Cys Thr Glu Leu Ile Asp Lys Leu Lys Val Glu
885 890 895
Gly Leu Asn Val Glu Glu Leu Lys Lys Leu Arg Leu Lys Asp Ile Asp
900 905 910
Thr Asp Thr Ala Lys Gln Glu Lys Asn Asn Ile Leu Asn Arg Val Met
915 920 925
Pro Met Gln Leu Pro Val Thr Val Tyr Glu Ile Asp Asp Ser His Asn
930 935 940

Ile Val Lys Asp Arg Pro Leu His Thr Val Tyr Ile Glu Glu Thr Lys
945 950 955 960
Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys Asp Arg
965 970 975
Arg Leu Asn Gly Leu Phe Ser Phe Val Asp Thr Ser Ser Glu Ala Glu
980 985 990
Leu Lys Ser Asn Pro Ile Ser Lys Ser Leu Val Glu Tyr Glu Leu Gly
995 1000 1005

Glu Tyr Gln Asn Ala Arg Ile Glu Thr Ile Lys Asp Met Leu Leu
1010 1015 1020
Leu Glu Glu Thr Leu Ile Glu Lys Tyr Lys Asn Leu Pro Thr Asp
1025 1030 1035
Asn Phe Ser Asp Met Leu Asn Gly Trp Leu Glu Gly Lys Asp Glu
1040 1045 1050
Ala Asp Lys Ala Arg Phe Gln Asn Asp Val Lys Leu Leu Val Ala
1055 1060 1065

Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Arg Asn Arg
1070 1075 1080
Ile Ala Phe Ala Asn Ile Asn Pro Phe Ser Leu Ser Ser Ala Asp
1085 1090 1095
Thr Ser Glu Glu Lys Lys Leu Asp Ile Ala Asn Gln Leu Lys Asp
1100 1105 1110
Lys Thr His Lys Ile Ile Lys Arg Ile Ile Glu Ile Glu Lys Pro

1115 1120 1125

Ile Glu Thr Lys Glu
1130

<210> 91
<211> 1127
<212> PRT
<213> Prevotella buccae
<400> 91

Met Gln Lys Gln Asp Lys Leu Phe Val Asp Arg Lys Lys Asn Ala Ile
1 5 10 15
Phe Ala Phe Pro Lys Tyr Ile Thr Ile Met Glu Asn Lys Glu Lys Pro
20 25 30
Glu Pro Ile Tyr Tyr Glu Leu Thr Asp Lys His Phe Trp Ala Ala Phe
35 40 45

Leu Asn Leu Ala Arg His Asn Val Tyr Thr Thr Ile Asn His Ile Asn
50 55 60
Arg Arg Leu Glu Ile Ala Glu Leu Lys Asp Asp Gly Tyr Met Met Gly
65 70 75 80
Ile Lys Gly Ser Trp Asn Glu Gln Ala Lys Lys Leu Asp Lys Lys Val
85 90 95
Arg Leu Arg Asp Leu Ile Met Lys His Phe Pro Phe Leu Glu Ala Ala
100 105 110

Ala Tyr Glu Met Thr Asn Ser Lys Ser Pro Asn Asn Lys Glu Gln Arg
115 120 125
Glu Lys Glu Gln Ser Glu Ala Leu Ser Leu Asn Asn Leu Lys Asn Val
130 135 140
Leu Phe Ile Phe Leu Glu Lys Leu Gln Val Leu Arg Asn Tyr Tyr Ser
145 150 155 160
His Tyr Lys Tyr Ser Glu Glu Ser Pro Lys Pro Ile Phe Glu Thr Ser
165 170 175

Leu Leu Lys Asn Met Tyr Lys Val Phe Asp Ala Asn Val Arg Leu Val

180 185 190

Lys Arg Asp Tyr Met His His Glu Asn Ile Asp Met Gln Arg Asp Phe

195 200 205

Thr His Leu Asn Arg Lys Lys Gln Val Gly Arg Thr Lys Asn Ile Ile

210 215 220

Asp Ser Pro Asn Phe His Tyr His Phe Ala Asp Lys Glu Gly Asn Met

225 230 235 240

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

245 250 255

Asp Ala Ile Trp Met Gln Lys Lys Leu Lys Gly Phe Lys Asp Gly Arg

260 265 270

Asn Leu Arg Glu Gln Met Thr Asn Glu Val Phe Cys Arg Ser Arg Ile

275 280 285

Ser Leu Pro Lys Leu Lys Leu Glu Asn Val Gln Thr Lys Asp Trp Met

290 295 300

Gln Leu Asp Met Leu Asn Glu Leu Val Arg Cys Pro Lys Ser Leu Tyr

305 310 315 320

Glu Arg Leu Arg Glu Lys Asp Arg Glu Ser Phe Lys Val Pro Phe Asp

325 330 335

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

340 345 350

Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Val Leu Arg Tyr Phe

355 360 365

Asp Leu Asn Glu Ile Phe Glu Gln Leu Arg Phe Gln Ile Asp Leu Gly

370 375 380

Thr Tyr His Phe Ser Ile Tyr Asn Lys Arg Ile Gly Asp Glu Asp Glu

385 390 395 400

Val Arg His Leu Thr His His Leu Tyr Gly Phe Ala Arg Ile Gln Asp

405 410 415

Phe Ala Pro Gln Asn Gln Pro Glu Glu Trp Arg Lys Leu Val Lys Asp

420 425 430

Leu Asp His Phe Glu Thr Ser Gln Glu Pro Tyr Ile Ser Lys Thr Ala
435 440 445
Pro His Tyr His Leu Glu Asn Glu Lys Ile Gly Ile Lys Phe Cys Ser
450 455 460
Ala His Asn Asn Leu Phe Pro Ser Leu Gln Thr Asp Lys Thr Cys Asn
465 470 475 480
Gly Arg Ser Lys Phe Asn Leu Gly Thr Gln Phe Thr Ala Glu Ala Phe
485 490 495

Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu
500 505 510
Thr Lys Asp Tyr Ser Arg Lys Glu Ser Ala Asp Lys Val Glu Gly Ile
515 520 525
Ile Arg Lys Glu Ile Ser Asn Ile Tyr Ala Ile Tyr Asp Ala Phe Ala
530 535 540
Asn Asn Glu Ile Asn Ser Ile Ala Asp Leu Thr Arg Arg Leu Gln Asn
545 550 555 560

Thr Asn Ile Leu Gln Gly His Leu Pro Lys Gln Met Ile Ser Ile Leu
565 570 575
Lys Gly Arg Gln Lys Asp Met Gly Lys Glu Ala Glu Arg Lys Ile Gly
580 585 590
Glu Met Ile Asp Asp Thr Gln Arg Arg Leu Asp Leu Leu Cys Lys Gln
595 600 605
Thr Asn Gln Lys Ile Arg Ile Gly Lys Arg Asn Ala Gly Leu Leu Lys
610 615 620

Ser Gly Lys Ile Ala Asp Trp Leu Val Asn Asp Met Met Arg Phe Gln
625 630 635 640
Pro Val Gln Lys Asp Gln Asn Asn Ile Pro Ile Asn Asn Ser Lys Ala
645 650 655
Asn Ser Thr Glu Tyr Arg Met Leu Gln Arg Ala Leu Ala Leu Phe Gly
660 665 670
Ser Glu Asn Phe Arg Leu Lys Ala Tyr Phe Asn Gln Met Asn Leu Val

675 680 685

Gly Asn Asp Asn Pro His Pro Phe Leu Ala Glu Thr Gln Trp Glu His
690 695 700
Gln Thr Asn Ile Leu Ser Phe Tyr Arg Asn Tyr Leu Glu Ala Arg Lys
705 710 715 720
Lys Tyr Leu Lys Gly Leu Lys Pro Gln Asn Trp Lys Gln Tyr Gln His
725 730 735
Phe Leu Ile Leu Lys Val Gln Lys Thr Asn Arg Asn Thr Leu Val Thr
740 745 750

Gly Trp Lys Asn Ser Phe Asn Leu Pro Arg Gly Ile Phe Thr Gln Pro
755 760 765
Ile Arg Glu Trp Phe Glu Lys His Asn Asn Ser Lys Arg Ile Tyr Asp
770 775 780
Gln Ile Leu Ser Phe Asp Arg Val Gly Phe Val Ala Lys Ala Ile Pro
785 790 795 800
Leu Tyr Phe Ala Glu Glu Tyr Lys Asp Asn Val Gln Pro Phe Tyr Asp
805 810 815

Tyr Pro Phe Asn Ile Gly Asn Arg Leu Lys Pro Lys Lys Arg Gln Phe
820 825 830
Leu Asp Lys Lys Glu Arg Val Glu Leu Trp Gln Lys Asn Lys Glu Leu
835 840 845
Phe Lys Asn Tyr Pro Ser Glu Lys Lys Lys Thr Asp Leu Ala Tyr Leu
850 855 860
Asp Phe Leu Ser Trp Lys Lys Phe Glu Arg Glu Leu Arg Leu Ile Lys
865 870 875 880

Asn Gln Asp Ile Val Thr Trp Leu Met Phe Lys Glu Leu Phe Asn Met
885 890 895
Ala Thr Val Glu Gly Leu Lys Ile Gly Glu Ile His Leu Arg Asp Ile
900 905 910
Asp Thr Asn Thr Ala Asn Glu Glu Ser Asn Asn Ile Leu Asn Arg Ile
915 920 925

Met Pro Met Lys Leu Pro Val Lys Thr Tyr Glu Thr Asp Asn Lys Gly
930 935 940

Asn Ile Leu Lys Glu Arg Pro Leu Ala Thr Phe Tyr Ile Glu Glu Thr
945 950 955 960
Glu Thr Lys Val Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys Asp
965 970 975
Arg Arg Leu Asn Gly Leu Phe Ser Phe Ala Glu Thr Thr Asp Leu Asn
980 985 990
Leu Glu Glu His Pro Ile Ser Lys Leu Ser Val Asp Leu Glu Leu Ile
995 1000 1005

Lys Tyr Gln Thr Thr Arg Ile Ser Ile Phe Glu Met Thr Leu Gly
1010 1015 1020
Leu Glu Lys Lys Leu Ile Asp Lys Tyr Ser Thr Leu Pro Thr Asp
1025 1030 1035
Ser Phe Arg Asn Met Leu Glu Arg Trp Leu Gln Cys Lys Ala Asn
1040 1045 1050
Arg Pro Glu Leu Lys Asn Tyr Val Asn Ser Leu Ile Ala Val Arg
1055 1060 1065

Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp Ala Thr Leu
1070 1075 1080
Phe Ala Glu Val Lys Lys Phe Thr Leu Phe Pro Ser Val Asp Thr
1085 1090 1095
Lys Lys Ile Glu Leu Asn Ile Ala Pro Gln Leu Leu Glu Ile Val
1100 1105 1110
Gly Lys Ala Ile Lys Glu Ile Glu Lys Ser Glu Asn Lys Asn
1115 1120 1125

<210> 92
<211> 1125
<212> PRT
<213> Prevotella pallens
<400> 92

Met Lys Glu Glu Glu Lys Gly Lys Thr Pro Val Val Ser Thr Tyr Asn
1 5 10 15
Lys Asp Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His
20 25 30
Asn Val Tyr Ile Thr Val Asn His Ile Asn Lys Ile Leu Gly Glu Gly
35 40 45
Glu Ile Asn Arg Asp Gly Tyr Glu Asn Thr Leu Glu Lys Ser Trp Asn

50 55 60
Glu Ile Lys Asp Ile Asn Lys Lys Asp Arg Leu Ser Lys Leu Ile Ile
65 70 75 80
Lys His Phe Pro Phe Leu Glu Val Thr Thr Tyr Gln Arg Asn Ser Ala
85 90 95
Asp Thr Thr Lys Gln Lys Glu Glu Lys Gln Ala Glu Ala Gln Ser Leu
100 105 110
Glu Ser Leu Lys Lys Ser Phe Phe Val Phe Ile Tyr Lys Leu Arg Asp

115 120 125
Leu Arg Asn His Tyr Ser His Tyr Lys His Ser Lys Ser Leu Glu Arg
130 135 140
Pro Lys Phe Glu Glu Asp Leu Gln Glu Lys Met Tyr Asn Ile Phe Asp
145 150 155 160
Ala Ser Ile Gln Leu Val Lys Glu Asp Tyr Lys His Asn Thr Asp Ile
165 170 175
Lys Thr Glu Glu Asp Phe Lys His Leu Asp Arg Lys Gly Gln Phe Lys

180 185 190
Tyr Ser Phe Ala Asp Asn Glu Gly Asn Ile Thr Glu Ser Gly Leu Leu
195 200 205
Phe Phe Val Ser Leu Phe Leu Glu Lys Lys Asp Ala Ile Trp Val Gln
210 215 220
Lys Lys Leu Glu Gly Phe Lys Cys Ser Asn Glu Ser Tyr Gln Lys Met
225 230 235 240
Thr Asn Glu Val Phe Cys Arg Ser Arg Met Leu Leu Pro Lys Leu Arg

245 250 255

Leu Gln Ser Thr Gln Thr Gln Asp Trp Ile Leu Leu Asp Met Leu Asn

260 265 270

Glu Leu Ile Arg Cys Pro Lys Ser Leu Tyr Glu Arg Leu Arg Glu Glu

275 280 285

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

290 295 300

Asp Ala Glu Gln Glu Pro Phe Lys Asn Ala Leu Val Arg His Gln Asp

305 310 315 320

Arg Phe Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Tyr Asn Glu Ile Phe

325 330 335

Thr Asn Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ser Ile

340 345 350

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

355 360 365

Lys Leu Tyr Gly Phe Glu Arg Ile Gln Glu Phe Thr Lys Gln Asn Arg

370 375 380

Pro Asp Glu Trp Arg Lys Phe Val Lys Thr Phe Asn Ser Phe Glu Thr

385 390 395 400

Ser Lys Glu Pro Tyr Ile Pro Glu Thr Thr Pro His Tyr His Leu Glu

405 410 415

Asn Gln Lys Ile Gly Ile Arg Phe Arg Asn Asp Asn Asp Lys Ile Trp

420 425 430

Pro Ser Leu Lys Thr Asn Ser Glu Lys Asn Glu Lys Ser Lys Tyr Lys

435 440 445

Leu Asp Lys Ser Phe Gln Ala Glu Ala Phe Leu Ser Val His Glu Leu

450 455 460

Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu Lys Thr Glu Asn Thr Asp

465 470 475 480

Asn Asp Asn Glu Ile Glu Thr Lys Lys Lys Glu Asn Lys Asn Asp Lys

485 490 495

Gln Glu Lys His Lys Ile Glu Glu Ile Ile Glu Asn Lys Ile Thr Glu

500 505 510

Ile Tyr Ala Leu Tyr Asp Ala Phe Ala Asn Gly Lys Ile Asn Ser Ile

515 520 525

Asp Lys Leu Glu Glu Tyr Cys Lys Gly Lys Asp Ile Glu Ile Gly His

530 535 540

Leu Pro Lys Gln Met Ile Ala Ile Leu Lys Ser Glu His Lys Asp Met

545 550 555 560

Ala Thr Glu Ala Lys Arg Lys Gln Glu Glu Met Leu Ala Asp Val Gln

565 570 575

Lys Ser Leu Glu Ser Leu Asp Asn Gln Ile Asn Glu Glu Ile Glu Asn

580 585 590

Val Glu Arg Lys Asn Ser Ser Leu Lys Ser Gly Glu Ile Ala Ser Trp

595 600 605

Leu Val Asn Asp Met Met Arg Phe Gln Pro Val Gln Lys Asp Asn Glu

610 615 620

Gly Asn Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met

625 630 635 640

Leu Gln Arg Ser Leu Ala Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg

645 650 655

Tyr Phe Arg Gln Val Asn Leu Ile Glu Ser Ser Asn Pro His Pro Phe

660 665 670

Leu Asn Asn Thr Glu Trp Glu Lys Cys Asn Asn Ile Leu Ser Phe Tyr

675 680 685

Arg Ser Tyr Leu Glu Ala Lys Lys Asn Phe Leu Glu Ser Leu Lys Pro

690 695 700

Glu Asp Trp Glu Lys Asn Gln Tyr Phe Leu Met Leu Lys Glu Pro Lys

705 710 715 720

Thr Asn Cys Glu Thr Leu Val Gln Gly Trp Lys Asn Gly Phe Asn Leu

725 730 735

Pro Arg Gly Ile Phe Thr Glu Pro Ile Arg Lys Trp Phe Met Glu His

740 745 750

Arg Lys Asn Ile Thr Val Ala Glu Leu Lys Arg Val Gly Leu Val Ala

755 760 765

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

770 775 780

Pro Phe Tyr Asn Tyr Leu Phe Asn Val Gly Asn Ile Asn Lys Pro Asp

785 790 795 800

Glu Lys Asn Phe Leu Asn Cys Glu Glu Arg Arg Glu Leu Leu Arg Lys

805 810 815

Lys Lys Asp Glu Phe Lys Lys Met Thr Asp Lys Glu Lys Glu Glu Asn

820 825 830

Pro Ser Tyr Leu Glu Phe Gln Ser Trp Asn Lys Phe Glu Arg Glu Leu

835 840 845

Arg Leu Val Arg Asn Gln Asp Ile Val Thr Trp Leu Leu Cys Met Glu

850 855 860

Leu Phe Asn Lys Lys Lys Ile Lys Glu Leu Asn Val Glu Lys Ile Tyr

865 870 875 880

Leu Lys Asn Ile Asn Thr Asn Thr Thr Lys Lys Glu Lys Asn Thr Glu

885 890 895

Glu Lys Asn Gly Glu Glu Lys Ile Ile Lys Glu Lys Asn Asn Ile Leu

900 905 910

Asn Arg Ile Met Pro Met Arg Leu Pro Ile Lys Val Tyr Gly Arg Glu

915 920 925

Asn Phe Ser Lys Asn Lys Lys Lys Lys Ile Arg Arg Asn Thr Phe Phe

930 935 940

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

945 950 955 960

Phe Lys Ala Leu Glu Arg Asp Arg Arg Leu Gly Gly Leu Phe Ser Phe

965 970 975

Val Lys Thr His Ser Lys Ala Glu Ser Lys Ser Asn Thr Ile Ser Lys

980 985 990

Ser Arg Val Glu Tyr Glu Leu Gly Glu Tyr Gln Lys Ala Arg Ile Glu
995 1000 1005
Ile Ile Lys Asp Met Leu Ala Leu Glu Glu Thr Leu Ile Asp Lys

1010 1015 1020
Tyr Asn Ser Leu Asp Thr Asp Asn Phe His Asn Met Leu Thr Gly
1025 1030 1035
Trp Leu Lys Leu Lys Asp Glu Pro Asp Lys Ala Ser Phe Gln Asn
1040 1045 1050
Asp Val Asp Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn
1055 1060 1065
Gln Tyr Pro Met Arg Asn Arg Ile Ala Phe Ala Asn Ile Asn Pro

1070 1075 1080
Phe Ser Leu Ser Ser Ala Asn Thr Ser Glu Glu Lys Gly Leu Gly
1085 1090 1095
Ile Ala Asn Gln Leu Lys Asp Lys Thr His Lys Thr Ile Glu Lys
1100 1105 1110
Ile Ile Glu Ile Glu Lys Pro Ile Glu Thr Lys Glu
1115 1120 1125

<210> 93
<211> 1157
<212> PRT
<213> Myroides odoratimimus
<400> 93

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

Tyr Ile Ile His Val Phe Lys Asp Glu Leu Ser Ile Ser Asp Phe Glu
65 70 75 80
Lys Arg Val Ala Ile Phe Ala Ser Tyr Phe Pro Ile Leu Glu Thr Val
85 90 95
Asp Lys Lys Ser Ile Lys Glu Arg Asn Arg Thr Ile Asp Leu Thr Leu
100 105 110
Ser Gln Arg Ile Arg Gln Phe Arg Glu Met Leu Ile Ser Leu Val Thr
115 120 125

Ala Val Asp Gln Leu Arg Asn Phe Tyr Thr His Tyr His His Ser Glu
130 135 140
Ile Val Ile Glu Asn Lys Val Leu Asp Phe Leu Asn Ser Ser Leu Val
145 150 155 160
Ser Thr Ala Leu His Val Lys Asp Lys Tyr Leu Lys Thr Asp Lys Thr
165 170 175
Lys Glu Phe Leu Lys Glu Thr Ile Ala Ala Glu Leu Asp Ile Leu Ile
180 185 190

Glu Ala Tyr Lys Lys Lys Gln Ile Glu Lys Lys Asn Thr Arg Phe Lys
195 200 205
Ala Asn Lys Arg Glu Asp Ile Leu Asn Ala Ile Tyr Asn Glu Ala Phe
210 215 220
Trp Ser Phe Ile Asn Asp Lys Asp Lys Asp Lys Glu Thr Val Val Ala
225 230 235 240
Lys Gly Ala Asp Ala Tyr Phe Glu Lys Asn His His Lys Ser Asn Asp
245 250 255

Pro Asp Phe Ala Leu Asn Ile Ser Glu Lys Gly Ile Val Tyr Leu Leu
260 265 270
Ser Phe Phe Leu Thr Asn Lys Glu Met Asp Ser Leu Lys Ala Asn Leu
275 280 285
Thr Gly Phe Lys Gly Lys Val Asp Arg Glu Ser Gly Asn Ser Ile Lys
290 295 300
Tyr Met Ala Thr Gln Arg Ile Tyr Ser Phe His Thr Tyr Arg Gly Leu

305 310 315 320

Lys Gln Lys Ile Arg Thr Ser Glu Glu Gly Val Lys Glu Thr Leu Leu
325 330 335

Met Gln Met Ile Asp Glu Leu Ser Lys Val Pro Asn Val Val Tyr Gln
340 345 350

His Leu Ser Thr Thr Gln Gln Asn Ser Phe Ile Glu Asp Trp Asn Glu
355 360 365

Tyr Tyr Lys Asp Tyr Glu Asp Asp Val Glu Thr Asp Asp Leu Ser Arg
370 375 380

Val Ile His Pro Val Ile Arg Lys Arg Tyr Glu Asp Arg Phe Asn Tyr
385 390 395 400

Phe Ala Ile Arg Phe Leu Asp Glu Phe Phe Asp Phe Pro Thr Leu Arg
405 410 415

Phe Gln Val His Leu Gly Asp Tyr Val His Asp Arg Arg Thr Lys Gln
420 425 430

Leu Gly Lys Val Glu Ser Asp Arg Ile Ile Lys Glu Lys Val Thr Val
435 440 445

Phe Ala Arg Leu Lys Asp Ile Asn Ser Ala Lys Ala Asn Tyr Phe His
450 455 460

Ser Leu Glu Glu Gln Asp Lys Glu Glu Leu Asp Asn Lys Trp Thr Leu
465 470 475 480

Phe Pro Asn Pro Ser Tyr Asp Phe Pro Lys Glu His Thr Leu Gln His
485 490 495

Gln Gly Glu Gln Lys Asn Ala Gly Lys Ile Gly Ile Tyr Val Lys Leu
500 505 510

Arg Asp Thr Gln Tyr Lys Glu Lys Ala Ala Leu Glu Glu Ala Arg Lys
515 520 525

Ser Leu Asn Pro Lys Glu Arg Ser Ala Thr Lys Ala Ser Lys Tyr Asp
530 535 540

Ile Ile Thr Gln Ile Ile Glu Ala Asn Asp Asn Val Lys Ser Glu Lys
545 550 555 560

Pro Leu Val Phe Thr Gly Gln Pro Ile Ala Tyr Leu Ser Met Asn Asp
565 570 575

Ile His Ser Met Leu Phe Ser Leu Leu Thr Asp Asn Ala Glu Leu Lys
580 585 590

Lys Thr Pro Glu Glu Val Glu Ala Lys Leu Ile Asp Gln Ile Gly Lys
595 600 605

Gln Ile Asn Glu Ile Leu Ser Lys Asp Thr Asp Thr Lys Ile Leu Lys
610 615 620

Lys Tyr Lys Asp Asn Asp Leu Lys Glu Thr Asp Thr Asp Lys Ile Thr
625 630 635 640

Arg Asp Leu Ala Arg Asp Lys Glu Glu Ile Glu Lys Leu Ile Leu Glu
645 650 655

Gln Lys Gln Arg Ala Asp Asp Tyr Asn Tyr Thr Ser Ser Thr Lys Phe
660 665 670

Asn Ile Asp Lys Ser Arg Lys Arg Lys His Leu Leu Phe Asn Ala Glu
675 680 685

Lys Gly Lys Ile Gly Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met
690 695 700

Thr Glu Glu Phe Lys Ser Lys Trp Lys Gly Tyr Gln His Thr Glu Leu
705 710 715 720

Gln Lys Leu Phe Ala Tyr Tyr Asp Thr Ser Lys Ser Asp Leu Asp Leu
725 730 735

Ile Leu Ser Asp Met Val Met Val Lys Asp Tyr Pro Ile Glu Leu Ile
740 745 750

Ala Leu Val Lys Lys Ser Arg Thr Leu Val Asp Phe Leu Asn Lys Tyr
755 760 765

Leu Glu Ala Arg Leu Gly Tyr Met Glu Asn Val Ile Thr Arg Val Lys
770 775 780

Asn Ser Ile Gly Thr Pro Gln Phe Lys Thr Val Arg Lys Glu Cys Phe
785 790 795 800

Thr Phe Leu Lys Lys Ser Asn Tyr Thr Val Val Ser Leu Asp Lys Gln

805 810 815

Val Glu Arg Ile Leu Ser Met Pro Leu Phe Ile Glu Arg Gly Phe Met

820 825 830

Asp Asp Lys Pro Thr Met Leu Glu Gly Lys Ser Tyr Gln Gln His Lys

835 840 845

Glu Lys Phe Ala Asp Trp Phe Val His Tyr Lys Glu Asn Ser Asn Tyr

850 855 860

Gln Asn Phe Tyr Asp Thr Glu Val Tyr Glu Ile Thr Thr Glu Asp Lys

865 870 875 880

Arg Glu Lys Ala Lys Val Thr Lys Lys Ile Lys Gln Gln Gln Lys Asn

885 890 895

Asp Val Phe Thr Leu Met Met Val Asn Tyr Met Leu Glu Glu Val Leu

900 905 910

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

915 920 925

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

930 935 940

Arg Asn Lys Asn Tyr Ile Trp Asn Lys Val Val Asp Leu Gln Leu Cys

945 950 955 960

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

965 970 975

Phe Arg Lys Tyr Glu Asn Asp Ser Arg Val Lys Glu Phe Leu Thr Tyr

980 985 990

Gln Ser Asp Ile Val Trp Ser Ala Tyr Leu Ser Asn Glu Val Asp Ser

995 1000 1005

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

1010 1015 1020

Ile Arg Ser Lys Glu Leu Leu Lys Glu Val Gln Glu Ile Glu Cys

1025 1030 1035

Ser Val Tyr Asn Gln Val Ala Asn Lys Glu Ser Leu Lys Gln Ser

1040 1045 1050

Gly Asn Glu Asn Phe Lys Gln Tyr Val Leu Gln Gly Leu Val Pro
1055 1060 1065
Ile Gly Met Asp Val Arg Glu Met Leu Ile Leu Ser Thr Asp Val
1070 1075 1080

Lys Phe Ile Lys Glu Glu Ile Ile Gln Leu Gly Gln Ala Gly Glu
1085 1090 1095
Val Glu Gln Asp Leu Tyr Ser Leu Ile Tyr Ile Arg Asn Lys Phe
1100 1105 1110
Ala His Asn Gln Leu Pro Ile Lys Glu Phe Phe Asp Phe Cys Glu
1115 1120 1125
Asn Asn Tyr Arg Ser Ile Ser Asp Asn Glu Tyr Tyr Ala Glu Tyr
1130 1135 1140

Tyr Met Glu Ile Phe Arg Ser Ile Lys Glu Lys Tyr Thr Ser
1145 1150 1155

<210> 94
<211> 1157
<212> PRT
<213> Myroides odoratimimus

<400> 94
Met Lys Asp Ile Leu Thr Thr Asp Thr Thr Glu Lys Gln Asn Arg Phe
1 5 10 15
Tyr Ser His Lys Ile Ala Asp Lys Tyr Phe Phe Gly Gly Tyr Phe Asn
20 25 30
Leu Ala Ser Asn Asn Ile Tyr Glu Val Phe Glu Glu Val Asn Lys Arg

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

Asp Lys Lys Ser Ile Lys Glu Arg Asn Arg Thr Ile Asp Leu Thr Leu

100 105 110

Ser Gln Arg Ile Arg Gln Phe Arg Glu Met Leu Ile Ser Leu Val Thr

115 120 125

Ala Val Asp Gln Leu Arg Asn Phe Tyr Thr His Tyr His His Ser Glu

130 135 140

Ile Val Ile Glu Asn Lys Val Leu Asp Phe Leu Asn Ser Ser Leu Val

145 150 155 160

Ser Thr Ala Leu His Val Lys Asp Lys Tyr Leu Lys Thr Asp Lys Thr

165 170 175

Lys Glu Phe Leu Lys Glu Thr Ile Ala Ala Glu Leu Asp Ile Leu Ile

180 185 190

Glu Ala Tyr Lys Lys Lys Gln Ile Glu Lys Lys Asn Thr Arg Phe Lys

195 200 205

Ala Asn Lys Arg Glu Asp Ile Leu Asn Ala Ile Tyr Asn Glu Ala Phe

210 215 220

Trp Ser Phe Ile Asn Asp Lys Asp Lys Asp Lys Glu Thr Val Val Ala

225 230 235 240

Lys Gly Ala Asp Ala Tyr Phe Glu Lys Asn His His Lys Ser Asn Asp

245 250 255

Pro Asp Phe Ala Leu Asn Ile Ser Glu Lys Gly Ile Val Tyr Leu Leu

260 265 270

Ser Phe Phe Leu Thr Asn Lys Glu Met Asp Ser Leu Lys Ala Asn Leu

275 280 285

Thr Gly Phe Lys Gly Lys Val Asp Arg Glu Ser Gly Asn Ser Ile Lys

290 295 300

Tyr Met Ala Thr Gln Arg Ile Tyr Ser Phe His Thr Tyr Arg Gly Leu

305 310 315 320

Lys Gln Lys Ile Arg Thr Ser Glu Glu Gly Val Lys Glu Thr Leu Leu

325 330 335

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

340 345 350

His Leu Ser Thr Thr Gln Gln Asn Ser Phe Ile Glu Asp Trp Asn Glu

355 360 365

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

370 375 380

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

385 390 395 400

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

405 410 415

Phe Gln Val His Leu Gly Asp Tyr Val His Asp Arg Arg Thr Lys Gln

420 425 430

Leu Gly Lys Val Glu Ser Asp Arg Ile Ile Lys Glu Lys Val Thr Val

435 440 445

Phe Ala Arg Leu Lys Asp Ile Asn Ser Ala Lys Ala Ser Tyr Phe His

450 455 460

Ser Leu Glu Glu Gln Asp Lys Glu Glu Leu Asp Asn Lys Trp Thr Leu

465 470 475 480

Phe Pro Asn Pro Ser Tyr Asp Phe Pro Lys Glu His Thr Leu Gln His

485 490 495

Gln Gly Glu Gln Lys Asn Ala Gly Lys Ile Gly Ile Tyr Val Lys Leu

500 505 510

Arg Asp Thr Gln Tyr Lys Glu Lys Ala Ala Leu Glu Glu Ala Arg Lys

515 520 525

Ser Leu Asn Pro Lys Glu Arg Ser Ala Thr Lys Ala Ser Lys Tyr Asp

530 535 540

Ile Ile Thr Gln Ile Ile Glu Ala Asn Asp Asn Val Lys Ser Glu Lys

545 550 555 560

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

565 570 575

Ile His Ser Met Leu Phe Ser Leu Leu Thr Asp Asn Ala Glu Leu Lys

580 585 590

Lys Thr Pro Glu Glu Val Glu Ala Lys Leu Ile Asp Gln Ile Gly Lys
595 600 605
Gln Ile Asn Glu Ile Leu Ser Lys Asp Thr Asp Thr Lys Ile Leu Lys

610 615 620
Lys Tyr Lys Asp Asn Asp Leu Lys Glu Thr Asp Thr Asp Lys Ile Thr
625 630 635 640
Arg Asp Leu Ala Arg Asp Lys Glu Glu Ile Glu Lys Leu Ile Leu Glu
645 650 655
Gln Lys Gln Arg Ala Asp Asp Tyr Asn Tyr Thr Ser Ser Thr Lys Phe
660 665 670
Asn Ile Asp Lys Ser Arg Lys Arg Lys His Leu Leu Phe Asn Ala Glu

675 680 685
Lys Gly Lys Ile Gly Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met
690 695 700
Phe Lys Glu Ser Lys Ser Lys Trp Lys Gly Tyr Gln His Thr Glu Leu
705 710 715 720
Gln Lys Leu Phe Ala Tyr Phe Asp Thr Ser Lys Ser Asp Leu Glu Leu
725 730 735
Ile Leu Ser Asp Met Val Met Val Lys Asp Tyr Pro Ile Glu Leu Ile

740 745 750
Asp Leu Val Arg Lys Ser Arg Thr Leu Val Asp Phe Leu Asn Lys Tyr
755 760 765
Leu Glu Ala Arg Leu Gly Tyr Ile Glu Asn Val Ile Thr Arg Val Lys
770 775 780
Asn Ser Ile Gly Thr Pro Gln Phe Lys Thr Val Arg Lys Glu Cys Phe
785 790 795 800
Ala Phe Leu Lys Glu Ser Asn Tyr Thr Val Ala Ser Leu Asp Lys Gln

805 810 815
Ile Glu Arg Ile Leu Ser Met Pro Leu Phe Ile Glu Arg Gly Phe Met
820 825 830
Asp Ser Lys Pro Thr Met Leu Glu Gly Lys Ser Tyr Gln Gln His Lys

835 840 845

Glu Asp Phe Ala Asp Trp Phe Val His Tyr Lys Glu Asn Ser Asn Tyr

850 855 860

Gln Asn Phe Tyr Asp Thr Glu Val Tyr Glu Ile Ile Thr Glu Asp Lys

865 870 875 880

Arg Glu Gln Ala Lys Val Thr Lys Lys Ile Lys Gln Gln Gln Lys Asn

885 890 895

Asp Val Phe Thr Leu Met Met Val Asn Tyr Met Leu Glu Glu Val Leu

900 905 910

Lys Leu Pro Ser Asn Asp Arg Leu Ser Leu Asn Glu Leu Tyr Gln Thr

915 920 925

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

930 935 940

Arg Asn Lys Asn Tyr Ile Trp Asn Lys Val Val Asp Leu Gln Leu Cys

945 950 955 960

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

965 970 975

Phe Arg Lys Tyr Glu Asn Asp Ser Arg Val Lys Glu Phe Leu Thr Tyr

980 985 990

Gln Ser Asp Ile Val Trp Ser Gly Tyr Leu Ser Asn Glu Val Asp Ser

995 1000 1005

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

1010 1015 1020

Ile Arg Ser Lys Glu Leu Leu Lys Glu Val Gln Glu Ile Glu Cys

1025 1030 1035

Ile Val Tyr Asn Gln Val Ala Asn Lys Glu Ser Leu Lys Gln Ser

1040 1045 1050

Gly Asn Glu Asn Phe Lys Gln Tyr Val Leu Gln Gly Leu Leu Pro

1055 1060 1065

Arg Gly Thr Asp Val Arg Glu Met Leu Ile Leu Ser Thr Asp Val

1070 1075 1080

Lys Phe Lys Lys Glu Glu Ile Met Gln Leu Gly Gln Val Arg Glu
1085 1090 1095
Val Glu Gln Asp Leu Tyr Ser Leu Ile Tyr Ile Arg Asn Lys Phe
1100 1105 1110
Ala His Asn Gln Leu Pro Ile Lys Glu Phe Phe Asp Phe Cys Glu

1115 1120 1125
Asn Asn Tyr Arg Pro Ile Ser Asp Asn Glu Tyr Tyr Ala Glu Tyr
1130 1135 1140
Tyr Met Glu Ile Phe Arg Ser Ile Lys Glu Lys Tyr Ala Ser
1145 1150 1155

<210> 95
<211> 1224
<212> PRT
<213> Bergeyella zoohelcum
<400> 95
Met Glu Asn Lys Thr Ser Leu Gly Asn Asn Ile Tyr Tyr Asn Pro Phe
1 5 10 15

Lys Pro Gln Asp Lys Ser Tyr Phe Ala Gly Tyr Phe Asn Ala Ala Met
20 25 30
Glu Asn Thr Asp Ser Val Phe Arg Glu Leu Gly Lys Arg Leu Lys Gly
35 40 45
Lys Glu Tyr Thr Ser Glu Asn Phe Phe Asp Ala Ile Phe Lys Glu Asn
50 55 60
Ile Ser Leu Val Glu Tyr Glu Arg Tyr Val Lys Leu Leu Ser Asp Tyr
65 70 75 80

Phe Pro Met Ala Arg Leu Leu Asp Lys Lys Glu Val Pro Ile Lys Glu
85 90 95
Arg Lys Glu Asn Phe Lys Lys Asn Phe Lys Gly Ile Ile Lys Ala Val
100 105 110
Arg Asp Leu Arg Asn Phe Tyr Thr His Lys Glu His Gly Glu Val Glu
115 120 125
Ile Thr Asp Glu Ile Phe Gly Val Leu Asp Glu Met Leu Lys Ser Thr

130 135 140

Val Leu Thr Val Lys Lys Lys Lys Val Lys Thr Asp Lys Thr Lys Glu
145 150 155 160
Ile Leu Lys Lys Ser Ile Glu Lys Gln Leu Asp Ile Leu Cys Gln Lys
165 170 175
Lys Leu Glu Tyr Leu Arg Asp Thr Ala Arg Lys Ile Glu Glu Lys Arg
180 185 190
Arg Asn Gln Arg Glu Arg Gly Glu Lys Glu Leu Val Ala Pro Phe Lys
195 200 205

Tyr Ser Asp Lys Arg Asp Asp Leu Ile Ala Ala Ile Tyr Asn Asp Ala
210 215 220
Phe Asp Val Tyr Ile Asp Lys Lys Lys Asp Ser Leu Lys Glu Ser Ser
225 230 235 240
Lys Ala Lys Tyr Asn Thr Lys Ser Asp Pro Gln Gln Glu Glu Gly Asp
245 250 255
Leu Lys Ile Pro Ile Ser Lys Asn Gly Val Val Phe Leu Leu Ser Leu
260 265 270

Phe Leu Thr Lys Gln Glu Ile His Ala Phe Lys Ser Lys Ile Ala Gly
275 280 285
Phe Lys Ala Thr Val Ile Asp Glu Ala Thr Val Ser Glu Ala Thr Val
290 295 300
Ser His Gly Lys Asn Ser Ile Cys Phe Met Ala Thr His Glu Ile Phe
305 310 315 320
Ser His Leu Ala Tyr Lys Lys Leu Lys Arg Lys Val Arg Thr Ala Glu
325 330 335

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

Glu Asp Trp Asn Glu Tyr Leu Lys Glu Asn Asn Gly Asp Val Gly Thr
385 390 395 400

Met Glu Glu Glu Gln Val Ile His Pro Val Ile Arg Lys Arg Tyr Glu
405 410 415
Asp Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Ala Gln
420 425 430
Phe Pro Thr Leu Arg Phe Gln Val His Leu Gly Asn Tyr Leu His Asp
435 440 445
Ser Arg Pro Lys Glu Asn Leu Ile Ser Asp Arg Arg Ile Lys Glu Lys
450 455 460

Ile Thr Val Phe Gly Arg Leu Ser Glu Leu Glu His Lys Lys Ala Leu
465 470 475 480
Phe Ile Lys Asn Thr Glu Thr Asn Glu Asp Arg Glu His Tyr Trp Glu
485 490 495
Ile Phe Pro Asn Pro Asn Tyr Asp Phe Pro Lys Glu Asn Ile Ser Val
500 505 510
Asn Asp Lys Asp Phe Pro Ile Ala Gly Ser Ile Leu Asp Arg Glu Lys
515 520 525

Gln Pro Val Ala Gly Lys Ile Gly Ile Lys Val Lys Leu Leu Asn Gln
530 535 540
Gln Tyr Val Ser Glu Val Asp Lys Ala Val Lys Ala His Gln Leu Lys
545 550 555 560
Gln Arg Lys Ala Ser Lys Pro Ser Ile Gln Asn Ile Ile Glu Glu Ile
565 570 575
Val Pro Ile Asn Glu Ser Asn Pro Lys Glu Ala Ile Val Phe Gly Gly
580 585 590

Gln Pro Thr Ala Tyr Leu Ser Met Asn Asp Ile His Ser Ile Leu Tyr
595 600 605
Glu Phe Phe Asp Lys Trp Glu Lys Lys Lys Glu Lys Leu Glu Lys Lys
610 615 620
Gly Glu Lys Glu Leu Arg Lys Glu Ile Gly Lys Glu Leu Glu Lys Lys

625 630 635 640

Ile Val Gly Lys Ile Gln Ala Gln Ile Gln Gln Ile Ile Asp Lys Asp

645 650 655

Thr Asn Ala Lys Ile Leu Lys Pro Tyr Gln Asp Gly Asn Ser Thr Ala

660 665 670

Ile Asp Lys Glu Lys Leu Ile Lys Asp Leu Lys Gln Glu Gln Asn Ile

675 680 685

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

690 695 700

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

705 710 715 720

Asp Arg Asn His Lys Gln Tyr Leu Lys Asp Asn Leu Lys Arg Lys Tyr

725 730 735

Pro Glu Ala Pro Ala Arg Lys Glu Val Leu Tyr Tyr Arg Glu Lys Gly

740 745 750

Lys Val Ala Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met Pro Thr

755 760 765

Asp Phe Lys Asn Glu Trp Lys Gly Glu Gln His Ser Leu Leu Gln Lys

770 775 780

Ser Leu Ala Tyr Tyr Glu Gln Cys Lys Glu Glu Leu Lys Asn Leu Leu

785 790 795 800

Pro Glu Lys Val Phe Gln His Leu Pro Phe Lys Leu Gly Gly Tyr Phe

805 810 815

Gln Gln Lys Tyr Leu Tyr Gln Phe Tyr Thr Cys Tyr Leu Asp Lys Arg

820 825 830

Leu Glu Tyr Ile Ser Gly Leu Val Gln Gln Ala Glu Asn Phe Lys Ser

835 840 845

Glu Asn Lys Val Phe Lys Lys Val Glu Asn Glu Cys Phe Lys Phe Leu

850 855 860

Lys Lys Gln Asn Tyr Thr His Lys Glu Leu Asp Ala Arg Val Gln Ser

865 870 875 880

Ile Leu Gly Tyr Pro Ile Phe Leu Glu Arg Gly Phe Met Asp Glu Lys
885 890 895
Pro Thr Ile Ile Lys Gly Lys Thr Phe Lys Gly Asn Glu Ala Leu Phe
900 905 910

Ala Asp Trp Phe Arg Tyr Tyr Lys Glu Tyr Gln Asn Phe Gln Thr Phe
915 920 925
Tyr Asp Thr Glu Asn Tyr Pro Leu Val Glu Leu Glu Lys Lys Gln Ala
930 935 940
Asp Arg Lys Arg Lys Thr Lys Ile Tyr Gln Gln Lys Lys Asn Asp Val
945 950 955 960
Phe Thr Leu Leu Met Ala Lys His Ile Phe Lys Ser Val Phe Lys Gln
965 970 975

Asp Ser Ile Asp Gln Phe Ser Leu Glu Asp Leu Tyr Gln Ser Arg Glu
980 985 990
Glu Arg Leu Gly Asn Gln Glu Arg Ala Arg Gln Thr Gly Glu Arg Asn
995 1000 1005
Thr Asn Tyr Ile Trp Asn Lys Thr Val Asp Leu Lys Leu Cys Asp
1010 1015 1020
Gly Lys Ile Thr Val Glu Asn Val Lys Leu Lys Asn Val Gly Asp
1025 1030 1035

Phe Ile Lys Tyr Glu Tyr Asp Gln Arg Val Gln Ala Phe Leu Lys
1040 1045 1050
Tyr Glu Glu Asn Ile Glu Trp Gln Ala Phe Leu Ile Lys Glu Ser
1055 1060 1065
Lys Glu Glu Glu Asn Tyr Pro Tyr Val Val Glu Arg Glu Ile Glu
1070 1075 1080
Gln Tyr Glu Lys Val Arg Arg Glu Glu Leu Leu Lys Glu Val His
1085 1090 1095

Leu Ile Glu Glu Tyr Ile Leu Glu Lys Val Lys Asp Lys Glu Ile
1100 1105 1110
Leu Lys Lys Gly Asp Asn Gln Asn Phe Lys Tyr Tyr Ile Leu Asn

1115 1120 1125

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

1130 1135 1140

Val Phe Asn Leu Asn Thr Glu Pro Glu Asp Val Asn Ile Asn Gln

1145 1150 1155

Leu Lys Gln Glu Ala Thr Asp Leu Glu Gln Lys Ala Phe Val Leu

1160 1165 1170

Thr Tyr Ile Arg Asn Lys Phe Ala His Asn Gln Leu Pro Lys Lys

1175 1180 1185

Glu Phe Trp Asp Tyr Cys Gln Glu Lys Tyr Gly Lys Ile Glu Lys

1190 1195 1200

Glu Lys Thr Tyr Ala Glu Tyr Phe Ala Glu Val Phe Lys Lys Glu

1205 1210 1215

Lys Glu Ala Leu Ile Lys

1220

<210> 96

<211> 1150

<212> PRT

<213> Prevotella saccharolytica

<400> 96

Met Met Glu Lys Glu Asn Val Gln Gly Ser His Ile Tyr Tyr Glu Pro

1 5 10 15

Thr Asp Lys Cys Phe Trp Ala Ala Phe Tyr Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Leu Thr Ile Ala His Ile Asn Ser Phe Val Asn Ser Lys Lys

35 40 45

Gly Ile Asn Asn Asp Asp Lys Val Leu Asp Ile Ile Asp Asp Trp Ser

50 55 60

Lys Phe Asp Asn Asp Leu Leu Met Gly Ala Arg Leu Asn Lys Leu Ile

65 70 75 80

Leu Lys His Phe Pro Phe Leu Lys Ala Pro Leu Tyr Gln Leu Ala Lys

85 90 95

Arg Lys Thr Arg Lys Gln Gln Gly Lys Glu Gln Gln Asp Tyr Glu Lys
100 105 110

Lys Gly Asp Glu Asp Pro Glu Val Ile Gln Glu Ala Ile Ala Asn Ala
115 120 125
Phe Lys Met Ala Asn Val Arg Lys Thr Leu His Ala Phe Leu Lys Gln
130 135 140
Leu Glu Asp Leu Arg Asn His Phe Ser His Tyr Asn Tyr Asn Ser Pro
145 150 155 160
Ala Lys Lys Met Glu Val Lys Phe Asp Asp Gly Phe Cys Asn Lys Leu
165 170 175

Tyr Tyr Val Phe Asp Ala Ala Leu Gln Met Val Lys Asp Asp Asn Arg
180 185 190
Met Asn Pro Glu Ile Asn Met Gln Thr Asp Phe Glu His Leu Val Arg
195 200 205
Leu Gly Arg Asn Arg Lys Ile Pro Asn Thr Phe Lys Tyr Asn Phe Thr
210 215 220
Asn Ser Asp Gly Thr Ile Asn Asn Asn Gly Leu Leu Phe Phe Val Ser
225 230 235 240

Leu Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Lys
245 250 255
Gly Phe Lys Gly Gly Thr Glu Asn Tyr Met Arg Met Thr Asn Glu Val
260 265 270
Phe Cys Arg Asn Arg Met Val Ile Pro Lys Leu Arg Leu Glu Thr Asp
275 280 285
Tyr Asp Asn His Gln Leu Met Phe Asp Met Leu Asn Glu Leu Val Arg
290 295 300

Cys Pro Leu Ser Leu Tyr Lys Arg Leu Lys Gln Glu Asp Gln Asp Lys
305 310 315 320
Phe Arg Val Pro Ile Glu Phe Leu Asp Glu Asp Asn Glu Ala Asp Asn
325 330 335
Pro Tyr Gln Glu Asn Ala Asn Ser Asp Glu Asn Pro Thr Glu Glu Thr

340 345 350

Asp Pro Leu Lys Asn Thr Leu Val Arg His Gln His Arg Phe Pro Tyr

355 360 365

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

370 375 380

Phe Gln Ile Asn Leu Gly Cys Tyr His Phe Ser Ile Tyr Asp Lys Thr

385 390 395 400

Ile Gly Glu Arg Thr Glu Lys Arg His Leu Thr Arg Thr Leu Phe Gly

405 410 415

Phe Asp Arg Leu Gln Asn Phe Ser Val Lys Leu Gln Pro Glu His Trp

420 425 430

Lys Asn Met Val Lys His Leu Asp Thr Glu Glu Ser Ser Asp Lys Pro

435 440 445

Tyr Leu Ser Asp Ala Met Pro His Tyr Gln Ile Glu Asn Glu Lys Ile

450 455 460

Gly Ile His Phe Leu Lys Thr Asp Thr Glu Lys Lys Glu Thr Val Trp

465 470 475 480

Pro Ser Leu Glu Val Glu Glu Val Ser Ser Asn Arg Asn Lys Tyr Lys

485 490 495

Ser Glu Lys Asn Leu Thr Ala Asp Ala Phe Leu Ser Thr His Glu Leu

500 505 510

Leu Pro Met Met Phe Tyr Tyr Gln Leu Leu Ser Ser Glu Glu Lys Thr

515 520 525

Arg Ala Ala Ala Gly Asp Lys Val Gln Gly Val Leu Gln Ser Tyr Arg

530 535 540

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

545 550 555 560

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

565 570 575

Gly Asn Met Pro Gln Gln Met Leu Ala Ile Leu Glu His Gln Glu Pro

580 585 590

Asp Met Glu Gln Lys Ala Lys Glu Lys Leu Asp Arg Leu Ile Thr Glu
595 600 605
Thr Lys Lys Arg Ile Gly Lys Leu Glu Asp Gln Phe Lys Gln Lys Val
610 615 620

Arg Ile Gly Lys Arg Arg Ala Asp Leu Pro Lys Val Gly Ser Ile Ala
625 630 635 640
Asp Trp Leu Val Asn Asp Met Met Arg Phe Gln Pro Ala Lys Arg Asn
645 650 655
Ala Asp Asn Thr Gly Val Pro Asp Ser Lys Ala Asn Ser Thr Glu Tyr
660 665 670
Arg Leu Leu Gln Glu Ala Leu Ala Phe Tyr Ser Ala Tyr Lys Asp Arg
675 680 685

Leu Glu Pro Tyr Phe Arg Gln Val Asn Leu Ile Gly Gly Thr Asn Pro
690 695 700
His Pro Phe Leu His Arg Val Asp Trp Lys Lys Cys Asn His Leu Leu
705 710 715 720
Ser Phe Tyr His Asp Tyr Leu Glu Ala Lys Glu Gln Tyr Leu Ser His
725 730 735
Leu Ser Pro Ala Asp Trp Gln Lys His Gln His Phe Leu Leu Leu Lys
740 745 750

Val Arg Lys Asp Ile Gln Asn Glu Lys Lys Asp Trp Lys Lys Ser Leu
755 760 765
Val Ala Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg Gly Leu Phe Thr
770 775 780
Glu Ser Ile Lys Thr Trp Phe Ser Thr Asp Ala Asp Lys Val Gln Ile
785 790 795 800
Thr Asp Thr Lys Leu Phe Glu Asn Arg Val Gly Leu Ile Ala Lys Leu
805 810 815

Ile Pro Leu Tyr Tyr Asp Lys Val Tyr Asn Asp Lys Pro Gln Pro Phe
820 825 830
Tyr Gln Tyr Pro Phe Asn Ile Asn Asp Arg Tyr Lys Pro Glu Asp Thr

835 840 845

Arg Lys Arg Phe Thr Ala Ala Ser Ser Lys Leu Trp Asn Glu Lys Lys

850 855 860

Met Leu Tyr Lys Asn Ala Gln Pro Asp Ser Ser Asp Lys Ile Glu Tyr

865 870 875 880

Pro Gln Tyr Leu Asp Phe Leu Ser Trp Lys Lys Leu Glu Arg Glu Leu

885 890 895

Arg Met Leu Arg Asn Gln Asp Met Met Val Trp Leu Met Cys Lys Asp

900 905 910

Leu Phe Ala Gln Cys Thr Val Glu Gly Val Glu Phe Ala Asp Leu Lys

915 920 925

Leu Ser Gln Leu Glu Val Asp Val Asn Val Gln Asp Asn Leu Asn Val

930 935 940

Leu Asn Asn Val Ser Ser Met Ile Leu Pro Leu Ser Val Tyr Pro Ser

945 950 955 960

Asp Ala Gln Gly Asn Val Leu Arg Asn Ser Lys Pro Leu His Thr Val

965 970 975

Tyr Val Gln Glu Asn Asn Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys

980 985 990

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

995 1000 1005

Ala Glu Gly Glu Asp Leu Gln Gln His Pro Leu Thr Lys Asn Arg

1010 1015 1020

Leu Glu Tyr Glu Leu Ser Ile Tyr Gln Thr Met Arg Ile Ser Val

1025 1030 1035

Phe Glu Gln Thr Leu Gln Leu Glu Lys Ala Ile Leu Thr Arg Asn

1040 1045 1050

Lys Thr Leu Cys Gly Asn Asn Phe Asn Asn Leu Leu Asn Ser Trp

1055 1060 1065

Ser Glu His Arg Thr Asp Lys Lys Thr Leu Gln Pro Asp Ile Asp

1070 1075 1080

Phe Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro
1085 1090 1095
Met Ser Thr Asn Thr Val Met Gln Gly Ile Glu Lys Phe Asn Ile
1100 1105 1110
Gln Thr Pro Lys Leu Glu Glu Lys Asp Gly Leu Gly Ile Ala Ser
1115 1120 1125

Gln Leu Ala Lys Lys Thr Lys Asp Ala Ala Ser Arg Leu Gln Asn
1130 1135 1140
Ile Ile Asn Gly Gly Thr Asn
1145 1150

<210> 97
<211> 1175
<212> PRT
<213> Porphyromonas gingivalis
<400> 97
Met Thr Glu Gln Asn Glu Lys Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15
Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn
20 25 30

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

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu
100 105 110
Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125
Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu
145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205

Arg Cys Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp
210 215 220

Arg Glu Glu Lys Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly
245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe
260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe
305 310 315 320

Arg Val Pro Val Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr
325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe
340 345 350

Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser
355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys
370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400
Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp
420 425 430
Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly
435 440 445
Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln Leu Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp
465 470 475 480

Lys Arg Phe Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Ala
500 505 510
Ser Ala Glu Arg Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525
Tyr Ala Val Tyr Asp Ala Phe Ala Arg Gly Glu Ile Asp Thr Leu Asp
530 535 540

Arg Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Arg Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asp Met Glu
565 570 575
Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg

705 710 715 720

Ser Asp Arg Val Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys

725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

Gly Leu Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg

835 840 845

Arg Ile Glu Asp Ala Phe Ala Gly Ile Glu Asn Ala Ser Arg Glu Asn

850 855 860

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

865 870 875 880

Phe Glu Ser Lys Leu Lys Val Lys Ala Asp Lys Ile Asn Ile Ala Lys
885 890 895
Leu Lys Lys Glu Ile Leu Glu Ala Lys Glu His Pro Tyr Leu Asp Phe
900 905 910
Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln
915 920 925

Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys
930 935 940
Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Thr
945 950 955 960
Asp Val His Glu Gln Gly Ser Leu Asn Val Leu Asn Arg Val Lys Pro
965 970 975
Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val
980 985 990

His Lys Glu Gln Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp
995 1000 1005
Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp
1010 1015 1020
Arg Arg Leu Asn Gly Leu Phe Ser Phe Val Asp Thr Gly Ala Leu
1025 1030 1035
Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu
1040 1045 1050

Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala Phe Glu Gln Thr
1055 1060 1065
Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro
1070 1075 1080
Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp Ser Asp Pro Leu
1085 1090 1095
Leu Asp Lys Trp Pro Asp Leu His Gly Asn Val Arg Leu Leu Ile
1100 1105 1110

Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp

1115 1120 1125

Glu Thr Leu Phe Ser Ser Ile Arg Lys Tyr Asp Pro Ser Ser Pro

1130 1135 1140

Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu

1145 1150 1155

Ser Glu Glu Val Lys Gln Ala Lys Glu Met Val Glu Arg Ile Ile

1160 1165 1170

Gln Ala

1175

<210> 98

<211> 1115

<212> PRT

<213> Bacteroides pyogenes

<400> 98

Met Glu Ser Ile Lys Asn Ser Gln Lys Ser Thr Gly Lys Thr Leu Gln

1 5 10 15

Lys Asp Pro Pro Tyr Phe Gly Leu Tyr Leu Asn Met Ala Leu Leu Asn

20 25 30

Val Arg Lys Val Glu Asn His Ile Arg Lys Trp Leu Gly Asp Val Ala

35 40 45

Leu Leu Pro Glu Lys Ser Gly Phe His Ser Leu Leu Thr Thr Asp Asn

50 55 60

Leu Ser Ser Ala Lys Trp Thr Arg Phe Tyr Tyr Lys Ser Arg Lys Phe

65 70 75 80

Leu Pro Phe Leu Glu Met Phe Asp Ser Asp Lys Lys Ser Tyr Glu Asn

85 90 95

Arg Arg Glu Thr Thr Glu Cys Leu Asp Thr Ile Asp Arg Gln Lys Ile

100 105 110

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

115 120 125

Ala Phe Ser His Tyr His Ile Asp Asp Gln Ser Val Lys His Thr Ala

130 135 140

Leu Ile Ile Ser Ser Glu Met His Arg Phe Ile Glu Asn Ala Tyr Ser
145 150 155 160
Phe Ala Leu Gln Lys Thr Arg Ala Arg Phe Thr Gly Val Phe Val Glu
165 170 175
Thr Asp Phe Leu Gln Ala Glu Glu Lys Gly Asp Asn Lys Lys Phe Phe

180 185 190
Ala Ile Gly Gly Asn Glu Gly Ile Lys Leu Lys Asp Asn Ala Leu Ile
195 200 205
Phe Leu Ile Cys Leu Phe Leu Asp Arg Glu Glu Ala Phe Lys Phe Leu
210 215 220
Ser Arg Ala Thr Gly Phe Lys Ser Thr Lys Glu Lys Gly Phe Leu Ala
225 230 235 240
Val Arg Glu Thr Phe Cys Ala Leu Cys Cys Arg Gln Pro His Glu Arg

245 250 255
Leu Leu Ser Val Asn Pro Arg Glu Ala Leu Leu Met Asp Met Leu Asn
260 265 270
Glu Leu Asn Arg Cys Pro Asp Ile Leu Phe Glu Met Leu Asp Glu Lys
275 280 285
Asp Gln Lys Ser Phe Leu Pro Leu Leu Gly Glu Glu Glu Gln Ala His
290 295 300
Ile Leu Glu Asn Ser Leu Asn Asp Glu Leu Cys Glu Ala Ile Asp Asp

305 310 315 320
Pro Phe Glu Met Ile Ala Ser Leu Ser Lys Arg Val Arg Tyr Lys Asn
325 330 335
Arg Phe Pro Tyr Leu Met Leu Arg Tyr Ile Glu Glu Lys Asn Leu Leu
340 345 350
Pro Phe Ile Arg Phe Arg Ile Asp Leu Gly Cys Leu Glu Leu Ala Ser
355 360 365
Tyr Pro Lys Lys Met Gly Glu Glu Asn Asn Tyr Glu Arg Ser Val Thr

370 375 380
Asp His Ala Met Ala Phe Gly Arg Leu Thr Asp Phe His Asn Glu Asp

385 390 395 400

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

405 410 415

Ser Leu Tyr Ala Pro Arg Tyr Ala Ile Tyr Asn Asn Lys Ile Gly Phe

420 425 430

Val Arg Thr Gly Gly Ser Asp Lys Ile Ser Phe Pro Thr Leu Lys Lys

435 440 445

Lys Gly Gly Glu Gly His Cys Val Ala Tyr Thr Leu Gln Asn Thr Lys

450 455 460

Ser Phe Gly Phe Ile Ser Ile Tyr Asp Leu Arg Lys Ile Leu Leu Leu

465 470 475 480

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

485 490 495

Glu Gln Cys Glu Lys His Trp Lys Asp Leu Ser Glu Asn Leu Phe Asp

500 505 510

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

515 520 525

Tyr Thr Leu Pro Arg Ser Lys Gly Gly Lys Leu Val Ser Ser Lys Leu

530 535 540

Ala Asp Lys Gln Glu Lys Tyr Glu Ser Glu Phe Glu Arg Arg Lys Glu

545 550 555 560

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

565 570 575

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

580 585 590

Glu Lys Lys Leu Lys Gly Tyr Val Glu Thr Leu Lys Leu Asp Cys Arg

595 600 605

Glu Arg Leu Arg Val Phe Glu Lys Arg Glu Lys Gly Glu His Pro Val

610 615 620

Pro Pro Arg Ile Gly Glu Met Ala Thr Asp Leu Ala Lys Asp Ile Ile

625 630 635 640

Arg Met Val Ile Asp Gln Gly Val Lys Gln Arg Ile Thr Ser Ala Tyr
645 650 655
Tyr Ser Glu Ile Gln Arg Cys Leu Ala Gln Tyr Ala Gly Asp Asp Asn
660 665 670
Arg Arg His Leu Asp Ser Ile Ile Arg Glu Leu Arg Leu Lys Asp Thr
675 680 685
Lys Asn Gly His Pro Phe Leu Gly Lys Val Leu Arg Pro Gly Leu Gly

690 695 700
His Thr Glu Lys Leu Tyr Gln Arg Tyr Phe Glu Glu Lys Lys Glu Trp
705 710 715 720
Leu Glu Ala Thr Phe Tyr Pro Ala Ala Ser Pro Lys Arg Val Pro Arg
725 730 735
Phe Val Asn Pro Pro Thr Gly Lys Gln Lys Glu Leu Pro Leu Ile Ile
740 745 750
Arg Asn Leu Met Lys Glu Arg Pro Glu Trp Arg Asp Trp Lys Gln Arg

755 760 765
Lys Asn Ser His Pro Ile Asp Leu Pro Ser Gln Leu Phe Glu Asn Glu
770 775 780
Ile Cys Arg Leu Leu Lys Asp Lys Ile Gly Lys Glu Pro Ser Gly Lys
785 790 795 800
Leu Lys Trp Asn Glu Met Phe Lys Leu Tyr Trp Asp Lys Glu Phe Pro
805 810 815
Asn Gly Met Gln Arg Phe Tyr Arg Cys Lys Arg Arg Val Glu Val Phe

820 825 830
Asp Lys Val Val Glu Tyr Glu Tyr Ser Glu Glu Gly Gly Asn Tyr Lys
835 840 845
Lys Tyr Tyr Glu Ala Leu Ile Asp Glu Val Val Arg Gln Lys Ile Ser
850 855 860
Ser Ser Lys Glu Lys Ser Lys Leu Gln Val Glu Asp Leu Thr Leu Ser
865 870 875 880
Val Arg Arg Val Phe Lys Arg Ala Ile Asn Glu Lys Glu Tyr Gln Leu

885 890 895

Arg Leu Leu Cys Glu Asp Asp Arg Leu Leu Phe Met Ala Val Arg Asp

900 905 910

Leu Tyr Asp Trp Lys Glu Ala Gln Leu Asp Leu Asp Lys Ile Asp Asn

915 920 925

Met Leu Gly Glu Pro Val Ser Val Ser Gln Val Ile Gln Leu Glu Gly

930 935 940

Gly Gln Pro Asp Ala Val Ile Lys Ala Glu Cys Lys Leu Lys Asp Val

945 950 955 960

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

965 970 975

Pro Tyr Phe Ala Asn His Glu Ala Thr Gln Glu Gln Val Glu Met Glu

980 985 990

Leu Arg His Tyr Glu Asp His Arg Arg Arg Val Phe Asn Trp Val Phe

995 1000 1005

Ala Leu Glu Lys Ser Val Leu Lys Asn Glu Lys Leu Arg Arg Phe

1010 1015 1020

Tyr Glu Glu Ser Gln Gly Gly Cys Glu His Arg Arg Cys Ile Asp

1025 1030 1035

Ala Leu Arg Lys Ala Ser Leu Val Ser Glu Glu Glu Tyr Glu Phe

1040 1045 1050

Leu Val His Ile Arg Asn Lys Ser Ala His Asn Gln Phe Pro Asp

1055 1060 1065

Leu Glu Ile Gly Lys Leu Pro Pro Asn Val Thr Ser Gly Phe Cys

1070 1075 1080

Glu Cys Ile Trp Ser Lys Tyr Lys Ala Ile Ile Cys Arg Ile Ile

1085 1090 1095

Pro Phe Ile Asp Pro Glu Arg Arg Phe Phe Gly Lys Leu Leu Glu

1100 1105 1110

Gln Lys

1115

<210> 99

<211> 1135

<212> PRT

<213> Porphyromonas gingivalis

<400> 99

Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys
35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu
50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr
65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp
85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile
100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His
115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp
130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala
145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala
165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys
180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg
195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu

275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu

290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro

420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Arg Ser Met

435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu

450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys
465 470 475 480
Ala Asn Arg Ile Leu Asp Glu Thr Ala Glu Gly Lys Leu Gln Phe Ser
485 490 495
Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro
500 505 510
Lys Ser Lys Asp Arg Arg Glu Lys Ala Glu Thr Thr Leu Glu Lys Tyr
515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu
530 535 540
Ser Ala Phe Asp Met Asp Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560
Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr
565 570 575
Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Gln Lys Phe
580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu
595 600 605
Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg
625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg His Gln Phe Arg Ala
645 650 655

Ile Val Ala Glu Leu Arg Leu Leu Asp Pro Ser Ser Gly His Pro Phe
660 665 670
Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr
675 680 685
Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr
690 695 700
Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg
725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met
755 760 765
Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe
770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr
785 790 795 800
Gly Leu Arg Arg Glu Leu Asn Ile His Gly Lys Ser Val Ser Tyr Ile
805 810 815
Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu
820 825 830
Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro
835 840 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala
850 855 860
Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg
865 870 875 880
Leu Met Leu Met Ala Ile Asn Lys Met Met Thr Asp Arg Glu Glu Asp
885 890 895
Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn
900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu
915 920 925
Gly Gly Asp Asn Ser Leu Ser Leu Val Pro Ala Thr Ile Glu Ile Lys
930 935 940
Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg
945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu
965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys
980 985 990
Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg
995 1000 1005
Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys
1010 1015 1020
Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys
1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg
1040 1045 1050
Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro
1055 1060 1065
Leu Ile Tyr Arg Asp Val Ser Ala Lys Val Gly Ser Ile Glu Gly
1070 1075 1080
Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser
1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile
1100 1105 1110
Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met
1115 1120 1125
Ser Gln Pro Ile Asn Asp Leu
1130 1135
<210> 100
<211> 1135
<212> PRT
<213> Porphyromonas gingivalis
<400> 100
Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30
Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys
35 40 45
Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu
50 55 60
Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr
65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp
85 90 95
His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile
100 105 110
Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His
115 120 125
Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp
130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala
145 150 155 160
Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala
165 170 175
Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys
180 185 190
Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg
195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr
210 215 220
Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys
225 230 235 240
Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala
245 250 255
Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu
275 280 285
Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu
290 295 300
Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr
305 310 315 320
Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe
325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu
340 345 350
Gly Glu Ile Glu Leu Asp Ser Tyr Ser Lys Lys Val Gly Arg Asn Gly
355 360 365
Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu
370 375 380
Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu
385 390 395 400

Ala Ser Tyr Pro Val Arg Phe Ser Leu Phe Ala Pro Arg Tyr Ala Ile
405 410 415
Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro
420 425 430
Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met
435 440 445
Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu
450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Gly Phe Leu Arg Lys
465 470 475 480
Ala Asn Arg Ile Leu Asp Glu Thr Ala Glu Gly Lys Leu Gln Phe Ser
485 490 495
Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro
500 505 510

Lys Ser Lys Asp Arg Arg Glu Lys Ala Glu Thr Thr Leu Glu Lys Tyr
515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu
530 535 540
Ser Ala Phe Asp Met Asn Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560
Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr
565 570 575
Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe
580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu
595 600 605
Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg
625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala
645 650 655

Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe
660 665 670
Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr
675 680 685
Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr
690 695 700
Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe
705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg
725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

```
        755                 760                 765
Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe
    770                 775                 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr
785                 790                 795                 800
Gly Leu Arg Arg Glu Leu Asn Ile His Gly Lys Ser Val Ser Tyr Ile
                805                 810                 815
Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu
            820                 825                 830
Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro
        835                 840                 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala
    850                 855                 860
Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg
865                 870                 875                 880
Leu Met Leu Met Ala Ile Asn Lys Met Met Thr Asp Arg Glu Glu Asp
                885                 890                 895
Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn
            900                 905                 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu
        915                 920                 925
Gly Gly Asp Asn Ser Leu Ser Leu Val Pro Ala Thr Ile Glu Ile Lys
    930                 935                 940
Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg
945                 950                 955                 960
Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu
                965                 970                 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys
            980                 985                 990
Ile Phe Asp Trp Ala Phe Ala Leu  Glu Gly Ala Ile Met  Ser Asp Arg
        995                 1000                1005
```

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys
1010 1015 1020
Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys
1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg
1040 1045 1050
Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro
1055 1060 1065
Leu Ile Tyr Arg Asp Val Ser Ala Lys Val Gly Ser Ile Glu Gly
1070 1075 1080
Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser
1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile
1100 1105 1110
Leu Asp His Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met
1115 1120 1125
Ser Gln Pro Ile Asn Asp Leu
1130 1135

<210> 101
<211> 1135
<212> PRT
<213> Porphyromonas gingivalis
<400> 101
Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30
Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys
35 40 45
Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu
50 55 60
Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr

65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp
85 90 95
His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile
100 105 110
Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His
115 120 125
Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp
130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala
145 150 155 160
Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala
165 170 175
Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys
180 185 190
Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg
195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr
210 215 220
Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys
225 230 235 240
Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala
245 250 255
Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu
260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu
275 280 285
Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu
290 295 300
Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr
305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe
325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu
340 345 350

Gly Glu Ile Glu Leu Asp Ser Tyr Ser Lys Lys Val Gly Arg Asn Gly
355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu
370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu
385 390 395 400

Ala Ser Tyr Pro Val Arg Phe Ser Leu Phe Ala Pro Arg Tyr Ala Ile
405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro
420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Arg Ser Met
435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu
450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys
465 470 475 480

Ala Asn Arg Ile Leu Asp Glu Thr Ala Glu Gly Lys Leu Gln Phe Ser
485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro
500 505 510

Lys Ser Lys Asp Arg Arg Glu Lys Ala Glu Thr Thr Leu Glu Lys Tyr
515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu
530 535 540

Ser Ala Phe Asp Met Asp Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560

Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575

Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Gln Lys Phe

580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu

595 600 605

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

610 615 620

Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640

Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg His Gln Phe Arg Ala

645 650 655

Ile Val Ala Glu Leu Arg Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670

Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr

675 680 685

Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700

Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg

725 730 735

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

740 745 750

Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Val Met

755 760 765

Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe

770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu
820 825 830
Lys Thr Val Arg Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro
835 840 845

Val Pro Pro Asp Leu Ala Ala Tyr Ile Lys Arg Ser Phe His Arg Ala
850 855 860
Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg
865 870 875 880
Leu Met Leu Met Ala Ile Asn Lys Ile Met Thr Asp Arg Glu Glu Asp
885 890 895
Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Lys Glu Asn
900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu
915 920 925
Gly Gly Asp Asn Ser Leu Ser Leu Val Pro Ala Thr Ile Glu Ile Lys
930 935 940
Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg
945 950 955 960
Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu
965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys
980 985 990
Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg
995 1000 1005
Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys
1010 1015 1020
Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys
1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg
1040 1045 1050
Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Ile Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 102

<211> 1125

<212> PRT

<213> Prevotella intermedia

<400> 102

Met Lys Met Glu Asp Asp Lys Lys Thr Thr Glu Ser Thr Asn Met Leu

1 5 10 15

Asp Asn Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Val Tyr Ile Thr Val Asn His Ile Asn Lys Val Leu Glu Leu Lys Asn

35 40 45

Lys Lys Asp Gln Asp Ile Ile Ile Asp Asn Asp Gln Asp Ile Leu Ala

50 55 60

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

65 70 75 80

Arg Leu Arg Glu Leu Met Thr Lys His Phe Pro Phe Leu Glu Thr Ala

85 90 95

Ile Tyr Thr Lys Asn Lys Glu Asp Lys Glu Glu Val Lys Gln Glu Lys

100 105 110

Gln Ala Glu Ala Gln Ser Leu Glu Ser Leu Lys Asp Cys Leu Phe Leu

115 120 125

Phe Leu Glu Lys Leu Gln Glu Ala Arg Asn Tyr Tyr Ser His Tyr Lys
130 135 140

Tyr Ser Glu Ser Thr Lys Glu Pro Met Leu Glu Glu Gly Leu Leu Glu
145 150 155 160
Lys Met Tyr Asn Ile Phe Asp Asp Asn Ile Gln Leu Val Ile Lys Asp
165 170 175
Tyr Gln His Asn Lys Asp Ile Asn Pro Asp Glu Asp Phe Lys His Leu
180 185 190
Asp Arg Lys Gly Gln Phe Lys Tyr Ser Phe Ala Asp Asn Glu Gly Asn
195 200 205

Ile Thr Glu Ser Gly Leu Leu Phe Phe Val Ser Leu Phe Leu Glu Lys
210 215 220
Lys Asp Ala Ile Trp Met Gln Gln Lys Leu Thr Gly Phe Lys Asp Asn
225 230 235 240
Arg Glu Ser Lys Lys Lys Met Thr His Glu Val Phe Cys Arg Arg Arg
245 250 255
Met Leu Leu Pro Lys Leu Arg Leu Glu Ser Thr Gln Thr Gln Asp Trp
260 265 270

Ile Leu Leu Asp Met Leu Asn Glu Leu Ile Arg Cys Pro Lys Ser Leu
275 280 285
Tyr Glu Arg Leu Gln Gly Glu Tyr Arg Lys Lys Phe Asn Val Pro Phe
290 295 300
Asp Ser Ala Asp Glu Asp Tyr Asp Ala Glu Gln Glu Pro Phe Lys Asn
305 310 315 320
Thr Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Ala Leu Arg Tyr
325 330 335

Phe Asp Tyr Asn Glu Ile Phe Thr Asn Leu Arg Phe Gln Ile Asp Leu
340 345 350
Gly Thr Tyr His Phe Ser Ile Tyr Lys Lys Leu Ile Gly Gly Gln Lys
355 360 365
Glu Asp Arg His Leu Thr His Lys Leu Tyr Gly Phe Glu Arg Ile Gln

370 375 380

Glu Phe Ala Lys Gln Asn Arg Pro Asp Glu Trp Lys Ala Leu Val Lys

385 390 395 400

Asp Leu Asp Thr Tyr Glu Thr Ser Asn Glu Arg Tyr Ile Ser Glu Thr

405 410 415

Thr Pro His Tyr His Leu Glu Asn Gln Lys Ile Gly Ile Arg Phe Arg

420 425 430

Asn Gly Asn Lys Glu Ile Trp Pro Ser Leu Lys Thr Asn Gly Glu Asn

435 440 445

Asn Glu Lys Ser Lys Tyr Lys Leu Asp Lys Pro Tyr Gln Ala Glu Ala

450 455 460

Phe Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu

465 470 475 480

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

485 490 495

Glu Gly Phe Ile Lys Arg Glu Ile Arg Asp Met Tyr Lys Leu Tyr Asp

500 505 510

Ala Phe Ala Asn Gly Glu Ile Asn Asn Ile Gly Asp Leu Glu Lys Tyr

515 520 525

Cys Glu Asp Lys Gly Ile Pro Lys Arg His Leu Pro Lys Gln Met Val

530 535 540

Ala Ile Leu Tyr Asp Glu Pro Lys Asp Met Val Lys Glu Ala Lys Arg

545 550 555 560

Lys Gln Lys Glu Met Val Lys Asp Thr Lys Lys Leu Leu Ala Thr Leu

565 570 575

Glu Lys Gln Thr Gln Glu Glu Ile Glu Asp Gly Gly Arg Asn Ile Arg

580 585 590

Leu Leu Lys Ser Gly Glu Ile Ala Arg Trp Leu Val Asn Asp Met Met

595 600 605

Arg Phe Gln Pro Val Gln Lys Asp Asn Glu Gly Asn Pro Leu Asn Asn

610 615 620

Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met Leu Gln Arg Ser Leu Ala
625 630 635 640
Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg Tyr Phe Arg Gln Val Asn
645 650 655

Leu Ile Asn Ser Ser Asn Pro His Pro Phe Leu Lys Trp Thr Lys Trp
660 665 670
Glu Glu Cys Asn Asn Ile Leu Ser Phe Tyr Arg Asn Tyr Leu Thr Lys
675 680 685
Lys Ile Glu Phe Leu Asn Lys Leu Lys Pro Glu Asp Trp Glu Lys Asn
690 695 700
Gln Tyr Phe Leu Lys Leu Lys Glu Pro Lys Thr Asn Arg Glu Thr Leu
705 710 715 720

Val Gln Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg Gly Ile Phe Thr
725 730 735
Glu Pro Ile Arg Glu Trp Phe Lys Arg His Gln Asn Asp Ser Lys Glu
740 745 750
Tyr Glu Lys Val Glu Ala Leu Lys Arg Val Gly Leu Val Thr Lys Val
755 760 765
Ile Pro Leu Phe Phe Lys Glu Glu Tyr Phe Lys Glu Asp Ala Gln Lys
770 775 780

Glu Ile Asn Asn Cys Val Gln Pro Phe Tyr Ser Phe Pro Tyr Asn Val
785 790 795 800
Gly Asn Ile His Lys Pro Asp Glu Lys Asp Phe Leu Pro Ser Glu Glu
805 810 815
Arg Lys Lys Leu Trp Gly Asp Lys Lys Asp Lys Phe Lys Gly Tyr Lys
820 825 830
Ala Lys Val Lys Ser Lys Lys Leu Thr Asp Lys Glu Lys Glu Glu Tyr
835 840 845

Arg Ser Tyr Leu Glu Phe Gln Ser Trp Asn Lys Phe Glu Arg Glu Leu
850 855 860
Arg Leu Val Arg Asn Gln Asp Ile Val Thr Trp Leu Leu Cys Thr Glu

865 870 875 880

Leu Ile Asp Lys Met Lys Val Glu Gly Leu Asn Val Glu Glu Leu Gln

885 890 895

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

900 905 910

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

915 920 925

Tyr Glu Ile Asp Asp Ser His Asn Ile Val Lys Asp Arg Pro Leu His

930 935 940

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

945 950 955 960

Phe Lys Ala Leu Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe

965 970 975

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

980 985 990

Ser Val Val Glu Tyr Glu Leu Gly Glu Tyr Gln Asn Ala Arg Ile Glu

995 1000 1005

Thr Ile Lys Asp Met Leu Leu Leu Glu Lys Thr Leu Ile Lys Lys

1010 1015 1020

Tyr Glu Lys Leu Pro Thr Asp Asn Phe Ser Asp Met Leu Asn Gly

1025 1030 1035

Trp Leu Glu Gly Lys Asp Glu Ser Asp Lys Ala Arg Phe Gln Asn

1040 1045 1050

Asp Val Lys Leu Leu Val Ala Val Arg Asn Ala Phe Ser His Asn

1055 1060 1065

Gln Tyr Pro Met Arg Asn Arg Ile Ala Phe Ala Asn Ile Asn Pro

1070 1075 1080

Phe Ser Leu Ser Ser Ala Asp Ile Ser Glu Glu Lys Lys Leu Asp

1085 1090 1095

Ile Ala Asn Gln Leu Lys Asp Lys Thr His Lys Ile Ile Lys Lys

1100 1105 1110

Ile Ile Glu Ile Glu Lys Pro Ile Glu Thr Lys Glu

1115 1120 1125

<210> 103

<211> 1161

<212> PRT

<213> Bacteroidetes bacterium

<400> 103

Met Glu Asn Gln Thr Gln Lys Gly Lys Gly Ile Tyr Tyr Tyr Tyr Thr

1 5 10 15

Lys Asn Glu Asp Lys His Tyr Phe Gly Ser Phe Leu Asn Leu Ala Asn

20 25 30

Asn Asn Ile Glu Gln Ile Ile Glu Glu Phe Arg Ile Arg Leu Ser Leu

35 40 45

Lys Asp Glu Lys Asn Ile Lys Glu Ile Ile Asn Asn Tyr Phe Thr Asp

50 55 60

Lys Lys Ser Tyr Thr Asp Trp Glu Arg Gly Ile Asn Ile Leu Lys Glu

65 70 75 80

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

85 90 95

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

100 105 110

Phe Arg Thr Asn Phe Ser Leu Leu Ile Asp Thr Ile Ile Asp Leu Arg

115 120 125

Asn Phe Tyr Thr His Tyr Phe His Lys Pro Ile Ser Ile Asn Pro Asp

130 135 140

Val Ala Lys Phe Leu Asp Lys Asn Leu Leu Asn Val Cys Leu Asp Ile

145 150 155 160

Lys Lys Gln Lys Met Lys Thr Asp Lys Thr Lys Gln Ala Leu Lys Asp

165 170 175

Gly Leu Asp Lys Glu Leu Lys Lys Leu Ile Glu Leu Lys Lys Ala Glu

180 185 190

Leu Lys Glu Lys Lys Ile Lys Thr Trp Asn Ile Thr Glu Asn Val Glu

195 200 205

Gly Ala Val Tyr Asn Asp Ala Phe Asn His Met Val Tyr Lys Asn Asn

210 215 220

Ala Gly Val Thr Ile Leu Lys Asp Tyr His Lys Ser Ile Leu Pro Asp

225 230 235 240

Asp Lys Ile Asp Ser Glu Leu Lys Leu Asn Phe Ser Ile Ser Gly Leu

245 250 255

Val Phe Leu Leu Ser Met Phe Leu Ser Lys Lys Glu Ile Glu Gln Phe

260 265 270

Lys Ser Asn Leu Glu Gly Phe Lys Gly Lys Val Ile Gly Glu Asn Gly

275 280 285

Glu Tyr Glu Ile Ser Lys Phe Asn Asn Ser Leu Lys Tyr Met Ala Thr

290 295 300

His Trp Ile Phe Ser Tyr Leu Thr Phe Lys Gly Leu Lys Gln Arg Val

305 310 315 320

Lys Asn Thr Phe Asp Lys Glu Thr Leu Leu Met Gln Met Ile Asp Glu

325 330 335

Leu Asn Lys Val Pro His Glu Val Tyr Gln Thr Leu Ser Lys Glu Gln

340 345 350

Gln Asn Glu Phe Leu Glu Asp Ile Asn Glu Tyr Val Gln Asp Asn Glu

355 360 365

Glu Asn Lys Lys Ser Met Glu Asn Ser Ile Val Val His Pro Val Ile

370 375 380

Arg Lys Arg Tyr Asp Asp Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu

385 390 395 400

Asp Glu Phe Ala Asn Phe Pro Thr Leu Lys Phe Phe Val Thr Ala Gly

405 410 415

Asn Phe Val His Asp Lys Arg Glu Lys Gln Ile Gln Gly Ser Met Leu

420 425 430

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

435 440 445

Thr Glu Ile Ala Lys Tyr Lys Ser Asp Tyr Phe Ser Asn Glu Asn Thr
450 455 460
Leu Glu Thr Ser Glu Trp Glu Leu Phe Pro Asn Pro Ser Tyr Leu Leu

465 470 475 480
Ile Gln Asn Asn Ile Pro Val His Ile Asp Leu Ile His Asn Thr Glu
485 490 495
Glu Ala Lys Gln Cys Gln Ile Ala Ile Asp Arg Ile Lys Cys Thr Thr
500 505 510
Asn Pro Ala Lys Lys Arg Asn Thr Arg Lys Ser Lys Glu Glu Ile Ile
515 520 525
Lys Ile Ile Tyr Gln Lys Asn Lys Asn Ile Lys Tyr Gly Asp Pro Thr

530 535 540
Ala Leu Leu Ser Ser Asn Glu Leu Pro Ala Leu Ile Tyr Glu Leu Leu
545 550 555 560
Val Asn Lys Lys Ser Gly Lys Glu Leu Glu Asn Ile Ile Val Glu Lys
565 570 575
Ile Val Asn Gln Tyr Lys Thr Ile Ala Gly Phe Glu Lys Gly Gln Asn
580 585 590
Leu Ser Asn Ser Leu Ile Thr Lys Lys Leu Lys Lys Ser Glu Pro Asn

595 600 605
Glu Asp Lys Ile Asn Ala Glu Lys Ile Ile Leu Ala Ile Asn Arg Glu
610 615 620
Leu Glu Ile Thr Glu Asn Lys Leu Asn Ile Ile Lys Asn Asn Arg Ala
625 630 635 640
Glu Phe Arg Thr Gly Ala Lys Arg Lys His Ile Phe Tyr Ser Lys Glu
645 650 655
Leu Gly Gln Glu Ala Thr Trp Ile Ala Tyr Asp Leu Lys Arg Phe Met

660 665 670
Pro Glu Ala Ser Arg Lys Glu Trp Lys Gly Phe His His Ser Glu Leu
675 680 685
Gln Lys Phe Leu Ala Phe Tyr Asp Arg Asn Lys Asn Asp Ala Lys Ala

690 695 700

Leu Leu Asn Met Phe Trp Asn Phe Asp Asn Asp Gln Leu Ile Gly Asn

705 710 715 720

Asp Leu Asn Ser Ala Phe Arg Glu Phe His Phe Asp Lys Phe Tyr Glu

725 730 735

Lys Tyr Leu Ile Lys Arg Asp Glu Ile Leu Glu Gly Phe Lys Ser Phe

740 745 750

Ile Ser Asn Phe Lys Asp Glu Pro Lys Leu Leu Lys Lys Gly Ile Lys

755 760 765

Asp Ile Tyr Arg Val Phe Asp Lys Arg Tyr Tyr Ile Ile Lys Ser Thr

770 775 780

Asn Ala Gln Lys Glu Gln Leu Leu Ser Lys Pro Ile Cys Leu Pro Arg

785 790 795 800

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

805 810 815

Ser Asn Ser Ala Leu Phe Ala Asp Trp Tyr Gln Tyr Thr Tyr Ser Asp

820 825 830

Lys His Glu Phe Gln Ser Phe Tyr Asp Met Pro Arg Asp Tyr Lys Glu

835 840 845

Gln Phe Glu Lys Phe Glu Leu Asn Asn Ile Lys Ser Ile Gln Asn Lys

850 855 860

Lys Asn Leu Asn Lys Ser Asp Lys Phe Ile Tyr Phe Arg Tyr Lys Gln

865 870 875 880

Asp Leu Lys Ile Lys Gln Ile Lys Ser Gln Asp Leu Phe Ile Lys Leu

885 890 895

Met Val Asp Glu Leu Phe Asn Val Val Phe Lys Asn Asn Ile Glu Leu

900 905 910

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

915 920 925

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

930 935 940

Asn Lys Met Asn Glu Asn Phe Ile Trp Asn Met Thr Ile Pro Leu Ser
945 950 955 960
Leu Cys Asn Gly Gln Ile Glu Glu Pro Lys Val Lys Leu Lys Asp Ile
965 970 975
Gly Lys Phe Arg Lys Leu Glu Thr Asp Asp Lys Val Ile Gln Leu Leu

980 985 990
Glu Tyr Asp Lys Ser Lys Val Trp Lys Lys Leu Glu Ile Glu Asp Glu
995 1000 1005
Leu Glu Asn Met Pro Asn Ser Tyr Glu Arg Ile Arg Arg Glu Lys
1010 1015 1020
Leu Leu Lys Gly Ile Gln Glu Phe Glu His Phe Leu Leu Glu Lys
1025 1030 1035
Glu Lys Phe Asp Gly Ile Asn His Pro Lys His Phe Glu Gln Asp

1040 1045 1050
Leu Asn Pro Asn Phe Lys Thr Tyr Val Ile Asn Gly Val Leu Arg
1055 1060 1065
Lys Asn Ser Lys Leu Asn Tyr Thr Glu Ile Asp Lys Leu Leu Asp
1070 1075 1080
Leu Glu His Ile Ser Ile Lys Asp Ile Glu Thr Ser Ala Lys Glu
1085 1090 1095
Ile His Leu Ala Tyr Phe Leu Ile His Val Arg Asn Lys Phe Gly

1100 1105 1110
His Asn Gln Leu Pro Lys Leu Glu Ala Phe Glu Leu Met Lys Lys
1115 1120 1125
Tyr Tyr Lys Lys Asn Asn Glu Glu Thr Tyr Ala Glu Tyr Phe His
1130 1135 1140
Lys Val Ser Ser Gln Ile Val Asn Glu Phe Lys Asn Ser Leu Glu
1145 1150 1155
Lys His Ser
1160
<210> 104

<211> 1165

<212> PRT

<213> Chryseobacterium jejuense

<400> 104

Met Glu Lys Thr Gln Thr Gly Leu Gly Ile Tyr Tyr Asp His Thr Lys
1 5 10 15

Leu Gln Asp Lys Tyr Phe Phe Gly Gly Phe Phe Asn Leu Ala Gln Asn
20 25 30

Asn Ile Asp Asn Val Ile Lys Ala Phe Ile Ile Lys Phe Phe Pro Glu
35 40 45

Arg Lys Asp Lys Asp Ile Asn Ile Ala Gln Phe Leu Asp Ile Cys Phe
50 55 60

Lys Asp Asn Asp Ala Asp Ser Asp Phe Gln Lys Lys Asn Lys Phe Leu
65 70 75 80

Arg Ile His Phe Pro Val Ile Gly Phe Leu Thr Ser Asp Asn Asp Lys
85 90 95

Ala Gly Phe Lys Lys Lys Phe Ala Leu Leu Leu Lys Thr Ile Ser Glu
100 105 110

Leu Arg Asn Phe Tyr Thr His Tyr Tyr His Lys Ser Ile Glu Phe Pro
115 120 125

Ser Glu Leu Phe Glu Leu Leu Asp Asp Ile Phe Val Lys Thr Thr Ser
130 135 140

Glu Ile Lys Lys Leu Lys Lys Lys Asp Asp Lys Thr Gln Gln Leu Leu
145 150 155 160

Asn Lys Asn Leu Ser Glu Glu Tyr Asp Ile Arg Tyr Gln Gln Gln Ile
165 170 175

Glu Arg Leu Lys Glu Leu Lys Ala Gln Gly Lys Arg Val Ser Leu Thr
180 185 190

Asp Glu Thr Ala Ile Arg Asn Gly Val Phe Asn Ala Ala Phe Asn His
195 200 205

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

210 215 220

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

225 230 235 240

Gln Asn Ser Ile Leu Phe Leu Leu Ser Met Phe Leu Glu Arg Lys Glu

245 250 255

Thr Glu Asp Leu Lys Ser Arg Val Lys Gly Phe Lys Ala Lys Ile Ile

260 265 270

Lys Gln Gly Glu Glu Gln Ile Ser Gly Leu Lys Phe Met Ala Thr His

275 280 285

Trp Val Phe Ser Tyr Leu Cys Phe Lys Gly Ile Lys Gln Lys Leu Ser

290 295 300

Thr Glu Phe His Glu Glu Thr Leu Leu Ile Gln Ile Ile Asp Glu Leu

305 310 315 320

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

325 330 335

Glu Lys Phe Leu Glu Asp Ile Asn Glu Tyr Met Lys Glu Gly Asn Ala

340 345 350

Asp Leu Ser Leu Glu Asp Ser Lys Val Ile His Pro Val Ile Arg Lys

355 360 365

Arg Tyr Glu Asn Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu

370 375 380

Tyr Leu Ser Ser Thr Ser Leu Lys Phe Gln Val His Val Gly Asn Tyr

385 390 395 400

Val His Asp Arg Arg Val Lys His Ile Asn Gly Thr Gly Phe Gln Thr

405 410 415

Glu Arg Ile Val Lys Asp Arg Ile Lys Val Phe Gly Arg Leu Ser Asn

420 425 430

Ile Ser Asn Leu Lys Ala Asp Tyr Ile Lys Glu Gln Leu Glu Leu Pro

435 440 445

Asn Asp Ser Asn Gly Trp Glu Ile Phe Pro Asn Pro Ser Tyr Ile Phe

450 455 460

Ile Asp Asn Asn Val Pro Ile His Val Leu Ala Asp Glu Ala Thr Lys
465 470 475 480
Lys Gly Ile Glu Leu Phe Lys Asp Lys Arg Arg Lys Glu Gln Pro Glu
485 490 495
Glu Leu Gln Lys Arg Lys Gly Lys Ile Ser Lys Tyr Asn Ile Val Ser

500 505 510
Met Ile Tyr Lys Glu Ala Lys Gly Lys Asp Lys Leu Arg Ile Asp Glu
515 520 525
Pro Leu Ala Leu Leu Ser Leu Asn Glu Ile Pro Ala Leu Leu Tyr Gln
530 535 540
Ile Leu Glu Lys Gly Ala Thr Pro Lys Asp Ile Glu Leu Ile Ile Lys
545 550 555 560
Asn Lys Leu Thr Glu Arg Phe Glu Lys Ile Lys Asn Tyr Asp Pro Glu

565 570 575
Thr Pro Ala Pro Ala Ser Gln Ile Ser Lys Arg Leu Arg Asn Asn Thr
580 585 590
Thr Ala Lys Gly Gln Glu Ala Leu Asn Ala Glu Lys Leu Ser Leu Leu
595 600 605
Ile Glu Arg Glu Ile Glu Asn Thr Glu Thr Lys Leu Ser Ser Ile Glu
610 615 620
Glu Lys Arg Leu Lys Ala Lys Lys Glu Gln Arg Arg Asn Thr Pro Gln

625 630 635 640
Arg Ser Ile Phe Ser Asn Ser Asp Leu Gly Arg Ile Ala Ala Trp Leu
645 650 655
Ala Asp Asp Ile Lys Arg Phe Met Pro Ala Glu Gln Arg Lys Asn Trp
660 665 670
Lys Gly Tyr Gln His Ser Gln Leu Gln Gln Ser Leu Ala Tyr Phe Glu
675 680 685
Lys Arg Pro Gln Glu Ala Phe Leu Leu Leu Lys Glu Gly Trp Asp Thr

690 695 700
Ser Asp Gly Ser Ser Tyr Trp Asn Asn Trp Val Met Asn Ser Phe Leu

705 710 715 720

Glu Asn Asn His Phe Glu Lys Phe Tyr Lys Asn Tyr Leu Met Lys Arg

725 730 735

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

740 745 750

Asn Thr Lys Phe Leu Arg Lys Phe Ile Lys Gln Gln Met Pro Ala Asp

755 760 765

Leu Phe Pro Lys Arg His Tyr Ile Leu Lys Asp Leu Glu Thr Glu Lys

770 775 780

Asn Lys Val Leu Ser Lys Pro Leu Val Phe Ser Arg Gly Leu Phe Asp

785 790 795 800

Asn Asn Pro Thr Phe Ile Lys Gly Val Lys Val Thr Glu Asn Pro Glu

805 810 815

Leu Phe Ala Glu Trp Tyr Ser Tyr Gly Tyr Lys Thr Glu His Val Phe

820 825 830

Gln His Phe Tyr Gly Trp Glu Arg Asp Tyr Asn Glu Leu Leu Asp Ser

835 840 845

Glu Leu Gln Lys Gly Asn Ser Phe Ala Lys Asn Ser Ile Tyr Tyr Asn

850 855 860

Arg Glu Ser Gln Leu Asp Leu Ile Lys Leu Lys Gln Asp Leu Lys Ile

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

Ala Gln Ser Leu Arg Glu Glu Gly Asp Asn Ser Pro Asn Ile Ile Lys

930 935 940

Asp Asp Phe Ile Trp Ser Lys Thr Ile Ala Phe Arg Ser Lys Gln Ile

945 950 955 960

Tyr Glu Pro Ala Ile Lys Leu Lys Asp Ile Gly Lys Phe Asn Arg Phe
965 970 975
Val Leu Asp Asp Glu Glu Ser Lys Ala Ser Lys Leu Leu Ser Tyr Asp
980 985 990
Lys Asn Lys Ile Trp Asn Lys Glu Gln Leu Glu Arg Glu Leu Ser Ile
995 1000 1005
Gly Glu Asn Ser Tyr Glu Val Ile Arg Arg Glu Lys Leu Phe Lys

1010 1015 1020
Glu Ile Gln Asn Leu Glu Leu Gln Ile Leu Ser Asn Trp Ser Trp
1025 1030 1035
Asp Gly Ile Asn His Pro Arg Glu Phe Glu Met Glu Asp Gln Lys
1040 1045 1050
Asn Thr Arg His Pro Asn Phe Lys Met Tyr Leu Val Asn Gly Ile
1055 1060 1065
Leu Arg Lys Asn Ile Asn Leu Tyr Lys Glu Asp Glu Asp Phe Trp

1070 1075 1080
Leu Glu Ser Leu Lys Glu Asn Asp Phe Lys Thr Leu Pro Ser Glu
1085 1090 1095
Val Leu Glu Thr Lys Ser Glu Met Val Gln Leu Leu Phe Leu Val
1100 1105 1110
Ile Leu Ile Arg Asn Gln Phe Ala His Asn Gln Leu Pro Glu Ile
1115 1120 1125
Gln Phe Tyr Asn Phe Ile Arg Lys Asn Tyr Pro Glu Ile Gln Asn

1130 1135 1140
Asn Thr Val Ala Glu Leu Tyr Leu Asn Leu Ile Lys Leu Ala Val
1145 1150 1155
Gln Lys Leu Lys Asp Asn Ser
1160 1165

<210> 105
<211> 1156
<212> PRT
<213> Chryseobacterium carnipullorum

<400> 105

Met Asn Thr Arg Val Thr Gly Met Gly Val Ser Tyr Asp His Thr Lys
1 5 10 15
Lys Glu Asp Lys His Phe Phe Gly Gly Phe Leu Asn Leu Ala Gln Asp

20 25 30
Asn Ile Thr Ala Val Ile Lys Ala Phe Cys Ile Lys Phe Asp Lys Asn
35 40 45
Pro Met Ser Ser Val Gln Phe Ala Glu Ser Cys Phe Thr Asp Lys Asp
50 55 60
Ser Asp Thr Asp Phe Gln Asn Lys Val Arg Tyr Val Arg Thr His Leu
65 70 75 80
Pro Val Ile Gly Tyr Leu Asn Tyr Gly Gly Asp Arg Asn Thr Phe Arg

85 90 95
Gln Lys Leu Ser Thr Leu Leu Lys Ala Val Asp Ser Leu Arg Asn Phe
100 105 110
Tyr Thr His Tyr Tyr His Ser Pro Leu Ala Leu Ser Thr Glu Leu Phe
115 120 125
Glu Leu Leu Asp Thr Val Phe Ala Ser Val Ala Val Glu Val Lys Gln
130 135 140
His Lys Met Lys Asp Asp Lys Thr Arg Gln Leu Leu Ser Lys Ser Leu

145 150 155 160
Ala Glu Glu Leu Asp Ile Arg Tyr Lys Gln Gln Leu Glu Arg Leu Lys
165 170 175
Glu Leu Lys Glu Gln Gly Lys Asn Ile Asp Leu Arg Asp Glu Ala Gly
180 185 190
Ile Arg Asn Gly Val Leu Asn Ala Ala Phe Asn His Leu Ile Tyr Lys
195 200 205
Glu Gly Glu Ile Ala Lys Pro Thr Leu Ser Tyr Ser Ser Phe Tyr Tyr

210 215 220
Gly Ala Asp Ser Ala Glu Asn Gly Ile Thr Ile Ser Gln Ser Gly Leu
225 230 235 240

Leu Phe Leu Leu Ser Met Phe Leu Gly Lys Lys Glu Ile Glu Asp Leu
245 250 255
Lys Ser Arg Ile Arg Gly Phe Lys Ala Lys Ile Val Arg Asp Gly Glu
260 265 270
Glu Asn Ile Ser Gly Leu Lys Phe Met Ala Thr His Trp Ile Phe Ser

275 280 285
Tyr Leu Ser Phe Lys Gly Met Lys Gln Arg Leu Ser Thr Asp Phe His
290 295 300
Glu Glu Thr Leu Leu Ile Gln Ile Ile Asp Glu Leu Ser Lys Val Pro
305 310 315 320
Asp Glu Val Tyr His Asp Phe Asp Thr Ala Thr Arg Glu Lys Phe Val
325 330 335
Glu Asp Ile Asn Glu Tyr Ile Arg Glu Gly Asn Glu Asp Phe Ser Leu

340 345 350
Gly Asp Ser Thr Ile Ile His Pro Val Ile Arg Lys Arg Tyr Glu Asn
355 360 365
Lys Phe Asn Tyr Phe Ala Val Arg Phe Leu Asp Glu Phe Ile Lys Phe
370 375 380
Pro Ser Leu Arg Phe Gln Val His Leu Gly Asn Phe Val His Asp Arg
385 390 395 400
Arg Ile Lys Asp Ile His Gly Thr Gly Phe Gln Thr Glu Arg Val Val

405 410 415
Lys Asp Arg Ile Lys Val Phe Gly Lys Leu Ser Glu Ile Ser Ser Leu
420 425 430
Lys Thr Glu Tyr Ile Glu Lys Glu Leu Asp Leu Asp Ser Asp Thr Gly
435 440 445
Trp Glu Ile Phe Pro Asn Pro Ser Tyr Val Phe Ile Asp Asn Asn Ile
450 455 460
Pro Ile Tyr Ile Ser Thr Asn Lys Thr Phe Lys Asn Gly Ser Ser Glu

465 470 475 480
Phe Ile Lys Leu Arg Arg Lys Glu Lys Pro Glu Glu Met Lys Met Arg

485 490 495

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

500 505 510

Ala Gly Ser Leu Asn Ser Lys Thr Pro Leu Ala Met Leu Ser Leu Asn

515 520 525

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

530 535 540

Glu Glu Ile Glu Leu Ile Ile Lys Glu Lys Leu Asp Ser His Phe Glu

545 550 555 560

Asn Ile Lys Asn Tyr Asp Pro Glu Lys Pro Leu Pro Ala Ser Gln Ile

565 570 575

Ser Lys Arg Leu Arg Asn Asn Thr Thr Asp Lys Gly Lys Lys Val Ile

580 585 590

Asn Pro Glu Lys Leu Ile His Leu Ile Asn Lys Glu Ile Asp Ala Thr

595 600 605

Glu Ala Lys Phe Ala Leu Leu Ala Lys Asn Arg Lys Glu Leu Lys Glu

610 615 620

Lys Phe Arg Gly Lys Pro Leu Arg Gln Thr Ile Phe Ser Asn Met Glu

625 630 635 640

Leu Gly Arg Glu Ala Thr Trp Leu Ala Asp Asp Ile Lys Arg Phe Met

645 650 655

Pro Asp Ile Leu Arg Lys Asn Trp Lys Gly Tyr Gln His Asn Gln Leu

660 665 670

Gln Gln Ser Leu Ala Phe Phe Asn Ser Arg Pro Lys Glu Ala Phe Thr

675 680 685

Ile Leu Gln Asp Gly Trp Asp Phe Ala Asp Gly Ser Ser Phe Trp Asn

690 695 700

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

705 710 715 720

Tyr Glu Ala Tyr Phe Glu Gly Arg Lys Glu Tyr Phe Ser Ser Leu Ala

725 730 735

Glu Asn Ile Lys Gln His Thr Ser Asn His Arg Asn Leu Arg Arg Phe
740 745 750
Ile Asp Gln Gln Met Pro Lys Gly Leu Phe Glu Asn Arg His Tyr Leu
755 760 765
Leu Glu Asn Leu Glu Thr Glu Lys Asn Lys Ile Leu Ser Lys Pro Leu
770 775 780
Val Phe Pro Arg Gly Leu Phe Asp Thr Lys Pro Thr Phe Ile Lys Gly

785 790 795 800
Ile Lys Val Asp Glu Gln Pro Glu Leu Phe Ala Glu Trp Tyr Gln Tyr
805 810 815
Gly Tyr Ser Thr Glu His Val Phe Gln Asn Phe Tyr Gly Trp Glu Arg
820 825 830
Asp Tyr Asn Asp Leu Leu Glu Ser Glu Leu Glu Lys Asp Asn Asp Phe
835 840 845
Ser Lys Asn Ser Ile His Tyr Ser Arg Thr Ser Gln Leu Glu Leu Ile

850 855 860
Lys Leu Lys Gln Asp Leu Lys Ile Lys Lys Ile Lys Ile Gln Asp Leu
865 870 875 880
Phe Leu Lys Leu Ile Ala Gly His Ile Phe Glu Asn Ile Phe Lys Tyr
885 890 895
Pro Ala Ser Phe Ser Leu Asp Glu Leu Tyr Leu Thr Gln Glu Glu Arg
900 905 910
Leu Asn Lys Glu Gln Glu Ala Leu Ile Gln Ser Gln Arg Lys Glu Gly

915 920 925
Asp His Ser Asp Asn Ile Ile Lys Asp Asn Phe Ile Gly Ser Lys Thr
930 935 940
Val Thr Tyr Glu Ser Lys Gln Ile Ser Glu Pro Asn Val Lys Leu Lys
945 950 955 960
Asp Ile Gly Lys Phe Asn Arg Phe Leu Leu Asp Asp Lys Val Lys Thr
965 970 975
Leu Leu Ser Tyr Asn Glu Asp Lys Val Trp Asn Lys Asn Asp Leu Asp

980 985 990

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

995 1000 1005

Lys Leu Phe Lys Lys Ile Gln Asn Phe Glu Leu Gln Thr Leu Thr

1010 1015 1020

Asp Trp Pro Trp Asn Gly Thr Asp His Pro Glu Glu Phe Gly Thr

1025 1030 1035

Thr Asp Asn Lys Gly Val Asn His Pro Asn Phe Lys Met Tyr Val

1040 1045 1050

Val Asn Gly Ile Leu Arg Lys His Thr Asp Trp Phe Lys Glu Gly

1055 1060 1065

Glu Asp Asn Trp Leu Glu Asn Leu Asn Glu Thr His Phe Lys Asn

1070 1075 1080

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

1085 1090 1095

Ala Phe Leu Ile Ile Met Ile Arg Asn Gln Phe Ala His Asn Gln

1100 1105 1110

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

1115 1120 1125

Glu Ile Gln Gly Ser Thr Thr Ser Glu Leu Tyr Leu Asn Phe Ile

1130 1135 1140

Asn Leu Ala Val Val Glu Leu Leu Glu Leu Leu Glu Lys

1145 1150 1155

<210> 106

<211> 1036

<212> PRT

<213> Chryseobacterium ureilyticum

<400> 106

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

1 5 10 15

Thr Glu Asp Lys His Phe Phe Gly Gly Phe Leu Asn Thr Ala Gln Asn

20 25 30

Asn Ile Asp Leu Leu Ile Lys Ala Tyr Ile Ser Lys Phe Glu Ser Ser
35 40 45
Pro Arg Lys Leu Asn Ser Val Gln Phe Pro Asp Val Cys Phe Lys Lys
50 55 60

Asn Asp Ser Asp Ala Asp Phe Gln His Lys Leu Gln Phe Ile Arg Lys
65 70 75 80
His Leu Pro Val Ile Gln Tyr Leu Lys Tyr Gly Gly Asn Arg Glu Val
85 90 95
Leu Lys Glu Lys Phe Arg Leu Leu Leu Gln Ala Val Asp Ser Leu Arg
100 105 110
Asn Phe Tyr Thr His Phe Tyr His Lys Pro Ile Gln Leu Pro Asn Glu
115 120 125

Leu Leu Thr Leu Leu Asp Thr Ile Phe Gly Glu Ile Gly Asn Glu Val
130 135 140
Arg Gln Asn Lys Met Lys Asp Asp Lys Thr Arg His Leu Leu Lys Lys
145 150 155 160
Asn Leu Ser Glu Glu Leu Asp Phe Arg Tyr Gln Glu Gln Leu Glu Arg
165 170 175
Leu Arg Lys Leu Lys Ser Glu Gly Lys Lys Val Asp Leu Arg Asp Thr
180 185 190

Glu Ala Ile Arg Asn Gly Val Leu Asn Ala Ala Phe Asn His Leu Ile
195 200 205
Phe Lys Asp Ala Glu Asp Phe Lys Pro Thr Val Ser Tyr Ser Ser Tyr
210 215 220
Tyr Tyr Asp Ser Asp Thr Ala Glu Asn Gly Ile Ser Ile Ser Gln Ser
225 230 235 240
Gly Leu Leu Phe Leu Leu Ser Met Phe Leu Gly Arg Arg Glu Met Glu
245 250 255

Asp Leu Lys Ser Arg Val Arg Gly Phe Lys Ala Arg Ile Ile Lys His
260 265 270
Glu Glu Gln His Val Ser Gly Leu Lys Phe Met Ala Thr His Trp Val

275 280 285

Phe Ser Glu Phe Cys Phe Lys Gly Ile Lys Thr Arg Leu Asn Ala Asp

290 295 300

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

305 310 315 320

Val Pro Asp Glu Leu Tyr Arg Ser Phe Asp Val Ala Thr Arg Glu Arg

325 330 335

Phe Ile Glu Asp Ile Asn Glu Tyr Ile Arg Asp Gly Lys Glu Asp Lys

340 345 350

Ser Leu Ile Glu Ser Lys Ile Val His Pro Val Ile Arg Lys Arg Tyr

355 360 365

Glu Ser Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Val

370 375 380

Asn Phe Pro Thr Leu Arg Phe Gln Val His Ala Gly Asn Tyr Val His

385 390 395 400

Asp Arg Arg Ile Lys Ser Ile Glu Gly Thr Gly Phe Lys Thr Glu Arg

405 410 415

Leu Val Lys Asp Arg Ile Lys Val Phe Gly Lys Leu Ser Thr Ile Ser

420 425 430

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

435 440 445

Thr Gly Trp Glu Leu Leu Pro His Pro Ser Tyr Val Phe Ile Asp Asn

450 455 460

Asn Ile Pro Ile His Leu Thr Val Asp Pro Ser Phe Lys Asn Gly Val

465 470 475 480

Lys Glu Tyr Gln Glu Lys Arg Lys Leu Gln Lys Pro Glu Glu Met Lys

485 490 495

Asn Arg Gln Gly Gly Asp Lys Met His Lys Pro Ala Ile Ser Ser Lys

500 505 510

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

515 520 525

Ser Met Asn Glu Ile Pro Ala Leu Leu Tyr Glu Ile Leu Val Lys Lys
530 535 540
Ala Ser Pro Glu Glu Val Glu Ala Lys Ile Arg Gln Lys Leu Thr Ala
545 550 555 560
Val Phe Glu Arg Ile Arg Asp Tyr Asp Pro Lys Val Pro Leu Pro Ala
565 570 575

Ser Gln Val Ser Lys Arg Leu Arg Asn Asn Thr Asp Thr Leu Ser Tyr
580 585 590
Asn Lys Glu Lys Leu Val Glu Leu Ala Asn Lys Glu Val Glu Gln Thr
595 600 605
Glu Arg Lys Leu Ala Leu Ile Thr Lys Asn Arg Arg Glu Cys Arg Glu
610 615 620
Lys Val Lys Gly Lys Phe Lys Arg Gln Lys Val Phe Lys Asn Ala Glu
625 630 635 640

Leu Gly Thr Glu Ala Thr Trp Leu Ala Asn Asp Ile Lys Arg Phe Met
645 650 655
Pro Glu Glu Gln Lys Lys Asn Trp Lys Gly Tyr Gln His Ser Gln Leu
660 665 670
Gln Gln Ser Leu Ala Phe Phe Glu Ser Arg Pro Gly Glu Ala Arg Ser
675 680 685
Leu Leu Gln Ala Gly Trp Asp Phe Ser Asp Gly Ser Ser Phe Trp Asn
690 695 700

Gly Trp Val Met Asn Ser Phe Ala Arg Asp Asn Thr Phe Asp Gly Phe
705 710 715 720
Tyr Glu Ser Tyr Leu Asn Gly Arg Met Lys Tyr Phe Leu Arg Leu Ala
725 730 735
Asp Asn Ile Ala Gln Gln Ser Ser Thr Asn Lys Leu Ile Ser Asn Phe
740 745 750
Ile Lys Gln Gln Met Pro Lys Gly Leu Phe Asp Arg Arg Leu Tyr Met
755 760 765

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

770 775 780

Ile Phe Pro Arg Gly Ile Phe Asp Asp Lys Pro Thr Phe Lys Lys Gly

785 790 795 800

Val Gln Val Ser Glu Glu Pro Glu Ala Phe Ala Asp Trp Tyr Ser Tyr

805 810 815

Gly Tyr Asp Val Lys His Lys Phe Gln Glu Phe Tyr Ala Trp Asp Arg

820 825 830

Asp Tyr Glu Glu Leu Leu Arg Glu Glu Leu Glu Lys Asp Thr Ala Phe

835 840 845

Thr Lys Asn Ser Ile His Tyr Ser Arg Glu Ser Gln Ile Glu Leu Leu

850 855 860

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

865 870 875 880

Tyr Leu Lys Leu Met Ala Glu Phe Leu Phe Glu Asn Val Phe Gly His

885 890 895

Glu Leu Ala Leu Pro Leu Asp Gln Phe Tyr Leu Thr Gln Glu Glu Arg

900 905 910

Leu Lys Gln Glu Gln Glu Ala Ile Val Gln Ser Gln Arg Pro Lys Gly

915 920 925

Asp Asp Ser Pro Asn Ile Val Lys Glu Asn Phe Ile Trp Ser Lys Thr

930 935 940

Ile Pro Phe Lys Ser Gly Arg Val Phe Glu Pro Asn Val Lys Leu Lys

945 950 955 960

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

965 970 975

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

980 985 990

Glu Leu Ile Ile Gly Ala Gly Ser Tyr Glu Phe Ile Arg Arg Glu Gln

995 1000 1005

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

1010 1015 1020

Val Arg Gly Met Gly Val Pro Ile Arg Leu Asn Leu Lys

1025 1030 1035

<210> 107

<211> 1127

<212> PRT

<213> Prevotella buccae

<400> 107

Met Gln Lys Gln Asp Lys Leu Phe Val Asp Arg Lys Lys Asn Ala Ile

1 5 10 15

Phe Ala Phe Pro Lys Tyr Ile Thr Ile Met Glu Asn Gln Glu Lys Pro

20 25 30

Glu Pro Ile Tyr Tyr Glu Leu Thr Asp Lys His Phe Trp Ala Ala Phe

35 40 45

Leu Asn Leu Ala Arg His Asn Val Tyr Thr Thr Ile Asn His Ile Asn

50 55 60

Arg Arg Leu Glu Ile Ala Glu Leu Lys Asp Asp Gly Tyr Met Met Asp

65 70 75 80

Ile Lys Gly Ser Trp Asn Glu Gln Ala Lys Lys Leu Asp Lys Lys Val

85 90 95

Arg Leu Arg Asp Leu Ile Met Lys His Phe Pro Phe Leu Glu Ala Ala

100 105 110

Ala Tyr Glu Ile Thr Asn Ser Lys Ser Pro Asn Asn Lys Glu Gln Arg

115 120 125

Glu Lys Glu Gln Ser Glu Ala Leu Ser Leu Asn Asn Leu Lys Asn Val

130 135 140

Leu Phe Ile Phe Leu Glu Lys Leu Gln Val Leu Arg Asn Tyr Tyr Ser

145 150 155 160

His Tyr Lys Tyr Ser Glu Glu Ser Pro Lys Pro Ile Phe Glu Thr Ser

165 170 175

Leu Leu Lys Asn Met Tyr Lys Val Phe Asp Ala Asn Val Arg Leu Val

180 185 190

Lys Arg Asp Tyr Met His His Glu Asn Ile Asp Met Gln Arg Asp Phe

195 200 205

Thr His Leu Asn Arg Lys Lys Gln Val Gly Arg Thr Lys Asn Ile Ile

210 215 220

Asp Ser Pro Asn Phe His Tyr His Phe Ala Asp Lys Glu Gly Asn Met

225 230 235 240

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

245 250 255

Asp Ala Ile Trp Met Gln Lys Lys Leu Lys Gly Phe Lys Asp Gly Arg

260 265 270

Asn Leu Arg Glu Gln Met Thr Asn Glu Val Phe Cys Arg Ser Arg Ile

275 280 285

Ser Leu Pro Lys Leu Lys Leu Glu Asn Val Gln Thr Lys Asp Trp Met

290 295 300

Gln Leu Asp Met Leu Asn Glu Leu Val Arg Cys Pro Lys Ser Leu Tyr

305 310 315 320

Glu Arg Leu Arg Glu Lys Asp Arg Glu Ser Phe Lys Val Pro Phe Asp

325 330 335

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

340 345 350

Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Val Leu Arg Tyr Phe

355 360 365

Asp Leu Asn Glu Ile Phe Glu Gln Leu Arg Phe Gln Ile Asp Leu Gly

370 375 380

Thr Tyr His Phe Ser Ile Tyr Asn Lys Arg Ile Gly Asp Glu Asp Glu

385 390 395 400

Val Arg His Leu Thr His His Leu Tyr Gly Phe Ala Arg Ile Gln Asp

405 410 415

Phe Ala Gln Gln Asn Gln Pro Glu Val Trp Arg Lys Leu Val Lys Asp

420 425 430

Leu Asp Tyr Phe Glu Ala Ser Gln Glu Pro Tyr Ile Pro Lys Thr Ala

435 440 445

Pro His Tyr His Leu Glu Asn Glu Lys Ile Gly Ile Lys Phe Cys Ser
450 455 460
Thr His Asn Asn Leu Phe Pro Ser Leu Lys Thr Glu Lys Thr Cys Asn
465 470 475 480
Gly Arg Ser Lys Phe Asn Leu Gly Thr Gln Phe Thr Ala Glu Ala Phe

485 490 495
Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu
500 505 510
Thr Lys Asp Tyr Ser Arg Lys Glu Ser Ala Asp Lys Val Glu Gly Ile
515 520 525
Ile Arg Lys Glu Ile Ser Asn Ile Tyr Ala Ile Tyr Asp Ala Phe Ala
530 535 540
Asn Gly Glu Ile Asn Ser Ile Ala Asp Leu Thr Cys Arg Leu Gln Lys

545 550 555 560
Thr Asn Ile Leu Gln Gly His Leu Pro Lys Gln Met Ile Ser Ile Leu
565 570 575
Glu Gly Arg Gln Lys Asp Met Glu Lys Glu Ala Glu Arg Lys Ile Gly
580 585 590
Glu Met Ile Asp Asp Thr Gln Arg Arg Leu Asp Leu Leu Cys Lys Gln
595 600 605
Thr Asn Gln Lys Ile Arg Ile Gly Lys Arg Asn Ala Gly Leu Leu Lys

610 615 620
Ser Gly Lys Ile Ala Asp Trp Leu Val Asn Asp Met Met Arg Phe Gln
625 630 635 640
Pro Val Gln Lys Asp Gln Asn Asn Ile Pro Ile Asn Asn Ser Lys Ala
645 650 655
Asn Ser Thr Glu Tyr Arg Met Leu Gln Arg Ala Leu Ala Leu Phe Gly
660 665 670
Ser Glu Asn Phe Arg Leu Lys Ala Tyr Phe Asn Gln Met Asn Leu Val

675 680 685
Gly Asn Asp Asn Pro His Pro Phe Leu Ala Glu Thr Gln Trp Glu His

690 695 700

Gln Thr Asn Ile Leu Ser Phe Tyr Arg Asn Tyr Leu Glu Ala Arg Lys

705 710 715 720

Lys Tyr Leu Lys Gly Leu Lys Pro Gln Asn Trp Lys Gln Tyr Gln His

725 730 735

Phe Leu Ile Leu Lys Val Gln Lys Thr Asn Arg Asn Thr Leu Val Thr

740 745 750

Gly Trp Lys Asn Ser Phe Asn Leu Pro Arg Gly Ile Phe Thr Gln Pro

755 760 765

Ile Arg Glu Trp Phe Glu Lys His Asn Asn Ser Lys Arg Ile Tyr Asp

770 775 780

Gln Ile Leu Ser Phe Asp Arg Val Gly Phe Val Ala Lys Ala Ile Pro

785 790 795 800

Leu Tyr Phe Ala Glu Glu Tyr Lys Asp Asn Val Gln Pro Phe Tyr Asp

805 810 815

Tyr Pro Phe Asn Ile Gly Asn Lys Leu Lys Pro Gln Lys Gly Gln Phe

820 825 830

Leu Asp Lys Lys Glu Arg Val Glu Leu Trp Gln Lys Asn Lys Glu Leu

835 840 845

Phe Lys Asn Tyr Pro Ser Glu Lys Lys Lys Thr Asp Leu Ala Tyr Leu

850 855 860

Asp Phe Leu Ser Trp Lys Lys Phe Glu Arg Glu Leu Arg Leu Ile Lys

865 870 875 880

Asn Gln Asp Ile Val Thr Trp Leu Met Phe Lys Glu Leu Phe Asn Met

885 890 895

Ala Thr Val Glu Gly Leu Lys Ile Gly Glu Ile His Leu Arg Asp Ile

900 905 910

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

915 920 925

Met Pro Met Lys Leu Pro Val Lys Thr Tyr Glu Thr Asp Asn Lys Gly

930 935 940

Asn Ile Leu Lys Glu Arg Pro Leu Ala Thr Phe Tyr Ile Glu Glu Thr
945 950 955 960
Glu Thr Lys Val Leu Lys Gln Gly Asn Phe Lys Val Leu Ala Lys Asp
965 970 975
Arg Arg Leu Asn Gly Leu Leu Ser Phe Ala Glu Thr Thr Asp Ile Asp
980 985 990
Leu Glu Lys Asn Pro Ile Thr Lys Leu Ser Val Asp His Glu Leu Ile

995 1000 1005
Lys Tyr Gln Thr Thr Arg Ile Ser Ile Phe Glu Met Thr Leu Gly
1010 1015 1020
Leu Glu Lys Lys Leu Ile Asn Lys Tyr Pro Thr Leu Pro Thr Asp
1025 1030 1035
Ser Phe Arg Asn Met Leu Glu Arg Trp Leu Gln Cys Lys Ala Asn
1040 1045 1050
Arg Pro Glu Leu Lys Asn Tyr Val Asn Ser Leu Ile Ala Val Arg

1055 1060 1065
Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp Ala Thr Leu
1070 1075 1080
Phe Ala Glu Val Lys Lys Phe Thr Leu Phe Pro Ser Val Asp Thr
1085 1090 1095
Lys Lys Ile Glu Leu Asn Ile Ala Pro Gln Leu Leu Glu Ile Val
1100 1105 1110
Gly Lys Ala Ile Lys Glu Ile Glu Lys Ser Glu Asn Lys Asn

1115 1120 1125
<210> 108
<211> 1135
<212> PRT
<213> Porphyromonas gingivalis
<400> 108
Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15
Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu

20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys

35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr

65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp

85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile

100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp

130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala

145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala

165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg

195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu
275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu
290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr
305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe
325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu
340 345 350

Gly Glu Ile Glu Leu Asp Ser Tyr Ser Lys Lys Val Gly Arg Asn Gly
355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu
370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu
385 390 395 400

Ala Ser Tyr Pro Val Arg Phe Ser Leu Phe Ala Pro Arg Tyr Ala Ile
405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro
420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met
435 440 445

Gly Phe Ile Ser Val His Asn Leu Arg Lys Leu Leu Leu Met Glu Leu
450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys
465 470 475 480

Ala Asn Arg Ile Leu Asp Glu Thr Ala Glu Gly Lys Leu Gln Phe Ser
485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro
500 505 510

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

515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu

530 535 540

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

545 550 555 560

Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575

Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe

580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu

595 600 605

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

610 615 620

Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640

Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala

645 650 655

Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670

Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr

675 680 685

Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700

Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg

725 730 735

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

740 745 750

Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765

Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe
770 775 780
Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr
785 790 795 800
Gly Leu Arg Arg Glu Leu Asn Ile His Gly Lys Ser Val Ser Tyr Ile
805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu
820 825 830
Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro
835 840 845
Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala
850 855 860
Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg
865 870 875 880

Leu Met Leu Met Ala Ile Asn Lys Met Met Thr Asp Arg Glu Glu Asp
885 890 895
Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn
900 905 910
Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu
915 920 925
Gly Gly Asp Asn Ser Leu Ser Leu Val Pro Ala Thr Ile Glu Ile Lys
930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg
945 950 955 960
Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu
965 970 975
Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys
980 985 990
Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg
995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 109

<211> 1175

<212> PRT

<213> Porphyromonas gingivalis

<400> 109

Met Thr Glu Gln Asn Glu Lys Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Ala His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95
Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu
100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125
Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140
Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu
145 150 155 160
Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190
Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205
Lys Tyr Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp
210 215 220
Arg Glu Glu Lys Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly
245 250 255
Phe Lys Gly Gly Thr Glu Ala Tyr Gln Gln Met Thr Asn Glu Val Phe
260 265 270
Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285
Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe
305 310 315 320
Arg Val Pro Val Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr

325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln Leu Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Phe Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Ala

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

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

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

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

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685

Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700

Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg
705 710 715 720

Ser Asp Arg Glu Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu
740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys
770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro
785 790 795 800

Phe Tyr Asp Tyr Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg

835 840 845

Arg Ile Glu Asp Ala Phe Val Gly Ile Glu Tyr Ala Ser Trp Glu Asn

850 855 860

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

865 870 875 880

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

885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Lys Glu His Pro Tyr His Asp Phe

900 905 910

Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln

915 920 925

Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys

930 935 940

Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Thr

945 950 955 960

Asp Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn His Val Lys Pro

965 970 975

Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val

980 985 990

His Lys Glu Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp

995 1000 1005

Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp

1010 1015 1020

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

1025 1030 1035

Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu

1040 1045 1050

Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala Phe Glu Gln Thr

1055 1060 1065

Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro
1070 1075 1080
Asp Glu Ser Phe Arg Glu Met Leu Glu Ser Trp Ser Asp Pro Leu
1085 1090 1095
Leu Asp Lys Trp Pro Asp Leu Gln Arg Glu Val Arg Leu Leu Ile
1100 1105 1110
Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp
1115 1120 1125

Glu Thr Ile Phe Ser Ser Ile Arg Lys Tyr Asp Pro Ser Ser Leu
1130 1135 1140
Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu
1145 1150 1155
Ser Glu Glu Val Lys Leu Ala Lys Glu Met Val Glu Arg Ile Ile
1160 1165 1170
Gln Ala
1175
<210> 110
<211> 1125
<212> PRT
<213> Prevotella pallens
<400> 110
Met Lys Glu Glu Glu Lys Gly Lys Thr Pro Val Val Ser Thr Tyr Asn

1 5 10 15
Lys Asp Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His
20 25 30
Asn Val Tyr Ile Thr Val Asn His Ile Asn Lys Ile Leu Gly Glu Gly
35 40 45
Glu Ile Asn Arg Asp Gly Tyr Glu Asn Thr Leu Glu Lys Ser Trp Asn
50 55 60
Glu Ile Lys Asp Ile Asn Lys Lys Asp Arg Leu Ser Lys Leu Ile Ile

65 70 75 80
Lys His Phe Pro Phe Leu Glu Val Thr Thr Tyr Gln Arg Asn Ser Ala

85 90 95

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

100 105 110

Glu Ser Leu Lys Lys Ser Phe Phe Val Phe Ile Tyr Lys Leu Arg Asp

115 120 125

Leu Arg Asn His Tyr Ser His Tyr Lys His Ser Lys Ser Leu Glu Arg

130 135 140

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

145 150 155 160

Ala Ser Ile Gln Leu Val Lys Glu Asp Tyr Lys His Asn Thr Asp Ile

165 170 175

Lys Thr Glu Glu Asp Phe Lys His Leu Asp Arg Lys Gly Gln Phe Lys

180 185 190

Tyr Ser Phe Ala Asp Asn Glu Gly Asn Ile Thr Glu Ser Gly Leu Leu

195 200 205

Phe Phe Val Ser Leu Phe Leu Glu Lys Lys Asp Ala Ile Trp Val Gln

210 215 220

Lys Lys Leu Glu Gly Phe Lys Cys Ser Asn Glu Ser Tyr Gln Lys Met

225 230 235 240

Thr Asn Glu Val Phe Cys Arg Ser Arg Met Leu Leu Pro Lys Leu Arg

245 250 255

Leu Gln Ser Thr Gln Thr Gln Asp Trp Ile Leu Leu Asp Met Leu Asn

260 265 270

Glu Leu Ile Arg Cys Pro Lys Ser Leu Tyr Glu Arg Leu Arg Glu Glu

275 280 285

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

290 295 300

Asp Ala Glu Gln Glu Pro Phe Lys Asn Ala Leu Val Arg His Gln Asp

305 310 315 320

Arg Phe Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Tyr Asn Glu Ile Phe

325 330 335

Thr Asn Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ser Ile
340 345 350
Tyr Lys Lys Gln Ile Gly Asp Tyr Lys Glu Ser His His Leu Thr His
355 360 365
Lys Leu Tyr Gly Phe Glu Arg Ile Gln Glu Phe Thr Lys Gln Asn Arg
370 375 380
Pro Asp Glu Trp Arg Lys Phe Val Lys Thr Phe Asn Ser Phe Glu Thr

385 390 395 400
Ser Lys Glu Pro Tyr Ile Pro Glu Thr Thr Pro His Tyr His Leu Glu
405 410 415
Asn Gln Lys Ile Gly Ile Arg Phe Arg Asn Asp Asn Asp Lys Ile Trp
420 425 430
Pro Ser Leu Lys Thr Asn Ser Glu Lys Asn Glu Lys Ser Lys Tyr Lys
435 440 445
Leu Asp Lys Ser Phe Gln Ala Glu Ala Phe Leu Ser Val His Glu Leu

450 455 460
Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu Lys Thr Glu Asn Thr Asp
465 470 475 480
Asn Asp Asn Glu Ile Glu Thr Lys Lys Lys Glu Asn Lys Asn Asp Lys
485 490 495
Gln Glu Lys His Lys Ile Glu Glu Ile Ile Glu Asn Lys Ile Thr Glu
500 505 510
Ile Tyr Ala Leu Tyr Asp Ala Phe Ala Asn Gly Lys Ile Asn Ser Ile

515 520 525
Asp Lys Leu Glu Glu Tyr Cys Lys Gly Lys Asp Ile Glu Ile Gly His
530 535 540
Leu Pro Lys Gln Met Ile Ala Ile Leu Lys Ser Glu His Lys Asp Met
545 550 555 560
Ala Thr Glu Ala Lys Arg Lys Gln Glu Glu Met Leu Ala Asp Val Gln
565 570 575
Lys Ser Leu Glu Ser Leu Asp Asn Gln Ile Asn Glu Glu Ile Glu Asn

580 585 590

Val Glu Arg Lys Asn Ser Ser Leu Lys Ser Gly Glu Ile Ala Ser Trp

595 600 605

Leu Val Asn Asp Met Met Arg Phe Gln Pro Val Gln Lys Asp Asn Glu

610 615 620

Gly Asn Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met

625 630 635 640

Leu Gln Arg Ser Leu Ala Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg

645 650 655

Tyr Phe Arg Gln Val Asn Leu Ile Glu Ser Ser Asn Pro His Pro Phe

660 665 670

Leu Asn Asn Thr Glu Trp Glu Lys Cys Asn Asn Ile Leu Ser Phe Tyr

675 680 685

Arg Ser Tyr Leu Glu Ala Lys Lys Asn Phe Leu Glu Ser Leu Lys Pro

690 695 700

Glu Asp Trp Glu Lys Asn Gln Tyr Phe Leu Met Leu Lys Glu Pro Lys

705 710 715 720

Thr Asn Cys Glu Thr Leu Val Gln Gly Trp Lys Asn Gly Phe Asn Leu

725 730 735

Pro Arg Gly Ile Phe Thr Glu Pro Ile Arg Lys Trp Phe Met Glu His

740 745 750

Arg Lys Asn Ile Thr Val Ala Glu Leu Lys Arg Val Gly Leu Val Ala

755 760 765

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

770 775 780

Pro Phe Tyr Asn Tyr Leu Phe Asn Val Gly Asn Ile Asn Lys Pro Asp

785 790 795 800

Glu Lys Asn Phe Leu Asn Cys Glu Glu Arg Arg Glu Leu Leu Arg Lys

805 810 815

Lys Lys Asp Glu Phe Lys Lys Met Thr Asp Lys Glu Lys Glu Glu Asn

820 825 830

Pro Ser Tyr Leu Glu Phe Gln Ser Trp Asn Lys Phe Glu Arg Glu Leu

835 840 845

Arg Leu Val Arg Asn Gln Asp Ile Val Thr Trp Leu Leu Cys Met Glu

850 855 860

Leu Phe Asn Lys Lys Lys Ile Lys Glu Leu Asn Val Glu Lys Ile Tyr

865 870 875 880

Leu Lys Asn Ile Asn Thr Asn Thr Thr Lys Lys Glu Lys Asn Thr Glu

885 890 895

Glu Lys Asn Gly Glu Glu Lys Ile Ile Lys Glu Lys Asn Asn Ile Leu

900 905 910

Asn Arg Ile Met Pro Met Arg Leu Pro Ile Lys Val Tyr Gly Arg Glu

915 920 925

Asn Phe Ser Lys Asn Lys Lys Lys Lys Ile Arg Arg Asn Thr Phe Phe

930 935 940

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

945 950 955 960

Phe Lys Ala Leu Glu Arg Asp Arg Arg Leu Gly Gly Leu Phe Ser Phe

965 970 975

Val Lys Thr His Ser Lys Ala Glu Ser Lys Ser Asn Thr Ile Ser Lys

980 985 990

Ser Arg Val Glu Tyr Glu Leu Gly Glu Tyr Gln Lys Ala Arg Ile Glu

995 1000 1005

Ile Ile Lys Asp Met Leu Ala Leu Glu Glu Thr Leu Ile Asp Lys

1010 1015 1020

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

1025 1030 1035

Trp Leu Lys Leu Lys Asp Glu Pro Asp Lys Ala Ser Phe Gln Asn

1040 1045 1050

Asp Val Asp Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn

1055 1060 1065

Gln Tyr Pro Met Arg Asn Arg Ile Ala Phe Ala Asn Ile Asn Pro

1070 1075 1080

Phe Ser Leu Ser Ser Ala Asn Thr Ser Glu Glu Lys Gly Leu Gly

1085 1090 1095

Ile Ala Asn Gln Leu Lys Asp Lys Thr His Lys Thr Ile Glu Lys

1100 1105 1110

Ile Ile Glu Ile Glu Lys Pro Ile Glu Thr Lys Glu

1115 1120 1125

<210> 111

<211> 1157

<212> PRT

<213> Myroides odoratimimus

<400> 111

Met Lys Asp Ile Leu Thr Thr Asp Thr Thr Glu Lys Gln Asn Arg Phe

1 5 10 15

Tyr Ser His Lys Ile Ala Asp Lys Tyr Phe Phe Gly Gly Tyr Phe Asn

20 25 30

Leu Ala Ser Asn Asn Ile Tyr Glu Val Phe Glu Glu Val Asn Lys Arg

35 40 45

Asn Thr Phe Gly Lys Leu Ala Lys Arg Asp Asn Gly Asn Leu Lys Asn

50 55 60

Tyr Ile Ile His Val Phe Lys Asp Glu Leu Ser Ile Ser Asp Phe Glu

65 70 75 80

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

85 90 95

Asp Lys Lys Ser Ile Lys Glu Arg Asn Arg Thr Ile Asp Leu Thr Leu

100 105 110

Ser Gln Arg Ile Arg Gln Phe Arg Glu Met Leu Ile Ser Leu Val Thr

115 120 125

Ala Val Asp Gln Leu Arg Asn Phe Tyr Thr His Tyr His His Ser Glu

130 135 140

Ile Val Ile Glu Asn Lys Val Leu Asp Phe Leu Asn Ser Ser Leu Val

145 150 155 160

Ser Thr Ala Leu His Val Lys Asp Lys Tyr Leu Lys Thr Asp Lys Thr

165 170 175

Lys Glu Phe Leu Lys Glu Thr Ile Ala Ala Glu Leu Asp Ile Leu Ile

180 185 190

Glu Ala Tyr Lys Lys Lys Gln Ile Glu Lys Lys Asn Thr Arg Phe Lys

195 200 205

Ala Asn Lys Arg Glu Asp Ile Leu Asn Ala Ile Tyr Asn Glu Ala Phe

210 215 220

Trp Ser Phe Ile Asn Asp Lys Asp Lys Asp Lys Glu Thr Val Val Ala

225 230 235 240

Lys Gly Ala Asp Ala Tyr Phe Glu Lys Asn His His Lys Ser Asn Asp

245 250 255

Pro Asp Phe Ala Leu Asn Ile Ser Glu Lys Gly Ile Val Tyr Leu Leu

260 265 270

Ser Phe Phe Leu Thr Asn Lys Glu Met Asp Ser Leu Lys Ala Asn Leu

275 280 285

Thr Gly Phe Lys Gly Lys Val Asp Arg Glu Ser Gly Asn Ser Ile Lys

290 295 300

Tyr Met Ala Thr Gln Arg Ile Tyr Ser Phe His Thr Tyr Arg Gly Leu

305 310 315 320

Lys Gln Lys Ile Arg Thr Ser Glu Glu Gly Val Lys Glu Thr Leu Leu

325 330 335

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

340 345 350

His Leu Ser Thr Thr Gln Gln Asn Ser Phe Ile Glu Asp Trp Asn Glu

355 360 365

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

370 375 380

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

385 390 395 400

Phe Ala Ile Arg Phe Leu Asp Glu Phe Phe Asp Phe Pro Thr Leu Arg
405 410 415
Phe Gln Val His Leu Gly Asp Tyr Val His Asp Arg Arg Thr Lys Gln
420 425 430
Leu Gly Lys Val Glu Ser Asp Arg Ile Ile Lys Glu Lys Val Thr Val
435 440 445
Phe Ala Arg Leu Lys Asp Ile Asn Ser Ala Lys Ala Asn Tyr Phe His
450 455 460

Ser Leu Glu Glu Gln Asp Lys Glu Glu Leu Asp Asn Lys Trp Thr Leu
465 470 475 480
Phe Pro Asn Pro Ser Tyr Asp Phe Pro Lys Glu His Thr Leu Gln His
485 490 495
Gln Gly Glu Gln Lys Asn Ala Gly Lys Ile Gly Ile Tyr Val Lys Leu
500 505 510
Arg Asp Thr Gln Tyr Lys Glu Lys Ala Ala Leu Glu Glu Ala Arg Lys
515 520 525

Ser Leu Asn Pro Lys Glu Arg Ser Ala Thr Lys Ala Ser Lys Tyr Asp
530 535 540
Ile Ile Thr Gln Ile Ile Glu Ala Asn Asp Asn Val Lys Ser Glu Lys
545 550 555 560
Pro Leu Val Phe Thr Gly Gln Pro Ile Ala Tyr Leu Ser Met Asn Asp
565 570 575
Ile His Ser Met Leu Phe Ser Leu Leu Thr Asp Asn Ala Glu Leu Lys
580 585 590

Lys Thr Pro Glu Glu Val Glu Ala Lys Leu Ile Asp Gln Ile Gly Lys
595 600 605
Gln Ile Asn Glu Ile Leu Ser Lys Asp Thr Asp Thr Lys Ile Leu Lys
610 615 620
Lys Tyr Lys Asp Asn Asp Leu Lys Glu Thr Asp Thr Asp Lys Ile Thr
625 630 635 640
Arg Asp Leu Ala Arg Asp Lys Glu Glu Ile Glu Lys Leu Ile Leu Glu

645 650 655

Gln Lys Gln Arg Ala Asp Asp Tyr Asn Tyr Thr Ser Ser Thr Lys Phe
660 665 670
Asn Ile Asp Lys Ser Arg Lys Arg Lys His Leu Leu Phe Asn Ala Glu
675 680 685
Lys Gly Lys Ile Gly Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met
690 695 700
Thr Glu Glu Phe Lys Ser Lys Trp Lys Gly Tyr Gln His Thr Glu Leu
705 710 715 720

Gln Lys Leu Phe Ala Tyr Tyr Asp Thr Ser Lys Ser Asp Leu Asp Leu
725 730 735
Ile Leu Ser Asp Met Val Met Val Lys Asp Tyr Pro Ile Glu Leu Ile
740 745 750
Ala Leu Val Lys Lys Ser Arg Thr Leu Val Asp Phe Leu Asn Lys Tyr
755 760 765
Leu Glu Ala Arg Leu Gly Tyr Met Glu Asn Val Ile Thr Arg Val Lys
770 775 780

Asn Ser Ile Gly Thr Pro Gln Phe Lys Thr Val Arg Lys Glu Cys Phe
785 790 795 800
Thr Phe Leu Lys Lys Ser Asn Tyr Thr Val Val Ser Leu Asp Lys Gln
805 810 815
Val Glu Arg Ile Leu Ser Met Pro Leu Phe Ile Glu Arg Gly Phe Met
820 825 830
Asp Asp Lys Pro Thr Met Leu Glu Gly Lys Ser Tyr Gln Gln His Lys
835 840 845

Glu Lys Phe Ala Asp Trp Phe Val His Tyr Lys Glu Asn Ser Asn Tyr
850 855 860
Gln Asn Phe Tyr Asp Thr Glu Val Tyr Glu Ile Thr Thr Glu Asp Lys
865 870 875 880
Arg Glu Lys Ala Lys Val Thr Lys Lys Ile Lys Gln Gln Gln Lys Asn
885 890 895

Asp Val Phe Thr Leu Met Met Val Asn Tyr Met Leu Glu Glu Val Leu
900 905 910

Lys Leu Ser Ser Asn Asp Arg Leu Ser Leu Asn Glu Leu Tyr Gln Thr
915 920 925
Lys Glu Glu Arg Ile Val Asn Lys Gln Val Ala Lys Asp Thr Gln Glu
930 935 940
Arg Asn Lys Asn Tyr Ile Trp Asn Lys Val Val Asp Leu Gln Leu Cys
945 950 955 960
Glu Gly Leu Val Arg Ile Asp Lys Val Lys Leu Lys Asp Ile Gly Asn
965 970 975

Phe Arg Lys Tyr Glu Asn Asp Ser Arg Val Lys Glu Phe Leu Thr Tyr
980 985 990
Gln Ser Asp Ile Val Trp Ser Ala Tyr Leu Ser Asn Glu Val Asp Ser
995 1000 1005
Asn Lys Leu Tyr Val Ile Glu Arg Gln Leu Asp Asn Tyr Glu Ser
1010 1015 1020
Ile Arg Ser Lys Glu Leu Leu Lys Glu Val Gln Glu Ile Glu Cys
1025 1030 1035

Ser Val Tyr Asn Gln Val Ala Asn Lys Glu Ser Leu Lys Gln Ser
1040 1045 1050
Gly Asn Glu Asn Phe Lys Gln Tyr Val Leu Gln Gly Leu Val Pro
1055 1060 1065
Ile Gly Met Asp Val Arg Glu Met Leu Ile Leu Ser Thr Asp Val
1070 1075 1080
Lys Phe Ile Lys Glu Glu Ile Ile Gln Leu Gly Gln Ala Gly Glu
1085 1090 1095

Val Glu Gln Asp Leu Tyr Ser Leu Ile Tyr Ile Arg Asn Lys Phe
1100 1105 1110
Ala His Asn Gln Leu Pro Ile Lys Glu Phe Phe Asp Phe Cys Glu
1115 1120 1125
Asn Asn Tyr Arg Ser Ile Ser Asp Asn Glu Tyr Tyr Ala Glu Tyr

1130 1135 1140

Tyr Met Glu Ile Phe Arg Ser Ile Lys Glu Lys Tyr Thr Ser

1145 1150 1155

<210> 112

<211> 1157

<212> PRT

<213> Myroides odoratimimus

<400> 112

Met Lys Asp Ile Leu Thr Thr Asp Thr Thr Glu Lys Gln Asn Arg Phe

1 5 10 15

Tyr Ser His Lys Ile Ala Asp Lys Tyr Phe Phe Gly Gly Tyr Phe Asn

20 25 30

Leu Ala Ser Asn Asn Ile Tyr Glu Val Phe Glu Glu Val Asn Lys Arg

35 40 45

Asn Thr Phe Gly Lys Leu Ala Lys Arg Asp Asn Gly Asn Leu Lys Asn

50 55 60

Tyr Ile Ile His Val Phe Lys Asp Glu Leu Ser Ile Ser Asp Phe Glu

65 70 75 80

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

85 90 95

Asp Lys Lys Ser Ile Lys Glu Arg Asn Arg Thr Ile Asp Leu Thr Leu

100 105 110

Ser Gln Arg Ile Arg Gln Phe Arg Glu Met Leu Ile Ser Leu Val Thr

115 120 125

Ala Val Asp Gln Leu Arg Asn Phe Tyr Thr His Tyr His His Ser Glu

130 135 140

Ile Val Ile Glu Asn Lys Val Leu Asp Phe Leu Asn Ser Ser Leu Val

145 150 155 160

Ser Thr Ala Leu His Val Lys Asp Lys Tyr Leu Lys Thr Asp Lys Thr

165 170 175

Lys Glu Phe Leu Lys Glu Thr Ile Ala Ala Glu Leu Asp Ile Leu Ile

180 185 190

Glu Ala Tyr Lys Lys Lys Gln Ile Glu Lys Lys Asn Thr Arg Phe Lys

195 200 205

Ala Asn Lys Arg Glu Asp Ile Leu Asn Ala Ile Tyr Asn Glu Ala Phe

210 215 220

Trp Ser Phe Ile Asn Asp Lys Asp Lys Asp Lys Glu Thr Val Val Ala

225 230 235 240

Lys Gly Ala Asp Ala Tyr Phe Glu Lys Asn His His Lys Ser Asn Asp

245 250 255

Pro Asp Phe Ala Leu Asn Ile Ser Glu Lys Gly Ile Val Tyr Leu Leu

260 265 270

Ser Phe Phe Leu Thr Asn Lys Glu Met Asp Ser Leu Lys Ala Asn Leu

275 280 285

Thr Gly Phe Lys Gly Lys Val Asp Arg Glu Ser Gly Asn Ser Ile Lys

290 295 300

Tyr Met Ala Thr Gln Arg Ile Tyr Ser Phe His Thr Tyr Arg Gly Leu

305 310 315 320

Lys Gln Lys Ile Arg Thr Ser Glu Glu Gly Val Lys Glu Thr Leu Leu

325 330 335

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

340 345 350

His Leu Ser Thr Thr Gln Gln Asn Ser Phe Ile Glu Asp Trp Asn Glu

355 360 365

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

370 375 380

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

385 390 395 400

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

405 410 415

Phe Gln Val His Leu Gly Asp Tyr Val His Asp Arg Arg Thr Lys Gln

420 425 430

Leu Gly Lys Val Glu Ser Asp Arg Ile Ile Lys Glu Lys Val Thr Val

435 440 445

Phe Ala Arg Leu Lys Asp Ile Asn Ser Ala Lys Ala Ser Tyr Phe His

450 455 460

Ser Leu Glu Glu Gln Asp Lys Glu Glu Leu Asp Asn Lys Trp Thr Leu

465 470 475 480

Phe Pro Asn Pro Ser Tyr Asp Phe Pro Lys Glu His Thr Leu Gln His

485 490 495

Gln Gly Glu Gln Lys Asn Ala Gly Lys Ile Gly Ile Tyr Val Lys Leu

500 505 510

Arg Asp Thr Gln Tyr Lys Glu Lys Ala Ala Leu Glu Glu Ala Arg Lys

515 520 525

Ser Leu Asn Pro Lys Glu Arg Ser Ala Thr Lys Ala Ser Lys Tyr Asp

530 535 540

Ile Ile Thr Gln Ile Ile Glu Ala Asn Asp Asn Val Lys Ser Glu Lys

545 550 555 560

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

565 570 575

Ile His Ser Met Leu Phe Ser Leu Leu Thr Asp Asn Ala Glu Leu Lys

580 585 590

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

595 600 605

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

610 615 620

Lys Tyr Lys Asp Asn Asp Leu Lys Glu Thr Asp Thr Asp Lys Ile Thr

625 630 635 640

Arg Asp Leu Ala Arg Asp Lys Glu Glu Ile Glu Lys Leu Ile Leu Glu

645 650 655

Gln Lys Gln Arg Ala Asp Asp Tyr Asn Tyr Thr Ser Ser Thr Lys Phe

660 665 670

Asn Ile Asp Lys Ser Arg Lys Arg Lys His Leu Leu Phe Asn Ala Glu

675 680 685

Lys Gly Lys Ile Gly Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met

690 695 700

Phe Lys Glu Ser Lys Ser Lys Trp Lys Gly Tyr Gln His Thr Glu Leu

705 710 715 720

Gln Lys Leu Phe Ala Tyr Phe Asp Thr Ser Lys Ser Asp Leu Glu Leu

725 730 735

Ile Leu Ser Asp Met Val Met Val Lys Asp Tyr Pro Ile Glu Leu Ile

740 745 750

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

755 760 765

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

770 775 780

Asn Ser Ile Gly Thr Pro Gln Phe Lys Thr Val Arg Lys Glu Cys Phe

785 790 795 800

Ala Phe Leu Lys Glu Ser Asn Tyr Thr Val Ala Ser Leu Asp Lys Gln

805 810 815

Ile Glu Arg Ile Leu Ser Met Pro Leu Phe Ile Glu Arg Gly Phe Met

820 825 830

Asp Ser Lys Pro Thr Met Leu Glu Gly Lys Ser Tyr Gln Gln His Lys

835 840 845

Glu Asp Phe Ala Asp Trp Phe Val His Tyr Lys Glu Asn Ser Asn Tyr

850 855 860

Gln Asn Phe Tyr Asp Thr Glu Val Tyr Glu Ile Ile Thr Glu Asp Lys

865 870 875 880

Arg Glu Gln Ala Lys Val Thr Lys Lys Ile Lys Gln Gln Gln Lys Asn

885 890 895

Asp Val Phe Thr Leu Met Met Val Asn Tyr Met Leu Glu Glu Val Leu

900 905 910

Lys Leu Pro Ser Asn Asp Arg Leu Ser Leu Asn Glu Leu Tyr Gln Thr

915 920 925

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

930 935 940

Arg Asn Lys Asn Tyr Ile Trp Asn Lys Val Val Asp Leu Gln Leu Cys

945 950 955 960

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

965 970 975

Phe Arg Lys Tyr Glu Asn Asp Ser Arg Val Lys Glu Phe Leu Thr Tyr

980 985 990

Gln Ser Asp Ile Val Trp Ser Gly Tyr Leu Ser Asn Glu Val Asp Ser

995 1000 1005

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

1010 1015 1020

Ile Arg Ser Lys Glu Leu Leu Lys Glu Val Gln Glu Ile Glu Cys

1025 1030 1035

Ile Val Tyr Asn Gln Val Ala Asn Lys Glu Ser Leu Lys Gln Ser

1040 1045 1050

Gly Asn Glu Asn Phe Lys Gln Tyr Val Leu Gln Gly Leu Leu Pro

1055 1060 1065

Arg Gly Thr Asp Val Arg Glu Met Leu Ile Leu Ser Thr Asp Val

1070 1075 1080

Lys Phe Lys Lys Glu Glu Ile Met Gln Leu Gly Gln Val Arg Glu

1085 1090 1095

Val Glu Gln Asp Leu Tyr Ser Leu Ile Tyr Ile Arg Asn Lys Phe

1100 1105 1110

Ala His Asn Gln Leu Pro Ile Lys Glu Phe Phe Asp Phe Cys Glu

1115 1120 1125

Asn Asn Tyr Arg Pro Ile Ser Asp Asn Glu Tyr Tyr Ala Glu Tyr

1130 1135 1140

Tyr Met Glu Ile Phe Arg Ser Ile Lys Glu Lys Tyr Ala Ser

1145 1150 1155

<210> 113

<211> 1127

<212> PRT

<213> Prevotella species

<400> 113

Met Gln Lys Gln Asp Lys Leu Phe Val Asp Arg Lys Lys Asn Ala Ile
1 5 10 15

Phe Ala Phe Pro Lys Tyr Ile Thr Ile Met Glu Asn Gln Glu Lys Pro
20 25 30

Glu Pro Ile Tyr Tyr Glu Leu Thr Asp Lys His Phe Trp Ala Ala Phe
35 40 45

Leu Asn Leu Ala Arg His Asn Val Tyr Thr Thr Ile Asn His Ile Asn
50 55 60

Arg Arg Leu Glu Ile Ala Glu Leu Lys Asp Asp Gly Tyr Met Met Gly
65 70 75 80

Ile Lys Gly Ser Trp Asn Glu Gln Ala Lys Lys Leu Asp Lys Lys Val
85 90 95

Arg Leu Arg Asp Leu Ile Met Lys His Phe Pro Phe Leu Glu Ala Ala
100 105 110

Ala Tyr Glu Ile Thr Asn Ser Lys Ser Pro Asn Asn Lys Glu Gln Arg
115 120 125

Glu Lys Glu Gln Ser Glu Ala Leu Ser Leu Asn Asn Leu Lys Asn Val
130 135 140

Leu Phe Ile Phe Leu Glu Lys Leu Gln Val Leu Arg Asn Tyr Tyr Ser
145 150 155 160

His Tyr Lys Tyr Ser Glu Glu Ser Pro Lys Pro Ile Phe Glu Thr Ser
165 170 175

Leu Leu Lys Asn Met Tyr Lys Val Phe Asp Ala Asn Val Arg Leu Val
180 185 190

Lys Arg Asp Tyr Met His His Glu Asn Ile Asp Met Gln Arg Asp Phe
195 200 205

Thr His Leu Asn Arg Lys Lys Gln Val Gly Arg Thr Lys Asn Ile Ile

210 215 220

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

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

Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Val Leu Arg Tyr Phe
355 360 365
Asp Leu Asn Glu Ile Phe Glu Gln Leu Arg Phe Gln Ile Asp Leu Gly
370 375 380
Thr Tyr His Phe Ser Ile Tyr Asn Lys Arg Ile Gly Asp Glu Asp Glu
385 390 395 400
Val Arg His Leu Thr His His Leu Tyr Gly Phe Ala Arg Ile Gln Asp
405 410 415

Phe Ala Pro Gln Asn Gln Pro Glu Glu Trp Arg Lys Leu Val Lys Asp
420 425 430
Leu Asp His Phe Glu Thr Ser Gln Glu Pro Tyr Ile Ser Lys Thr Ala
435 440 445
Pro His Tyr His Leu Glu Asn Glu Lys Ile Gly Ile Lys Phe Cys Ser
450 455 460

Thr His Asn Asn Leu Phe Pro Ser Leu Lys Arg Glu Lys Thr Cys Asn
465 470 475 480

Gly Arg Ser Lys Phe Asn Leu Gly Thr Gln Phe Thr Ala Glu Ala Phe
485 490 495

Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu Leu Leu
500 505 510

Thr Lys Asp Tyr Ser Arg Lys Glu Ser Ala Asp Lys Val Glu Gly Ile
515 520 525

Ile Arg Lys Glu Ile Ser Asn Ile Tyr Ala Ile Tyr Asp Ala Phe Ala
530 535 540

Asn Asn Glu Ile Asn Ser Ile Ala Asp Leu Thr Cys Arg Leu Gln Lys
545 550 555 560

Thr Asn Ile Leu Gln Gly His Leu Pro Lys Gln Met Ile Ser Ile Leu
565 570 575

Glu Gly Arg Gln Lys Asp Met Glu Lys Glu Ala Glu Arg Lys Ile Gly
580 585 590

Glu Met Ile Asp Asp Thr Gln Arg Arg Leu Asp Leu Leu Cys Lys Gln
595 600 605

Thr Asn Gln Lys Ile Arg Ile Gly Lys Arg Asn Ala Gly Leu Leu Lys
610 615 620

Ser Gly Lys Ile Ala Asp Trp Leu Val Ser Asp Met Met Arg Phe Gln
625 630 635 640

Pro Val Gln Lys Asp Thr Asn Asn Ala Pro Ile Asn Asn Ser Lys Ala
645 650 655

Asn Ser Thr Glu Tyr Arg Met Leu Gln His Ala Leu Ala Leu Phe Gly
660 665 670

Ser Glu Ser Ser Arg Leu Lys Ala Tyr Phe Arg Gln Met Asn Leu Val
675 680 685

Gly Asn Ala Asn Pro His Pro Phe Leu Ala Glu Thr Gln Trp Glu His
690 695 700

Gln Thr Asn Ile Leu Ser Phe Tyr Arg Asn Tyr Leu Glu Ala Arg Lys

705 710 715 720

Lys Tyr Leu Lys Gly Leu Lys Pro Gln Asn Trp Lys Gln Tyr Gln His

725 730 735

Phe Leu Ile Leu Lys Val Gln Lys Thr Asn Arg Asn Thr Leu Val Thr

740 745 750

Gly Trp Lys Asn Ser Phe Asn Leu Pro Arg Gly Ile Phe Thr Gln Pro

755 760 765

Ile Arg Glu Trp Phe Glu Lys His Asn Asn Ser Lys Arg Ile Tyr Asp

770 775 780

Gln Ile Leu Ser Phe Asp Arg Val Gly Phe Val Ala Lys Ala Ile Pro

785 790 795 800

Leu Tyr Phe Ala Glu Glu Tyr Lys Asp Asn Val Gln Pro Phe Tyr Asp

805 810 815

Tyr Pro Phe Asn Ile Gly Asn Lys Leu Lys Pro Gln Lys Gly Gln Phe

820 825 830

Leu Asp Lys Lys Glu Arg Val Glu Leu Trp Gln Lys Asn Lys Glu Leu

835 840 845

Phe Lys Asn Tyr Pro Ser Glu Lys Asn Lys Thr Asp Leu Ala Tyr Leu

850 855 860

Asp Phe Leu Ser Trp Lys Lys Phe Glu Arg Glu Leu Arg Leu Ile Lys

865 870 875 880

Asn Gln Asp Ile Val Thr Trp Leu Met Phe Lys Glu Leu Phe Lys Thr

885 890 895

Thr Thr Val Glu Gly Leu Lys Ile Gly Glu Ile His Leu Arg Asp Ile

900 905 910

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

915 920 925

Met Pro Met Lys Leu Pro Val Lys Thr Tyr Glu Thr Asp Asn Lys Gly

930 935 940

Asn Ile Leu Lys Glu Arg Pro Leu Ala Thr Phe Tyr Ile Glu Glu Thr

945 950 955 960

Glu Thr Lys Val Leu Lys Gln Gly Asn Phe Lys Val Leu Ala Lys Asp
965 970 975
Arg Arg Leu Asn Gly Leu Leu Ser Phe Ala Glu Thr Thr Asp Ile Asp
980 985 990

Leu Glu Lys Asn Pro Ile Thr Lys Leu Ser Val Asp Tyr Glu Leu Ile
995 1000 1005
Lys Tyr Gln Thr Thr Arg Ile Ser Ile Phe Glu Met Thr Leu Gly
1010 1015 1020
Leu Glu Lys Lys Leu Ile Asp Lys Tyr Ser Thr Leu Pro Thr Asp
1025 1030 1035
Ser Phe Arg Asn Met Leu Glu Arg Trp Leu Gln Cys Lys Ala Asn
1040 1045 1050

Arg Pro Glu Leu Lys Asn Tyr Val Asn Ser Leu Ile Ala Val Arg
1055 1060 1065
Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp Ala Thr Leu
1070 1075 1080
Phe Ala Glu Val Lys Lys Phe Thr Leu Phe Pro Ser Val Asp Thr
1085 1090 1095
Lys Lys Ile Glu Leu Asn Ile Ala Pro Gln Leu Leu Glu Ile Val
1100 1105 1110

Gly Lys Ala Ile Lys Glu Ile Glu Lys Ser Glu Asn Lys Asn
1115 1120 1125
<210> 114
<211> 1119
<212> PRT
<213> Porphyromonas gingivalis
<400> 114
Met Thr Glu Gln Asn Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15
Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn
20 25 30

Ala Tyr Ile Thr Leu Ala His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu

145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Tyr Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp

210 215 220

Arg Glu Glu Lys Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr

325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Asp Glu Ala

500 505 510

Ser Ala Glu Arg Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

Tyr Ala Val Tyr Asp Ala Phe Ala Arg Gly Glu Ile Asn Thr Arg Asp
530 535 540
Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560
Pro Arg Gln Met Ile Gly Ile Leu Ser Gln Glu His Lys Asp Met Glu
565 570 575
Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Val Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg
705 710 715 720
Ser Asp Arg Val Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu

740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765
Gly Leu Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830

Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala

835 840 845

Lys Glu His Pro Tyr Leu Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg

850 855 860

Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Ile Cys

865 870 875 880

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

885 890 895

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

900 905 910

Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr

915 920 925

Arg Ala Asp Ser Arg Gly His Val His Lys Glu Gln Ala Pro Leu Ala

930 935 940

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

945 950 955 960

Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe

965 970 975

Val Asp Thr Gly Ala Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu

980 985 990

Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala

995 1000 1005

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

1010 1015 1020

Pro His Leu Pro Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp
1025 1030 1035
Ser Asp Pro Leu Leu Asp Lys Trp Pro Asp Leu His Gly Asn Val
1040 1045 1050
Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr

1055 1060 1065
Pro Met Tyr Asp Glu Ala Val Phe Ser Ser Ile Arg Lys Tyr Asp
1070 1075 1080
Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile
1085 1090 1095
Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Met Ala
1100 1105 1110
Glu Arg Ile Ile Gln Ala
1115

<210> 115
<211> 1154
<212> PRT
<213> Paludibacter propionicigenes
<400> 115
Met Lys Thr Ser Ala Asn Asn Ile Tyr Phe Asn Gly Ile Asn Ser Phe
1 5 10 15
Lys Lys Ile Phe Asp Ser Lys Gly Ala Ile Ala Pro Ile Ala Glu Lys
20 25 30
Ser Cys Arg Asn Phe Asp Ile Lys Ala Gln Asn Asp Val Asn Lys Glu
35 40 45
Gln Arg Ile His Tyr Phe Ala Val Gly His Thr Phe Lys Gln Leu Asp

50 55 60
Thr Glu Asn Leu Phe Glu Tyr Val Leu Asp Glu Asn Leu Arg Ala Lys
65 70 75 80
Arg Pro Thr Arg Phe Ile Ser Leu Gln Gln Phe Asp Lys Glu Phe Ile
85 90 95
Glu Asn Ile Lys Arg Leu Ile Ser Asp Ile Arg Asn Ile Asn Ser His

100 105 110

Tyr Ile His Arg Phe Asp Pro Leu Lys Ile Asp Ala Val Pro Thr Asn

115 120 125

Ile Ile Asp Phe Leu Lys Glu Ser Phe Glu Leu Ala Val Ile Gln Ile

130 135 140

Tyr Leu Lys Glu Lys Gly Ile Asn Tyr Leu Gln Phe Ser Glu Asn Pro

145 150 155 160

His Ala Asp Gln Lys Leu Val Ala Phe Leu His Asp Lys Phe Leu Pro

165 170 175

Leu Asp Glu Lys Lys Thr Ser Met Leu Gln Asn Glu Thr Pro Gln Leu

180 185 190

Lys Glu Tyr Lys Glu Tyr Arg Lys Tyr Phe Lys Thr Leu Ser Lys Gln

195 200 205

Ala Ala Ile Asp Gln Leu Leu Phe Ala Glu Lys Glu Thr Asp Tyr Ile

210 215 220

Trp Asn Leu Phe Asp Ser His Pro Val Leu Thr Ile Ser Ala Gly Lys

225 230 235 240

Tyr Leu Ser Phe Tyr Ser Cys Leu Phe Leu Leu Ser Met Phe Leu Tyr

245 250 255

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

260 265 270

Asn Thr Thr Glu Glu Glu Lys Ser Lys Arg Glu Ile Phe Thr Phe Phe

275 280 285

Ser Lys Arg Phe Asn Ser Met Asp Ile Asp Ser Glu Glu Asn Gln Leu

290 295 300

Val Lys Phe Arg Asp Leu Ile Leu Tyr Leu Asn His Tyr Pro Val Ala

305 310 315 320

Trp Asn Lys Asp Leu Glu Leu Asp Ser Ser Asn Pro Ala Met Thr Asp

325 330 335

Lys Leu Lys Ser Lys Ile Ile Glu Leu Glu Ile Asn Arg Ser Phe Pro

340 345 350

Leu Tyr Glu Gly Asn Glu Arg Phe Ala Thr Phe Ala Lys Tyr Gln Ile
355 360 365
Trp Gly Lys Lys His Leu Gly Lys Ser Ile Glu Lys Glu Tyr Ile Asn

370 375 380
Ala Ser Phe Thr Asp Glu Glu Ile Thr Ala Tyr Thr Tyr Glu Thr Asp
385 390 395 400
Thr Cys Pro Glu Leu Lys Asp Ala His Lys Lys Leu Ala Asp Leu Lys
405 410 415
Ala Ala Lys Gly Leu Phe Gly Lys Arg Lys Glu Lys Asn Glu Ser Asp
420 425 430
Ile Lys Lys Thr Glu Thr Ser Ile Arg Glu Leu Gln His Glu Pro Asn

435 440 445
Pro Ile Lys Asp Lys Leu Ile Gln Arg Ile Glu Lys Asn Leu Leu Thr
450 455 460
Val Ser Tyr Gly Arg Asn Gln Asp Arg Phe Met Asp Phe Ser Ala Arg
465 470 475 480
Phe Leu Ala Glu Ile Asn Tyr Phe Gly Gln Asp Ala Ser Phe Lys Met
485 490 495
Tyr His Phe Tyr Ala Thr Asp Glu Gln Asn Ser Glu Leu Glu Lys Tyr

500 505 510
Glu Leu Pro Lys Asp Lys Lys Lys Tyr Asp Ser Leu Lys Phe His Gln
515 520 525
Gly Lys Leu Val His Phe Ile Ser Tyr Lys Glu His Leu Lys Arg Tyr
530 535 540
Glu Ser Trp Asp Asp Ala Phe Val Ile Glu Asn Asn Ala Ile Gln Leu
545 550 555 560
Lys Leu Ser Phe Asp Gly Val Glu Asn Thr Val Thr Ile Gln Arg Ala

565 570 575
Leu Leu Ile Tyr Leu Leu Glu Asp Ala Leu Arg Asn Ile Gln Asn Asn
580 585 590
Thr Ala Glu Asn Ala Gly Lys Gln Leu Leu Gln Glu Tyr Tyr Ser His

595 600 605

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

610 615 620

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

625 630 635 640

Arg Leu Leu Asn Asn Tyr Ser Pro Ala Ile Asn His Leu Gln Thr Pro

645 650 655

His Ser Ser Leu Pro Leu Ile Leu Glu Lys Ala Leu Leu Ala Glu Lys

660 665 670

Arg Tyr Cys Ser Leu Val Val Lys Ala Lys Ala Glu Gly Asn Tyr Asp

675 680 685

Asp Phe Ile Lys Arg Asn Lys Gly Lys Gln Phe Lys Leu Gln Phe Ile

690 695 700

Arg Lys Ala Trp Asn Leu Met Tyr Phe Arg Asn Ser Tyr Leu Gln Asn

705 710 715 720

Val Gln Ala Ala Gly His His Lys Ser Phe His Ile Glu Arg Asp Glu

725 730 735

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

740 745 750

Tyr Lys Tyr Tyr Leu Asn Glu Met Phe Glu Lys Lys Gly Phe Phe Glu

755 760 765

Asn Asn Glu Phe Lys Ile Leu Phe Gln Ser Gly Thr Ser Leu Glu Asn

770 775 780

Leu Tyr Glu Lys Thr Lys Gln Lys Phe Glu Ile Trp Leu Ala Ser Asn

785 790 795 800

Thr Ala Lys Thr Asn Lys Pro Asp Asn Tyr His Leu Asn Asn Tyr Glu

805 810 815

Gln Gln Phe Ser Asn Gln Leu Phe Phe Ile Asn Leu Ser His Phe Ile

820 825 830

Asn Tyr Leu Lys Ser Thr Gly Lys Leu Gln Thr Asp Ala Asn Gly Gln

835 840 845

Ile Ile Tyr Glu Ala Leu Asn Asn Val Gln Tyr Leu Ile Pro Glu Tyr
850 855 860
Tyr Tyr Thr Asp Lys Pro Glu Arg Ser Glu Ser Lys Ser Gly Asn Lys
865 870 875 880
Leu Tyr Asn Lys Leu Lys Ala Thr Lys Leu Glu Asp Ala Leu Leu Tyr

885 890 895
Glu Met Ala Met Cys Tyr Leu Lys Ala Asp Lys Gln Ile Ala Asp Lys
900 905 910
Ala Lys His Pro Ile Thr Lys Leu Leu Thr Ser Asp Val Glu Phe Asn
915 920 925
Ile Thr Asn Lys Glu Gly Ile Gln Leu Tyr His Leu Leu Val Pro Phe
930 935 940
Lys Lys Ile Asp Ala Phe Ile Gly Leu Lys Met His Lys Glu Gln Gln

945 950 955 960
Asp Lys Lys His Pro Thr Ser Phe Leu Ala Asn Ile Val Asn Tyr Leu
965 970 975
Glu Leu Val Lys Asn Asp Lys Asp Ile Arg Lys Thr Tyr Glu Ala Phe
980 985 990
Ser Thr Asn Pro Val Lys Arg Thr Leu Thr Tyr Asp Asp Leu Ala Lys
995 1000 1005
Ile Asp Gly His Leu Ile Ser Lys Ser Ile Lys Phe Thr Asn Val

1010 1015 1020
Thr Leu Glu Leu Glu Arg Tyr Phe Ile Phe Lys Glu Ser Leu Ile
1025 1030 1035
Val Lys Lys Gly Asn Asn Ile Asp Phe Lys Tyr Ile Lys Gly Leu
1040 1045 1050
Arg Asn Tyr Tyr Asn Asn Glu Lys Lys Lys Asn Glu Gly Ile Arg
1055 1060 1065
Asn Lys Ala Phe His Phe Gly Ile Pro Asp Ser Lys Ser Tyr Asp

1070 1075 1080
Gln Leu Ile Arg Asp Ala Glu Val Met Phe Ile Ala Asn Glu Val

1085 1090 1095

Lys Pro Thr His Ala Thr Lys Tyr Thr Asp Leu Asn Lys Gln Leu

1100 1105 1110

His Thr Val Cys Asp Lys Leu Met Glu Thr Val His Asn Asp Tyr

1115 1120 1125

Phe Ser Lys Glu Gly Asp Gly Lys Lys Lys Arg Glu Ala Ala Gly

1130 1135 1140

Gln Lys Tyr Phe Glu Asn Ile Ile Ser Ala Lys

1145 1150

<210> 116

<211> 1119

<212> PRT

<213> Porphyromonas gingivalis

<400> 116

Met Thr Glu Gln Asn Glu Lys Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Ala His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu
145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205

Arg Tyr Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp
210 215 220

Arg Glu Gly Thr Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly
245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe
260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe
305 310 315 320

Arg Val Pro Val Asp Ile Leu Ser Asp Glu Glu Asp Thr Asp Gly Ala
325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe
340 345 350

Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser
355 360 365

Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys
370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Phe Thr Ala Glu Ala Phe Leu Ser Ala His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Ala

500 505 510

Ser Ala Glu Arg Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

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

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

Pro Arg Gln Met Ile Gly Ile Leu Ser Gln Glu His Lys Asp Met Glu

565 570 575

Glu Lys Ile Arg Lys Lys Leu Gln Glu Met Met Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg
705 710 715 720
Ser Asp Arg Val Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu
740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765
Gly Leu Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys
770 775 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Trp Val Gln Pro
785 790 795 800

Phe Tyr Asn Tyr Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815
Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly
820 825 830
Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala
835 840 845
Lys Glu His Pro Tyr Leu Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg
850 855 860

Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Ile Cys
865 870 875 880
Gly Asp Leu Met Glu Glu Asn Lys Val Glu Gly Leu Asp Thr Gly Thr

885 890 895

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

900 905 910

Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr

915 920 925

Arg Ala Asp Ser Arg Gly His Val His Lys Glu Gln Ala Pro Leu Ala

930 935 940

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

945 950 955 960

Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe

965 970 975

Val Asp Thr Gly Ala Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu

980 985 990

Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala

995 1000 1005

Phe Glu Gln Thr Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Cys

1010 1015 1020

Pro His Leu Pro Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp

1025 1030 1035

Ser Asp Pro Leu Leu Asp Lys Trp Pro Asp Leu His Arg Lys Val

1040 1045 1050

Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr

1055 1060 1065

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

1070 1075 1080

Pro Ser Phe Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile

1085 1090 1095

Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val

1100 1105 1110

Glu Arg Ile Ile Gln Ala

1115

<210> 117

<211> 1179

<212> PRT

<213> Flavobacterium columnare

<400> 117

Met Ser Ser Lys Asn Glu Ser Tyr Asn Lys Gln Lys Thr Phe Asn His
1 5 10 15

Tyr Lys Gln Glu Asp Lys Tyr Phe Phe Gly Gly Phe Leu Asn Asn Ala
20 25 30

Asp Asp Asn Leu Arg Gln Val Gly Lys Glu Phe Lys Thr Arg Ile Asn
35 40 45

Phe Asn His Asn Asn Asn Glu Leu Ala Ser Val Phe Lys Asp Tyr Phe
50 55 60

Asn Lys Glu Lys Ser Val Ala Lys Arg Glu His Ala Leu Asn Leu Leu
65 70 75 80

Ser Asn Tyr Phe Pro Val Leu Glu Arg Ile Gln Lys His Thr Asn His
85 90 95

Asn Phe Glu Gln Thr Arg Glu Ile Phe Glu Leu Leu Leu Asp Thr Ile
100 105 110

Lys Lys Leu Arg Asp Tyr Tyr Thr His His Tyr His Lys Pro Ile Thr
115 120 125

Ile Asn Pro Lys Ile Tyr Asp Phe Leu Asp Asp Thr Leu Leu Asp Val
130 135 140

Leu Ile Thr Ile Lys Lys Lys Lys Val Lys Asn Asp Thr Ser Arg Glu
145 150 155 160

Leu Leu Lys Glu Lys Leu Arg Pro Glu Leu Thr Gln Leu Lys Asn Gln
165 170 175

Lys Arg Glu Glu Leu Ile Lys Lys Gly Lys Lys Leu Leu Glu Glu Asn
180 185 190

Leu Glu Asn Ala Val Phe Asn His Cys Leu Arg Pro Phe Leu Glu Glu
195 200 205

Asn Lys Thr Asp Asp Lys Gln Asn Lys Thr Val Ser Leu Arg Lys Tyr

210 215 220

Arg Lys Ser Lys Pro Asn Glu Glu Thr Ser Ile Thr Leu Thr Gln Ser
225 230 235 240

Gly Leu Val Phe Leu Met Ser Phe Phe Leu His Arg Lys Glu Phe Gln
245 250 255

Val Phe Thr Ser Gly Leu Glu Gly Phe Lys Ala Lys Val Asn Thr Ile
260 265 270

Lys Glu Glu Glu Ile Ser Leu Asn Lys Asn Asn Ile Val Tyr Met Ile
275 280 285

Thr His Trp Ser Tyr Ser Tyr Tyr Asn Phe Lys Gly Leu Lys His Arg
290 295 300

Ile Lys Thr Asp Gln Gly Val Ser Thr Leu Glu Gln Asn Asn Thr Thr
305 310 315 320

His Ser Leu Thr Asn Thr Asn Thr Lys Glu Ala Leu Leu Thr Gln Ile
325 330 335

Val Asp Tyr Leu Ser Lys Val Pro Asn Glu Ile Tyr Glu Thr Leu Ser
340 345 350

Glu Lys Gln Gln Lys Glu Phe Glu Glu Asp Ile Asn Glu Tyr Met Arg
355 360 365

Glu Asn Pro Glu Asn Glu Asp Ser Thr Phe Ser Ser Ile Val Ser His
370 375 380

Lys Val Ile Arg Lys Arg Tyr Glu Asn Lys Phe Asn Tyr Phe Ala Met
385 390 395 400

Arg Phe Leu Asp Glu Tyr Ala Glu Leu Pro Thr Leu Arg Phe Met Val
405 410 415

Asn Phe Gly Asp Tyr Ile Lys Asp Arg Gln Lys Lys Ile Leu Glu Ser
420 425 430

Ile Gln Phe Asp Ser Glu Arg Ile Ile Lys Lys Glu Ile His Leu Phe
435 440 445

Glu Lys Leu Ser Leu Val Thr Glu Tyr Lys Lys Asn Val Tyr Leu Lys
450 455 460

Glu Thr Ser Asn Ile Asp Leu Ser Arg Phe Pro Leu Phe Pro Asn Pro
465 470 475 480
Ser Tyr Val Met Ala Asn Asn Asn Ile Pro Phe Tyr Ile Asp Ser Arg
485 490 495

Ser Asn Asn Leu Asp Glu Tyr Leu Asn Gln Lys Lys Lys Ala Gln Ser
500 505 510
Gln Asn Lys Lys Arg Asn Leu Thr Phe Glu Lys Tyr Asn Lys Glu Gln
515 520 525
Ser Lys Asp Ala Ile Ile Ala Met Leu Gln Lys Glu Ile Gly Val Lys
530 535 540
Asp Leu Gln Gln Arg Ser Thr Ile Gly Leu Leu Ser Cys Asn Glu Leu
545 550 555 560

Pro Ser Met Leu Tyr Glu Val Ile Val Lys Asp Ile Lys Gly Ala Glu
565 570 575
Leu Glu Asn Lys Ile Ala Gln Lys Ile Arg Glu Gln Tyr Gln Ser Ile
580 585 590
Arg Asp Phe Thr Leu Asp Ser Pro Gln Lys Asp Asn Ile Pro Thr Thr
595 600 605
Leu Ile Lys Thr Ile Asn Thr Asp Ser Ser Val Thr Phe Glu Asn Gln
610 615 620

Pro Ile Asp Ile Pro Arg Leu Lys Asn Ala Ile Gln Lys Glu Leu Thr
625 630 635 640
Leu Thr Gln Glu Lys Leu Leu Asn Val Lys Glu His Glu Ile Glu Val
645 650 655
Asp Asn Tyr Asn Arg Asn Lys Asn Thr Tyr Lys Phe Lys Asn Gln Pro
660 665 670
Lys Asn Lys Val Asp Asp Lys Lys Leu Gln Arg Lys Tyr Val Phe Tyr
675 680 685

Arg Asn Glu Ile Arg Gln Glu Ala Asn Trp Leu Ala Ser Asp Leu Ile
690 695 700
His Phe Met Lys Asn Lys Ser Leu Trp Lys Gly Tyr Met His Asn Glu

705 710 715 720

Leu Gln Ser Phe Leu Ala Phe Phe Glu Asp Lys Lys Asn Asp Cys Ile

725 730 735

Ala Leu Leu Glu Thr Val Phe Asn Leu Lys Glu Asp Cys Ile Leu Thr

740 745 750

Lys Gly Leu Lys Asn Leu Phe Leu Lys His Gly Asn Phe Ile Asp Phe

755 760 765

Tyr Lys Glu Tyr Leu Lys Leu Lys Glu Asp Phe Leu Asn Thr Glu Ser

770 775 780

Thr Phe Leu Glu Asn Gly Leu Ile Gly Leu Pro Pro Lys Ile Leu Lys

785 790 795 800

Lys Glu Leu Ser Lys Arg Phe Lys Tyr Ile Phe Ile Val Phe Gln Lys

805 810 815

Arg Gln Phe Ile Ile Lys Glu Leu Glu Glu Lys Lys Asn Asn Leu Tyr

820 825 830

Ala Asp Ala Ile Asn Leu Ser Arg Gly Ile Phe Asp Glu Lys Pro Thr

835 840 845

Met Ile Pro Phe Lys Lys Pro Asn Pro Asp Glu Phe Ala Ser Trp Phe

850 855 860

Val Ala Ser Tyr Gln Tyr Asn Asn Tyr Gln Ser Phe Tyr Glu Leu Thr

865 870 875 880

Pro Asp Ile Val Glu Arg Asp Lys Lys Lys Lys Tyr Lys Asn Leu Arg

885 890 895

Ala Ile Asn Lys Val Lys Ile Gln Asp Tyr Tyr Leu Lys Leu Met Val

900 905 910

Asp Thr Leu Tyr Gln Asp Leu Phe Asn Gln Pro Leu Asp Lys Ser Leu

915 920 925

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

930 935 940

Ala Lys Ala Tyr Gln Lys Arg Asn Asp Ser Ser Leu Trp Asn Lys Val

945 950 955 960

Ile His Leu Ser Leu Gln Asn Asn Arg Ile Thr Ala Asn Pro Lys Leu
965 970 975
Lys Asp Ile Gly Lys Tyr Lys Arg Ala Leu Gln Asp Glu Lys Ile Ala
980 985 990
Thr Leu Leu Thr Tyr Asp Asp Arg Thr Trp Thr Tyr Ala Leu Gln Lys
995 1000 1005

Pro Glu Lys Glu Asn Glu Asn Asp Tyr Lys Glu Leu His Tyr Thr
1010 1015 1020
Ala Leu Asn Met Glu Leu Gln Glu Tyr Glu Lys Val Arg Ser Lys
1025 1030 1035
Glu Leu Leu Lys Gln Val Gln Glu Leu Glu Lys Gln Ile Leu Glu
1040 1045 1050
Glu Tyr Thr Asp Phe Leu Ser Thr Gln Ile His Pro Ala Asp Phe
1055 1060 1065

Glu Arg Glu Gly Asn Pro Asn Phe Lys Lys Tyr Leu Ala His Ser
1070 1075 1080
Ile Leu Glu Asn Glu Asp Asp Leu Asp Lys Leu Pro Glu Lys Val
1085 1090 1095
Glu Ala Met Arg Glu Leu Asp Glu Thr Ile Thr Asn Pro Ile Ile
1100 1105 1110
Lys Lys Ala Ile Val Leu Ile Ile Ile Arg Asn Lys Met Ala His
1115 1120 1125

Asn Gln Tyr Pro Pro Lys Phe Ile Tyr Asp Leu Ala Asn Arg Phe
1130 1135 1140
Val Pro Lys Lys Glu Glu Glu Tyr Phe Ala Thr Tyr Phe Asn Arg
1145 1150 1155
Val Phe Glu Thr Ile Thr Lys Glu Leu Trp Glu Asn Lys Glu Lys
1160 1165 1170
Lys Asp Lys Thr Gln Val
1175
<210> 118
<211> 1145

<212> PRT

<213> Psychroflexus torquis

<400> 118

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

Met Ser Ile Val Asp Tyr Glu Lys Lys Ile Ser Ile Leu Asn Gly Tyr
65 70 75 80
Leu Pro Ile Ile Asp Phe Leu Asp Asp Glu Leu Glu Asn Asn Leu Asn
85 90 95
Thr Arg Val Lys Asn Phe Lys Lys Asn Phe Ile Ile Leu Ala Glu Ala
100 105 110
Ile Glu Lys Leu Arg Asp Tyr Tyr Thr His Phe Tyr His Asp Pro Ile
115 120 125

Thr Phe Glu Asp Asn Lys Glu Pro Leu Leu Glu Leu Leu Asp Glu Val
130 135 140
Leu Leu Lys Thr Ile Leu Asp Val Lys Lys Lys Tyr Leu Lys Thr Asp
145 150 155 160
Lys Thr Lys Glu Ile Leu Lys Asp Ser Leu Arg Glu Glu Met Asp Leu
165 170 175
Leu Val Ile Arg Lys Thr Asp Glu Leu Arg Glu Lys Lys Lys Thr Asn
180 185 190

Pro Lys Ile Gln His Thr Asp Ser Ser Gln Ile Lys Asn Ser Ile Phe
195 200 205
Asn Asp Ala Phe Gln Gly Leu Leu Tyr Glu Asp Lys Gly Asn Asn Lys
210 215 220

Lys Thr Gln Val Ser His Arg Ala Lys Thr Arg Leu Asn Pro Lys Asp
225 230 235 240
Ile His Lys Gln Glu Glu Arg Asp Phe Glu Ile Pro Leu Ser Thr Ser
245 250 255

Gly Leu Val Phe Leu Met Ser Leu Phe Leu Ser Lys Lys Glu Ile Glu
260 265 270
Asp Phe Lys Ser Asn Ile Lys Gly Phe Lys Gly Lys Val Val Lys Asp
275 280 285
Glu Asn His Asn Ser Leu Lys Tyr Met Ala Thr His Arg Val Tyr Ser
290 295 300
Ile Leu Ala Phe Lys Gly Leu Lys Tyr Arg Ile Lys Thr Asp Thr Phe
305 310 315 320

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

Glu Asp Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Ala
385 390 395 400
Asn Phe Lys Thr Leu Lys Phe Gln Val Phe Met Gly Tyr Tyr Ile His
405 410 415
Asp Gln Arg Thr Lys Thr Ile Gly Thr Thr Asn Ile Thr Thr Glu Arg
420 425 430
Thr Val Lys Glu Lys Ile Asn Val Phe Gly Lys Leu Ser Lys Met Asp
435 440 445

Asn Leu Lys Lys His Phe Phe Ser Gln Leu Ser Asp Asp Glu Asn Thr
450 455 460
Asp Trp Glu Phe Phe Pro Asn Pro Ser Tyr Asn Phe Leu Thr Gln Ala

465 470 475 480

Asp Asn Ser Pro Ala Asn Asn Ile Pro Ile Tyr Leu Glu Leu Lys Asn

485 490 495

Gln Gln Ile Ile Lys Glu Lys Asp Ala Ile Lys Ala Glu Val Asn Gln

500 505 510

Thr Gln Asn Arg Asn Pro Asn Lys Pro Ser Lys Arg Asp Leu Leu Asn

515 520 525

Lys Ile Leu Lys Thr Tyr Glu Asp Phe His Gln Gly Asp Pro Thr Ala

530 535 540

Ile Leu Ser Leu Asn Glu Ile Pro Ala Leu Leu His Leu Phe Leu Val

545 550 555 560

Lys Pro Asn Asn Lys Thr Gly Gln Gln Ile Glu Asn Ile Ile Arg Ile

565 570 575

Lys Ile Glu Lys Gln Phe Lys Ala Ile Asn His Pro Ser Lys Asn Asn

580 585 590

Lys Gly Ile Pro Lys Ser Leu Phe Ala Asp Thr Asn Val Arg Val Asn

595 600 605

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

610 615 620

Lys Lys His Ile Ala Phe Lys Glu Asn Gln Lys Ala Ser Ser Asn Tyr

625 630 635 640

Asp Lys Leu Leu Lys Glu His Gln Phe Thr Pro Lys Asn Lys Arg Pro

645 650 655

Glu Leu Arg Lys Tyr Val Phe Tyr Lys Ser Glu Lys Gly Glu Glu Ala

660 665 670

Thr Trp Leu Ala Asn Asp Ile Lys Arg Phe Met Pro Lys Asp Phe Lys

675 680 685

Thr Lys Trp Lys Gly Cys Gln His Ser Glu Leu Gln Arg Lys Leu Ala

690 695 700

Phe Tyr Asp Arg His Thr Lys Gln Asp Ile Lys Glu Leu Leu Ser Gly

705 710 715 720

Cys Glu Phe Asp His Ser Leu Leu Asp Ile Asn Ala Tyr Phe Gln Lys
725 730 735

Asp Asn Phe Glu Asp Phe Phe Ser Lys Tyr Leu Glu Asn Arg Ile Glu
740 745 750

Thr Leu Glu Gly Val Leu Lys Lys Leu His Asp Phe Lys Asn Glu Pro
755 760 765

Thr Pro Leu Lys Gly Val Phe Lys Asn Cys Phe Lys Phe Leu Lys Arg
770 775 780

Gln Asn Tyr Val Thr Glu Ser Pro Glu Ile Ile Lys Lys Arg Ile Leu
785 790 795 800

Ala Lys Pro Thr Phe Leu Pro Arg Gly Val Phe Asp Glu Arg Pro Thr
805 810 815

Met Lys Lys Gly Lys Asn Pro Leu Lys Asp Lys Asn Glu Phe Ala Glu
820 825 830

Trp Phe Val Glu Tyr Leu Glu Asn Lys Asp Tyr Gln Lys Phe Tyr Asn
835 840 845

Ala Glu Glu Tyr Arg Met Arg Asp Ala Asp Phe Lys Lys Asn Ala Val
850 855 860

Ile Lys Lys Gln Lys Leu Lys Asp Phe Tyr Thr Leu Gln Met Val Asn
865 870 875 880

Tyr Leu Leu Lys Glu Val Phe Gly Lys Asp Glu Met Asn Leu Gln Leu
885 890 895

Ser Glu Leu Phe Gln Thr Arg Gln Glu Arg Leu Lys Leu Gln Gly Ile
900 905 910

Ala Lys Lys Gln Met Asn Lys Glu Thr Gly Asp Ser Ser Glu Asn Thr
915 920 925

Arg Asn Gln Thr Tyr Ile Trp Asn Lys Asp Val Pro Val Ser Phe Phe
930 935 940

Asn Gly Lys Val Thr Ile Asp Lys Val Lys Leu Lys Asn Ile Gly Lys
945 950 955 960

Tyr Lys Arg Tyr Glu Arg Asp Glu Arg Val Lys Thr Phe Ile Gly Tyr

965 970 975

Glu Val Asp Glu Lys Trp Met Met Tyr Leu Pro His Asn Trp Lys Asp

980 985 990

Arg Tyr Ser Val Lys Pro Ile Asn Val Ile Asp Leu Gln Ile Gln Glu

995 1000 1005

Tyr Glu Glu Ile Arg Ser His Glu Leu Leu Lys Glu Ile Gln Asn

1010 1015 1020

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

1025 1030 1035

Leu Gln Asp Gly Asn Pro Asn Phe Lys Met Tyr Val Leu Asn Gly

1040 1045 1050

Leu Leu Ile Gly Ile Lys Gln Val Asn Ile Pro Asp Phe Ile Val

1055 1060 1065

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

1070 1075 1080

Ala Ser Cys Ser Glu Leu Glu Lys Lys Thr Ile Ile Leu Ile Ala

1085 1090 1095

Ile Arg Asn Lys Phe Ala His Asn Gln Leu Pro Asn Lys Met Ile

1100 1105 1110

Tyr Asp Leu Ala Asn Glu Phe Leu Lys Ile Glu Lys Asn Glu Thr

1115 1120 1125

Tyr Ala Asn Tyr Tyr Leu Lys Val Leu Lys Lys Met Ile Ser Asp

1130 1135 1140

Leu Ala

1145

<210> 119

<211> 948

<212> PRT

<213> Riemerella anatipestifer

<400> 119

Met Phe Phe Ser Phe His Asn Ala Gln Arg Val Ile Phe Lys His Leu

1 5 10 15

Tyr Lys Ala Phe Asp Ala Ser Leu Arg Met Val Lys Glu Asp Tyr Lys
20 25 30
Ala His Phe Thr Val Asn Leu Thr Arg Asp Phe Ala His Leu Asn Arg
35 40 45
Lys Gly Lys Asn Lys Gln Asp Asn Pro Asp Phe Asn Arg Tyr Arg Phe

50 55 60
Glu Lys Asp Gly Phe Phe Thr Glu Ser Gly Leu Leu Phe Phe Thr Asn
65 70 75 80
Leu Phe Leu Asp Lys Arg Asp Ala Tyr Trp Met Leu Lys Lys Val Ser
85 90 95
Gly Phe Lys Ala Ser His Lys Gln Arg Glu Lys Met Thr Thr Glu Val
100 105 110
Phe Cys Arg Ser Arg Ile Leu Leu Pro Lys Leu Arg Leu Glu Ser Arg

115 120 125
Tyr Asp His Asn Gln Met Leu Leu Asp Met Leu Ser Glu Leu Ser Arg
130 135 140
Cys Pro Lys Leu Leu Tyr Glu Lys Leu Ser Glu Glu Asn Lys Lys His
145 150 155 160
Phe Gln Val Glu Ala Asp Gly Phe Leu Asp Glu Ile Glu Glu Glu Gln
165 170 175
Asn Pro Phe Lys Asp Thr Leu Ile Arg His Gln Asp Arg Phe Pro Tyr

180 185 190
Phe Ala Leu Arg Tyr Leu Asp Leu Asn Glu Ser Phe Lys Ser Ile Arg
195 200 205
Phe Gln Val Asp Leu Gly Thr Tyr His Tyr Cys Ile Tyr Asp Lys Lys
210 215 220
Ile Gly Asp Glu Gln Glu Lys Arg His Leu Thr Arg Thr Leu Leu Ser
225 230 235 240
Phe Gly Arg Leu Gln Asp Phe Thr Glu Ile Asn Arg Pro Gln Glu Trp

245 250 255
Lys Ala Leu Thr Lys Asp Leu Asp Tyr Lys Glu Thr Ser Asn Gln Pro

260 265 270

Phe Ile Ser Lys Thr Thr Pro His Tyr His Ile Thr Asp Asn Lys Ile

275 280 285

Gly Phe Arg Leu Gly Thr Ser Lys Glu Leu Tyr Pro Ser Leu Glu Ile

290 295 300

Lys Asp Gly Ala Asn Arg Ile Ala Lys Tyr Pro Tyr Asn Ser Gly Phe

305 310 315 320

Val Ala His Ala Phe Ile Ser Val His Glu Leu Leu Pro Leu Met Phe

325 330 335

Tyr Gln His Leu Thr Gly Lys Ser Glu Asp Leu Leu Lys Glu Thr Val

340 345 350

Arg His Ile Gln Arg Ile Tyr Lys Asp Phe Glu Glu Glu Arg Ile Asn

355 360 365

Thr Ile Glu Asp Leu Glu Lys Ala Asn Gln Gly Arg Leu Pro Leu Gly

370 375 380

Ala Phe Pro Lys Gln Met Leu Gly Leu Leu Gln Asn Lys Gln Pro Asp

385 390 395 400

Leu Ser Glu Lys Ala Lys Ile Lys Ile Glu Lys Leu Ile Ala Glu Thr

405 410 415

Lys Leu Leu Ser His Arg Leu Asn Thr Lys Leu Lys Ser Ser Pro Lys

420 425 430

Leu Gly Lys Arg Arg Glu Lys Leu Ile Lys Thr Gly Val Leu Ala Asp

435 440 445

Trp Leu Val Lys Asp Phe Met Arg Phe Gln Pro Val Ala Tyr Asp Ala

450 455 460

Gln Asn Gln Pro Ile Lys Ser Ser Lys Ala Asn Ser Thr Glu Phe Trp

465 470 475 480

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

485 490 495

Glu Gly Tyr Phe Lys Gln Thr Asn Leu Ile Gly Asn Thr Asn Pro His

500 505 510

Pro Phe Leu Asn Lys Phe Asn Trp Lys Ala Cys Arg Asn Leu Val Asp
515 520 525
Phe Tyr Gln Gln Tyr Leu Glu Gln Arg Glu Lys Phe Leu Glu Ala Ile
530 535 540
Lys His Gln Pro Trp Glu Pro Tyr Gln Tyr Cys Leu Leu Leu Lys Val
545 550 555 560
Pro Lys Glu Asn Arg Lys Asn Leu Val Lys Gly Trp Glu Gln Gly Gly

565 570 575
Ile Ser Leu Pro Arg Gly Leu Phe Thr Glu Ala Ile Arg Glu Thr Leu
580 585 590
Ser Lys Asp Leu Thr Leu Ser Lys Pro Ile Arg Lys Glu Ile Lys Lys
595 600 605
His Gly Arg Val Gly Phe Ile Ser Arg Ala Ile Thr Leu Tyr Phe Lys
610 615 620
Glu Lys Tyr Gln Asp Lys His Gln Ser Phe Tyr Asn Leu Ser Tyr Lys

625 630 635 640
Leu Glu Ala Lys Ala Pro Leu Leu Lys Lys Glu Glu His Tyr Glu Tyr
645 650 655
Trp Gln Gln Asn Lys Pro Gln Ser Pro Thr Glu Ser Gln Arg Leu Glu
660 665 670
Leu His Thr Ser Asp Arg Trp Lys Asp Tyr Leu Leu Tyr Lys Arg Trp
675 680 685
Gln His Leu Glu Lys Lys Leu Arg Leu Tyr Arg Asn Gln Asp Ile Met

690 695 700
Leu Trp Leu Met Thr Leu Glu Leu Thr Lys Asn His Phe Lys Glu Leu
705 710 715 720
Asn Leu Asn Tyr His Gln Leu Lys Leu Glu Asn Leu Ala Val Asn Val
725 730 735
Gln Glu Ala Asp Ala Lys Leu Asn Pro Leu Asn Gln Thr Leu Pro Met
740 745 750
Val Leu Pro Val Lys Val Tyr Pro Thr Thr Ala Phe Gly Glu Val Gln

755 760 765

Tyr His Glu Thr Pro Ile Arg Thr Val Tyr Ile Arg Glu Glu Gln Thr

770 775 780

Lys Ala Leu Lys Met Gly Asn Phe Lys Ala Leu Val Lys Asp Arg Arg

785 790 795 800

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

805 810 815

His Pro Ile Ser Gln Leu Arg Leu Arg Arg Glu Leu Glu Ile Tyr Gln

820 825 830

Ser Leu Arg Val Asp Ala Phe Lys Glu Thr Leu Ser Leu Glu Glu Lys

835 840 845

Leu Leu Asn Lys His Ala Ser Leu Ser Ser Leu Glu Asn Glu Phe Arg

850 855 860

Thr Leu Leu Glu Glu Trp Lys Lys Lys Tyr Ala Ala Ser Ser Met Val

865 870 875 880

Thr Asp Lys His Ile Ala Phe Ile Ala Ser Val Arg Asn Ala Phe Cys

885 890 895

His Asn Gln Tyr Pro Phe Tyr Lys Glu Thr Leu His Ala Pro Ile Leu

900 905 910

Leu Phe Thr Val Ala Gln Pro Thr Thr Glu Glu Lys Asp Gly Leu Gly

915 920 925

Ile Ala Glu Ala Leu Leu Lys Val Leu Arg Glu Tyr Cys Glu Ile Val

930 935 940

Lys Ser Gln Ile

945

<210> 120

<211> 1139

<212> PRT

<213> Prevotella pleuritidis

<400> 120

Met Glu Asn Asp Lys Arg Leu Glu Glu Ser Ala Cys Tyr Thr Leu Asn

1 5 10 15

Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val
20 25 30
Tyr Ile Thr Val Asn His Ile Asn Lys Thr Leu Glu Leu Lys Asn Lys
35 40 45
Lys Asn Gln Glu Ile Ile Ile Asp Asn Asp Gln Asp Ile Leu Ala Ile

50 55 60
Lys Thr His Trp Ala Lys Val Asn Gly Asp Leu Asn Lys Thr Asp Arg
65 70 75 80
Leu Arg Glu Leu Met Ile Lys His Phe Pro Phe Leu Glu Ala Ala Ile
85 90 95
Tyr Ser Asn Asn Lys Glu Asp Lys Glu Glu Val Lys Glu Glu Lys Gln
100 105 110
Ala Lys Ala Gln Ser Phe Lys Ser Leu Lys Asp Cys Leu Phe Leu Phe

115 120 125
Leu Glu Lys Leu Gln Glu Ala Arg Asn Tyr Tyr Ser His Tyr Lys Tyr
130 135 140
Ser Glu Ser Ser Lys Glu Pro Glu Phe Glu Glu Gly Leu Leu Glu Lys
145 150 155 160
Met Tyr Asn Thr Phe Asp Ala Ser Ile Arg Leu Val Lys Glu Asp Tyr
165 170 175
Gln Tyr Asn Lys Asp Ile Asp Pro Glu Lys Asp Phe Lys His Leu Glu

180 185 190
Arg Lys Glu Asp Phe Asn Tyr Leu Phe Thr Asp Lys Asp Asn Lys Gly
195 200 205
Lys Ile Thr Lys Asn Gly Leu Leu Phe Phe Val Ser Leu Phe Leu Glu
210 215 220
Lys Lys Asp Ala Ile Trp Met Gln Gln Lys Phe Arg Gly Phe Lys Asp
225 230 235 240
Asn Arg Gly Asn Lys Glu Lys Met Thr His Glu Val Phe Cys Arg Ser

245 250 255
Arg Met Leu Leu Pro Lys Ile Arg Leu Glu Ser Thr Gln Thr Gln Asp

260 265 270

Trp Ile Leu Leu Asp Met Leu Asn Glu Leu Ile Arg Cys Pro Lys Ser

275 280 285

Leu Tyr Glu Arg Leu Gln Gly Ala Tyr Arg Glu Lys Phe Lys Val Pro

290 295 300

Phe Asp Ser Ile Asp Glu Asp Tyr Asp Ala Glu Gln Glu Pro Phe Arg

305 310 315 320

Asn Thr Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Ala Leu Arg

325 330 335

Tyr Phe Asp Tyr Asn Glu Ile Phe Lys Asn Leu Arg Phe Gln Ile Asp

340 345 350

Leu Gly Thr Tyr His Phe Ser Ile Tyr Lys Lys Leu Ile Gly Gly Lys

355 360 365

Lys Glu Asp Arg His Leu Thr His Lys Leu Tyr Gly Phe Glu Arg Ile

370 375 380

Gln Glu Phe Thr Lys Gln Asn Arg Pro Asp Lys Trp Gln Ala Ile Ile

385 390 395 400

Lys Asp Leu Asp Thr Tyr Glu Thr Ser Asn Glu Arg Tyr Ile Ser Glu

405 410 415

Thr Thr Pro His Tyr His Leu Glu Asn Gln Lys Ile Gly Ile Arg Phe

420 425 430

Arg Asn Asp Asn Asn Asp Ile Trp Pro Ser Leu Lys Thr Asn Gly Glu

435 440 445

Lys Asn Glu Lys Ser Lys Tyr Asn Leu Asp Lys Pro Tyr Gln Ala Glu

450 455 460

Ala Phe Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr Tyr Leu

465 470 475 480

Leu Leu Lys Met Glu Asn Thr Asp Asn Asp Lys Glu Asp Asn Glu Val

485 490 495

Gly Thr Lys Lys Lys Gly Asn Lys Asn Asn Lys Gln Glu Lys His Lys

500 505 510

Ile Glu Glu Ile Ile Glu Asn Lys Ile Lys Asp Ile Tyr Ala Leu Tyr
515 520 525
Asp Ala Phe Thr Asn Gly Glu Ile Asn Ser Ile Asp Glu Leu Ala Glu
530 535 540
Gln Arg Glu Gly Lys Asp Ile Glu Ile Gly His Leu Pro Lys Gln Leu
545 550 555 560
Ile Val Ile Leu Lys Asn Lys Ser Lys Asp Met Ala Glu Lys Ala Asn

565 570 575
Arg Lys Gln Lys Glu Met Ile Lys Asp Thr Lys Lys Arg Leu Ala Thr
580 585 590
Leu Asp Lys Gln Val Lys Gly Glu Ile Glu Asp Gly Gly Arg Asn Ile
595 600 605
Arg Leu Leu Lys Ser Gly Glu Ile Ala Arg Trp Leu Val Asn Asp Met
610 615 620
Met Arg Phe Gln Pro Val Gln Lys Asp Asn Glu Gly Lys Pro Leu Asn

625 630 635 640
Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Met Leu Gln Arg Ser Leu
645 650 655
Ala Leu Tyr Asn Lys Glu Glu Lys Pro Thr Arg Tyr Phe Arg Gln Val
660 665 670
Asn Leu Ile Lys Ser Ser Asn Pro His Pro Phe Leu Glu Asp Thr Lys
675 680 685
Trp Glu Glu Cys Tyr Asn Ile Leu Ser Phe Tyr Arg Asn Tyr Leu Lys

690 695 700
Ala Lys Ile Lys Phe Leu Asn Lys Leu Lys Pro Glu Asp Trp Lys Lys
705 710 715 720
Asn Gln Tyr Phe Leu Met Leu Lys Glu Pro Lys Thr Asn Arg Lys Thr
725 730 735
Leu Val Gln Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg Gly Ile Phe
740 745 750
Thr Glu Pro Ile Lys Glu Trp Phe Lys Arg His Gln Asn Asp Ser Glu

```
        755                 760                 765
Glu Tyr Lys Lys Val Glu Ala Leu Asp Arg Val Gly Leu Val Ala Lys
    770                 775                 780
Val Ile Pro Leu Phe Phe Lys Glu Glu Tyr Phe Lys Glu Asp Ala Gln
785                 790                 795                 800
Lys Glu Ile Asn Asn Cys Val Gln Pro Phe Tyr Ser Phe Pro Tyr Asn
                805                 810                 815
Val Gly Asn Ile His Lys Pro Glu Glu Lys Asn Phe Leu His Cys Glu

            820                 825                 830
Glu Arg Arg Lys Leu Trp Asp Lys Lys Lys Asp Lys Phe Lys Gly Tyr
        835                 840                 845
Lys Ala Lys Glu Lys Ser Lys Lys Met Thr Asp Lys Glu Lys Glu Glu
    850                 855                 860
His Arg Ser Tyr Leu Glu Phe Gln Ser Trp Asn Lys Phe Glu Arg Glu
865                 870                 875                 880
Leu Arg Leu Val Arg Asn Gln Asp Ile Leu Thr Trp Leu Leu Cys Thr

                885                 890                 895
Lys Leu Ile Asp Lys Leu Lys Ile Asp Glu Leu Asn Ile Glu Glu Leu
            900                 905                 910
Gln Lys Leu Arg Leu Lys Asp Ile Asp Thr Asp Thr Ala Lys Lys Glu
        915                 920                 925
Lys Asn Asn Ile Leu Asn Arg Val Met Pro Met Arg Leu Pro Val Thr
    930                 935                 940
Val Tyr Glu Ile Asp Lys Ser Phe Asn Ile Val Lys Asp Lys Pro Leu

945                 950                 955                 960
His Thr Val Tyr Ile Glu Glu Thr Gly Thr Lys Leu Leu Lys Gln Gly
                965                 970                 975
Asn Phe Lys Ala Leu Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser
            980                 985                 990
Phe Val Lys Thr Ser Ser Glu Ala  Glu Ser Lys Ser Lys  Pro Ile Ser
        995                 1000                1005
```

Lys Leu Arg Val Glu Tyr Glu Leu Gly Ala Tyr Gln Lys Ala Arg

1010 1015 1020

Ile Asp Ile Ile Lys Asp Met Leu Ala Leu Glu Lys Thr Leu Ile

1025 1030 1035

Asp Asn Asp Glu Asn Leu Pro Thr Asn Lys Phe Ser Asp Met Leu

1040 1045 1050

Lys Ser Trp Leu Lys Gly Lys Gly Glu Ala Asn Lys Ala Arg Leu

1055 1060 1065

Gln Asn Asp Val Gly Leu Leu Val Ala Val Arg Asn Ala Phe Ser

1070 1075 1080

His Asn Gln Tyr Pro Met Tyr Asn Ser Glu Val Phe Lys Gly Met

1085 1090 1095

Lys Leu Leu Ser Leu Ser Ser Asp Ile Pro Glu Lys Glu Gly Leu

1100 1105 1110

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

1115 1120 1125

Arg Ile Ile Glu Ile Glu Lys Glu Ile Arg Asn

1130 1135

<210> 121

<211> 1135

<212> PRT

<213> Porphyromonas gingivalis

<400> 121

Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val

1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu

20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys

35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr
65 70 75 80
Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp
85 90 95
His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile
100 105 110
Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

115 120 125
Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp
130 135 140
Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala
145 150 155 160
Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala
165 170 175
Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190
Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg
195 200 205
Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr
210 215 220
Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys
225 230 235 240
Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255
Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu
260 265 270
Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu
275 280 285
Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu
290 295 300
Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro

420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Arg Ser Met

435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu

450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys

465 470 475 480

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

485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510

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

515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu

530 535 540

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

545 550 555 560

Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575

Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Gln Lys Phe

580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu

595 600 605

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

610 615 620

Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640

Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg His Gln Phe Arg Ala

645 650 655

Ile Val Ala Glu Leu Arg Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670

Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr

675 680 685

Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700

Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg

725 730 735

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

740 745 750

Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765

Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe

770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830

Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro

835 840 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn

900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu

915 920 925

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

930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 122

<211> 1135

<212> PRT

<213> Porphyromonas gingivalis

<400> 122

Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val

1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu

20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys

35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr

65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp

85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile

100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp

130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala

145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala

165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg

195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asn Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu

275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu

290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro

420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met

435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu

450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Gly Phe Leu Arg Lys

465 470 475 480

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

485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510

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

515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu

530 535 540

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

545 550 555 560

Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575

Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe

580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu

595 600 605

Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala
645 650 655
Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe
660 665 670
Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr
675 680 685
Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700
Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe
705 710 715 720
Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg
725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765
Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe
770 775 780
Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr
785 790 795 800
Gly Leu Arg Arg Glu Leu Asn Ile His Gly Lys Ser Val Ser Tyr Ile
805 810 815
Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830
Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro
835 840 845
Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Lys Glu Asn

900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu

915 920 925

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

930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile
1100 1105 1110
Leu Asp His Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met
1115 1120 1125
Ser Gln Pro Ile Asn Asp Leu
1130 1135

<210> 123
<211> 1135
<212> PRT
<213> Porphyromonas gingivalis
<400> 123
Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15
Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30
Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys
35 40 45
Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60
Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr
65 70 75 80
Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp
85 90 95
His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile
100 105 110
Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

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

165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg

195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu

275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu

290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro
420 425 430
Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met

435 440 445
Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu
450 455 460
Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Gly Phe Leu Arg Lys
465 470 475 480
Ala Asn Arg Ile Leu Asp Glu Thr Ala Glu Gly Lys Leu Gln Phe Ser
485 490 495
Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510
Lys Ser Lys Asp Arg Arg Glu Lys Ala Glu Thr Thr Leu Glu Lys Tyr
515 520 525
Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu
530 535 540
Ser Ala Phe Asp Met Asn Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560
Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575
Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe
580 585 590
Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu
595 600 605
Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala
645 650 655
Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670

Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr

675 680 685

Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700

Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg

725 730 735

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

740 745 750

Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765

Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe

770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830

Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro

835 840 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn

900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu

915 920 925

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

930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp His Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 124

<211> 1135

<212> PRT

<213> Porphyromonas gingivalis

<400> 124

Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys
35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu
50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr
65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp
85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile
100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His
115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp
130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala
145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala
165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys
180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg
195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr
210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys
225 230 235 240
Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255
Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu
260 265 270
Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu
275 280 285
Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu
290 295 300
Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320
Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe
325 330 335
Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu
340 345 350
Gly Glu Ile Glu Leu Asp Ser Tyr Ser Lys Lys Val Gly Arg Asn Gly
355 360 365
Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380
Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu
385 390 395 400
Ala Ser Tyr Pro Val Arg Phe Ser Leu Phe Ala Pro Arg Tyr Ala Ile
405 410 415
Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro
420 425 430
Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Arg Ser Met

435 440 445
Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu
450 455 460
Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys

465 470 475 480

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

485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510

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

515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu

530 535 540

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

545 550 555 560

Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575

Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Gln Lys Phe

580 585 590

Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu

595 600 605

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

610 615 620

Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640

Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg His Gln Phe Arg Ala

645 650 655

Ile Val Ala Glu Leu Arg Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670

Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr

675 680 685

Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700

Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe

705 710 715 720

Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg
725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Val Met

755 760 765
Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe
770 775 780
Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr
785 790 795 800
Gly Leu Arg Arg Glu Leu Asn Ile His Gly Lys Ser Val Ser Tyr Ile
805 810 815
Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830
Lys Thr Val Arg Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro
835 840 845
Val Pro Pro Asp Leu Ala Ala Tyr Ile Lys Arg Ser Phe His Arg Ala
850 855 860
Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg
865 870 875 880
Leu Met Leu Met Ala Ile Asn Lys Ile Met Thr Asp Arg Glu Glu Asp

885 890 895
Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Lys Glu Asn
900 905 910
Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu
915 920 925
Gly Gly Asp Asn Ser Leu Ser Leu Val Pro Ala Thr Ile Glu Ile Lys
930 935 940
Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960
Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Ile Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 125

<211> 1135

<212> PRT

<213> Porphyromonas gingivalis

<400> 125

Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val

1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu

20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys

35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr

65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp

85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile

100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp

130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala

145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala

165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg

195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu

275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu

290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro

420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Arg Ser Met

435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu

450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys

465 470 475 480

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

485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510

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

515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu
530 535 540
Ser Ala Phe Asp Met Asn Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560
Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575
Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe
580 585 590
Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu
595 600 605
Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala
645 650 655
Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe
660 665 670
Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr
675 680 685
Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700
Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe
705 710 715 720
Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg
725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765
Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe

770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830

Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro

835 840 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn

900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu

915 920 925

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

930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 126

<211> 1133

<212> PRT

<213> Prevotella falsenii

<400> 126

Met Lys Asn Asp Asn Asn Ser Thr Lys Ser Thr Asp Tyr Thr Leu Gly

1 5 10 15

Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val

20 25 30

Tyr Ile Thr Val Asn His Ile Asn Lys Val Leu Glu Leu Lys Asn Lys

35 40 45

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

50 55 60

Lys Thr Leu Trp Gly Lys Val Asp Thr Asp Ile Asn Lys Lys Asp Arg

65 70 75 80

Leu Arg Glu Leu Ile Met Lys His Phe Pro Phe Leu Glu Ala Ala Thr

85 90 95

Tyr Gln Gln Ser Ser Thr Asn Asn Thr Lys Gln Lys Glu Glu Glu Gln

100 105 110

Ala Lys Ala Gln Ser Phe Glu Ser Leu Lys Asp Cys Leu Phe Leu Phe

115 120 125

Leu Glu Lys Leu Arg Glu Ala Arg Asn Tyr Tyr Ser His Tyr Lys His

130 135 140

Ser Lys Ser Leu Glu Glu Pro Lys Leu Glu Glu Lys Leu Leu Glu Asn

145 150 155 160

Met Tyr Asn Ile Phe Asp Thr Asn Val Gln Leu Val Ile Lys Asp Tyr

165 170 175

Glu His Asn Lys Asp Ile Asn Pro Glu Glu Asp Phe Lys His Leu Gly

180 185 190

Arg Ala Glu Gly Glu Phe Asn Tyr Tyr Phe Thr Arg Asn Lys Lys Gly

195 200 205

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

210 215 220

Lys Lys Asp Ala Ile Trp Ala Gln Thr Lys Ile Lys Gly Phe Lys Asp

225 230 235 240

Asn Arg Glu Asn Lys Gln Lys Met Thr His Glu Val Phe Cys Arg Ser

245 250 255

Arg Met Leu Leu Pro Lys Leu Arg Leu Glu Ser Thr Gln Thr Gln Asp

260 265 270

Trp Ile Leu Leu Asp Met Leu Asn Glu Leu Ile Arg Cys Pro Lys Ser

275 280 285

Leu Tyr Lys Arg Leu Gln Gly Glu Lys Arg Glu Lys Phe Arg Val Pro

290 295 300

Phe Asp Pro Ala Asp Glu Asp Tyr Asp Ala Glu Gln Glu Pro Phe Lys

305 310 315 320

Asn Thr Leu Val Arg His Gln Asp Arg Phe Pro Tyr Phe Ala Leu Arg

325 330 335

Tyr Phe Asp Tyr Asn Glu Ile Phe Thr Asn Leu Arg Phe Gln Ile Asp
340 345 350
Leu Gly Thr Tyr His Phe Ser Ile Tyr Lys Lys Gln Ile Gly Asp Lys
355 360 365
Lys Glu Asp Arg His Leu Thr His Lys Leu Tyr Gly Phe Glu Arg Ile

370 375 380
Gln Glu Phe Ala Lys Glu Asn Arg Pro Asp Glu Trp Lys Ala Leu Val
385 390 395 400
Lys Asp Leu Asp Thr Phe Glu Glu Ser Asn Glu Pro Tyr Ile Ser Glu
405 410 415
Thr Thr Pro His Tyr His Leu Glu Asn Gln Lys Ile Gly Ile Arg Asn
420 425 430
Lys Asn Lys Lys Lys Lys Lys Thr Ile Trp Pro Ser Leu Glu Thr Lys

435 440 445
Thr Thr Val Asn Glu Arg Ser Lys Tyr Asn Leu Gly Lys Ser Phe Lys
450 455 460
Ala Glu Ala Phe Leu Ser Val His Glu Leu Leu Pro Met Met Phe Tyr
465 470 475 480
Tyr Leu Leu Leu Asn Lys Glu Glu Pro Asn Asn Gly Lys Ile Asn Ala
485 490 495
Ser Lys Val Glu Gly Ile Ile Glu Lys Lys Ile Arg Asp Ile Tyr Lys

500 505 510
Leu Tyr Gly Ala Phe Ala Asn Glu Glu Ile Asn Asn Glu Glu Glu Leu
515 520 525
Lys Glu Tyr Cys Glu Gly Lys Asp Ile Ala Ile Arg His Leu Pro Lys
530 535 540
Gln Met Ile Ala Ile Leu Lys Asn Glu Tyr Lys Asp Met Ala Lys Lys
545 550 555 560
Ala Glu Asp Lys Gln Lys Lys Met Ile Lys Asp Thr Lys Lys Arg Leu

565 570 575
Ala Ala Leu Asp Lys Gln Val Lys Gly Glu Val Glu Asp Gly Gly Arg

580 585 590

Asn Ile Lys Pro Leu Lys Ser Gly Arg Ile Ala Ser Trp Leu Val Asn

595 600 605

Asp Met Met Arg Phe Gln Pro Val Gln Arg Asp Arg Asp Gly Tyr Pro

610 615 620

Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Gln Leu Leu Gln Arg

625 630 635 640

Thr Leu Ala Leu Phe Gly Ser Glu Arg Glu Arg Leu Ala Pro Tyr Phe

645 650 655

Arg Gln Met Asn Leu Ile Gly Lys Asp Asn Pro His Pro Phe Leu Lys

660 665 670

Asp Thr Lys Trp Lys Glu His Asn Asn Ile Leu Ser Phe Tyr Arg Ser

675 680 685

Tyr Leu Glu Ala Lys Lys Asn Phe Leu Gly Ser Leu Lys Pro Glu Asp

690 695 700

Trp Lys Lys Asn Gln Tyr Phe Leu Lys Leu Lys Glu Pro Lys Thr Asn

705 710 715 720

Arg Glu Thr Leu Val Gln Gly Trp Lys Asn Gly Phe Asn Leu Pro Arg

725 730 735

Gly Ile Phe Thr Glu Pro Ile Arg Glu Trp Phe Ile Arg His Gln Asn

740 745 750

Glu Ser Glu Glu Tyr Lys Lys Val Lys Asp Phe Asp Arg Ile Gly Leu

755 760 765

Val Ala Lys Val Ile Pro Leu Phe Phe Lys Glu Asp Tyr Gln Lys Glu

770 775 780

Ile Glu Asp Tyr Val Gln Pro Phe Tyr Gly Tyr Pro Phe Asn Val Gly

785 790 795 800

Asn Ile His Asn Ser Gln Glu Gly Thr Phe Leu Asn Lys Lys Glu Arg

805 810 815

Glu Glu Leu Trp Lys Gly Asn Lys Thr Lys Phe Lys Asp Tyr Lys Thr

820 825 830

Lys Glu Lys Asn Lys Glu Lys Thr Asn Lys Asp Lys Phe Lys Lys Lys
835 840 845
Thr Asp Glu Glu Lys Glu Glu Phe Arg Ser Tyr Leu Asp Phe Gln Ser
850 855 860
Trp Lys Lys Phe Glu Arg Glu Leu Arg Leu Val Arg Asn Gln Asp Ile
865 870 875 880
Val Thr Trp Leu Leu Cys Met Glu Leu Ile Asp Lys Leu Lys Ile Asp

885 890 895
Glu Leu Asn Ile Glu Glu Leu Gln Lys Leu Arg Leu Lys Asp Ile Asp
900 905 910
Thr Asp Thr Ala Lys Lys Glu Lys Asn Asn Ile Leu Asn Arg Ile Met
915 920 925
Pro Met Glu Leu Pro Val Thr Val Tyr Glu Thr Asp Asp Ser Asn Asn
930 935 940
Ile Ile Lys Asp Lys Pro Leu His Thr Ile Tyr Ile Lys Glu Ala Glu

945 950 955 960
Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys Asp Arg
965 970 975
Arg Leu Asn Gly Leu Phe Ser Phe Val Glu Thr Ser Ser Glu Ala Glu
980 985 990
Leu Lys Ser Lys Pro Ile Ser Lys Ser Leu Val Glu Tyr Glu Leu Gly
995 1000 1005
Glu Tyr Gln Arg Ala Arg Val Glu Ile Ile Lys Asp Met Leu Arg

1010 1015 1020
Leu Glu Glu Thr Leu Ile Gly Asn Asp Glu Lys Leu Pro Thr Asn
1025 1030 1035
Lys Phe Arg Gln Met Leu Asp Lys Trp Leu Glu His Lys Lys Glu
1040 1045 1050
Thr Asp Asp Thr Asp Leu Lys Asn Asp Val Lys Leu Leu Thr Glu
1055 1060 1065
Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Arg Asp Arg

1070 1075 1080

Ile Ala Phe Ala Asn Ile Lys Pro Phe Ser Leu Ser Ser Ala Asn

1085 1090 1095

Thr Ser Asn Glu Glu Gly Leu Gly Ile Ala Lys Lys Leu Lys Asp

1100 1105 1110

Lys Thr Lys Glu Thr Ile Asp Arg Ile Ile Glu Ile Glu Glu Gln

1115 1120 1125

Thr Ala Thr Lys Arg

1130

<210> 127

<211> 1116

<212> PRT

<213> Prevotella pleuritidis

<400> 127

Met Glu Asn Asp Lys Arg Leu Glu Glu Ser Thr Cys Tyr Thr Leu Asn

1 5 10 15

Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn Val

20 25 30

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

35 40 45

Asp Asn Asp Glu Lys Val Leu Asp Ile Lys Ala Leu Trp Gln Lys Val

50 55 60

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

65 70 75 80

His Phe Pro Phe Leu Glu Ala Ala Ile Tyr Ser Asn Asn Lys Glu Asp

85 90 95

Lys Glu Glu Val Lys Glu Glu Lys Gln Ala Lys Ala Gln Ser Phe Lys

100 105 110

Ser Leu Lys Asp Cys Leu Phe Leu Phe Leu Glu Lys Leu Gln Glu Ala

115 120 125

Arg Asn Tyr Tyr Ser His Tyr Lys Ser Ser Glu Ser Ser Lys Glu Pro

130 135 140

Glu Phe Glu Glu Gly Leu Leu Glu Lys Met Tyr Asn Thr Phe Gly Val
145 150 155 160
Ser Ile Arg Leu Val Lys Glu Asp Tyr Gln Tyr Asn Lys Asp Ile Asp
165 170 175
Pro Glu Lys Asp Phe Lys His Leu Glu Arg Lys Glu Asp Phe Asn Tyr

180 185 190
Leu Phe Thr Asp Lys Asp Asn Lys Gly Lys Ile Thr Lys Asn Gly Leu
195 200 205
Leu Phe Phe Val Ser Leu Phe Leu Glu Lys Lys Asp Ala Ile Trp Met
210 215 220
Gln Gln Lys Leu Arg Gly Phe Lys Asp Asn Arg Gly Asn Lys Glu Lys
225 230 235 240
Met Thr His Glu Val Phe Cys Arg Ser Arg Met Leu Leu Pro Lys Ile

245 250 255
Arg Leu Glu Ser Thr Gln Thr Gln Asp Trp Ile Leu Leu Asp Met Leu
260 265 270
Asn Glu Leu Ile Arg Cys Pro Lys Ser Leu Tyr Glu Arg Leu Gln Gly
275 280 285
Ala Tyr Arg Glu Lys Phe Lys Val Pro Phe Asp Ser Ile Asp Glu Asp
290 295 300
Tyr Asp Ala Glu Gln Glu Pro Phe Arg Asn Thr Leu Val Arg His Gln

305 310 315 320
Asp Arg Phe Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Tyr Asn Glu Ile
325 330 335
Phe Lys Asn Leu Arg Phe Gln Ile Asp Leu Gly Thr Tyr His Phe Ser
340 345 350
Ile Tyr Lys Lys Leu Ile Gly Asp Asn Lys Glu Asp Arg His Leu Thr
355 360 365
His Lys Leu Tyr Gly Phe Glu Arg Ile Gln Glu Phe Ala Lys Gln Lys

370 375 380
Arg Pro Asn Glu Trp Gln Ala Leu Val Lys Asp Leu Asp Ile Tyr Glu

385 390 395 400

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

405 410 415

Glu Asn Gln Lys Ile Gly Ile Arg Phe Lys Asn Lys Lys Asp Lys Ile

420 425 430

Trp Pro Ser Leu Glu Thr Asn Gly Lys Glu Asn Glu Lys Ser Lys Tyr

435 440 445

Asn Leu Asp Lys Ser Phe Gln Ala Glu Ala Phe Leu Ser Ile His Glu

450 455 460

Leu Leu Pro Met Met Phe Tyr Asp Leu Leu Leu Lys Lys Glu Glu Pro

465 470 475 480

Asn Asn Asp Glu Lys Asn Ala Ser Ile Val Glu Gly Phe Ile Lys Lys

485 490 495

Glu Ile Lys Arg Met Tyr Ala Ile Tyr Asp Ala Phe Ala Asn Glu Glu

500 505 510

Ile Asn Ser Lys Glu Gly Leu Glu Glu Tyr Cys Lys Asn Lys Gly Phe

515 520 525

Gln Glu Arg His Leu Pro Lys Gln Met Ile Ala Ile Leu Thr Asn Lys

530 535 540

Ser Lys Asn Met Ala Glu Lys Ala Lys Arg Lys Gln Lys Glu Met Ile

545 550 555 560

Lys Asp Thr Lys Lys Arg Leu Ala Thr Leu Asp Lys Gln Val Lys Gly

565 570 575

Glu Ile Glu Asp Gly Gly Arg Asn Ile Arg Leu Leu Lys Ser Gly Glu

580 585 590

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

595 600 605

Lys Asp Lys Glu Gly Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr

610 615 620

Glu Tyr Gln Met Leu Gln Arg Ser Leu Ala Leu Tyr Asn Lys Glu Gln

625 630 635 640

Lys Pro Thr Pro Tyr Phe Ile Gln Val Asn Leu Ile Lys Ser Ser Asn
645 650 655
Pro His Pro Phe Leu Glu Glu Thr Lys Trp Glu Glu Cys Asn Asn Ile
660 665 670
Leu Ser Phe Tyr Arg Ser Tyr Leu Glu Ala Lys Lys Asn Phe Leu Glu
675 680 685
Ser Leu Lys Pro Glu Asp Trp Lys Lys Asn Gln Tyr Phe Leu Met Leu

690 695 700
Lys Glu Pro Lys Thr Asn Arg Lys Thr Leu Val Gln Gly Trp Lys Asn
705 710 715 720
Gly Phe Asn Leu Pro Arg Gly Ile Phe Thr Glu Pro Ile Lys Glu Trp
725 730 735
Phe Lys Arg His Gln Asn Asp Ser Glu Glu Tyr Lys Lys Val Glu Ala
740 745 750
Leu Asp Arg Val Gly Leu Val Ala Lys Val Ile Pro Leu Phe Phe Lys

755 760 765
Glu Glu Tyr Phe Lys Glu Asp Ala Gln Lys Glu Ile Asn Asn Cys Val
770 775 780
Gln Pro Phe Tyr Ser Phe Pro Tyr Asn Val Gly Asn Ile His Lys Pro
785 790 795 800
Glu Glu Lys Asn Phe Leu His Cys Glu Glu Arg Arg Lys Leu Trp Asp
805 810 815
Lys Lys Lys Asp Lys Phe Lys Gly Tyr Lys Ala Lys Glu Lys Ser Lys

820 825 830
Lys Met Thr Asp Lys Glu Lys Glu Glu His Arg Ser Tyr Leu Glu Phe
835 840 845
Gln Ser Trp Asn Lys Phe Glu Arg Glu Leu Arg Leu Val Arg Asn Gln
850 855 860
Asp Ile Val Thr Trp Leu Leu Cys Thr Glu Leu Ile Asp Lys Leu Lys
865 870 875 880
Ile Asp Glu Leu Asn Ile Glu Glu Leu Gln Lys Leu Arg Leu Lys Asp

885 890 895

Ile Asp Thr Asp Thr Ala Lys Lys Glu Lys Asn Asn Ile Leu Asn Arg

900 905 910

Ile Met Pro Met Gln Leu Pro Val Thr Val Tyr Glu Ile Asp Lys Ser

915 920 925

Phe Asn Ile Val Lys Asp Lys Pro Leu His Thr Ile Tyr Ile Glu Glu

930 935 940

Thr Gly Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ala Leu Val Lys

945 950 955 960

Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe Val Lys Thr Ser Ser Glu

965 970 975

Ala Glu Ser Lys Ser Lys Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu

980 985 990

Leu Gly Ala Tyr Gln Lys Ala Arg Ile Asp Ile Ile Lys Asp Met Leu

995 1000 1005

Ala Leu Glu Lys Thr Leu Ile Asp Asn Asp Glu Asn Leu Pro Thr

1010 1015 1020

Asn Lys Phe Ser Asp Met Leu Lys Ser Trp Leu Lys Gly Lys Gly

1025 1030 1035

Glu Ala Asn Lys Ala Arg Leu Gln Asn Asp Val Asp Leu Leu Val

1040 1045 1050

Ala Ile Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asn

1055 1060 1065

Ser Glu Val Phe Lys Gly Met Lys Leu Leu Ser Leu Ser Ser Asp

1070 1075 1080

Ile Pro Glu Lys Glu Gly Leu Gly Ile Ala Lys Gln Leu Lys Asp

1085 1090 1095

Lys Ile Lys Glu Thr Ile Glu Arg Ile Ile Glu Ile Glu Lys Glu

1100 1105 1110

Ile Arg Asn

1115

<210> 128

<211> 1119

<212> PRT

<213> Porphyromonas gingivalis

<400> 128

Met Thr Glu Gln Asn Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn
20 25 30

Ala Tyr Ile Thr Leu Ala His Ile Asp Arg Gln Leu Ala Tyr Ser Lys
35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp
50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu
65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu
100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu
145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205

Arg Tyr Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp

210 215 220

Arg Glu Gly Thr Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly
245 250 255

Phe Lys Gly Gly Thr Glu Ala Tyr Gln Gln Met Thr Asn Glu Val Phe
260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe
305 310 315 320

Arg Val Pro Ile Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr
325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe
340 345 350

Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser
355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys
370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp
420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly
435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp
465 470 475 480
Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val
500 505 510
Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525

Tyr Ala Val Tyr Asp Ala Phe Ala Arg Gly Glu Ile Asp Thr Leu Asp
530 535 540
Arg Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Arg Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asp Met Glu
565 570 575
Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg

705 710 715 720

Ser Asp Arg Glu Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu
740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765
Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys
770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro
785 790 795 800
Phe Tyr Asp Tyr Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815
Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly
820 825 830
Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala
835 840 845

Lys Glu His Pro Tyr Leu Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg
850 855 860
Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Met Cys
865 870 875 880
Arg Asp Leu Met Glu Glu Asn Lys Val Glu Gly Leu Asp Thr Gly Thr
885 890 895
Leu Tyr Leu Lys Asp Ile Arg Thr Asp Val His Glu Gln Gly Ser Leu
900 905 910

Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr
915 920 925
Arg Ala Asp Ser Arg Gly His Val His Lys Glu Gln Ala Pro Leu Ala
930 935 940
Thr Val Tyr Ile Glu Glu Arg Asp Thr Lys Leu Leu Lys Gln Gly Asn
945 950 955 960

Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe
965 970 975

Val Asp Thr Gly Ala Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu
980 985 990
Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala
995 1000 1005
Phe Glu Gln Thr Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr
1010 1015 1020
Pro His Leu Pro Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp
1025 1030 1035

Ser Asp Pro Leu Leu Asp Lys Trp Pro Asp Leu His Arg Lys Val
1040 1045 1050
Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr
1055 1060 1065
Pro Met Tyr Asp Glu Ala Val Phe Ser Ser Ile Arg Lys Tyr Asp
1070 1075 1080
Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile
1085 1090 1095

Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Met Ala
1100 1105 1110
Glu Arg Ile Ile Gln Val
1115
<210> 129
<211> 1175
<212> PRT
<213> Porphyromonas gulae
<400> 129
Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15
Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn
20 25 30

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

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

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

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

275 280 285

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

Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser
355 360 365
Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys
370 375 380
Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400
Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp
420 425 430
Lys Pro Tyr Ile Ser Gln Thr Ser Pro His Tyr His Ile Glu Lys Gly
435 440 445
Lys Ile Gly Leu Arg Phe Met Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp
465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val
500 505 510
Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525

Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys
530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Lys Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asn Met Glu
565 570 575
Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Ala Ser Gly
625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Asp Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Arg Ala Arg Lys Ala Phe Leu Glu Arg Ile Gly Arg
705 710 715 720
Ser Asp Arg Met Glu Asn Arg Pro Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu
740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765
Gly Tyr Asp Glu Val Gly Ser Tyr Arg Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro
785 790 795 800

Phe Tyr Asp Ser Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815

Gly Arg Phe Leu Ser Lys Glu Glu Arg Ala Glu Glu Trp Glu Arg Gly
820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg
835 840 845

Arg Ile Glu Asp Ala Phe Ala Gly Ile Glu Tyr Ala Ser Pro Gly Asn
850 855 860

Lys Lys Lys Ile Glu Gln Leu Leu Arg Asp Leu Ser Leu Trp Glu Ala
865 870 875 880

Phe Glu Ser Lys Leu Lys Val Arg Ala Asp Lys Ile Asn Leu Ala Lys
885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Gln Glu His Pro Tyr His Asp Phe
900 905 910

Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln
915 920 925

Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys
930 935 940

Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Thr
945 950 955 960

Asn Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn His Val Lys Pro
965 970 975

Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val
980 985 990

His Lys Glu Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp
995 1000 1005

Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp
1010 1015 1020

Arg Arg Leu Asn Gly Leu Phe Ser Phe Val Asp Thr Gly Gly Leu
1025 1030 1035
Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu
1040 1045 1050

Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala Phe Glu Gln Thr
1055 1060 1065
Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro
1070 1075 1080
Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp Ser Asp Pro Leu
1085 1090 1095
Leu Ala Lys Trp Pro Glu Leu His Gly Lys Val Arg Leu Leu Ile
1100 1105 1110

Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp
1115 1120 1125
Glu Ala Val Phe Ser Ser Ile Arg Lys Tyr Asp Pro Ser Ser Pro
1130 1135 1140
Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu
1145 1150 1155
Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val Glu Arg Ile Ile
1160 1165 1170

Gln Ala
1175

<210> 130
<211> 1119
<212> PRT
<213> Porphyromonas gulae
<400> 130

Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15
Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn
20 25 30
Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Gln Asp Val Leu Ser Phe Lys Ala Leu Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Gly Ser Ser Glu Leu

145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp

210 215 220

Gly Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300
Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Lys Asp Arg Ala Arg Phe

305 310 315 320
Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly
325 330 335
Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe
340 345 350
Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser
355 360 365
Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380
Lys Val Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400
Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415
Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp
420 425 430
Lys Pro Tyr Ile Ser Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445
Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp
465 470 475 480
Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510
Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525
Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Arg Asp

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

Pro Lys Gln Met Ile Gly Ile Leu Ser Gln Glu His Lys Asn Met Glu

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu Asp Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

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

705 710 715 720

Ser Asp Arg Val Glu Asn Arg Pro Phe Leu Leu Leu Lys Glu Pro Lys

725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro
785 790 795 800
Phe Tyr Asp Ser Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815
Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830
Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala
835 840 845
Gln Glu His Pro Tyr His Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg
850 855 860
Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Met Cys
865 870 875 880
Arg Asp Leu Met Glu Glu Asn Lys Val Glu Gly Leu Asp Thr Gly Thr

885 890 895
Leu Tyr Leu Lys Asp Ile Arg Pro Asn Val Gln Glu Gln Gly Ser Leu
900 905 910
Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr
915 920 925
Arg Ala Asp Ser Arg Gly His Val His Lys Glu Glu Ala Pro Leu Ala
930 935 940
Thr Val Tyr Ile Glu Glu Arg Asp Thr Lys Leu Leu Lys Gln Gly Asn

945 950 955 960
Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe
965 970 975
Val Asp Thr Gly Gly Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu
980 985 990
Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Val
995 1000 1005
Phe Glu Leu Thr Leu Arg Leu Glu Glu Ser Leu Leu Ser Arg Tyr

1010 1015 1020
Pro His Leu Pro Asp Glu Ser Phe Arg Glu Met Leu Glu Ser Trp

1025 1030 1035

Ser Asp Pro Leu Leu Ala Lys Trp Pro Glu Leu His Gly Lys Val

1040 1045 1050

Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr

1055 1060 1065

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

1070 1075 1080

Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile

1085 1090 1095

Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val

1100 1105 1110

Glu Arg Ile Ile Gln Ala

1115

<210> 131

<211> 1119

<212> PRT

<213> Porphyromonas gulae

<400> 131

Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Gln Asp Val Leu Ser Phe Lys Ala Leu Trp

50 55 60

Lys Asn Phe Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125
Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140
Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Gly Ser Ser Glu Leu
145 150 155 160
Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175
Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190
Asp Pro His Tyr His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205
Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp
210 215 220
Ser Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240
Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly
245 250 255
Phe Lys Gly Gly Thr Gly Pro Tyr Glu Gln Met Thr Asn Glu Val Phe

260 265 270
Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285
Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300
Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Lys Asp Arg Ala Cys Phe
305 310 315 320
Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly

325 330 335
Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe
340 345 350
Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Ser Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Met Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Lys Asp Val

515 520 525

Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys

530 535 540

Glu Leu Asp Ala Cys Ser Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

Pro Lys Gln Met Ile Gly Ile Leu Ser Gln Glu His Lys Asn Met Glu

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu Asp Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

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

705 710 715 720

Ser Asp Arg Val Glu Asn Arg Pro Phe Leu Leu Leu Lys Glu Pro Lys

725 730 735

Asn Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830

Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala

835 840 845

Lys Glu His Pro Tyr His Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg

850 855 860

Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Met Cys

865 870 875 880

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

885 890 895

Leu Tyr Leu Lys Asp Ile Arg Thr Asp Val His Glu Gln Gly Ser Leu

900 905 910

Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr

915 920 925

Arg Ala Asp Ser Arg Gly His Val His Lys Glu Gln Ala Pro Leu Ala

930 935 940

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

945 950 955 960

Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe

965 970 975

Val Asp Thr Gly Gly Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu

980 985 990

Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala

995 1000 1005

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

1010 1015 1020

Pro His Leu Pro Asp Glu Asn Phe Arg Glu Met Leu Glu Ser Trp

1025 1030 1035

Ser Asp Pro Leu Leu Gly Lys Trp Pro Asp Leu His Gly Lys Val

1040 1045 1050

Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr

1055 1060 1065

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

1070 1075 1080

Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile

1085 1090 1095

Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val

1100 1105 1110

Glu Arg Ile Ile Gln Ala

1115

<210> 132

<211> 1175

<212> PRT

<213> Porphyromonas gulae

<400> 132

Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Gln Asp Val Leu Ser Phe Lys Ala Leu Trp

50 55 60

Lys Asn Phe Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Glu Ser Ser Glu Leu

145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp

210 215 220

Gly Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Asp Asp Arg Ala Cys Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly

325 330 335

Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Ser Gln Thr Ser Pro His Tyr His Ile Glu Lys Gly
435 440 445
Lys Ile Gly Leu Arg Phe Met Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480
Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val
500 505 510
Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525
Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys

530 535 540
Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Lys Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asp Met Glu
565 570 575
Glu Lys Ile Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Lys Arg Leu Thr Pro

660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

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

705 710 715 720

Ser Asp Arg Met Glu Asn Arg Pro Phe Leu Leu Leu Lys Glu Pro Lys

725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

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

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Glu Arg Ala Glu Glu Trp Glu Arg Gly

820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg

835 840 845

Arg Ile Glu Asp Ala Phe Ala Gly Ile Glu Tyr Ala Ser Pro Gly Asn

850 855 860

Lys Lys Lys Ile Glu Gln Leu Leu Arg Asp Leu Ser Leu Trp Glu Ala

865 870 875 880

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

885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Gln Glu His Pro Tyr His Asp Phe

900 905 910

Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln

915 920 925

```
Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys
    930                 935                 940
Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Pro
945                 950                 955                 960
Asn Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn Arg Val Lys Pro
                965                 970                 975
Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val

            980                 985                 990
His Lys Glu Glu Ala Pro Leu Ala  Thr Val Tyr Ile Glu  Glu Arg Asp
        995                 1000                1005
Thr Lys  Leu Leu Lys Gln Gly  Asn Phe Lys Ser Phe  Val Lys Asp
    1010                1015                1020
Arg Arg  Leu Asn Gly Leu Phe  Ser Phe Val Asp Thr  Gly Gly Leu
    1025                1030                1035
Ala Met  Glu Gln Tyr Pro Ile  Ser Lys Leu Arg Val  Glu Tyr Glu

    1040                1045                1050
Leu Ala  Lys Tyr Gln Thr Ala  Arg Val Cys Val Phe  Glu Leu Thr
    1055                1060                1065
Leu Arg  Leu Glu Glu Ser Leu  Leu Thr Arg Tyr Pro  His Leu Pro
    1070                1075                1080
Asp Glu  Ser Phe Arg Lys Met  Leu Glu Ser Trp Ser  Asp Pro Leu
    1085                1090                1095
Leu Ala  Lys Trp Pro Glu Leu  His Gly Lys Val Arg  Leu Leu Ile

    1100                1105                1110
Ala Val  Arg Asn Ala Phe Ser  His Asn Gln Tyr Pro  Met Tyr Asp
    1115                1120                1125
Glu Ala  Val Phe Ser Ser Ile  Arg Lys Tyr Asp Pro  Ser Ser Pro
    1130                1135                1140
Asp Ala  Ile Glu Glu Arg Met  Gly Leu Asn Ile Ala  His Arg Leu
    1145                1150                1155
Ser Glu  Glu Val Lys Gln Ala  Lys Glu Thr Val Glu  Arg Ile Ile
```

1160 1165 1170

Gln Val

1175

<210> 133

<211> 1119

<212> PRT

<213> Porphyromonas gulae

<400> 133

Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn
20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys
35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp
50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu
65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95

Phe Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu
100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Gly Ser Ser Glu Leu
145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Glu Val
180 185 190

Asp Pro His Tyr His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205
Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp
210 215 220
Gly Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly
245 250 255
Phe Lys Gly Gly Thr Glu Pro Tyr Glu Gln Met Thr Asn Glu Val Phe
260 265 270
Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285
Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300

Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Lys Asp Arg Ala Cys Phe
305 310 315 320
Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly
325 330 335
Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe
340 345 350
Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Leu Lys Lys Val Phe Thr Ser
355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys
370 375 380
Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400
Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415
Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp
420 425 430

Lys Pro Tyr Ile Ser Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

Pro Lys Gln Met Ile Gly Ile Leu Ser Gln Glu Arg Lys Asp Met Glu

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Arg Ala Arg Lys Ala Phe Leu Glu Arg Ile Gly Arg
705 710 715 720
Ser Asp Arg Val Glu Asn Cys Pro Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu
740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765
Gly Tyr Asp Glu Val Gly Ser Tyr Arg Glu Val Gly Phe Met Ala Lys
770 775 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro
785 790 795 800
Phe Tyr Asp Ser Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly
820 825 830
Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala
835 840 845
Gln Glu His Pro Tyr His Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg
850 855 860
Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Met Cys
865 870 875 880

Arg Asp Leu Met Glu Glu Asn Lys Val Glu Gly Leu Asp Thr Gly Thr
885 890 895
Leu Tyr Leu Lys Asp Ile Arg Pro Asn Val Gln Glu Gln Gly Ser Leu
900 905 910
Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr
915 920 925
Arg Ala Asp Ser Arg Gly His Val His Lys Glu Glu Ala Pro Leu Ala

930 935 940

Thr Val Tyr Ile Glu Glu Arg Asp Thr Lys Leu Leu Lys Gln Gly Asn
945 950 955 960
Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe
965 970 975
Val Asp Thr Gly Ala Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu
980 985 990
Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala
995 1000 1005

Phe Glu Gln Thr Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr
1010 1015 1020
Pro His Leu Pro Asp Glu Ser Phe Arg Glu Met Leu Glu Ser Trp
1025 1030 1035
Ser Asp Pro Leu Leu Thr Lys Trp Pro Glu Leu His Gly Lys Val
1040 1045 1050
Arg Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr
1055 1060 1065

Pro Met Tyr Asp Glu Ala Val Phe Ser Ser Ile Trp Lys Tyr Asp
1070 1075 1080
Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile
1085 1090 1095
Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Ile
1100 1105 1110
Glu Arg Ile Ile Gln Ala
1115

<210> 134
<211> 1174
<212> PRT
<213> Porphyromonas gulae

<400> 134
Met Thr Glu Gln Ser Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Leu Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Gln Asp Val Leu Ser Phe Lys Ala Leu Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Gly Ser Ser Glu Leu

145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Tyr His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp

210 215 220

Ser Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Gly Pro Tyr Glu Gln Met Thr Asn Glu Val Phe
260 265 270
Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285
Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys
290 295 300
Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Lys Asp Arg Ala Cys Phe
305 310 315 320

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

Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400
Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415
Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Leu Glu Thr Gly Asp
420 425 430
Lys Pro Tyr Ile Ser Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly
435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Cys Ala Gln Asp
465 470 475 480
Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525
Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys
530 535 540
Glu Leu Asp Thr Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Lys Gln Met Ile Thr Ile Leu Ser Gln Glu Arg Lys Asp Met Lys
565 570 575

Glu Lys Ile Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Ala Ser Gly
625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700

Arg Ser Tyr Leu Arg Ala Arg Lys Ala Phe Leu Glu Arg Ile Gly Arg
705 710 715 720
Ser Asp Arg Val Glu Asn Cys Pro Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Asp Glu Phe His Leu
740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Arg Glu Val Gly Phe Met Ala Lys
770 775 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro
785 790 795 800
Phe Tyr Asp Ser Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815
Gly Arg Phe Leu Ser Lys Glu Asp Arg Ala Glu Glu Trp Glu Arg Gly
820 825 830

Met Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg
835 840 845
Arg Ile Lys Asp Ala Phe Ala Gly Ile Glu Tyr Ala Ser Pro Gly Asn
850 855 860
Lys Lys Lys Ile Glu Gln Leu Leu Arg Asp Leu Ser Leu Trp Glu Ala
865 870 875 880
Phe Glu Ser Lys Leu Lys Val Arg Ala Asp Lys Ile Asn Leu Ala Lys
885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Gln Glu His Pro Tyr His Asp Phe
900 905 910
Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln
915 920 925
Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys
930 935 940
Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Pro
945 950 955 960

Asn Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn Arg Val Lys Pro
965 970 975
Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val
980 985 990
His Lys Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asn Thr

995 1000 1005

Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp Arg

1010 1015 1020

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

1025 1030 1035

Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu Leu

1040 1045 1050

Ala Lys Tyr Gln Thr Ala Arg Val Cys Val Phe Glu Leu Thr Leu

1055 1060 1065

Arg Leu Glu Glu Ser Leu Leu Ser Arg Tyr Pro His Leu Pro Asp

1070 1075 1080

Glu Ser Phe Arg Glu Met Leu Glu Ser Trp Ser Asp Pro Leu Leu

1085 1090 1095

Ala Lys Trp Pro Glu Leu His Gly Lys Val Arg Leu Leu Ile Ala

1100 1105 1110

Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp Glu

1115 1120 1125

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

1130 1135 1140

Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu Ser

1145 1150 1155

Glu Glu Val Lys Gln Ala Lys Glu Thr Val Glu Arg Ile Ile Gln

1160 1165 1170

Ala

<210> 135

<211> 1135

<212> PRT

<213> Porphyromonas gulae

<400> 135

Met Asn Thr Val Pro Ala Thr Glu Asn Lys Gly Gln Ser Arg Thr Val

1 5 10 15

Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu

20 25 30

Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys

35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60

Leu Ser Ile Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr

65 70 75 80

Leu Pro Phe Leu His Cys Phe Asp Pro Asp Ser Gly Ile Glu Lys Asp

85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile

100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Lys Val Ser Pro Asp

130 135 140

Ile Ser Ser Phe Ile Thr Gly Ala Tyr Thr Phe Ala Cys Glu Arg Ala

145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Leu Leu Ala

165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190

Leu Thr Val Ser Gly Phe Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg

195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu

275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu

290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro

420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met

435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu

450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Asp Phe Leu Arg Lys

465 470 475 480

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

485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510

Lys Ser Lys Asp Arg Arg Glu Lys Ala Glu Thr Thr Leu Glu Lys Tyr
515 520 525
Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu

530 535 540
Ser Ala Phe Asp Met Asn Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560
Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr
565 570 575
Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe
580 585 590
Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu

595 600 605
Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg
625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala
645 650 655
Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe

660 665 670
Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr
675 680 685
Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr
690 695 700
Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe
705 710 715 720
Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg

725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765

Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe

770 775 780

Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830

Lys Thr Val Arg Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro

835 840 845

Val Pro Pro Asp Leu Ala Ala Tyr Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Glu Glu Asn

900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu

915 920 925

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

930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp His Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 136

<211> 1218

<212> PRT

<213> Capnocytophaga cynodegmi

<400> 136

Met Glu Asn Lys Thr Ser Leu Gly Asn Asn Ile Tyr Tyr Asn Pro Phe

1 5 10 15

Lys Pro Gln Asp Lys Ser Tyr Phe Ala Gly Tyr Leu Asn Ala Ala Met

20 25 30

Glu Asn Ile Asp Ser Val Phe Arg Glu Leu Gly Lys Arg Leu Lys Gly

35 40 45

Lys Glu Tyr Thr Ser Glu Asn Phe Phe Asp Ala Ile Phe Lys Glu Asn

50 55 60

Ile Ser Leu Val Glu Tyr Glu Arg Tyr Val Lys Leu Leu Ser Asp Tyr

65 70 75 80

Phe Pro Met Ala Arg Leu Leu Asp Lys Lys Glu Val Pro Ile Lys Glu

85 90 95

Arg Lys Glu Asn Phe Lys Lys Asn Phe Arg Gly Ile Ile Lys Ala Val

100 105 110

Arg Asp Leu Arg Asn Phe Tyr Thr His Lys Glu His Gly Glu Val Glu

115 120 125

Ile Thr Asp Glu Ile Phe Gly Val Leu Asp Glu Met Leu Lys Ser Thr

130 135 140

Val Leu Thr Val Lys Lys Lys Lys Ile Lys Thr Asp Lys Thr Lys Glu

145 150 155 160

Ile Leu Lys Lys Ser Ile Glu Lys Gln Leu Asp Ile Leu Cys Gln Lys

165 170 175

Lys Leu Glu Tyr Leu Lys Asp Thr Ala Arg Lys Ile Glu Glu Lys Arg

180 185 190

Arg Asn Gln Arg Glu Arg Gly Glu Lys Lys Leu Val Pro Arg Phe Glu

195 200 205

Tyr Ser Asp Arg Arg Asp Asp Leu Ile Ala Ala Ile Tyr Asn Asp Ala

210 215 220

Phe Asp Val Tyr Ile Asp Lys Lys Lys Asp Ser Leu Lys Glu Ser Ser

225 230 235 240

Lys Thr Lys Tyr Asn Thr Glu Ser Tyr Pro Gln Gln Glu Glu Gly Asp

245 250 255

Leu Lys Ile Pro Ile Ser Lys Asn Gly Val Val Phe Leu Leu Ser Leu

260 265 270

Phe Leu Ser Lys Gln Glu Val His Ala Phe Lys Ser Lys Ile Ala Gly

275 280 285

Phe Lys Ala Thr Val Ile Asp Glu Ala Thr Val Ser His Arg Lys Asn

290 295 300

Ser Ile Cys Phe Met Ala Thr His Glu Ile Phe Ser His Leu Ala Tyr

305 310 315 320

Lys Lys Leu Lys Arg Lys Val Arg Thr Ala Glu Ile Asn Tyr Ser Glu
325 330 335
Ala Glu Asn Ala Glu Gln Leu Ser Ile Tyr Ala Lys Glu Thr Leu Met

340 345 350
Met Gln Met Leu Asp Glu Leu Ser Lys Val Pro Asp Val Val Tyr Gln
355 360 365
Asn Leu Ser Glu Asp Val Gln Lys Thr Phe Ile Glu Asp Trp Asn Glu
370 375 380
Tyr Leu Lys Glu Asn Asn Gly Asp Val Gly Thr Met Glu Glu Glu Gln
385 390 395 400
Val Ile His Pro Val Ile Arg Lys Arg Tyr Glu Asp Lys Phe Asn Tyr

405 410 415
Phe Ala Ile Arg Phe Leu Asp Glu Phe Ala Gln Phe Pro Thr Leu Arg
420 425 430
Phe Gln Val His Leu Gly Asn Tyr Leu His Asp Ser Arg Pro Lys Glu
435 440 445
His Leu Ile Ser Asp Arg Arg Ile Lys Glu Lys Ile Thr Val Phe Gly
450 455 460
Arg Leu Ser Glu Leu Glu His Lys Lys Ala Leu Phe Ile Lys Asn Thr

465 470 475 480
Glu Thr Asn Glu Asp Arg Lys His Tyr Trp Glu Val Phe Pro Asn Pro
485 490 495
Asn Tyr Asp Phe Pro Lys Glu Asn Ile Ser Val Asn Asp Lys Asp Phe
500 505 510
Pro Ile Ala Gly Ser Ile Leu Asp Arg Glu Lys Gln Pro Thr Ala Gly
515 520 525
Lys Ile Gly Ile Lys Val Asn Leu Leu Asn Gln Lys Tyr Ile Ser Glu

530 535 540
Val Asp Lys Ala Val Lys Ala His Gln Leu Lys Gln Arg Asn Asn Lys
545 550 555 560
Pro Ser Ile Gln Asn Ile Ile Glu Glu Ile Val Pro Ile Asn Gly Ser

565 570 575

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

580 585 590

Ser Met Asn Asp Ile His Ser Ile Leu Tyr Glu Phe Phe Asp Lys Trp

595 600 605

Glu Lys Lys Lys Glu Lys Leu Glu Lys Lys Gly Glu Lys Glu Leu Arg

610 615 620

Lys Glu Ile Gly Lys Glu Leu Glu Glu Lys Ile Val Gly Lys Ile Gln

625 630 635 640

Thr Gln Ile Gln Gln Ile Ile Asp Lys Asp Ile Asn Ala Lys Ile Leu

645 650 655

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

660 665 670

Ile Lys Asp Leu Lys Gln Glu Gln Lys Ile Leu Gln Lys Leu Lys Asn

675 680 685

Glu Gln Thr Ala Arg Glu Lys Glu Tyr Gln Glu Cys Ile Ala Tyr Gln

690 695 700

Glu Glu Ser Arg Lys Ile Lys Arg Ser Asp Lys Ser Arg Gln Lys Tyr

705 710 715 720

Leu Arg Asn Gln Leu Lys Arg Lys Tyr Pro Glu Val Pro Thr Arg Lys

725 730 735

Glu Ile Leu Tyr Tyr Gln Glu Lys Gly Lys Val Ala Val Trp Leu Ala

740 745 750

Asn Asp Ile Lys Arg Phe Met Pro Thr Asp Phe Lys Asn Glu Trp Lys

755 760 765

Gly Glu Gln His Ser Leu Leu Gln Lys Ser Leu Ala Tyr Tyr Glu Gln

770 775 780

Cys Lys Glu Glu Leu Lys Asn Leu Leu Pro Gln Gln Lys Val Phe Lys

785 790 795 800

His Leu Pro Phe Glu Leu Gly Gly His Phe Gln Gln Lys Tyr Leu Tyr

805 810 815

Gln Phe Tyr Thr Arg Tyr Leu Asp Lys Arg Leu Glu His Ile Ser Gly
820 825 830
Leu Val Gln Gln Ala Glu Asn Phe Lys Asn Glu Asn Lys Val Phe Lys
835 840 845
Lys Val Glu Asn Glu Cys Phe Lys Phe Leu Lys Lys Gln Asn Tyr Thr

850 855 860
His Lys Gly Leu Asp Ala Gln Ala Gln Ser Val Leu Gly Tyr Pro Ile
865 870 875 880
Phe Leu Glu Arg Gly Phe Met Asp Glu Lys Pro Thr Ile Ile Lys Gly
885 890 895
Lys Thr Phe Lys Gly Asn Glu Ser Leu Phe Thr Asp Trp Phe Arg Tyr
900 905 910
Tyr Lys Glu Tyr Gln Asn Phe Gln Thr Phe Tyr Asp Thr Glu Asn Tyr

915 920 925
Pro Leu Val Glu Leu Glu Lys Lys Gln Ala Asp Arg Lys Arg Glu Thr
930 935 940
Lys Ile Tyr Gln Gln Lys Lys Asn Asp Val Phe Thr Leu Leu Met Ala
945 950 955 960
Lys His Ile Phe Lys Ser Val Phe Lys Gln Asp Ser Ile Asp Arg Phe
965 970 975
Ser Leu Glu Asp Leu Tyr Gln Ser Arg Glu Glu Arg Leu Glu Asn Gln

980 985 990
Glu Lys Ala Lys Gln Thr Gly Glu Arg Asn Thr Asn Tyr Ile Trp Asn
995 1000 1005
Lys Thr Val Asp Leu Asn Leu Cys Asp Gly Lys Val Thr Val Glu
1010 1015 1020
Asn Val Lys Leu Lys Asn Val Gly Asn Phe Ile Lys Tyr Glu Tyr
1025 1030 1035
Asp Gln Arg Val Gln Thr Phe Leu Lys Tyr Glu Glu Asn Ile Lys

1040 1045 1050
Trp Gln Ala Phe Leu Ile Lys Glu Ser Lys Glu Glu Glu Asn Tyr

1055 1060 1065

Pro Tyr Ile Val Glu Arg Glu Ile Glu Gln Tyr Glu Lys Val Arg

1070 1075 1080

Arg Glu Glu Leu Leu Lys Glu Val His Leu Ile Glu Glu Tyr Ile

1085 1090 1095

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

1100 1105 1110

Gln Asn Phe Lys Tyr Tyr Ile Leu Asn Gly Leu Leu Lys Gln Leu

1115 1120 1125

Lys Asn Glu Asp Val Glu Ser Tyr Lys Val Phe Asn Leu Asn Thr

1130 1135 1140

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

1145 1150 1155

Asp Leu Glu Gln Lys Ala Phe Val Leu Thr Tyr Ile Arg Asn Lys

1160 1165 1170

Phe Ala His Asn Gln Leu Pro Lys Lys Glu Phe Trp Asp Tyr Cys

1175 1180 1185

Gln Glu Lys Tyr Gly Lys Ile Glu Lys Glu Lys Thr Tyr Ala Glu

1190 1195 1200

Tyr Phe Ala Glu Val Phe Lys Arg Glu Lys Glu Ala Leu Met Lys

1205 1210 1215

<210> 137

<211> 1090

<212> PRT

<213> Prevotella species

<400> 137

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

1 5 10 15

Tyr Ser Val Met Ala Met Leu Asn Ala Gln Thr Val Leu Asp His Ile

20 25 30

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

35 40 45

Leu Trp Phe His Pro Val Met Ser His Leu Tyr Asn Ala Lys Asn Gly
50 55 60

Tyr Asp Lys Gln Pro Glu Lys Thr Met Phe Ile Ile Glu Arg Leu Gln
65 70 75 80
Ser Tyr Phe Pro Phe Leu Lys Ile Met Ala Glu Asn Gln Arg Glu Tyr
85 90 95
Ser Asn Gly Lys Tyr Lys Gln Asn Arg Val Glu Val Asn Ser Asn Asp
100 105 110
Ile Phe Glu Val Leu Lys Arg Ala Phe Gly Val Leu Lys Met Tyr Arg
115 120 125

Asp Leu Thr Asn His Tyr Lys Thr Tyr Glu Glu Lys Leu Ile Asp Gly
130 135 140
Cys Glu Phe Leu Thr Ser Thr Glu Gln Pro Leu Ser Gly Met Ile Ser
145 150 155 160
Lys Tyr Tyr Thr Val Ala Leu Arg Asn Thr Lys Glu Arg Tyr Gly Tyr
165 170 175
Lys Thr Glu Asp Leu Ala Phe Ile Gln Asp Asn Ile Lys Lys Ile Thr
180 185 190

Lys Asp Ala Tyr Gly Lys Arg Lys Ser Gln Val Asn Thr Gly Phe Phe
195 200 205
Leu Ser Leu Gln Asp Tyr Asn Gly Asp Thr Gln Lys Lys Leu His Leu
210 215 220
Ser Gly Val Gly Ile Ala Leu Leu Ile Cys Leu Phe Leu Asp Lys Gln
225 230 235 240
Tyr Ile Asn Ile Phe Leu Ser Arg Leu Pro Ile Phe Ser Ser Tyr Asn
245 250 255

Ala Gln Ser Glu Glu Arg Arg Ile Ile Ile Arg Ser Phe Gly Ile Asn
260 265 270
Ser Ile Lys Leu Pro Lys Asp Arg Ile His Ser Glu Lys Ser Asn Lys
275 280 285
Ser Val Ala Met Asp Met Leu Asn Glu Val Lys Arg Cys Pro Asp Glu

290 295 300

Leu Phe Thr Thr Leu Ser Ala Glu Lys Gln Ser Arg Phe Arg Ile Ile

305 310 315 320

Ser Asp Asp His Asn Glu Val Leu Met Lys Arg Ser Thr Asp Arg Phe

325 330 335

Val Pro Leu Leu Leu Gln Tyr Ile Asp Tyr Gly Lys Leu Phe Asp His

340 345 350

Ile Arg Phe His Val Asn Met Gly Lys Leu Arg Tyr Leu Leu Lys Ala

355 360 365

Asp Lys Thr Cys Ile Asp Gly Gln Thr Arg Val Arg Val Ile Glu Gln

370 375 380

Pro Leu Asn Gly Phe Gly Arg Leu Glu Glu Ala Glu Thr Met Arg Lys

385 390 395 400

Gln Glu Asn Gly Thr Phe Gly Asn Ser Gly Ile Arg Ile Arg Asp Phe

405 410 415

Glu Asn Val Lys Arg Asp Asp Ala Asn Pro Ala Asn Tyr Pro Tyr Ile

420 425 430

Val Asp Thr Tyr Thr His Tyr Ile Leu Glu Asn Asn Lys Val Glu Met

435 440 445

Phe Ile Ser Asp Lys Gly Ser Ser Ala Pro Leu Leu Pro Leu Ile Glu

450 455 460

Asp Asp Arg Tyr Val Val Lys Thr Ile Pro Ser Cys Arg Met Ser Thr

465 470 475 480

Leu Glu Ile Pro Ala Met Ala Phe His Met Phe Leu Phe Gly Ser Lys

485 490 495

Lys Thr Glu Lys Leu Ile Val Asp Val His Asn Arg Tyr Lys Arg Leu

500 505 510

Phe Gln Ala Met Gln Lys Glu Glu Val Thr Ala Glu Asn Ile Ala Ser

515 520 525

Phe Gly Ile Ala Glu Ser Asp Leu Pro Gln Lys Ile Leu Asp Leu Ile

530 535 540

Ser Gly Asn Ala His Gly Lys Asp Val Asp Ala Phe Ile Arg Leu Thr
545 550 555 560
Val Asp Asp Met Leu Thr Asp Thr Glu Arg Arg Ile Lys Arg Phe Lys
565 570 575

Asp Asp Arg Lys Ser Ile Arg Ser Ala Asp Asn Lys Met Gly Lys Arg
580 585 590
Gly Phe Lys Gln Ile Ser Thr Gly Lys Leu Ala Asp Phe Leu Ala Lys
595 600 605
Asp Ile Val Leu Phe Gln Pro Ser Val Asn Asp Gly Glu Asn Lys Ile
610 615 620
Thr Gly Leu Asn Tyr Arg Ile Met Gln Ser Ala Ile Ala Val Tyr Asp
625 630 635 640

Ser Gly Asp Asp Tyr Glu Ala Lys Gln Gln Phe Lys Leu Met Phe Glu
645 650 655
Lys Ala Arg Leu Ile Gly Lys Gly Thr Thr Glu Pro His Pro Phe Leu
660 665 670
Tyr Lys Val Phe Ala Arg Ser Ile Pro Ala Asn Ala Val Asp Phe Tyr
675 680 685
Glu Arg Tyr Leu Ile Glu Arg Lys Phe Tyr Leu Thr Gly Leu Cys Asn
690 695 700

Glu Ile Lys Arg Gly Asn Arg Val Asp Val Pro Phe Ile Arg Arg Asp
705 710 715 720
Gln Asn Lys Trp Lys Thr Pro Ala Met Lys Thr Leu Gly Arg Ile Tyr
725 730 735
Ser Glu Asp Leu Pro Val Glu Leu Pro Arg Gln Met Phe Asp Asn Glu
740 745 750
Ile Lys Ser His Leu Lys Ser Leu Pro Gln Met Glu Gly Ile Asp Phe
755 760 765

Asn Asn Ala Asn Val Thr Tyr Leu Ile Ala Glu Tyr Met Lys Arg Val
770 775 780
Leu Asn Asp Asp Phe Gln Thr Phe Tyr Gln Trp Lys Arg Asn Tyr His

785 790 795 800

Tyr Met Asp Met Leu Lys Gly Glu Tyr Asp Arg Lys Gly Ser Leu Gln

805 810 815

His Cys Phe Thr Ser Val Glu Glu Arg Glu Gly Leu Trp Lys Glu Arg

820 825 830

Ala Ser Arg Thr Glu Arg Tyr Arg Lys Leu Ala Ser Asn Lys Ile Arg

835 840 845

Ser Asn Arg Gln Met Arg Asn Ala Ser Ser Glu Glu Ile Glu Thr Ile

850 855 860

Leu Asp Lys Arg Leu Ser Asn Cys Arg Asn Glu Tyr Gln Lys Ser Glu

865 870 875 880

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

885 890 895

Ala Lys Lys Thr Leu Thr Glu Leu Ala Asp Phe Asp Gly Glu Arg Phe

900 905 910

Lys Leu Lys Glu Ile Met Pro Asp Ala Glu Lys Gly Ile Leu Ser Glu

915 920 925

Ile Met Pro Met Ser Phe Thr Phe Glu Lys Gly Gly Lys Lys Tyr Thr

930 935 940

Ile Thr Ser Glu Gly Met Lys Leu Lys Asn Tyr Gly Asp Phe Phe Val

945 950 955 960

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

965 970 975

Asp Ile Val Ser Lys Glu Asp Ile Met Glu Glu Phe Asn Lys Tyr Asp

980 985 990

Gln Cys Arg Pro Glu Ile Ser Ser Ile Val Phe Asn Leu Glu Lys Trp

995 1000 1005

Ala Phe Asp Thr Tyr Pro Glu Leu Ser Ala Arg Val Asp Arg Glu

1010 1015 1020

Glu Lys Val Asp Phe Lys Ser Ile Leu Lys Ile Leu Leu Asn Asn

1025 1030 1035

Lys Asn Ile Asn Lys Glu Gln Ser Asp Ile Leu Arg Lys Ile Arg
1040 1045 1050

Asn Ala Phe Asp His Asn Asn Tyr Pro Asp Lys Gly Ile Val Glu
1055 1060 1065

Ile Lys Ala Leu Pro Glu Ile Ala Met Ser Ile Lys Lys Ala Phe
1070 1075 1080

Gly Glu Tyr Ala Ile Met Lys
1085 1090

<210> 138

<211> 1090

<212> PRT

<213> Prevotella species

<400> 138

Met Asn Ile Pro Ala Leu Val Glu Asn Gln Lys Lys Tyr Phe Gly Thr
1 5 10 15

Tyr Ser Val Met Ala Met Leu Asn Ala Gln Thr Val Leu Asp His Ile
20 25 30

Gln Lys Val Ala Asp Ile Glu Gly Glu Gln Asn Glu Asn Asn Glu Asn
35 40 45

Leu Trp Phe His Pro Val Met Ser His Leu Tyr Asn Ala Lys Asn Gly
50 55 60

Tyr Asp Lys Gln Pro Glu Lys Thr Met Phe Ile Ile Glu Arg Leu Gln
65 70 75 80

Ser Tyr Phe Pro Phe Leu Lys Ile Met Ala Glu Asn Gln Arg Glu Tyr
85 90 95

Ser Asn Gly Lys Tyr Lys Gln Asn Arg Val Glu Val Asn Ser Asn Asp
100 105 110

Ile Phe Glu Val Leu Lys Arg Ala Phe Gly Val Leu Lys Met Tyr Arg
115 120 125

Asp Gln Ala Ser His Tyr Lys Thr Tyr Asp Glu Lys Leu Ile Asp Gly
130 135 140

Cys Glu Phe Leu Thr Ser Thr Glu Gln Pro Leu Ser Gly Met Ile Asn

145 150 155 160

Asn Tyr Tyr Thr Val Ala Leu Arg Asn Met Asn Glu Arg Tyr Gly Tyr

165 170 175

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

180 185 190

Lys Asp Ala Tyr Gly Lys Lys Lys Ser Gln Val Asn Thr Gly Phe Phe

195 200 205

Leu Ser Leu Gln Asp Tyr Asn Gly Asp Thr Gln Lys Lys Leu His Leu

210 215 220

Ser Gly Val Gly Ile Ala Leu Leu Ile Cys Leu Phe Leu Asp Lys Gln

225 230 235 240

Tyr Ile Asn Ile Phe Leu Ser Arg Leu Pro Ile Phe Ser Ser Tyr Asn

245 250 255

Ala Gln Ser Glu Glu Arg Arg Ile Ile Ile Arg Ser Phe Gly Ile Asn

260 265 270

Ser Ile Lys Gln Pro Lys Asp Arg Ile His Ser Glu Lys Ser Asn Lys

275 280 285

Ser Val Ala Met Asp Met Leu Asn Glu Ile Lys Arg Cys Pro Asn Glu

290 295 300

Leu Phe Glu Thr Leu Ser Ala Glu Lys Gln Ser Arg Phe Arg Ile Ile

305 310 315 320

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

325 330 335

Val Pro Leu Leu Leu Gln Tyr Ile Asp Tyr Gly Lys Leu Phe Asp His

340 345 350

Ile Arg Phe His Val Asn Met Gly Lys Leu Arg Tyr Leu Leu Lys Ala

355 360 365

Asp Lys Thr Cys Ile Asp Gly Gln Thr Arg Val Arg Val Ile Glu Gln

370 375 380

Pro Leu Asn Gly Phe Gly Arg Leu Glu Glu Val Glu Thr Met Arg Lys

385 390 395 400

Gln Glu Asn Gly Thr Phe Gly Asn Ser Gly Ile Arg Ile Arg Asp Phe
405 410 415
Glu Asn Met Lys Arg Asp Asp Ala Asn Pro Ala Asn Tyr Pro Tyr Ile
420 425 430

Val Asp Thr Tyr Thr His Tyr Ile Leu Glu Asn Asn Lys Val Glu Met
435 440 445
Phe Ile Ser Asp Glu Glu Thr Pro Ala Pro Leu Leu Pro Val Ile Glu
450 455 460
Asp Asp Arg Tyr Val Val Lys Thr Ile Pro Ser Cys Arg Met Ser Thr
465 470 475 480
Leu Glu Ile Pro Ala Met Ala Phe His Met Phe Leu Phe Gly Ser Lys
485 490 495

Lys Thr Glu Lys Leu Ile Val Asp Val His Asn Arg Tyr Lys Arg Leu
500 505 510
Phe Lys Ala Met Gln Lys Glu Glu Val Thr Ala Glu Asn Ile Ala Ser
515 520 525
Phe Gly Ile Ala Glu Ser Asp Leu Pro Gln Lys Ile Ile Asp Leu Ile
530 535 540
Ser Gly Asn Ala His Gly Lys Asp Val Asp Ala Phe Ile Arg Leu Thr
545 550 555 560

Val Asp Asp Met Leu Ala Asp Thr Glu Arg Arg Ile Lys Arg Phe Lys
565 570 575
Asp Asp Arg Lys Ser Ile Arg Ser Ala Asp Asn Lys Met Gly Lys Arg
580 585 590
Gly Phe Lys Gln Ile Ser Thr Gly Lys Leu Ala Asp Phe Leu Ala Lys
595 600 605
Asp Ile Val Leu Phe Gln Pro Ser Val Asn Asp Gly Glu Asn Lys Ile
610 615 620

Thr Gly Leu Asn Tyr Arg Ile Met Gln Ser Ala Ile Ala Val Tyr Asn
625 630 635 640
Ser Gly Asp Asp Tyr Glu Ala Lys Gln Gln Phe Lys Leu Met Phe Glu

645 650 655

Lys Ala Arg Leu Ile Gly Lys Gly Thr Thr Glu Pro His Pro Phe Leu

660 665 670

Tyr Lys Val Phe Val Arg Ser Ile Pro Ala Asn Ala Val Asp Phe Tyr

675 680 685

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

690 695 700

Glu Ile Lys Lys Gly Asn Arg Val Asp Val Pro Phe Ile Arg Arg Asp

705 710 715 720

Gln Asn Lys Trp Lys Thr Pro Ala Met Lys Thr Leu Gly Arg Ile Tyr

725 730 735

Asp Glu Asp Leu Pro Val Glu Leu Pro Arg Gln Met Phe Asp Asn Glu

740 745 750

Ile Lys Ser His Leu Lys Ser Leu Pro Gln Met Glu Gly Ile Asp Phe

755 760 765

Asn Asn Ala Asn Val Thr Tyr Leu Ile Ala Glu Tyr Met Lys Arg Val

770 775 780

Leu Asn Asp Asp Phe Gln Thr Phe Tyr Gln Trp Lys Arg Asn Tyr Arg

785 790 795 800

Tyr Met Asp Met Leu Arg Gly Glu Tyr Asp Arg Lys Gly Ser Leu Gln

805 810 815

Ser Cys Phe Thr Ser Val Glu Glu Arg Glu Gly Leu Trp Lys Glu Arg

820 825 830

Ala Ser Arg Thr Glu Arg Tyr Arg Lys Leu Ala Ser Asn Lys Ile Arg

835 840 845

Ser Asn Arg Gln Met Arg Asn Ala Ser Ser Glu Glu Ile Glu Thr Ile

850 855 860

Leu Asp Lys Arg Leu Ser Asn Ser Arg Asn Glu Tyr Gln Lys Ser Glu

865 870 875 880

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

885 890 895

Ala Lys Lys Thr Leu Thr Glu Leu Ala Asp Phe Asp Gly Glu Arg Phe
900 905 910
Lys Leu Lys Glu Ile Met Pro Asp Ala Glu Lys Gly Ile Leu Ser Glu
915 920 925
Ile Met Pro Met Ser Phe Thr Phe Glu Lys Gly Gly Lys Lys Tyr Thr
930 935 940

Ile Thr Ser Glu Gly Met Lys Leu Lys Asn Tyr Gly Asp Phe Phe Val
945 950 955 960
Leu Ala Ser Asp Lys Arg Ile Gly Asn Leu Leu Glu Leu Val Gly Ser
965 970 975
Asp Thr Val Ser Lys Glu Asp Ile Met Glu Glu Phe Lys Lys Tyr Asp
980 985 990
Gln Cys Arg Pro Glu Ile Ser Ser Ile Val Phe Asn Leu Glu Lys Trp
995 1000 1005

Ala Phe Asp Thr Tyr Pro Glu Leu Ser Ala Arg Val Asp Arg Glu
1010 1015 1020
Glu Lys Val Asp Phe Lys Ser Ile Leu Lys Ile Leu Leu Asn Asn
1025 1030 1035
Lys Asn Ile Asn Lys Glu Gln Ser Asp Ile Leu Arg Lys Ile Arg
1040 1045 1050
Asn Ala Phe Asp His Asn Asn Tyr Pro Asp Lys Gly Val Val Glu
1055 1060 1065

Ile Arg Ala Leu Pro Glu Ile Ala Met Ser Ile Lys Lys Ala Phe
1070 1075 1080
Gly Glu Tyr Ala Ile Met Lys
1085 1090

<210> 139
<211> 1090
<212> PRT
<213> Prevotella species
<400> 139
Met Asn Ile Pro Ala Leu Val Glu Asn Gln Lys Lys Tyr Phe Gly Thr

1 5 10 15

Tyr Ser Val Met Ala Met Leu Asn Ala Gln Thr Val Leu Asp His Ile

20 25 30

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

35 40 45

Leu Trp Phe His Pro Val Met Ser His Leu Tyr Asn Ala Lys Asn Gly

50 55 60

Tyr Asp Lys Gln Pro Glu Lys Thr Met Phe Ile Ile Glu Arg Leu Gln

65 70 75 80

Ser Tyr Phe Pro Phe Leu Lys Ile Met Ala Glu Asn Gln Arg Glu Tyr

85 90 95

Ser Asn Gly Lys Tyr Lys Gln Asn Arg Val Glu Val Asn Ser Asn Asp

100 105 110

Ile Phe Glu Val Leu Lys Arg Ala Phe Gly Val Leu Lys Met Tyr Arg

115 120 125

Asp Leu Thr Asn His Tyr Lys Thr Tyr Glu Glu Lys Leu Ile Asp Gly

130 135 140

Cys Glu Phe Leu Thr Ser Thr Glu Gln Pro Phe Ser Gly Met Ile Ser

145 150 155 160

Lys Tyr Tyr Thr Val Ala Leu Arg Asn Thr Lys Glu Arg Tyr Gly Tyr

165 170 175

Lys Ala Glu Asp Leu Ala Phe Ile Gln Asp Asn Arg Tyr Lys Phe Thr

180 185 190

Lys Asp Ala Tyr Gly Lys Arg Lys Ser Gln Val Asn Thr Gly Ser Phe

195 200 205

Leu Ser Leu Gln Asp Tyr Asn Gly Asp Thr Thr Lys Lys Leu His Leu

210 215 220

Ser Gly Val Gly Ile Ala Leu Leu Ile Cys Leu Phe Leu Asp Lys Gln

225 230 235 240

Tyr Ile Asn Leu Phe Leu Ser Arg Leu Pro Ile Phe Ser Ser Tyr Asn

245 250 255

Ala Gln Ser Glu Glu Arg Arg Ile Ile Ile Arg Ser Phe Gly Ile Asn
260 265 270
Ser Ile Lys Gln Pro Lys Asp Arg Ile His Ser Glu Lys Ser Asn Lys
275 280 285

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

Ile Arg Phe His Val Asn Met Gly Lys Leu Arg Tyr Leu Leu Lys Ala
355 360 365
Asp Lys Thr Cys Ile Asp Gly Gln Thr Arg Val Arg Val Ile Glu Gln
370 375 380
Pro Leu Asn Gly Phe Gly Arg Leu Glu Glu Val Glu Thr Met Arg Lys
385 390 395 400
Gln Glu Asn Gly Thr Phe Gly Asn Ser Gly Ile Arg Ile Arg Asp Phe
405 410 415

Glu Asn Met Lys Arg Asp Asp Ala Asn Pro Ala Asn Tyr Pro Tyr Ile
420 425 430
Val Glu Thr Tyr Thr His Tyr Ile Leu Glu Asn Asn Lys Val Glu Met
435 440 445
Phe Ile Ser Asp Glu Glu Asn Pro Thr Pro Leu Leu Pro Val Ile Glu
450 455 460
Asp Asp Arg Tyr Val Val Lys Thr Ile Pro Ser Cys Arg Met Ser Thr
465 470 475 480

Leu Glu Ile Pro Ala Met Ala Phe His Met Phe Leu Phe Gly Ser Glu
485 490 495
Lys Thr Glu Lys Leu Ile Ile Asp Val His Asp Arg Tyr Lys Arg Leu

500 505 510

Phe Gln Ala Met Gln Lys Glu Glu Val Thr Ala Glu Asn Ile Ala Ser

515 520 525

Phe Gly Ile Ala Glu Ser Asp Leu Pro Gln Lys Ile Met Asp Leu Ile

530 535 540

Ser Gly Asn Ala His Gly Lys Asp Val Asp Ala Phe Ile Arg Leu Thr

545 550 555 560

Val Asp Asp Met Leu Thr Asp Thr Glu Arg Arg Ile Lys Arg Phe Lys

565 570 575

Asp Asp Arg Lys Ser Ile Arg Ser Ala Asp Asn Lys Met Gly Lys Arg

580 585 590

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

595 600 605

Asp Ile Val Leu Phe Gln Pro Ser Val Asn Asp Gly Glu Asn Lys Ile

610 615 620

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

625 630 635 640

Ser Gly Asp Asp Tyr Glu Ala Lys Gln Gln Phe Lys Leu Met Phe Glu

645 650 655

Lys Ala Arg Leu Ile Gly Lys Gly Thr Thr Glu Pro His Pro Phe Leu

660 665 670

Tyr Lys Val Phe Val Arg Ser Ile Pro Ala Asn Ala Val Asp Phe Tyr

675 680 685

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

690 695 700

Glu Ile Lys Lys Gly Asn Arg Val Asp Val Pro Phe Ile Arg Arg Asp

705 710 715 720

Gln Asn Lys Trp Lys Thr Pro Ala Met Lys Thr Leu Gly Arg Ile Tyr

725 730 735

Ser Glu Asp Leu Pro Val Glu Leu Pro Arg Gln Met Phe Asp Asn Glu

740 745 750

Ile Lys Ser His Leu Lys Ser Leu Pro Gln Met Glu Gly Ile Asp Phe
755 760 765

Asn Asn Ala Asn Val Thr Tyr Leu Ile Ala Glu Tyr Met Lys Arg Val
770 775 780

Leu Asn Asp Asp Phe Gln Thr Phe Tyr Gln Trp Lys Arg Asn Tyr Arg
785 790 795 800

Tyr Met Asp Met Leu Arg Gly Glu Tyr Asp Arg Lys Gly Ser Leu Gln
805 810 815

His Cys Phe Thr Ser Ile Glu Glu Arg Glu Gly Leu Trp Lys Glu Arg
820 825 830

Ala Ser Arg Thr Glu Arg Tyr Arg Lys Leu Ala Ser Asn Lys Ile Arg
835 840 845

Ser Asn Arg Gln Met Arg Asn Ala Ser Ser Glu Glu Ile Glu Thr Ile
850 855 860

Leu Asp Lys Arg Leu Ser Asn Cys Arg Asn Glu Tyr Gln Lys Ser Glu
865 870 875 880

Lys Ile Ile Arg Arg Tyr Arg Val Gln Asp Ala Leu Leu Phe Leu Leu
885 890 895

Ala Lys Lys Thr Leu Thr Glu Leu Ala Asp Phe Asp Gly Glu Arg Phe
900 905 910

Lys Leu Lys Glu Ile Met Pro Asp Ala Glu Lys Gly Ile Leu Ser Glu
915 920 925

Ile Met Pro Met Ser Phe Thr Phe Glu Lys Gly Gly Lys Ile Tyr Thr
930 935 940

Ile Thr Ser Gly Gly Met Lys Leu Lys Asn Tyr Gly Asp Phe Phe Val
945 950 955 960

Leu Ala Ser Asp Lys Arg Ile Gly Asn Leu Leu Glu Leu Val Gly Ser
965 970 975

Asn Thr Val Ser Lys Glu Asp Ile Met Glu Glu Phe Lys Lys Tyr Asp
980 985 990

Gln Cys Arg Pro Glu Ile Ser Ser Ile Val Phe Asn Leu Glu Lys Trp

995 1000 1005

Ala Phe Asp Thr Tyr Pro Glu Leu Pro Ala Arg Val Asp Arg Lys

1010 1015 1020

Glu Lys Val Asp Phe Trp Ser Ile Leu Asp Val Leu Ser Asn Asn

1025 1030 1035

Lys Asp Ile Asn Asn Glu Gln Ser Tyr Ile Leu Arg Lys Ile Arg

1040 1045 1050

Asn Ala Phe Asp His Asn Asn Tyr Pro Asp Lys Gly Ile Val Glu

1055 1060 1065

Ile Lys Ala Leu Pro Glu Ile Ala Met Ser Ile Lys Lys Ala Phe

1070 1075 1080

Gly Glu Tyr Ala Ile Met Lys

1085 1090

<210> 140

<211> 1131

<212> PRT

<213> Phaeodactylibacter xiamenensis

<400> 140

Met Thr Asn Thr Pro Lys Arg Arg Thr Leu His Arg His Pro Ser Tyr

1 5 10 15

Phe Gly Ala Phe Leu Asn Ile Ala Arg His Asn Ala Phe Met Ile Met

20 25 30

Glu His Leu Ser Thr Lys Tyr Asp Met Glu Asp Lys Asn Thr Leu Asp

35 40 45

Glu Ala Gln Leu Pro Asn Ala Lys Leu Phe Gly Cys Leu Lys Lys Arg

50 55 60

Tyr Gly Lys Pro Asp Val Thr Glu Gly Val Ser Arg Asp Leu Arg Arg

65 70 75 80

Tyr Phe Pro Phe Leu Asn Tyr Pro Leu Phe Leu His Leu Glu Lys Gln

85 90 95

Gln Asn Ala Glu Gln Ala Ala Thr Tyr Asp Ile Asn Pro Glu Asp Ile

100 105 110

Glu Phe Thr Leu Lys Gly Phe Phe Arg Leu Leu Asn Gln Met Arg Asn
115 120 125
Asn Tyr Ser His Tyr Ile Ser Asn Thr Asp Tyr Gly Lys Phe Asp Lys
130 135 140

Leu Pro Val Gln Asp Ile Tyr Glu Ala Ala Ile Phe Arg Leu Leu Asp
145 150 155 160
Arg Gly Lys His Thr Lys Arg Phe Asp Val Phe Glu Ser Lys His Thr
165 170 175
Arg His Leu Glu Ser Asn Asn Ser Glu Tyr Arg Pro Arg Ser Leu Ala
180 185 190
Asn Ser Pro Asp His Glu Asn Thr Val Ala Phe Val Thr Cys Leu Phe
195 200 205

Leu Glu Arg Lys Tyr Ala Phe Pro Phe Leu Ser Arg Leu Asp Cys Phe
210 215 220
Arg Ser Thr Asn Asp Ala Ala Glu Gly Asp Pro Leu Ile Arg Lys Ala
225 230 235 240
Ser His Glu Cys Tyr Thr Met Phe Cys Cys Arg Leu Pro Gln Pro Lys
245 250 255
Leu Glu Ser Ser Asp Ile Leu Leu Asp Met Val Asn Glu Leu Gly Arg
260 265 270

Cys Pro Ser Ala Leu Tyr Asn Leu Leu Ser Glu Glu Asp Gln Ala Arg
275 280 285
Phe His Ile Lys Arg Glu Glu Ile Thr Gly Phe Glu Glu Asp Pro Asp
290 295 300
Glu Glu Leu Glu Gln Glu Ile Val Leu Lys Arg His Ser Asp Arg Phe
305 310 315 320
Pro Tyr Phe Ala Leu Arg Tyr Phe Asp Asp Thr Glu Ala Phe Gln Thr
325 330 335

Leu Arg Phe Asp Val Tyr Leu Gly Arg Trp Arg Thr Lys Pro Val Tyr
340 345 350
Lys Lys Arg Ile Tyr Gly Gln Glu Arg Asp Arg Val Leu Thr Gln Ser

355 360 365

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

370 375 380

Val Lys His Asp Ala Val Arg Gln Asn Glu Glu Asp Gly Lys Leu Val

385 390 395 400

Asn Pro Asp Val Thr Ser Gln Phe His Lys Ser Trp Ile Gln Ile Glu

405 410 415

Ser Asp Asp Arg Ala Phe Leu Ser Asp Arg Ile Glu His Phe Ser Pro

420 425 430

His Tyr Asn Phe Gly Asp Gln Val Ile Gly Leu Lys Phe Ile Asn Pro

435 440 445

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

450 455 460

Glu Lys Lys Asp Lys Asp Ala Lys Leu Val Asn Glu Thr Ala Asp Ala

465 470 475 480

Ile Ile Ser Thr His Glu Ile Arg Ser Leu Phe Leu Tyr His Tyr Leu

485 490 495

Ser Lys Lys Pro Ile Ser Ala Gly Asp Glu Arg Arg Phe Ile Gln Val

500 505 510

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

515 520 525

Phe Glu Asp Ile Lys Ser Gly Glu Leu Gln Pro Ile Ala Asp Pro Pro

530 535 540

Asn Tyr Gln Lys Asn Glu Pro Leu Pro Tyr Val Arg Gly Asp Lys Glu

545 550 555 560

Lys Thr Gln Glu Glu Arg Ala Gln Tyr Arg Glu Arg Gln Lys Glu Ile

565 570 575

Lys Glu Arg Arg Lys Glu Leu Asn Thr Leu Leu Gln Asn Arg Tyr Gly

580 585 590

Leu Ser Ile Gln Tyr Ile Pro Ser Arg Leu Arg Glu Tyr Leu Leu Gly

595 600 605

Tyr Lys Lys Val Pro Tyr Glu Lys Leu Ala Leu Gln Lys Leu Arg Ala
610 615 620
Gln Arg Lys Glu Val Lys Lys Arg Ile Lys Asp Ile Glu Lys Met Arg
625 630 635 640
Thr Pro Arg Val Gly Glu Gln Ala Thr Trp Leu Ala Glu Asp Ile Val
645 650 655

Phe Leu Thr Pro Pro Lys Met His Thr Pro Glu Arg Lys Thr Thr Lys
660 665 670
His Pro Gln Lys Leu Asn Asn Asp Gln Phe Arg Ile Met Gln Ser Ser
675 680 685
Leu Ala Tyr Phe Ser Val Asn Lys Lys Ala Ile Lys Lys Phe Phe Gln
690 695 700
Lys Glu Thr Gly Ile Gly Leu Ser Asn Arg Glu Thr Ser His Pro Phe
705 710 715 720

Leu Tyr Arg Ile Asp Val Gly Arg Cys Arg Gly Ile Leu Asp Phe Tyr
725 730 735
Thr Gly Tyr Leu Lys Tyr Lys Met Asp Trp Leu Asp Asp Ala Ile Lys
740 745 750
Lys Val Asp Asn Arg Lys His Gly Lys Lys Glu Ala Lys Lys Tyr Glu
755 760 765
Lys Tyr Leu Pro Ser Ser Ile Gln His Lys Thr Pro Leu Glu Leu Asp
770 775 780

Tyr Thr Arg Leu Pro Val Tyr Leu Pro Arg Gly Leu Phe Lys Lys Ala
785 790 795 800
Ile Val Lys Ala Leu Ala Ala His Ala Asp Phe Gln Val Glu Pro Glu
805 810 815
Glu Asp Asn Val Ile Phe Cys Leu Asp Gln Leu Leu Asp Gly Asp Thr
820 825 830
Gln Asp Phe Tyr Asn Trp Gln Arg Tyr Tyr Arg Ser Ala Leu Thr Glu
835 840 845

Lys Glu Thr Asp Asn Gln Leu Val Leu Ala His Pro Tyr Ala Glu Gln

850 855 860

Ile Leu Gly Thr Ile Lys Thr Leu Glu Gly Lys Gln Lys Asn Asn Lys

865 870 875 880

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

885 890 895

Lys Arg Ala Lys Arg Arg Leu Leu Asp Arg Glu Gln Tyr Leu Arg Ala

900 905 910

Val Gln Ala Glu Asp Arg Ala Leu Trp Leu Met Ile Gln Glu Arg Gln

915 920 925

Lys Gln Lys Ala Glu His Glu Glu Ile Ala Phe Asp Gln Leu Asp Leu

930 935 940

Lys Asn Ile Thr Lys Ile Leu Thr Glu Ser Ile Asp Ala Arg Leu Arg

945 950 955 960

Ile Pro Asp Thr Lys Val Asp Ile Thr Asp Lys Leu Pro Leu Arg Arg

965 970 975

Tyr Gly Asp Leu Arg Arg Val Ala Lys Asp Arg Arg Leu Val Asn Leu

980 985 990

Ala Ser Tyr Tyr His Val Ala Gly Leu Ser Glu Ile Pro Tyr Asp Leu

995 1000 1005

Val Lys Lys Glu Leu Glu Glu Tyr Asp Arg Arg Arg Val Ala Phe

1010 1015 1020

Phe Glu His Val Tyr Gln Phe Glu Lys Glu Val Tyr Asp Arg Tyr

1025 1030 1035

Ala Ala Glu Leu Arg Asn Glu Asn Pro Lys Gly Glu Ser Thr Tyr

1040 1045 1050

Phe Ser His Trp Glu Tyr Val Ala Val Ala Val Lys His Ser Ala

1055 1060 1065

Asp Thr His Phe Asn Glu Leu Phe Lys Glu Lys Val Met Gln Leu

1070 1075 1080

Arg Asn Lys Phe His His Asn Glu Phe Pro Tyr Phe Asp Trp Leu

1085 1090 1095

Leu Pro Glu Val Glu Lys Ala Ser Ala Ala Leu Tyr Ala Asp Arg

1100 1105 1110

Val Phe Asp Val Ala Glu Gly Tyr Tyr Gln Lys Met Arg Lys Leu

1115 1120 1125

Met Arg Gln

1130

<210> 141

<211> 1155

<212> PRT

<213> Flavobacterium species

<400> 141

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

1 5 10 15

Thr Tyr Asn His Asp Lys Val Glu Asp Lys His Tyr Phe Gly Gly Phe

20 25 30

Phe Asn Leu Ala Gln Asn Asn Ile Asp Leu Val Ala Gln Glu Phe Lys

35 40 45

Lys Arg Leu Leu Ile Gln Gly Lys Asp Ser Ile Asn Ile Phe Ala Asn

50 55 60

Tyr Phe Ser Asp Gln Cys Ser Ile Thr Asn Leu Glu Arg Gly Ile Lys

65 70 75 80

Ile Leu Ala Glu Tyr Phe Pro Val Val Ser Tyr Ile Asp Leu Asp Glu

85 90 95

Lys Asn Lys Ser Lys Ser Ile Arg Glu His Leu Ile Leu Leu Leu Glu

100 105 110

Thr Ile Asn Asn Leu Arg Asn Tyr Tyr Thr His Tyr Tyr His Lys Lys

115 120 125

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

130 135 140

Lys Val Val Leu Glu Ile Lys Lys Lys Lys Leu Lys Glu Asp Lys Thr

145 150 155 160

Lys Gln Leu Leu Lys Lys Gly Leu Glu Lys Glu Met Thr Ile Leu Phe

165 170 175

Asn Leu Met Lys Ala Glu Gln Lys Glu Lys Lys Ile Lys Gly Trp Asn

180 185 190

Ile Asp Glu Asn Ile Lys Gly Ala Val Leu Asn Arg Ala Phe Ser His

195 200 205

Leu Leu Tyr Asn Asp Glu Leu Ser Asp Tyr Arg Lys Ser Lys Tyr Asn

210 215 220

Thr Glu Asp Glu Thr Leu Lys Asp Thr Leu Thr Glu Ser Gly Ile Leu

225 230 235 240

Phe Leu Leu Ser Phe Phe Leu Asn Lys Lys Glu Gln Glu Gln Leu Lys

245 250 255

Ala Asn Ile Lys Gly Tyr Lys Gly Lys Ile Ala Ser Ile Pro Asp Glu

260 265 270

Glu Ile Thr Leu Lys Asn Asn Ser Leu Arg Asn Met Ala Thr His Trp

275 280 285

Thr Tyr Ser His Leu Thr Tyr Lys Gly Leu Lys His Arg Ile Lys Thr

290 295 300

Asp His Glu Lys Glu Thr Leu Leu Val Asn Met Val Asp Tyr Leu Ser

305 310 315 320

Lys Val Pro His Glu Ile Tyr Gln Asn Leu Ser Glu Gln Asn Lys Ser

325 330 335

Leu Phe Leu Glu Asp Ile Asn Glu Tyr Met Arg Asp Asn Glu Glu Asn

340 345 350

His Asp Ser Ser Glu Ala Ser Arg Val Ile His Pro Val Ile Arg Lys

355 360 365

Arg Tyr Glu Asn Lys Phe Ala Tyr Phe Ala Ile Arg Phe Leu Asp Glu

370 375 380

Phe Ala Glu Phe Pro Thr Leu Arg Phe Met Val Asn Val Gly Asn Tyr

385 390 395 400

Ile His Asp Asn Arg Lys Lys Asp Ile Gly Gly Thr Ser Leu Ile Thr

405 410 415

Asn Arg Thr Ile Lys Gln Gln Ile Asn Val Phe Gly Asn Leu Thr Glu
420 425 430
Ile His Lys Lys Lys Asn Asp Tyr Phe Glu Lys Glu Glu Asn Lys Glu
435 440 445
Lys Thr Leu Glu Trp Glu Leu Phe Pro Asn Pro Ser Tyr His Phe Gln
450 455 460
Lys Glu Asn Ile Pro Ile Phe Ile Asp Leu Glu Lys Ser Lys Glu Thr

465 470 475 480
Asn Asp Leu Ala Lys Glu Tyr Ala Lys Glu Lys Lys Lys Ile Phe Gly
485 490 495
Ser Ser Arg Lys Lys Gln Gln Asn Thr Ala Lys Lys Asn Arg Glu Thr
500 505 510
Ile Ile Asn Leu Val Phe Asp Lys Tyr Lys Thr Ser Asp Arg Lys Thr
515 520 525
Val Thr Phe Glu Gln Pro Thr Ala Leu Leu Ser Phe Asn Glu Leu Asn

530 535 540
Ser Phe Leu Tyr Ala Phe Leu Val Glu Asn Lys Thr Gly Lys Glu Leu
545 550 555 560
Glu Lys Ile Ile Ile Glu Lys Ile Ala Asn Gln Tyr Gln Ile Leu Lys
565 570 575
Asn Cys Ser Ser Thr Val Asp Lys Thr Asn Asp Asn Ile Pro Lys Ser
580 585 590
Ile Lys Lys Ile Val Asn Thr Thr Thr Asp Ser Phe Tyr Phe Glu Gly

595 600 605
Lys Lys Ile Asp Ile Glu Lys Leu Glu Lys Asp Ile Thr Ile Glu Ile
610 615 620
Glu Lys Thr Asn Glu Lys Leu Glu Thr Ile Lys Glu Asn Glu Glu Ser
625 630 635 640
Ala Gln Asn Tyr Lys Arg Asn Glu Arg Asn Thr Gln Lys Arg Lys Leu
645 650 655
Tyr Arg Lys Tyr Val Phe Phe Thr Asn Glu Ile Gly Ile Glu Ala Thr

660 665 670

Trp Ile Thr Asn Asp Ile Leu Arg Phe Leu Asp Asn Lys Glu Asn Trp

675 680 685

Lys Gly Tyr Gln His Ser Glu Leu Gln Lys Phe Ile Ser Gln Tyr Asp

690 695 700

Asn Tyr Lys Lys Glu Ala Leu Gly Leu Leu Glu Ser Glu Trp Asn Leu

705 710 715 720

Glu Ser Asp Ala Phe Phe Gly Gln Asn Leu Lys Arg Met Phe Gln Ser

725 730 735

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

740 745 750

Asn Thr Leu Glu Thr Tyr Leu Ser Ala Ile Glu Asn Leu Lys Thr Met

755 760 765

Thr Asp Val Arg Pro Lys Val Leu Lys Lys Lys Trp Thr Glu Leu Phe

770 775 780

Arg Phe Phe Asp Lys Lys Ile Tyr Leu Leu Ser Thr Ile Glu Thr Lys

785 790 795 800

Ile Asn Glu Leu Ile Thr Lys Pro Ile Asn Leu Ser Arg Gly Ile Phe

805 810 815

Glu Glu Lys Pro Thr Phe Ile Asn Gly Lys Asn Pro Asn Lys Glu Asn

820 825 830

Asn Gln His Leu Phe Ala Asn Trp Phe Ile Tyr Ala Lys Lys Gln Thr

835 840 845

Ile Leu Gln Asp Phe Tyr Asn Leu Pro Leu Glu Gln Pro Lys Ala Ile

850 855 860

Thr Asn Leu Lys Lys His Lys Tyr Lys Leu Glu Arg Ser Ile Asn Asn

865 870 875 880

Leu Lys Ile Glu Asp Ile Tyr Ile Lys Gln Met Val Asp Phe Leu Tyr

885 890 895

Gln Lys Leu Phe Glu Gln Ser Phe Ile Gly Ser Leu Gln Asp Leu Tyr

900 905 910

Thr Ser Lys Glu Lys Arg Glu Ile Glu Lys Gly Lys Ala Lys Asn Glu

915 920 925

Gln Thr Pro Asp Glu Ser Phe Ile Trp Lys Lys Gln Val Glu Ile Asn

930 935 940

Thr His Asn Gly Arg Ile Ile Ala Lys Thr Lys Ile Lys Asp Ile Gly

945 950 955 960

Lys Phe Lys Asn Leu Leu Thr Asp Asn Lys Ile Ala His Leu Ile Ser

965 970 975

Tyr Asp Asp Arg Ile Trp Asp Phe Ser Leu Asn Asn Asp Gly Asp Ile

980 985 990

Thr Lys Lys Leu Tyr Ser Ile Asn Thr Glu Leu Glu Ser Tyr Glu Thr

995 1000 1005

Ile Arg Arg Glu Lys Leu Leu Lys Gln Ile Gln Gln Phe Glu Gln

1010 1015 1020

Phe Leu Leu Glu Gln Glu Thr Glu Tyr Ser Ala Glu Arg Lys His

1025 1030 1035

Pro Glu Lys Phe Glu Lys Asp Cys Asn Pro Asn Phe Lys Lys Tyr

1040 1045 1050

Ile Ile Glu Gly Val Leu Asn Lys Ile Ile Pro Asn His Glu Ile

1055 1060 1065

Glu Glu Ile Glu Ile Leu Lys Ser Lys Glu Asp Val Phe Lys Ile

1070 1075 1080

Asn Phe Ser Asp Ile Leu Ile Leu Asn Asn Asp Asn Ile Lys Lys

1085 1090 1095

Gly Tyr Leu Leu Ile Met Ile Arg Asn Lys Phe Ala His Asn Gln

1100 1105 1110

Leu Ile Asp Lys Asn Leu Phe Asn Phe Ser Leu Gln Leu Tyr Ser

1115 1120 1125

Lys Asn Glu Asn Glu Asn Phe Ser Glu Tyr Leu Asn Lys Val Cys

1130 1135 1140

Gln Asn Ile Ile Gln Glu Phe Lys Glu Lys Leu Lys

1145 1150 1155

<210> 142

<211> 1175

<212> PRT

<213> Porphyromonas gulae

<400> 142

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

Lys Asn Phe Asp Asn Asp Leu Glu Arg Lys Ser Arg Leu Arg Ser Leu
65 70 75 80
Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95
Leu Phe Glu Ser Lys Ser Ser Gly Asn Lys Ser Ser Lys Asn Lys Glu
100 105 110
Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140
Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Ser Glu Ser Ser Glu Leu
145 150 155 160
Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175
Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Tyr Gly His Asn Asp Asn Pro Ser Phe Lys His His Phe Val Asp

210 215 220

Ser Glu Gly Met Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Pro Leu Tyr Asp Arg Leu Arg Glu Lys Asp Arg Ala Arg Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Pro Asp Glu Asp Asp Thr Asp Gly Gly

325 330 335

Gly Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Met Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Ser Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Met Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Thr Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp
465 470 475 480
Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro
485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val
500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525
Tyr Ala Ile Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Leu Lys
530 535 540
Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Lys Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asp Met Glu
565 570 575

Glu Lys Ile Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590
Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655
Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Lys Arg Leu Thr Pro
660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

Arg Ser Tyr Leu Arg Ala Arg Lys Ala Phe Leu Glu Arg Ile Gly Arg
705 710 715 720
Ser Asp Arg Met Glu Asn Arg Pro Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu
740 745 750
Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Arg Glu Val Gly Phe Met Ala Lys
770 775 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Glu Asp Arg Val Gln Pro
785 790 795 800
Phe Tyr Asp Ser Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys
805 810 815
Gly Arg Phe Leu Ser Lys Glu Glu Arg Ala Glu Glu Trp Glu Arg Gly
820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg
835 840 845
Arg Ile Glu Asp Ala Phe Ala Gly Ile Glu Tyr Ala Ser Pro Gly Asn
850 855 860
Lys Lys Lys Ile Glu Gln Leu Leu Arg Asp Leu Ser Leu Trp Glu Ala
865 870 875 880
Phe Glu Ser Lys Leu Lys Val Arg Ala Asp Lys Ile Asn Leu Ala Lys
885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Gln Glu His Pro Tyr His Asp Phe
900 905 910
Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln
915 920 925
Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys
930 935 940

Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Pro
945 950 955 960

Asn Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn Arg Val Lys Pro
965 970 975
Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val
980 985 990
His Lys Glu Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp
995 1000 1005
Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp
1010 1015 1020

Arg Arg Leu Asn Gly Leu Phe Ser Phe Val Asp Thr Gly Gly Leu
1025 1030 1035
Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu
1040 1045 1050
Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Val Phe Glu Leu Thr
1055 1060 1065
Leu Arg Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro
1070 1075 1080

Asp Glu Ser Phe Arg Lys Met Leu Glu Ser Trp Ser Asp Pro Leu
1085 1090 1095
Leu Ala Lys Trp Pro Glu Leu His Gly Lys Val Arg Leu Leu Ile
1100 1105 1110
Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp
1115 1120 1125
Glu Ala Val Phe Ser Ser Ile Arg Lys Tyr Asp Pro Ser Ser Pro
1130 1135 1140

Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu
1145 1150 1155
Ser Glu Glu Val Lys Gln Ala Lys Glu Thr Val Glu Arg Ile Ile
1160 1165 1170
Gln Val

1175

<210> 143

<211> 1145

<212> PRT

<213> Chryseobacterium species

<400> 143

Met Glu Thr Gln Thr Ile Gly His Gly Ile Ala Tyr Asp His Ser Lys
1 5 10 15

Ile Gln Asp Lys His Phe Phe Gly Gly Phe Leu Asn Leu Ala Glu Asn

20 25 30

Asn Ile Lys Ala Val Leu Lys Ala Phe Ser Glu Lys Phe Asn Val Gly
35 40 45

Asn Val Asp Val Lys Gln Phe Ala Asp Val Ser Leu Lys Asp Asn Leu
50 55 60

Pro Asp Asn Asp Phe Gln Lys Arg Val Ser Phe Leu Lys Met Tyr Phe
65 70 75 80

Pro Val Val Asp Phe Ile Asn Ile Pro Asn Asn Arg Ala Lys Phe Arg

85 90 95

Ser Asp Leu Thr Thr Leu Phe Lys Ser Val Asp Gln Leu Arg Asn Phe
100 105 110

Tyr Thr His Tyr Tyr His Lys Pro Leu Asp Phe Asp Ala Ser Leu Phe
115 120 125

Ile Leu Leu Asp Asp Ile Phe Ala Arg Thr Ala Lys Glu Val Arg Asp
130 135 140

Gln Lys Met Lys Asp Asp Lys Thr Arg Gln Leu Leu Ser Lys Ser Leu

145 150 155 160

Ser Glu Glu Leu Gln Lys Gly Tyr Glu Leu Gln Leu Glu Arg Leu Lys
165 170 175

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

Ile Lys Asn Gly Val Leu Asn Asn Ala Phe Asn His Leu Ile Tyr Lys
195 200 205

Asp Gly Glu Ser Phe Lys Thr Lys Leu Thr Tyr Ser Ser Ala Leu Thr

210 215 220

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

275 280 285

Tyr Leu Cys Phe Lys Gly Leu Lys Ser Lys Leu Ser Thr Glu Phe His
290 295 300
Glu Glu Thr Leu Leu Ile Gln Ile Ile Asp Glu Leu Ser Lys Val Pro
305 310 315 320
Asp Glu Leu Tyr Cys Ala Phe Asp Lys Glu Thr Arg Asp Lys Phe Ile
325 330 335
Glu Asp Ile Asn Glu Tyr Val Lys Glu Gly His Gln Asp Phe Ser Leu

340 345 350

Glu Asp Ala Lys Val Ile His Pro Val Ile Arg Lys Arg Tyr Glu Asn
355 360 365
Lys Phe Asn Tyr Phe Ala Ile Arg Phe Leu Asp Glu Phe Val Lys Phe
370 375 380
Pro Ser Leu Arg Phe Gln Val His Val Gly Asn Tyr Val His Asp Arg
385 390 395 400
Arg Ile Lys Asn Ile Asp Gly Thr Thr Phe Glu Thr Glu Arg Val Val

405 410 415

Lys Asp Arg Ile Lys Val Phe Gly Arg Leu Ser Glu Ile Ser Ser Tyr
420 425 430
Lys Ala Gln Tyr Leu Ser Ser Val Ser Asp Lys His Asp Glu Thr Gly
435 440 445
Trp Glu Ile Phe Pro Asn Pro Ser Tyr Val Phe Ile Asn Asn Asn Ile

450 455 460

Pro Ile His Ile Ser Val Asp Thr Ser Phe Lys Lys Glu Ile Ala Asp

465 470 475 480

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

485 490 495

Gly Ala Glu Lys Lys Arg Lys Phe Glu Ile Thr Gln Met Ile Gly Ser

500 505 510

Lys Ser Val Leu Asn Gln Glu Glu Pro Ile Ala Leu Leu Ser Leu Asn

515 520 525

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

530 535 540

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

545 550 555 560

Val Ile Lys Asn Tyr Asn Pro Glu Asn Trp Leu Pro Ala Ser Gln Ile

565 570 575

Ser Arg Arg Leu Arg Ser Asn Lys Gly Glu Arg Ile Ile Asn Thr Asp

580 585 590

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

595 600 605

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

610 615 620

Gly Lys Tyr Ile Arg Lys Phe Ile Phe Thr Asn Ser Glu Leu Gly Arg

625 630 635 640

Glu Ala Ile Trp Leu Ala Asp Asp Ile Lys Arg Phe Met Pro Ala Asp

645 650 655

Val Arg Lys Glu Trp Lys Gly Tyr Gln His Ser Gln Leu Gln Gln Ser

660 665 670

Leu Ala Phe Tyr Asn Ser Arg Pro Lys Glu Ala Leu Ala Ile Leu Glu

675 680 685

Ser Ser Trp Asn Leu Lys Asp Glu Lys Ile Ile Trp Asn Glu Trp Ile

690 695 700

Leu Lys Ser Phe Thr Gln Asn Lys Phe Phe Asp Ala Phe Tyr Asn Glu
705 710 715 720
Tyr Leu Lys Gly Arg Lys Lys Tyr Phe Ala Phe Leu Ser Glu His Ile

725 730 735
Val Gln Tyr Thr Ser Asn Ala Lys Asn Leu Gln Lys Phe Ile Lys Gln
740 745 750
Gln Met Pro Lys Asp Leu Phe Glu Lys Arg His Tyr Ile Ile Glu Asp
755 760 765
Leu Gln Thr Glu Lys Asn Lys Ile Leu Ser Lys Pro Phe Ile Phe Pro
770 775 780
Arg Gly Ile Phe Asp Lys Lys Pro Thr Phe Ile Lys Gly Val Lys Val

785 790 795 800
Glu Asp Ser Pro Glu Ser Phe Ala Asn Trp Tyr Gln Tyr Gly Tyr Gln
805 810 815
Lys Asp His Gln Phe Gln Lys Phe Tyr Asp Trp Lys Arg Asp Tyr Ser
820 825 830
Asp Val Phe Leu Glu His Leu Gly Lys Pro Phe Ile Asn Asn Gly Asp
835 840 845
Arg Arg Thr Leu Gly Met Glu Glu Leu Lys Glu Arg Ile Ile Ile Lys

850 855 860
Gln Asp Leu Lys Ile Lys Lys Ile Lys Ile Gln Asp Leu Phe Leu Arg
865 870 875 880
Leu Ile Ala Glu Asn Leu Phe Gln Lys Val Phe Lys Tyr Ser Ala Lys
885 890 895
Leu Pro Leu Ser Asp Phe Tyr Leu Thr Gln Glu Glu Arg Met Glu Lys
900 905 910
Glu Asn Met Ala Ala Leu Gln Asn Val Arg Glu Glu Gly Asp Lys Ser

915 920 925
Pro Asn Ile Ile Lys Asp Asn Phe Ile Trp Ser Lys Met Ile Pro Tyr
930 935 940
Lys Lys Gly Gln Ile Ile Glu Asn Ala Val Lys Leu Lys Asp Ile Gly

945 950 955 960

Lys Leu Asn Val Leu Ser Leu Asp Asp Lys Val Gln Thr Leu Leu Ser

965 970 975

Tyr Asp Asp Ala Lys Pro Trp Ser Lys Ile Ala Leu Glu Asn Glu Phe

980 985 990

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

995 1000 1005

Lys Glu Ile Gln Gln Phe Glu Ser Glu Ile Leu Phe Arg Ser Gly

1010 1015 1020

Trp Asp Gly Ile Asn His Pro Ala Gln Leu Glu Asp Asn Arg Asn

1025 1030 1035

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

1040 1045 1050

Ala Gly Leu Tyr Ser Gln Gly Glu Asp Ile Trp Phe Glu Tyr Asn

1055 1060 1065

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

1070 1075 1080

Glu Leu Val Gln Leu Ala Phe Leu Val Thr Ala Ile Arg Asn Lys

1085 1090 1095

Phe Ala His Asn Gln Leu Pro Ala Lys Glu Phe Tyr Phe Tyr Ile

1100 1105 1110

Arg Ala Lys Tyr Gly Phe Ala Asp Glu Pro Ser Val Ala Leu Val

1115 1120 1125

Tyr Leu Asn Phe Thr Lys Tyr Ala Ile Asn Glu Phe Lys Lys Val

1130 1135 1140

Met Ile

1145

<210> 144

<211> 948

<212> PRT

<213> Riemerella anatipestifer

<400> 144

Met Phe Phe Ser Phe His Asn Ala Gln Arg Val Ile Phe Lys His Leu
1 5 10 15

Tyr Lys Ala Phe Asp Ala Ser Leu Arg Met Val Lys Glu Asp Tyr Lys
20 25 30

Ala His Phe Thr Val Asn Leu Thr Arg Asp Phe Ala His Leu Asn Arg
35 40 45

Lys Gly Lys Asn Lys Gln Asp Asn Pro Asp Phe Asn Arg Tyr Arg Phe
50 55 60

Glu Lys Asp Gly Phe Phe Thr Glu Ser Gly Leu Leu Phe Phe Thr Asn
65 70 75 80

Leu Phe Leu Asp Lys Arg Asp Ala Tyr Trp Met Leu Lys Lys Val Ser
85 90 95

Gly Phe Lys Ala Ser His Lys Gln Arg Glu Lys Met Thr Thr Glu Val
100 105 110

Phe Cys Arg Ser Arg Ile Leu Leu Pro Lys Leu Arg Leu Glu Ser Arg
115 120 125

Tyr Asp His Asn Gln Met Leu Leu Asp Met Leu Ser Glu Leu Ser Arg
130 135 140

Cys Pro Lys Leu Leu Tyr Glu Lys Leu Ser Glu Glu Asn Lys Lys His
145 150 155 160

Phe Gln Val Glu Ala Asp Gly Phe Leu Asp Glu Ile Glu Glu Glu Gln
165 170 175

Asn Pro Phe Lys Asp Thr Leu Ile Arg His Gln Asp Arg Phe Pro Tyr
180 185 190

Phe Ala Leu Arg Tyr Leu Asp Leu Asn Glu Ser Phe Lys Ser Ile Arg
195 200 205

Phe Gln Val Asp Leu Gly Thr Tyr His Tyr Cys Ile Tyr Asp Lys Lys
210 215 220

Ile Gly Asp Glu Gln Glu Lys Arg His Leu Thr Arg Thr Leu Leu Ser
225 230 235 240

Phe Gly Arg Leu Gln Asp Phe Thr Glu Ile Asn Arg Pro Gln Glu Trp

245 250 255

Lys Ala Leu Thr Lys Asp Leu Asp Tyr Lys Glu Thr Ser Asn Gln Pro

260 265 270

Phe Ile Ser Lys Thr Thr Pro His Tyr His Ile Thr Asp Asn Lys Ile

275 280 285

Gly Phe Arg Leu Gly Thr Ser Lys Glu Leu Tyr Pro Ser Leu Glu Ile

290 295 300

Lys Asp Gly Ala Asn Arg Ile Ala Lys Tyr Pro Tyr Asn Ser Gly Phe

305 310 315 320

Val Ala His Ala Phe Ile Ser Val His Glu Leu Leu Pro Leu Met Phe

325 330 335

Tyr Gln His Leu Thr Gly Lys Ser Glu Asp Leu Leu Lys Glu Thr Val

340 345 350

Arg His Ile Gln Arg Ile Tyr Lys Asp Phe Glu Glu Glu Arg Ile Asn

355 360 365

Thr Ile Glu Asp Leu Glu Lys Ala Asn Gln Gly Arg Leu Pro Leu Gly

370 375 380

Ala Phe Pro Lys Gln Met Leu Gly Leu Leu Gln Asn Lys Gln Pro Asp

385 390 395 400

Leu Ser Glu Lys Ala Lys Ile Lys Ile Glu Lys Leu Ile Ala Glu Thr

405 410 415

Lys Leu Leu Ser His Arg Leu Asn Thr Lys Leu Lys Ser Ser Pro Lys

420 425 430

Leu Gly Lys Arg Arg Glu Lys Leu Ile Lys Thr Gly Val Leu Ala Asp

435 440 445

Trp Leu Val Lys Asp Phe Met Arg Phe Gln Pro Val Ala Tyr Asp Ala

450 455 460

Gln Asn Gln Pro Ile Lys Ser Ser Lys Ala Asn Ser Thr Glu Phe Trp

465 470 475 480

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

485 490 495

Glu Gly Tyr Phe Lys Gln Thr Asn Leu Ile Gly Asn Thr Asn Pro His
500 505 510
Pro Phe Leu Asn Lys Phe Asn Trp Lys Ala Cys Arg Asn Leu Val Asp
515 520 525

Phe Tyr Gln Gln Tyr Leu Glu Gln Arg Glu Lys Phe Leu Glu Ala Ile
530 535 540
Lys Asn Gln Pro Trp Glu Pro Tyr Gln Tyr Cys Leu Leu Leu Lys Ile
545 550 555 560
Pro Lys Glu Asn Arg Lys Asn Leu Val Lys Gly Trp Glu Gln Gly Gly
565 570 575
Ile Ser Leu Pro Arg Gly Leu Phe Thr Glu Ala Ile Arg Glu Thr Leu
580 585 590

Ser Glu Asp Leu Met Leu Ser Lys Pro Ile Arg Lys Glu Ile Lys Lys
595 600 605
His Gly Arg Val Gly Phe Ile Ser Arg Ala Ile Thr Leu Tyr Phe Lys
610 615 620
Glu Lys Tyr Gln Asp Lys His Gln Ser Phe Tyr Asn Leu Ser Tyr Lys
625 630 635 640
Leu Glu Ala Lys Ala Pro Leu Leu Lys Arg Glu Glu His Tyr Glu Tyr
645 650 655

Trp Gln Gln Asn Lys Pro Gln Ser Pro Thr Glu Ser Gln Arg Leu Glu
660 665 670
Leu His Thr Ser Asp Arg Trp Lys Asp Tyr Leu Leu Tyr Lys Arg Trp
675 680 685
Gln His Leu Glu Lys Lys Leu Arg Leu Tyr Arg Asn Gln Asp Val Met
690 695 700
Leu Trp Leu Met Thr Leu Glu Leu Thr Lys Asn His Phe Lys Glu Leu
705 710 715 720

Asn Leu Asn Tyr His Gln Leu Lys Leu Glu Asn Leu Ala Val Asn Val
725 730 735
Gln Glu Ala Asp Ala Lys Leu Asn Pro Leu Asn Gln Thr Leu Pro Met

740 745 750

Val Leu Pro Val Lys Val Tyr Pro Ala Thr Ala Phe Gly Glu Val Gln

755 760 765

Tyr His Lys Thr Pro Ile Arg Thr Val Tyr Ile Arg Glu Glu His Thr

770 775 780

Lys Ala Leu Lys Met Gly Asn Phe Lys Ala Leu Val Lys Asp Arg Arg

785 790 795 800

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

805 810 815

His Pro Ile Ser Gln Leu Arg Leu Arg Arg Glu Leu Glu Ile Tyr Gln

820 825 830

Ser Leu Arg Val Asp Ala Phe Lys Glu Thr Leu Ser Leu Glu Glu Lys

835 840 845

Leu Leu Asn Lys His Thr Ser Leu Ser Ser Leu Glu Asn Glu Phe Arg

850 855 860

Ala Leu Leu Glu Glu Trp Lys Lys Glu Tyr Ala Ala Ser Ser Met Val

865 870 875 880

Thr Asp Glu His Ile Ala Phe Ile Ala Ser Val Arg Asn Ala Phe Cys

885 890 895

His Asn Gln Tyr Pro Phe Tyr Lys Glu Ala Leu His Ala Pro Ile Pro

900 905 910

Leu Phe Thr Val Ala Gln Pro Thr Thr Glu Glu Lys Asp Gly Leu Gly

915 920 925

Ile Ala Glu Ala Leu Leu Lys Val Leu Arg Glu Tyr Cys Glu Ile Val

930 935 940

Lys Ser Gln Ile

945

<210> 145

<211> 1175

<212> PRT

<213> Porphyromonas gingivalis

<400> 145

Met Thr Glu Gln Asn Glu Lys Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn
20 25 30

Ala Tyr Ile Thr Leu Ala His Ile Asp Arg Gln Leu Ala Tyr Ser Lys
35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp
50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu
65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu
100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu
145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205

Lys Tyr Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp
210 215 220

Arg Glu Glu Lys Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu
225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Ala Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Leu Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr

325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln Leu Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Phe Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Ala
500 505 510
Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val
515 520 525

Tyr Ala Val Tyr Asp Ala Phe Ala Arg Asp Glu Ile Asn Thr Arg Asp
530 535 540
Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu
545 550 555 560
Pro Arg Gln Met Ile Ala Ile Leu Ser Gln Glu His Lys Asp Met Glu
565 570 575
Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His
580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly
595 600 605
Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu
610 615 620
Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly
625 630 635 640
Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu
645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro
660 665 670
Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe
675 680 685
Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr
690 695 700
Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg
705 710 715 720

Ser Asp Arg Glu Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys
725 730 735
Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830

Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg

835 840 845

Arg Ile Glu Asp Ala Phe Val Gly Ile Glu Tyr Ala Ser Trp Glu Asn

850 855 860

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

865 870 875 880

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

885 890 895

Leu Lys Lys Glu Ile Leu Glu Ala Lys Glu His Pro Tyr His Asp Phe

900 905 910

Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln

915 920 925

Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys

930 935 940

Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Thr

945 950 955 960

Asp Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn His Val Lys Pro

965 970 975

Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val

980 985 990

His Lys Glu Glu Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp
995 1000 1005

Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp
1010 1015 1020

Arg Arg Leu Asn Gly Leu Phe Ser Phe Val Asp Thr Gly Ala Leu
1025 1030 1035

Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu
1040 1045 1050

Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala Phe Glu Gln Thr
1055 1060 1065

Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro
1070 1075 1080

Asp Glu Ser Phe Arg Glu Met Leu Glu Ser Trp Ser Asp Pro Leu
1085 1090 1095

Leu Asp Lys Trp Pro Asp Leu Gln Arg Glu Val Arg Leu Leu Ile
1100 1105 1110

Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp
1115 1120 1125

Glu Thr Ile Phe Ser Ser Ile Arg Lys Tyr Asp Pro Ser Ser Leu
1130 1135 1140

Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu
1145 1150 1155

Ser Glu Glu Val Lys Leu Ala Lys Glu Met Val Glu Arg Ile Ile
1160 1165 1170

Gln Ala
1175

<210> 146

<211> 1175

<212> PRT

<213> Porphyromonas gingivalis

<400> 146

Met Thr Glu Gln Asn Glu Lys Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu

1 5 10 15

Lys Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn

20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys

35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp

50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu

65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys

85 90 95

Leu Phe Glu Ser Gln Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu

100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu

115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp

130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu

145 150 155 160

Pro Met Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp

165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val

180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp

195 200 205

Arg Cys Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp

210 215 220

Arg Glu Gly Lys Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Thr Tyr Gln Gln Met Thr Asn Glu Val Phe
260 265 270
Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg
275 280 285
Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

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

355 360 365
Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys
370 375 380
Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu
385 390 395 400
Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu
405 410 415
Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Cys Phe Glu Thr Gly Asp

420 425 430
Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly
435 440 445
Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser
450 455 460
Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp
465 470 475 480
Lys Arg Phe Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495
Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Arg Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

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

530 535 540

Glu Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

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

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg

705 710 715 720

Ser Asp Arg Val Glu Asn His Arg Phe Leu Leu Leu Lys Glu Pro Lys

725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Gly Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met
755 760 765
Gly Leu Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys
770 775 780
Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro
785 790 795 800
Phe Tyr Asp Tyr Pro Phe Asn Val Gly Asn Ser Leu Lys Pro Lys Lys

805 810 815
Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly
820 825 830
Lys Glu Arg Phe Arg Asp Leu Glu Ala Trp Ser His Ser Ala Ala Arg
835 840 845
Arg Ile Glu Asp Ala Phe Ala Gly Ile Glu Asn Ala Ser Arg Glu Asn
850 855 860
Lys Lys Lys Ile Glu Gln Leu Leu Gln Asp Leu Ser Leu Trp Glu Thr

865 870 875 880
Phe Glu Ser Lys Leu Lys Val Lys Ala Asp Lys Ile Asn Ile Ala Lys
885 890 895
Leu Lys Lys Glu Ile Leu Glu Ala Lys Glu His Pro Tyr Leu Asp Phe
900 905 910
Lys Ser Trp Gln Lys Phe Glu Arg Glu Leu Arg Leu Val Lys Asn Gln
915 920 925
Asp Ile Ile Thr Trp Met Met Cys Arg Asp Leu Met Glu Glu Asn Lys

930 935 940
Val Glu Gly Leu Asp Thr Gly Thr Leu Tyr Leu Lys Asp Ile Arg Thr
945 950 955 960
Asp Val Gln Glu Gln Gly Ser Leu Asn Val Leu Asn His Val Lys Pro
965 970 975
Met Arg Leu Pro Val Val Val Tyr Arg Ala Asp Ser Arg Gly His Val
980 985 990
His Lys Glu Gln Ala Pro Leu Ala Thr Val Tyr Ile Glu Glu Arg Asp

995 1000 1005

Thr Lys Leu Leu Lys Gln Gly Asn Phe Lys Ser Phe Val Lys Asp

1010 1015 1020

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

1025 1030 1035

Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu Arg Val Glu Tyr Glu

1040 1045 1050

Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala Phe Glu Gln Thr

1055 1060 1065

Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Tyr Pro His Leu Pro

1070 1075 1080

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

1085 1090 1095

Leu Asp Lys Trp Pro Asp Leu His Arg Lys Val Arg Leu Leu Ile

1100 1105 1110

Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr Pro Met Tyr Asp

1115 1120 1125

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

1130 1135 1140

Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile Ala His Arg Leu

1145 1150 1155

Ser Glu Glu Val Lys Gln Ala Lys Glu Met Ala Glu Arg Ile Ile

1160 1165 1170

Gln Ala

1175

<210> 147

<211> 1213

<212

> PRT

<213> Flavobacterium columnare

<400> 147

Met Ser Ser Lys Asn Glu Ser Tyr Asn Lys Gln Lys Thr Phe Asn His

1 5 10 15

Tyr Lys Gln Glu Asp Lys Tyr Phe Phe Gly Gly Phe Leu Asn Asn Ala
20 25 30
Asp Asp Asn Leu Arg Gln Val Gly Lys Glu Phe Lys Thr Arg Ile Asn
35 40 45
Phe Asn His Asn Asn Asn Glu Leu Ala Ser Val Phe Lys Asp Tyr Phe

50 55 60
Asn Lys Glu Lys Ser Val Ala Lys Arg Glu His Ala Leu Asn Leu Leu
65 70 75 80
Ser Asn Tyr Phe Pro Val Leu Glu Arg Ile Gln Lys His Thr Asn His
85 90 95
Asn Phe Glu Gln Thr Arg Glu Ile Phe Glu Leu Leu Leu Asp Thr Ile
100 105 110
Lys Lys Leu Arg Asp Tyr Tyr Thr His His Tyr His Lys Pro Ile Thr

115 120 125
Ile Asn Pro Lys Val Tyr Asp Phe Leu Asp Asp Thr Leu Leu Asp Val
130 135 140
Leu Ile Thr Ile Lys Lys Lys Lys Val Lys Asn Asp Thr Ser Arg Glu
145 150 155 160
Leu Leu Lys Glu Lys Phe Arg Pro Glu Leu Thr Gln Leu Lys Asn Gln
165 170 175
Lys Arg Glu Glu Leu Ile Lys Lys Gly Lys Lys Leu Leu Glu Glu Asn

180 185 190
Leu Glu Asn Ala Val Phe Asn His Cys Leu Arg Pro Phe Leu Glu Glu
195 200 205
Asn Lys Thr Asp Asp Lys Gln Asn Lys Thr Val Ser Leu Arg Lys Tyr
210 215 220
Arg Lys Ser Lys Pro Asn Glu Glu Thr Ser Ile Thr Leu Thr Gln Ser
225 230 235 240
Gly Leu Val Phe Leu Ile Ser Phe Phe Leu His Arg Lys Glu Phe Gln

245 250 255
Val Phe Thr Ser Gly Leu Glu Gly Phe Lys Ala Lys Val Asn Thr Ile

260 265 270

Lys Glu Glu Glu Ile Ser Leu Asn Lys Asn Asn Ile Val Tyr Met Ile

275 280 285

Thr His Trp Ser Tyr Ser Tyr Tyr Asn Phe Lys Gly Leu Lys His Arg

290 295 300

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

305 310 315 320

His Ser Leu Thr Asn Thr Asn Thr Lys Glu Ala Leu Leu Thr Gln Ile

325 330 335

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

340 345 350

Glu Lys Gln Gln Lys Glu Phe Glu Glu Asp Ile Asn Glu Tyr Met Arg

355 360 365

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

370 375 380

Lys Val Ile Arg Lys Arg Tyr Glu Asn Lys Phe Asn Tyr Phe Ala Met

385 390 395 400

Arg Phe Leu Asp Glu Tyr Ala Glu Leu Pro Thr Leu Arg Phe Met Val

405 410 415

Asn Phe Gly Asp Tyr Ile Lys Asp Arg Gln Lys Lys Ile Leu Glu Ser

420 425 430

Ile Gln Phe Asp Ser Glu Arg Ile Ile Lys Lys Glu Ile His Leu Phe

435 440 445

Glu Lys Leu Gly Leu Val Thr Glu Tyr Lys Lys Asn Val Tyr Leu Lys

450 455 460

Glu Thr Ser Asn Ile Asp Leu Ser Arg Phe Pro Leu Phe Pro Ser Pro

465 470 475 480

Ser Tyr Val Met Ala Asn Asn Asn Ile Pro Phe Tyr Ile Asp Ser Arg

485 490 495

Ser Asn Asn Leu Asp Glu Tyr Leu Asn Gln Lys Lys Lys Ala Gln Ser

500 505 510

Gln Asn Arg Lys Arg Asn Leu Thr Phe Glu Lys Tyr Asn Lys Glu Gln
515 520 525
Ser Lys Asp Ala Ile Ile Ala Met Leu Gln Lys Glu Ile Gly Val Lys
530 535 540
Asp Leu Gln Gln Arg Ser Thr Ile Gly Leu Leu Ser Cys Asn Glu Leu
545 550 555 560
Pro Ser Met Leu Tyr Glu Val Ile Val Lys Asp Ile Lys Gly Ala Glu

565 570 575
Leu Glu Asn Lys Ile Ala Gln Lys Ile Arg Glu Gln Tyr Gln Ser Ile
580 585 590
Arg Asp Phe Thr Leu Asp Ser Pro Gln Lys Asp Asn Ile Pro Thr Thr
595 600 605
Leu Thr Lys Thr Ile Ser Thr Asp Thr Ser Val Thr Phe Glu Asn Gln
610 615 620
Pro Ile Asp Ile Pro Arg Leu Lys Asn Ala Leu Gln Lys Glu Leu Thr

625 630 635 640
Leu Thr Gln Glu Lys Leu Leu Asn Val Lys Gln His Glu Ile Glu Val
645 650 655
Asp Asn Tyr Asn Arg Asn Lys Asn Thr Tyr Lys Phe Lys Asn Gln Pro
660 665 670
Lys Asp Lys Val Asp Asp Asn Lys Leu Gln Arg Lys Tyr Val Phe Tyr
675 680 685
Arg Asn Glu Ile Gly Gln Glu Ala Asn Trp Leu Ala Ser Asp Leu Ile

690 695 700
His Phe Met Lys Asn Lys Ser Leu Trp Lys Gly Tyr Met His Asn Glu
705 710 715 720
Leu Gln Ser Phe Leu Ala Phe Phe Glu Asp Lys Lys Asn Asp Cys Ile
725 730 735
Ala Leu Leu Glu Thr Val Phe Asn Leu Lys Glu Asp Cys Ile Leu Thr
740 745 750
Lys Asp Leu Lys Asn Leu Phe Leu Lys His Gly Asn Phe Ile Asp Phe

```
        755                 760                 765
Tyr Lys Glu Tyr Leu Lys Leu Lys Glu Asp Phe Leu Asn Thr Glu Ser
    770                 775                 780
Thr Phe Leu Glu Asn Gly Phe Ile Gly Leu Pro Pro Lys Ile Leu Lys
785                 790                 795                 800
Lys Glu Leu Ser Lys Arg Leu Asn Tyr Ile Phe Ile Val Phe Gln Lys
                805                 810                 815
Arg Gln Phe Ile Ile Lys Glu Leu Glu Glu Lys Lys Asn Asn Leu Tyr

            820                 825                 830
Ala Asp Ala Ile Asn Leu Ser Arg Gly Ile Phe Asp Glu Lys Pro Thr
        835                 840                 845
Met Ile Pro Phe Lys Lys Pro Asn Pro Asp Glu Phe Ala Ser Trp Phe
    850                 855                 860
Val Ala Ser Tyr Gln Tyr Asn Asn Tyr Gln Ser Phe Tyr Glu Leu Thr
865                 870                 875                 880
Pro Asp Lys Ile Glu Asn Asp Lys Lys Lys Lys Tyr Lys Asn Leu Arg

                885                 890                 895
Ala Ile Asn Lys Val Lys Ile Gln Asp Tyr Tyr Leu Lys Leu Met Val
            900                 905                 910
Asp Thr Leu Tyr Gln Asp Leu Phe Asn Gln Pro Leu Asp Lys Ser Leu
        915                 920                 925
Ser Asp Phe Tyr Val Ser Lys Thr Asp Arg Glu Lys Ile Lys Ala Asp
    930                 935                 940
Ala Lys Ala Tyr Gln Lys Arg Asn Asp Ser Phe Leu Trp Asn Lys Val

945                 950                 955                 960
Ile His Leu Ser Leu Gln Asn Asn Arg Ile Thr Ala Asn Pro Lys Leu
                965                 970                 975
Lys Asp Ile Gly Lys Tyr Lys Arg Ala Leu Gln Asp Glu Lys Ile Ala
            980                 985                 990
Thr Leu Leu Thr Tyr Asp Asp Arg  Thr Trp Thr Tyr Ala  Leu Gln Lys
        995                 1000                1005
```

Pro Glu Lys Glu Asn Glu Asn Asp Tyr Lys Glu Leu His Tyr Thr

1010 1015 1020

Ala Leu Asn Met Glu Leu Gln Glu Tyr Glu Lys Val Arg Ser Lys

1025 1030 1035

Lys Leu Leu Lys Gln Val Gln Glu Leu Glu Lys Gln Ile Leu Asp

1040 1045 1050

Lys Phe Tyr Asp Phe Ser Asn Asn Ala Thr His Pro Glu Asp Leu

1055 1060 1065

Glu Ile Glu Asp Lys Lys Gly Lys Arg His Pro Asn Phe Lys Leu

1070 1075 1080

Tyr Ile Thr Lys Ala Leu Leu Lys Asn Glu Ser Glu Ile Ile Asn

1085 1090 1095

Leu Glu Asn Ile Asp Ile Glu Ile Leu Ile Lys Tyr Tyr Asp Tyr

1100 1105 1110

Asn Thr Glu Lys Leu Lys Glu Lys Ile Lys Asn Met Asp Glu Asp

1115 1120 1125

Glu Lys Ala Lys Ile Val Asn Thr Lys Glu Asn Tyr Asn Lys Ile

1130 1135 1140

Thr Asn Val Leu Ile Lys Lys Ala Leu Val Leu Ile Ile Ile Arg

1145 1150 1155

Asn Lys Met Ala His Asn Gln Tyr Pro Pro Lys Phe Ile Tyr Asp

1160 1165 1170

Leu Ala Thr Arg Phe Val Pro Lys Lys Glu Glu Glu Tyr Phe Ala

1175 1180 1185

Cys Tyr Phe Asn Arg Val Phe Glu Thr Ile Thr Thr Glu Leu Trp

1190 1195 1200

Glu Asn Lys Lys Lys Ala Lys Glu Ile Val

1205 1210

<210> 148

<211> 1119

<212> PRT

<213> Porphyromonas gingivalis

<400> 148

Met Thr Glu Gln Asn Glu Arg Pro Tyr Asn Gly Thr Tyr Tyr Thr Leu
1 5 10 15

Glu Asp Lys His Phe Trp Ala Ala Phe Phe Asn Leu Ala Arg His Asn
20 25 30

Ala Tyr Ile Thr Leu Thr His Ile Asp Arg Gln Leu Ala Tyr Ser Lys
35 40 45

Ala Asp Ile Thr Asn Asp Glu Asp Ile Leu Phe Phe Lys Gly Gln Trp
50 55 60

Lys Asn Leu Asp Asn Asp Leu Glu Arg Lys Ala Arg Leu Arg Ser Leu
65 70 75 80

Ile Leu Lys His Phe Ser Phe Leu Glu Gly Ala Ala Tyr Gly Lys Lys
85 90 95

Leu Phe Glu Asn Lys Ser Ser Gly Asn Lys Ser Ser Lys Lys Lys Glu
100 105 110

Leu Thr Lys Lys Glu Lys Glu Glu Leu Gln Ala Asn Ala Leu Ser Leu
115 120 125

Asp Asn Leu Lys Ser Ile Leu Phe Asp Phe Leu Gln Lys Leu Lys Asp
130 135 140

Phe Arg Asn Tyr Tyr Ser His Tyr Arg His Pro Glu Ser Ser Glu Leu
145 150 155 160

Pro Leu Phe Asp Gly Asn Met Leu Gln Arg Leu Tyr Asn Val Phe Asp
165 170 175

Val Ser Val Gln Arg Val Lys Arg Asp His Glu His Asn Asp Lys Val
180 185 190

Asp Pro His Arg His Phe Asn His Leu Val Arg Lys Gly Lys Lys Asp
195 200 205

Arg Cys Gly Asn Asn Asp Asn Pro Phe Phe Lys His His Phe Val Asp
210 215 220

Arg Glu Gly Lys Val Thr Glu Ala Gly Leu Leu Phe Phe Val Ser Leu

225 230 235 240

Phe Leu Glu Lys Arg Asp Ala Ile Trp Met Gln Lys Lys Ile Arg Gly

245 250 255

Phe Lys Gly Gly Thr Glu Ala Tyr Gln Gln Met Thr Asn Glu Val Phe

260 265 270

Cys Arg Ser Arg Ile Ser Leu Pro Lys Leu Lys Leu Glu Ser Leu Arg

275 280 285

Thr Asp Asp Trp Met Leu Leu Asp Met Leu Asn Glu Leu Val Arg Cys

290 295 300

Pro Lys Ser Leu Tyr Asp Arg Leu Arg Glu Glu Asp Arg Ala Arg Phe

305 310 315 320

Arg Val Pro Val Asp Ile Leu Ser Asp Glu Asp Asp Thr Asp Gly Thr

325 330 335

Glu Glu Asp Pro Phe Lys Asn Thr Leu Val Arg His Gln Asp Arg Phe

340 345 350

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

355 360 365

Leu Arg Phe His Ile Asp Leu Gly Thr Tyr His Phe Ala Ile Tyr Lys

370 375 380

Lys Asn Ile Gly Glu Gln Pro Glu Asp Arg His Leu Thr Arg Asn Leu

385 390 395 400

Tyr Gly Phe Gly Arg Ile Gln Asp Phe Ala Glu Glu His Arg Pro Glu

405 410 415

Glu Trp Lys Arg Leu Val Arg Asp Leu Asp Tyr Phe Glu Thr Gly Asp

420 425 430

Lys Pro Tyr Ile Thr Gln Thr Thr Pro His Tyr His Ile Glu Lys Gly

435 440 445

Lys Ile Gly Leu Arg Phe Val Pro Glu Gly Gln His Leu Trp Pro Ser

450 455 460

Pro Glu Val Gly Ala Thr Arg Thr Gly Arg Ser Lys Tyr Ala Gln Asp

465 470 475 480

Lys Arg Leu Thr Ala Glu Ala Phe Leu Ser Val His Glu Leu Met Pro

485 490 495

Met Met Phe Tyr Tyr Phe Leu Leu Arg Glu Lys Tyr Ser Glu Glu Val

500 505 510

Ser Ala Glu Lys Val Gln Gly Arg Ile Lys Arg Val Ile Glu Asp Val

515 520 525

Tyr Ala Val Tyr Asp Ala Phe Ala Arg Gly Glu Ile Asp Thr Leu Asp

530 535 540

Arg Leu Asp Ala Cys Leu Ala Asp Lys Gly Ile Arg Arg Gly His Leu

545 550 555 560

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

565 570 575

Glu Lys Val Arg Lys Lys Leu Gln Glu Met Ile Ala Asp Thr Asp His

580 585 590

Arg Leu Asp Met Leu Asp Arg Gln Thr Asp Arg Lys Ile Arg Ile Gly

595 600 605

Arg Lys Asn Ala Gly Leu Pro Lys Ser Gly Val Ile Ala Asp Trp Leu

610 615 620

Val Arg Asp Met Met Arg Phe Gln Pro Val Ala Lys Asp Thr Ser Gly

625 630 635 640

Lys Pro Leu Asn Asn Ser Lys Ala Asn Ser Thr Glu Tyr Arg Met Leu

645 650 655

Gln Arg Ala Leu Ala Leu Phe Gly Gly Glu Lys Glu Arg Leu Thr Pro

660 665 670

Tyr Phe Arg Gln Met Asn Leu Thr Gly Gly Asn Asn Pro His Pro Phe

675 680 685

Leu His Glu Thr Arg Trp Glu Ser His Thr Asn Ile Leu Ser Phe Tyr

690 695 700

Arg Ser Tyr Leu Lys Ala Arg Lys Ala Phe Leu Gln Ser Ile Gly Arg

705 710 715 720

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

725 730 735

Thr Asp Arg Gln Thr Leu Val Ala Gly Trp Lys Ser Glu Phe His Leu

740 745 750

Pro Arg Gly Ile Phe Thr Glu Ala Val Arg Asp Cys Leu Ile Glu Met

755 760 765

Gly Tyr Asp Glu Val Gly Ser Tyr Lys Glu Val Gly Phe Met Ala Lys

770 775 780

Ala Val Pro Leu Tyr Phe Glu Arg Ala Cys Lys Asp Arg Val Gln Pro

785 790 795 800

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

805 810 815

Gly Arg Phe Leu Ser Lys Glu Lys Arg Ala Glu Glu Trp Glu Ser Gly

820 825 830

Lys Glu Arg Phe Arg Leu Ala Lys Leu Lys Lys Glu Ile Leu Glu Ala

835 840 845

Lys Glu His Pro Tyr Leu Asp Phe Lys Ser Trp Gln Lys Phe Glu Arg

850 855 860

Glu Leu Arg Leu Val Lys Asn Gln Asp Ile Ile Thr Trp Met Met Cys

865 870 875 880

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

885 890 895

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

900 905 910

Asn Val Leu Asn Arg Val Lys Pro Met Arg Leu Pro Val Val Val Tyr

915 920 925

Arg Ala Asp Ser Arg Gly His Val His Lys Glu Gln Ala Pro Leu Ala

930 935 940

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

945 950 955 960

Phe Lys Ser Phe Val Lys Asp Arg Arg Leu Asn Gly Leu Phe Ser Phe

965 970 975

Val Asp Thr Gly Gly Leu Ala Met Glu Gln Tyr Pro Ile Ser Lys Leu
980 985 990
Arg Val Glu Tyr Glu Leu Ala Lys Tyr Gln Thr Ala Arg Val Cys Ala

995 1000 1005
Phe Glu Gln Thr Leu Glu Leu Glu Glu Ser Leu Leu Thr Arg Cys
1010 1015 1020
Pro His Leu Pro Asp Lys Asn Phe Arg Lys Met Leu Glu Ser Trp
1025 1030 1035
Ser Asp Pro Leu Leu Asp Lys Trp Pro Asp Leu Gln Arg Glu Val
1040 1045 1050
Trp Leu Leu Ile Ala Val Arg Asn Ala Phe Ser His Asn Gln Tyr

1055 1060 1065
Pro Met Tyr Asp Glu Ala Val Phe Ser Ser Ile Arg Lys Tyr Asp
1070 1075 1080
Pro Ser Ser Pro Asp Ala Ile Glu Glu Arg Met Gly Leu Asn Ile
1085 1090 1095
Ala His Arg Leu Ser Glu Glu Val Lys Gln Ala Lys Glu Met Ala
1100 1105 1110
Glu Arg Ile Ile Gln Ala
1115

<210> 149
<211> 1135
<212> PRT
<213> Porphyromonas gingivalis
<400> 149
Met Asn Thr Val Pro Ala Ser Glu Asn Lys Gly Gln Ser Arg Thr Val
1 5 10 15
Glu Asp Asp Pro Gln Tyr Phe Gly Leu Tyr Leu Asn Leu Ala Arg Glu
20 25 30
Asn Leu Ile Glu Val Glu Ser His Val Arg Ile Lys Phe Gly Lys Lys
35 40 45

Lys Leu Asn Glu Glu Ser Leu Lys Gln Ser Leu Leu Cys Asp His Leu

50 55 60

Leu Ser Val Asp Arg Trp Thr Lys Val Tyr Gly His Ser Arg Arg Tyr

65 70 75 80

Leu Pro Phe Leu His Tyr Phe Asp Pro Asp Ser Gln Ile Glu Lys Asp

85 90 95

His Asp Ser Lys Thr Gly Val Asp Pro Asp Ser Ala Gln Arg Leu Ile

100 105 110

Arg Glu Leu Tyr Ser Leu Leu Asp Phe Leu Arg Asn Asp Phe Ser His

115 120 125

Asn Arg Leu Asp Gly Thr Thr Phe Glu His Leu Glu Val Ser Pro Asp

130 135 140

Ile Ser Ser Phe Ile Thr Gly Thr Tyr Ser Leu Ala Cys Gly Arg Ala

145 150 155 160

Gln Ser Arg Phe Ala Asp Phe Phe Lys Pro Asp Asp Phe Val Leu Ala

165 170 175

Lys Asn Arg Lys Glu Gln Leu Ile Ser Val Ala Asp Gly Lys Glu Cys

180 185 190

Leu Thr Val Ser Gly Leu Ala Phe Phe Ile Cys Leu Phe Leu Asp Arg

195 200 205

Glu Gln Ala Ser Gly Met Leu Ser Arg Ile Arg Gly Phe Lys Arg Thr

210 215 220

Asp Glu Asn Trp Ala Arg Ala Val His Glu Thr Phe Cys Asp Leu Cys

225 230 235 240

Ile Arg His Pro His Asp Arg Leu Glu Ser Ser Asn Thr Lys Glu Ala

245 250 255

Leu Leu Leu Asp Met Leu Asn Glu Leu Asn Arg Cys Pro Arg Ile Leu

260 265 270

Tyr Asp Met Leu Pro Glu Glu Glu Arg Ala Gln Phe Leu Pro Ala Leu

275 280 285

Asp Glu Asn Ser Met Asn Asn Leu Ser Glu Asn Ser Leu Asn Glu Glu

290 295 300

Ser Arg Leu Leu Trp Asp Gly Ser Ser Asp Trp Ala Glu Ala Leu Thr

305 310 315 320

Lys Arg Ile Arg His Gln Asp Arg Phe Pro Tyr Leu Met Leu Arg Phe

325 330 335

Ile Glu Glu Met Asp Leu Leu Lys Gly Ile Arg Phe Arg Val Asp Leu

340 345 350

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

355 360 365

Glu Tyr Asp Arg Thr Ile Thr Asp His Ala Leu Ala Phe Gly Lys Leu

370 375 380

Ser Asp Phe Gln Asn Glu Glu Glu Val Ser Arg Met Ile Ser Gly Glu

385 390 395 400

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

405 410 415

Tyr Asp Asn Lys Ile Gly Tyr Cys His Thr Ser Asp Pro Val Tyr Pro

420 425 430

Lys Ser Lys Thr Gly Glu Lys Arg Ala Leu Ser Asn Pro Gln Ser Met

435 440 445

Gly Phe Ile Ser Val His Asp Leu Arg Lys Leu Leu Leu Met Glu Leu

450 455 460

Leu Cys Glu Gly Ser Phe Ser Arg Met Gln Ser Gly Phe Leu Arg Lys

465 470 475 480

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

485 490 495

Ala Leu Phe Pro Glu Met Arg His Arg Phe Ile Pro Pro Gln Asn Pro

500 505 510

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

515 520 525

Lys Gln Glu Ile Lys Gly Arg Lys Asp Lys Leu Asn Ser Gln Leu Leu

530 535 540

Ser Ala Phe Asp Met Asn Gln Arg Gln Leu Pro Ser Arg Leu Leu Asp
545 550 555 560
Glu Trp Met Asn Ile Arg Pro Ala Ser His Ser Val Lys Leu Arg Thr

565 570 575
Tyr Val Lys Gln Leu Asn Glu Asp Cys Arg Leu Arg Leu Arg Lys Phe
580 585 590
Arg Lys Asp Gly Asp Gly Lys Ala Arg Ala Ile Pro Leu Val Gly Glu
595 600 605
Met Ala Thr Phe Leu Ser Gln Asp Ile Val Arg Met Ile Ile Ser Glu
610 615 620
Glu Thr Lys Lys Leu Ile Thr Ser Ala Tyr Tyr Asn Glu Met Gln Arg

625 630 635 640
Ser Leu Ala Gln Tyr Ala Gly Glu Glu Asn Arg Arg Gln Phe Arg Ala
645 650 655
Ile Val Ala Glu Leu His Leu Leu Asp Pro Ser Ser Gly His Pro Phe
660 665 670
Leu Ser Ala Thr Met Glu Thr Ala His Arg Tyr Thr Glu Asp Phe Tyr
675 680 685
Lys Cys Tyr Leu Glu Lys Lys Arg Glu Trp Leu Ala Lys Thr Phe Tyr

690 695 700
Arg Pro Glu Gln Asp Glu Asn Thr Lys Arg Arg Ile Ser Val Phe Phe
705 710 715 720
Val Pro Asp Gly Glu Ala Arg Lys Leu Leu Pro Leu Leu Ile Arg Arg
725 730 735
Arg Met Lys Glu Gln Asn Asp Leu Gln Asp Trp Ile Arg Asn Lys Gln
740 745 750
Ala His Pro Ile Asp Leu Pro Ser His Leu Phe Asp Ser Lys Ile Met

755 760 765
Glu Leu Leu Lys Val Lys Asp Gly Lys Lys Lys Trp Asn Glu Ala Phe
770 775 780
Lys Asp Trp Trp Ser Thr Lys Tyr Pro Asp Gly Met Gln Pro Phe Tyr

785 790 795 800

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

805 810 815

Pro Ser Asp Gly Lys Lys Phe Ala Asp Cys Tyr Thr His Leu Met Glu

820 825 830

Lys Thr Val Gln Asp Lys Lys Arg Glu Leu Arg Thr Ala Gly Lys Pro

835 840 845

Val Pro Pro Asp Leu Ala Ala Asp Ile Lys Arg Ser Phe His Arg Ala

850 855 860

Val Asn Glu Arg Glu Phe Met Leu Arg Leu Val Gln Glu Asp Asp Arg

865 870 875 880

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

885 890 895

Ile Leu Pro Gly Leu Lys Asn Ile Asp Ser Ile Leu Asp Lys Glu Asn

900 905 910

Gln Phe Ser Leu Ala Val His Ala Lys Val Leu Glu Lys Glu Gly Glu

915 920 925

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

930 935 940

Ser Lys Arg Lys Asp Trp Ser Lys Tyr Ile Arg Tyr Arg Tyr Asp Arg

945 950 955 960

Arg Val Pro Gly Leu Met Ser His Phe Pro Glu His Lys Ala Thr Leu

965 970 975

Asp Glu Val Lys Thr Leu Leu Gly Glu Tyr Asp Arg Cys Arg Ile Lys

980 985 990

Ile Phe Asp Trp Ala Phe Ala Leu Glu Gly Ala Ile Met Ser Asp Arg

995 1000 1005

Asp Leu Lys Pro Tyr Leu His Glu Ser Ser Ser Arg Glu Gly Lys

1010 1015 1020

Ser Gly Glu His Ser Thr Leu Val Lys Met Leu Val Glu Lys Lys

1025 1030 1035

Gly Cys Leu Thr Pro Asp Glu Ser Gln Tyr Leu Ile Leu Ile Arg

1040 1045 1050

Asn Lys Ala Ala His Asn Gln Phe Pro Cys Ala Ala Glu Met Pro

1055 1060 1065

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

1070 1075 1080

Ser Ser Ala Lys Asp Leu Pro Glu Gly Ser Ser Leu Val Asp Ser

1085 1090 1095

Leu Trp Lys Lys Tyr Glu Met Ile Ile Arg Lys Ile Leu Pro Ile

1100 1105 1110

Leu Asp Pro Glu Asn Arg Phe Phe Gly Lys Leu Leu Asn Asn Met

1115 1120 1125

Ser Gln Pro Ile Asn Asp Leu

1130 1135

<210> 150

<211> 950

<212> PRT

<213> Riemerella anatipestifer

<400> 150

Met Phe Phe Ser Phe His Asn Ala Gln Arg Val Ile Phe Lys His Leu

1 5 10 15

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

20 25 30

Ala His Phe Thr Val Asn Leu Thr Arg Asp Phe Ala His Leu Asn Arg

35 40 45

Lys Gly Lys Asn Lys Gln Asp Asn Pro Asp Phe Asn Arg Tyr Arg Phe

50 55 60

Glu Lys Asp Gly Phe Phe Thr Glu Ser Gly Leu Leu Phe Phe Thr Asn

65 70 75 80

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

85 90 95

Gly Phe Lys Ala Ser His Lys Gln Ser Glu Lys Met Thr Thr Glu Val

100 105 110

Phe Cys Arg Ser Arg Ile Leu Leu Pro Lys Leu Arg Leu Glu Ser Arg

115 120 125

Tyr Asp His Asn Gln Met Leu Leu Asp Met Leu Ser Glu Leu Ser Arg

130 135 140

Cys Pro Lys Leu Leu Tyr Glu Lys Leu Ser Glu Lys Asp Lys Lys Cys

145 150 155 160

Phe Gln Val Glu Ala Asp Gly Phe Leu Asp Glu Ile Glu Glu Glu Gln

165 170 175

Asn Pro Phe Lys Asp Thr Leu Ile Arg His Gln Asp Arg Phe Pro Tyr

180 185 190

Phe Ala Leu Arg Tyr Leu Asp Leu Asn Glu Ser Phe Lys Ser Ile Arg

195 200 205

Phe Gln Val Asp Leu Gly Thr Tyr His Tyr Cys Ile Tyr Asp Lys Lys

210 215 220

Ile Gly Tyr Glu Gln Glu Lys Arg His Leu Thr Arg Thr Leu Leu Asn

225 230 235 240

Phe Gly Arg Leu Gln Asp Phe Thr Glu Ile Asn Arg Pro Gln Glu Trp

245 250 255

Lys Ala Leu Thr Lys Asp Leu Asp Tyr Asn Glu Thr Ser Asn Gln Pro

260 265 270

Phe Ile Ser Lys Thr Thr Pro His Tyr His Ile Thr Asp Asn Lys Ile

275 280 285

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

290 295 300

Lys Asp Gly Ala Asn Arg Ile Ala Lys Tyr Pro Tyr Asn Ser Asp Phe

305 310 315 320

Val Ala His Ala Phe Ile Ser Ile Ser Val His Glu Leu Leu Pro Leu

325 330 335

Met Phe Tyr Gln His Leu Thr Gly Lys Ser Glu Asp Leu Leu Lys Glu

340 345 350

Thr Val Arg His Ile Gln Arg Ile Tyr Lys Asp Phe Glu Glu Glu Arg

355 360 365

Ile Asn Thr Ile Glu Asp Leu Glu Lys Ala Asn Gln Gly Arg Leu Pro

370 375 380

Leu Gly Ala Phe Pro Lys Gln Met Leu Gly Leu Leu Gln Asn Lys Gln

385 390 395 400

Pro Asp Leu Ser Glu Lys Ala Lys Ile Lys Ile Glu Lys Leu Ile Ala

405 410 415

Glu Thr Lys Leu Leu Ser His Arg Leu Asn Thr Lys Leu Lys Ser Ser

420 425 430

Pro Lys Leu Gly Lys Arg Arg Glu Lys Leu Ile Lys Thr Gly Val Leu

435 440 445

Ala Asp Trp Leu Val Lys Asp Phe Met Arg Phe Gln Pro Val Val Tyr

450 455 460

Asp Ala Gln Asn Gln Pro Ile Lys Ser Ser Lys Ala Asn Ser Thr Glu

465 470 475 480

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

485 490 495

Arg Leu Glu Gly Tyr Phe Lys Gln Thr Asn Leu Ile Gly Asn Thr Asn

500 505 510

Pro His Pro Phe Leu Asn Lys Phe Asn Trp Lys Ala Cys Arg Asn Leu

515 520 525

Val Asp Phe Tyr Gln Gln Tyr Leu Glu Gln Arg Glu Lys Phe Leu Glu

530 535 540

Ala Ile Lys His Gln Pro Trp Glu Pro Tyr Gln Tyr Cys Leu Leu Leu

545 550 555 560

Lys Val Pro Lys Glu Asn Arg Lys Asn Leu Val Lys Gly Trp Glu Gln

565 570 575

Gly Gly Ile Ser Leu Pro Arg Gly Leu Phe Thr Glu Ala Ile Arg Glu

580 585 590

Thr Leu Ser Lys Asp Leu Thr Leu Ser Lys Pro Ile Arg Lys Glu Ile

595 600 605

Lys Lys His Gly Arg Val Gly Phe Ile Ser Arg Ala Ile Thr Leu Tyr

610 615 620

Phe Lys Glu Lys Tyr Gln Asp Lys His Gln Ser Phe Tyr Asn Leu Ser

625 630 635 640

Tyr Lys Leu Glu Ala Lys Ala Pro Leu Leu Lys Lys Glu Glu His Tyr

645 650 655

Glu Tyr Trp Gln Gln Asn Lys Pro Gln Ser Pro Thr Glu Ser Gln Arg

660 665 670

Leu Glu Leu His Thr Ser Asp Arg Trp Lys Asp Tyr Leu Leu Tyr Lys

675 680 685

Arg Trp Gln His Leu Glu Lys Lys Leu Arg Leu Tyr Arg Asn Gln Asp

690 695 700

Ile Met Leu Trp Leu Met Thr Leu Glu Leu Thr Lys Asn His Phe Lys

705 710 715 720

Glu Leu Asn Leu Asn Tyr His Gln Leu Lys Leu Glu Asn Leu Ala Val

725 730 735

Asn Val Gln Glu Ala Asp Ala Lys Leu Asn Pro Leu Asn Gln Thr Leu

740 745 750

Pro Met Val Leu Pro Val Lys Val Tyr Pro Thr Thr Ala Phe Gly Glu

755 760 765

Val Gln Tyr His Glu Thr Pro Ile Arg Thr Val Tyr Ile Arg Glu Glu

770 775 780

Gln Thr Lys Ala Leu Lys Met Gly Asn Phe Lys Ala Leu Val Lys Asp

785 790 795 800

Arg His Leu Asn Gly Leu Phe Ser Phe Ile Lys Glu Glu Asn Asp Thr

805 810 815

Gln Lys His Pro Ile Ser Gln Leu Arg Leu Arg Arg Glu Leu Glu Ile

820 825 830

Tyr Gln Ser Leu Arg Val Asp Ala Phe Lys Glu Thr Leu Ser Leu Glu

835 840 845

Glu Lys Leu Leu Asn Lys His Ala Ser Leu Ser Ser Leu Glu Asn Glu
850 855 860
Phe Arg Thr Leu Leu Glu Glu Trp Lys Lys Lys Tyr Ala Ala Ser Ser
865 870 875 880
Met Val Thr Asp Lys His Ile Ala Phe Ile Ala Ser Val Arg Asn Ala

885 890 895
Phe Cys His Asn Gln Tyr Pro Phe Tyr Lys Glu Thr Leu His Ala Pro
900 905 910
Ile Leu Leu Phe Thr Val Ala Gln Pro Thr Thr Glu Glu Lys Asp Gly
915 920 925
Leu Gly Ile Ala Glu Ala Leu Leu Arg Val Leu Arg Glu Tyr Cys Glu
930 935 940
Ile Val Lys Ser Gln Ile
945 950

<210> 151
<211> 1213
<212> PRT
<213> Flavobacterium columnare
<400> 151
Met Ser Ser Lys Asn Glu Ser Tyr Asn Lys Gln Lys Thr Phe Asn His
1 5 10 15
Tyr Lys Gln Glu Asp Lys Tyr Phe Phe Gly Gly Phe Leu Asn Asn Ala
20 25 30
Asp Asp Asn Leu Arg Gln Val Gly Lys Glu Phe Lys Thr Arg Ile Asn
35 40 45
Phe Asn His Asn Asn Asn Glu Leu Ala Ser Val Phe Lys Asp Tyr Phe

50 55 60
Asn Lys Glu Lys Ser Val Ala Lys Arg Glu His Ala Leu Asn Leu Leu
65 70 75 80
Ser Asn Tyr Phe Pro Val Leu Glu Arg Ile Gln Lys His Thr Asn His
85 90 95
Asn Phe Glu Gln Thr Arg Glu Ile Phe Glu Leu Leu Leu Asp Thr Ile

100 105 110

Lys Lys Leu Arg Asp Tyr Tyr Thr His His Tyr His Lys Pro Ile Thr

115 120 125

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

130 135 140

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

145 150 155 160

Leu Leu Lys Glu Lys Leu Arg Pro Glu Leu Thr Gln Leu Lys Asn Gln

165 170 175

Lys Arg Glu Glu Leu Ile Lys Lys Gly Lys Lys Leu Leu Glu Glu Asn

180 185 190

Leu Glu Asn Ala Val Phe Asn His Cys Leu Arg Pro Phe Leu Glu Glu

195 200 205

Asn Lys Thr Asp Asp Lys Gln Asn Lys Thr Val Ser Leu Arg Lys Tyr

210 215 220

Arg Lys Ser Lys Pro Asn Glu Glu Thr Ser Ile Thr Leu Thr Gln Ser

225 230 235 240

Gly Leu Val Phe Leu Met Ser Phe Phe Leu His Arg Lys Glu Phe Gln

245 250 255

Val Phe Thr Ser Gly Leu Glu Gly Phe Lys Ala Lys Val Asn Thr Ile

260 265 270

Lys Glu Glu Lys Ile Ser Leu Asn Lys Asn Asn Ile Val Tyr Met Ile

275 280 285

Thr His Trp Ser Tyr Ser Tyr Tyr Asn Phe Lys Gly Leu Lys His Arg

290 295 300

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

305 310 315 320

His Ser Leu Thr Asn Thr Asn Thr Lys Glu Ala Leu Leu Thr Gln Ile

325 330 335

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

340 345 350

Glu Lys Gln Gln Lys Glu Phe Glu Glu Asp Ile Asn Glu Tyr Met Arg
355 360 365
Glu Asn Pro Glu Asn Glu Asp Ser Thr Phe Ser Ser Ile Val Ser His

370 375 380
Lys Val Ile Arg Lys Arg Tyr Glu Asn Lys Phe Asn Tyr Phe Ala Met
385 390 395 400
Arg Phe Leu Asp Glu Tyr Ala Glu Leu Pro Thr Leu Arg Phe Met Val
405 410 415
Asn Phe Gly Asp Tyr Ile Lys Asp Arg Gln Lys Lys Ile Leu Glu Ser
420 425 430
Ile Gln Phe Asp Ser Glu Arg Ile Ile Lys Lys Glu Ile His Leu Phe

435 440 445
Glu Lys Leu Gly Leu Val Thr Glu Tyr Lys Lys Asn Val Tyr Leu Lys
450 455 460
Glu Thr Ser Asn Ile Asp Leu Ser Arg Phe Pro Leu Phe Pro Ser Pro
465 470 475 480
Ser Tyr Val Met Ala Asn Asn Asn Ile Pro Phe Tyr Ile Asp Ser Arg
485 490 495
Ser Asn Asn Leu Asp Glu Tyr Leu Asn Gln Lys Lys Lys Ala Gln Ser

500 505 510
Gln Asn Arg Lys Arg Asn Leu Thr Phe Glu Lys Tyr Asn Lys Glu Gln
515 520 525
Ser Lys Asp Ala Ile Ile Ala Met Leu Gln Lys Glu Ile Gly Val Lys
530 535 540
Asp Leu Gln Gln Arg Ser Thr Ile Gly Leu Leu Ser Cys Asn Glu Leu
545 550 555 560
Pro Ser Met Leu Tyr Glu Val Ile Val Lys Asp Ile Lys Gly Ala Glu

565 570 575
Leu Glu Asn Lys Ile Ala Gln Lys Ile Arg Glu Gln Tyr Gln Ser Ile
580 585 590
Arg Asp Phe Thr Leu Asn Ser Pro Gln Lys Asp Asn Ile Pro Thr Thr

595 600 605

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

610 615 620

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

625 630 635 640

Leu Thr Gln Glu Lys Leu Leu Asn Val Lys Gln His Glu Ile Glu Val

645 650 655

Asn Asn Tyr Asn Arg Asn Lys Asn Thr Tyr Lys Phe Lys Asn Gln Pro

660 665 670

Lys Asp Lys Val Asp Asp Asn Lys Leu Gln Arg Lys Tyr Val Phe Tyr

675 680 685

Arg Asn Glu Ile Gly Gln Glu Ala Asn Trp Leu Ala Ser Asp Leu Ile

690 695 700

His Phe Met Lys Asn Lys Ser Leu Trp Lys Gly Tyr Met His Asn Glu

705 710 715 720

Leu Gln Ser Phe Leu Ala Phe Phe Glu Asp Lys Lys Asn Asp Cys Ile

725 730 735

Ala Leu Leu Glu Thr Val Phe Asn Leu Lys Glu Asp Cys Ile Leu Thr

740 745 750

Lys Asp Leu Lys Asn Leu Phe Leu Lys His Gly Asn Phe Ile Asp Phe

755 760 765

Tyr Lys Glu Tyr Leu Lys Leu Lys Glu Asp Phe Leu Asn Thr Glu Ser

770 775 780

Thr Phe Leu Glu Asn Gly Phe Ile Gly Leu Pro Pro Lys Ile Leu Lys

785 790 795 800

Lys Glu Leu Ser Lys Arg Leu Asn Tyr Ile Phe Ile Val Phe Gln Lys

805 810 815

Arg Gln Phe Ile Ile Lys Glu Leu Glu Glu Lys Lys Asn Asn Leu Tyr

820 825 830

Ala Asp Ala Ile Asn Leu Ser Arg Gly Ile Phe Asp Glu Lys Pro Thr

835 840 845

Met Ile Pro Phe Lys Lys Pro Asn Pro Asp Glu Phe Ala Ser Trp Phe
850 855 860
Val Ala Ser Tyr Gln Tyr Asn Asn Tyr Gln Ser Phe Tyr Glu Leu Thr
865 870 875 880
Pro Asp Lys Ile Glu Asn Asp Lys Lys Lys Lys Tyr Lys Asn Leu Arg

885 890 895
Ala Ile Asn Lys Val Lys Ile Gln Asp Tyr Tyr Leu Lys Leu Met Val
900 905 910
Asp Thr Leu Tyr Gln Asp Leu Phe Asn Gln Pro Leu Asp Lys Ser Leu
915 920 925
Ser Asp Phe Tyr Val Ser Lys Thr Asp Arg Glu Lys Ile Lys Ala Asp
930 935 940
Ala Lys Ala Tyr Gln Lys Arg Asn Asp Ser Phe Leu Trp Asn Lys Val

945 950 955 960
Ile His Leu Ser Leu Gln Asn Asn Arg Ile Thr Ala Asn Pro Lys Leu
965 970 975
Lys Asp Ile Gly Lys Tyr Lys Arg Ala Leu Gln Asp Glu Lys Ile Ala
980 985 990
Thr Leu Leu Thr Tyr Asp Asp Arg Thr Trp Thr Tyr Ala Leu Gln Lys
995 1000 1005
Pro Glu Lys Glu Asn Glu Asn Asp Tyr Lys Glu Leu His Tyr Thr

1010 1015 1020
Ala Leu Asn Met Glu Leu Gln Glu Tyr Glu Lys Val Arg Ser Lys
1025 1030 1035
Lys Leu Leu Lys Gln Val Gln Glu Leu Glu Lys Gln Ile Leu Asp
1040 1045 1050
Lys Phe Tyr Asp Phe Ser Asn Asn Ala Thr His Pro Glu Asp Leu
1055 1060 1065
Glu Ile Glu Asp Lys Lys Gly Lys Arg His Pro Asn Phe Lys Leu

1070 1075 1080
Tyr Ile Thr Lys Ala Leu Leu Lys Asn Glu Ser Glu Ile Ile Asn

1085 1090 1095

Leu Glu Asn Ile Asp Ile Glu Ile Leu Ile Lys Tyr Tyr Asp Tyr

1100 1105 1110

Asn Thr Glu Lys Leu Lys Glu Lys Ile Lys Asn Met Asp Glu Asp

1115 1120 1125

Glu Lys Ala Lys Ile Val Asn Thr Lys Glu Asn Tyr Asn Lys Ile

1130 1135 1140

Thr Asn Val Leu Ile Lys Lys Ala Leu Val Leu Ile Ile Ile Arg

1145 1150 1155

Asn Lys Met Ala His Asn Gln Tyr Pro Pro Lys Phe Ile Tyr Asp

1160 1165 1170

Leu Ala Thr Arg Phe Val Pro Lys Lys Glu Glu Glu Tyr Phe Ala

1175 1180 1185

Cys Tyr Phe Asn Arg Val Phe Glu Thr Ile Thr Thr Glu Leu Trp

1190 1195 1200

Glu Asn Lys Lys Lys Ala Lys Glu Ile Val

1205 1210

<210> 152

<211> 1095

<212> PRT

<213> Riemerella anatipestifer

<400> 152

Met Glu Lys Pro Leu Pro Pro Asn Val Tyr Thr Leu Lys His Lys Phe

1 5 10 15

Phe Trp Gly Ala Phe Leu Asn Ile Ala Arg His Asn Ala Phe Ile Thr

20 25 30

Ile Cys His Ile Asn Glu Gln Leu Gly Leu Thr Thr Pro Pro Asn Asp

35 40 45

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

50 55 60

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

65 70 75 80

Lys His Phe Pro Phe Leu Ala Ala Met Cys Tyr His Pro Pro Lys Lys
85 90 95
Glu Gly Lys Lys Lys Gly Ser Gln Lys Glu Gln Gln Lys Glu Lys Glu

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

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

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

290 295 300
Lys Lys Arg Phe Gln Val Glu Ala Asp Gly Phe Leu Asp Glu Ile Glu
305 310 315 320
Glu Glu Gln Asn Pro Phe Lys Asp Thr Leu Ile Arg His Gln Asp Arg

325 330 335

Phe Pro Tyr Phe Ala Leu Arg Tyr Leu Asp Leu Asn Glu Ser Phe Lys

340 345 350

Ser Ile Arg Phe Gln Val Asp Leu Gly Thr Tyr His Tyr Cys Ile Tyr

355 360 365

Asp Lys Lys Ile Gly Asp Glu Gln Glu Lys Arg His Leu Thr Arg Thr

370 375 380

Leu Leu Ser Phe Gly Arg Leu Gln Asp Phe Thr Glu Ile Asn Arg Pro

385 390 395 400

Gln Glu Trp Lys Ala Leu Thr Lys Asp Leu Asp Tyr Lys Glu Thr Ser

405 410 415

Lys Gln Pro Phe Ile Ser Lys Thr Thr Pro His Tyr His Ile Thr Asp

420 425 430

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

435 440 445

Leu Glu Val Lys Asp Gly Ala Asn Arg Ile Ala Gln Tyr Pro Tyr Asn

450 455 460

Ser Asp Phe Val Ala His Ala Phe Ile Ser Val His Glu Leu Leu Pro

465 470 475 480

Leu Met Phe Tyr Gln His Leu Thr Gly Lys Ser Glu Asp Leu Leu Lys

485 490 495

Glu Thr Val Arg His Ile Gln Arg Ile Tyr Lys Asp Phe Glu Glu Glu

500 505 510

Arg Ile Asn Thr Ile Glu Asp Leu Glu Lys Ala Asn Gln Gly Arg Leu

515 520 525

Pro Leu Gly Ala Phe Pro Lys Gln Met Leu Gly Leu Leu Gln Asn Lys

530 535 540

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

545 550 555 560

Ala Glu Thr Lys Leu Leu Ser His Arg Leu Asn Thr Lys Leu Lys Ser

565 570 575

Ser Pro Lys Leu Gly Lys Arg Arg Glu Lys Leu Ile Lys Thr Gly Val
580 585 590
Leu Ala Asp Trp Leu Val Lys Asp Phe Met Arg Phe Gln Pro Val Ala
595 600 605
Tyr Asp Ala Gln Asn Gln Pro Ile Glu Ser Ser Lys Ala Asn Ser Thr

610 615 620
Glu Phe Gln Leu Ile Gln Arg Ala Leu Ala Leu Tyr Gly Gly Glu Lys
625 630 635 640
Asn Arg Leu Glu Gly Tyr Phe Lys Gln Thr Asn Leu Ile Gly Asn Thr
645 650 655
Asn Pro His Pro Phe Leu Asn Lys Phe Asn Trp Lys Ala Cys Arg Asn
660 665 670
Leu Val Asp Phe Tyr Gln Gln Tyr Leu Glu Gln Arg Glu Lys Phe Leu

675 680 685
Glu Ala Ile Lys Asn Gln Pro Trp Glu Pro Tyr Gln Tyr Cys Leu Leu
690 695 700
Leu Lys Ile Pro Lys Glu Asn Arg Lys Asn Leu Val Lys Gly Trp Glu
705 710 715 720
Gln Gly Gly Ile Ser Leu Pro Arg Gly Leu Phe Thr Glu Ala Ile Arg
725 730 735
Glu Thr Leu Ser Lys Asp Leu Thr Leu Ser Lys Pro Ile Arg Lys Glu

740 745 750
Ile Lys Lys His Gly Arg Val Gly Phe Ile Ser Arg Ala Ile Thr Leu
755 760 765
Tyr Phe Arg Glu Lys Tyr Gln Asp Asp His Gln Ser Phe Tyr Asp Leu
770 775 780
Pro Tyr Lys Leu Glu Ala Lys Ala Ser Pro Leu Pro Lys Lys Glu His
785 790 795 800
Tyr Glu Tyr Trp Gln Gln Asn Lys Pro Gln Ser Pro Thr Glu Leu Gln

805 810 815
Arg Leu Glu Leu His Thr Ser Asp Arg Trp Lys Asp Tyr Leu Leu Tyr

820 825 830

Lys Arg Trp Gln His Leu Glu Lys Lys Leu Arg Leu Tyr Arg Asn Gln

835 840 845

Asp Val Met Leu Trp Leu Met Thr Leu Glu Leu Thr Lys Asn His Phe

850 855 860

Lys Glu Leu Asn Leu Asn Tyr His Gln Leu Lys Leu Glu Asn Leu Ala

865 870 875 880

Val Asn Val Gln Glu Ala Asp Ala Lys Leu Asn Pro Leu Asn Gln Thr

885 890 895

Leu Pro Met Val Leu Pro Val Lys Val Tyr Pro Ala Thr Ala Phe Gly

900 905 910

Glu Val Gln Tyr Gln Glu Thr Pro Ile Arg Thr Val Tyr Ile Arg Glu

915 920 925

Glu Gln Thr Lys Ala Leu Lys Met Gly Asn Phe Lys Ala Leu Val Lys

930 935 940

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

945 950 955 960

Thr Gln Lys His Pro Ile Ser Gln Leu Arg Leu Arg Arg Glu Leu Glu

965 970 975

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

980 985 990

Glu Glu Lys Leu Leu Lys Lys His Thr Ser Leu Ser Ser Val Glu Asn

995 1000 1005

Lys Phe Arg Ile Leu Leu Glu Glu Trp Lys Lys Glu Tyr Ala Ala

1010 1015 1020

Ser Ser Met Val Thr Asp Glu His Ile Ala Phe Ile Ala Ser Val

1025 1030 1035

Arg Asn Ala Phe Cys His Asn Gln Tyr Pro Phe Tyr Glu Glu Ala

1040 1045 1050

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

1055 1060 1065

Glu Glu Lys Asp Gly Leu Gly Ile Ala Glu Ala Leu Leu Arg Val
1070 1075 1080
Leu Arg Glu Tyr Cys Glu Ile Val Lys Ser Gln Ile
1085 1090 1095

<210> 153
<211> 1149
<212> PRT
<213> Sinomicrobium oceani

<400> 153

Met Glu Ser Thr Thr Thr Leu Gly Leu His Leu Lys Tyr Gln His Asp
1 5 10 15

Leu Phe Glu Asp Lys His Tyr Phe Gly Gly Gly Val Asn Leu Ala Val
20 25 30
Gln Asn Ile Glu Ser Ile Phe Gln Ala Phe Ala Glu Arg Tyr Gly Ile
35 40 45
Gln Asn Pro Leu Arg Lys Asn Gly Val Pro Ala Ile Asn Asn Ile Phe
50 55 60
His Asp Asn Ile Ser Ile Ser Asn Tyr Lys Glu Tyr Leu Lys Phe Leu
65 70 75 80

Lys Gln Tyr Leu Pro Val Val Gly Phe Leu Glu Lys Ser Asn Glu Ile
85 90 95
Asn Ile Phe Glu Phe Arg Glu Asp Phe Glu Ile Leu Ile Asn Ala Ile
100 105 110
Tyr Lys Leu Arg His Phe Tyr Thr His Tyr Tyr His Ser Pro Ile Lys
115 120 125
Leu Glu Asp Arg Phe Tyr Thr Cys Leu Asn Glu Leu Phe Val Ala Val
130 135 140

Ala Ile Gln Val Lys Lys His Lys Met Lys Ser Asp Lys Thr Arg Gln
145 150 155 160
Leu Leu Asn Lys Asn Leu His Gln Leu Leu Gln Gln Leu Ile Glu Gln
165 170 175
Lys Arg Glu Lys Leu Lys Asp Lys Lys Ala Glu Gly Glu Lys Val Ser

180 185 190

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

195 200 205

His Leu Leu Asp Lys Asp Glu Asn Ile Arg Leu Asn Tyr Ser Ser Arg

210 215 220

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

225 230 235 240

Gly Leu Leu Phe Leu Leu Ser Leu Phe Leu Gln Arg Lys Glu Ala Glu

245 250 255

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

260 265 270

Arg Phe Met Ala Thr His Trp Val Phe Ser Tyr Leu Asn Val Lys Arg

275 280 285

Ile Lys His Arg Leu Asn Thr Asp Phe Gln Lys Glu Thr Leu Leu Ile

290 295 300

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

305 310 315 320

Leu Asp His Glu Asn Arg Ser Lys Phe Leu Glu Asp Ile Asn Glu Tyr

325 330 335

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

340 345 350

His Gly Val Ile Arg Lys Arg Tyr Glu Asn Lys Phe His Tyr Leu Val

355 360 365

Leu Arg Tyr Leu Asp Glu Phe Val Asp Phe Pro Ser Leu Arg Phe Gln

370 375 380

Val His Leu Gly Asn Tyr Ile His Asp Arg Arg Asp Lys Val Ile Asp

385 390 395 400

Gly Thr Asn Phe Ile Thr Asn Arg Val Ile Lys Glu Pro Ile Lys Val

405 410 415

Phe Gly Lys Leu Ser His Val Ser Lys Leu Lys Ser Asp Tyr Met Glu

420 425 430

Ser Leu Ser Arg Glu His Lys Asn Gly Trp Asp Val Phe Pro Asn Pro
435 440 445
Ser Tyr Asn Phe Val Gly His Asn Ile Pro Ile Phe Ile Asn Leu Arg
450 455 460

Ser Ala Ser Ser Lys Gly Lys Glu Leu Tyr Arg Asp Leu Met Lys Ile
465 470 475 480
Lys Ser Glu Lys Lys Lys Lys Ser Arg Glu Glu Gly Ile Pro Met Glu
485 490 495
Arg Arg Asp Gly Lys Pro Thr Lys Ile Glu Ile Ser Asn Gln Ile Asp
500 505 510
Arg Asn Ile Lys Asp Asn Asn Phe Lys Asp Ile Tyr Pro Gly Glu Pro
515 520 525

Leu Ala Met Leu Ser Leu Asn Glu Leu Pro Ala Leu Leu Phe Glu Leu
530 535 540
Leu Arg Arg Pro Ser Ile Thr Pro Gln Asp Ile Glu Asp Arg Met Val
545 550 555 560
Glu Lys Leu Tyr Glu Arg Phe Gln Ile Ile Arg Asp Tyr Lys Pro Gly
565 570 575
Asp Gly Leu Ser Thr Ser Lys Ile Ser Lys Lys Leu Arg Lys Ala Asp
580 585 590

Asn Ser Thr Arg Leu Asp Gly Lys Lys Leu Leu Arg Ala Ile Gln Thr
595 600 605
Glu Thr Arg Asn Ala Arg Glu Lys Leu His Thr Leu Glu Glu Asn Lys
610 615 620
Ala Leu Gln Lys Asn Arg Lys Arg Arg Thr Val Tyr Thr Thr Arg Glu
625 630 635 640
Gln Gly Arg Glu Ala Ser Trp Leu Ala Gln Asp Leu Lys Arg Phe Met
645 650 655

Pro Ile Ala Ser Arg Lys Glu Trp Arg Gly Tyr His His Ser Gln Leu
660 665 670
Gln Gln Ile Leu Ala Phe Tyr Asp Gln Asn Pro Lys Gln Pro Leu Glu

675 680 685

Leu Leu Glu Gln Phe Trp Asp Leu Lys Glu Asp Thr Tyr Val Trp Asn

690 695 700

Ser Trp Ile His Lys Ser Leu Ser Gln His Asn Gly Phe Val Pro Met

705 710 715 720

Tyr Glu Gly Tyr Leu Lys Gly Arg Leu Gly Tyr Tyr Lys Lys Leu Glu

725 730 735

Ser Asp Ile Ile Gly Phe Leu Glu Glu His Lys Val Leu Lys Arg Tyr

740 745 750

Tyr Thr Gln Gln His Leu Asn Val Ile Phe Arg Glu Arg Leu Tyr Phe

755 760 765

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

770 775 780

Val Phe Pro Arg Gly Ile Phe Asp Asp Lys Pro Thr Phe Val Gln Asp

785 790 795 800

Lys Lys Val Val Asp His Pro Glu Leu Phe Ala Asp Trp Tyr Val Tyr

805 810 815

Ser Tyr Lys Asp Asp His Ser Phe Gln Glu Phe Tyr His Tyr Lys Arg

820 825 830

Asp Tyr Asn Glu Ile Phe Glu Thr Glu Leu Ser Trp Asp Ile Asp Phe

835 840 845

Lys Asp Asn Lys Arg Gln Leu Asn Pro Ser Glu Gln Met Asp Leu Phe

850 855 860

Arg Met Lys Trp Asp Leu Lys Ile Lys Lys Ile Lys Ile Gln Asp Ile

865 870 875 880

Phe Leu Lys Ile Val Ala Glu Asp Ile Tyr Leu Lys Ile Phe Gly His

885 890 895

Lys Ile Pro Leu Ser Leu Ser Asp Phe Tyr Ile Ser Arg Gln Glu Arg

900 905 910

Leu Thr Leu Asp Glu Gln Ala Val Ala Gln Ser Met Arg Leu Pro Gly

915 920 925

Asp Thr Ser Glu Asn Gln Ile Lys Glu Ser Asn Leu Trp Gln Thr Thr
930 935 940

Val Pro Tyr Glu Lys Glu Gln Ile Arg Glu Pro Lys Ile Lys Leu Lys
945 950 955 960

Asp Ile Gly Lys Phe Lys Tyr Phe Leu Gln Gln Gln Lys Val Leu Asn
965 970 975

Leu Leu Lys Tyr Asp Pro Gln His Val Trp Thr Lys Ala Glu Leu Glu
980 985 990

Glu Glu Leu Tyr Ile Gly Lys His Ser Tyr Glu Val Val Arg Arg Glu
995 1000 1005

Met Leu Leu Gln Lys Cys His Gln Leu Glu Lys His Ile Leu Glu
1010 1015 1020

Gln Phe Arg Phe Asp Gly Ser Asn His Pro Arg Glu Leu Glu Gln
1025 1030 1035

Gly Asn His Pro Asn Phe Lys Met Tyr Ile Val Asn Gly Ile Leu
1040 1045 1050

Thr Lys Arg Gly Glu Leu Glu Ile Glu Ala Glu Asn Trp Trp Leu
1055 1060 1065

Glu Leu Gly Asn Ser Lys Asn Ser Leu Asp Lys Val Glu Val Glu
1070 1075 1080

Leu Leu Thr Met Lys Thr Ile Pro Glu Gln Lys Ala Phe Leu Leu
1085 1090 1095

Ile Leu Ile Arg Asn Lys Phe Ala His Asn Gln Leu Pro Ala Asp
1100 1105 1110

Asn Tyr Phe His Tyr Ala Ser Asn Leu Met Asn Leu Lys Lys Ser
1115 1120 1125

Asp Thr Tyr Ser Leu Phe Trp Phe Thr Val Ala Asp Thr Ile Val
1130 1135 1140

Gln Glu Phe Met Ser Leu
1145

<210> 154

<211> 1133

<212> PRT

<213> Reichenbachiella agariperforans

<400> 154

Met Lys Thr Asn Pro Leu Ile Ala Ser Ser Gly Glu Lys Pro Asn Tyr
1 5 10 15

Lys Lys Phe Asn Thr Glu Ser Asp Lys Ser Phe Lys Lys Ile Phe Gln
20 25 30

Asn Lys Gly Ser Ile Ala Pro Ile Ala Glu Lys Ala Cys Lys Asn Phe
35 40 45

Glu Ile Lys Ser Lys Ser Pro Val Asn Arg Asp Gly Arg Leu His Tyr
50 55 60

Phe Ser Val Gly His Ala Phe Lys Asn Ile Asp Ser Lys Asn Val Phe
65 70 75 80

Arg Tyr Glu Leu Asp Glu Ser Gln Met Asp Met Lys Pro Thr Gln Phe
85 90 95

Leu Ala Leu Gln Lys Glu Phe Phe Asp Phe Gln Gly Ala Leu Asn Gly
100 105 110

Leu Leu Lys His Ile Arg Asn Val Asn Ser His Tyr Val His Thr Phe
115 120 125

Glu Lys Leu Glu Ile Gln Ser Ile Asn Gln Lys Leu Ile Thr Phe Leu
130 135 140

Ile Glu Ala Phe Glu Leu Ala Val Ile His Ser Tyr Leu Asn Glu Glu
145 150 155 160

Glu Leu Ser Tyr Glu Ala Tyr Lys Asp Asp Pro Gln Ser Gly Gln Lys
165 170 175

Leu Val Gln Phe Leu Cys Asp Lys Phe Tyr Pro Asn Lys Glu His Glu
180 185 190

Val Glu Glu Arg Lys Thr Ile Leu Ala Lys Asn Lys Arg Gln Ala Leu
195 200 205

Glu His Leu Leu Phe Ile Glu Val Thr Ser Asp Ile Asp Trp Lys Leu
210 215 220

Phe Glu Lys His Lys Val Phe Thr Ile Ser Asn Gly Lys Tyr Leu Ser
225 230 235 240
Phe His Ala Cys Leu Phe Leu Leu Ser Leu Phe Leu Tyr Lys Ser Glu
245 250 255

Ala Asn Gln Leu Ile Ser Lys Ile Lys Gly Phe Lys Arg Asn Asp Asp
260 265 270
Asn Gln Tyr Arg Ser Lys Arg Gln Ile Phe Thr Phe Phe Ser Lys Lys
275 280 285
Phe Thr Ser Gln Asp Val Asn Ser Glu Glu Gln His Leu Val Lys Phe
290 295 300
Arg Asp Val Ile Gln Tyr Leu Asn His Tyr Pro Ser Ala Trp Asn Lys
305 310 315 320

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

Gln Glu Gln Lys Gly Phe Ser Tyr Glu Ile Asn Thr Ser Ala Glu Ile
385 390 395 400
Lys Asp Ile Glu Thr Lys Leu Lys Ala Leu Val Leu Lys Gly Pro Leu
405 410 415
Asn Phe Lys Glu Lys Lys Glu Gln Asn Arg Leu Glu Lys Asp Leu Arg
420 425 430
Arg Glu Lys Lys Glu Gln Pro Thr Asn Arg Val Lys Glu Lys Leu Leu
435 440 445

Thr Arg Ile Gln His Asn Met Leu Tyr Val Ser Tyr Gly Arg Asn Gln
450 455 460
Asp Arg Phe Met Asp Phe Ala Ala Arg Phe Leu Ala Glu Thr Asp Tyr

465 470 475 480

Phe Gly Lys Asp Ala Lys Phe Lys Met Tyr Gln Phe Tyr Thr Ser Asp

485 490 495

Glu Gln Arg Asp His Leu Lys Glu Gln Lys Lys Glu Leu Pro Lys Lys

500 505 510

Glu Phe Glu Lys Leu Lys Tyr His Gln Ser Lys Leu Val Asp Tyr Phe

515 520 525

Thr Tyr Ala Glu Gln Gln Ala Arg Tyr Pro Asp Trp Asp Thr Pro Phe

530 535 540

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

545 550 555 560

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

565 570 575

Glu Asp Ala Leu Tyr Ser Glu Lys Arg Glu Asn Ala Gly Lys Gly Leu

580 585 590

Ile Ser Gly Tyr Phe Val His His Gln Lys Glu Leu Lys Asp Gln Leu

595 600 605

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

610 615 620

Phe Lys Lys Leu Leu Pro Lys Arg Leu Leu His Arg Tyr Ser Pro Ala

625 630 635 640

Gln Ile Asn Asp Thr Thr Glu Trp Asn Pro Met Glu Val Ile Leu Glu

645 650 655

Glu Ala Lys Ala Gln Glu Gln Arg Tyr Gln Leu Leu Leu Glu Lys Ala

660 665 670

Ile Leu His Gln Thr Glu Glu Asp Phe Leu Lys Arg Asn Lys Gly Lys

675 680 685

Gln Phe Lys Leu Arg Phe Val Arg Lys Ala Trp His Leu Met Tyr Leu

690 695 700

Lys Glu Leu Tyr Met Asn Lys Val Ala Glu His Gly His His Lys Ser

705 710 715 720

Phe His Ile Thr Lys Glu Glu Phe Asn Asp Phe Cys Arg Trp Met Phe
725 730 735
Ala Phe Asp Glu Val Pro Lys Tyr Lys Glu Tyr Leu Cys Asp Tyr Phe
740 745 750
Ser Gln Lys Gly Phe Phe Asn Asn Ala Glu Phe Lys Asp Leu Ile Glu
755 760 765

Ser Ser Thr Ser Leu Asn Asp Leu Tyr Glu Lys Thr Lys Gln Arg Phe
770 775 780
Glu Gly Trp Ser Lys Asp Leu Thr Lys Gln Ser Asp Glu Asn Lys Tyr
785 790 795 800
Leu Leu Ala Asn Tyr Glu Ser Met Leu Lys Asp Asp Met Leu Tyr Val
805 810 815
Asn Ile Ser His Phe Ile Ser Tyr Leu Glu Ser Lys Gly Lys Ile Asn
820 825 830

Arg Asn Ala His Gly His Ile Ala Tyr Lys Ala Leu Asn Asn Val Pro
835 840 845
His Leu Ile Glu Glu Tyr Tyr Tyr Lys Asp Arg Leu Ala Pro Glu Glu
850 855 860
Tyr Lys Ser His Gly Lys Leu Tyr Asn Lys Leu Lys Thr Val Lys Leu
865 870 875 880
Glu Asp Ala Leu Leu Tyr Glu Met Ala Met His Tyr Leu Ser Leu Glu
885 890 895

Pro Ala Leu Val Pro Lys Val Lys Thr Lys Val Lys Asp Ile Leu Ser
900 905 910
Ser Asn Ile Ala Phe Asp Ile Lys Asp Ala Ala Gly His His Leu Tyr
915 920 925
His Leu Leu Ile Pro Phe His Lys Ile Asp Ser Phe Val Ala Leu Ile
930 935 940
Asn His Gln Ser Gln Gln Glu Lys Asp Pro Asp Lys Thr Ser Phe Leu
945 950 955 960

Ala Lys Ile Gln Pro Tyr Leu Glu Lys Val Lys Asn Ser Lys Asp Leu

965 970 975

Lys Ala Val Tyr His Tyr Tyr Lys Asp Thr Pro His Thr Leu Arg Tyr

980 985 990

Glu Asp Leu Asn Met Ile His Ser His Ile Val Ser Gln Ser Val Gln

995 1000 1005

Phe Thr Lys Val Ala Leu Lys Leu Glu Glu Tyr Phe Ile Ala Lys

1010 1015 1020

Lys Ser Ile Thr Leu Gln Ile Ala Arg Gln Ile Ser Tyr Ser Glu

1025 1030 1035

Ile Ala Asp Leu Ser Asn Tyr Phe Thr Asp Glu Val Arg Asn Thr

1040 1045 1050

Ala Phe His Phe Asp Val Pro Glu Thr Ala Tyr Ser Met Ile Leu

1055 1060 1065

Gln Gly Ile Glu Ser Glu Phe Leu Asp Arg Glu Ile Lys Pro Gln

1070 1075 1080

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

1085 1090 1095

Cys Thr Ala Phe Leu Glu Thr Leu His Asn Asn Leu Phe Asp Arg

1100 1105 1110

Lys Asp Asp Lys Lys Glu Arg Leu Ser Lys Ala Arg Glu Arg Tyr

1115 1120 1125

Phe Glu Gln Ile Asn

1130

<210> 155

<211> 1115

<212> PRT

<213> Fusobacterium necrophorum subsp. funduliforme

<400> 155

Met Glu Lys Phe Arg Arg Gln Asn Arg Asn Ser Ile Ile Lys Ile Ile

1 5 10 15

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

20 25 30

Arg Lys Gln Ala Gln Leu Asp Thr Phe Ile Ile Lys Thr Glu Ile Val
35 40 45
Asn Asn Asp Ile Phe Ile Lys Ser Ile Ile Glu Lys Ala Arg Glu Lys
50 55 60

Tyr Arg Tyr Ser Phe Leu Phe Asp Gly Glu Glu Lys Tyr His Phe Lys
65 70 75 80
Asn Lys Ser Ser Val Glu Ile Val Lys Lys Asp Ile Phe Ser Gln Thr
85 90 95
Pro Asp Asn Met Ile Arg Asn Tyr Lys Ile Thr Leu Lys Ile Ser Glu
100 105 110
Lys Asn Pro Arg Val Val Glu Ala Glu Ile Glu Asp Leu Met Asn Ser
115 120 125

Thr Ile Leu Lys Asp Gly Arg Arg Ser Ala Arg Arg Glu Lys Ser Met
130 135 140
Thr Glu Arg Lys Leu Ile Glu Glu Lys Val Ala Lys Asn Tyr Ser Leu
145 150 155 160
Leu Ala Asn Cys Pro Met Glu Glu Val Asp Ser Ile Lys Ile Tyr Lys
165 170 175
Ile Lys Arg Phe Leu Thr Tyr Arg Ser Asn Met Leu Leu Tyr Phe Ala
180 185 190

Ser Ile Asn Ser Phe Leu Cys Glu Gly Ile Lys Gly Lys Asp Asn Glu
195 200 205
Thr Glu Glu Ile Trp His Leu Lys Asp Asn Asp Val Arg Lys Glu Lys
210 215 220
Val Arg Glu Asn Phe Lys Asn Lys Leu Ile Gln Ser Thr Glu Asn Tyr
225 230 235 240
Asn Ser Ser Leu Lys Asn Gln Ile Glu Glu Lys Glu Lys Leu Leu Arg
245 250 255

Lys Glu Phe Lys Lys Gly Ala Phe Tyr Arg Thr Ile Ile Lys Lys Leu
260 265 270
Gln Gln Glu Arg Ile Lys Glu Leu Ser Glu Lys Ser Leu Thr Glu Asp

275 280 285

Cys Glu Lys Ile Ile Lys Leu Tyr Ser Lys Leu Arg His Ser Leu Met

290 295 300

His Tyr Asp Tyr Gln Tyr Phe Glu Asn Leu Phe Glu Asn Lys Lys Asn

305 310 315 320

Asp Asp Leu Met Lys Asp Leu Asn Leu Asp Leu Phe Lys Ser Leu Pro

325 330 335

Leu Ile Arg Lys Met Lys Leu Asn Asn Lys Val Asn Tyr Leu Glu Asp

340 345 350

Gly Asp Thr Leu Phe Val Leu Gln Lys Thr Lys Lys Ala Lys Thr Leu

355 360 365

Tyr Gln Ile Tyr Asp Ala Leu Cys Glu Gln Lys Asn Gly Phe Asn Lys

370 375 380

Phe Ile Asn Asp Phe Phe Val Ser Asp Gly Glu Glu Asn Thr Val Phe

385 390 395 400

Lys Gln Ile Ile Asn Glu Lys Phe Gln Ser Glu Met Glu Phe Leu Glu

405 410 415

Lys Arg Ile Ser Glu Ser Glu Lys Lys Asn Glu Lys Leu Lys Lys Lys

420 425 430

Leu Asp Ser Met Lys Ala His Phe Arg Asn Ile Asn Ser Glu Asp Thr

435 440 445

Lys Glu Ala Tyr Phe Trp Asp Ile His Ser Ser Arg Asn Tyr Lys Thr

450 455 460

Lys Tyr Asn Glu Arg Lys Asn Leu Val Asn Glu Tyr Thr Glu Leu Leu

465 470 475 480

Gly Ser Ser Lys Glu Lys Lys Leu Leu Arg Glu Glu Ile Thr Lys Ile

485 490 495

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

500 505 510

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

515 520 525

Phe Leu Phe Cys Glu Phe Asp Gly Asn Ile Ser Lys Phe Lys Asp Glu
530 535 540
Phe Asp Ala Ser Asn Gln Glu Lys Ile Ile Gln Tyr His Lys Asn Gly
545 550 555 560
Glu Lys Tyr Leu Thr Ser Phe Leu Lys Glu Glu Glu Lys Glu Lys Phe
565 570 575

Asn Leu Glu Lys Met Gln Lys Ile Ile Gln Lys Thr Glu Glu Glu Asp
580 585 590
Trp Leu Leu Pro Glu Thr Lys Asn Asn Leu Phe Lys Phe Tyr Leu Leu
595 600 605
Thr Tyr Leu Leu Leu Pro Tyr Glu Leu Lys Gly Asp Phe Leu Gly Phe
610 615 620
Val Lys Lys His Tyr Tyr Asp Ile Lys Asn Val Asp Phe Ile Asp Glu
625 630 635 640

Asn Gln Asn Asn Ile Gln Val Ser Gln Thr Val Glu Lys Gln Glu Asp
645 650 655
Tyr Phe Tyr His Lys Ile Arg Leu Phe Glu Lys Asn Thr Lys Lys Tyr
660 665 670
Glu Ile Val Lys Tyr Ser Ile Val Pro Asn Glu Lys Leu Lys Gln Tyr
675 680 685
Phe Glu Asp Leu Gly Ile Asp Ile Lys Tyr Leu Thr Val Glu Gln Lys
690 695 700

Ser Glu Val Ser Glu Glu Lys Asn Lys Lys Val Ser Leu Lys Asn Asn
705 710 715 720
Gly Met Phe Asn Lys Thr Ile Leu Leu Phe Val Phe Lys Tyr Tyr Gln
725 730 735
Ile Ala Phe Lys Leu Phe Asn Asp Ile Glu Leu Tyr Ser Leu Phe Phe
740 745 750
Leu Arg Glu Lys Ser Gly Lys Pro Leu Glu Ile Phe Arg Lys Glu Leu
755 760 765

Glu Ser Lys Met Lys Asp Gly Tyr Leu Asn Phe Gly Gln Leu Leu Tyr

770 775 780

Val Val Tyr Glu Val Leu Val Lys Asn Lys Asp Leu Asp Lys Ile Leu

785 790 795 800

Ser Lys Lys Ile Asp Tyr Arg Lys Asp Lys Ser Phe Ser Pro Glu Ile

805 810 815

Ala Tyr Leu Arg Asn Phe Leu Ser His Leu Asn Tyr Ser Lys Phe Leu

820 825 830

Asp Asn Phe Met Lys Ile Asn Thr Asn Lys Ser Asp Glu Asn Lys Glu

835 840 845

Val Leu Ile Pro Ser Ile Lys Ile Gln Lys Met Ile Gln Phe Ile Glu

850 855 860

Lys Cys Asn Leu Gln Asn Gln Ile Asp Phe Asp Phe Asn Phe Val Asn

865 870 875 880

Asp Phe Tyr Met Arg Lys Glu Lys Met Phe Phe Ile Gln Leu Lys Gln

885 890 895

Ile Phe Pro Asp Ile Asn Ser Thr Glu Lys Gln Lys Met Asn Glu Lys

900 905 910

Glu Glu Ile Leu Arg Asn Arg Tyr His Leu Thr Asp Lys Lys Asn Glu

915 920 925

Gln Ile Lys Asp Glu His Glu Ala Gln Ser Gln Leu Tyr Glu Lys Ile

930 935 940

Leu Ser Leu Gln Lys Ile Tyr Ser Ser Asp Lys Asn Asn Phe Tyr Gly

945 950 955 960

Arg Leu Lys Glu Glu Lys Leu Leu Phe Leu Glu Lys Gln Gly Lys Lys

965 970 975

Lys Leu Ser Met Glu Glu Ile Lys Asp Lys Ile Ala Gly Asp Ile Ser

980 985 990

Asp Leu Leu Gly Ile Leu Lys Lys Glu Ile Thr Arg Asp Ile Lys Asp

995 1000 1005

Lys Leu Thr Glu Lys Phe Arg Tyr Cys Glu Glu Lys Leu Leu Asn

1010 1015 1020

Leu Ser Phe Tyr Asn His Gln Asp Lys Lys Lys Glu Glu Ser Ile
1025 1030 1035
Arg Val Phe Leu Ile Arg Asp Lys Asn Ser Asp Asn Phe Lys Phe
1040 1045 1050
Glu Ser Ile Leu Asp Asp Gly Ser Asn Lys Ile Phe Ile Ser Lys
1055 1060 1065
Asn Gly Lys Glu Ile Thr Ile Gln Cys Cys Asp Lys Val Leu Glu
1070 1075 1080

Thr Leu Ile Ile Glu Lys Asn Thr Leu Lys Ile Ser Ser Asn Gly
1085 1090 1095
Lys Ile Ile Ser Leu Ile Pro His Tyr Ser Tyr Ser Ile Asp Val
1100 1105 1110
Lys Tyr
1115

<210> 156
<211> 1111
<212> PRT
<213> Fusobacterium necrophorum
<400> 156

Met Glu Lys Phe Arg Arg Gln Asn Arg Ser Ser Ile Ile Lys Ile Ile
1 5 10 15
Ile Ser Asn Tyr Asp Thr Lys Gly Ile Lys Glu Leu Lys Val Arg Tyr
20 25 30
Arg Lys Gln Ala Gln Leu Asp Thr Phe Ile Ile Lys Thr Glu Ile Val
35 40 45
Asn Asn Asp Ile Phe Ile Lys Ser Ile Ile Glu Lys Ala Arg Glu Lys
50 55 60
Tyr Arg Tyr Ser Phe Leu Phe Asp Gly Glu Glu Lys Tyr His Phe Lys
65 70 75 80
Asn Lys Ser Ser Val Glu Ile Val Lys Lys Asp Ile Phe Ser Gln Thr
85 90 95
Pro Asp Asn Met Ile Arg Asn Tyr Lys Ile Thr Leu Lys Ile Ser Glu

100 105 110

Lys Asn Pro Arg Val Val Glu Ala Glu Ile Glu Asp Leu Met Asn Ser

115 120 125

Thr Ile Leu Lys Asp Gly Arg Arg Ser Ala Arg Arg Glu Lys Ser Met

130 135 140

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

145 150 155 160

Leu Ala Asn Cys Pro Met Glu Glu Val Asp Ser Ile Lys Ile Tyr Lys

165 170 175

Ile Lys Arg Phe Leu Thr Tyr Arg Ser Asn Met Leu Leu Tyr Phe Ala

180 185 190

Ser Ile Asn Ser Phe Leu Cys Glu Gly Ile Lys Gly Lys Asp Asn Glu

195 200 205

Thr Glu Glu Ile Trp His Leu Lys Asp Asn Asp Val Arg Lys Glu Lys

210 215 220

Val Lys Glu Asn Phe Lys Asn Lys Leu Ile Gln Ser Thr Glu Asn Tyr

225 230 235 240

Asn Ser Ser Leu Lys Asn Gln Ile Glu Glu Lys Glu Lys Leu Leu Arg

245 250 255

Lys Glu Ser Lys Lys Gly Ala Phe Tyr Arg Thr Ile Ile Lys Lys Leu

260 265 270

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

275 280 285

Cys Glu Lys Ile Ile Lys Leu Tyr Ser Glu Leu Arg His Pro Leu Met

290 295 300

His Tyr Asp Tyr Gln Tyr Phe Glu Asn Leu Phe Glu Asn Lys Glu Asn

305 310 315 320

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

325 330 335

Leu Val Arg Lys Met Lys Leu Asn Asn Lys Val Asn Tyr Leu Glu Asp

340 345 350

Asn Asp Thr Leu Phe Val Leu Gln Lys Thr Lys Lys Ala Lys Thr Leu
355 360 365
Tyr Gln Ile Tyr Asp Ala Leu Cys Glu Gln Lys Asn Gly Phe Asn Lys
370 375 380
Phe Ile Asn Asp Phe Phe Val Ser Asp Gly Glu Glu Asn Thr Val Phe
385 390 395 400
Lys Gln Ile Ile Asn Glu Lys Phe Gln Ser Glu Ile Glu Phe Leu Glu

405 410 415
Lys Arg Ile Ser Glu Ser Glu Lys Lys Asn Glu Lys Leu Lys Lys Lys
420 425 430
Leu Asp Ser Met Lys Ala His Phe Arg Asn Ile Asn Ser Glu Asp Thr
435 440 445
Lys Glu Ala Tyr Phe Trp Asp Ile His Ser Ser Arg Asn Tyr Lys Thr
450 455 460
Lys Tyr Asn Glu Arg Lys Asn Leu Val Asn Glu Tyr Thr Glu Leu Leu

465 470 475 480
Gly Ser Ser Lys Glu Lys Lys Leu Leu Arg Glu Glu Ile Thr Lys Ile
485 490 495
Asn Arg Gln Leu Leu Lys Leu Lys Gln Glu Met Glu Glu Ile Thr Lys
500 505 510
Lys Asn Ser Leu Phe Arg Leu Glu Tyr Lys Met Lys Met Ala Phe Gly
515 520 525
Phe Leu Phe Cys Glu Phe Asp Gly Asn Ile Ser Arg Phe Lys Asp Glu

530 535 540
Phe Asp Ala Ser Asn Gln Glu Lys Ile Ile Gln Tyr His Lys Asn Gly
545 550 555 560
Glu Lys Tyr Leu Thr Tyr Phe Leu Lys Glu Glu Glu Lys Glu Lys Phe
565 570 575
Asn Leu Lys Lys Leu Gln Glu Thr Ile Gln Lys Thr Gly Glu Glu Asn
580 585 590
Trp Leu Leu Pro Gln Asn Lys Asn Asn Leu Phe Lys Phe Tyr Leu Leu

595 600 605

Thr Tyr Leu Leu Leu Pro Tyr Glu Leu Lys Gly Asp Phe Leu Gly Phe

610 615 620

Val Lys Lys His Tyr Tyr Asp Ile Lys Asn Val Asp Phe Met Asp Glu

625 630 635 640

Asn Gln Ser Ser Lys Ile Ile Glu Ser Lys Glu Asp Asp Phe Tyr His

645 650 655

Lys Ile Arg Leu Phe Glu Lys Asn Thr Lys Lys Tyr Glu Ile Val Lys

660 665 670

Tyr Ser Ile Val Pro Asp Lys Lys Leu Lys Gln Tyr Phe Lys Asp Leu

675 680 685

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

690 695 700

Gly Glu Lys Asn Lys Lys Val Ser Leu Lys Asn Asn Gly Met Phe Asn

705 710 715 720

Lys Thr Ile Leu Leu Phe Val Phe Lys Tyr Tyr Gln Ile Ala Phe Lys

725 730 735

Leu Phe Asn Asp Ile Glu Leu Tyr Ser Leu Phe Phe Leu Arg Glu Lys

740 745 750

Ser Gly Lys Pro Phe Glu Val Phe Leu Lys Glu Leu Lys Asp Lys Met

755 760 765

Ile Gly Lys Gln Leu Asn Phe Gly Gln Leu Leu Tyr Val Val Tyr Glu

770 775 780

Val Leu Val Lys Asn Lys Asp Leu Ser Glu Ile Leu Ser Glu Arg Ile

785 790 795 800

Asp Tyr Arg Lys Asp Met Cys Phe Ser Ala Glu Ile Ala Asp Leu Arg

805 810 815

Asn Phe Leu Ser His Leu Asn Tyr Ser Lys Phe Leu Asp Asn Phe Met

820 825 830

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

835 840 845

Ser Ile Lys Ile Gln Lys Met Ile Lys Phe Ile Glu Glu Cys Asn Leu

850 855 860

Gln Ser Gln Ile Asp Phe Asp Phe Asn Phe Val Asn Asp Phe Tyr Met

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

Glu His Glu Ala Gln Ser Gln Leu Tyr Glu Lys Ile Leu Ser Leu Gln

930 935 940

Lys Ile Tyr Ser Ser Asp Lys Asn Asn Phe Tyr Gly Arg Leu Lys Glu

945 950 955 960

Glu Lys Leu Leu Phe Leu Glu Lys Gln Glu Lys Lys Lys Leu Ser Met

965 970 975

Glu Glu Ile Lys Asp Lys Ile Ala Gly Asp Ile Ser Asp Leu Leu Gly

980 985 990

Ile Leu Lys Lys Glu Ile Thr Arg Asp Ile Lys Asp Lys Leu Thr Glu

995 1000 1005

Lys Phe Arg Tyr Cys Glu Glu Lys Leu Leu Asn Leu Ser Phe Tyr

1010 1015 1020

Asn His Gln Asp Lys Lys Lys Glu Glu Ser Ile Arg Val Phe Leu

1025 1030 1035

Ile Arg Asp Lys Asn Ser Asp Asn Phe Lys Phe Glu Ser Ile Leu

1040 1045 1050

Asp Asp Gly Ser Asn Lys Ile Phe Ile Ser Lys Asn Gly Lys Glu

1055 1060 1065

Ile Thr Ile Gln Cys Cys Asp Lys Val Leu Glu Thr Leu Ile Ile

1070 1075 1080

Glu Lys Asn Thr Leu Lys Ile Ser Ser Asn Gly Lys Ile Ile Ser

1085 1090 1095

Leu Ile Pro His Tyr Ser Tyr Ser Ile Asp Val Lys Tyr

1100 1105 1110

<210> 157

<211> 1110

<212> PRT

<213> Fusobacterium necrophorum

<400> 157

Met Lys Val Arg Tyr Arg Lys Gln Ala Gln Leu Asp Thr Phe Ile Ile
1 5 10 15

Lys Thr Glu Ile Val Asn Asn Asp Ile Phe Ile Lys Ser Ile Ile Glu
20 25 30

Lys Ala Arg Glu Lys Tyr Arg Tyr Ser Phe Leu Phe Asp Gly Glu Glu
35 40 45

Lys Tyr His Phe Lys Asn Lys Ser Ser Val Glu Ile Val Lys Asn Asp
50 55 60

Ile Phe Ser Gln Thr Pro Asp Asn Met Ile Arg Asn Tyr Lys Ile Thr
65 70 75 80

Leu Lys Ile Ser Glu Lys Asn Pro Arg Val Val Glu Ala Glu Ile Glu
85 90 95

Asp Leu Met Asn Ser Thr Ile Leu Lys Asp Gly Arg Arg Ser Ala Arg
100 105 110

Arg Glu Lys Ser Met Thr Glu Arg Lys Leu Ile Glu Glu Lys Val Ala
115 120 125

Glu Asn Tyr Ser Leu Leu Ala Asn Cys Pro Ile Glu Glu Val Asp Ser
130 135 140

Ile Lys Ile Tyr Lys Ile Lys Arg Phe Leu Thr Tyr Arg Ser Asn Met
145 150 155 160

Leu Leu Tyr Phe Ala Ser Ile Asn Ser Phe Leu Cys Glu Gly Ile Lys
165 170 175

Gly Lys Asp Asn Glu Thr Glu Glu Ile Trp His Leu Lys Asp Asn Asp

180 185 190

Val Arg Lys Glu Lys Val Lys Glu Asn Phe Lys Asn Lys Leu Ile Gln

195 200 205

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

210 215 220

Glu Lys Leu Ser Ser Lys Glu Phe Lys Lys Gly Ala Phe Tyr Arg Thr

225 230 235 240

Ile Ile Lys Lys Leu Gln Gln Glu Arg Ile Lys Glu Leu Ser Glu Lys

245 250 255

Ser Leu Thr Glu Asp Cys Glu Lys Ile Ile Lys Leu Tyr Ser Glu Leu

260 265 270

Arg His Pro Leu Met His Tyr Asp Tyr Gln Tyr Phe Glu Asn Leu Phe

275 280 285

Glu Asn Lys Glu Asn Ser Glu Leu Thr Lys Asn Leu Asn Leu Asp Ile

290 295 300

Phe Lys Ser Leu Pro Leu Val Arg Lys Met Lys Leu Asn Asn Lys Val

305 310 315 320

Asn Tyr Leu Glu Asp Asn Asp Thr Leu Phe Val Leu Gln Lys Thr Lys

325 330 335

Lys Ala Lys Thr Leu Tyr Gln Ile Tyr Asp Ala Leu Cys Glu Gln Lys

340 345 350

Asn Gly Phe Asn Lys Phe Ile Asn Asp Phe Phe Val Ser Asp Gly Glu

355 360 365

Glu Asn Thr Val Phe Lys Gln Ile Ile Asn Glu Lys Phe Gln Ser Glu

370 375 380

Met Glu Phe Leu Glu Lys Arg Ile Ser Glu Ser Glu Lys Lys Asn Glu

385 390 395 400

Lys Leu Lys Lys Lys Leu Asp Ser Met Lys Ala His Phe Arg Asn Ile

405 410 415

Asn Ser Glu Asp Thr Lys Glu Ala Tyr Phe Trp Asp Ile His Ser Ser

420 425 430

Arg Asn Tyr Lys Thr Lys Tyr Asn Glu Arg Lys Asn Leu Val Asn Glu
435 440 445
Tyr Thr Lys Leu Leu Gly Ser Ser Lys Glu Lys Lys Leu Leu Arg Glu
450 455 460
Glu Ile Thr Lys Ile Asn Arg Gln Leu Leu Lys Leu Lys Gln Glu Met
465 470 475 480
Glu Glu Ile Thr Lys Lys Asn Ser Leu Phe Arg Leu Glu Tyr Lys Met
485 490 495

Lys Ile Ala Phe Gly Phe Leu Phe Cys Glu Phe Asp Gly Asn Ile Ser
500 505 510
Lys Phe Lys Asp Glu Phe Asp Ala Ser Asn Gln Glu Lys Ile Ile Gln
515 520 525
Tyr His Lys Asn Gly Glu Lys Tyr Leu Thr Ser Phe Leu Lys Glu Glu
530 535 540
Glu Lys Glu Lys Phe Asn Leu Glu Lys Met Gln Lys Ile Ile Gln Lys
545 550 555 560

Thr Glu Glu Glu Asp Trp Leu Leu Pro Glu Thr Lys Asn Asn Leu Phe
565 570 575
Lys Phe Tyr Leu Leu Thr Tyr Leu Leu Leu Pro Tyr Glu Leu Lys Gly
580 585 590
Asp Phe Leu Gly Phe Val Lys Lys His Tyr Tyr Asp Ile Lys Asn Val
595 600 605
Asp Phe Met Asp Glu Asn Gln Asn Asn Ile Gln Val Ser Gln Thr Val
610 615 620

Glu Lys Gln Glu Asp Tyr Phe Tyr His Lys Ile Arg Leu Phe Glu Lys
625 630 635 640
Asn Thr Lys Lys Tyr Glu Ile Val Lys Tyr Ser Ile Val Pro Asn Glu
645 650 655
Lys Leu Lys Gln Tyr Phe Glu Asp Leu Gly Ile Asp Ile Lys Tyr Leu
660 665 670
Thr Gly Ser Val Glu Ser Gly Glu Lys Trp Leu Gly Glu Asn Leu Gly

675 680 685

Ile Asp Ile Lys Tyr Leu Thr Val Glu Gln Lys Ser Glu Val Ser Glu
690 695 700
Glu Lys Asn Lys Lys Val Ser Leu Lys Asn Asn Gly Met Phe Asn Lys
705 710 715 720
Thr Ile Leu Leu Phe Val Phe Lys Tyr Tyr Gln Ile Ala Phe Lys Leu
725 730 735
Phe Asn Asp Ile Glu Leu Tyr Ser Leu Phe Phe Leu Arg Glu Lys Ser
740 745 750

Glu Lys Pro Phe Glu Val Phe Leu Glu Glu Leu Lys Asp Lys Met Ile
755 760 765
Gly Lys Gln Leu Asn Phe Gly Gln Leu Leu Tyr Val Val Tyr Glu Val
770 775 780
Leu Val Lys Asn Lys Asp Leu Asp Lys Ile Leu Ser Lys Lys Ile Asp
785 790 795 800
Tyr Arg Lys Asp Lys Ser Phe Ser Pro Glu Ile Ala Tyr Leu Arg Asn
805 810 815

Phe Leu Ser His Leu Asn Tyr Ser Lys Phe Leu Asp Asn Phe Met Lys
820 825 830
Ile Asn Thr Asn Lys Ser Asp Glu Asn Lys Glu Val Leu Ile Pro Ser
835 840 845
Ile Lys Ile Gln Lys Met Ile Gln Phe Ile Glu Lys Cys Asn Leu Gln
850 855 860
Asn Gln Ile Asp Phe Asp Phe Asn Phe Val Asn Asp Phe Tyr Met Arg
865 870 875 880

Lys Glu Lys Met Phe Phe Ile Gln Leu Lys Gln Ile Phe Pro Asp Ile
885 890 895
Asn Ser Thr Glu Lys Gln Lys Lys Ser Glu Lys Glu Glu Ile Leu Arg
900 905 910
Lys Arg Tyr His Leu Ile Asn Lys Lys Asn Glu Gln Ile Lys Asp Glu
915 920 925

His Glu Ala Gln Ser Gln Leu Tyr Glu Lys Ile Leu Ser Leu Gln Lys
930 935 940

Ile Phe Ser Cys Asp Lys Asn Asn Phe Tyr Arg Arg Leu Lys Glu Glu
945 950 955 960
Lys Leu Leu Phe Leu Glu Lys Gln Gly Lys Lys Lys Ile Ser Met Lys
965 970 975
Glu Ile Lys Asp Lys Ile Ala Ser Asp Ile Ser Asp Leu Leu Gly Ile
980 985 990
Leu Lys Lys Glu Ile Thr Arg Asp Ile Lys Asp Lys Leu Thr Glu Lys
995 1000 1005

Phe Arg Tyr Cys Glu Glu Lys Leu Leu Asn Ile Ser Phe Tyr Asn
1010 1015 1020
His Gln Asp Lys Lys Lys Glu Glu Gly Ile Arg Val Phe Leu Ile
1025 1030 1035
Arg Asp Lys Asn Ser Asp Asn Phe Lys Phe Glu Ser Ile Leu Asp
1040 1045 1050
Asp Gly Ser Asn Lys Ile Phe Ile Ser Lys Asn Gly Lys Glu Ile
1055 1060 1065

Thr Ile Gln Cys Cys Asp Lys Val Leu Glu Thr Leu Met Ile Glu
1070 1075 1080
Lys Asn Thr Leu Lys Ile Ser Ser Asn Gly Lys Ile Ile Ser Leu
1085 1090 1095
Ile Pro His Tyr Ser Tyr Ser Ile Asp Val Lys Tyr
1100 1105 1110

<210> 158
<211> 668
<212> PRT
<213> Fusobacterium necrophorum subsp. funduliforme
<400> 158

Met Thr Glu Lys Lys Ser Ile Ile Phe Lys Asn Lys Ser Ser Val Glu
1 5 10 15

Ile Val Lys Lys Asp Ile Phe Ser Gln Thr Pro Asp Asn Met Ile Arg
20 25 30
Asn Tyr Lys Ile Thr Leu Lys Ile Ser Glu Lys Asn Pro Arg Val Val
35 40 45
Glu Ala Glu Ile Glu Asp Leu Met Asn Ser Thr Ile Leu Lys Asp Gly
50 55 60
Arg Arg Ser Ala Arg Arg Glu Lys Ser Met Thr Glu Arg Lys Leu Ile

65 70 75 80
Glu Glu Lys Val Ala Glu Asn Tyr Ser Leu Leu Ala Asn Cys Pro Met
85 90 95
Glu Glu Val Asp Ser Ile Lys Ile Tyr Lys Ile Lys Arg Phe Leu Thr
100 105 110
Tyr Arg Ser Asn Met Leu Leu Tyr Phe Ala Ser Ile Asn Ser Phe Leu
115 120 125
Cys Glu Gly Ile Lys Gly Lys Asp Asn Glu Thr Glu Glu Ile Trp His

130 135 140
Leu Lys Asp Asn Asp Val Arg Lys Glu Lys Val Lys Glu Asn Phe Lys
145 150 155 160
Asn Lys Leu Ile Gln Ser Thr Glu Asn Tyr Asn Ser Ser Leu Lys Asn
165 170 175
Gln Ile Glu Glu Lys Glu Lys Leu Leu Arg Lys Glu Ser Lys Lys Gly
180 185 190
Ala Phe Tyr Arg Thr Ile Ile Lys Lys Leu Gln Gln Glu Arg Ile Lys

195 200 205
Glu Leu Ser Glu Lys Ser Leu Thr Glu Asp Cys Glu Lys Ile Ile Lys
210 215 220
Leu Tyr Ser Glu Leu Arg His Pro Leu Met His Tyr Asp Tyr Gln Tyr
225 230 235 240
Phe Glu Asn Leu Phe Glu Asn Lys Glu Asn Ser Glu Leu Thr Lys Asn
245 250 255
Leu Asn Leu Asp Ile Phe Lys Ser Leu Pro Leu Val Arg Lys Met Lys

260 265 270

Leu Asn Asn Lys Val Asn Tyr Leu Glu Asp Asn Asp Thr Leu Phe Val

275 280 285

Leu Gln Lys Thr Lys Lys Ala Lys Thr Leu Tyr Gln Ile Tyr Asp Ala

290 295 300

Leu Cys Glu Gln Lys Asn Gly Phe Asn Lys Phe Ile Asn Asp Phe Phe

305 310 315 320

Val Ser Asp Gly Glu Glu Asn Thr Val Phe Lys Gln Ile Ile Asn Glu

325 330 335

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

340 345 350

Glu Lys Lys Asn Glu Lys Leu Lys Lys Lys Phe Asp Ser Met Lys Ala

355 360 365

His Phe His Asn Ile Asn Ser Glu Asp Thr Lys Glu Ala Tyr Phe Trp

370 375 380

Asp Ile His Ser Ser Ser Asn Tyr Lys Thr Lys Tyr Asn Glu Arg Lys

385 390 395 400

Asn Leu Val Asn Glu Tyr Thr Glu Leu Leu Gly Ser Ser Lys Glu Lys

405 410 415

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

420 425 430

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

435 440 445

Leu Glu Tyr Lys Met Lys Ile Ala Phe Gly Phe Leu Phe Cys Glu Phe

450 455 460

Asp Gly Asn Ile Ser Lys Phe Lys Asp Glu Phe Asp Ala Ser Asn Gln

465 470 475 480

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

485 490 495

Phe Leu Lys Glu Glu Glu Lys Glu Lys Phe Asn Leu Glu Lys Met Gln

500 505 510

Lys Ile Ile Gln Lys Thr Glu Glu Glu Asp Trp Leu Leu Pro Glu Thr

515 520 525

Lys Asn Asn Leu Phe Lys Phe Tyr Leu Leu Thr Tyr Leu Leu Leu Pro

530 535 540

Tyr Glu Leu Lys Gly Asp Phe Leu Gly Phe Val Lys Lys His Tyr Tyr

545 550 555 560

Asp Ile Lys Asn Val Asp Phe Met Asp Glu Asn Gln Asn Asn Ile Gln

565 570 575

Val Ser Gln Thr Val Glu Lys Gln Glu Asp Tyr Phe Tyr His Lys Ile

580 585 590

Arg Leu Phe Glu Lys Asn Thr Lys Lys Tyr Glu Ile Val Lys Tyr Ser

595 600 605

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

610 615 620

Asp Ile Lys Tyr Leu Thr Gly Ser Val Glu Ser Gly Glu Lys Trp Leu

625 630 635 640

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

645 650 655

Ser Glu Val Ser Glu Glu Lys Ile Lys Lys Phe Leu

660 665

<210> 159

<211> 1121

<212> PRT

<213> Fusobacterium perfoetens

<400> 159

Met Gly Lys Pro Asn Arg Ser Ser Ile Ile Lys Ile Ile Ile Ser Asn

1 5 10 15

Tyr Asp Asn Lys Gly Ile Lys Glu Val Lys Val Arg Tyr Asn Lys Gln

20 25 30

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

35 40 45

Phe Ile Leu Tyr Ser Ile Val Asp Lys Ala Arg Glu Lys Tyr Arg Tyr
50 55 60
Ser Phe Glu Ile Asp Lys Thr Asn Ile Asn Lys Asn Glu Ile Leu Ile
65 70 75 80
Ile Lys Lys Asp Ile Tyr Ser Asn Lys Glu Asp Lys Val Ile Arg Lys
85 90 95

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

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

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

Glu Lys Asp Lys Ile Glu Leu Ser Asp Glu Glu Ile Arg Glu Asp Ile

290 295 300

Glu Lys Ile Leu Lys Ile Tyr Ser Asp Leu Arg His Lys Leu Met His

305 310 315 320

Tyr Asn Tyr Gln Tyr Phe Glu Asn Leu Phe Glu Asn Lys Lys Ile Ser

325 330 335

Lys Glu Lys Asn Glu Asp Val Asn Leu Thr Glu Leu Leu Asp Leu Asn

340 345 350

Leu Phe Arg Tyr Leu Pro Leu Val Arg Gln Leu Lys Leu Glu Asn Lys

355 360 365

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

370 375 380

Asp Ser Ala Ile Lys Tyr Tyr Ser Tyr Tyr Asn Phe Leu Cys Glu Gln

385 390 395 400

Lys Asn Gly Phe Asn Asn Phe Ile Asn Ser Phe Phe Ser Asn Asp Gly

405 410 415

Glu Glu Asn Lys Ser Phe Lys Glu Lys Ile Asn Leu Ser Leu Glu Lys

420 425 430

Glu Ile Glu Ile Met Glu Lys Glu Thr Asn Glu Lys Ile Lys Glu Ile

435 440 445

Asn Lys Asn Glu Leu Gln Leu Met Lys Glu Gln Lys Glu Leu Gly Thr

450 455 460

Ala Tyr Val Leu Asp Ile His Ser Leu Asn Asp Tyr Lys Ile Ser His

465 470 475 480

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

485 490 495

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

500 505 510

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

515 520 525

Leu Lys Tyr Lys Leu Gln Val Ala Tyr Gly Phe Leu Met Glu Glu Tyr

530 535 540

Lys Gly Asn Ile Lys Lys Phe Lys Asp Glu Phe Asp Ile Ser Lys Glu
545 550 555 560
Lys Ile Lys Ser Tyr Lys Ser Lys Gly Glu Lys Tyr Leu Glu Val Lys
565 570 575
Ser Glu Lys Lys Tyr Ile Thr Lys Ile Leu Asn Ser Ile Glu Asp Ile
580 585 590
His Asn Ile Thr Trp Leu Lys Asn Gln Glu Glu Asn Asn Leu Phe Lys
595 600 605

Phe Tyr Val Leu Thr Tyr Ile Leu Leu Pro Phe Glu Phe Arg Gly Asp
610 615 620
Phe Leu Gly Phe Val Lys Lys His Tyr Tyr Asp Ile Lys Asn Val Glu
625 630 635 640
Phe Leu Asp Glu Asn Asn Asp Arg Leu Thr Pro Glu Gln Leu Glu Lys
645 650 655
Met Lys Asn Asp Ser Phe Phe Asn Lys Ile Arg Leu Phe Glu Lys Asn
660 665 670

Ser Lys Lys Tyr Asp Ile Leu Lys Glu Ser Ile Leu Thr Ser Glu Arg
675 680 685
Ile Gly Lys Tyr Phe Ser Leu Leu Asn Thr Gly Ala Lys Tyr Phe Glu
690 695 700
Tyr Gly Gly Glu Glu Asn Arg Gly Ile Phe Asn Lys Asn Ile Ile Ile
705 710 715 720
Pro Ile Phe Lys Tyr Tyr Gln Ile Val Leu Lys Leu Tyr Asn Asp Val
725 730 735

Glu Leu Ala Met Leu Leu Thr Leu Ser Glu Ser Asp Glu Lys Asp Ile
740 745 750
Asn Lys Ile Lys Glu Leu Val Thr Leu Lys Glu Lys Val Ser Pro Lys
755 760 765
Lys Ile Asp Tyr Glu Lys Lys Tyr Lys Phe Ser Val Leu Leu Asp Cys
770 775 780
Phe Asn Arg Ile Ile Asn Leu Gly Lys Lys Asp Phe Leu Ala Ser Glu

785 790 795 800

Glu Val Lys Glu Val Ala Lys Thr Phe Thr Asn Leu Ala Tyr Leu Arg
805 810 815
Asn Lys Ile Cys His Leu Asn Tyr Ser Lys Phe Ile Asp Asp Leu Leu
820 825 830
Thr Ile Asp Thr Asn Lys Ser Thr Thr Asp Ser Glu Gly Lys Leu Leu
835 840 845
Ile Asn Asp Arg Ile Arg Lys Leu Ile Lys Phe Ile Arg Glu Asn Asn
850 855 860

Gln Lys Met Asn Ile Ser Ile Asp Tyr Asn Tyr Ile Asn Asp Tyr Tyr
865 870 875 880
Met Lys Lys Glu Lys Phe Ile Phe Gly Gln Arg Lys Gln Ala Lys Thr
885 890 895
Ile Ile Asp Ser Gly Lys Lys Ala Asn Lys Arg Asn Lys Ala Glu Glu
900 905 910
Leu Leu Lys Met Tyr Arg Val Lys Lys Glu Asn Ile Asn Leu Ile Tyr
915 920 925

Glu Leu Ser Lys Lys Leu Asn Glu Leu Thr Lys Ser Glu Leu Phe Leu
930 935 940
Leu Asp Lys Lys Leu Leu Lys Asp Ile Asp Phe Thr Asp Val Lys Ile
945 950 955 960
Lys Asn Lys Ser Phe Phe Glu Leu Lys Asn Asp Val Lys Glu Val Ala
965 970 975
Asn Ile Lys Gln Ala Leu Gln Lys His Ser Ser Glu Leu Ile Gly Ile
980 985 990

Tyr Lys Lys Glu Val Ile Met Ala Ile Lys Arg Ser Ile Val Ser Lys
995 1000 1005
Leu Ile Tyr Asp Glu Glu Lys Val Leu Ser Ile Ile Ile Tyr Asp
1010 1015 1020
Lys Thr Asn Lys Lys Tyr Glu Asp Phe Leu Leu Glu Ile Arg Arg
1025 1030 1035

Glu Arg Asp Ile Asn Lys Phe Gln Phe Leu Ile Asp Glu Lys Lys
1040 1045 1050

Glu Lys Leu Gly Tyr Glu Lys Ile Ile Glu Thr Lys Glu Lys Lys
1055 1060 1065
Lys Val Val Val Lys Ile Gln Asn Asn Ser Glu Leu Val Ser Glu
1070 1075 1080
Pro Arg Ile Ile Lys Asn Lys Asp Lys Lys Lys Ala Lys Thr Pro
1085 1090 1095
Glu Glu Ile Ser Lys Leu Gly Ile Leu Asp Leu Thr Asn His Tyr
1100 1105 1110

Cys Phe Asn Leu Lys Ile Thr Leu
1115 1120
<210> 160
<211> 897
<212> PRT
<213> Fusobacterium ulcerans
<400> 160
Met Glu Asn Lys Gly Asn Asn Lys Lys Ile Asp Phe Asp Glu Asn Tyr
1 5 10 15
Asn Ile Leu Val Ala Gln Ile Lys Glu Tyr Phe Thr Lys Glu Ile Glu
20 25 30
Asn Tyr Asn Asn Arg Ile Asp Asn Ile Ile Asp Lys Lys Glu Leu Leu
35 40 45

Lys Tyr Ser Glu Lys Lys Glu Glu Ser Glu Lys Asn Lys Lys Leu Glu
50 55 60
Glu Leu Asn Lys Leu Lys Ser Gln Lys Leu Lys Ile Leu Thr Asp Glu
65 70 75 80
Glu Ile Lys Ala Asp Val Ile Lys Ile Ile Lys Ile Phe Ser Asp Leu
85 90 95
Arg His Ser Leu Met His Tyr Glu Tyr Lys Tyr Phe Glu Asn Leu Phe
100 105 110

Glu Asn Lys Lys Asn Glu Glu Leu Ala Glu Leu Leu Asn Leu Asn Leu
115 120 125
Phe Lys Asn Leu Thr Leu Leu Arg Gln Met Lys Ile Glu Asn Lys Thr
130 135 140
Asn Tyr Leu Glu Gly Arg Glu Glu Phe Asn Ile Ile Gly Lys Asn Ile
145 150 155 160
Lys Ala Lys Glu Val Leu Gly His Tyr Asn Leu Leu Ala Glu Gln Lys
165 170 175

Asn Gly Phe Asn Asn Phe Ile Asn Ser Phe Phe Val Gln Asp Gly Thr
180 185 190
Glu Asn Leu Glu Phe Lys Lys Leu Ile Asp Glu His Phe Val Asn Ala
195 200 205
Lys Lys Arg Leu Glu Arg Asn Ile Lys Lys Ser Lys Lys Leu Glu Lys
210 215 220
Glu Leu Glu Lys Met Glu Gln His Tyr Gln Arg Leu Asn Cys Ala Tyr
225 230 235 240

Val Trp Asp Ile His Thr Ser Thr Thr Tyr Lys Lys Leu Tyr Asn Lys
245 250 255
Arg Lys Ser Leu Ile Glu Glu Tyr Asn Lys Gln Ile Asn Glu Ile Lys
260 265 270
Asp Lys Glu Val Ile Thr Ala Ile Asn Val Glu Leu Leu Arg Ile Lys
275 280 285
Lys Glu Met Glu Glu Ile Thr Lys Ser Asn Ser Leu Phe Arg Leu Lys
290 295 300

Tyr Lys Met Gln Ile Ala Tyr Ala Phe Leu Glu Ile Glu Phe Gly Gly
305 310 315 320
Asn Ile Ala Lys Phe Lys Asp Glu Phe Asp Cys Ser Lys Met Glu Glu
325 330 335
Val Gln Lys Tyr Leu Lys Lys Gly Val Lys Tyr Leu Lys Tyr Tyr Lys
340 345 350
Asp Lys Glu Ala Gln Lys Asn Tyr Glu Phe Pro Phe Glu Glu Ile Phe

355 360 365

Glu Asn Lys Asp Thr His Asn Glu Glu Trp Leu Glu Asn Thr Ser Glu
370 375 380
Asn Asn Leu Phe Lys Phe Tyr Ile Leu Thr Tyr Leu Leu Leu Pro Met
385 390 395 400
Glu Phe Lys Gly Asp Phe Leu Gly Val Val Lys Lys His Tyr Tyr Asp
405 410 415
Ile Lys Asn Val Asp Phe Thr Asp Glu Ser Glu Lys Glu Leu Ser Gln
420 425 430

Val Gln Leu Asp Lys Met Ile Gly Asp Ser Phe Phe His Lys Ile Arg
435 440 445
Leu Phe Glu Lys Asn Thr Lys Arg Tyr Glu Ile Ile Lys Tyr Ser Ile
450 455 460
Leu Thr Ser Asp Glu Ile Lys Arg Tyr Phe Arg Leu Leu Glu Leu Asp
465 470 475 480
Val Pro Tyr Phe Glu Tyr Glu Lys Gly Thr Asp Glu Ile Gly Ile Phe
485 490 495

Asn Lys Asn Ile Ile Leu Thr Ile Phe Lys Tyr Tyr Gln Ile Ile Phe
500 505 510
Arg Leu Tyr Asn Asp Leu Glu Ile His Gly Leu Phe Asn Ile Ser Ser
515 520 525
Asp Leu Asp Lys Ile Leu Arg Asp Leu Lys Ser Tyr Gly Asn Lys Asn
530 535 540
Ile Asn Phe Arg Glu Phe Leu Tyr Val Ile Lys Gln Asn Asn Asn Ser
545 550 555 560

Ser Thr Glu Glu Glu Tyr Arg Lys Ile Trp Glu Asn Leu Glu Ala Lys
565 570 575
Tyr Leu Arg Leu His Leu Leu Thr Pro Glu Lys Glu Glu Ile Lys Thr
580 585 590
Lys Thr Lys Glu Glu Leu Glu Lys Leu Asn Glu Ile Ser Asn Leu Arg
595 600 605

Asn Gly Ile Cys His Leu Asn Tyr Lys Glu Ile Ile Glu Glu Ile Leu
610 615 620

Lys Thr Glu Ile Ser Glu Lys Asn Lys Glu Ala Thr Leu Asn Glu Lys
625 630 635 640
Ile Arg Lys Val Ile Asn Phe Ile Lys Glu Asn Glu Leu Asp Lys Val
645 650 655
Glu Leu Gly Phe Asn Phe Ile Asn Asp Phe Phe Met Lys Lys Glu Gln
660 665 670
Phe Met Phe Gly Gln Ile Lys Gln Val Lys Glu Gly Asn Ser Asp Ser
675 680 685

Ile Thr Thr Glu Arg Glu Arg Lys Glu Lys Asn Asn Lys Lys Leu Lys
690 695 700
Glu Thr Tyr Glu Leu Asn Cys Asp Asn Leu Ser Glu Phe Tyr Glu Thr
705 710 715 720
Ser Asn Asn Leu Arg Glu Arg Ala Asn Ser Ser Ser Leu Leu Glu Asp
725 730 735
Ser Ala Phe Leu Lys Lys Ile Gly Leu Tyr Lys Val Lys Asn Asn Lys
740 745 750

Val Asn Ser Lys Val Lys Asp Glu Glu Lys Arg Ile Glu Asn Ile Lys
755 760 765
Arg Lys Leu Leu Lys Asp Ser Ser Asp Ile Met Gly Met Tyr Lys Ala
770 775 780
Glu Val Val Lys Lys Leu Lys Glu Lys Leu Ile Leu Ile Phe Lys His
785 790 795 800
Asp Glu Glu Lys Arg Ile Tyr Val Thr Val Tyr Asp Thr Ser Lys Ala
805 810 815

Val Pro Glu Asn Ile Ser Lys Glu Ile Leu Val Lys Arg Asn Asn Ser
820 825 830
Lys Glu Glu Tyr Phe Phe Glu Asp Asn Asn Lys Lys Tyr Val Thr Glu
835 840 845
Tyr Tyr Thr Leu Glu Ile Thr Glu Thr Asn Glu Leu Lys Val Ile Pro

850 855 860

Ala Lys Lys Leu Glu Gly Lys Glu Phe Lys Thr Glu Lys Asn Lys Glu
865 870 875 880

Asn Lys Leu Met Leu Asn Asn His Tyr Cys Phe Asn Val Lys Ile Ile
885 890 895

Tyr

<210> 161
<211> 1111
<212> PRT
<213> Anaerosalibacter species
<400> 161

Met Lys Ser Gly Arg Arg Glu Lys Ala Lys Ser Asn Lys Ser Ser Ile
1 5 10 15

Val Arg Val Ile Ile Ser Asn Phe Asp Asp Lys Gln Val Lys Glu Ile
20 25 30

Lys Val Leu Tyr Thr Lys Gln Gly Gly Ile Asp Val Ile Lys Phe Lys
35 40 45

Ser Thr Glu Lys Asp Glu Lys Gly Arg Met Lys Phe Asn Phe Asp Cys
50 55 60

Ala Tyr Asn Arg Leu Glu Glu Glu Glu Phe Asn Ser Phe Gly Gly Lys
65 70 75 80

Gly Lys Gln Ser Phe Phe Val Thr Thr Asn Glu Asp Leu Thr Glu Leu
85 90 95

His Val Thr Lys Arg His Lys Thr Thr Gly Glu Ile Ile Lys Asp Tyr
100 105 110

Thr Ile Gln Gly Lys Tyr Thr Pro Ile Lys Gln Asp Arg Thr Lys Val
115 120 125

Thr Val Ser Ile Thr Asp Asn Lys Asp His Phe Asp Ser Asn Asp Leu
130 135 140

Gly Asp Lys Ile Arg Leu Ser Arg Ser Leu Thr Gln Tyr Thr Asn Arg
145 150 155 160

Ile Leu Leu Asp Ala Asp Val Met Lys Asn Tyr Arg Glu Ile Val Cys

165 170 175

Ser Asp Ser Glu Lys Val Asp Glu Thr Ile Asn Ile Asp Ser Gln Glu

180 185 190

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

195 200 205

Tyr Tyr Gln Met Ile Asn Asn Phe Leu Leu His Tyr Asp Gly Glu Glu

210 215 220

Asp Lys Gly Gly Asn Asp Ser Ile Asn Leu Ile Asn Glu Ile Trp Lys

225 230 235 240

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

245 250 255

Tyr Lys Ser Ile Glu Lys Ser Ile Asn Gln Tyr Ile Leu Asn His Asn

260 265 270

Thr Glu Val Glu Ser Gly Asp Lys Glu Lys Lys Ile Asp Ile Ser Glu

275 280 285

Glu Arg Ile Lys Glu Asp Leu Lys Lys Thr Phe Ile Leu Phe Ser Arg

290 295 300

Leu Arg His Tyr Met Val His Tyr Asn Tyr Lys Phe Tyr Glu Asn Leu

305 310 315 320

Tyr Ser Gly Lys Asn Phe Ile Ile Tyr Asn Lys Asp Lys Ser Lys Ser

325 330 335

Arg Arg Phe Ser Glu Leu Leu Asp Leu Asn Ile Phe Lys Glu Leu Ser

340 345 350

Lys Ile Lys Leu Val Lys Asn Arg Ala Val Ser Asn Tyr Leu Asp Lys

355 360 365

Lys Thr Thr Ile His Val Leu Asn Lys Asn Ile Asn Ala Ile Lys Leu

370 375 380

Leu Asp Ile Tyr Arg Asp Ile Cys Glu Thr Lys Asn Gly Phe Asn Asn

385 390 395 400

Phe Ile Asn Asn Met Met Thr Ile Ser Gly Glu Glu Asp Lys Glu Tyr

405 410 415

Lys Glu Met Val Thr Lys His Phe Asn Glu Asn Met Asn Lys Leu Ser

420 425 430

Ile Tyr Leu Glu Asn Phe Lys Lys His Ser Asp Phe Lys Thr Asn Asn

435 440 445

Lys Lys Lys Glu Thr Tyr Asn Leu Leu Lys Gln Glu Leu Asp Glu Gln

450 455 460

Lys Lys Leu Arg Leu Trp Phe Asn Ala Pro Tyr Val Tyr Asp Ile His

465 470 475 480

Ser Ser Lys Lys Tyr Lys Glu Leu Tyr Val Glu Arg Lys Lys Tyr Val

485 490 495

Asp Ile His Ser Lys Leu Ile Glu Ala Gly Ile Asn Asn Asp Asn Lys

500 505 510

Lys Lys Leu Asn Glu Ile Asn Val Lys Leu Cys Glu Leu Asn Thr Glu

515 520 525

Met Lys Glu Met Thr Lys Leu Asn Ser Lys Tyr Arg Leu Gln Tyr Lys

530 535 540

Leu Gln Leu Ala Phe Gly Phe Ile Leu Glu Glu Phe Asn Leu Asp Ile

545 550 555 560

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

565 570 575

Lys Phe Met Glu Lys Arg Glu Thr Tyr Leu Ser Lys Ser Leu Asp Arg

580 585 590

Arg Asp Asn Arg Phe Lys Lys Leu Ile Lys Asp Tyr Lys Phe Arg Asp

595 600 605

Thr Glu Asp Ile Phe Cys Ser Asp Arg Glu Asn Asn Leu Val Lys Leu

610 615 620

Tyr Ile Leu Met Tyr Ile Leu Leu Pro Val Glu Ile Arg Gly Asp Phe

625 630 635 640

Leu Gly Phe Val Lys Lys Asn Tyr Tyr Asp Leu Lys His Val Asp Phe

645 650 655

Ile Asp Lys Arg Asn Asn Asp Asn Lys Asp Thr Phe Phe His Asp Leu
660 665 670
Arg Leu Phe Glu Lys Asn Val Lys Arg Leu Glu Val Thr Ser Tyr Ser

675 680 685
Leu Ser Asp Gly Phe Leu Gly Lys Lys Ser Arg Glu Lys Phe Gly Lys
690 695 700
Glu Leu Glu Lys Phe Ile Tyr Lys Asn Val Ser Ile Ala Leu Pro Thr
705 710 715 720
Asn Ile Asp Ile Lys Glu Phe Asn Lys Ser Leu Val Leu Pro Met Met
725 730 735
Lys Asn Tyr Gln Ile Ile Phe Lys Leu Leu Asn Asp Ile Glu Ile Ser

740 745 750
Ala Leu Phe Leu Ile Ala Lys Lys Glu Gly Asn Glu Gly Ser Ile Thr
755 760 765
Phe Lys Lys Val Ile Asp Lys Val Arg Lys Glu Asp Met Asn Gly Asn
770 775 780
Ile Asn Phe Ser Gln Val Met Lys Met Ala Leu Asn Glu Lys Val Asn
785 790 795 800
Cys Gln Ile Arg Asn Ser Ile Ala His Ile Asn Met Lys Gln Leu Tyr

805 810 815
Ile Glu Pro Leu Asn Ile Tyr Ile Asn Asn Asn Gln Asn Lys Lys Thr
820 825 830
Ile Ser Glu Gln Met Glu Glu Ile Ile Asp Ile Cys Ile Thr Lys Gly
835 840 845
Leu Thr Gly Lys Glu Leu Asn Lys Asn Ile Ile Asn Asp Tyr Tyr Met
850 855 860
Lys Lys Glu Lys Leu Val Phe Asn Leu Lys Leu Arg Lys Arg Asn Asn

865 870 875 880
Leu Val Ser Ile Asp Ala Gln Gln Lys Asn Met Lys Glu Lys Ser Ile
885 890 895
Leu Asn Lys Tyr Asp Leu Asn Tyr Lys Asp Glu Asn Leu Asn Ile Lys

900 905 910

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

915 920 925

Lys Glu Thr Thr Glu Gly Glu Ser Asn Tyr Lys Asn Ala Leu Ser Lys

930 935 940

Asp Ile Leu Leu Leu Asn Gly Ile Ile Arg Lys Asn Ile Asn Phe Lys

945 950 955 960

Ile Lys Glu Met Ile Leu Gly Ile Ile Gln Gln Asn Glu Tyr Arg Tyr

965 970 975

Val Asn Ile Asn Ile Tyr Asp Lys Ile Arg Lys Glu Asp His Asn Ile

980 985 990

Asp Leu Lys Ile Asn Asn Lys Tyr Ile Glu Ile Ser Cys Tyr Glu Asn

995 1000 1005

Lys Ser Asn Glu Ser Thr Asp Glu Arg Ile Asn Phe Lys Ile Lys

1010 1015 1020

Tyr Met Asp Leu Lys Val Lys Asn Glu Leu Leu Val Pro Ser Cys

1025 1030 1035

Tyr Glu Asp Ile Tyr Ile Lys Lys Lys Ile Asp Leu Glu Ile Arg

1040 1045 1050

Tyr Ile Glu Asn Cys Lys Val Val Tyr Ile Asp Ile Tyr Tyr Lys

1055 1060 1065

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

1070 1075 1080

Lys Phe Asn Lys Asp Val Lys Lys Asn Asn Gln Lys Val Asn Leu

1085 1090 1095

Glu Ser Asn Tyr Ile Gln Asn Ile Lys Phe Ile Val Ser

1100 1105 1110

<210> 162

<211> 45

<212> DNA

<213> Artificial Sequence

<220

><223> Synthetic Oligonucleotide

<400> 162

catagaatgt tctaaaccat cctgcggcct ctactctgca ttcaa 45

<210> 163

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 163

catagaatgt tcgaaaccat cctgcggcct ctactctgca ttcaa 45

<210> 164

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 164

catagaatgg tcgaaaccat cctgcggcct ctactctgca ttcaa 45

<210> 165

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 165

catagaatgt tcaaaaccat cctgcggcct ctactctgca ttcaa 45

<210> 166

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 166

catagaatga tcaaaaccat cctgcggcct ctactctgca ttcaa 45

<210> 167

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 167

ggtttctcca tcaattacca cttgcttcct gtaggaatcc tctatt 46

<210> 168

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 168

ggtttctcca tcaatgacca cttgcttcct gtaggaatcc tctatt 46

<210> 169

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 169

ggtttctcca tcaaggacca cttgcttcct gtaggaatcc tctatt 46

<210> 170

<211> 46

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 170

ggtttctcca tcaataacca cttgcttcct gtaggaatcc tctatt 46

<210> 171

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 171

ggtttctcca tcaaaaacca cttgcttcct gtaggaatcc tctatt 46

<210> 172

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 172

gcctttctct cctgtagcca aggccacaaa attatccact gttttt 46

<210> 173

<211> 46

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 173

Gly Cys Cys Thr Thr Thr Cys Thr Cys Thr Cys Cys Thr Gly Gly Ala
1 5 10 15

Gly Cys Cys Ala Ala Gly Gly Cys Cys Ala Cys Ala Ala Ala Ala Thr
20 25 30

Thr Ala Thr Cys Cys Ala Cys Thr Gly Thr Thr Thr Thr Thr
35 40 45

<210> 174

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 174

gcctttctct cctgaagcca aggccacaaa attatccact gttttt 46

<210> 175

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 175

gcatagaatg ttctaaacca tcctgcggcc tctactctgc attcaa 46

<210> 176

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 176

gcatagaatg ttctaaacga tcctgcggcc tctactctgc attcaa 46

<210> 177

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 177

gcatagaatg ttctaaacta tcctgcggcc tctactctgc attcaa 46

<210> 178

<211> 46

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 178

Gly Cys Ala Thr Ala Gly Ala Ala Thr Gly Thr Thr Cys Thr Ala Ala

1 5 10 15

Ala Cys Ala Ala Thr Cys Cys Thr Gly Cys Gly Gly Cys Cys Thr Cys

20 25 30

Thr Ala Cys Thr Cys Thr Gly Cys Ala Thr Thr Cys Ala Ala

35 40 45

<210> 179

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 179

gcatcctgcg gcctctactc tgcattcaat t 31

<210> 180

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 180

gaccatcctg cggcctctac tctgcattca a 31

<210> 181

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 181

gaaaccatcc tgcggcctct actctgcatt c 31

<210> 182

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 182

gctaaaccat cctgcggcct ctactctgca t 31

<210> 183

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 183

gttctaaacc atcctgcggc ctctactctg c 31

<210> 184

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 184

gtgttctaaa ccatcctgcg gcctctactc t 31

<210> 185

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 185

gaatgttcta aaccatcctg cggcctctac t 31

<210> 186

<211> 31

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 186

gagaatgttc taaaccatcc tgcggcctct a 31

<210> 187

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 187

gatagaatgt tctaaaccat cctgcggcct c 31

<210> 188

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 188

gccatagaat gttctaaacc atcctgcggc c 31

<210> 189

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 189

gttccataga atgttctaaa ccatcctgcg g 31

<210> 190

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 190

gctttccata gaatgttcta aaccatcctg c 31

<210> 191

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 191

gctctttcca tagaatgttc taaaccatcc t 31

<210> 192

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 192

gatctctttc catagaatgt tctaaaccat c 31

<210> 193

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 193

ggaatctctt tccatagaat gttctaaacc a 31

<210> 194

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 194

gcatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc a 51

<210> 195

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 195

gaccatcctg cggcctctac tctgcattca attacatact gacacattcg g 51

<210> 196

<211> 51

<212> DNA

<213> Artificial SequenceSynthetic Oligonucleotide

<220><223> Synthetic Oligonucleotide

<400> 196

gaaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt c 51

<210> 197

<211> 51

<212> DNA

<213> Artificial SequenceSynthetic Oligonucleotide

<220><223> Synthetic Oligonucleotide

<400> 197

gctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca t 51

<210> 198

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 198

gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca c 51

<210> 199

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 199

gtgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga c 51

<210> 200

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 200

gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact g 51

<210> 201

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 201

gagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata c 51

<210> 202

<211> 51

<212

> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 202

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca t 51

<210> 203

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 203

gccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta c 51

<210> 204

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 204

gttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat t 51

<210> 205

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 205

gctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca a 51

<210> 206

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 206

gctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt c 51

<210> 207

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 207

gatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca t 51

<210> 208

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 208

ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg c 51

<210> 209

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 209

gtggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc t 51

<210> 210

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 210

gactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac t 51

<210> 211

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 211

ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct a 51

<210> 212

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 212

gtggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct c 51

<210> 213

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 213

gcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc c 51

<210> 214

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 214

gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg g 51

<210> 215

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 215

gggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg c 51

<210> 216

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 216

gcaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc t 51

<210> 217

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 217

gaccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat c 51

<210> 218

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 218

ggtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc a 51

<210> 219

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 219

gcatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc aacatgtttt 60

tcctggttta t 71

<210> 220

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 220

gaccatcctg cggcctctac tctgcattca attacatact gacacattcg gcaacatgtt 60

tttcctggtt t 71

<210> 221

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 221

gaaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt cggcaacatg 60

tttttcctgg t 71

<210> 222

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Oligonucleotide

<400> 222

gctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca ttcggcaaca 60

tgtttttcct g 71

<210> 223

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 223

gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca cattcggcaa 60

catgtttttc c 71

<210> 224

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 224

gtgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc 60

aacatgtttt t 71

<210> 225

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 225

gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact gacacattcg 60

gcaacatgtt t 71

<210

> 226

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 226

gagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt 60

cggcaacatg t 71

<210> 227

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 227

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca 60

ttcggcaaca t 71

<210> 228

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 228

gccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca 60

cattcggcaa c 71

<210> 229

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 229

gttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga 60

cacattcggc a 71

<210> 230

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 230

gctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact 60

gacacattcg g 71

<210> 231

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 231

gctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata 60

ctgacacatt c 71

<210> 232

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 232

gatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca 60

tactgacaca t 71

<210> 233

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 233

ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta 60

catactgaca c 71

<210> 234

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 234

gtggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat 60

tacatactga c 71

<210> 235

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 235

gactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca 60

attacatact g 71

<210> 236

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 236

ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt 60

caattacata c 71

<210> 237

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 237

gtggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca 60

ttcaattaca t 71

<210> 238

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 238

gcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg 60

cattcaatta c 71

<210> 239

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 239

gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc 60

tgcattcaat t 71

<210> 240

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 240

gggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac 60

tctgcattca a 71

<210> 241

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 241

gcaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct 60

actctgcatt c 71

<210> 242

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 242

gaccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct 60

ctactctgca t 71

<210> 243

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 243

ggtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc 60

ctctactctg c 71

<210> 244

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 244

gatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg 60

gcctctactc t 71

<210> 245

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 245

ggtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg 60

cggcctctac t 71

<210> 246

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 246

gacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc 60

tgcggcctct a 71

<210> 247

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 247

gacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat 60

cctgcggcct c 71

<210> 248

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 248

gcaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc 60

atcctgcggc c 71

<210> 249

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 249

gcccaacacg tatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa 60

ccatcctgcg g 71

<210> 250

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 250

ggacccaaca cgtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta 60

aaccatcctg c 71

<210> 251

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 251

gttgacccaa cacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc 60

taaaccatcc t 71

<210> 252

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 252

gccttgaccc aacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt 60

tctaaaccat c 71

<210> 253

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 253

gttccttgac ccaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat 60

gttctaaacc a 71

<210> 254

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 254

gcatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc aacatgtttt 60

tcctggttta ttttcacaca gtcca 85

<210> 255

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 255

gaccatcctg cggcctctac tctgcattca attacatact gacacattcg gcaacatgtt 60

tttcctggtt tattttcaca cagtc 85

<210> 256

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 256

gaaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt cggcaacatg 60

tttttcctgg tttattttca cacag 85

<210> 257

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 257

gctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca ttcggcaaca 60

tgtttttcct ggtttatttt cacac 85

<210> 258

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 258

gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca cattcggcaa 60

catgtttttc ctggtttatt ttcac 85

<210> 259

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 259

gtgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc 60

aacatgtttt tcctggttta ttttc 85

<210> 260

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 260

gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact gacacattcg 60

gcaacatgtt tttcctggtt tattt 85

<210> 261

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 261

gagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt 60

cggcaacatg tttttcctgg tttat 85

<210> 262

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 262

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca 60

ttcggcaaca tgtttttcct ggttt 85

<210> 263

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 263

gccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca 60

cattcggcaa catgtttttc ctggt 85

<210> 264

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 264

gttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga 60

cacattcggc aacatgtttt tcctg 85

<210> 265

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 265

gctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact 60

gacacattcg gcaacatgtt tttcc 85

<210> 266

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 266

gctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata 60

ctgacacatt cggcaacatg ttttt 85

<210> 267

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 267

gatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca 60

tactgacaca ttcggcaaca tgttt 85

<210> 268

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 268

ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta 60

catactgaca cattcggcaa catgt 85

<210> 269

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 269

gtggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat 60

tacatactga cacattcggc aacat 85

<210> 270

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 270

gactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca 60

attacatact gacacattcg gcaac 85

<210> 271

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 271

ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt 60

caattacata ctgacacatt cggca 85

<210> 272

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 272

gtggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca 60

ttcaattaca tactgacaca ttcgg 85

<210> 273

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 273

gcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg 60

cattcaatta catactgaca cattc 85

<210> 274

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 274

gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc 60

tgcattcaat tacatactga cacat 85

<210> 275

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 275

gggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac 60

tctgcattca attacatact gacac 85

<210> 276

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 276

gcaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct 60

actctgcatt caattacata ctgac 85

<210> 277

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 277

gaccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct 60

ctactctgca ttcaattaca tactg 85

<210> 278

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 278

ggtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc 60

ctctactctg cattcaatta catac 85

<210> 279

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 279

gatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg 60

gcctctactc tgcattcaat tacat 85

<210> 280

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 280

ggtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg 60

cggcctctac tctgcattca attac 85

<210> 281

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 281

gacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc 60

tgcggcctct actctgcatt caatt 85

<210> 282

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 282

gacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat 60

cctgcggcct ctactctgca ttcaa 85

<210> 283

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 283

gcaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc 60

atcctgcggc ctctactctg cattc 85

<210> 284

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 284

gcccaacacg tatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa 60

ccatcctgcg gcctctactc tgcat 85

<210> 285

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 285

ggacccaaca cgtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta 60

aaccatcctg cggcctctac tctgc 85

<210> 286

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 286

gttgacccaa cacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc 60

taaaccatcc tgcggcctct actct 85

<210> 287

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 287

gccttgaccc aacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt 60

tctaaaccat cctgcggcct ctact 85

<210> 288

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 288

gttccttgac ccaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat 60

gttctaaacc atcctgcggc ctcta 85

<210> 289

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 289

gggttccttg acccaacacg tatgtaccag gttcctggaa ctggaatctc tttccataga 60

atgttctaaa ccatcctgcg gcctc 85

<210> 290

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 290

gttggttcct tgacccaaca cgtatgtacc aggttcctgg aactggaatc tctttccata 60

gaatgttcta aaccatcctg cggcc 85

<210> 291

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 291

gccttggttc cttgacccaa cacgtatgta ccaggttcct ggaactggaa tctctttcca 60

tagaatgttc taaaccatcc tgcgg 85

<210> 292

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 292

ggcccttggt tccttgaccc aacacgtatg taccaggttc ctggaactgg aatctctttc 60

catagaatgt tctaaaccat cctgc 85

<210> 293

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 293

gccgcccttg gttccttgac ccaacacgta tgtaccaggt tcctggaact ggaatctctt 60

tccatagaat gttctaaacc atcct 85

<210> 294

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 294

gcgccgccct tggttccttg acccaacacg tatgtaccag gttcctggaa ctggaatctc 60

tttccataga atgttctaaa ccatc 85

<210> 295

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 295

ggtcgccgcc cttggttcct tgacccaaca cgtatgtacc aggttcctgg aactggaatc 60

tctttccata gaatgttcta aacca 85

<210> 296

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 296

Gly Ser Gly Gly Gly Gly Ser

1 5

<210> 297

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 297

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu

1 5 10 15

Ala Ala Ala Lys

20

<210> 298

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 298

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly

1 5 10 15

Gly Ser

<210> 299

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 299

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser

1 5 10 15

<210> 300

<211> 28

<212> RNA

<213> Artificial Sequence

<220

><223> Synthetic Oligonucleotide

<400> 300

gccagcuuuc cgggcauugg cuuccauc 28

<210> 301

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 301

gccagctttc cgggcattgg cttccatc 28

<210> 302

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 302

acccaggaat ctcaggaatg tcgacgat 28

<210> 303

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 303

agggttttcc cagtcacgac gttgtaaa 28

<210> 304

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 304

gggcattggc ttccatctct ttgagcacct 30

<210> 305

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 305

gugcagccag cuuuccgggc auuggcuucc 30

<210> 306

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 306

gcagccagcu uuccgggcau uggcuuccau 30

<210> 307

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 307

ggcuuccauc ucuuugagca ccuccagcgg 30

<210> 308

<211> 30

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 308

ggaaugucga cgaucgccuc gccuaugccg 30

<210> 309

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 309

gaaugucgac gaucgccucg ccuaugccgc 30

<210> 310

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 310

gaccugugcg augaacugcu ccaugggcuc 30

<210> 311

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 311

guguggcagc guccugggau gaacuucuuc 30

<210> 312

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 312

guggcagcgu ccugggauga acuucuucau 30

<210> 313

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 313

gcuucuugcc gggcaacuuc ccgcggucag 30

<210> 314

<211> 31

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 314

gcaggguuuu cccagucacg acguuguaaa a 31

<210> 315

<211> 31

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 315

gcaggguuuu cccagucacg acguuguaaa a 31

<210> 316

<211> 28

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 316

acccaggaau cucaggaaug ucgacgau 28

<210> 317

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 317

cagctttccg ggcattggct t 21

<210> 318

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 318

ccgcuggagg ugcucaaaga gauggaagcc 30

<210> 319

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 319

gccagctttc cgggcattgg cttccatc 28

<210> 320

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 320

acccaggaat ctcaggaatg tcgacgat 28

<210> 321

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 321

gggcattggc ttccatctct ttgagcacct 30

<210> 322

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 322

gaaugucgac gaucgccucg ccuaugccgc 30

<210> 323

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 323

caaggcactc ttgcctacgc caccagct 28

<210> 324

<211> 28

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 324

tcatattcgt ccacaaaatg attctgaa 28

<210> 325

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 325

attatttatg gcaaatacac aaagaaag 28

<210> 326

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 326

gaatatcttc aaatgattta gtattatt 28

<210> 327

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 327

accataggta catcttcaga gtccttaa 28

<210> 328

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 328

gtcaaggcac tcttgcctac gccaccagct 30

<210> 329

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 329

gatcatattc gtccacaaaa tgattctgaa 30

<210> 330

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 330

gtattattta tggcaaatac acaaagaaag 30

<210> 331

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 331

gtgaatatct tcaaatgatt tagtattatt 30

<210> 332

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 332

ggaccatagg tacatcttca gagtccttaa 30

<210> 333

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 333

aagagtgcct tgacgataca gcctcgaggc tgtatcgtca aggcactctt 50

<210> 334

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 334

aatcattttg tggacgaata tctcgagata ttcgtccaca aaatgatt 48

<210> 335

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 335

aaataatact aaatcatttg actcgagtca aatgatttag tattattt 48

<210> 336

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 336

aataatacta aatcatttga actcgagttc aaatgattta gtattatt 48

<210> 337

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 337

aaggactctg aagatgtacc tctcgagagg tacatcttca gagtcctt 48

<210> 338

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 338

gagatcaggg caaacagaac tttgactccc 30

<210> 339

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 339

ggatgcagat cagggcaaac agaactttga 30

<210> 340

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 340

gcacagcgat gcagatcagg gcaaacagaa 30

<210> 341

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 341

gctcggccac agcgatgcag atcagggcaa 30

<210> 342

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 342

ggggcttggc ctcggccaca gcgatgcaga 30

<210> 343

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 343

gtgggcttgg cctcggccac agcgatgcag 30

<210> 344

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 344

gtctcggtgg gcttggcctc ggccacagcg 30

<210> 345

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 345

gttcgttgtt ctcggtgggc ttggcctcgg 30

<210> 346

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 346

ggaagtcttc gttgttctcg gtgggcttgg 30

<210> 347

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 347

gatgttgaag tcttcgttgt tctcggtggg 30

<210> 348

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 348

gcggccacga tgttgaagtc ttcgttgttc 30

<210> 349

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 349

gtggccacgg ccacgatgtt gaagtcttcg 30

<210> 350

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 350

ggttgctggc cacggccacg atgttgaagt 30

<210> 351

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 351

gtcgcgaagt tgctggccac ggccacgatg 30

<210> 352

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 352

gccgtggtcg cgaagttgct ggccacggcc 30

<210> 353

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 353

gcgagatccg tggtcgcgaa gttgctggcc 30

<210> 354

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 354

gcagcatcga gatccgtggt cgcgaagttg 30

<210> 355

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 355

gggtcagcat cgagatccgt ggtcgcgaag 30

<210> 356

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 356

gcttcccgcg gtcagcatcg agatccgtgg 30

<210> 357

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 357

ggggcaactt cccgcggtca gcatcgagat 30

<210> 358

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 358

gtcttgccgg gcaacttccc gcggtcagca 30

<210> 359

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 359

ggcagcttct tgccgggcaa cttcccgcgg 30

<210> 360

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 360

gccagcggca gcttcttgcc gggcaacttc 30

<210> 361

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 361

gcacctccag cggcagcttc ttgccgggca 30

<210> 362

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 362

gctttgagca cctccagcgg cagcttcttg 30

<210> 363

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 363

gcatctcttt gagcacctcc agcggcagct 30

<210> 364

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 364

gtccatctct ttgagcacct ccagcggcag 30

<210> 365

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 365

gggcattggc ttccatctct ttgagcacct 30

<210> 366

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 366

gtccgggcat tggcttccat ctctttgagc 30

<210> 367

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 367

ggccagcttt ccgggcattg gcttccatct 30

<210> 368

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 368

gggtgcagcc agctttccgg gcattggctt 30

<210> 369

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 369

gagcccctgg tgcagccagc tttccgggca 30

<210> 370

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 370

gatcagacag cccctggtgc agccagcttt 30

<210> 371

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 371

ggcagatcag acagcccctg gtgcagccag 30

<210> 372

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 372

gacaggcaga tcagacagcc cctggtgcag 30

<210> 373

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 373

gtgatgtggg acaggcagat cagacagccc 30

<210> 374

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 374

gacttgatgt gggacaggca gatcagacag 30

<210> 375

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 375

ggggcgtgca cttgatgtgg gacaggcaga 30

<210> 376

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 376

gcttcatctt gggcgtgcac ttgatgtggg 30

<210> 377

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 377

gtgaacttct tcatcttggg cgtgcacttg 30

<210> 378

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 378

gggatgaact tcttcatctt gggcgtgcac 30

<210> 379

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 379

gtgggatgaa cttcttcatc ttgggcgtgc 30

<210> 380

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 380

gggcagcgtc ctgggatgaa cttcttcatc 30

<210> 381

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 381

gggtgtggca gcgtcctggg atgaacttct 30

<210> 382

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 382

gttcgtaggt gtggcagcgt cctgggatga 30

<210> 383

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 383

gcgccttcgt aggtgtggca gcgtcctggg 30

<210> 384

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 384

gtctttgtcg ccttcgtagg tgtggcagcg 30

<210> 385

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 385

gctttgtcgc cttcgtaggt gtggcagcgt 30

<210> 386

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 386

gtgccgccct gtgcggactc tttgtcgcct 30

<210> 387

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 387

gtatgccgcc ctgtgcggac tctttgtcgc 30

<210> 388

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 388

gcctcgccta tgccgccctg tgcggactct 30

<210> 389

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 389

ggatcgcctc gcctatgccg ccctgtgcgg 30

<210> 390

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 390

gatgtcgacg atcgcctcgc ctatgccgcc 30

<210> 391

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 391

gcaggaatgt cgacgatcgc ctcgcctatg 30

<210> 392

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 392

gaatctcagg aatgtcgacg atcgcctcgc 30

<210> 393

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 393

gcccaggaat ctcaggaatg tcgacgatcg 30

<210> 394

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 394

gccttgaacc caggaatctc aggaatgtcg 30

<210> 395

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 395

gccaagtcct tgaacccagg aatctcagga 30

<210> 396

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 396

gtgggctcca agtccttgaa cccaggaatc 30

<210> 397

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 397

gccatgggct ccaagtcctt gaacccagga 30

<210> 398

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 398

ggaactgctc catgggctcc aagtccttga 30

<210> 399

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 399

gtgcgatgaa ctgctccatg ggctccaagt 30

<210> 400

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 400

ggacctgtgc gatgaactgc tccatgggct 30

<210> 401

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 401

gacagatcga cctgtgcgat gaactgctcc 30

<210> 402

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 402

gacacacaga tcgacctgtg cgatgaactg 30

<210> 403

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 403

gtgcagtcca cacacagatc gacctgtgcg 30

<210> 404

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 404

gccagttgtg cagtccacac acagatcgac 30

<210> 405

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 405

gggcagccag ttgtgcagtc cacacacaga 30

<210> 406

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 406

gtttgaggca gccagttgtg cagtccacac 30

<210> 407

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 407

gaagcccttt gaggcagcca gttgtgcagt 30

<210> 408

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 408

gcacgttggc aagccctttg aggcagccag 30

<210> 409

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 409

gactgcacgt tggcaagccc tttgaggcag 30

<210> 410

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 410

gggtcagaac actgcacgtt ggcaagccct 30

<210> 411

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 411

gcaggtcaga acactgcacg ttggcaagcc 30

<210> 412

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 412

gagcaggtca gaacactgca cgttggcaag 30

<210> 413

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 413

ggccacttct tgagcaggtc agaacactgc 30

<210> 414

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 414

gcggcagcca cttcttgagc aggtcagaac 30

<210> 415

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 415

gtgcggcagc cacttcttga gcaggtcaga 30

<210> 416

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 416

gagcgttgcg gcagccactt cttgagcagg 30

<210> 417

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 417

gaaaggtcgc acagcgttgc ggcagccact 30

<210> 418

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 418

gctggcaaag gtcgcacagc gttgcggcag 30

<210> 419

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 419

gggcaaaggt cgcacagcgt tgcggcagcc 30

<210> 420

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 420

gtggatcttg ctggcaaagg tcgcacagcg 30

<210> 421

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 421

gcacctggcc ctggatcttg ctggcaaagg 30

<210> 422

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 422

gtggccctgg atcttgctgg caaaggtcgc 30

<210> 423

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 423

gtgatcttgt ccacctggcc ctggatcttg 30

<210> 424

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 424

gccccttgat cttgtccacc tggccctgga 30

<210> 425

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 425

gcccttgatc ttgtccacct ggccctggat 30

<210> 426

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 426

gccttgatct tgtccacctg gccctggatc 30

<210> 427

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 427

ggcaaaggtc gcacagcgtt gcggcagcca 30

<210> 428

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 428

gcaaaggtcg cacagcgttg cggcagccac 30

<210> 429

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 429

gaaggtcgca cagcgttgcg gcagccactt 30

<210> 430

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 430

gaggtcgcac agcgttgcgg cagccacttc 30

<210> 431

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 431

ggtaatgcct ggcttgtcga cgcatagtct g 31

<210> 432

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 432

gggaaccttg gccgttataa agtctgacca g 31

<210> 433

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 433

ggagggtgag aatttagaac caagattgtt g 31

<210> 434

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 434

gagtcctggc aatgaacagt ggcgcagtag 30

<210> 435

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 435

gggtgccaca gctgctatca atacattctc 30

<210> 436

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 436

gttacatact gacacattcg gcaacatgtt 30

<210> 437

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 437

gtatgtacca ggttcctgga actggaatct 30

<210> 438

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 438

gccttggttc catccaggtt ctccagggtg 30

<210> 439

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 439

gcagtgatgg gattctcagt agcttgagcg 30

<210> 440

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 440

gagcctggca tctcaacaac agcgatggtg 30

<210> 441

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 441

gtgtctgggg cgattcttac agatcttcct 30

<210> 442

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 442

gctggatctg aagtgaagtc tgtatcttcc 30

<210> 443

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 443

ggcaacgtca tcaggatttc catagagtgg 30

<210> 444

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 444

gaggcgcagg agatggtgta gtagtagaag 30

<210> 445

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 445

gagggaccct ggaattggta tcttgctttg 30

<210> 446

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 446

ggtaagagtc aacattcctg tgtgaaacct 30

<210> 447

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 447

gaccagaatc tgttttccat caacaatgag 30

<210> 448

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 448

gatggctgta gtcagtatgt caccatcttg 30

<210> 449

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 449

gtaccatcga atggatctct aatatgtacg 30

<210> 450

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 450

gagatcacag gctccttcag catcaaaaga 30

<210> 451

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 451

gctttgaccg gcgaagagac tattgcagag 30

<210> 452

<211> 30

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 452

gcccctcagg caatactcgt acatgcatcg 30

<210> 453

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 453

gctggtactt ctagggtgtc tccatgcttt 30

<210> 454

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 454

ggtaatgcct ggcttgtcga cgcatagtct g 31

<210> 455

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 455

gggaaccttg gccgttataa agtctgacca g 31

<210> 456

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 456

ggagggtgag aatttagaac caagattgtt g 31

<210> 457

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 457

gcatcctgcg gcctctactc tgcattcaat t 31

<210> 458

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 458

gaccatcctg cggcctctac tctgcattca a 31

<210> 459

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 459

gaaaccatcc tgcggcctct actctgcatt c 31

<210> 460

<211> 31

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 460

gctaaaccat cctgcggcct ctactctgca t 31

<210> 461

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 461

gttctaaacc atcctgcggc ctctactctg c 31

<210> 462

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 462

gtgttctaaa ccatcctgcg gcctctactc t 31

<210> 463

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 463

gaatgttcta aaccatcctg cggcctctac t 31

<210> 464

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 464

gagaatgttc taaaccatcc tgcggcctct a 31

<210> 465

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 465

gatagaatgt tctaaaccat cctgcggcct c 31

<210> 466

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 466

gccatagaat gttctaaacc atcctgcggc c 31

<210> 467

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 467

gttccataga atgttctaaa ccatcctgcg g 31

<210> 468

<211> 31

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 468

gctttccata gaatgttcta aaccatcctg c 31

<210> 469

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 469

gctctttcca tagaatgttc taaaccatcc t 31

<210> 470

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 470

gatctctttc catagaatgt tctaaaccat c 31

<210> 471

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 471

ggaatctctt tccatagaat gttctaaacc a 31

<210> 472

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 472

gcatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc a 51

<210> 473

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 473

gaccatcctg cggcctctac tctgcattca attacatact gacacattcg g 51

<210> 474

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 474

gaaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt c 51

<210> 475

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 475

gctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca t 51

<210> 476

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 476

gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca c 51

<210> 477

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 477

gtgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga c 51

<210> 478

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 478

gaatgttcta aaccatcctg cggcctctac tctgcattca ttacatact g 51

<210> 479

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 479

gagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata c 51

<210> 480

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 480

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca t 51

<210> 481

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 481

gccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta c 51

<210> 482

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 482

gttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat t 51

<210> 483

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 483

gctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca a 51

<210> 484

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 484

gctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt c 51

<210> 485

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 485

gatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca t 51

<210> 486

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 486

ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg c 51

<210> 487

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 487

gtggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc t 51

<210> 488

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 488

gactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac t 51

<210> 489

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 489

ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct a 51

<210> 490

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 490

gtggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct c 51

<210> 491

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 491

gcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc c 51

<210> 492

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 492

gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg g 51

<210> 493

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 493

gggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg c 51

<210> 494

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 494

gcaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc t 51

<210> 495

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 495

gaccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat c 51

<210> 496

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 496

ggtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc a 51

<210> 497

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 497

gcatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc aacatgtttt 60

tcctggttta t 71

<210> 498

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 498

gaccatcctg cggcctctac tctgcattca attacatact gacacattcg gcaacatgtt 60

tttcctggtt t 71

<210> 499

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 499

gaaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt cggcaacatg 60

tttttcctgg t 71

<210> 500

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 500

gctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca ttcggcaaca 60

tgtttttcct g 71

<210> 501

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Oligonucleotide

<400> 501

gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca cattcggcaa 60

catgtttttc c 71

<210> 502

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 502

gtgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc 60

aacatgtttt t 71

<210> 503

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 503

gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact gacacattcg 60

gcaacatgtt t 71

<210> 504

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 504

gagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt 60

cggcaacatg t 71

<210

> 505

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 505

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca 60

ttcggcaaca t 71

<210> 506

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 506

gccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca 60

cattcggcaa c 71

<210> 507

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 507

gttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga 60

cacattcggc a 71

<210> 508

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 508

gctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact 60

gacacattcg g 71

<210> 509

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 509

gctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata 60

ctgacacatt c 71

<210> 510

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 510

gatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca 60

tactgacaca t 71

<210> 511

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 511

ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta 60

catactgaca c 71

<210> 512

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 512

gtggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat 60

tacatactga c 71

<210> 513

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 513

gactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca 60

attacatact g 71

<210> 514

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 514

ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt 60

caattacata c 71

<210> 515

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 515

gtggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca 60

ttcaattaca t 71

<210> 516

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 516

gcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg 60

cattcaatta c 71

<210> 517

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 517

gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc 60

tgcattcaat t 71

<210> 518

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 518

gggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac 60

tctgcattca a 71

<210> 519

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 519

gcaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct 60

actctgcatt c 71

<210> 520

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 520

gaccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct 60

ctactctgca t 71

<210> 521

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 521

ggtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc 60

ctctactctg c 71

<210> 522

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 522

gatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg 60

gcctctactc t 71

<210> 523

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 523

ggtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg 60

cggcctctac t 71

<210> 524

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 524

gacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc 60

tgcggcctct a 71

<210> 525

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 525

gacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat 60

cctgcggcct c 71

<210> 526

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 526

gcaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc 60

atcctgcggc c 71

<210> 527

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 527

gcccaacacg tatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa 60

ccatcctgcg g 71

<210> 528

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 528

ggacccaaca cgtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta 60

aaccatcctg c 71

<210> 529

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 529

gttgacccaa cacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc 60

taaaccatcc t 71

<210> 530

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 530

gccttgaccc aacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt 60

tctaaaccat c 71

<210> 531

<211> 71

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 531

gttccttgac ccaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat 60

gttctaaacc a 71

<210> 532

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 532

gcatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc aacatgtttt 60

tcctggttta ttttcacaca gtcca 85

<210> 533

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 533

gaccatcctg cggcctctac tctgcattca attacatact gacacattcg gcaacatgtt 60

tttcctggtt tattttcaca cagtc 85

<210> 534

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 534

gaaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt cggcaacatg 60

ttttcctgg tttattttca cacag 85

<210> 535

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 535

gctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca ttcggcaaca 60

tgtttttcct ggtttatttt cacac 85

<210> 536

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 536

gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca cattcggcaa 60

catgtttttc ctggtttatt ttcac 85

<210> 537

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 537

gtgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga cacattcggc 60

aacatgtttt tcctggttta ttttc 85

<210> 538

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 538

gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact gacacattcg 60

gcaacatgtt tttcctggtt tattt 85

<210> 539

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 539

gagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata ctgacacatt 60

cggcaacatg tttttcctgg tttat 85

<210> 540

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 540

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca tactgacaca 60

ttcggcaaca tgtttttcct ggttt 85

<210> 541

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 541

gccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta catactgaca 60

cattcggcaa catgtttttc ctggt 85

<210> 542

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 542

gttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat tacatactga 60

cacattcggc aacatgtttt tcctg 85

<210> 543

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 543

gctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca attacatact 60

gacacattcg gcaacatgtt tttcc 85

<210> 544

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 544

gctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt caattacata 60

ctgacacatt cggcaacatg ttttt 85

<210> 545

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 545

gatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca 60

tactgacaca ttcggcaaca tgttt 85

<210> 546

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 546

ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg cattcaatta 60

catactgaca cattcggcaa catgt 85

<210> 547

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 547

gtggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc tgcattcaat 60

tacatactga cacattcggc aacat 85

<210> 548

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 548

gactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac tctgcattca 60

attacatact gacacattcg gcaac 85

<210> 549

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 549

ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct actctgcatt 60

caattacata ctgacacatt cggca 85

<210> 550

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 550

gtggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct ctactctgca 60

ttcaattaca tactgacaca ttcgg 85

<210> 551

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 551

gcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc ctctactctg 60

cattcaatta catactgaca cattc 85

<210> 552

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 552

gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg gcctctactc 60

tgcattcaat tacatactga cacat 85

<210> 553

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 553

gggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg cggcctctac 60

tctgcattca attacatact gacac 85

<210> 554

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 554

gcaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc tgcggcctct 60

actctgcatt caattacata ctgac 85

<210> 555

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 555

gaccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat cctgcggcct 60

ctactctgca ttcaattaca tactg 85

<210> 556

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 556

ggtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc atcctgcggc 60

ctctactctg cattcaatta catac 85

<210> 557

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 557

gatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa ccatcctgcg 60

gcctctactc tgcattcaat tacat 85

<210> 558

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 558

ggtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta aaccatcctg 60

cggcctctac tctgcattca attac 85

<210> 559

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 559

gacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc taaaccatcc 60

tgcggcctct actctgcatt caatt 85

<210> 560

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 560

gacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt tctaaaccat 60

cctgcggcct ctactctgca ttcaa 85

<210> 561

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 561

gcaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat gttctaaacc 60

atcctgcggc ctctactctg cattc 85

<210> 562

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 562

gcccaacacg tatgtaccag gttcctggaa ctggaatctc tttccataga atgttctaaa 60

ccatcctgcg gcctctactc tgcat 85

<210> 563

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 563

ggacccaaca cgtatgtacc aggttcctgg aactggaatc tctttccata gaatgttcta 60

aaccatcctg cggcctctac tctgc 85

<210> 564

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 564

gttgacccaa cacgtatgta ccaggttcct ggaactggaa tctctttcca tagaatgttc 60

taaaccatcc tgcggcctct actct 85

<210> 565

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 565

gccttgaccc aacacgtatg taccaggttc ctggaactgg aatctctttc catagaatgt 60

tctaaaccat cctgcggcct ctact 85

<210> 566

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 566

gttccttgac ccaacacgta tgtaccaggt tcctggaact ggaatctctt tccatagaat 60

gttctaaacc atcctgcggc ctcta 85

<210> 567

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 567

gggttccttg acccaacacg tatgtaccag gttcctggaa ctggaatctc tttccataga 60

atgttctaaa ccatcctgcg gcctc 85

<210> 568

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 568

gttggttcct tgacccaaca cgtatgtacc aggttcctgg aactggaatc tctttccata 60

gaatgttcta aaccatcctg cggcc 85

<210> 569

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 569

gccttggttc cttgacccaa cacgtatgta ccaggttcct ggaactggaa tctctttcca 60

tagaatgttc taaaccatcc tgcgg 85

<210> 570

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 570

ggcccttggt tccttgaccc aacacgtatg taccaggttc ctggaactgg aatctctttc 60

catagaatgt tctaaaccat cctgc 85

<210> 571

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 571

gccgcccttg gttccttgac ccaacacgta tgtaccaggt tcctggaact ggaatctctt 60

tccatagaat gttctaaacc atcct 85

<210> 572

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 572

gcgccgccct tggttccttg acccaacacg tatgtaccag gttcctggaa ctggaatctc 60

tttccataga atgttctaaa ccatc 85

<210> 573

<211> 85

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 573

ggtcgccgcc cttggttcct tgacccaaca cgtatgtacc aggttcctgg aactggaatc 60

tctttccata gaatgttcta aacca 85

<210> 574

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 574

gtaatgcctg gcttgtcgac gcatagtctg 30

<210> 575

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 575

gaaaacgcag gttcctccag tttcgggagc agcgcacgtc tccctgtagt c 51

<210> 576

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 576

gacgcaggtt cctctagctt cgggagcagc gcacgtctcc ctgtagtcaa g 51

<210> 577

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 577

gtaatgcctg gcttgtcgac gcatagtctg 30

<210> 578

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 578

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca t 51

<210> 579

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 579

gtaatgcctg gcttgtcgac gcatagtctg 30

<210> 580

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 580

gcaaggccac aaaattatcc actgtttttg gaacagtctt tccgaagaga c 51

<210> 581

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 581

gcctgtagcc aaggccacaa aattatccac tgtttttgga acagtctttc c 51

<210> 582

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 582

gctttctctc ctgtagccaa ggccacaaaa ttatccactg tttttggaac a 51

<210> 583

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 583

ggccaaatcc tttctctcct gtagccaagg ccacaaaatt atccactgtt t 51

<210> 584

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 584

gtttttgtag ccaaatcctt tctctcctgt agccaaggcc acaaaattat c 51

<210> 585

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 585

gatttgctgt ttttgtagcc aaatcctttc tctcctgtag ccaaggccac a 51

<210> 586

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 586

gacgatggaa tttgctgttt ttgtagccaa atcctttctc tcctgtagcc a 51

<210> 587

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 587

gatagaatgt tctaaacgat cctgcggcct ctactctgca ttcaattaca t 51

<210> 588

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 588

gatagaatgt tctaaacaat cctgcggcct ctactctgca ttcaattaca t 51

<210> 589

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 589

gatagaatgt tctaaactat cctgcggcct ctactctgca ttcaattaca t 51

<210> 590

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 590

ggtcccacgc ggcccacagc tgcaccagga agaagggtgc ccagcacagc a 51

<210> 591

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 591

ggggtcccac gcggcccaca gctgcaccag gaagaagggt gcccagcaca g 51

<210> 592

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 592

gccgggtccc acgcggccca cagctgcacc aggaagaagg gtgcccagca c 51

<210> 593

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 593

gggtgatgac atcccaggcg atcgtgtggc ctccaggagc ccagagcagg a 51

<210> 594

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 594

gagggtgatg acatcccagg cgatcgtgtg gcctccagga gcccagagca g 51

<210> 595

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 595

gatcagggtg atgacatccc aggcgatcgt gtggcctcca ggagcccaga g 51

<210> 596

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 596

ggtggctcca ttcactccaa tgctgagcac ttccacagag tgggttaaag c 51

<210> 597

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 597

gtttctaata tattttgcca gactgatgga ctattctcaa ttaataatga t 51

<210> 598

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 598

gagatgttgc tgtggatcca gtccacagcc agcccgtcgg gggcctggat g 51

<210> 599

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 599

gcaggccggc ccagctgcca ggtgcacctg ctcggagcat cgggccggat c 51

<210> 600

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 600

gcaaagaacc tctgggtcca agggtagacc accagcagcc tgcccagggc c 51

<210> 601

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 601

gaagagaaac ttagtttcca gggctttggt agagggcaaa ggttgatagc a 51

<210> 602

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 602

gtcagcctag tgcagagcca ctggtagttg gtggttagag tttcaagttc c 51

<210> 603

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 603

ggctcattgt gaacaggcca gtaatgtccg ggatggggcg gcataggcgg g 51

<210> 604

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 604

gtagctaaag aacttgacca agacatatca ggatccacct cagctcctag a 51

<210> 605

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 605

ggggcattgt tctgtgccca gtcctgctgg tagacctgct ccccggtggc t 51

<210> 606

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 606

gagaagtcgt tcatgtgcca ccgtgggagc gtacagtcat cattgatctt g 51

<210> 607

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 607

gggattaatg ctgaacgcac caaagttcat cccaccaccc atattactac c 51

<210> 608

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 608

gctccaaagg ctttcctcca ctgttgcaaa gttattgaat cccaagacac a 51

<210> 609

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 609

gatgaatgaa cgatttccca gaactcccta atcagaacag agtccctggt a 51

<210> 610

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 610

ggagcctctg ccggagccca gagaacccga gagtcagaca gagccagcgc c 51

<210> 611

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 611

ggcttccgtg gagacaccca atcaatttga agagatcttg aagtgatgcc a 51

<210> 612

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 612

gtgggactgc cctcctccca tttgcagatg ccgtcgtaga atcgcagcag g 51

<210> 613

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 613

gcttcttcaa tagttctcca gctacactgg caggcatatg cccgtgttcc t 51

<210> 614

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 614

gattcctttt cttcgtccca attcacctca gtggctagtc gaagaatgaa g 51

<210> 615

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 615

gcagcttcag caccttccag tcagactcct gcttcaagca ctgcagcagg a 51

<210> 616

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 616

gccatttgct tgcagtgcca ctccagagga ttccggattg ccataaatac t 51

<210> 617

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 617

gttcaatagt tttggtccag tatcgtttac agcccttctt ggtagatttc a 51

<210> 618

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 618

ggcaaccgtc ttctgaccaa atggcagaac atttgtcccc aactttccac t 51

<210> 619

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 619

gcgactttcc aatgaaccac tgaagcccag gtatgacaaa gccgatgatc t 51

<210> 620

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 620

gtttacactc atgcttccac agctttaaca gatcatttgg ttccttgatg a 51

<210> 621

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 621

gcttaagctt ccgtgtccag ccttcaggca gggtggggtc atcatacatg g 51

<210> 622

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 622

ggacagctgg gctgatccat gatgtcatcc agaaacactg gggaccctca g 51

<210> 623

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 623

gtctcatctc aactttccat atccgtatca tggaatcata gcatcctgta a 51

<210> 624

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 624

gcatgcagac gcggttccac tcgcagccac agttccagca ccactcgagc c 51

<210> 625

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 625

gttggttagg gtcaacccag tattctccac tcttgagttc aggatggcag a 51

<210> 626

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 626

gctacactgt ccaacaccca ctctcgggtc accacaggtg cctcacacat c 51

<210> 627

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 627

gctgcactgt gtacccccag agctccgtgt tgccgacatc ctggggtggc t 51

<210> 628

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 628

gagcttcctg ccactcccaa caggtttcac agtaagcgcg tatctgttcc a 51

<210> 629

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 629

gacggcaaga gcttacccag tcacttgtgt ggagacttaa atacttgcat a 51

<210> 630

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 630

gcaaggccac aaaattatcc actgtttttg gaacagtctt tccgaagaga c 51

<210> 631

<211> 51

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 631

gcctgtagcc aaggccacaa aattatccac tgtttttgga acagtctttc c 51

<210> 632

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 632

gctttctctc ctgtagccaa ggccacaaaa ttatccactg tttttggaac a 51

<210> 633

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 633

ggccaaatcc tttctctcct gtagccaagg ccacaaaatt atccactgtt t 51

<210> 634

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 634

gtttttgtag ccaaatcctt tctctcctgt agccaaggcc acaaaattat c 51

<210> 635

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 635

gatttgctgt ttttgtagcc aaatcctttc tctcctgtag ccaaggccac a 51

<210> 636

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 636

gacgatggaa tttgctgttt ttgtagccaa atcctttctc tcctgtagcc a 51

<210> 637

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 637

gtaatgcctg gcttgtcgac gcatagtctg 30

<210> 638

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 638

gatagaatgt tctaaaccat cctgcggcct ctactctgca ttcaattaca t 51

<210> 639

<211> 34

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 639

gcagggtttt cccagtcacg acgttgtaaa gttg 34

<210> 640

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 640

gtcaaggcac tcttgcccac gccaccagct ccaactacca caagtttata t 51

<210> 641

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 641

gcaaagatca cccggcccac atcttcatct ccaattcgta ggtcaaaata c 51

<210> 642

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 642

gcgccaccag ctccaaccac cacaagttta tattcagtca ttttcagcag g 51

<210> 643

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 643

gtttctccat caattaccac ttgcttcctg taggaatcct ctattgttgg a 51

<210> 644

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 644

gctttctctc ctgtagccaa ggccacaaaa ttatccactg tttttggaac a 51

<210> 645

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 645

gtaatgcctg gcttgtcgac gcatagtctg 30

<210> 646

<211> 83

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 646

tctttccata ggccctgaaa aagggcctgt tctaaaccat cctgcggcct ctactcggcc 60

ctgaaaaagg gccattcaat tac 83

<210> 647

<211> 83

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 647

cagctggcga ggccctgaaa aagggccggg gatgtgccgc aaggcgatta agttggggcc 60

ctgaaaaagg gccacgccag ggt 83

<210> 648

<211> 79

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 648

gtggaatagt ataacaatat gctaaatgtt gttatagtat cccactctaa accatcctgc 60

ggggccctct tcagggccc 79

<210> 649

<211> 79

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 649

gtggaatagt ataacaatat gctaaatgtt gttatagtat cccacaccct ggcgttaccc 60

agggccctct tcagggccc 79

<210> 650

<211> 385

<212> PRT

<213> Homo sapiens

<400> 650

Gln Leu His Leu Pro Gln Val Leu Ala Asp Ala Val Ser Arg Leu Val
1 5 10 15

Leu Gly Lys Phe Gly Asp Leu Thr Asp Asn Phe Ser Ser Pro His Ala

20 25 30

Arg Arg Lys Val Leu Ala Gly Val Val Met Thr Thr Gly Thr Asp Val
35 40 45

Lys Asp Ala Lys Val Ile Ser Val Ser Thr Gly Thr Lys Cys Ile Asn
50 55 60

Gly Glu Tyr Met Ser Asp Arg Gly Leu Ala Leu Asn Asp Cys His Ala
65 70 75 80

Glu Ile Ile Ser Arg Arg Ser Leu Leu Arg Phe Leu Tyr Thr Gln Leu

85 90 95

Glu Leu Tyr Leu Asn Asn Lys Asp Asp Gln Lys Arg Ser Ile Phe Gln
100 105 110

Lys Ser Glu Arg Gly Gly Phe Arg Leu Lys Glu Asn Val Gln Phe His
115 120 125

Leu Tyr Ile Ser Thr Ser Pro Cys Gly Asp Ala Arg Ile Phe Ser Pro
130 135 140

His Glu Pro Ile Leu Glu Glu Pro Ala Asp Arg His Pro Asn Arg Lys

145 150 155 160

Ala Arg Gly Gln Leu Arg Thr Lys Ile Glu Ser Gly Gln Gly Thr Ile
165 170 175

Pro Val Arg Ser Asn Ala Ser Ile Gln Thr Trp Asp Gly Val Leu Gln
180 185 190

Gly Glu Arg Leu Leu Thr Met Ser Cys Ser Asp Lys Ile Ala Arg Trp
195 200 205

Asn Val Val Gly Ile Gln Gly Ser Leu Leu Ser Ile Phe Val Glu Pro

210 215 220

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

275 280 285
Val Gly Asp Ser Ala Ile Glu Val Ile Asn Ala Thr Thr Gly Lys Asp
290 295 300
Glu Leu Gly Arg Ala Ser Arg Leu Cys Lys His Ala Leu Tyr Cys Arg
305 310 315 320
Trp Met Arg Val His Gly Lys Val Pro Ser His Leu Leu Arg Ser Lys
325 330 335
Ile Thr Lys Pro Asn Val Tyr His Glu Ser Lys Leu Ala Ala Lys Glu

340 345 350
Tyr Gln Ala Ala Lys Ala Arg Leu Phe Thr Ala Phe Ile Lys Ala Gly
355 360 365
Leu Gly Ala Trp Val Glu Lys Pro Thr Glu Gln Asp Gln Phe Ser Leu
370 375 380
Thr
385

<210> 651
<211> 336
<212> PRT
<213> Homo sapiens
<400> 651

Ser Leu Gly Thr Gly Asn Arg Cys Val Lys Gly Asp Ser Leu Ser Leu
1 5 10 15

Lys Gly Glu Thr Val Asn Asp Cys His Ala Glu Ile Ile Ser Arg Arg
20 25 30
Gly Phe Ile Arg Phe Leu Tyr Ser Glu Leu Met Lys Tyr Asn Ser Gln

35 40 45

Thr Ala Lys Asp Ser Ile Phe Glu Pro Ala Lys Gly Gly Glu Lys Leu

50 55 60

Gln Ile Lys Lys Thr Val Ser Phe His Leu Tyr Ile Ser Thr Ala Pro

65 70 75 80

Cys Gly Asp Gly Ala Leu Phe Asp Lys Ser Cys Ser Asp Arg Ala Met

85 90 95

Glu Ser Thr Glu Ser Arg His Tyr Pro Val Phe Glu Asn Pro Lys Gln

100 105 110

Gly Lys Leu Arg Thr Lys Val Glu Asn Gly Gln Gly Thr Ile Pro Val

115 120 125

Glu Ser Ser Asp Ile Val Pro Thr Trp Asp Gly Ile Arg Leu Gly Glu

130 135 140

Arg Leu Arg Thr Met Ser Cys Ser Asp Lys Ile Leu Arg Trp Asn Val

145 150 155 160

Leu Gly Leu Gln Gly Ala Leu Leu Thr His Phe Leu Gln Pro Ile Tyr

165 170 175

Leu Lys Ser Val Thr Leu Gly Tyr Leu Phe Ser Gln Gly His Leu Thr

180 185 190

Arg Ala Ile Cys Cys Arg Val Thr Arg Asp Gly Ser Ala Phe Glu Asp

195 200 205

Gly Leu Arg His Pro Phe Ile Val Asn His Pro Lys Val Gly Arg Val

210 215 220

Ser Ile Tyr Asp Ser Lys Arg Gln Ser Gly Lys Thr Lys Glu Thr Ser

225 230 235 240

Val Asn Trp Cys Leu Ala Asp Gly Tyr Asp Leu Glu Ile Leu Asp Gly

245 250 255

Thr Arg Gly Thr Val Asp Gly Pro Arg Asn Glu Leu Ser Arg Val Ser

260 265 270

Lys Lys Asn Ile Phe Leu Leu Phe Lys Lys Leu Cys Ser Phe Arg Tyr

275 280 285

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

<210> 652
<211> 340
<212> PRT
<213> Octopus vulgaris
<400> 652
Ser Val Gly Thr Gly Asn Arg Cys Leu Thr Gly Asp His Leu Ser Leu
1 5 10 15
Glu Gly Asn Ser Val Asn Asp Ser His Ala Glu Met Ile Thr Arg Arg
20 25 30
Gly Phe Leu Arg Tyr Leu Tyr Arg His Leu Leu Glu Tyr Asp Ala Glu
35 40 45
Val Pro Asn Asp Leu Phe Glu Lys Gly Glu Arg Ser Ile Cys Arg Ile

50 55 60
Lys Thr Asn Ile Thr Phe His Leu Tyr Ile Ser Thr Ala Pro Cys Gly
65 70 75 80
Asp Gly Ala Leu Phe Ser Pro Arg Asp Thr Asp Ser Ser Asn Ala Lys
85 90 95
Met Glu Glu Glu Asn Lys His Ile His Asn Pro Thr Phe Ser Ser Ser
100 105 110
Val Gln Gly Leu Leu Arg Thr Lys Val Glu Gly Gly Gln Gly Thr Ile

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

165 170 175

Ile Tyr Leu Asp Ser Leu Thr Leu Gly Tyr Leu Tyr Asp His Gly His

180 185 190

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

195 200 205

Asn Gln Leu Leu Pro Glu Gly Tyr Arg Leu Asn His Pro Trp Leu Gly

210 215 220

Arg Val Thr Ala Cys Asp Pro Pro Arg Glu Thr Gln Lys Thr Lys Ser

225 230 235 240

Leu Ser Ile Asn Trp Cys Tyr Asp Asp Glu Lys Ser Glu Val Leu Asp

245 250 255

Gly Thr Ala Gly Ile Cys Tyr Thr Ala Ile Glu Lys Asn Leu Phe Ser

260 265 270

Arg Leu Thr Lys His Asn Leu Tyr Glu Glu Phe Lys Arg Val Cys Arg

275 280 285

Lys Phe Asp Arg Asn Asp Leu Leu Thr Ala Pro Ser Tyr Asn Lys Ala

290 295 300

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

305 310 315 320

Lys Leu Lys Glu Asn Asn Cys Gly Thr Trp Val Ser Lys Pro Ile Glu

325 330 335

Glu Glu Met Phe

340

<210> 653

<211> 340

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 653

Ser Val Gly Thr Gly Asn Arg Cys Leu Thr Gly Asp Arg Leu Ser Leu

1 5 10 15

Glu Gly Asn Ser Val Asn Asp Ser His Ala Glu Met Val Thr Arg Arg

20 25 30

Gly Phe Leu Arg Tyr Leu Tyr Lys His Leu Leu Glu Tyr Asp Pro Glu

35 40 45

Lys Pro His Asp Leu Phe Glu Lys Gly Glu Arg Ser Leu Cys Arg Ile

50 55 60

Lys Thr Asn Ile Thr Phe His Leu Tyr Ile Ser Thr Ala Pro Cys Gly

65 70 75 80

Asp Gly Ala Leu Phe Ser Pro Arg Asp Thr Asp Ser Ser Asn Val Lys

85 90 95

Val Asp Glu Glu Asn Lys His Val His Asn Pro Thr Phe Ser Ser Ser

100 105 110

Val Gln Gly Leu Leu Arg Thr Lys Val Glu Gly Gly Gln Gly Thr Ile

115 120 125

Pro Ile Asp Ala Asp Phe Thr Glu Gln Thr Trp Asp Gly Ile Gln Arg

130 135 140

Gly Glu Arg Leu Arg Thr Met Ser Cys Ser Asp Lys Ile Cys Arg Trp

145 150 155 160

Asn Val Val Gly Leu Gln Gly Ala Leu Leu Ser His Phe Val Glu Pro

165 170 175

Ile Tyr Leu Glu Ser Leu Thr Leu Gly Tyr Leu Tyr Asp His Gly His

180 185 190

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

195 200 205

Asn Gln Leu Leu Pro Glu Gly Tyr Arg Leu Asn His Pro Trp Leu Gly

210 215 220

Arg Val Thr Ala Cys Asp Pro Pro Arg Glu Thr Gln Lys Thr Lys Ser

225 230 235 240

Leu Ser Ile Asn Trp Cys Tyr Asp Asp Glu Lys Ser Glu Val Leu Asp

245 250 255

Gly Thr Ala Gly Ile Cys Tyr Thr Ala Ile Glu Lys Asn Leu Phe Ser

260 265 270

Arg Leu Thr Lys His Ser Leu Tyr Glu Glu Phe Lys Lys Val Cys Gln

275 280 285

Lys Phe Glu Arg Glu Asp Leu Leu Asn Val Thr Ser Tyr Asn Lys Ala

290 295 300

Lys Met Met Ala Ile Pro Phe Gln Thr Ala Lys Asn Val Met Leu Lys

305 310 315 320

Lys Leu Lys Glu Asn Asn Cys Gly Thr Trp Val Ser Lys Pro Ile Glu

325 330 335

Glu Glu Met Phe

340

<210> 654

<211> 328

<212> PRT

<213> Octopus vulgaris

<400> 654

Gly Ile Gly Thr Gly Thr Lys Cys Ile Asn Gly Glu His Met Ser Asp

1 5 10 15

Arg Gly Phe Gly Val Asn Asp Cys His Ala Glu Ile Ile Ala Arg Arg

20 25 30

Cys Phe Leu Arg Tyr Ile Tyr Asp Gln Leu Glu Leu His Leu Ser Asp

35 40 45

Asn Ser Asp Val Arg Asn Ser Ser Ile Phe Glu Leu Arg Asp Lys Gly

50 55 60

Gly Tyr Gln Leu Lys Glu Asn Ile Gln Phe His Leu Tyr Ile Ser Thr

65 70 75 80

Ala Pro Cys Gly Asp Ala Arg Ile Phe Ser Pro His Gly Gln Asp Val

85 90 95

Glu Thr Gly Asp Arg His Pro Asn Arg Lys Ala Arg Gly Gln Leu Arg

100 105 110

Thr Lys Ile Glu Ser Gly Gln Gly Thr Ile Pro Val Arg Thr Ser Gly

115 120 125

Val Ile Gln Thr Trp Asp Gly Val Leu Glu Gly Glu Arg Leu Leu Thr

130 135 140

Met Ser Cys Ser Asp Lys Ile Ala Arg Trp Asn Val Leu Gly Ile Gln

145 150 155 160

Gly Ser Leu Leu Ser His Phe Met Asn Pro Ile Tyr Leu Glu Ser Ile

165 170 175

Ile Leu Gly Ser Leu Tyr His Ser Asp His Leu Ser Arg Ala Met Tyr

180 185 190

Ser Arg Ile Ser Ile Ile Glu Asn Leu Pro Glu Pro Phe His Leu Asn

195 200 205

Arg Pro Phe Leu Ser Gly Ile Ser Ser Pro Glu Ser Arg Gln Pro Gly

210 215 220

Lys Ala Pro Asn Phe Gly Ile Asn Trp Arg Lys Glu Asp Glu Thr Phe

225 230 235 240

Glu Val Ile Asn Ala Met Thr Gly Arg Val Glu Gly Gly Ser Val Ser

245 250 255

Arg Ile Cys Lys Gln Ala Leu Phe Gly Arg Phe Met Ser Leu Tyr Gly

260 265 270

Lys Leu Ser Ser Leu Thr Gly Gln Ser Val Thr Thr Arg Pro Thr His

275 280 285

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

290 295 300

Cys Val Phe Gln Ala Phe Gln Lys Ala Gly Leu Gly Asn Trp Val Gln

305 310 315 320

Lys Pro Ile Glu Gln Asp Gln Phe

325

<210> 655

<211> 328

<212> PRT

<213> Sepia mutant

<400> 655

Gly Ile Gly Thr Gly Thr Lys Cys Ile Asn Gly Glu Tyr Met Asn Asp
1 5 10 15
Arg Gly Phe Ala Val Asn Asp Cys His Ala Glu Ile Ile Ala Arg Arg
20 25 30
Cys Phe Leu Arg Phe Ile Tyr Asp Gln Leu Glu Met His Leu Ser Glu

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

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

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

225 230 235 240
Glu Val Ile Asn Ala Met Thr Gly Arg Val Glu Gly Gly Asn Met Ser

245 250 255

Arg Ile Cys Lys Gln Ala Leu Phe Asp Arg Phe Met Asn Leu Tyr Gly

260 265 270

Arg Leu Ser Ser Leu Thr Gly Gln Ser Val Thr Thr Arg Pro Thr Leu

275 280 285

Tyr Ser Glu Ala Lys Ala Ala Val Met Glu Tyr Gln Leu Ala Lys Gln

290 295 300

Cys Val Phe Gln Ala Phe Gln Lys Ala Gly Leu Gly Asn Trp Val Gln

305 310 315 320

Lys Pro Ile Glu Gln Asp Gln Phe

325

<210> 656

<211> 328

<212> PRT

<213> Doryteusthis opalescens

<400> 656

Gly Ile Gly Thr Gly Thr Lys Cys Ile Asn Gly Glu Tyr Met Asn Asp

1 5 10 15

Arg Gly Phe Ala Val Asn Asp Cys His Ala Glu Ile Ile Ala Arg Arg

20 25 30

Cys Phe Leu Arg Phe Ile Tyr Asp Gln Leu Glu Leu His Leu Ser Asp

35 40 45

Asn Ala Glu Val Arg Gly Gln Ser Ile Phe Glu Leu Arg Asp Ala Gly

50 55 60

Gly Tyr Lys Leu Lys Pro Asn Ile Gln Phe His Leu Tyr Ile Ser Thr

65 70 75 80

Ala Pro Cys Gly Asp Ala Arg Ile Phe Ser Pro His Gly Gln Asp Val

85 90 95

Glu Thr Gly Asp Arg His Pro Asn Arg Lys Ala Arg Gly Gln Leu Arg

100 105 110

Thr Lys Ile Glu Ser Gly Gln Gly Thr Ile Pro Val Arg Ser Ser Gly

115 120 125

Phe Ile Gln Thr Trp Asp Gly Val Leu Glu Gly Glu Arg Leu Leu Thr
130 135 140
Met Ser Cys Ser Asp Lys Ile Ala Arg Trp Asn Val Leu Gly Val Gln

145 150 155 160
Gly Ala Leu Leu Cys His Phe Met His Pro Ile Tyr Leu Glu Ser Ile
165 170 175
Ile Leu Gly Ser Leu Tyr His Ser Asp His Leu Ser Arg Ala Val Tyr
180 185 190
Cys Arg Ile Ala Ala Ile Glu Asn Leu Pro Asp Leu Phe Arg Leu Asn
195 200 205
Arg Pro Phe Leu Ser Gly Ile Ser Ser Pro Glu Ser Arg Gln Pro Gly

210 215 220
Lys Ala Pro Asn Phe Gly Ile Asn Trp Arg Arg Asn Asp Asp Ser Phe
225 230 235 240
Glu Val Ile Asn Ala Met Thr Gly Arg Val Glu Gly Gly Ser Met Ser
245 250 255
Arg Ile Cys Lys Gln Ala Leu Phe Asp Arg Phe Met Asn Leu Tyr Gly
260 265 270
Lys Leu Ser Ser Leu Thr Gly Gln Ser Val Thr Thr Arg Pro Ala Leu

275 280 285
Tyr Ser Glu Ala Lys Ala Thr Val Met Glu Tyr Gln Leu Ala Lys Gln
290 295 300
Cys Val Phe Gln Ala Phe Gln Lys Ala Gly Leu Gly Asn Trp Val Gln
305 310 315 320
Lys Pro Ile Glu Gln Asp Gln Phe
325

<210> 657
<211> 57
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic Oligonucleotide
<400>

657

gaatgcagag tagaggccgc aggatagttt agaacattct atggaaagag attccag 57

<210> 658

<211> 86

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 658

ttgccgaatg tgtcagtatg taattgaatg cagagtagag gccgcaggat agtttagaac 60

attctatgga aagagattcc agttcc 86

<210> 659

<211> 86

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 659

aacggcttac acagtcatac attaacttac gtctcatctc cggcgtccta tcaaatcttg 60

taagatacct ttctctaagg tcaagg 86

<210> 660

<211> 51

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 660

cttacgtctc atctccggcg tcctatcaaa tcttgtaaga tacctttctc t 51

<210> 661

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 661

augacaugga cugauuugc 19

<210> 662

<211> 19

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 662

uacugcaccg gacgaaacg 19

<210> 663

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (6)..(6)

<223> n = inosine

<400> 663

augacnugga cugauuugc 19

<210> 664

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 664

uacugcaccg gacgaaacg 19

<210> 665

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 665

augacaugga cugauuugc 19

<210> 666

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 666

uacugcacca gacaaaacg 19

<210> 667

<211> 19

<212

> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (6)..(6)

<223> n = inosine

<400> 667

augacnugga cugauuugc 19

<210> 668

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 668

uacugcacca gacaaaacg 19

<210> 669

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 669

augacaugga cugauuugc 19

<210> 670

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 670

uacugcaccg gaccaaacg 19

<210> 671

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (6)..(6)

<223> n = inosine

<220><221> misc_feature

<222> (14)..(14)

<223> n = inosine

<400> 671

augacnugga cugnuuugc 19

<210> 672

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 672

uacugcaccg gaccaaacg 19

<210> 673

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 673

tggaggtgct caaagagatg gaagccaatg cccgga 36

<210> 674

<211> 36

<212

> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<400> 674

nggaggtgct caaagagatg gaagccaatg cccgga 36

<210> 675

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (2)..(2)

<223> n = any nucleotide

<400> 675

tngaggtgct caaagagatg gaagccaatg cccgga 36

<210

> 676

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (3)..(3)

<223> n = any nucleotide

<400> 676

tgnaggtgct caaagagatg gaagccaatg cccgga 36

<210> 677

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<400> 677

tggnggtgct caaagagatg gaagccaatg cccgga 36

<210> 678

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<220><221> misc_feature

<222> (36)..(36)

<223> n = any nucleotide

<400> 678

tggnggtgct caaagagatg gaagccaatg cccggn 36

<210> 679

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(2)

<223> n = any nucleotide

<400> 679

nngaggtgct caaagagatg gaagccaatg cccgga 36

<210> 680

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (3)..(3)

<223> n = any nucleotide

<400> 680

ngnaggtgct caaagagatg gaagccaatg cccgga 36

<210> 681

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<400> 681

nggnggtgct caaagagatg gaagccaatg cccgga 36

<210> 682

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (5)..(5)

<223> n = any nucleotide

<400> 682

nggangtgct caaagagatg gaagccaatg cccgga 36

<210> 683

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (35)..(36)

<223> n = any nucleotide

<400> 683

tggaggtgct caaagagatg gaagccaatg cccgnn 36

<210> 684

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 684

agagatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 685

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<400> 685

ngagatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 686

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (2)..(2)

<223> n = any nucleotide

<400> 686

anagatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 687

<211> 34

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (3)..(3)

<223> n = any nucleotide

<400> 687

agngatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 688

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<400> 688

aganatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 689

<211

> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (34)..(34)

<223> n = any nucleotide

<400> 689

agagatggaa gccaatgccc ggaaagctgg ctgn 34

<210> 690

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(2)

<223> n = any nucleotide

<400> 690

nnagatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 691

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (3)..(3)

<223> n = any nucleotide

<400> 691

ngngatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 692

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<400> 692

nganatggaa gccaatgccc ggaaagctgg ctgc 34

<210> 693

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (5)..(5)

<223> n = any nucleotide

<400> 693

ngagntggaa gccaatgccc ggaaagctgg ctgc 34

<210> 694

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (33)..(34)

<223> n = any nucleotide

<400> 694

agagatggaa gccaatgccc ggaaagctgg ctnn 34

<210> 695

<211> 55

<212> RNA

<213> Cypridina luciferase

<400> 695

aaugcagagu agaggccgca ggauaguuua gaacauucua uggaaagaga uucca 55

<210> 696

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 696

Cys Arg Met Cys Gln Tyr Val Ile Glu Cys Arg Val Glu Ala Ala Gly
1 5 10 15

Phe Arg Thr Phe Tyr Gly Lys Arg Phe Gln Phe
20 25

<210> 697

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 697

cgagcctcac acgtgctgct tgactacagg gagacgtgcg ctgctccc 48

<210> 698

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223>
Synthetic Peptide

<400> 698

Arg Ala Ser His Val Leu Leu Asp Tyr Arg Glu Thr Cys Ala Ala Pro
1 5 10 15

<210> 699

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 699

cgagcctcac acgtggaaag ggactacagg gagacgtgcg ctgctccc 48

<210> 700

<211> 45

<212> RNA

<213> Cypridina luciferase

<400> 700

auugaaugca gaguagaggc cgcaggauag uuuagaacau ucuau 45

<210> 701

<211> 66

<212> RNA

<213> Cypridina luciferase

<400> 701

ucuaaaccau ccugcggccu cuacucugca guuguggaag guccaguuuu gggggcuauu 60

acaaca 66

<210> 702

<211> 105

<212> DNA

<213> Cypridinia luciferase

<400> 702

tgccgaaugu gucaguaugu aauugaaugc agaguagagg ccgcaggaua guuuagaaca 60

uucuauggaa agagauucca guuccaggaa ccugguacau acgug 105

<210> 703

<211> 48

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 703

cgagcctcac acaagaccct tgactacagg gagacgtgcg ctgctccc 48

<210> 704

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 704

uuccccacau ugacguuca 19

<210> 705

<211> 65

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 705

ucgucuaugu gcugugcugg gcacccuucu uccuggugca gcuguaggcc gcgugggacc 60

cggag 65

<210> 706

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 706

gucccacgcg gcccacagcu gcaccaggaa gaagggugcc cagcacagca 50

<210> 707

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 707

gggucccacg cggcccacag cugcaccagg aagaagggug cccagcacag 50

<210> 708

<211> 50

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 708

ccggguccca cgcggcccac agcugcacca ggaagaaggg ugcccagcac 50

<210> 709

<211> 65

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 709

ccucaacuuc cugcucuggg cuccuggagg ccacacgauc gccuaggaug ucaucacccu 60

gaugg 65

<210> 710

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 710

ggugaugaca ucccaggcga ucguguggcc uccaggagcc cagagcagga 50

<210> 711

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 711

agggugauga caucccaggc gaucgugugg ccuccaggag cccagagcag 50

<210> 712

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 712

aucaggguga ugacauccca ggcgaucgug uggccuccag gagcccagag 50

<210> 713

<211> 99

<212

> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 713

cccaggugcg ggagagaggc cugcugaaaa ugacugaaua uaaacuugug guaguuggag 60

cugguggcgu aggcaagagu gccuugacga uacagcuaa 99

<210> 714

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 714

caccacaagu uuauauucag ucauuuucag caggccucuc ucccgcaccu 50

<210> 715

<211> 50

<212> RNA

<213> Artificial Sequence

<220

><223> Synthetic Oligonucleotide

<400> 715

gcuccaacca ccacaaguuu auauucaguc auuuucagca ggccucucuc 50

<210> 716

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 716

cgccaccagc uccaaccacc acaaguuuau auucagucau uuucagcagg 50

<210> 717

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 717

cuugccuacg ccaccagcuc caaccaccac aaguuuauau ucagucauuu 50

<210> 718

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 718

aaggcacucu ugccuacgcc accagcucca accaccacaa guuuauauuc 50

<210> 719

<211> 49

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 719

guacgucaag gcacucuugc cuacgccacc agcuccaacc accacaagu 49

<210> 720

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 720

aauuagcugu aucgucaagg cacucuugcc uacgccacca gcuccaacca 50

<210> 721

<211> 44

<212> DNA

<213> Cypridinia luciferase

<400> 721

tatgtaattg aatgcagagt agaggccgca ggatagttta gaac 44

<210> 722

<211> 39

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (4)..(5)

<223> n = any nucleotide

<220><221> misc_feature

<222> (36)..(39)

<223> n = any nucleotide

<400> 722

augnnuugcc accgacgaca ccgugaggau aaauunnnn 39

<210> 723

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 723

uggaggugcu caaagagaug gaagccaaug cccgga 36

<210> 724

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<400> 724

nggaggugcu caaagagaug gaagccaaug cccgga 36

<210> 725

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (2)..(2)

<223> n = any nucleotide

<400> 725

ungaggugcu caaagagaug gaagccaaug cccgga 36

<210> 726

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (3)..(3)

<223> n = any nucleotide

<400> 726

ugnaggugcu caaagagaug gaagccaaug cccgga 36

<210> 727

<211> 36

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<400> 727

uggnggugcu caaagagaug gaagccaaug cccgga 36

<210> 728

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (18)..(18)

<223> n = any nucleotide

<400> 728

uggaggugcu caaagagnug gaagccaaug cccgga 36

<210> 729

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (36)..(36)

<223> n = any nucleotide

<400> 729

uggaggugcu caaagagaug gaagccaaug cccggn 36

<210> 730

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(2)

<223> n = any nucleotide

<400> 730

nngaggugcu caaagagaug gaagccaaug cccgga 36

<210> 731

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (3)..(3)

<223> n = any nucleotide

<400> 731

ngnaggugcu caaagagaug gaagccaaug cccgga 36

<210> 732

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (4)..(4)

<223> n = any nucleotide

<400> 732

nggnggugcu caaagagaug gaagccaaug cccgga 36

<210> 733

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(1)

<223> n = any nucleotide

<220><221> misc_feature

<222> (5)..(5)

<223> n = any nucleotide

<400> 733

nggangugcu caaagagaug gaagccaaug cccgga 36

<210> 734

<211> 36

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (14)..(14)

<223> n = any nucleotide

<220><221> misc_feature

<222> (18)..(18)

<223> n = any nucleotide

<400> 734

uggaggugcu caangagnug gaagccaaug cccgga 36

<210> 735

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 735

aaggggugua accagcaagu 20

<210> 736

<211> 34

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 736

agagauggaa gccaaugccc ggaaagcugg cugc 34

<210> 737

<211> 24

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (1)..(4)

<223> n = any nucleotide

<400> 737

nnnngacuac agucccgaaa cuag 24

<210> 738

<211> 19

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 738

uuccccacau ugacguuca 19

<210> 739

<211> 20

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 739

aaggggugua accggcaagu 20

<210> 740

<211> 36

<212> DNA

<213> Cypridinia luciferase

<400> 740

tatgtaattg aatgcagccg caggatggtt tggagc 36

<210> 741

<211> 37

<212

> DNA

<213> Cypridinia luciferase

<400> 741

tatgtaattg aatgcagacc gcaggatagt ttagaac 37

<210> 742

<211> 38

<212> DNA

<213> Cypridinia luciferase

<400> 742

tatgtaattg aatgcagagc cgcaggatgg tttggaac 38

<210> 743

<211> 38

<212> DNA

<213> Cypridinia luciferase

<400> 743

tatgtaattg aatgcagagt ggaaggatgg tttagaac 38

<210> 744

<211> 38

<212> DNA

<213> Cypridinia luciferase

<400> 744

tatgtaattg aatgcagagt ggaaggatgg tttagaac 38

<210> 745

<211> 38

<212> DNA

<213> Cypridinia luciferase

<400> 745

tgtgtaattg aatgcagagt ggaaggatgg tttagaac 38

<210> 746

<211> 104

<212> RNA

<213> Homo sapiens

<400> 746

uuggucucuu cggaaagacu guuccaaaaa caguggauaa uuuuguggcc uuagcuacag 60

gagagaaagg auuuggcuac aaaaacagca aauuccaucg ugua 104

<210> 747

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 747

cagagaagcc uuucugacaa gguuuuuguc accuauuaaa acaccggaac 50

<210> 748

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 748

ccuuucugac aagguuuuug ucaccuauua aaacaccgga accgaugucc 50

<210> 749

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 749

acaagguuuu ugucaccuau uaaaacaccg gaaccgaugu ccucucuuuc 50

<210> 750

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 750

uuugucaccu auuaaaacac cggaaccgau guccucucuu uccuaaaccg 50

<210> 751

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 751

cuauuaaaac accggaaccg auguccucuc uuuccuaaac cgauguuuuu 50

<210> 752

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 752

acaccggaac cgauguccuc ucuuuccuaa accgauguuu uugucguuua 50

<210> 753

<211> 50

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 753

accgaugucc ucucuuuccu aaaccgaugu uuuugucguu uaagguagca 50

<210> 754

<211> 56

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 754

aguauguaau ugaaugcaga guagaggccg caggauaguu uagaacauuc uaugga 56

<210> 755

<211> 50

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<220><221> misc_feature

<222> (17)..(17)

<223> n is a, c, g, or u

<400> 755

auagaauguu cuaaacnauc cugcggccuc uacucugcau ucaauuacau 50

<210> 756

<211> 59

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 756

auguaauuga augcagagua gaggccgcag gauaguuuag aacauucuau ggaaagaga 59

<210> 757

<211> 83

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 757

ucuuuccaua ggcccugaaa aagggccugu ucuaaaccau ccugcggccu cuacucggcc 60

cugaaaaagg gccauucaau uac 83

<210> 758

<211> 21

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 758

ggccgcagga uaguuuagaa c 21

<210> 759

<211> 55

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 759

aaauguuguu auaguauccc accucuaaac cauccugcgg ggcccuaggg cccuu 55

<210> 760

<211> 26

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 760

accgguggaa uaguauaaca auaugc 26

<210> 761

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 761

Gly Ser Gly Gly Gly Gly Ser

1 5

<210> 762

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 762

tatgtaattg aatgcagagt ggaagggtgg tttggaac 38

<210> 763

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 763

tatgtaattg aatgcagagt agaaggatgg tttagaac 38

<210> 764

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 764

tatgtaattg aatgcagagt agagaggatg gtttggaac 39

<210> 765

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 765

tgtgtaattg aatgcagagt ggagaggatg gtttagaac 39

<210> 766

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 766

tatgtaattg aatgcagagt ggagaggatg gtttagagc 39

<210> 767

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 767

tatgtaattg aatgcagagt agagaggatg gtttagagc 39

<210> 768

<211> 39

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 768

tatgtaattg aatgcagagt ggaggccgcg gtttggaac 39

<210> 769

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 769

tatgtaattg aatgcagagt agagaccgcg gtttggagc 39

<210> 770

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 770

tatgtaattg aatgcagagt ggagggcgca gtttggagc 39

<210> 771

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 771

tatgtaattg aatgcagagt agaggccgca ggttggaac 39

<210> 772

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 772

tgtgtaattg aatgcagagt ggaggccgca ggttagaac 39

<210> 773

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 773

tatgtaattg aatgcagagt ggaggccgca ggttggaac 39

<210> 774

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 774

tatgtaattg aatgcagagt ggaggccgca ggttggaac 39

<210> 775

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 775

tatgtaattg aatgcagagt agagggcgca ggtggagc 38

<210> 776

<211> 39

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 776

tatgtaattg aatgcagagt ccgcaggatg gtttagaac 39

<210> 777

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 777

tatgtaattg aatgcagagt agcaggatag tttagaac 38

<210> 778

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 778

tatgtaattg aatgcagagt agaaggatag tttagaac 38

<210> 779

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 779

tatgtaattg aatgcagagt agagaggatg gtttagaac 39

<210> 780

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 780

tatgtaattg aatgcagagt agagaggatg gtttagaac 39

<210> 781

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 781

tatgtaattg aatgcagagt agagaggatg gtttagaac 39

<210> 782

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 782

tatgtaattg aatgcagagt agagaggata gtttagaac 39

<210> 783

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 783

tatgtaattg aatgcagagt agaggcatgg tttagaac 38

<210> 784

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 784

tatgtaattg aatgcagagt agaggccatg gtttagaac 39

<210> 785

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 785

tatgtaattg aatgcagagt agaggccatg gtttagaac 39

<210> 786

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 786

tatgtaattg aatgcagagt agaggcctag tttagaac 38

<210> 787

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 787

tatgtaattg aatgcagagt agaggcctgg tttagaac 38

<210> 788

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 788

tatgtaattg aatgcagagt agagggccgc agtttagaac 40

<210> 789

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 789

tatgtaattg aatgcagagt agaggccgca ggttagaac 39

<210> 790

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 790

tatgtaattg aatgcagagt agaggccgca ccttagaac 39

<210> 791

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 791

taattgaatg cagagtagag gccgcaggat ggtttag 37

<210> 792

<211> 37

<212> DNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 792

attgaatgca gagtagaggc cgcaggatgc tttagaa 37

<210> 793

<211> 106

<212> RNA

<213> Cypridinia luciferase

<400> 793

gccgaaugug ucaguaugua auugaaugca gaguagaggc cgcaggauag uuuagaacau 60

ucuauggaaa gagauuccag uuccaggaac cugguacaua cguguu 106

<210> 794

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 794

uuaacuuacg ucucaucucc ggcguccuac 30

<210> 795

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 795

cuuacgucuc aucuccggcg uccuaccaaa 30

<210> 796

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 796

cgucucaucu ccggcguccu accaaaucuu 30

<210> 797

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 797

ucaucuccgg cguccuacca aaucuuguaa 30

<210> 798

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 798

cuccggcguc cuaccaaauc uuguaagaua 30

<210> 799

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 799

ggcguccuac caaaucuugu aagauaccuu 30

<210> 800

<211> 30

<212> RNA

<213

> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 800

uccuaccaaa ucuuguaaga uaccuuucuc 30

<210> 801

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> Synthetic Oligonucleotide

<400> 801

accaaaucuu guaagauacc uuucucuaag 30

<210> 802

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 802

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 803

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 803

Glu Ala Ala Ala Lys

1 5

<210> 804

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 804

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly

1 5 10 15

Gly Ser

<210> 805

<211> 16

<212

> PRT

<213> Artificial Sequence

<220><223> Synthetic Peptide

<400> 805

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser

1 5 10 15