(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0078200
(43) 공개일자 2020년07월01일

(51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) *C12N 15/10* (2017.01)
C12N 15/113 (2010.01) *C12N 15/90* (2006.01)
(52) CPC특허분류
C12N 9/22 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2018-0167870
(22) 출원일자 2018년12월21일
심사청구일자 없음

(71) 출원인
(주)지플러스 생명과학
서울특별시 관악구 낙성대로 38, 1층(봉천동)
(72) 발명자
박종진
서울특별시 관악구 낙성대로 38(봉천동)
윤지영
서울특별시 동작구 사당로20길 119, 101동 302호 (사당동, 르메이에르)
(뒷면에 계속)
(74) 대리인
제일특허법인(유)

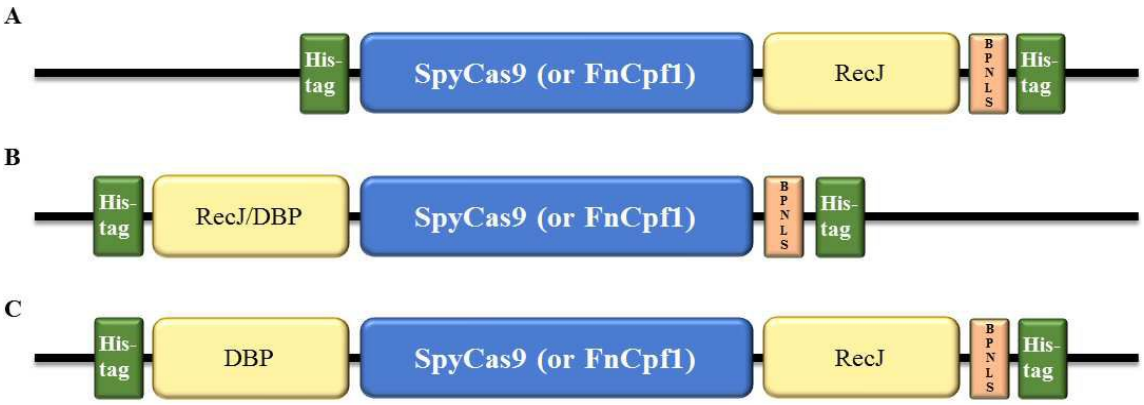
전체 청구항 수 : 총 22 항

(54) 발명의 명칭 **크리스퍼 연관 단백질 및 엑소뉴클레아제를 포함하는 변형된 크리스퍼 연관 단백질**

(57) 요 약

본원 발명은 크리스퍼 연관 단백질 및 엑소뉴클레아제인 RecJ를 포함하는 유전자 편집 기능이 향상된 크리스퍼 플러스 단백질에 관한 것이다. 본원 발명의 일 구체예로 RecJ가 크리스퍼 연관 단백질에 결합된 크리스퍼 플러스 단백질은 종래의 크리스 연관 단백질에 비해 녹아웃 및 녹인 효율이 증가하였다. 따라서, 본 발명의 크리스퍼 플러스는 유전자 편집 효율이 향상된 유전자 가위를 제공하며, 이를 이용할 경우 보다 안정적이고 효율적으로 유전자를 편집할 수 있다. 따라서, 상기 크리스퍼 플러스 단백질은 종래의 크리스퍼 연관 단백질을 대체할 수 있는 효율적인 유전자 편집 기능을 보유한 단백질로서 사업성이 우수할 것으로 예상된다.

대 표 도 - 도12



(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 15/907 (2013.01)
C07K 2319/09 (2013.01)
C07K 2319/21 (2013.01)
C12N 2310/20 (2017.05)

(72) 발명자
최선미
서울특별시 강서구 마곡중앙로 36, 1502동 1604호 (마곡동, 마곡엠밸리15단지)
박미진
서울특별시 동작구 상도로38길 73, 202(상도동)
박슬기
서울특별시 관악구 대학14길 90, 110호(신림동)
박영훈
서울특별시 관악구 낙성대역4길 12-22, 105호(봉천동)
이정혁
서울특별시 관악구 낙성대역16길 10, 304호(봉천동)
임정학
서울특별시 관악구 관악로14길 23, 303호(봉천동, 현대원룸)
김동욱
서울특별시 관악구 남현7길 3, 806호(남현동, 예성그린캐슬2차아파트)
최성화
서울특별시 서초구 양재대로2길 90, 204동 1703호 (우면동, 서초힐스)

명 세 서

청구범위

청구항 1

크리스퍼 연관 단백질 및 엑소뉴클레아제인 RecJ를 포함하는 크리스퍼 플러스 단백질.

청구항 2

제1항에 있어서,

상기 크리스퍼 플러스 단백질은 하기 구조식 1 내지 4에서 선택되는 어느 하나의 구조를 갖는 것인, 크리스퍼 플러스 단백질로서,

<구조식 1>

크리스퍼 연관 단백질-RecJ

<구조식 2>

크리스퍼 연관 단백질-링커 1-RecJ

<구조식 3>

RecJ-크리스퍼 연관 단백질

<구조식 4>

RecJ-링커 1-크리스퍼 연관 단백질

상기 링커 1은 1 내지 20개의 아미노산으로 구성된 것일 수 있다.

청구항 3

제2항에 있어서,

상기 링커 1은 서열번호 17, 18 , 19, 20, 21, 22, 23 또는 24의 아미노산 서열을 가지는 것인, 크리스퍼 플러스 단백질.

청구항 4

제1항에 있어서,

추가적으로 DNA 결합 단백질(DNA binding protein, DBP)를 더 포함하는 것인, 크리스퍼 플러스 단백질.

청구항 5

제4항에 있어서,

상기 DBP는 SSB(single stranded binding protein) 또는 DSB(double stranded binding protein)인 것인, 크리스퍼 플러스 단백질.

청구항 6

제1항에 있어서,

상기 RecJ는 크리스퍼 연관 단백질의 N 말단 또는 C 말단에 결합되는 것인, 크리스퍼 플러스 단백질.

청구항 7

제4항에 있어서,

상기 크리스퍼 플러스 단백질는 하기 구조식 5 내지 7에서 선택되는 어느 하나의 구조를 갖는 것인, 크리스퍼

플러스 단백질로서,

<구조식 5>

DBP-크리스퍼 연관 단백질-RecJ

<구조식 6>

DBP-링커 2-크리스퍼 연관 단백질-RecJ

<구조식 7>

DBP-링커 2-크리스퍼 연관 단백질-링커 1-RecJ

이때, DBP는 SSB 또는 DSB이며,

상기 링커 1은 1 내지 20개의 아미노산으로 구성된 것일 수 있으며,

상기 링커 2는 1 내지 20개의 아미노산으로 구성된 것일 수 있다.

청구항 8

제1항에 있어서,

추가적으로 핵 위치화 시그널(nuclear localization signal)을 더 포함하는 것인, 크리스퍼 플러스 단백질.

청구항 9

제8항에 있어서,

상기 핵 위치화 시그널은 융합 단백질의 C 말단에 위치하는 것인, 크리스퍼 플러스 단백질.

청구항 10

제1항에 있어서,

추가적으로 His tag을 포함하는 것인, 크리스퍼 플러스 단백질.

청구항 11

제1항에 있어서,

상기 크리스퍼 연관 단백질은 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cas12a, Cas12b, Cas12c, Cas12d, Cas12e, Cas13a, Cas13b, Cas13c, Cas13d, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3 및 Csf4로 구성된 군으로부터 선택되는 어느 하나인 것인, 크리스퍼 플러스 단백질.

청구항 12

제1항에 있어서,

상기 RecJ는 서열번호 1, 3, 5, 7, 9, 11, 13 또는 15로 이루어진 군에서 선택되는 어느 하나의 아미노산 서열을 가지는 것인, 크리스퍼 플러스 단백질.

청구항 13

제1항 내지 제12항에 있어서,

상기 크리스퍼 플러스 단백질은 서열번호 31, 41, 49, 53, 57, 71 또는 77로 구성된 군에서 선택되는 어느 하나의 아미노산 서열을 가지는 것인, 크리스퍼 플러스 단백질.

청구항 14

제12항의 크리스퍼 플러스 단백질을 코딩하는 폴리뉴클레오티드.

청구항 15

제14항에 있어서,

상기 폴리뉴클레오티드는 서열번호 32, 42, 50, 54, 58, 72 또는 78로 구성된 군에서 선택되는 어느 하나의 핵산 서열을 가지는 것인, 폴리뉴클레오티드.

청구항 16

제14항의 폴리뉴클레오티드가 적재된 벡터.

청구항 17

제16항에 있어서,

상기 벡터는 플라스미드인, 벡터.

청구항 18

제16항의 벡터가 도입된 숙주 세포.

청구항 19

제1항 내지 제12항의 크리스퍼 플러스 단백질 및 crRNA가 결합된 RNP 복합체.

청구항 20

제19항의 RNP 복합체를 포함하는 유전자 편집용 조성물.

청구항 21

제1항 내지 제12항의 크리스퍼 플러스 단백질 및 crRNA를 유효성분으로 포함하는 유전자 편집용 키트.

청구항 22

제19항에 있어서,

상기 crRNA는 표적 유전자에 상보적으로 결합 가능한 10 내지 30개의 뉴클레오티드로 구성된, 유전자 편집용 키트.

발명의 설명

기 술 분 야

[0001] 본 발명은 엑소뉴클레아제가 크리스퍼 연관 단백질에 결합된 변형된 크리스퍼 연관 단백질 및 이의 용도에 관한 것이다.

배 경 기 술

[0002] 유전체 교정(Genome editing)이란 생명체의 유전정보를 자유롭게 교정하는 기술이다. 생명과학분야의 진보와 유전체 서열 분석 기술의 발전을 통해 우리는 다양한 유전정보에 대해 폭넓게 이해할 수 있게 되었다. 예를 들어 동식물의 번식, 질병과 성장, 다양한 인간 유전질병을 일으키는 유전자 변이, 바이오연료의 생산 등을 위한 유전자에 대한 이해는 이미 확보된 상황이지만 이를 직접적으로 활용하여 생명체를 개선하고, 인간 질병을 치료하는 수준에까지 이르기 위해서는 그 이상의 기술 진보가 필수적이다.

[0003] 유전체 교정 기술은 인간을 포함하여 동물, 식물, 미생물의 유전정보를 변화시켜 그 활용범위를 획기적으로 확장시킬 수 있다. 유전자가위는 원하는 유전정보를 정확히 자를 수 있도록 설계되어 만들어지는 분자 도구로 유전체 교정 기술에서 핵심역할을 하고 있다. 유전자 서열분석 분야를 한 차원 발전시켰던 차세대시퀀싱(Next generation sequencing) 기술과 같이, 유전자가위는 유전정보 활용의 속도와 그 범위를 확장시키고 새로운 산업분야를 창출해 내는 핵심 기술이 되고 있다.

[0004] 지금까지 개발된 유전자 가위는 그 순서에 따라 3세대로 나눌 수 있다. 1세대 유전자 가위는 ZFN(Zinc Finger

Nuclease), 2세대 유전자 가위는 TALEN(Transcription Activator-Like Effector Nuclease), 가장 최근에 연구된 CRISPR(Clustered regularly interspaced short palindromic repeat)-Cas(CRISPR-associated) 9은 3세대 유전자 가위다.

[0005] CRISPRs는 유전자 서열이 밝혀진 박테리아의 대략 40% 및 유전자 서열이 밝혀진 고세균의 90%의 유전체에서 발견되는 여러 짧은 직접 반복을 포함하는 좌위이다. Cas 단백질은 CRISPR RNA(crRNA) 및 trans-activating crRNA(tracrRNA)로 명명된 두 개의 RNA와 복합체를 형성했을 때, 활성 엔도뉴클레아제(endonuclease)를 형성하고, 그렇게 함으로써 파지 또는 플라스미드의 침입에서 외부 유전적 요소를 분해하여 숙주 세포를 보호한다. crRNA는 전달에 외부 침입자로부터 점유되었던 숙주 유전체의 CRISPR 요소로부터 전사된다.

[0006] 이러한 CRISPR-Cas 시스템 유래의 RNA-가이드 뉴클레아제(RNA-guided nuclease)는 유전체를 교정할 수 있는 수단을 제공해준다. 특히, 단일 가이드 RNA(sgRNA)와 Cas 단백질을 이용하여 세포 및 기관의 유전체를 편집할 수 있는 기술과 관련된 연구들이 활발히 진행되어왔다. 최근, Cpf1 단백질(Prevotella 및 Francisella 1 유래)이 CRISPR-Cas 시스템의 또 다른 뉴클레아제 단백질로서 보고되었고(B. Zetsche, et al., 2015), 이에 따라 유전체 교정에 있어서 선택의 폭이 넓어졌다. Cpf1은 5' 돌출부 형태로 절단하는 점, 가이드 RNA의 길이가 더 짧은 점, 시드(seed) 서열 및 절단 위치 사이가 더 길다는 점에서 여러 장점을 가진다. 그럼에도 여전히 보다 우수한 유전자 가위를 개발하기 위하여 많은 연구가 이루어지 있는 실정이다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) B. Zetsche, et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system, Cell, 2015, 163, 759-771.

발명의 내용

해결하려는 과제

[0008] 본 발명자들은 유전체 교정에 있어서 기존에 알려진 크리스퍼 연관 단백질보다 더 효과적인 단백질을 개발하기 위하여 지속적으로 연구한 결과, 표적 핵산 서열을 인식하여 절단하는 효율이 향상될 뿐 아니라, 녹인 효율이 증가된, 엑소뉴클레아제를 포함하는 변형된 크리스퍼 연관 단백질을 개발하여 본 발명을 완성하였다. 따라서, 본 발명의 목적은 표적 핵산 서열을 인식하여 절단하는 효율이 향상된 변형된 크리스퍼 연관 단백질을 제공하는 것이다.

과제의 해결 수단

[0009] 상기 목적을 달성하기 위하여, 크리스퍼 연관 단백질에 엑소뉴클레아제 중 RecJ가 결합된 변형된 크리스퍼 연관 단백질을 제공한다.

발명의 효과

[0010] RecJ가 크리스퍼 연관 단백질에 결합된 크리스퍼 플러스 단백질은 크리스 연관 단백질에 비해 녹아웃(Knockout) 및 녹인(Knockin) 효율이 증가하였다. 따라서, 본원의 변형된 크리스퍼 연관 단백질은 유전자 편집 효율이 향상된 유전자 가위를 제공하며, 이를 이용할 경우 보다 안정적이고 효율적으로 유전자를 편집할 수 있다.

도면의 간단한 설명

[0011] 도 1은 변형된 Cas9 단백질을 체외에서 절단하여 분석한 것이다. 각 RNP의 농도는 25 nM이다. 별표(*)는 약 600 bp에서 나타나는 절단된 템플릿(template)을 의미한다. 각 겔 이미지의 하단 부분에 절단 효율을 표시하였다.

도 2는 SpCas9 및 SpCas9-RecJ에 의한 원형 dsDNA 분해를 확인한 것이다. 선형 dsDNA는 DNA 전기 영동에서 뚜렷하고 두꺼운 dsDNA 크기를 보였다. SpCas9/gRNA C apoproteins를 배양한 후, SpCas9은 시간에 상관없이 단일의 선명하고 두꺼운 dsDNA를 나타내었다. 그러나, SpCas9-RecJ/gRNA C는 60-180분 배양 후에 흐릿하며, 약하고, 얇은, 꼬리가 끌리는 dsDNA형태를 나타내었는데, 이는 DNA 분해를 의미한다.

도 3은 반응 시간에 따른 유전자 편집 효율을 나타낸 것이다. 이는 T7E1 엔도뉴클레아제 활성의 산물이며, 삽입(insertion)-결실률(deletion)(% indel)과 비교하여 측정하였다.

도 4는 CCR5 및 DHCR7에 대한 SpyCas9, SpyCas9-RecJ 및 SpyCas9-T5 간의 HDR 효율을 비교하여 나타낸 것이다. CCR5 및 DHCR7 유전자자리(locus)에서 SpyCas9, SpyCas9-RecJ 및 SpyCas9-T5의 HDR 효율을 나타낸 것이다.

도 5 및 도 6은 유전자 편집 효율을 확인하기 위하여, HEK293 세포에서 SpCas9 및 SpCas9-RecJ의 In vitro 절단 에세이를 수행한 결과를 나타낸다.

도 7은 Cas9 및 Cas9-RecJ이 인간 PD-1 및 CCR5의 유전자에 미치는 온-타겟 효과를 나타낸 것이다.

도 8은 인간 CCR5에 대한 Cas9 및 Cas9-RecJ의 오프 타겟 효과를 확인한 것이다.

도 9는 5개의 오프 타겟팅된 유전자의 In vitro 절단 에세이를 나타낸 것이다.

도 10은 선형 dsDNA를 이용한 In vitro 절단 에세이를 통하여, FnCpf1, FnCpf1-RecJ 및 SpCas9이 가지고 있는 엑소뉴클레아제 기능을 확인한 것이다.

도 11은 선형 dsDNA를 이용하여 FnCpf1 또는 FnCpf1-RecJ 엑소뉴클레아제에 의한 In vitro 절단 에세이 결과를 나타낸 것이다.

도 12는 RecJ를 포함하는 변형된 크리스퍼 연관 단백질의 구조를 나타낸 것이다. A는 SpyCas9 또는 FnCpf1의 C 말단에 RecJ가 융합된 형태이다. B는 SpyCas9 또는 FnCpf1의 N 말단에 RecJ가 결합된 형태이다. C는 SpyCas9 또는 FnCpf1의 C 말단에는 RecJ가 N 말단에는 DBP(DNA Binding Protein)이 결합된 형태이다. 이때, His tag은 6개의 His로 구성되며, BPNLS는 핵 위치화 시그널(nuclear localization signal)을 의미한다.

도 13은 CCR5 및 DHCR7에 대한 SpyCas9의 sgRNA 결합 사이트를 나타낸 것이다. 인간 CCR5(A) 및 인간 DHCR7(B)에 대한 SpyCas9의 sgRNA 결합 영역을 각각 노란색과 핑크색으로 나타낸 것이다. PAM 서열은 녹색으로 표시하였다. 화살표는 뉴클레아제 절단 위치를 나타낸다.

도 14는 CCR5 및 DHCR7에 대한 FnCpf1의 crRNA 결합 사이트를 나타낸 것이다. 인간 CCR5(A) 및 인간 DMNT1(B)에 대한 FnCpf1의 crRNA 결합 영역을 각각 붉은색과 오렌지색으로 나타낸 것이다. PAM 서열은 파란색으로 표시하였다. 화살표는 뉴클레아제 절단 위치를 나타낸다.

도 15 및 도 16은 SpyCas9과 SpyCas9의 C 말단에 다양한 단백질이 결합된 변형된 SpyCas9 사이의 편집 효율을 비교한 것이다. SpyCas9의 C 말단에 GFP, hTdT, 및 엑소뉴클레아제로 알려진 RecJ, RecE, lambda, mungbean, and T5가 결합된 형태이다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율이 서로 상이하다.

도 17은 SpyCas9과 SpyCas9-RecJ의 Knock-out 편집 효율을 비교한 것이다. A는 SpyCas9-RecJ의 구조를 나타낸 것이다. B는 대조군, SpyCas9 및 SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 18은 SpyCas9과 SpyCas9-RecJ의 Knock-in 편집 효율을 비교한 것이다. A는 SpyCas9-RecJ의 구조를 나타낸 것이다. B는 대조군, SpyCas9 및 SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 19는 및 20은 SpyCas9과 SpyCas9의 N 말단 또는 N/C 양말단에 다양한 단백질이 결합된 변형된 SpyCas9 사이의 편집 효율을 비교한 것이다. 변형된 SpyCas9은 SpyCas9의 N 말단에 RecJ, RecE, GFP, 단일 가닥 DNA 결합 단백질(single stranded DNA binding protein, SSB) 및 이중 가닥 DNA 결합 단백질(double stranded DNA binding protein, DSB)이 결합되어 있다. 또한, 변형된 SpyCas9은 SpyCas9의 C 말단에 RecJ가 결합되어 있고, N 말단에 SSB 또는 DSB이 결합되어 있다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율은 서로 상이하다.

도 21은 SpyCas9과 RecJ-SpyCas9 간의 Knock-out 편집 효율을 비교한 것이다. A는 RecJ-SpyCas9의 구조를 나타낸 것이다. B는 대조군, SpyCas9 및 RecJ-SpyCas9의 유전자 편집 효율을 나타낸 것이다. C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 RecJ-SpyCas9의 편집 효율 변화를 백분율로 나타낸 것이다.

도 22는 SpyCas9과 RecJ-SpyCas9의 Knock-in 편집 효율을 비교한 것이다. A는 RecJ-SpyCas9의 구조를 나타낸 것이다. B는 대조군, SpyCas9 및 RecJ-SpyCas9 간의 유전자 편집 효율을 나타낸 것이다. C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 RecJ-SpyCas9의 편집 효율 변화를 백분율로 나타낸 것이다.

도 23은 SpyCas9과 SSB-SpyCas9-RecJ의 Knock-out 편집 효율을 비교한 것이다. A는 SSB-SpyCas9-RecJ의 구조를 나타낸 것이다. B는 대조군, SpyCas9 및 SSB-SpyCas9-RecJ의 유전자 편집 효율을 나타낸 것이다. C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 SSB-SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 24는 SpyCas9과 DSB-SpyCas9-RecJ 간의 Knock-out 편집 효율을 비교한 것이다. A는 DSB-SpyCas9-RecJ의 구조를 나타낸 것이다. B는 대조군, SpyCas9 및 DSB-SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 DSB-SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 25 및 도 26은 FnCpf1과 FnCpf1의 C 말단에 다양한 단백질이 결합된 변형된 FnCpf1 사이의 편집 효율을 비교한 것이다. FnCpf1의 C 말단에 GFP, hTdT, 및 엑소뉴클레아제로 알려진 RecJ, RecE, lambda, mungbean, 및 T5가 결합된 형태이다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율은 서로 상이하다.

도 27은 FnCpf1 및 FnCpf1-RecJ의 Knock-out 편집 효율을 비교한 것이다. A는 FnCpf1-RecJ의 구조를 나타낸 것이다. B는 대조군, FnCpf1 및 FnCpf1-RecJ의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 28은 FnCpf1 및 FnCpf1-RecJ의 Knock-in 편집 효율을 비교한 것이다. A는 FnCpf1-RecJ의 구조를 나타낸 것이다. B는 대조군, FnCpf1 및 FnCpf1-RecJ 간의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 29 및 30은 FnCpf1과 FnCpf1의 N 말단 또는 N/C 양말단에 다양한 단백질이 결합된 변형된 FnCpf1 사이의 편집 효율을 비교한 것이다. 변형된 FnCpf1은 FnCpf1의 N 말단에 RecJ, RecE, GFP, 단일 가닥 DNA 결합 단백질(single stranded DNA binding protein, SSB) 및 이중 가락 DNA 결합 단백질(double stranded DNA binding protein, DSB)이 결합되어 있다. 또한, 변형된 FnCpf1은 FnCpf1의 C 말단에 RecJ가 결합되어 있고, N 말단에 SSB가 결합되어 있다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율은 서로 상이하다.

도 31은 FnCpf1 및 RecJ-FnCpf1의 Knock-out 편집 효율을 비교한 것이다. A는 RecJ-FnCpf1의 구조를 나타낸 것이다. B는 대조군, RecJ-FnCpf1 간의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 RecJ-FnCpf1의 편집 효율 변화를 백분율로 나타낸 것이다.

도 32는 FnCpf1 및 RecJ-FnCpf1 간의 Knock-in 편집 효율을 비교한 것이다. A는 RecJ-FnCpf1의 구조를 나타낸 것이다. B는 대조군, RecJ-FnCpf1의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 RecJ-FnCpf1의 편집 효율 변화를 백분율로 나타낸 것이다.

도 33은 FnCpf1과 SSB-FnCpf1-RecJ의 Knock-out 편집 효율을 비교한 것이다. A는 SSB-FnCpf1-RecJ의 구조를 나타낸 것이다. B는 대조군, FnCpf1 및 SSB-FnCpf1-RecJ의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 SSB-FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 34는 FnCpf1과 SSB-FnCpf1-RecJ의 Knock-in 편집 효율을 비교한 것이다. A는 SSB-FnCpf1-RecJ의 구조를 나타낸 것이다. B는 대조군, FnCpf1 및 SSB-FnCpf1-RecJ의 유전자 편집 효율을 나타낸 것이다. C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 SSB-FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

도 35a, 도 35b 및 도 35c는 SpyCas9 및 SpyCas9의 C 말단에 단백질이 결합된 변형된 SpyCas9에 유도되는 비상동적인 말단 연결에 있어서, ssODN 및 노코다졸(nocodazole)이 미치는 영향을 확인한 것이다. 도 35a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 35b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 35c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, SpyCas9 및 변형된 SpyCas9의 효율에는 큰 영향이 없음을 확인하였다.

도 36a, 도 36b 및 도 36c는 SpyCas9 및 SpyCas9의 N 말단 또는 N/C 양말단에 단백질이 결합된 변형된 SpyCas9에 유도되는 비상동적인 말단 연결에 있어서, ssODN 및 노코다졸이 미치는 영향을 확인한 것이다. 도 36a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 36b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 36c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, SpyCas9 및 변형된 SpyCas9의 효율에는 큰 영향

이 없음을 확인하였다.

도 37a, 도 37b 및 도 37c는 FnCpf1 및 FnCpf1의 C 말단에 단백질이 결합된 변형된 FnCpf1에 유도되는 비상동적인 말단 연결에 있어서, ssODN 및 노코다졸(nocodazole)이 미치는 영향을 확인한 것이다. 도 37a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 37b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 37c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, FnCpf1 및 변형된 FnCpf1의 효율을 향상 시킨다는 점을 확인하였다.

도 38a, 도 38b 및 도 38c는 FnCpf1 및 FnCpf1의 N 말단 또는 N/C 양말단에 단백질이 결합된 변형된 FnCpf1에 유도되는 비상동적인 말단 연결에 있어서, ssODN 및 노코다졸이 미치는 영향을 확인한 것이다. 도 38a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 38b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 38c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, FnCpf1 및 변형된 FnCpf1의 효율에는 큰 영향이 없음을 확인하였다.

도 39a 및 도 39b는 각각의 실험에 대한 타당성을 나타낸 것이다. 최상단에는 on-target 서열을 나타낸 것이며, 미스매치된 핵산은 다른 색으로 표시하였다. GFLAS01은 음성대조군이며; GFLAS02 및 GFLAS05는 SpyCas9이며; GFLAS03, GFLAS04, GFLAS06 및 GFLAS07은 Cas9-RecJ를 이용하여 실험한 결과를 나타낸다.

도 40은 Cas9 및 Cas9-RecJ간의 오프-타겟팅 효율을 비교한 것이다. 5개의 인간 CCR5로 오프-타겟을 테스트하였다. A는 CCR5에 대한 오프-타겟 후보 서열을 나타낸 것이다. B는 딥 시퀀싱(deep sequencing)을 하기 위한 증폭물을 나타낸 것이다.

도 41은 Cas9 및 Cas9-RecJ의 증폭물의 딥 시퀀싱 결과를 나타낸 것이다. Cas9-RecJ은 KCNJ6, CNTPNA2, 및 Ch.5에서는 Cas9에 비해 오프-타겟이 나타나는 비율이 낮다는 것을 확인하였다.

도 42는 SpyCas9의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 43은 SpyCas9-RecJ의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 44는 RecJ-SpyCas9의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 45는 SSB-SpyCas9-RecJ의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 46은 DSB-SpyCas9-RecJ의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 47은 FnCpf1의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 48은 FnCpf1-RecJ의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 49는 RecJ-FnCpf1의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 50은 SSB-FnCpf1-RecJ의 단백질 서열 및 핵산 서열을 도식화한 것이다.

도 51은 변형된 크리스퍼 연관 단백질 및 sgRNA를 이용하여 식물체의 유전자를 편집하는 실험 방법의 개요를 나타낸 것이다.

도 52는 유전자 편집을 하기 위해 준비된 식물체를 나타낸 것이다.

도 53은 식물체의 타겟 위치를 도식화한 것이다.

도 54는 차세대 염기서열 분석법(Next-Generation Sequencing, NGS)을 통해 유전자 편집 후의 식물체의 핵산 서열을 분석한 것이다.

도 55는 변형된 크리스퍼 연관 단백질의 삽입-결실 효율을 나타낸 것이다.

도 56은 변형된 크리스퍼 연관 단백질의 삽입-결실 효율을 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

[0012] 이하 본원 발명을 상세히 설명한다.

[0013] 본 발명은 일 측면으로 크리스퍼 연관 단백질 및 엑소뉴클레아제인 RecJ를 포함하는 변형된 크리스퍼 연관 단백

질을 제공한다. 이때, 상기 변형된 크리스퍼 연관 단백질은 이하에서 크리스퍼 플러스 단백질이라고 지칭할 수 있다.

[0014] 본 명세서에서 사용된 용어, "크리스퍼 플러스 단백질"이란, 크리스퍼 연관 단백질의 유전자 편집 효율이 개선된 크리스퍼 연관 단백질을 지칭한다. 이때, 크리스퍼 연관 단백질의 편집 효율의 개선은 크리스퍼 연관 단백질 이외의 단백질을 결합시킴으로써 달성할 수 있다.

[0015] 상기 크리스퍼 플러스 단백질은 유전자 편집 기능이 향상되었다. 이때, 유전자 편집 기능이란 녹아웃(Knockout) 및 녹인(Knockin) 기능을 의미한다. 녹아웃은 크리스퍼 연관 단백질이 타겟 핵산 서열과 결합하였을 시에, 타겟 핵산의 일부를 결실시키거나, 타겟 핵산의 일부에 뉴클레오티드를 삽입한 경우를 의미한다. 또한, 녹인이란 외래 유전자가 타겟 부위에 삽입되는 것을 의미한다.

[0016] 상기 RecJ는 크리스퍼 연관 단백질의 N 말단 또는 C 말단에 결합될 수 있다. 또한, 크리스퍼 연관 단백질과 RecJ는 직접 결합될 수도 있다. 또한, 상기 크리스퍼 연관 단백질 및 RecJ는 링커 1을 통해 연결되는 것일 수 있다.

[0017] 구체적으로, 상기 크리스퍼 플러스 단백질은 다음 구조 중 어느 하나일 수 있다:

[0018] **<구조식 1>**

[0019] 크리스퍼 연관 단백질-RecJ

[0020] **<구조식 2>**

[0021] 크리스퍼 연관 단백질-링커 1-RecJ

[0022] **<구조식 3>**

[0023] RecJ-크리스퍼 연관 단백질

[0024] **<구조식 4>**

[0025] RecJ-링커 1-크리스퍼 연관 단백질

[0026] 상기 구조식 1 내지 4의 융합 단백질은 N 말단에서 C 말단으로 결합된 것을 나타낸다. 이때, 상기 구조식 1 내지 4에 표시된 링커 1은 1 내지 20개의 아미노산으로 구성된 것일 수 있다. 또한, 상기 링커 1은 5 내지 10개의 아미노산일 수 있으며, 6개, 7개, 8개, 9개, 10개의 아미노산으로 구성될 수 있다. 일 구체예의 링커 1은 서열번호 17, 18, 19, 20, 21, 22, 23 또는 24로 표시되는 아미노산을 가질 수 있다.

[0027] 본 명세서에서 사용된 용어, "크리스퍼 연관 단백질"은 Cas 효소, CRISPR 효소, CRISPR 단백질, Cas 단백질 및 CRISPR Cas는 일반적으로 상호교환적으로 사용될 수 있다.

[0028] 또한, 상기 크리스퍼 연관 단백질은 스트렙토코커스(Streptococcus), 캄필로박터(Campylobacter), 니트라티프락토르(Nitratifractor), 스타필로코커스(Staphylococcus), 파르비바쿨룸(Parvibaculum), 로세부리아(Roseburia), 네이세리아(Neisseria), 글루콘아세토박터(Gluconacetobacter), 아조스피릴룸(Azospirillum), 스파에로카에타(Sphaerochaeta), 락토바실러스(Lactobacillus), 유박테리움(Eubacterium), 코리네박터(Corynebacter), 카르노박테리움(Carnobacterium), 로도박터(Rhodobacter), 리스테리아(Listeria), 팔루디박터(Paludibacter), 클로스트리디움(Clostridium), 라크노스피라세아에(Lachnospiraceae), 클로스트리디아리디움(Clostridiaridium), 렙토트리키아(Leptotrichia), 프란시셀라(Francisella), 레지오넬라(Legionella), 알리사이클로바실러스(Alicyclobacillus), 메타노메티오필러스(Methanomethyophilus), 포르피로모나스(Porphyromonas), 프레보텔라(Prevotella), 박테로이데테스(Bacteroidetes), 헬코코커스(Helcococcus), 레토스피라(Letospira), 데설포비브리오(Desulfovibrio), 데설포나트로눔(Desulfonatronum), 오피투타세아에(Opitutaceae), 투베리바실러스(Tuberibacillus), 바실러스(Bacillus), 브레비바실러스(Brevibacilus), 메틸로박테리움(Methylobacterium) 또는 아시다미노코커스(Acidaminococcus)를 포함하는 속에서 유래된 것일 수 있다.

[0029] 또한, 상기 크리스퍼 연관 단백질은 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cas12a, Cas12b, Cas12c, Cas12d, Cas12e, Cas13a, Cas13b, Cas13c, Cas13d, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3 및 Csf4로 구성된 군으로부터 선택되는 어느 하나 일 수 있다. 일 구체예로 상기 크리스퍼 연관 단백질은 Cas9 또는 Cas12a 일 수 있다. 일 실시

예로 Cas9은 SpyCas9일 수 있다. 또한, Cpf1은 FnCpf1, AsCpf1 또는 LbCpf1 일 수 있으며, 일 실시예로 Cpf1은 FnCpf1 일 수 있다. 상기 크리스퍼 연관 단백질은 종래에 알려진 서열과 95%, 96%, 97%, 98%, 99%의 상동성을 가질 수 있다.

[0030] 이때, 상기 RecJ는 엑소뉴클레아제로 알려진 단백질이다. 이때, 상기 RecJ의 일 구체예로는 에스케리치아(Escherichia) 속 유래의 RecJ(서열번호 1), 스트렙토코코스(Streptococcus) 속 유래의 RecJ(서열번호 3), 락토코코스(Lactococcus) 속 유래의 RecJ(서열번호 5), 바실러스(Bacillus) 속 유래의 RecJ(서열번호 7), 에르리치아(Ehrlichia) 속 유래의 RecJ(서열번호 9), 클라미디아(Chlamydia) 속 유래의 RecJ(서열번호 11), 피네골디아(Finegoldia) 속 유래의 RecJ(서열번호 13), 또는 스트렙토코코스(Streptococcus) 속 유래의 RecJ(서열번호 15)일 수 있다.

[0031] 본 발명은 또 다른 측면으로, 크리스퍼 연관 단백질, RecJ 및 DNA 결합 단백질(DNA binding protein, DBP)를 포함하는 크리스퍼 플러스 단백질 일 수다.

[0032] 상기 RecJ 및 DBP는 크리스퍼 연관 단백질의 N 말단 또는 C 말단에 결합될 수 있다. 또한, 이때, 상기 크리스퍼 연관 단백질, RecJ 및 DBP는 링커을 통해 연결되는 것일 수 있다. 이때, DBP는 SSB(single strand binding protein) 또는 DSB(double strand binding protein) 일 수 있다. 상기 크리스퍼 플러스 단백질의 일시예는 다음 구조 중 어느 하나일 수 있다:

[0033] **<구조식 5>**

[0034] DBP-크리스퍼 연관 단백질-RecJ

[0035] **<구조식 6>**

[0036] DBP-링커 2-크리스퍼 연관 단백질-RecJ

[0037] **<구조식 7>**

[0038] DBP-링커 2-크리스퍼 연관 단백질-링커 1-RecJ

[0039] 상기 구조식 5 내지 7의 융합 단백질은 N 말단에서 C 말단으로 결합된 것이다. 크리스퍼 연관 단백질, RecJ 및 DBP는 직접 결합될 수도 있으며, 링커를 통해 연결될 수 있다. 이때, 링커 1은 상술한 바와 같다. 또한, 상기 구조식 5 내지 7에 표시된 링커 2는 1 내지 20개의 아미노산으로 구성된 것일 수 있다. 또한, 상기 링커 2는 5 내지 10개의 아미노산일 수 있으며, 6개, 7개, 8개, 9개, 10개의 아미노산으로 구성될 수 있다. 일 구체예의 링커 2는 서열번호 17, 18, 19, 20, 21, 22, 23 또는 24로 표시되는 아미노산 서열을 가질 수 있다.

[0040] 본 발명은 또 다른 측면으로, 크리스퍼 연관 단백질 및 RecJ를 포함하는 크리스퍼 플러스 단백질 및 크리스퍼 연관 단백질, RecJ 및 DBP를 포함하는 크리스퍼 플러스 단백질에 추가적으로 추가적으로 핵 위치화 시그널(nuclear localization signal, NLS)을 적어도 하나를 더 포함하는 크리스퍼 플러스 단백질 일 수 있다. 이때, 핵 위치화 시그널은 bpNLS(biparticle nuclear localization sequence)일 수 있다. 적어도 하나 이상의 NLS가 C-말단 또는 N-말단 부착될 수 있다. 이때, 일 구체예에서는 상기 핵 위치화 시그널은 융합 단백질의 C 말단에 위치할 수 있다. 상기 핵 위치화 시그널은 크리스퍼 플러스 단백질의 활성에는 영향을 미치지 않으나, 크리스퍼 플러스 단백질이 세포의 핵내로 이동할 수 있도록 도와주는 펩타이드를 의미한다.

[0041] 또한, 본 발명의 다른 측면은, 상기 크리스퍼 플러스 단백질에 추가적으로 추가적으로 His tag를 더 포함할 수 있다. 이때, 상기 His tag는 크리스퍼 플러스 단백질의 C 말단에 위치할 수 있다. 상기 His tag는 크리스퍼 플러스 단백질의 활성에는 영향을 미치지 않으나, 크리스퍼 플러스 단백질의 분리 정제에 사용될 수 있다.

[0042] 또한, 상기 크리스퍼 플러스 단백질은 서열번호 31, 41, 49, 53, 57, 71 또는 77으로 표시되는 아미노산 서열을 가지는 것 일 수 있다.

[0043] 본 명세서에서 사용된 용어, "PAM"은 프로토스페이서 인접 모티프(protospacer adjacent motif) 또는 PAM-유사 모티프는 크리스퍼 연관 단백질 복합체가 타겟 유전자좌에 결합하는 것을 지시한다.

[0044] 본 발명의 일 구체예에서, 크리스퍼 연관 단백질은 FnCpf1과 관련이 있는 PAM은 5' TTN이며, 여기서 N은 A/C/G 또는 T이다. PAM은 타겟 유전자의 5' 말단의 상류에 위치할 수 있다. Cpf1 패밀리의 T-풍부 PAM은 AT-풍부 게놈의 표적화 및 편집을 가능하게 한다. 반면, Cas9 단백질의 PAM은 타겟 유전자의 3' 말단 하류에 위치할 수 있다.

[0045] 본 발명은 또 다른 측면으로, 상술한 크리스퍼 플러스 단백질을 코딩하는 폴리뉴클레오티드를 제공한다.

[0046] 상기 뉴클레오티드는 신호서열(signal sequence) 또는 리더 서열(leader sequence)을 추가적으로 포함할 수 있다. 여기에서 사용된 용어 "신호서열"은 목적 단백질의 분비를 지시하는 신호펩타이드를 코딩하는 핵산을 의미한다. 상기 신호펩타이드는 숙주 세포에서 번역된 후에 절단된다. 구체적으로, 본원 발명의 신호서열은 ER(endoplasmic reticulum) 막을 관통하는 단백질의 이동을 개시하는 아미노산 서열을 코딩하는 뉴클레오티드일 수 있다.

[0047] 신호서열은 당업계에 그 특징이 잘 알려져 있으며, 통상 16 내지 30개의 아미노산 잔기를 포함하나, 그보다 더 많거나 적은 아미노산 잔기를 포함할 수 있다. 통상적인 신호 펩타이드는 기본 N-말단 영역, 중심 소수성 영역, 및 보다 극성인(polar) C-말단 영역의 세 영역으로 구성된다. 중심 소수성 영역은 미성숙 폴리펩티드가 이동하는 동안 막지질 이중층을 통하여 신호서열을 고정시키는 4 내지 12개의 소수성 잔기를 포함할 수 있다.

[0048] 이때, 상기 크리스퍼 플러스 단백질을 코딩하는 뉴클레오티드의 일 실시예는 서열번호 32, 42, 50, 54, 58, 72 또는 72로 표시되는 서열을 가질 수 있다.

[0049] 본 발명의 또 다른 측면으로, 상기 폴리뉴클레오티드가 적재된 벡터를 제공한다.

[0050] 본원 발명에서 사용된 용어 "벡터"는 숙주 세포에 도입되어 숙주 세포 유전체 내로 재조합 및 삽입될 수 있다. 또는 상기 벡터는 에피좀으로서 자발적으로 복제될 수 있는 뉴클레오티드 서열을 포함하는 핵산 수단으로 이해된다. 상기 벡터는 선형 핵산, 플라스미드, 파지미드, 코스미드, RNA 벡터, 바이러스 벡터 및 이의 유사체들을 포함한다. 바이러스 벡터의 예로는 레트로바이러스, 아데노바이러스, 및 아데노-관련 바이러스를 포함하나 이에 제한되지 않는다. 또한, 상기 플라스미드는 항생제 내성 유전자와 같은 선별 마커를 포함할 수 있고, 플라스미드를 유지하는 숙주 세포는 선택적인 조건하에서 배양될 수 있다.

[0051] 본 발명의 또 다른 측면으로, 상기 벡터가 도입된 숙주 세포를 제공한다.

[0052] 상기 숙주 세포는 벡터를 도입하여 본원의 크리스퍼 플러스 단백질을 생산할 수 있다면 종류에 한정되는 것은 아니다. 일 구체예로 상기 숙주세포는 대장균일 수 있다.

[0053] 본 발명의 또 다른 측면으로, 상기 크리스퍼 플러스 단백질 및 crRNA가 결합된 RNP 복합체를 제공한다.

[0054] 일 구체예에서, 상기 복합체의 핵산 성분는 직접 반복 서열에 연결된 가이드 서열을 포함할 수 있다. 상기 직접 반복 서열은 하나 이상의 줄기 루프(stem loop) 또는 최적화된 2차 구조를 포함할 수 있다. 바람직한 구현예에서, 직접 반복은 최소한의 길이가 16 nt이고 단일 줄기 루프를 가질 수 있다. 추가 구현예에서, 직접 반복은 길이가 16 nt보다 더 길며, 바람직하게는 17 nt 초과이고, 하나 초과의 줄기 루프 또는 최적화된 2차 구조를 가질 수 있다. 바람직한 구현예에서, 직접 반복은 변형되어 하나 이상의 단백질-결합 RNA 압타머를 포함할 수 있다.

[0055] 이때, 상기 crRNA는 타겟 유전자에 상보적으로 결합할 수 있는 핵산 서열을 갖는다. 이때, 타겟 유전자는 세포 내 DNA 분자에 포함될 수 있다. 세포는 원핵 세포 또는 진핵 세포일 수 있다. 세포는 포유동물 세포일 수 있다. 포유동물 세포는 비인간 영장류, 소, 돼지, 설치류 또는 마우스 세포일 수 있다. 세포는 비포유동물 진핵 세포, 예컨대 가금류, 어류 또는 새우일 수 있다. 세포는 또한 식물 세포일 수 있다. 식물 세포는 카사바, 옥수수, 수수, 밀, 담배, 애기장대, 채소 또는 쌀 일 수 있다.

[0056] 본 발명에 의해 세포에 도입된 변형은, 세포 및 세포 의 자손이 생물학적 생성물, 예컨대 항체, 전분, 알코올 또는 기타 다른 원하는 세포 산출물의 개선된 생성을 위하여 변경되도록 하는 것일 수 있다. 본 발명에 의해 세포에 도입된 변형은 세포 및 세포의 자손이 생성된 생물학적 생성물을 변화시키는 변경을 포함하도록 하는 것일 수 있다.

[0057] 또한, 본 발명의 또 다른 측면으로, 상기 RNP 복합체를 유효성분으로 포함하는 유전자 편집용 조성물 또는 유전자 편집용 키트를 제공한다.

[0058] 이하, 본원 발명을 하기 실시예에 의하여 더욱 상세하게 설명한다. 단, 하기 실시예는 본원 발명을 예시하기 위한 것일 뿐, 본원 발명의 범위가 이들만으로 한정되는 것은 아니다.

[0059] **I. 크리스퍼 플러스 단백질의 제조 및 활성 확인**

[0060] **제조예 1. Cpf1 및 RecJ외의 다른 변형된 단백질이 결합된 변형된 크리스퍼 연관 단백질 제조**

[0061] 이콜라이 단백질 발현 벡터 pET28a (addgene, Watertown, MA, USA) 를 NotI 제한 효소를 사용하여서 FnCpf1

(3,902 bp)의 정방향 프라이머 (CGTCGACAAGCTTGCGGCCGCATGTCAATTTATCAAGAATT) 와 역방향 프라이머 (TGGTGGTGCTCGAGTGCGGCCGCTTgGTTATTCCTATTCTGCAC) 를 이용하여 pY002 (addgene, Watertown, MA, USA)를 중합효소 연쇄반응으로 얻은 산물을 깁슨 어셈블리 킷트 (NEB, Ipswich, MA, USA) 방법으로 50도에 25분 인큐베이션으로 통하여 클로닝하였다. 결과적으로 얻어진 pET28a-FnCpf1의 말단에 PspXI 제한효소를 사용하여서 RecJ (1,731 bp), RecE (2,598 bp), SSB (348 bp)는 이콜라이 지놈에서 증폭하였고, mungbean (1,065 bp)은 숙주 나물에서 mRNA 산물을 얻어서 cDNA로 전환한 다음 코딩 영역을 얻을 수 있었다. T5 (828 bp), DSB (267 bp)는 GenScript (Piscataway, NJ, USA)에서 합성하여서 사용하였다. GFP (714 bp) 서열은 생명공학연구원에서 개발한 pCTV-GFP 벡터에서 증폭하여서 사용하였고, hTdT (1,527 bp)는 GenScript에서 DNTT(NM_004088) ORF Clone을 주문하여서 증폭하였다. 증폭된 중합효소반응 산물은 깁슨 어셈블리 방법을 통해서 Cpf1의 말단에 삽입되었다.

표 1

[0062]

name	sequence	size	function
RecJF1	GCGGCCGCACTCGACctgcagGTGAAACAACAGATACAACTTCGT	45	RecJ full-length in back
RecJR1	ccggccttttcgtggccgccggccttttctgcagAATTGGCCAGATATTGTCGATGATA	60	RecJ full-length in back
RecJR2	TCAGTGGTGGTGGTGGTGctttttctttttgcctggccggccttttcgtggccgc	60	RecJ full-length in back
MungbeanF1	CGGCCGCACTCGACctgcagATGCAAACGTTACAGATGAGT	41	Mungbean Nuclease in back
MungbeanR1	cgtggccgccggccttttctgcagACCATTGTAAGAGATAGGTC	44	Mungbean Nuclease in back
T5exonuF1	CGGCCGCACTCGACctgcagATGGCTTCCCGTCGTAATCTAATG	44	T5 exonuclease in back
T5exonuR5	gtggccgccggccttttctgcagTTGTTCTGCAATCTCCAAAATA	45	T5 exonuclease in back
SSBF1	GCGGCaGCACTCGACctgcagATGGCCAGCAGAGGCGTAAACAAGG	46	SSB isolation in back
SSBR1	tcgtggccgccggccttttctgcagACGACCACCCAGCATCTGCATGGT	49	SSB isolation in back
RecEF1	GCGGGCGCACTCGACctgcagATGAGCACAAAACCACTCTTCC	43	RecE in back
RecER1	ttcgtggccgccggccttttctgcagGTCATTTGCATATTCCTTAGCCCA	50	RecE in back
DSBF1	AGCGGCCGCACTCGACctcgagATGGCTAAAAAAGAAATGGTTGAATT	21	with DSBR1 in back
DSBR1	ctttttcgtggccgccggccttttctcgagTTTAGAGAAAACTGTGTCA	37	with DSBF1 in back
T5exoF5	GGTGGACAGCAAATGGGTCGCGGATCCATGGCTTCCCGTCGTAATCTAA	49	T5 full-length in back
T5exoR5	GCTTGTCGACGGAGCTCGAATTCGGATCCTTGTTCTGCAATCTCCAAA	48	T5 full-length in back
pAGROF2	gtctcagctgggaggcgacgaaATGAGTAAAGGAGAAGAACTTTT	45	GFP in back
pAGROR2	gcggaggggttggatcaaagtgaaAATAACCTCTCCTTCTTTTTC	45	GFP in back
HsTdTF2	GCGGCCGCACTCGACctcgagATGGATCCACCACGAGCGTCCCAC	45	HsTdT full length in back
HsTdTR2	ttcgtggccgccggccttttctcgagGGCATTTCTTTCCCACGGTTCAATAT	52	HsTdT full length in back

[0063] **제조예 2. 타겟 sgRNA 또는 crRNA의 제작**

[0064] **제조예 2.1. CCR5 및 DHCR7을 타겟팅하는 sgRNA의 제작**

[0065] CCR5 및 DHCR7을 타겟팅하며, SpyCas9와 함께 이용될 수 있는 sgRNA를 제작하였다. sgRNA 결합사이트는 도면 13에 나타낸 바와 같다. 인간 CCR5(A) 및 인간 DHCR7(B)에 대한 SpyCas9의 sgRNA 결합 영역을 각각 노란색과 핑크색으로 나타낸 것이다. PAM 서열은 녹색으로 표시하였다. 화살표는 뉴클레아제 절단 위치를 나타낸다.

[0066] **제조예 2.2. CCR5 및 DHCR7을 타겟팅하는 crRNA의 제작**

[0067] CCR5 및 DHCR7을 타겟팅하며, FnCpf1과 함께 이용될 수 있는 crRNA를 제작하였다. crRNA 결합사이트는 도면 14에 나타낸 바와 같다. 인간 CCR5(A) 및 인간 DMNT1(B)에 대한 FnCpf1의 crRNA 결합 영역을 각각 붉은색과 오렌지색으로 나타낸 것이다. PAM 서열은 파란색으로 표시하였다. 화살표는 뉴클레아제 절단 위치를 나타낸다.

[0068] **실시예 1. Cas9 및 RecJ가 결합된 변형된 크리스퍼 연관 단백질 제조**

[0069] **실시예 1.1. SpCas9 제조**

[0070] 이콜라이 단백질 발현 벡터 pET28a를 NotI 제한 효소를 사용하여서 SpCas9(4,104 bp)의 정방향 프라이머(CCGTCGACAAGCTTGCGGCCGCATGGACAAGAAGTACAGCATCGGCCTGGA) 와 역방향 프라이머(TGAGCCAGCTGGGCGGCGACGCGGCCGCACTCGACctcgagaaaaggccggcggccacgaaaaaggccggccaggcaaaaaagaaaaagCACCACCACCACCACCACTGA) 를 이용하여 pET28a/Cas9-cys (addgene, Watertown, MA, USA)에서 중합효소 연쇄반응으로 얻은 산물을 깁슨 어셈블리 킷트 (NEB, Ipswich, MA, USA) 방법으로 50도에 25분 인큐베이션으로 통하여 클로닝하였다. 서열 정보는 도42-도50에 명시하고 있다.

[0071] **실시예 1.2. RecJ이 결합된 SpCas9의 제조**

[0072] 결과적으로 얻어진 pET28a-SpCas9의 말단에 PspXI 제한효소를 사용하고 SpCas9 앞쪽에는 BamHI 제한 효소를 사용하여서 RecJ (1,731 bp), RecE (2,598 bp), SSB (348 bp)는 이콜라이 지놈에서 증폭하였고, mungbean (1,065 bp)은 숙주 나물에서 mRNA 산물을 얻어서 cDNA로 전환한 다음 코딩 영역을 얻을 수 있었다. T5 (828 bp), DSB (267 bp)는 GenScript (Piscataway, NJ, USA)에서 합성하여서 사용하였다. GFP (714 bp) 서열은 생명공학연구원에서 개발한 pCTV-GFP 벡터에서 증폭하여서 사용하였고, hTdT (1,527 bp)는 GenScript에서 DNTT(NM_004088) ORF Clone을 주문하여서 증폭하였다. 증폭된 중합효소반응 산물은 깁슨 어셈블리 방법을 통해서 50도에 25분 인큐베이션을 통해서 SpCas9의 말단과 상단에 각각 삽입되었다.

표 2

[0073]

name	sequence	size	function
RecJF2	GGACAGCAAATGGGTCGCGGATCCGTGAAACAACAGATACAACTTCGT	45	RecJ full-length in front
RecJR2	TTGTCGACGGAGCTCGAATTCGGATCCAATTGGCCAGATATTGTCGATGATA	60	RecJ full-length in front
MungbeanF2	GGACAGCAAATGGGTCGCGGATCCATGCAAACGTTACAGATGAGT	41	Mungbean Nuclease in front
MungbeanR2	TTGTCGACGGAGCTCGAATTCGGATCCACCATTGTAAGAGATAGGTC	44	Mungbean Nuclease in front
T5exonuF2	GGACAGCAAATGGGTCGCGGATCCATGGCTTCCCGTCGTAATCTAATG	44	T5 exonuclease in front
T5exonuR6	TTGTCGACGGAGCTCGAATTCGGATCCTTGTTCTGCAATCTCCAAAATA	45	T5 exonuclease in front
SSBF2	GGACAGCAAATGGGTCGCGGATCCATGGCCAGCAGAGGCGTAAACAAGG	46	SSB isolation in front
SSBR2	TTGTCGACGGAGCTCGAATTCGGATCCACGACCACCCAGCATCTGCATGGT	49	SSB isolation in front

RecEF2	GGACAGCAAATGGGTCGCGGATCCATGAGCACAAAACCACTCTTCC	43	RecE in front
RecER2	TTGTCGACGGAGCTCGAATTCGGATCCGTCATTTGCATATTCCTTAGCCCA	50	RecE in front
DSBF2	GGACAGCAAATGGGTCGCGGATCCATGGCTAAAAAAGAAATGGTT	21	with DSBR1 in front
DSBR2	TTGTCGACGGAGCTCGAATTCGGATCCTTTAGAGAAAACTGTGT	37	with DSBF1 in front
pAGROF3	gtctcagctgggaggcgacgaaATGAGTAAAGGAGAAGAACTTTT	45	GFP in front
pAGROR3	TTGTCGACGGAGCTCGAATTCGGATCCAATAACCTCTCCTTCTTTTTC	45	GFP in front
HsTdTF3	GCGGCCGCACTCGACctcgagATGGATCCACCACGAGCGTCCCAC	45	HsTdT full length in front
HsTdTR3	TTGTCGACGGAGCTCGAATTCGGATCCGGCATTTCTTTCCCACGGTTCAATAT	52	HsTdT full length in front
RecJF1	GCGGCCGCACTCGACctgcagGTGAAACAACAGATACAACTTCGT	45	RecJ full-length in back
RecJR1	ccggccttttt cgtggccgccggccttttctgcagAATTGGCCAGATATTGTCGATGATA	60	RecJ full-length in back
RecJR2	TCAGTGGTGGTGGTGGTGctttttctttttgcctggccggccttttt cgtggccgc	60	RecJ full-length in back
MungbeanF1	CGGCCGCACTCGACctgcagATGCAAACGTTACAGATGAGT	41	Mungbean Nuclease in back
MungbeanR1	cgtggccgccggcctttt ctgcagACCATTGTAAGAGATAGGTC	44	Mungbean Nuclease in back
T5exonuF1	CGGCCGCACTCGACctgcagATGGCTTCCCGTCGTAATCTAATG	44	T5 exonuclease in back
T5exonuR5	gtggccgccggcctttt ctgcagTTGTTCTGCAATCTCCAAAATA	45	T5 exonuclease in back
SSBF1	GCGGCaGCACTCGACctgcagATGGCCAGCAGAGGCGTAAACAAGG	46	SSB isolation in back
SSBR1	tcgtggccgccggcctttt ctgcagACGACCACCCAGCATCTGCATGGT	49	SSB isolation in back
RecEF1	GCGGGCGCACTCGACctgcagATGAGCACAAAACCACTCTTCC	43	RecE in back
RecER1	ttcgtggccgccggcctttt ctgcagGTCATTTGCATATTCCTTAGCCCA	50	RecE in back
DSBF1	AGCGGCCGCACTCGACctcgagATGGCTAAAAAAGAAATGGTTGAATT	21	with DSBR1 in back
DSBR1	ctttttcgtggccgccggcctttt ctcgagTTTAGAGAAAACTGTGTCA	37	with DSBF1 in back
T5exoF5	GGTGGACAGCAAATGGGTCGCGGATCCATGGCTTCCCGTCGTAATCTAA	49	T5 full-length in back
T5exoR5	GCTTGTCGACGGAGCTCGAATTCGGATCCTTGTTCTGCAATCTCCAAA	48	T5 full-length in back
pAGROF2	gtctcagctgggaggcgacgaaATGAGTAAAGGAGAAGAACTTTT	45	GFP in back
pAGROR2	gcggaggggttggatcaaagtgaaAATAACCTCTCCTTCTTTTTC	45	GFP in back
HsTdTF2	GCGGCCGCACTCGACctcgagATGGATCCACCACGAGCGTCCCAC	45	HsTdT full length in back
HsTdTR2	ttcgtggccgccggcctttt ctcgagGGCATTTCTTTCCCACGGTTCAATAT	52	HsTdT full length in back

[0074] **실시예 2. Cpf1 및 RecJ가 결합된 변형된 크리스퍼 연관 단백질 제조**

[0075] **실시예 2.1. FnCpf1 제조**

[0076] 이콜라이 단백질 발현 벡터 pET28a를 NotI 제한 효소를 사용하여서 FnCpf1 (3,902 bp)의 정방향 프라이머 (CGTCGACAAGCTTGCGGCCGCATGTCAATTTATCAAGAATT) 와 역방향 프라이머 (TGGTGGTGCTCGAGTGCGGCCGCTTgGTTATTCCTATTCTGCAC) 를 이용하여 pY002 (addgene, Watertown, MA, USA) 을 중합효소 연쇄반응으로 얻은 산물을 깁슨 어셈블리 킷트 (NEB, Ipswich, MA, USA) 방법으로 50도에 25분 인큐베이션으로 통하여 클로닝하였다. 서열 정보는 도42-도50에 명시하고 있다.

[0077] **실시예 2.2. RecJ이 결합된 변형된 FnCpf1의 제조**

[0078] 해당 플라스미드를 지닌 대장균 세포를 카나마이신 항생제가 포함된 LB 배지에 접종하여 37℃에서 밤새 배양한다. 다음 날 배양액을 LB 20 mL에 접종하여 2~3시간 키운 뒤 이를 LB 1L에 넣어 OD_{600}이 0.5~0.7이 될 때까지 37℃에서 배양한다. 1 mM IPTG를 넣고 18℃에서 20시간 배양한다. 다음 날 배양액을 4000 rpm에서 30분간 원심분리하여 상층액은 버리고 펠릿에 용해 완충액(lysis buffer, 5 mM imidazole, 20 mM Tris, 500 mM sodium chloride, 1 mM DTT, 1 mM PMSF, pH 8.0)을 넣어 현탁한다. 현탁액을 초음파-파쇄(sonication)한 후 15000 xG에서 1시간 동안 원심분리하여 상층액만을 분리하고 히스티딘 태깅 컬럼을 사용한 FPLC에 주입하여 단백질을 분리한다. 분리한 단백질 분획을 다시 탈염(desalting) 컬럼을 사용한 FPLC에 주입하여 단백질을 분리 후 해당 분획을 농축한다. 농축된 단백질을 Bradford assay로 정량한 뒤 -80℃에 분주하여 보관한다. 변형된 FnCpf1의 구체예를 도 12에 나타내었다.

[0079] **실험예 1. 변형된 Cas9 단백질의 절단 효율 확인 : 선형 dsDNA**

[0080] 단백질과 gRNA는 몰농도 비율이 1:1.2가 되도록 준비하였다. 반응 부피 20 μL로 하여 단백질 25 nM에 gRNA 30 nM을 섞은 후 20분 상온에 두고 RNP가 형성되도록 하였다. SpCas9-BPNLS의 분자량을 기준으로 각 단백질의 분자량을 반영하여 넣어줄 단백질의 양을 계산하였다. SpCas9-BPNLS의 분자량은 166.580 kDa이며 25 nM에 해당하는 양은 83.5 g이다. gRNA는 CCR5 유전자를 타겟으로하며 약 33,500 g/mol의 농도를 가지고 있고 30 nM에 해당하는 양은 20 ng이다. 타겟 유전자인 CCR5 주형의 크기는 약 1.4 kb이다.

[0081] 1.5 kb 크기의 타겟 유전자인 CCR5 주형 DNA 500 ng을 RNP 복합체와 섞은 후 37℃에서 1시간 반응시켰다. 2% agarose gel에 전기영동한 후 겔 이미지를 수득하였다. 이를 이용하여 주형 DNA와 RNP 복합체가 작용하고 남은 주형 DNA의 비율을 계산하여, 단백질의 뉴클레아제 활성 효율을 수득하였다. 절단 후에는 약 600 bp 정도의 밴드가 나타났다(도 1).

[0082] **실험예 2. 변형된 Cas9 단백질의 절단 효율 확인 : 원형 dsDNA**

[0083] TAAGG 또는 AAAC 오버행을 가지는 목표서열 20개의 정방향과 역방향 단일 가닥 올리고뉴클레오타이드 2개로 이루어져 있으며, 95℃에서 5분 후, 55℃에서 10분의 인큐베이션하여 올리고뉴클레오타이드가 이차구조를 가지지 않도록 하면서 20개의 정방향과 역방향의 올리고뉴클레오타이드의 단일 가닥을 상보적으로 결합시켜서 20개 베이스페어를 가지는 이중 가닥의 선형으로 정렬시켰다. 이중 가닥의 올리고뉴클레오타이드는 Cas9에 대한 gRNA 지지체를 포함하는 클로닝 벡터의 2개 BsaI 사이트 (A↓TAGGTGAGACCGCAGGTCTCG↓GTTTT)에 클로닝하였다.

[0084] 결과적으로, sgRNA는 pGEM_447 vector(서열번호 102)를 목표로하는 20개의 목표서열 (목표서열; AAGCCCGGCCGAAATGACCA) 이 T7 프로모터와 sgRNA 지지체 사이에 위치하는 클로닝 벡터에서 T7 프로모터를 주형으로 하는 정방향 프라이머 (정방향 프라이머 서열; ATTCTAATACGACTCACTATAGG)와 sgRNA 지지체의 마지막 20 개의 서열을 주형으로 하는 역방향 프라이머 (역방향 프라이머 서열; GCACCGACTCGGTGCCACTT)를 사용하여서 중합효소 연쇄방반응으로 얻어지는 PCR 산물을 MEGAshortscript™ High Yield Transcription Kit(Invitrogen, Grand Island, NY)를 이용하여 RNA 산물인 sgRNA 산물을 얻었다. RNA 산물은 DNase I (NEB, Ipswich, MA)을 이용하여 dsDNA 주형을 제거하기 위해 절단하였고, MEGAclear™ 전사 Clean-Up kit (Invitrogen, Grand Island, NY)를 이용하여 정제하였다.

[0085] 절단 시험 전의 Cas9-sgRNA 복합체는 NEB2.1 완충액(NEB, Ipswich, MA)을 이용하여 37℃에서 10분간 배양하였다. 절단 시험은 20 μL의 반응 내 0.2-1.6 μg Cas9 단백질, 500 ng sgRNA 및 500 ng dsDNA 기질로 구성되어 있으며, 37℃에서 0 내지 180분 동안 수행되었다. 절단된 dsDNA는 40분 동안 DNA 전기 영동으로 확인하였다(도

2).

[0086] **실험예 3. 시간에 따른 변형된 Cas9 단백질의 편집 효율 확인**

[0087] HEK293T 세포를 10% 태아 소혈청인 FBS와 P/S가 함유된 DMEM 배지에서 37℃, 5% CO_2 배양기에서 유지시켰다. 트랜스펙션 전날 150,000 세포를 24 well 플레이트로 옮겼다. Cas9 변이 단백질을 세포에 트렌스펙션 시키기 위해 각각 정제된 Cas9 변이체 단백질을 25 ㎕ Opti-MEM 혈청 환원 배지에 첨가한 후 120 ng의 가이드 RNA를 첨가하였다. Cas9 단백질에 대한 가이드 RNA의 몰비는 거의 1.2 : 1의 비율로 유지하였다. 혼합물을 부드럽게 두드린 후 잘 혼합하여 10분 동안 실온에서 배양하였다. 다른 튜브에는 2 ㎕의 lipofectamine 3000을 25 ㎕ Opti-MEM 혈청 환원 배지에 넣어주었고, 희석된 25 ㎕의 시약은 가이드 RNA와 Cas9 단백질 복합체에 넣고 15분 동안 상온에서 배양한 뒤 총 50 ㎕의 혼합물을 세포에 넣어주었다.

[0088] 게놈 DNA 추출은 제조사의 지시에 따라 PureLink Genomic DNA 키트를 사용하여 진행하였다. 간단하게 세포를 PBS로 1회 세척하고 트립신으로 37℃에서 5분 동안 인큐베이션하였다. 그 후, 세포를 얻기 위해 1.2 ml의 성장 배지를 세포에 첨가 한 후 상온, 250xg에서 5분 동안 스핀다운시켰다. 게놈상 인간 DHCR7의 유전자 증폭을 위해, 두 개의 상이한 PCR 프라이머 쌍이 사용되었다. 전방 프라이머 1: 5'-CAGTAGAGCAGGCATGTTGAGT-3'와 역방향 프라이머 1 : 5'-GTGAAGGTGTATCAAACGCTGA-3' 를 사용하였고, 이는 T7E1 분석 후 크기가 각각 427 및 205의 절단된 2개의 DNA 밴드를 나타낸다. 반면, 다음의 다른 프라이머 쌍을 사용할 때는 절단된 단일 DNA 단편만 생성된다 : 정방향 프라이머 2 : 5'-GGGAAACCACTGGCCTTGG-3'; 역방향 프라이머 2 : 5'-GAGCCAGGATCCATGTCCCA-3'.

[0089] 게놈 DNA를 증폭 후 Wizard® SV 젤 및 PCR 클린업 시스템을 사용하여 정제하였다. 200 ng의 DNA를 PCR 튜브에 넣고 NEB 완충액 2번을 1 ㎕와 증류수 9 ㎕를 넣었다. 이후, 혼합물을 다음 조건을 사용하여 PCR 기계에서 재구성하였다. 98℃에서 5분 동안 변성 후 -2℃/초로 85℃ 온도를 낮추고 -0.1℃/초로 25℃까지 온도를 낮춘다. 이후 이형이중가닥의 변성을 막기 위해 4℃로 유지시킨다. 1 ㎕의 T7E1 엔도뉴클레아제를 혼합물에 첨가하고 37℃에서 60분 동안 배양한다.

[0090] 유전자 변형의 분석을 위해, 2% 겔을 사용하여 이미지를 얻고, Alphaimager 2200 이미지 분석 소프트웨어를 분석에 사용하였다. 유전자 변형에 대한 비율은 1- [(1-fraction cleaved) 1/2]식을 이용하여 구했다. 그 결과를 도 3에 나타내었다.

[0091] **실험예 4. 인간 유전자 CCR5 및 DHCR7에 대한 SpyCas9, SpyCas9-RecJ 및 SpyCas9-T5의 HDR(Homology Dependent Repair) 효율 분석**

[0092] HEK293T 세포를 10% 태아 소혈청인 FBS와 P/S가 함유된 DMEM 배지에서 37℃, 5% CO_2 배양기에서 유지시켰다.

HEK293T 세포를 $3x10^6$ 밀도로 10 cm 배양접시에 배양하였다. 세포의 증식 정도는 70% 이하로 유지하였다. 이후 형질전환 16시간 전 200 ng/ml의 노코다졸(Sigma)을 세포에 처리하였고, 동기화된 세포를 유전자 도입 실험에 사용하였다.

[0093] 외부 유전자인 ssODN(single-stranded oligodeoxynucleotides)을 도입시키기 위해 $2x10^5$의 HEK293T 세포를 형질전환 실험에 이용하였다. 각각 33 pmole의 SpCas9, SpCas9-RecJ, SpCas9-T5 단백질과 66 pmole의 CCR5 sgRNA를 이용하였다. 도입 유전자로는 ssODN 20 pmole을 사용하였다. 형질전환을 위해 Lonza사의 4D-Nucleofector 기기를 이용한 전기천공법을 사용하였다. 24시간 후 invitrogen사의 PureLink™ Genomic DNA Mini Kit를 사용하여 genomic DNA를 추출하였다.

[0094] 편집된 부분을 가지고 있는 gDNA와 그렇지 않은 gDNA 샘플을 이용하여 표적 부위를 증폭하였다. 이후 일루미나의 프로토콜대로 PCR과 정제 과정을 거친 후 MiniSeq 장비를 이용하여 표적 부위에 대한 딥 시퀀싱(deep sequencing) 분석을 진행하였다. HDR 분석 결과는 도 4에 나타내었다. 이때, 분석에 이용된 프라이서 서열은 표 3에 기재된 바와 같다.

표 3

[0095]

Proteins	Genes	Adapter primer sequences (5' -3')
SpyCas9	CCR5	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGAGGGCAACTAAATACATTCT (서열번호 81)
		GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAACACCAGTGAGTAGAGCGG (서열번호 82)
	DHCR7	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTTTGAGCAACAGTTCTCC (서열번호 83)
		GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGACGATGTCCACCACAG (서열번호 84)

FnCpf1	CCR5	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTATTTCTGTTCAGATCAC (서열번호 85)
		GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCCCATCAATTATAGAAAGCC (서열번호 86)
	DNMT1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGCACACAGCAGGCCTTTG (서열번호 87)
		GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCCAATAAGTGGCAGAGTGC (서열번호 88)

[0096] **실험예 5. SpCas9 및 SpCas9-RecJ에 의한 오프-타겟 효과 확인**

[0097] SpCas9과 SpCas9-RecJ 단백질을 이용하여 인간 CCR5 유전자를 표적으로 했을 때 표적 부분 외 다른 부분에 대한 영향을 알아보기 위해 스페이서(spacer) 서열 부분이 유사한 다음의 4가지 유전자를 선택하였고 서열은 하기 표 4와 같다.

표 4

[0098]

Target	Gene	Spacer sequence
On	CCR5	TGACATCAATTATTATACAT (서열번호 89)
off#1	ADCY5	TGACATCAATTATTATAgAT (서열번호 90)
off#2	KCNJ6	TGACATCAcTTATTATgCAT (서열번호 91)
off#3	CNTNAP2	TGACATaAATTATTcTACAT (서열번호 92)
off#4	Chr. 5 N/A	TGAaATCAATTATcATAgAT (서열번호 93)

[0099] Mock, SpCas9, SpCas9-RecJ를 처리한 gDNA로부터 표적 외 부분 ADCY5, KCNJ6, CNTNAP2, Chr.5 N/A에 대한 각각의 프라이머를 이용하여 증폭시켰다. 이후 정제 과정과 제조사의 지시에 따라 시퀀싱 라이브러리를 제조하였고 일루미나 MiniSeq 장비를 이용하여 표적 부위에 대한 딥 시퀀싱 분석을 진행하였다. 그 결과를 도 5 및 도 6에 나타내었다. 이때, 분석에 이용된 프라이서 서열은 표 5에 기재된 바와 같다.

표 5

[0100]

Gene	Target specific primer sequences	Adapter primer sequences
ADCY5	GTCCCATGACAGGCGTGTAT	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAAGAGTGTAGAGGAGGACAG
	GCTCCCACCTTAGTGCTCTG	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAAAAGGAGATGCTTGGCAC
KCNJ6	TGGAGCCATTGGTTTGCATC	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGACCAAAAAACCTCATGGAT
	GCCTGGCCAAGTTTCAGTTA	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTTGAGGAAGACAAGTCATGT
CNTNAP2	TGCAGTGCAGACTCTTTCCA	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGCAGACTCTTTCCATACTGT
	AAGGACACAGGGCAACTGAA	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAAATCAAATTTCATGTGTCCA
Chr. 5 N/A	TGTGGAACGAGTGGTGACAG	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCAGAAGTTCCTGAAATTATCC
	GACCAAACCACATTCTTCTCAC	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACATGAGAGATAGTGATATGT

[0101] **실시예 6. Cas9 및 Cas9-RecJ의 PD-1 및 CCR5에 대한 온-타겟 효과 확인**

[0102] 인간 PD-1 및 CCR5의 동일한 유전자 자리에 대한 Cas9 및 Cas9-RecJ 아포단백질(apoproteins)을 지질 매개 형질감염(lipid mediated transfection) 방법으로 HEK293 세포에 형질감염 시켰다. 게놈 DNA 추출을 위해 0시간 내지 48시간의 다른 시점에서 세포를 수득하였다. 이때, 0 시간은 대조군이다. CCR5를 확인하기 위한 544 bp 밴드 및 PD-1을 확인하기 위한 524 bp 밴드를 생성하기 위하여 PCR을 수행하였다. T7E1 엔도뉴클레아제 어세이를 수행할 경우, CCR5에서는 274 bp 및 271 bp의 작은 단편이 나타나고, PD-1에서는 319 bp 및 205 bp의 작은 단편이 나타났다. 이러한 것을 통하여 삽입-결실율(indel)을 측정할 수 있었다. 이에 대한 결과를 도 7에 나타내었다.

[0103] **실험예 7. Cas9 및 Cas9-RecJ에 의한 오프-타겟 효과 확인**

[0104] 인간 CCR5 sgRNA와 Cas9 및 Cas9-RecJ를 이용시 5개의 오프 타겟팅된 유전자를 나타낸 것이다. ADCY5 641bps; KCNJ6, 566 bps; CNTPNA2, 300 bps; Chr.5, 605bps. 오프-타겟 효과를 도 8에 나타내었다.

[0105] 또한, 5개의 오프-타겟된 유전자의 In vitro 절단 에세이를 분석하였다. 예상된 절단된 DNA의 크기는 ADCY5에서 377 bp 및 264 bp이고; KCNJ6에서 352 bp 및 214 bp이며; CNTPNA2에서 183 bp 및 117 bp이며; Ch. 5에서 355 bp 및 250 bp이다. 절단은 오프 타겟팅 효과를 확인하기에는 명확하지 않았다. In vitro 에세이를 근거로, SpCas9-RecJ의 오프 타겟팅 정도는 SpCas9과 차이가 없었다. 이에 대한 결과를 도 9에 나타내었다.

[0106] **실험예 8. 변형된 FnCpf1 단백질의 절단 효율 확인**

[0107] 단백질과 gRNA는 몰농도 비율이 1:1.2가 되도록 준비하였다. 반응 부피 20 μL로 하여 단백질 50 nM에 crRNA 60 nM을 섞은 후 20분 상온에 두고 RNP가 형성되도록 하였다. crRNA는 CCR5 유전자를 타겟으로하며, 타겟 유전자인 CCR5 주형의 크기는 약 1.4 kb이다.

[0108] 1.5 kb 크기의 타겟 유전자인 CCR5 주형 DNA 500 ng을 RNP 복합체와 섞은 후 37℃에서 1시간 반응시킨다. 2% agarose gel에 전기영동한 후 겔 이미지를 수득하였다. 이를 이용하여 주형 DNA와 RNP 복합체가 작용하고 남은 주형 DNA의 비율을 계산하여, 단백질의 뉴클레아제 활성 효율을 얻었다.

[0109] FnCpf1, FnCpf1-RecJ, 또는 SpCas9에 의한 선형 dsDNA가 절단됨을 확인하기 위한 실험을 수행하였다. 선형 dsDNA는 PCR을 통해 증폭되고 정제되었다. Cas9 sgRNA 및 Cpf1 sgRNA 사이트는 서로 근접한 곳에 위치한다. SpCas9 및 FnCpf1 단백질에 사용된 sgRNA는 각각 5'-NGG-3' PAM(protospacer adjacent motif)이 상단(upstream)에는 존재하며, 5'-TTTN-3' PAM이 하단(downstream)에 존재한다. 절단되기 전의 온전한 dsDNA는 1,200 bp 위치에서 단일 밴드로 관찰되었다. 상기 dsDNA는 절단된 후 두개 모두 600 bp 밴드에 위치하였다. 그러나, FnCpf1-RecJ를 이용할 경우, RecJ의 엑소뉴클레아제 활성에 의해 상기 절단된 밴드가 사라짐을 확인하였다(도 10).

[0110] FnCpf1 또는 FnCpf1-RecJ 엑소뉴클레아제에 의한 선형 dsDNA In vitro 절단 에세이를 수행하였다. 선형 dsDNA는 FnCpf1 또는 FnCpf1-RecJ에 의해 절단되었다. 선형 dsDNA는 플라스미드로부터 PCR을 이용하여 증폭되고 분리되었다. sgRNA는 FnCpf1 단백질을 타겟 사이트로 위치시키기 위해 사용되었으며, 5'-TTTN-3' PAM이 하단(downstream)에 존재한다. 절단되기 전의 온전한 dsDNA는 1,200 bp 위치에서 단일 밴드로 관찰되었다. 상기 dsDNA는 절단되면 두개 모두 600 bp 밴드에 위치하였다. 그러나, FnCpf1-RecJ를 이용할 경우, RecJ의 엑소뉴클레아제 활성에 의해 상기 절단된 밴드가 사라졌다. 반응 시간을 증가시킴에 따라, FnCpf1 및 FnCpf1-RecJ 반응 혼합물에서 원래의 밴드가 흐려짐을 확인하였다(도 11).

[0111] **실험예 9. 변형된 Cas9 단백질의 편집 효율 확인**

[0112] HEK293T 세포를 10% 태아 소혈청인 FBS와 P/S가 함유된 DMEM 배지에서 37℃, 5% CO_2 배양기에서 유지시켰다. 녹인 실험의 효율을 높이기 위해 HEK293T 세포를 $3x10^6$ 밀도로 10 cm 배양접시에 배양한 후 형질전환 16시간 전 200 ng/ml의 노코다졸(Sigma)을 세포에 처리하였고, 동기화된 세포를 유전자 도입 실험에 사용하였다. 사용한 외부 유전자인 ssODN(single-stranded oligodeoxynucleotides)의 서열은 다음과 같다:

[0113] CCR5: 5' TGGAACAAGATGGATTATCAAGTGTCAAGTCCAATCTATGACATCAATTATTAT

[0114] ACATATGCATCGGAGCCCTGCCAAAAAATCAATGTGAAGCAAATCGCAGCCC

[0115] DHCR7: 5' ATGGGCCCCAGTGTGACTGCCTGCATCCGTCCTCGCAGGGAGGTGGACTGGT

[0116] TTTGAATTCCACTGGCGAGCGTCATCTTCCTACTGCTGTTCGCCCCCTTCATCG

[0117] 외부 유전자인 ssODN(single-stranded oligodeoxynucleotides)을 도입시키기 위해 $2x10^5$의 HEK293T 세포를 형질전환 실험에 이용하였다. 각각 33 pmole의 단백질과 66 pmole의 sgRNA를 이용하였다. 도입 유전자로는 ssODN 20 pmole을 사용하였다. 형질전환을 위해 Lonza사의 4D-Nucleofector 기기를 이용한 전기천공법을 사용하였다. 24시간 후 invitrogen사의 PureLink™ Genomic DNA Mini Kit를 사용하여 genomic DNA를 추출하였다.

[0118] 편집된 부분을 가지고 있는 gDNA와 그렇지 않은 gDNA 샘플을 이용하여 표적 부위를 증폭하였다. 이후 일루미나의 프로토콜대로 PCR과 정제 과정을 거친 후 MiniSeq 장비를 이용하여 표적 부위에 대한 딥 시퀀싱(deep sequencing) 분석을 진행하였다. 시퀀싱을 위해 사용된 어댑터 프라이머 서열은 다음과 같다.

[0119] CCR5 (F): 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGAGGGCAACTAAATACATTCT

[0120] CCR5 (R): 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAACACCAGTGAGTAGAGCGG

[0121] DHCR7 (F): 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTTTGAGCAACAGTTCTCC

[0122] DHCR7 (R): 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGACGATGTCCACCACAG

[0123] SpyCas9과 SpyCas9의 C 말단에 다양한 단백질이 결합된 변형된 SpyCas9의 CCR5 및 DHCR7에 대한 유전자 편집 효율을 도 15 및 도 16에 나타내었다. SpyCas9의 C 말단에 GFP, hTdT, 및 엑소뉴클레아제로 알려진 RecJ, RecE, lambda, mungbean, and T5가 결합된 형태이다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율이 서로 상이하였다.

[0124] 사용한 변형된 Cas9-RecJ의 구조 및 이의 Knock-out 및 Knock-in 효율을 도 17 및 도 18에 나타내었다. 도 17A는 SpyCas9-RecJ의 구조를 나타낸 것이다. 도 17B는 대조군, SpyCas9 및 SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. 도 17C는 CCR5 및 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0125] 도 18A는 SpyCas9-RecJ의 구조를 나타낸 것이다. 도 18B는 대조군, SpyCas9 및 SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. 도 18C는 CCR5 및 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0126] 또한, SpyCas9과 SpyCas9의 N 말단 또는 N/C 양말단에 다양한 단백질이 결합된 변형된 SpyCas9의 DHCR7에 대한 유전자 편집 효율을 도 19는 및 20에 나타내었다. 변형된 SpyCas9은 SpyCas9의 N 말단에 RecJ, RecE, GFP, 단일 가닥 DNA 결합 단백질(SSB) 및 이중 가락 DNA 결합 단백질(DSB)이 결합되어 있다. 또한, 변형된 SpyCas9은 SpyCas9의 C 말단에 RecJ가 결합되어 있고, N 말단에 SSB 또는 DSB이 결합되어 있다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율이 서로 상이하였다.

[0127] 도 21A는 RecJ-SpyCas9의 구조를 나타낸 것이다. 도 21B는 대조군, SpyCas9 및 RecJ-SpyCas9 간의 유전자 편집 효율을 나타낸 것이다. 도 21C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 RecJ-SpyCas9의 편집 효율 변화를 백분율로 나타낸 것이다.

[0128] 도 22A는 RecJ-SpyCas9의 구조를 나타낸 것이다. 도 22B는 대조군, SpyCas9 및 RecJ-SpyCas9 간의 유전자 편집 효율을 나타낸 것이다. 도 22C는 CCR5 및 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 RecJ-SpyCas9의 편집 효율 변화를 백분율로 나타낸 것이다.

[0129] 도 23A는 SSB-SpyCas9-RecJ의 구조를 나타낸 것이다. 도 23B는 대조군, SpyCas9 및 SSB-SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. 도 23C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 SSB-SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0130] 도 24A는 DSB-SpyCas9-RecJ의 구조를 나타낸 것이다. 도 24B는 대조군, SpyCas9 및 DSB-SpyCas9-RecJ 간의 유전자 편집 효율을 나타낸 것이다. 도 24C는 DHCR7 유전자의 편집 정도를 확인하여, SpyCas9 대비 DSB-SpyCas9-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0131] **실험예 10. 변형된 Cpf1 단백질의 편집 효율 확인**

[0132] HEK293T 세포를 10% 태아 소혈청인 FBS와 P/S가 함유된 DMEM 배지에서 37℃, 5% CO_2 배양기에서 유지시켰다. 녹인 실험의 효율을 높이기 위해 HEK293T 세포를 $3x10^6$ 밀도로 10 cm 배양접시에 배양한 후 형질전환 16시간 전 200 ng/ml의 노코다졸(Sigma)을 세포에 처리하였고, 동기화된 세포를 유전자 도입 실험에 사용하였다. 사용한 외부 유전자인 ssODN(single-stranded oligodeoxynucleotides)의 서열은 다음과 같다.

[0133] CCR5: 5' AATTCTCTGAGGCTTTCTTTTAAATATACATAAGGAACTTTCGGAGTGAAGGG

[0134] AGAGTTTCATATGGTCAATAACTTGATGCATGTGAAGGGGAGATAAAAAGGTT

[0135] DNMT1: 5' TGGCCCTGGGGCCGTTTCCCTCACTCCTGCTCGGTGAATTTGGCTCAGCAGGC

[0136] ACCTGCCGAATTCTCAGCTGCTCACTTGAGCCTCTGGGTCTAGAACCCTCTGG

[0137] 외부 유전자인 ssODN(single-stranded oligodeoxynucleotides)을 도입시키기 위해 $2x10^5$의 HEK293T 세포를 형

질전환 실험에 이용하였다. 각각 33 pmole의 단백질과 66 pmole의 sgRNA를 이용하였다. 도입 유전자로는 ssODN 20 pmole을 사용하였다. 형질전환을 위해 Lonza사의 4D-Nucleofector 기기를 이용한 전기천공법을 사용하였다. 24시간 후 invitrogen사의 PureLink™ Genomic DNA Mini Kit를 사용하여 genomic DNA를 추출하였다.

[0138] 편집된 부분을 가지고 있는 gDNA와 그렇지 않은 gDNA 샘플을 이용하여 표적 부위를 증폭하였다. 이후 일루미나의 프로토콜대로 PCR과 정제 과정을 거친 후 MiniSeq 장비를 이용하여 표적 부위에 대한 딥 시퀀싱(deep sequencing) 분석을 진행하였다. 시퀀싱을 위해 사용된 어댑터 프라이머 서열은 다음과 같다.

[0139] CCR5 (F): 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTATTTCTGTTCAGATCAC

[0140] CCR5 (R): 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCCCATCAATTATAGAAAGCC

[0141] DNMT1 (F): 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGCACACAGCAGGCCTTTG

[0142] DNMT1 (R): 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCCAATAAGTGGCAGAGTGC

[0143] FnCpf1과 FnCpf1의 C 말단에 다양한 단백질이 결합된 변형된 FnCpf1의 CCR5 및 DNMT1 유전자에 대한 유전자 편집 효율을 도 25 및 도 26에 나타내었다. FnCpf1의 C 말단에 GFP, hTdT, 및 엑소뉴클레아제로 알려진 RecJ, RecE, lambda, mungbean, 및 T5이 결합된 형태이다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율이 서로 상이하다.

[0144] 도 27A는 FnCpf1-RecJ의 구조를 나타낸 것이다. 도 27B는 대조군, FnCpf1 및 FnCpf1-RecJ 간의 유전자 편집 효율을 나타낸 것이다. 도 27C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0145] 도 28A는 FnCpf1-RecJ의 구조를 나타낸 것이다. 도 28B는 대조군, FnCpf1 및 FnCpf1-RecJ 간의 유전자 편집 효율을 나타낸 것이다. 도 28C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0146] FnCpf1과 FnCpf1의 N 말단 또는 N/C 양말단에 다양한 단백질이 결합된 변형된 FnCpf1의 유전자 편집 효율을 도 29 및 도 30에 나타내었다. 변형된 FnCpf1은 FnCpf1의 N 말단에 RecJ, RecE, GFP, 단일 가닥 DNA 결합 단백질(single stranded DNA binding protein, SSB) 및 이중 가락 DNA 결합 단백질(double stranded DNA binding protein, DSB)이 결합되어 있다. 또한, 변형된 FnCpf1은 FnCpf1의 C 말단에 RecJ가 결합되어 있고, N 말단에 SSB이 결합되어 있다. 도면에서 나타낸 바와 같이 Knock-out 및 Knock-in 효율이 서로 상이하였다.

[0147] 도 31A는 RecJ-FnCpf1의 구조를 나타낸 것이다. 도 31B는 대조군, RecJ-FnCpf1 간의 유전자 편집 효율을 나타낸 것이다. 도 31C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 RecJ-FnCpf1의 편집 효율 변화를 백분율로 나타낸 것이다.

[0148] 도 32A는 RecJ-FnCpf1의 구조를 나타낸 것이다. 도 32B는 대조군, RecJ-FnCpf1 간의 유전자 편집 효율을 나타낸 것이다. 도 32C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 RecJ-FnCpf1의 편집 효율 변화를 백분율로 나타낸 것이다.

[0149] 도 33A는 SSB-FnCpf1-RecJ의 구조를 나타낸 것이다. 도 33B는 대조군, FnCpf1 및 SSB-FnCpf1-RecJ 간의 Knock-out 유전자 편집 효율을 나타낸 것이다. 도 33C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 SSB-FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0150] 도 34A는 SSB-FnCpf1-RecJ의 구조를 나타낸 것이다. 도 34B는 대조군, FnCpf1 및 SSB-FnCpf1-RecJ 간의 Knock-in 유전자 편집 효율을 나타낸 것이다. 도 34C는 CCR5 및 DNMT1 유전자의 편집 정도를 확인하여, FnCpf1 대비 SSB-FnCpf1-RecJ의 편집 효율 변화를 백분율로 나타낸 것이다.

[0151] **실험예 11. 변형된 크리스퍼 연관 단백질의 편집 효율에 미치는 조건 확인**

[0152] ssODN 및 노코다졸(Nocodazole)의 존재 유무에 따른 변형된 Cas9 단백질의 활성 변화를 확인하기 위한 실험을 수행하였다.

[0153] 크리스퍼 플러스 매개 녹아웃을 확인하기 위하여, 노코다졸이 처리되지 않은 세포를 사용하였다. 녹인의 효율을 증가시키기 위하여, HEK293T 세포를 10-cm 배양 접시에 3 x 10^6 세포 밀도로 접종하였다. 상기 세포를 노코다졸(200 ng/ml)로 16시간 배양한 후, 전기 영동을 수행하였다. 하기 표 6에 기재한 ssODN 주형을 녹인을 위해 사용

하였다:

[0154] 크리스퍼 플러스 단백질을 세포에 트랜스펙션하기 전에, 분리정제된 SpyCas9 또는 FnCpf1 변이체 단백질(33 pmol) 및 sgRNA 또는 crRNA(66 pmol)을 상온에서 20분간 인큐베이션시켜 RNP를 복합체를 형성하였다. HDR 주형(20 pmol ssODN 또는 2 pmol dsDNA)을 RNP 복합체에 첨가하였다. Lonza를 이용하여 HEK293T 세포에 뉴클레오펙션(Nucleofection)을 수행하였다. 그 후, 각각의 뉴클레오펙션은 20 μl의 뉴클레오펙션 시약에 있는 2 x 10^5 세포와 RNP:DNA 10 μl을 혼합였다.

표 6

[0155]

Proteins	Genes	ssODN donor sequences (5' -3')
SpyCas9	CCR5	TGGAACAAGATGGATTATCAAGTGTCAAGTCCAATCTATGACATCAATTATTAT (서열번호 94) ACATATGCATCGGAGCCCTGCCAAAAAATCAATGTGAAGCAAATCGCAGCCC (서열번호 95)
	DHCR7	ATGGGCCCCAGTGTGACTGCCTGCATCCGTCCTCGCAGGGAGGTGGACTGGT (서열번호 96) TTTGAATTCCACTGGCGAGCGTCATCTTCCTACTGCTGTTCGCCCCCTTCATCG (서열번호 97)
FnCpf1	CCR5	AATTCTCTGAGGCTTTCTTTTAAATATACATAAGGAACTTTCGGAGTGAAGGG (서열번호 98) AGAGTTTCATATGGTCAATAACTTGATGCATGTGAAGGGGAGATAAAAAGGTT (서열번호 99)
	DNMT1	TGGCCCTGGGGCCGTTTCCCTCACTCCTGCTCGGTGAATTTGGCTCAGCAGGC (서열번호 100) ACCTGCCGAATTCTCAGCTGCTCACTTGAGCCTCTGGGTCTAGAACCCTCTGG (서열번호 101)

[0156] **실험예 11.1. 변형된 Cas9의 조건 확인**

[0157] 도 35a 및 도 36a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 35b 및 도 36b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 35c 및 도 36c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, SpyCas9 및 변형된 SpyCas9의 효율에는 큰 영향이 없음을 확인하였다.

[0158] **실험예 11.2. 변형된 Cpf1의 조건 확인**

[0159] 크리스퍼 플러스 매개 녹아웃을 확인하기 위하여, 노코다졸이 처리되지 않은 세포를 사용하였다. 녹인의 효율을 증가시키기 위하여, HEK293T 세포를 10-cm 배양 접시에 3 x 10^6 세포 밀도로 접종하였다. 상기 세포를 노코다졸(200 ng/ml)로 16시간 배양한 후, 전기 영동을 수행하였다. 상기 표 6에 기재한 ssODN 주형을 녹인을 위해 사용하였다:

[0160] 크리스퍼 플러스 단백질을 세포에 트랜스펙션하기 전에, 분리정제된 SpyCas9 또는 FnCpf1 변이체 단백질(33 pmol) 및 sgRNA 또는 crRNA(66 pmol)을 상온에서 20분간 인큐베이션시켜 RNP를 복합체를 형성하였다. HDR 주형(20 pmol ssODN 또는 2 pmol dsDNA)을 RNP 복합체에 첨가하였다. Lonza를 이용하여 HEK293T 세포에 뉴클레오펙션(Nucleofection)을 수행하였다. 그 후, 각각의 뉴클레오펙션은 20 μl의 뉴클레오펙션 시약에 있는 2 x 10^5 세포와 RNP:DNA 10 μl을 혼합하였다.

[0161] FnCpf1 및 FnCpf1의 C 말단에 단백질이 결합된 변형된 FnCpf1에 유도되는 비상동적인 말단 연결에 있어서, ssODN 및 노코다졸(nocodazole)이 미치는 영향을 확인한 것을 도 37에 나타내었다. 도 37a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 37b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 37c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, FnCpf1 및 변형된 FnCpf1의 효율을 향상 시킨다는 점을 확인하였다.

[0162] FnCpf1 및 FnCpf1의 N 말단 또는 N/C 양말단에 단백질이 결합된 변형된 FnCpf1에 유도되는 비상동적인 말단 연결에 있어서, ssODN 및 노코다졸이 미치는 영향을 도 38에 나타내었다. 도 38a는 ssODN 및 노코다졸이 처리되지 않았을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 38b는 ssODN 및 노코다졸이 처리되었을 때의 비상동적인 말단 연결 효율을 나타낸다. 도 38c는 ssODN 및 노코다졸의 존재 유무에 따른 효율을 비교한 것이다. ssODN 및 노코다졸은 Knock-out의 경우, FnCpf1 및 변형된 FnCpf1의 효율에는 큰 영향이 없음을 확인하였다.

[0163] **실험예 12. 크리스퍼 플러스 단백질의 유전자 편집 효능 검증**

[0164] **제조예 1. 식물 유전자 편집을 위한 용액 조성물 제조**

[0165] 종자 파종 배지의 조성은 다음과 같다.

표 7

[0166]

종자 파종 배지	1 L
1/2 MS Powder	2.2 g
Sucrose	10 g
Plant agar	7 g
pH	5.7 (1N NaOH 또는 1N HCl 로 측정함)

[0167] Enzyme solution의 조성은 다음과 같다. 단, 사용 직전에 만들어 사용함을 원칙으로 한다.

표 8

[0168]

Enzyme Solution	20 mL
1.0% Cellulase R10	200 mg
0.5% Macerozyme R10	100 mg
0.4 M Mannitol	10 mL (0.8 M Mannitol Stock solution)
20 mM MES, pH 5.7	4 mL (100 mM MES Stock solution, pH 5.7)
20 mM KCl	200 μL (2 M KCl Stock solution)
위 시약들 조합한 후 60도에서 10분간 배양한 후 아래 시약을 조합한다.	
10 mM $CaCl_2 \cdot 2H_2O$	200 μL (1 M $CaCl_2 \cdot 2H_2O$ Stock solution)
0.1 % BSA	200 μL (10% BSA Stock solution)

[0169] MMG solution의 조성은 다음과 같다. 단, 사용 직전에 만들어 사용함을 원칙으로 한다.

표 9

[0170]

MMG Solution	10 mL
0.4 M Mannitol	5 mL (0.8 M Mannitol Stock solution)
4 mM MES, pH 5.7	400 μL (0.1 M MES Stock solution, pH5.7)
15 mM $MgCl_2$	150 μL (1 M $MgCl_2$ Stock solution)
Nuclease-free water	4.45 mL

[0171] PEG4000 solution의 조성은 다음과 같다. 단, 사용 직전에 만드는 것을 원칙으로 한다.

표 10

[0172]

PEG Solution	5 mL
0.2 M Mannitol	1.25 mL (0.8 M Mannitol Stock solution)
40%W/V PEG-4000	2 g (Polyethylene glycol 4000)
100 mM $CaCl_2 \cdot 2H_2O$	500 μL (1 M $CaCl_2$*?*$2H_2O$ Stock solution)
Nuclease-free water	1.5 mL

[0173] W5 solution의 조성은 다음과 같다.

표 11

[0174]

W5 Solution	50 mL
154 mM NaCl	3.85 mL (2 M NaCl Stock solution)
125 mM $CaCl_2 \cdot 2H_2O$	6.25 mL (1 M $CaCl_2 \cdot 2H_2O$ Stock solution)
5 mM KCl	125 μL (2 M KCl Stock solution)
2 mM MES, pH 5.7	500 μL (0.1 M MES Stock solution)
Nuclease-free water	39.275 mL

[0175] 상기 Enzyme Solution, MMG Solution, PEG Solution, W5 Solution은 모두 0.45 μm Syringe-filter로 필터링하였다..

[0176] **실험예 12.1. 실험 식물체의 준비**

[0177] 도 51에 나타낸 바와 같이 실험을 수행하였다. 구체적으로, 담배 종자 파종 및 원형질체 분리를 다음과 같이 수행하였다. 먼저, 담배 종자를 50% Clorox로 1분간 처리하였다. 그 후, Clean bench로 옮겨 멸균수로 5번 씻어, 필터페이퍼에서 20분간 건조하였다. 그 후, 멸균된 종자를 종자파종 배지에 올려 일주일간 배양한 후 마젠타박스에 옮겨 3주동안 키웠다. 배양 빛조건은 광배양 16시간, 암배양 8시간이며, 온도는 25℃ 내지 28℃에서 키웠다. 이때, 식물은 4~6주 자란 잎을 사용해야하며, 8주 이상 자란 잎은 사용하지 않음을 원칙으로 한다.

[0178] 4~6주 자란 후, 담배 어린잎을 유리판에 올려 잎병과 잎끝부분을 자른 후 잎 안쪽부분만을 잘랐다. 0.5 mm 이하로 자르고, 자른 잎은 Enzyme solution 10 mL에 올려 상온에서 암 상태로 3~4시간 orbital shaker(50 rpm) 위에 올려서 배양하였다(도 52).

[0179] 배양 후 W5 solution 10mL을 첨가한 후 조심스럽게 잘 섞어주었다. 그 후, Cell strainer(70 μm)를 이용해서 enzyme solution 내에 존재하는 원형질체를 걸러주었다. 그 후, 걸러진 원형질체를 100 g으로 6분 원심분리하였다. 상층액을 버리고 MMG solution을 첨가해서 조심스럽게 원형질체 pellet을 풀어 주었다. 그 후, Ice에서 10 ~ 30분 정도 방치하였다. 이 중 일부를 hemocytometer 이용해 현미경으로 원형질체의 수를 계수하였다. MMG solution을 더 첨가하여 원형질체의 농도가 2 X 10^6 cells/mL이 되도록 희석하였다.

[0180] **실험예 12.2. 크리스퍼 플러스 단백질을 이용한 식물체 유전자 편집**

[0181] 2 mL e-tube에 sgRNA와 Cas9 protein 그리고 NEB buffer 3.1을 최종 volume 20 μL 가 되게 한 후 상온에서10분간 반응시켰다. Folding을 방지함으로써 DNA binding 효율을 높여 주기 위하여, gRNA는 사용 전에 95℃에서 3분, ice에서 3분 놓아준다.

표 12

[0182]

SgRNA	2.5ug	5ug	10ug
Cas9	10ug	20ug	30ug
10x NEB buffer (3.1)	2 μL	2 μL	2 μL
Nuclease-free water	Up to 20 μL		

[0183] E-tube (2 mL)에 원형질체 200 μL (5 X 10^5 cells)와 PFT3 sgRNA 서열은 CAAGGGCTTCTAAAGCTTGCAAA과 PFT1 sgRNA 서열 GGCTTTTGAGAATTCTAATG 이 사용되었다. PFT3은 담배의 알파 1,3-푸코스트랜스퍼래이즈의 3번 도메인을 목표하고 PFT1은 담배의 알파 1,3-푸코스트랜스퍼래이즈의 4번 도메인을 목표하였다. Cas9 protein (volume 20 μL)을 넣고 잘 섞은 후, Clean bench에서 10분간 배양하였다. 앞에서 배양한 volume과 동일한 PEG solution 220 μL를 넣고 조심스럽게 섞어주었다. 상온에서 15분간 배양 후, W5 solution 840 μL을 첨가하여 잘 혼합하였다. 100 g에서 2분간 원심분리한 후, 상등액을 제거하였다. 형질전환한 원형질체에 W5 Solution 1000 μL을 넣어 재현탁하였다. 100g에서 2분간 원심분리한 후, 상등액을 제거하였다. W5 Solution 1000 μL을 넣어 Resuspension한 후, 6 또는 12 well plate에 넣어 30℃ 암배양기에서 48시간 배양 후 수거하였다. 수거된 세포는 DNA를 추출 후 타겟 시퀀스를 PCR 한 후 NGS 분석에 들어간다.

[0184] **실험예 12.3. NGS 분석**

[0185] 분석장비(일루미나제품, MiniSeq (SY-420-1001), Miniseq mid output kit (300 cycles, FC-420-1004))를 이용하여 분석하였다. Amplicon primers 조건은 아래와 같으며, 타겟 specific 시퀀스 앞 뒤로 붙여서 제작하였다.

[0186] 타겟 PFT1에 사용한 시퀀스 정보는 다음과 같다:

[0187] F:5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-ACTTCTCTTGGGCTGAGTATGA

[0188] R:5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-CAGTAAGTTTGGATATTTGAAA

[0189] 타겟 PFT3에 사용한 시퀀스 정보는 다음과 같다:

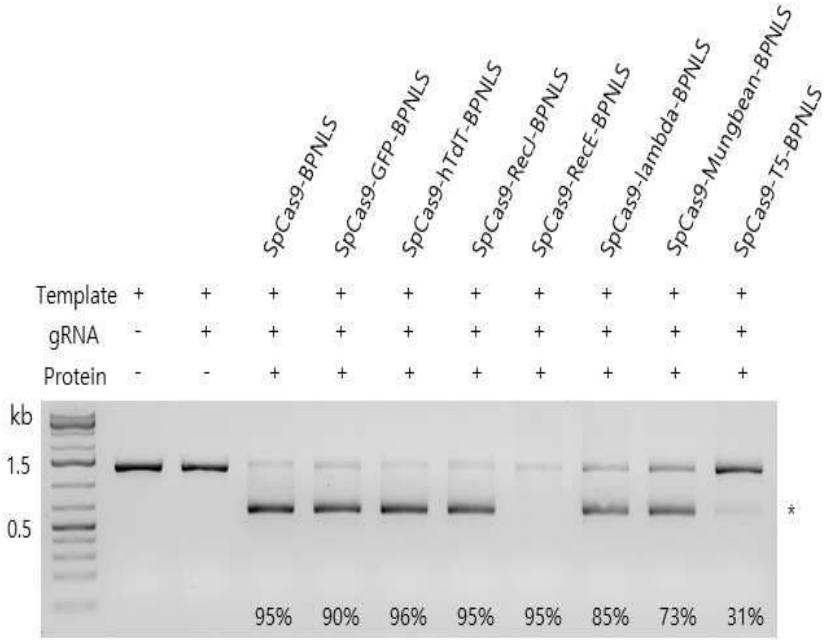
[0190] F:5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-GCATTTGGTGTAGGTTTAGGCT

[0191] R:5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-AATTCTGAAAATCCAAGTCTAT

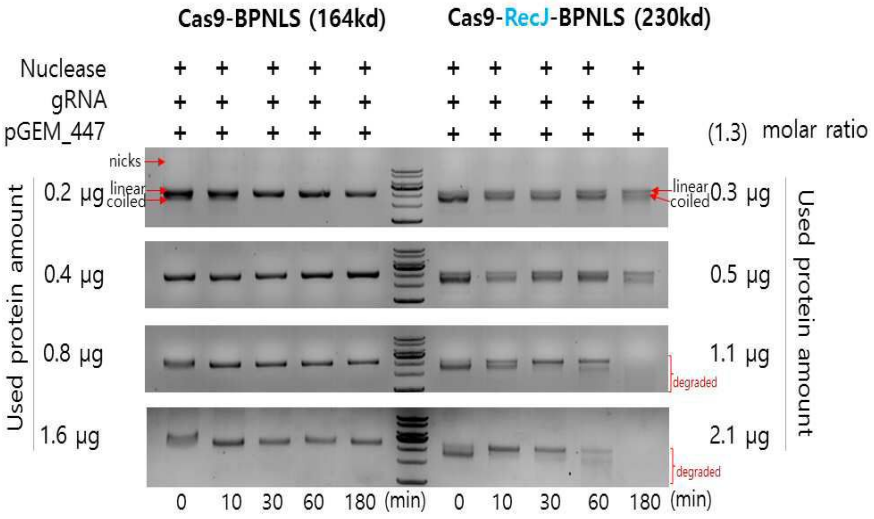
[0192] 타겟 영역을 amplicon primer를 이용해 1차 PCR을 수행하였다. 생산된 PCR product는 반드시 clean-up을 수행하였다. 타겟 영역에 amplicon이 포함된 DNA에 index를 붙여서 library를 만들었다. 수행한 2차 PCR을 Bioanalyzer로 QC를 수행 후 시퀀싱을 수행하였다.

도면

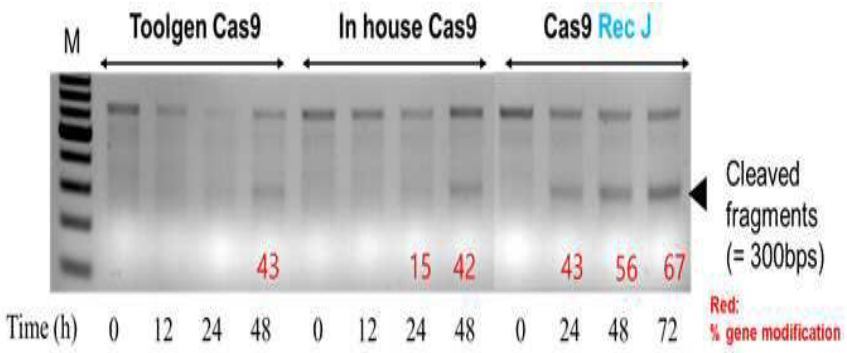
도면1



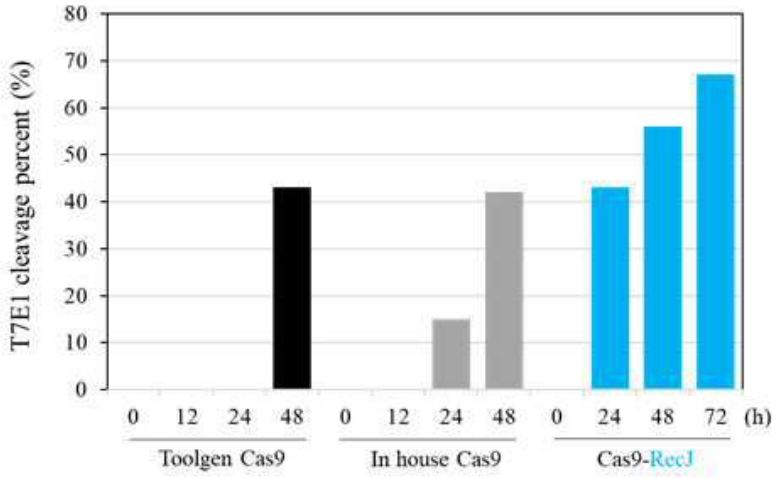
도면2



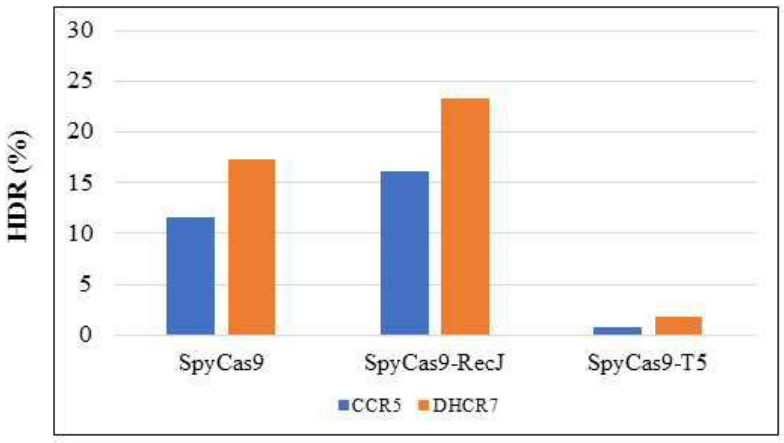
도면3a



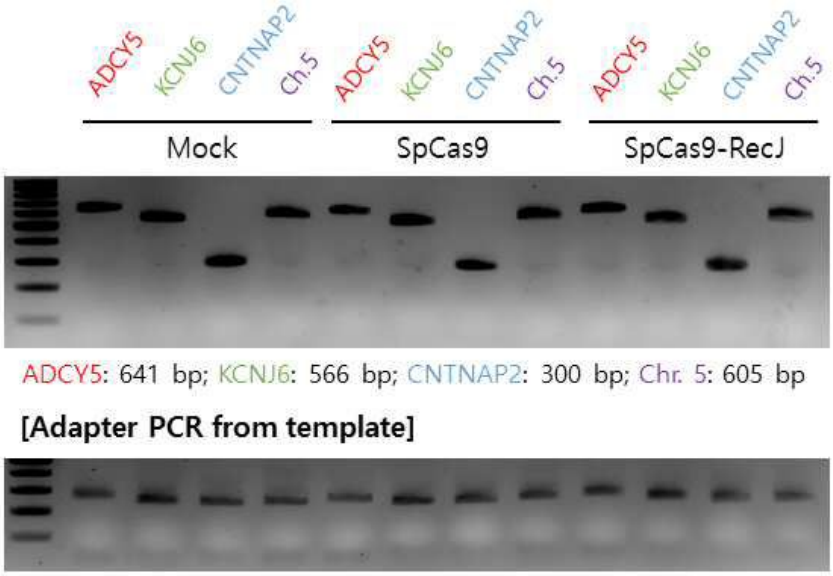
도면3b



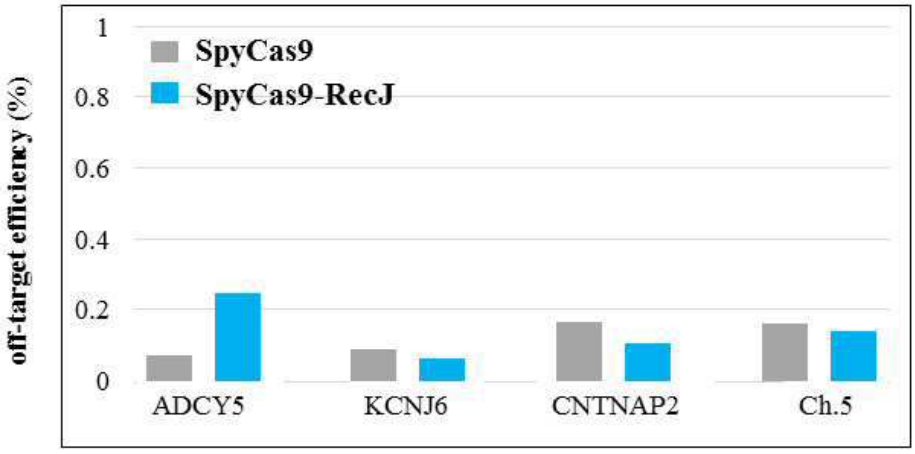
도면4



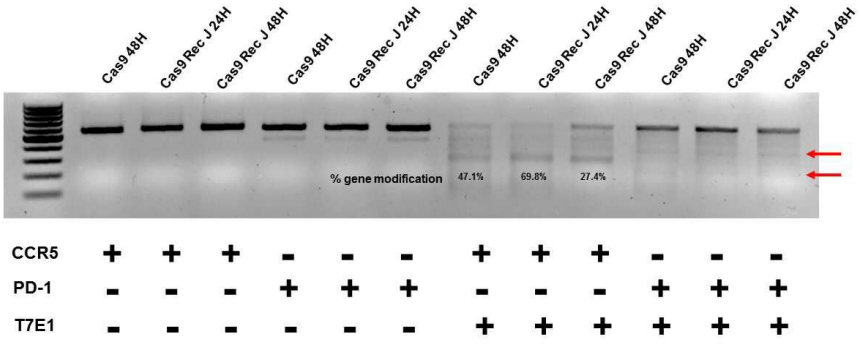
도면5


ADCY5
KCNJ6
CNTNAP2
Ch.5
ADCY5
KCNJ6
CNTNAP2
Ch.5
ADCY5
KCNJ6
CNTNAP2
Ch.5
Mock
SpCas9
SpCas9-RecJ
ADCY5: 641 bp; KCNJ6: 566 bp; CNTNAP2: 300 bp; Chr. 5: 605 bp
[Adapter PCR from template]

도면6

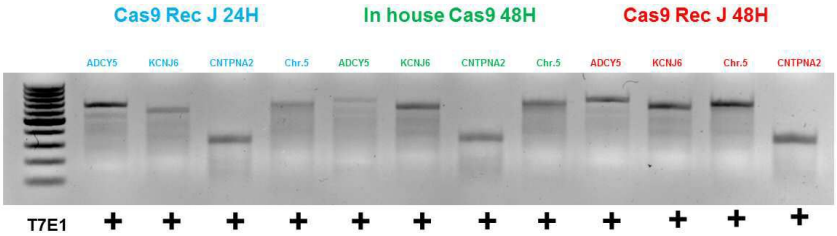

off-target efficiency (%)
1
0.8
0.6
0.4
0.2
0
SpyCas9
SpyCas9-RecJ
ADCY5
KCNJ6
CNTNAP2
Ch.5

도면7

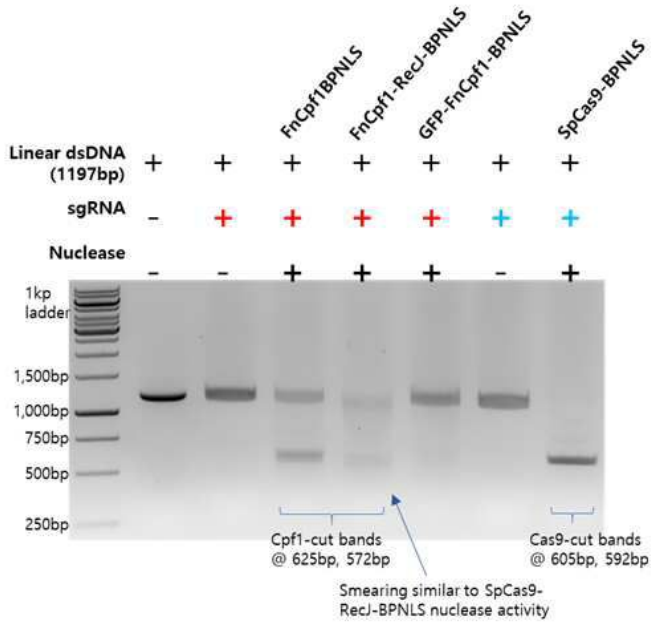

Cas9 48H
Cas9 Rec J 24H
Cas9 Rec J 48H
Cas9 48H
Cas9 Rec J 24H
Cas9 Rec J 48H
Cas9 48H
Cas9 Rec J 24H
Cas9 Rec J 48H
Cas9 48H
Cas9 Rec J 24H
Cas9 Rec J 48H
% gene modification
47.1%
69.8%
27.4%
CCR5 + + + - - - + + + - - -
PD-1 - - - + + + - - - + + +
T7E1 - - - - - - + + + + + +

도면8

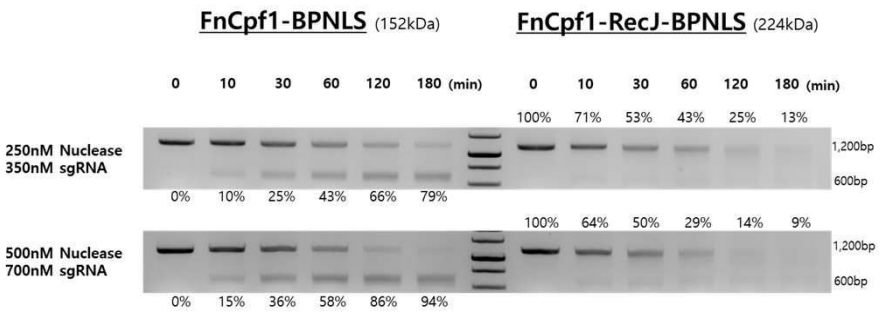
도면9



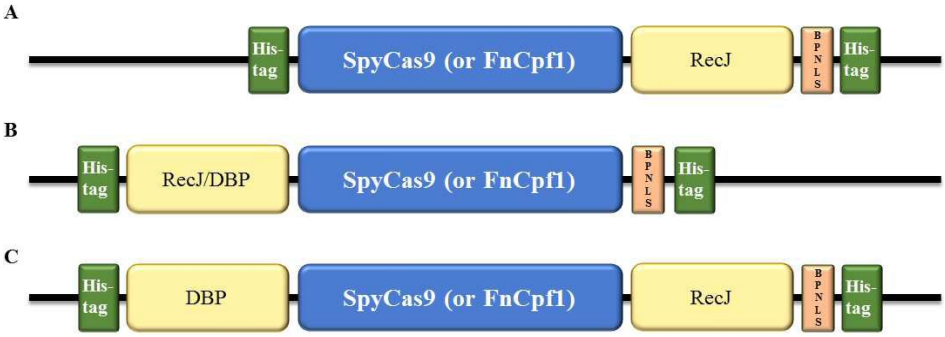
도면10



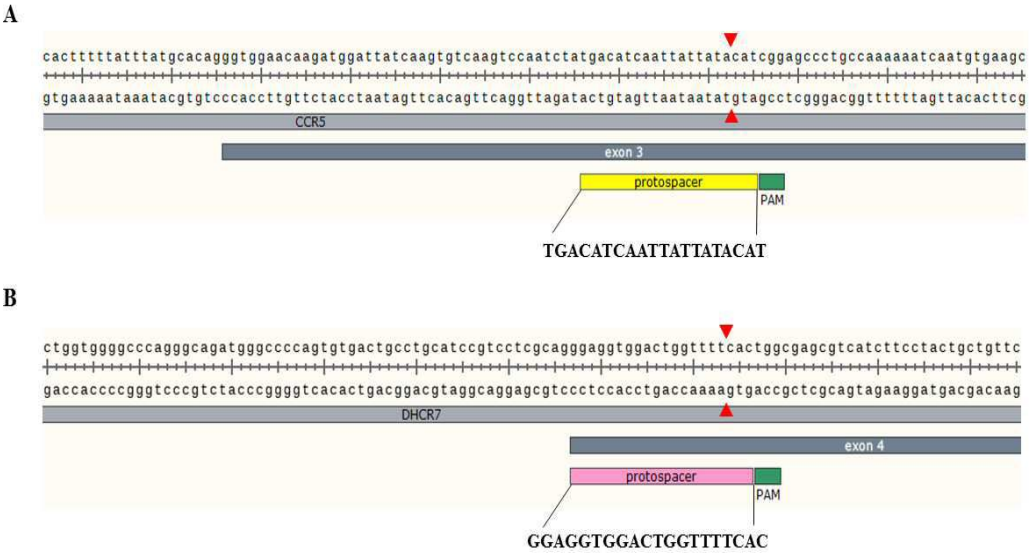
도면11



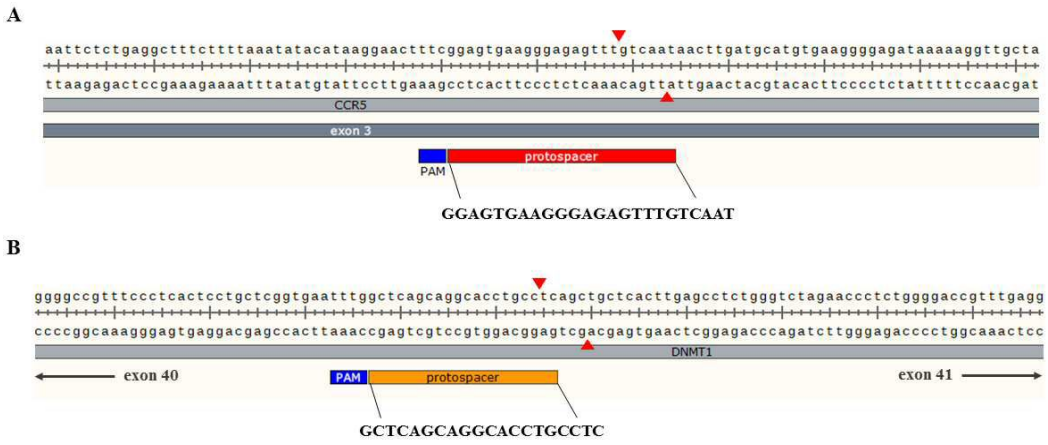
도면12



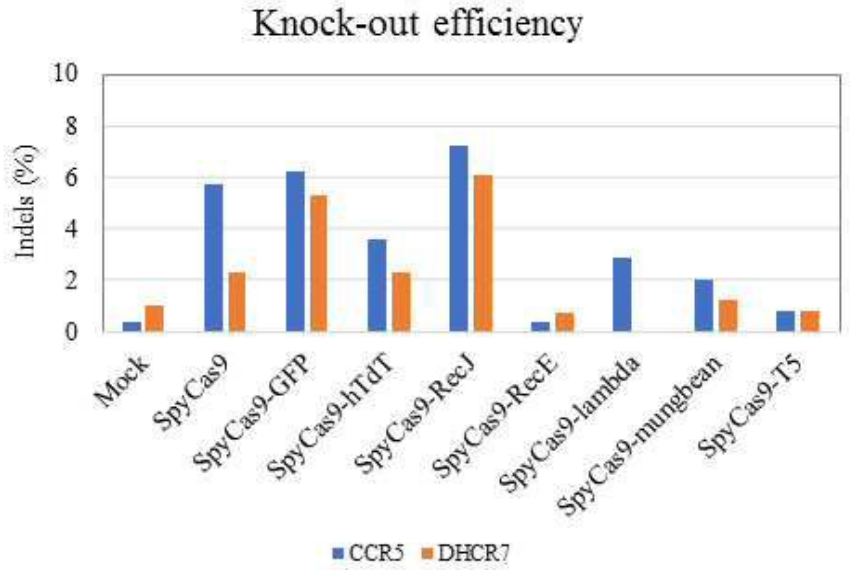
도면13



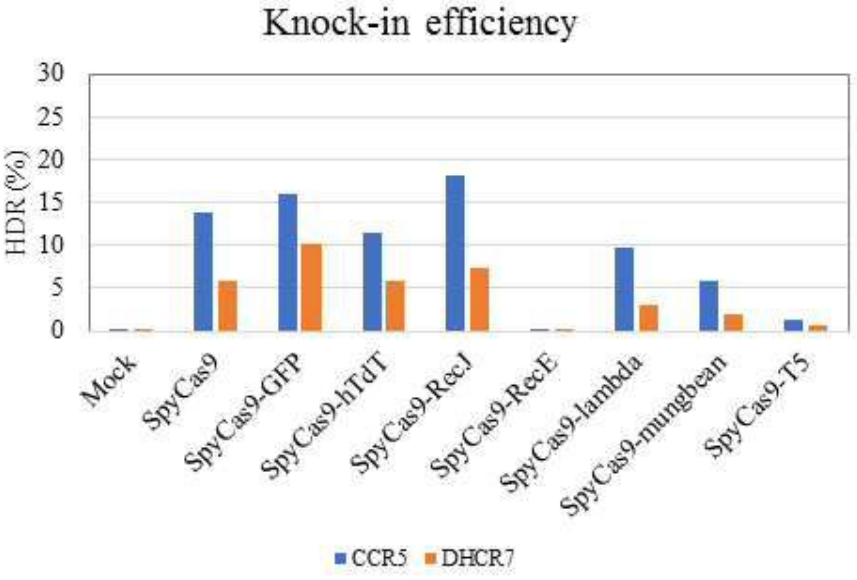
도면14



도면15


Knock-out efficiency
Indels (%)
10
8
6
4
2
0
Mock
SpyCas9
SpyCas9-GFP
SpyCas9-hTdT
SpyCas9-RecJ
SpyCas9-RecE
SpyCas9-lambda
SpyCas9-mungbean
SpyCas9-T5
CCR5
DHCR7

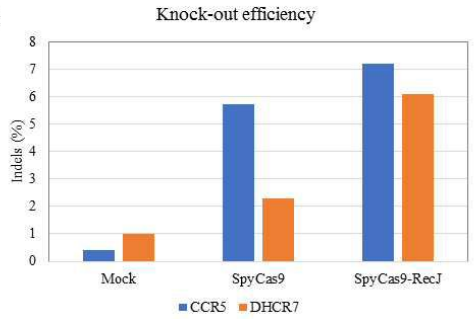
도면16


Knock-in efficiency
HDR (%)
30
25
20
15
10
5
0
Mock
SpyCas9
SpyCas9-GFP
SpyCas9-hTdT
SpyCas9-RecJ
SpyCas9-RecE
SpyCas9-lambda
SpyCas9-mungbean
SpyCas9-T5
CCR5
DHCR7

도면17

A

B



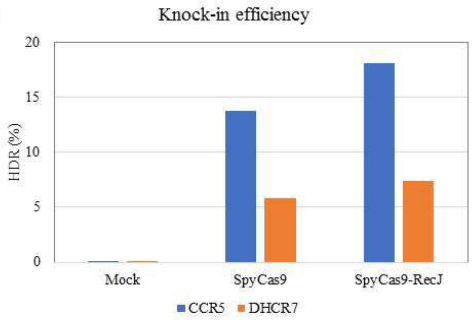
C

Percent change in efficiency of SpyCas9-RecJ over SpyCas9		
	CCR5	DHCR7
SpyCas9-RecJ	+30%	+170%

도면18

A

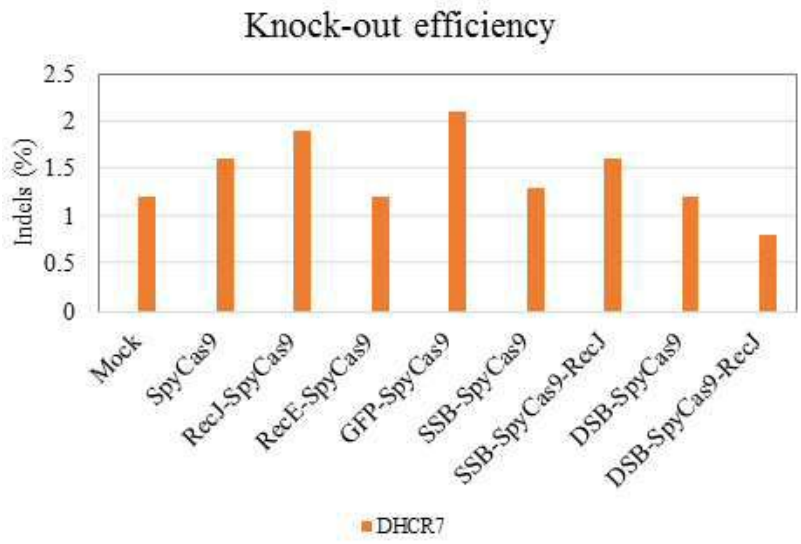
B



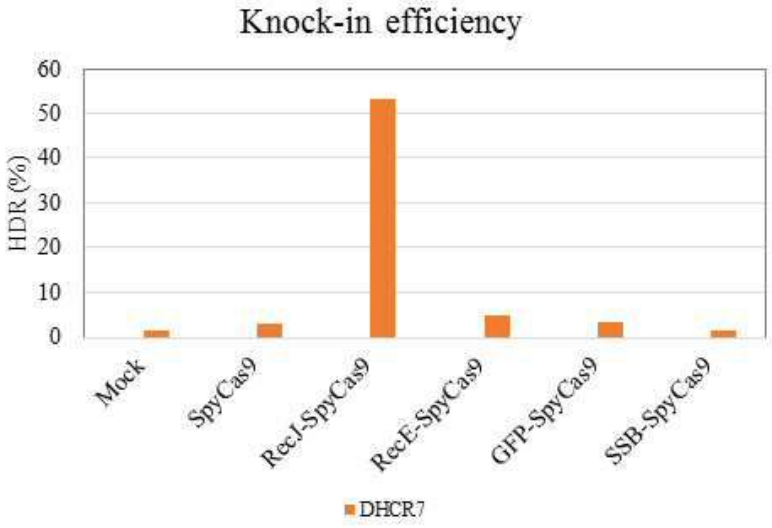
C

Percent change in efficiency of SpyCas9-RecJ over SpyCas9		
	CCR5	DHCR7
SpyCas9-RecJ	+30%	+30%

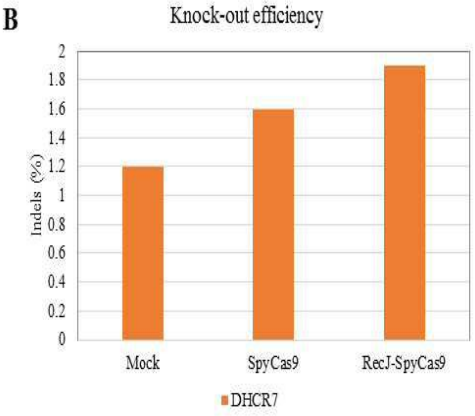
도면19



도면20



도면21



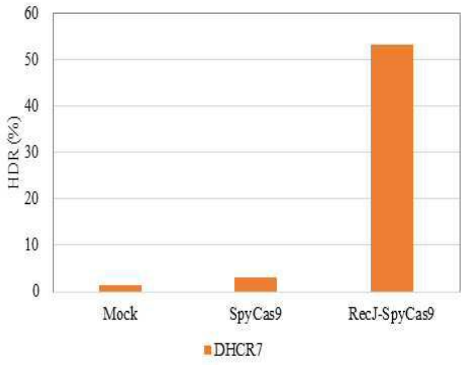
C

Percent change in efficiency of RecJ-SpyCas9 over SpyCas9	
	DHCR7
RecJ-SpyCas9	+20%

도면22

A

B



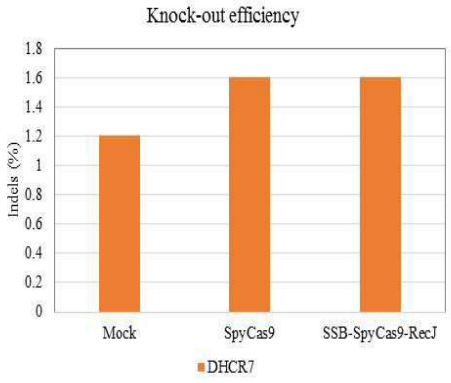
C

Percent change in efficiency of RecJ-SpyCas9 over SpyCas9	
	DHCR7
RecJ-SpyCas9	+1770%

도면23

A

B



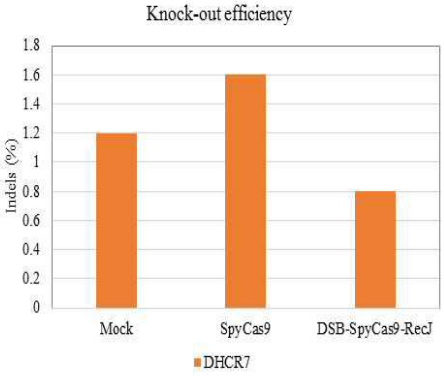
C

Percent change in efficiency of SSB-SpyCas9-RecJ over SpyCas9	
	DHCR7
SSB-SpyCas9-RecJ	0%

도면24

A

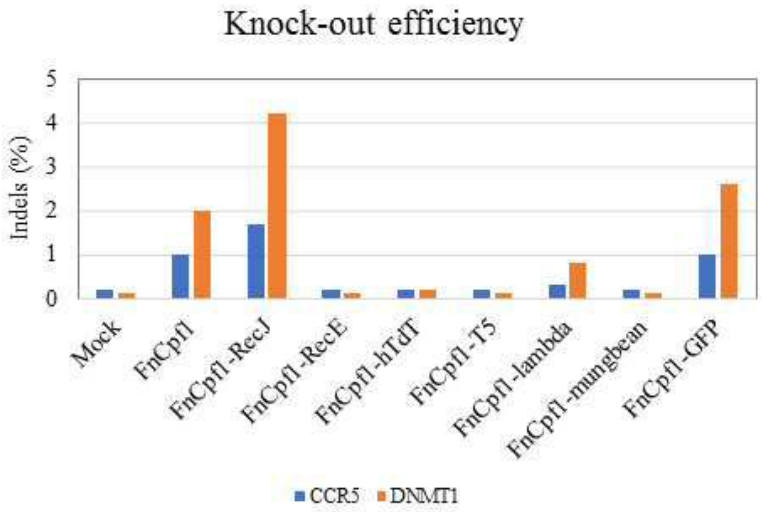
B



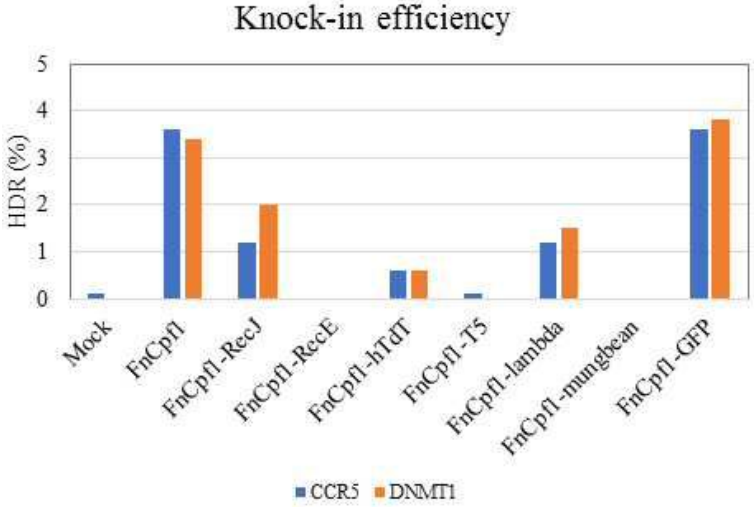
C

Percent change in efficiency of DSB-SpyCas9-RecJ over SpyCas9	
	DHCR7
DSB-SpyCas9-RecJ	-50%

도면25



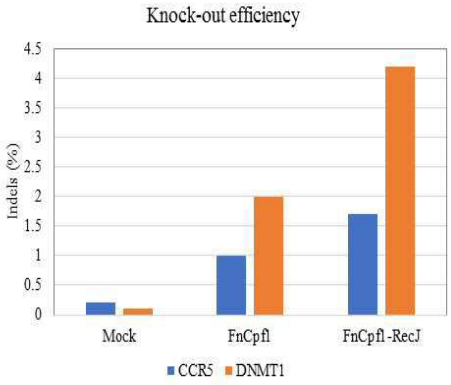
도면26



도면27

A

B



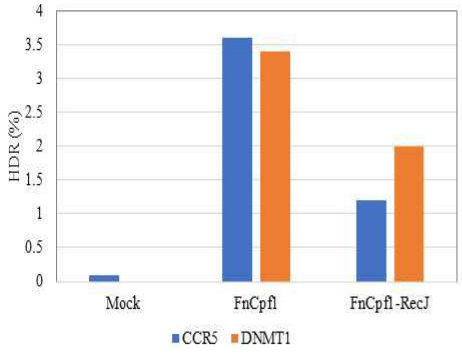
C

Percent change in efficiency of FnCpf1-RecJ over FnCpf1		
	CCR5	DNMT1
FnCpf1-RecJ	+70%	+110%

도면28

A

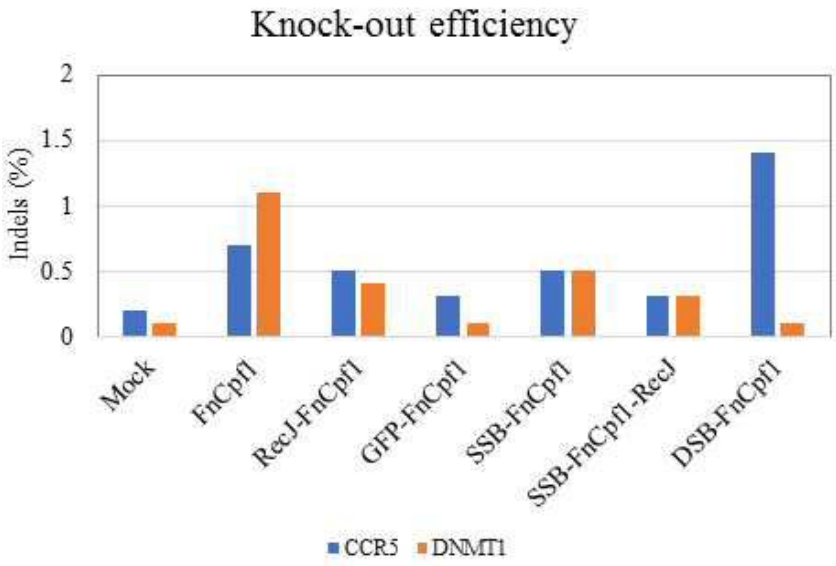
B



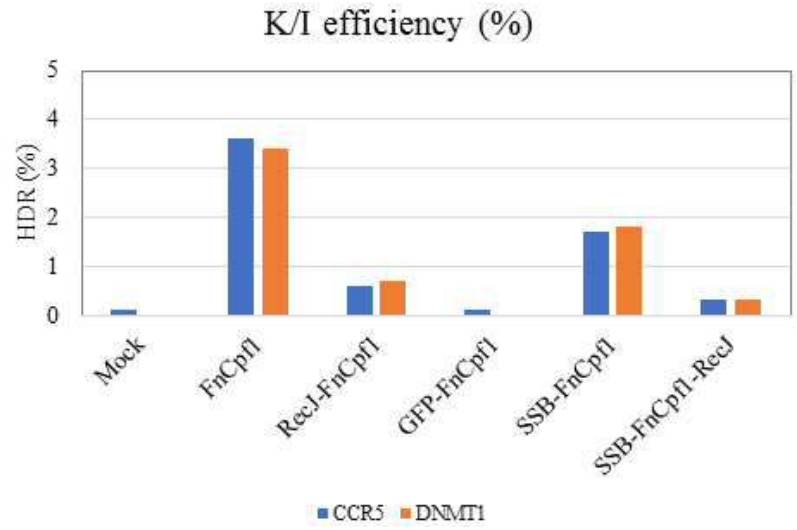
C

Percent change in efficiency of FnCpf1-RecJ over FnCpf1		
	CCR5	DNMT1
FnCpf1-RecJ	-70%	-40%

도면29



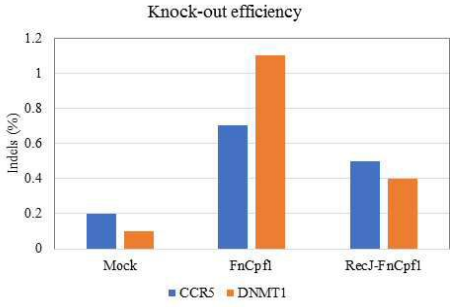
도면30



도면31

A

B



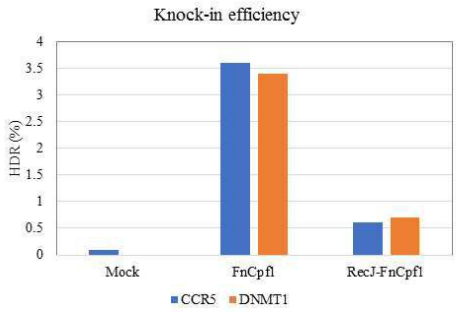
C

Percent change in efficiency of RecJ-FnCpf1 over FnCpf1		
	CCR5	DNMT1
RecJ-FnCpf1	-30%	-60%

도면32

A

B



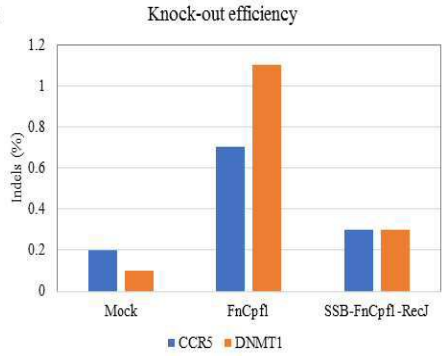
C

Percent change in efficiency of RecJ-FnCpf1 over FnCpf1		
	CCR5	DNMT1
RecJ-FnCpf1	-80%	-80%

도면33

A

B



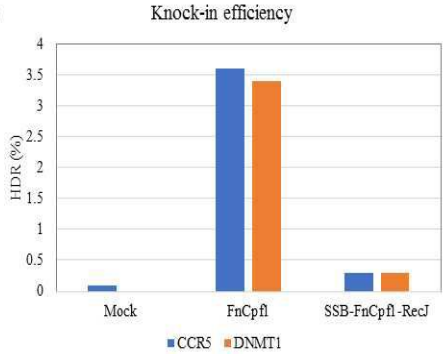
C

Percent change in efficiency of SSB-FnCpf1-RecJ over FnCpf1		
	CCR5	DNMT1
SSB-FnCpf1-RecJ	-60%	-70%

도면34

A

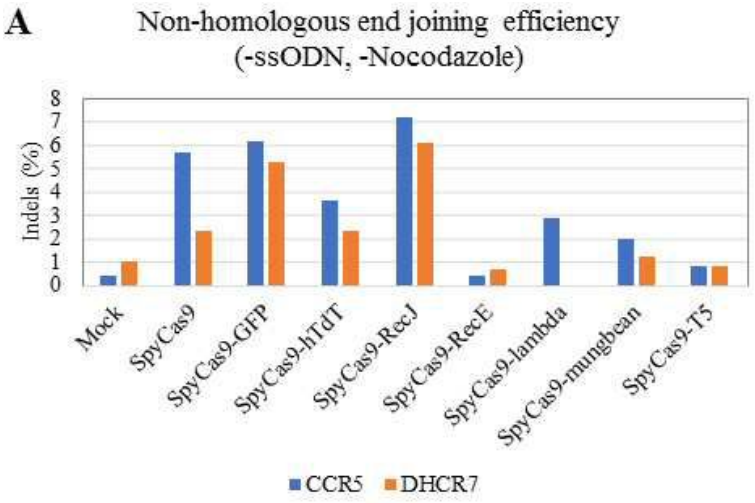
B



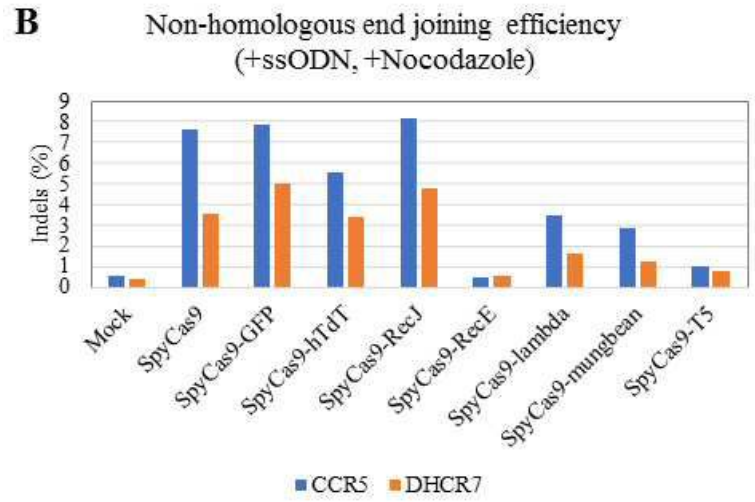
C

Percent change in efficiency of SSB-FnCpf1-RecJ over FnCpf1		
	CCR5	DNMT1
SSB-FnCpf1-RecJ	-90%	-90%

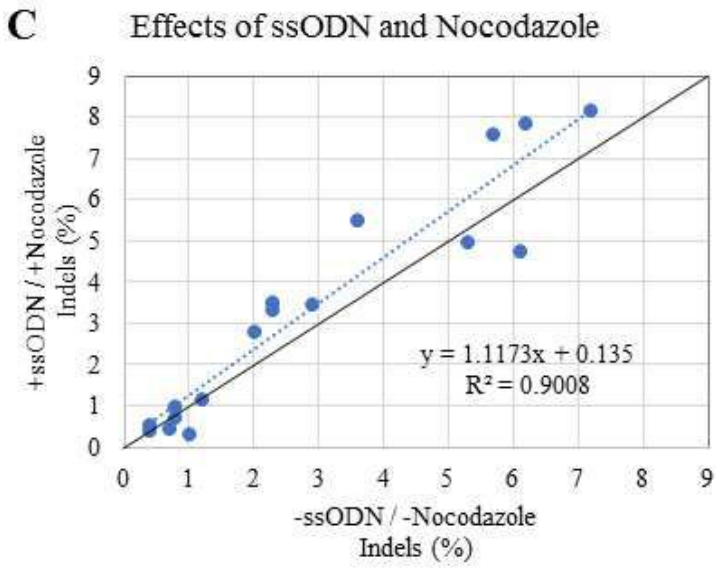
도면35a



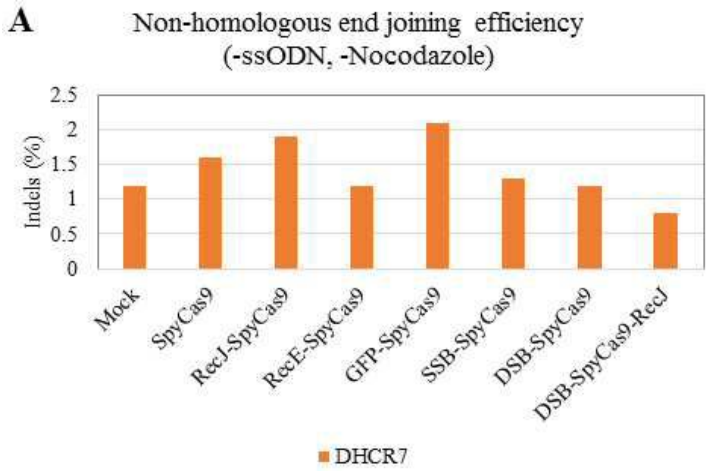
도면35b



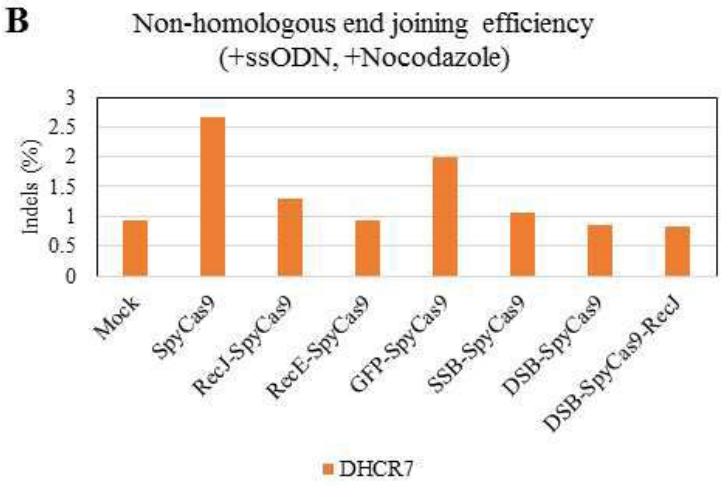
도면35c



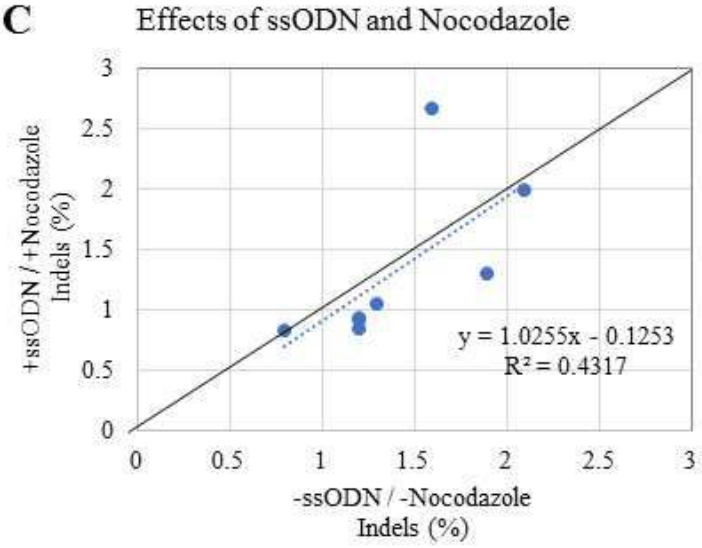
도면36a



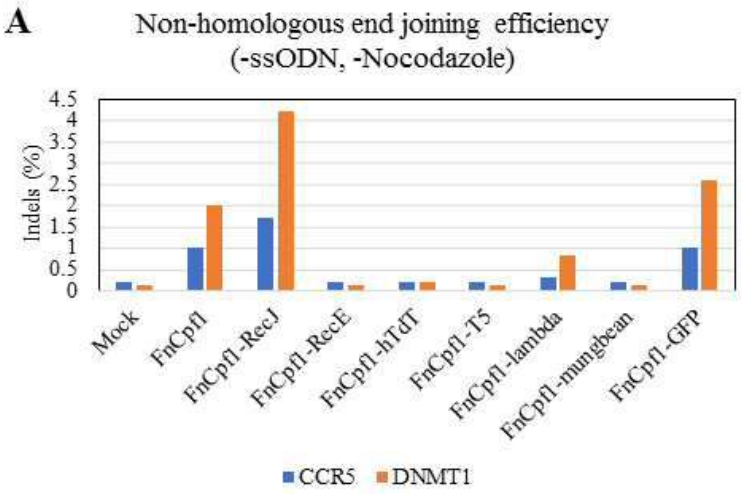
도면36b



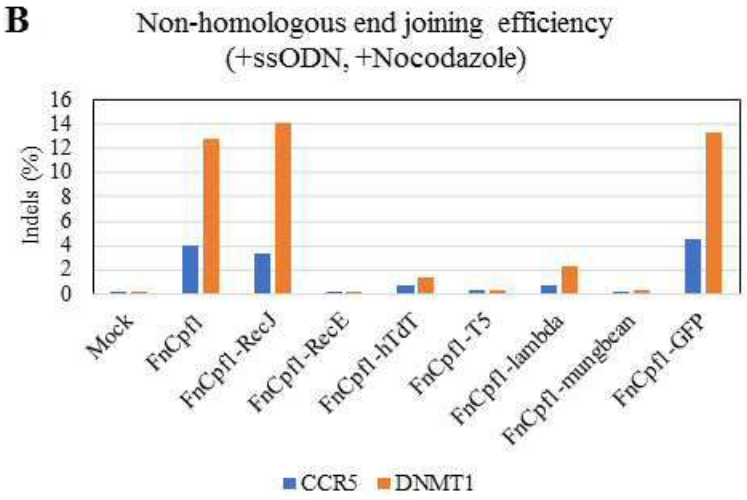
도면36c



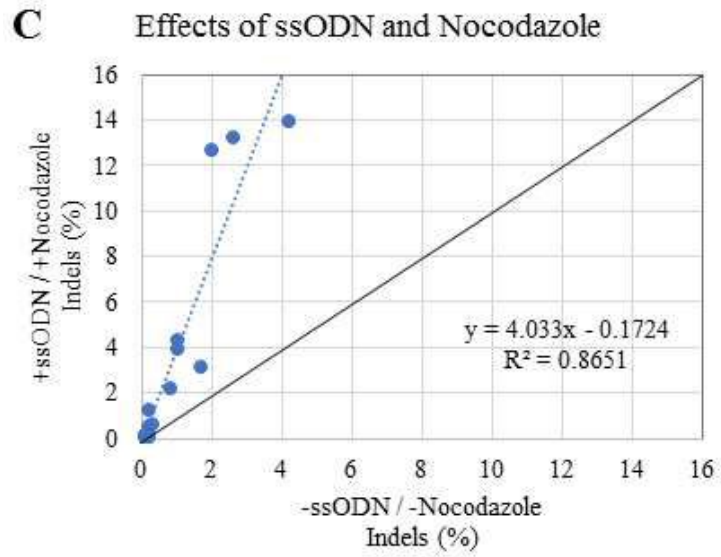
도면37a



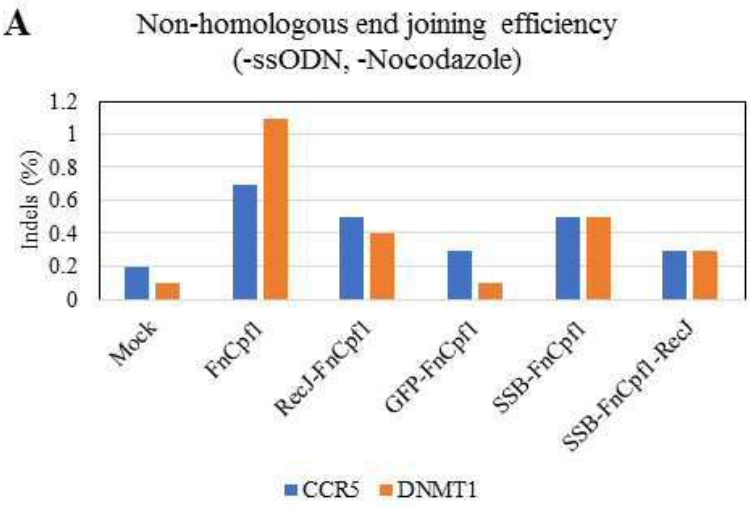
도면37b



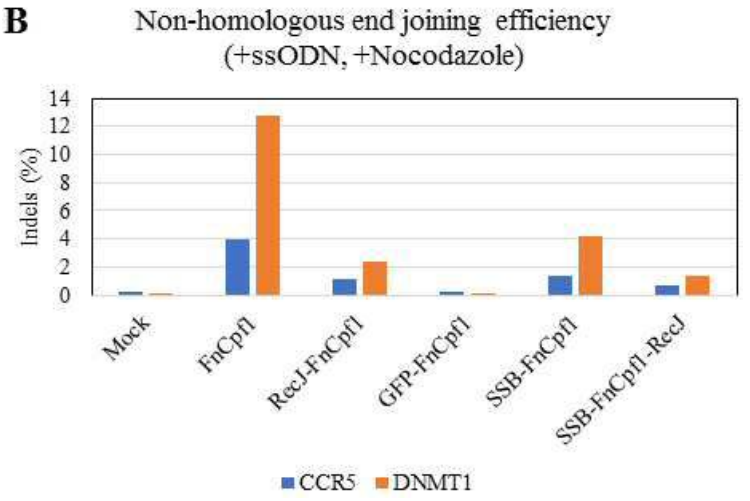
도면37c



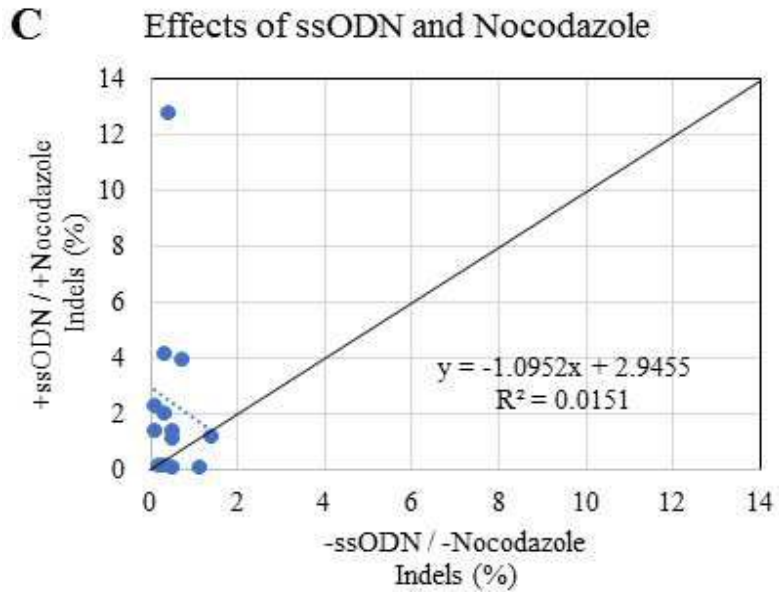
도면38a



도면38b

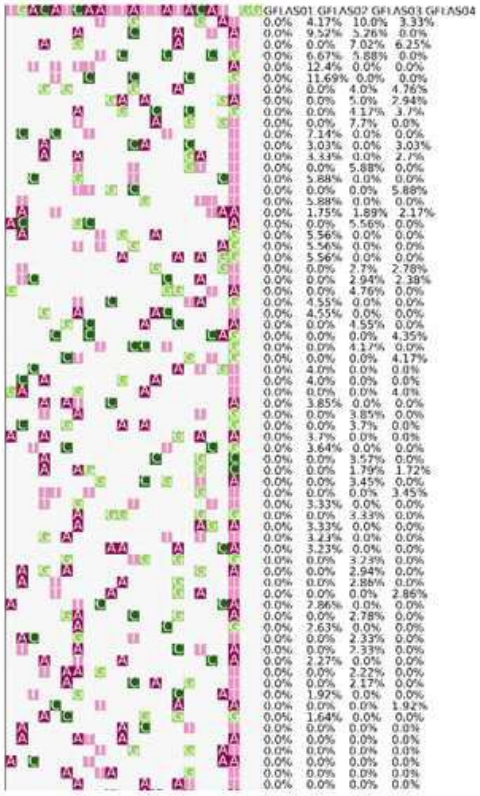


도면38c



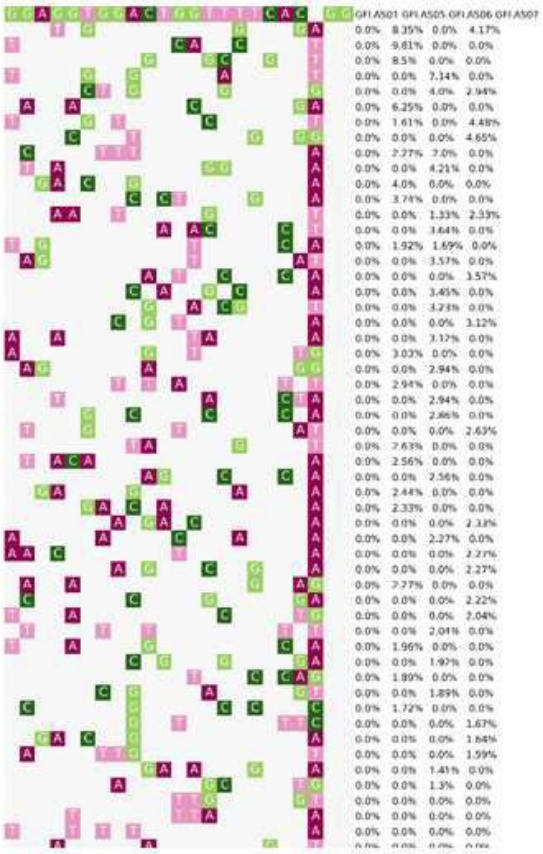
도면39a

Off-Target effect of Crispr on CCR5



도면39b

Off-Target effect of Crispr on DHCR7



도면40

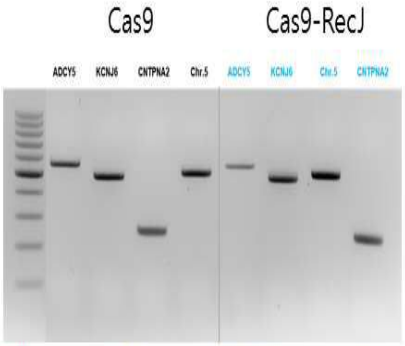
Human CCR 5 off-target candidate

A

Human CCR 5 off-target sequence list

Target	Gene	Spacer sequence
On	CCR5	TGACATCAATTATTATACAT
Off #1	ADCY5	TGACATCAATTATTATAgAT
Off #2	KCNJ6	TGACATCAcTTATTATgCAT
Off #3	CNTNAP2	TGACATaAATTATTcTACAT
Off #4	Chr. 5 N/A	TGAaATCAATTATcATAgAT

B



ADCY5: 641 bp; KCNJ6: 566 bp; CNTNAP2: 300 bp; Chr. 5: 605 bp

도면41

off-targets of CCR5	Cas9		Cas9-RecJ	
	reads (372,767)	percent (%)	reads (406,009)	percent (%)
ADCY5	269	0.0007	1009	0.0025
KCNJ6	339	0.0009	254	0.0006
CNTPNA2	624	0.0017	437	0.0011
Ch.5	601	0.0016	573	0.0014

(by amplicon deep sequencing)

도면42

SpyCas9

Nucleotide sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA

Amino acid sequence

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFELRRQACGRMDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDAAALDLQKDQQDLRRREVDETADLPALERLPPLLRRYVASRGVRSAQELSRSVKGHPQLIKSLKDVQAVEELRSVVKGMLPNVQQLSGVDKNAVELVNAFRGTRNVVGDFDADQAISTALSVLLAMRSLGCSNIDYLVPNERFDGVFGLSPEVVDQAHAKRGAQALNT

도면43

SpyCas9-RecJ

Nucleotide sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGA

Amino acid sequence

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFELRRQACGRMDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDAAALDLQKDQQDLRRREVDETADLPALERLPPLLRRYVASRGVRSAQELSRSVKGHPQLIKSLKDVQAVEELRSVVKGMLPNVQQLSGVDKNAVELVNAFRGTRNVVGDFDADQAISTALSVLLAMRSLGCSNIDYLVPNERFDGVFGLSPEVVDQAHAKRGAQALNTQPTRKEIEQGMQVEALTCEKLERSRDTLPGGLAMYHPEWHQQVVVGLASRIKERFHRPVIAFAPAGDGTLKSGSRGIQGLHMVRDALERLDTLPPSLLQGEVVSDGPLSPAEMTMEVAQLLRDAGPWGQMPPEPLPDGHPRLLQQRLVGERHLKVMVEPVQGGPLLDIGRAPNVDTALWPDNGVRBVQLAYYLDINHPGNRSLQIIEDNIWPLKQKPAATKKAGQAKKKKHHHHHH.

도면44

도면45

도면46

도면47

FnCpf1

Nucleotide sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCATGTCAATTTATCAAGAATTTGTTAATAAATATAGTTTAAGTAAAA
CTCTAAGATTTGAGTTAATCCCACAGGGTAAAACACTTGAAAACATAAAAGCAAGAGGTTTGATTTTAGATGATGAGAAAAGAGCTAAAGACTACAAAAAGGCTAAACAAATAATTGATAAATATCATCAGTTTTTTATAGAGGAGATATTAAGTTCGGTTTGTATTAGCGAAGATTTATTACA
AAACTATTCTGATGTTTATTTTAAACTTAAAAAGAGTGATGATGATAATCTACAAAAAGATTTTAAAAGTGCAAAAGATACGATAAAGAAACAAATATCTGAATATATAAAGGACTCAGAGAAATTTAAGAATTTGTTTAATCAAAACCTTATCGATGCTAAAAAAGGGCAAGAGTCAGATTTA
ATTCTATGGCTAAAGCAATCTAAGGATAATGGTATAGAACTATTTAAAGCCAATAGTGATATCACAGATATAGATGAGGCGTTAGAAATAATCAAATCTTTTAAAGGTTGGACAACTTATTTTAAGGGTTTTCATGAAAATAGAAAAAATGTTTATAGTAGCAATGATATTCCTACATCTATTATT
TATAGGATAGTAGATGATAATTTGCCTAAATTTCTAGAAAATAAAGCTAAGTATGAGAGTTTAAAAGACAAAGCTCCAGAAGCTATAAACTATGAACAAATTAAAAAAGATTTGGCAGAAGAGCTAACCTTTGATATTGACTACAAAACATCTGAAGTTAATCAAAGAGTTTTTTCACTTGATG
AAGTTTTTGAGATAGCAAACTTTAATAATTATCTAAATCAAAGTGGTATTACTAAATTTAATACTATTATTGGTGGTAAATTTGTAAATGGTGAAAATACAAAGAGAAAAGGTATAAATGAATATATAAATCTATACTCACAGCAAATAAATGATAAAACACTCAAAAAAATATAAAATGAGTGTT
TTATTTAAGCAAATTTTAAGTGATACAGAATCTAAATCTTTTGTAATTGATAAGTTAGAAGATGATAGTGATGTAGTTACAACGATGCAAAGTTTTTATGAGCAAATAGCAGCTTTTAAAACAGTAGAAGAAAAAATCTATTAAAGAAACACTATCTTTATTATTTGATGATTTAAAAGCTCAAAAA
CTTGATTTGAGTAAAATTTATTTTAAAAATGATAAATCTCTTACTGATCTATCACAACAAGTTTTTGATGATTATAGTGTTATTGGTACAGCGGTACTAGAATATATAACTCAACAAATAGCACCTAAAAATCTTGATAACCCTAGTAAGAAAGAGCAAGAATTAATAGCCAAAAAAACTGAAAA
AGCAAAATACTTATCTCTAGAAACTATAAAGCTTGCCTTAGAAGAATTTAATAAGCATAGAGATATAGATAAACAGTGTAGGTTTGAAGAAATACTTGCAAACTTTGCGGCTATTCCGATGATATTTGATGAAATAGCTCAAAACAAAGACAATTTGGCACAGATATCTATCAAATATCAAAAT
CAAGGTAAAAAAGACCTACTTCAAGCTAGTGCGGAAGATGATGTTAAAGCTATCAAGGATCTTTTAGATCAAACTAATAATCTCTTACATAAACTAAAAAATATTTCATATTAGTCAGTCAGAAGATAAGGCAAATATTTTAGACAAGGATGAGCATTTTTATCTAGTATTTGAGGAGTGCTACTT
TGAGCTAGCGAATATAGTGCCTCTTTATAACAAAATTAGAAACTATATAACTCAAAAGCCATATAGTGATGAGAAATTTAAGCTCAATTTTGAGAACTCGACTTTGGCTAATGGTTGGGATAAAAATAAAGAGCCTGACAATACGGCAATTTTATTTATCAAAGATGATAAATATTATCTGGGT
GTGATGAATAAGAAAAATAACAAAATATTTGATGATAAAGCTATCAAAGAAAATAAAGGCGAGGGTTATAAAAAAATTGTTTATAAACTTTTACCTGGCGCAAATAAAATGTTACCTAAGGTTTTCTTTTCTGCTAAATCTATAAAATTTTATAATCCTAGTGAAGATATACTTAGAATAAGAAA
TCATTCCACACATACAAAAAATGGTAGTCCTCAAAAAGGATATGAAAAATTTGAGTTTAATATTGAAGATTGCCGAAAATTTATAGATTTTTATAAACAGTCTATAAGTAAGCATCCGGAGTGGAAAGATTTTGGATTTAGATTTTCTGATACTCAAAGATATAATTCTATAGATGAATTTTATAG
AGAAGTTGAAAATCAAGGCTACAAACTAACTTTTGAAAATATATCAGAGAGCTATATTGATAGCGTAGTTAATCAGGGTAAATTGTACCTATTCCAAATCTATAATAAAGATTTTTCAGCTTATAGCAAAGGGCGACCAAATCTACATACTTTATATTGGAAAGCGCTGTTTGATGAGAGAAAT
CTTCAAGATGTGGTTTATAAGCTAAATGGTGAGGCAGAGCTTTTTTATCGTAAACAATCAATACCTAAAAAAATCACTCACCCAGCTAAAGAGGCAATAGCTAATAAAAACAAAGATAATCCTAAAAAAGAGAGTGTTTTTGAATATGATTTAATCAAAGATAAACGCTTTACTGAAGATAAGT
TTTTCTTTCACTGTCCTATTACAATCAATTTTAAATCTAGTGGAGCTAATAAGTTTAATGATGAAATCAATTTATTGCTAAAAGAAAAAGCAAATGATGTTCATATATTAAGTATAGATAGAGGTGAAAGACATTTAGCTTACTATACTTTGGTAGATGGTAAAGGCAATATCATCAAACAAGATA
CTTTCAACATCATTGGTAATGATAGAATGAAAACAAACTACCATGATAAGCTTGCTGCAATAGAGAAAGATAGGGATTCAGCTAGGAAAGACTGGAAAAAGATAAATAACATCAAAGAGATGAAAGAGGGCTATCTATCTCAGGTAGTTCATGAAATAGCTAAGCTAGTTATAGAGTATAAT
GCTATTGTGGTTTTTGAGGATTTAAATTTTGGATTTAAAAGAGGGCGTTTCAAGGTAGAGAAGCAGGTCTATCAAAAGTTAGAAAAAATGCTAATTGAGAAACTAAACTATCTAGTTTTCAAAGATAATGAGTTTGATAAAACTGGGGGAGTGCTTAGAGCTTATCAGCTAACAGCACCTTTTG
AGACTTTTAAAAAGATGGGTAAACAAACAGGTATTATCTACTATGTACCAGCTGGTTTTACTTCAAAAATTTGTCCTGTAACTGGTTTTGTAAATCAGTTATATCCTAAGTATGAAAGTGTCAGCAAATCTCAAGAGTTCTTTAGTAAGTTTGACAAGATTTGTTATAACCTTGATAAGGGCTATTT
TGAGTTTAGTTTTGATTATAAAAACTTTGGTGACAAGGCTGCCAAAGGCAAGTGGACTATAGCTAGCTTTGGGAGTAGATTGATTAACTTTAGAAATTCAGATAAAAATCATAATTGGGATACTCGAGAAGTTTATCCAACTAAAGAGTTGGAGAAATTGCTAAAAGATTATTCTATCGAATAT
GGGCATGGCGAATGTATCAAAGCAGCTATTTGCGGTGAGAGCGACAAAAAGTTTTTTGCTAAGCTAACTAGTGTCCTAAATACTATCTTACAAATGCGTAACTCAAAAACAGGTACTGAGTTAGATTATCTAATTTCACCAGTAGCAGATGTAAATGGCAATTTCTTTGATTCGCGACAGGCG
CCAAAAAATATGCCTCAAGATGCTGATGCCAATGGTGCTTATCATATTGGGCTAAAAGGTCTGATGCTACTAGGTAGGATCAAAAATAATCAAGAGGGCAAAAAACTCAATTTGGTTATCAAAAATGAAGAGTATTTTGAGTTCGTGCAGAATAGGAATAACCAAGCGGCCGCACTCGAGAA
AAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGTCGACACCACCACCACCACCACTGAGATCCGGCTGCTAA

Amino acid sequence

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFELRRQACGRMSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDN
GIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDV
VTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQ
SEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSD
TQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDT
FNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGD
KAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNNQAAALEKRPAATKKAGQAKKKKSTP
PPPPLRSGC

도면48

FnCpf1-RecJ

Nucleotide sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCATGTCAATTTATCAAGAATTTGTTAATAAATATAGTTTAAGTAAAA
CTCTAAGATTTGAGTTAATCCCACAGGGTAAAACACTTGAAAACATAAAAGCAAGAGGTTTGATTTTAGATGATGAGAAAAGAGCTAAAGACTACAAAAAGGCTAAACAAATAATTGATAAATATCATCAGTTTTTTATAGAGGAGATATTAAGTTCGGTTTGTATTAGCGAAGATTTATTACA
AAACTATTCTGATGTTTATTTTAAACTTAAAAAGAGTGATGATGATAATCTACAAAAAGATTTTAAAAGTGCAAAAGATACGATAAAGAAACAAATATCTGAATATATAAAGGACTCAGAGAAATTTAAGAATTTGTTTAATCAAAACCTTATCGATGCTAAAAAAGGGCAAGAGTCAGATTTA
ATTCTATGGCTAAAGCAATCTAAGGATAATGGTATAGAACTATTTAAAGCCAATAGTGATATCACAGATATAGATGAGGCGTTAGAAATAATCAAATCTTTTAAAGGTTGGACAACTTATTTTAAGGGTTTTCATGAAAATAGAAAAAATGTTTATAGTAGCAATGATATTCCTACATCTATTATT
TATAGGATAGTAGATGATAATTTGCCTAAATTTCTAGAAAATAAAGCTAAGTATGAGAGTTTAAAAGACAAAGCTCCAGAAGCTATAAACTATGAACAAATTAAAAAAGATTTGGCAGAAGAGCTAACCTTTGATATTGACTACAAAACATCTGAAGTTAATCAAAGAGTTTTTTCACTTGATG
AAGTTTTTGAGATAGCAAACTTTAATAATTATCTAAATCAAAGTGGTATTACTAAATTTAATACTATTATTGGTGGTAAATTTGTAAATGGTGAAAATACAAAGAGAAAAGGTATAAATGAATATATAAATCTATACTCACAGCAAATAAATGATAAAACACTCAAAAAAATATAAAATGAGTGTT
TTATTTAAGCAAATTTTAAGTGATACAGAATCTAAATCTTTTGTAATTGATAAGTTAGAAGATGATAGTGATGTAGTTACAACGATGCAAAGTTTTTATGAGCAAATAGCAGCTTTTAAAACAGTAGAAGAAAAAATCTATTAAAGAAACACTATCTTTATTATTTGATGATTTAAAAGCTCAAAAA
CTTGATTTGAGTAAAATTTATTTTAAAAATGATAAATCTCTTACTGATCTATCACAACAAGTTTTTGATGATTATAGTGTTATTGGTACAGCGGTACTAGAATATATAACTCAACAAATAGCACCTAAAAATCTTGATAACCCTAGTAAGAAAGAGCAAGAATTAATAGCCAAAAAAACTGAAAA
AGCAAAATACTTATCTCTAGAAACTATAAAGCTTGCCTTAGAAGAATTTAATAAGCATAGAGATATAGATAAACAGTGTAGGTTTGAAGAAATACTTGCAAACTTTGCGGCTATTCCGATGATATTTGATGAAATAGCTCAAAACAAAGACAATTTGGCACAGATATCTATCAAATATCAAAAT
CAAGGTAAAAAAGACCTACTTCAAGCTAGTGCGGAAGATGATGTTAAAGCTATCAAGGATCTTTTAGATCAAACTAATAATCTCTTACATAAACTAAAAAATATTTCATATTAGTCAGTCAGAAGATAAGGCAAATATTTTAGACAAGGATGAGCATTTTTATCTAGTATTTGAGGAGTGCTACTT
TGAGCTAGCGAATATAGTGCCTCTTTATAACAAAATTAGAAACTATATAACTCAAAAGCCATATAGTGATGAGAAATTTAAGCTCAATTTTGAGAACTCGACTTTGGCTAATGGTTGGGATAAAAATAAAGAGCCTGACAATACGGCAATTTTATTTATCAAAGATGATAAATATTATCTGGGT
GTGATGAATAAGAAAAATAACAAAATATTTGATGATAAAGCTATCAAAGAAAATAAAGGCGAGGGTTATAAAAAAATTGTTTATAAACTTTTACCTGGCGCAAATAAAATGTTACCTAAGGTTTTCTTTTCTGCTAAATCTATAAAATTTTATAATCCTAGTGAAGATATACTTAGAATAAGAAA
TCATTCCACACATACAAAAAATGGTAGTCCTCAAAAAGGATATGAAAAATTTGAGTTTAATATTGAAGATTGCCGAAAATTTATAGATTTTTATAAACAGTCTATAAGTAAGCATCCGGAGTGGAAAGATTTTGGATTTAGATTTTCTGATACTCAAAGATATAATTCTATAGATGAATTTTATAG
AGAAGTTGAAAATCAAGGCTACAAACTAACTTTTGAAAATATATCAGAGAGCTATATTGATAGCGTAGTTAATCAGGGTAAATTGTACCTATTCCAAATCTATAATAAAGATTTTTCAGCTTATAGCAAAGGGCGACCAAATCTACATACTTTATATTGGAAAGCGCTGTTTGATGAGAGAAAT
CTTCAAGATGTGGTTTATAAGCTAAATGGTGAGGCAGAGCTTTTTTATCGTAAACAATCAATACCTAAAAAAATCACTCACCCAGCTAAAGAGGCAATAGCTAATAAAAACAAAGATAATCCTAAAAAAGAGAGTGTTTTTGAATATGATTTAATCAAAGATAAACGCTTTACTGAAGATAAGT
TTTTCTTTCACTGTCCTATTACAATCAATTTTAAATCTAGTGGAGCTAATAAGTTTAATGATGAAATCAATTTATTGCTAAAAGAAAAAGCAAATGATGTTCATATATTAAGTATAGATAGAGGTGAAAGACATTTAGCTTACTATACTTTGGTAGATGGTAAAGGCAATATCATCAAACAAGATA
CTTTCAACATCATTGGTAATGATAGAATGAAAACAAACTACCATGATAAGCTTGCTGCAATAGAGAAAGATAGGGATTCAGCTAGGAAAGACTGGAAAAAGATAAATAACATCAAAGAGATGAAAGAGGGCTATCTATCTCAGGTAGTTCATGAAATAGCTAAGCTAGTTATAGAGTATAAT
GCTATTGTGGTTTTTGAGGATTTAAATTTTGGATTTAAAAGAGGGCGTTTCAAGGTAGAGAAGCAGGTCTATCAAAAGTTAGAAAAAATGCTAATTGAGAAACTAAACTATCTAGTTTTCAAAGATAATGAGTTTGATAAAACTGGGGGAGTGCTTAGAGCTTATCAGCTAACAGCACCTTTTG
AGACTTTTAAAAAGATGGGTAAACAAACAGGTATTATCTACTATGTACCAGCTGGTTTTACTTCAAAAATTTGTCCTGTAACTGGTTTTGTAAATCAGTTATATCCTAAGTATGAAAGTGTCAGCAAATCTCAAGAGTTCTTTAGTAAGTTTGACAAGATTTGTTATAACCTTGATAAGGGCTATTT
TGAGTTTAGTTTTGATTATAAAAACTTTGGTGACAAGGCTGCCAAAGGCAAGTGGACTATAGCTAGCTTTGGGAGTAGATTGATTAACTTTAGAAATTCAGATAAAAATCATAATTGGGATACTCGAGAAGTTTATCCAACTAAAGAGTTGGAGAAATTGCTAAAAGATTATTCTATCGAATAT
GGGCATGGCGAATGTATCAAAGCAGCTATTTGCGGTGAGAGCGACAAAAAGTTTTTTGCTAAGCTAACTAGTGTCCTAAATACTATCTTACAAATGCGTAACTCAAAAACAGGTACTGAGTTAGATTATCTAATTTCACCAGTAGCAGATGTAAATGGCAATTTCTTTGATTCGCGACAGGCG
CCAAAAAATATGCCTCAAGATGCTGATGCCAATGGTGCTTATCATATTGGGCTAAAAGGTCTGATGCTACTAGGTAGGATCAAAAATAATCAAGAGGGCAAAAAACTCAATTTGGTTATCAAAAATGAAGAGTATTTTGAGTTCGTGCAGAATAGGAATAACCAAGCGGCCGCACTCGACCT
GCAGGTGAAACAACAGATACAACTTCGTCGCCGTGAAGTCGATGAAACGGCAGACTTGCCCGCTGAATTGCCTCCCTTGCTGCGCCGTTTATACGCCAGCCGGGGCGTGCGCAGTGCGCAAGAACTGGAACGCAGTGTTAAAGGTATGTTGCCCTGGCAGCAACTGAGCGGCGTCGAAA
AGGCCGTTGAGATCCTTTACAACGCTTTTCGCGAAGGAACGCGGATTATTGTGGTCGGCGATTTTGACGCCGACGGCGCGACCAGCACGGCTCTAAGCGTGCTGGCGATGCGCTCGCTTGGTTGCAGCAATATCGACTATCTGGTACCAAACCGTTTCGAAGACGGTTACGGCTTAAGCCC
GGAAGTAGTCGATCAGGCCCATGCCCGTGGCGCGCAGTTAATTGTCACGGTGGATAACGGTATTTCCTCCCATGCGGGCGTTGAACACGCTCGCTCGTTGGGCATTCCGGTTATTGTTACCGATCACCATTTGCCGGGCGAAACATTACCCGCAGCGGAAGCGATCATTAACCCTAACTTGC
GCGACTGTAATTTCCCGTCGAAATCACTGGCAGGCGTGGGTGTGGCGTTTTATCTGATGCTGGCGCTGCGCACCTTTTTGCGCGATCAGGGCTGGTTTGATGAGCGTGGCATCGCAATTCCTAACCTGGCAGAACTGCTGGATCTGGTCGCGCTGGGAACAGTGGCGGACGTCGTGCCGCT
GGACGCTAATAATCGCATTCTGACCTGGCAGGGGATGAGTCGCATCCGTGCCGGAAAGTGCCGTCCAGGGATTAAAGCGCTGCTGGAAGTGGCAAACCGTGATGCACAAAAACTCGCCGCCAGCGATTTAGGTTTTGCGCTGGGGCCACGTCTCAATGCTGCCGGACGACTGGACGATAT
GTCCGTCGGTGTGGCGCTCTTGCTGTGCGACAACATCGGCGAAGCGCGCGTGCTGGCAAATGAACTCGATGCGCTAAACCAGACGCGAAAAGAGATCGAACAAGGAATGCAAGTTGAAGCCCTGACCCTGTGCGAGAAACTGGAGCGAAGTCGCGACACGCTACCCGGCGGGCTGGCA
ATGTATCACCCCGAATGGCATCAGGGCGTTGTCGGTATTCTGGCTTCGCGCATCAAAGAGCGTTTTCACCGTCCGGTTATCGCCTTTGCGCCAGCAGGTGATGGTACGCTGAAAGGTTCAGGTCGCTCCATTCAGGGGCTGCATATGCGTGATGCACTGGAGCGATTAGACACACTCTACCC
TGGCATGATACTGAAGTTTGGCGGTCATGCGATGGCGGCGGGTTTGTCGCTGGAAGAGGATAAATTCGAACTCTTTCAACAACGGTTTGGCGAGCTGGTTACCGAGTGGCTGGACCCTTCGCTATTGCAAGGCGAAGTGGTGTCAGACGGCCCGTTAAGCCCGGCCGAAATGACCATGGA
AGTGGCGCAGCTGCTGCGCGATGCTGGCCCGTGGGGGCAGATGTTCCCGGAGCCGCTGTTTGATGGTCATTTCCGTCTGCTGCAACAGCGGCTGGTGGGCGAACGTCATTTGAAAGTCATGGTCGAACCGGTCGGCGGCGGTCCGCTGCTGGATGGTATTGCTTTTAATGTCGATACCGCC
CTCTGGCCGGATAACGGCGTGCGCGAAGTGCAACTGGCTTACAAGCTCGATATCAACGAGTTTCGCGGCAACCGCAGCCTGCAAATTATCATCGACAATATCTGGCCAATTCTGCAGAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGCACCACCACCACC
ACCACTGA

Amino acid sequence

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFELRRQACGRMSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQSKDN
GIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDV
VTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQ
SEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSD
TQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDT
FNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGD
KAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNNQAAALDLQVKQQIQLRRREVDETAD
LPAELPPLLRRLYASRGVRSAQELERSVKGMLPWQQLSGVEKAVEILYNAFREGTRIIVVGDFDADGATSTALSVLAMRSLGCSNIDYLVPNRFEDGYGLSPEVVDQAHARGAQLIVTVDNGISSHAGVEHARSLGIPVIVTDHHLPGETLPAAEAIINPNLRDCNFPSKSLAGVGVAFYLMLALRTFLRDQ
GWFDERGIAIPNLAELLDLVALGTVADVVPLDANNRILTWQGMSRIRAGKCRPGIKALLEVANRDAQKLAASDLGFALGPRLNAAGRLDDMSVGVALLLCDNIGEARVLANELDALNQTRKEIEQGMQVEALTLCEKLERSRDTLPGGLAMYHPEWHQGVVGILASRIKERFHRPVIAFAPAGDGTLKG
SGRSIQGLHMRDALERLDTLYPGMILKFGGHAMAAGLSLEEDKFELFQQRFGELVTEWLDPSLLQGEVVSDGPLSPAEMTMEVAQLLRDAGPWGQMFPEPLFDGHFRLLQQRLVGERHLKVMVEPVGGGPLLDGIAFNVDTALWPDNGVREVQLAYKLDINEFRGNRSLQIIDNIWPILQKRPAATK
KAGQAKKKKHHHHHH

도면49

RecJ-FnCpf1

Nucleotide sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGTGAAACAACAGATACAACTTCGTCGCCGTGAAGTCGATGAAACGGCAGACTTGCCCGCTGAATTGCCTCCCTTGCTG
CGCCGTTTATACGCCAGCCGGGGCGTGCGCAGTGCGCAAGAACTGGAACGCAGTGTTAAAGGTATGTTGCCCTGGCAGCAACTGAGCGGCGTCGAAAAGGCCGTTGAGATCCTTTACAACGCTTTTCGCGAAGGAACGCGGATTATTGTGGTCGGCGATTTTGACGCCGACGGCGCGACC
AGCACGGCTCTAAGCGTGCTGGCGATGCGCTCGCTTGGTTGCAGCAATATCGACTATCTGGTACCAAACCGTTTCGAAGACGGTTACGGCTTAAGCCCGGAAGTAGTCGATCAGGCCCATGCCCGTGGCGCGCAGTTAATTGTCACGGTGGATAACGGTATTTCCTCCCATGCGGGCGTTG
AACACGCTCGCTCGTTGGGCATTCCGGTTATTGTTACCGATCACCATTTGCCGGGCGAAACATTACCCGCAGCGGAAGCGATCATTAACCCTAACTTGCGCGACTGTAATTTCCCGTCGAAATCACTGGCAGGCGTGGGTGTGGCGTTTTATCTGATGCTGGCGCTGCGCACCTTTTTGCGCG
ATCAGGGCTGGTTTGATGAGCGTGGCATCGCAATTCCTAACCTGGCAGAACTGCTGGATCTGGTCGCGCTGGGAACAGTGGCGGACGTCGTGCCGCTGGACGCTAATAATCGCATTCTGACCTGGCAGGGGATGAGTCGCATCCGTGCCGGAAAGTGCCGTCCAGGGATTAAAGCGCTGC
TGGAAGTGGCAAACCGTGATGCACAAAAACTCGCCGCCAGCGATTTAGGTTTTGCGCTGGGGCCACGTCTCAATGCTGCCGGACGACTGGACGATATGTCCGTCGGTGTGGCGCTCTTGCTGTGCGACAACATCGGCGAAGCGCGCGTGCTGGCAAATGAACTCGATGCGCTAAACCAGA
CGCGAAAAGAGATCGAACAAGGAATGCAAGTTGAAGCCCTGACCCTGTGCGAGAAACTGGAGCGAAGTCGCGACACGCTACCCGGCGGGCTGGCAATGTATCACCCCGAATGGCATCAGGGCGTTGTCGGTATTCTGGCTTCGCGCATCAAAGAGCGTTTTCACCGTCCGGTTATCGCCT
TTGCGCCAGCAGGTGATGGTACGCTGAAAGGTTCAGGTCGCTCCATTCAGGGGCTGCATATGCGTGATGCACTGGAGCGATTAGACACACTCTACCCTGGCATGATACTGAAGTTTGGCGGTCATGCGATGGCGGCGGGTTTGTCGCTGGAAGAGGATAAATTCGAACTCTTTCAACAACG
GTTTGGCGAGCTGGTTACCGAGTGGCTGGACCCTTCGCTATTGCAAGGCGAAGTGGTGTCAGACGGCCCGTTAAGCCCGGCCGAAATGACCATGGAAGTGGCGCAGCTGCTGCGCGATGCTGGCCCGTGGGGGCAGATGTTCCCGGAGCCGCTGTTTGATGGTCATTTCCGTCTGCTGCA
ACAGCGGCTGGTGGGCGAACGTCATTTGAAAGTCATGGTCGAACCGGTCGGCGGCGGTCCGCTGCTGGATGGTATTGCTTTTAATGTCGATACCGCCCTCTGGCCGGATAACGGCGTGCGCGAAGTGCAACTGGCTTACAAGCTCGATATCAACGAGTTTCGCGGCAACCGCAGCCTGCA
AATTATCATCGACAATATCTGGCCAATTGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCATGTCAATTTATCAAGAATTTGTTAATAAATATAGTTTAAGTAAAACTCTAAGATTTGAGTTAATCCCACAGGGTAAAACACTTGAAAACATAAAAGCAAGAGGTTTGATTTTAGAT
GATGAGAAAAGAGCTAAAGACTACAAAAAGGCTAAACAAATAATTGATAAATATCATCAGTTTTTTATAGAGGAGATATTAAGTTCGGTTTGTATTAGCGAAGATTTATTACAAAACTATTCTGATGTTTATTTTAAACTTAAAAAGAGTGATGATGATAATCTACAAAAAGATTTTAAAAGTGC
AAAAGATACGATAAAGAAACAAATATCTGAATATATAAAGGACTCAGAGAAATTTAAGAATTTGTTTAATCAAAACCTTATCGATGCTAAAAAAGGGCAAGAGTCAGATTTAATTCTATGGCTAAAGCAATCTAAGGATAATGGTATAGAACTATTTAAAGCCAATAGTGATATCACAGATAT
AGATGAGGCGTTAGAAATAATCAAATCTTTTAAAGGTTGGACAACTTATTTTAAGGGTTTTCATGAAAATAGAAAAAATGTTTATAGTAGCAATGATATTCCTACATCTATTATTTATAGGATAGTAGATGATAATTTGCCTAAATTTCTAGAAAATAAAGCTAAGTATGAGAGTTTAAAAGACAA
AGCTCCAGAAGCTATAAACTATGAACAAATTAAAAAAGATTTGGCAGAAGAGCTAACCTTTGATATTGACTACAAAACATCTGAAGTTAATCAAAGAGTTTTTTCACTTGATGAAGTTTTTGAGATAGCAAACTTTAATAATTATCTAAATCAAAGTGGTATTACTAAATTTAATACTATTATTGG
TGGTAAATTTGTAAATGGTGAAAATACAAAGAGAAAAGGTATAAATGAATATATAAATCTATACTCACAGCAAATAAATGATAAAACACTCAAAAAAATATAAAATGAGTGTTTTATTTAAGCAAATTTTAAGTGATACAGAATCTAAATCTTTTGTAATTGATAAGTTAGAAGATGATAGTGATG
TAGTTACAACGATGCAAAGTTTTTATGAGCAAATAGCAGCTTTTAAAACAGTAGAAGAAAAATCTATTAAAGAAACACTATCTTTATTATTTGATGATTTAAAAGCTCAAAAACTTGATTTGAGTAAAATTTATTTTAAAAATGATAAATCTCTTACTGATCTATCACAACAAGTTTTTGATGATTA
TAGTGTTATTGGTACAGCGGTACTAGAATATATAACTCAACAAATAGCACCTAAAAATCTTGATAACCCTAGTAAGAAAGAGCAAGAATTAATAGCCAAAAAAACTGAAAAAGCAAAATACTTATCTCTAGAAACTATAAAGCTTGCCTTAGAAGAATTTAATAAGCATAGAGATATAGATAA
ACAGTGTAGGTTTGAAGAAATACTTGCAAACTTTGCGGCTATTCCGATGATATTTGATGAAATAGCTCAAAACAAAGACAATTTGGCACAGATATCTATCAAATATCAAAATCAAGGTAAAAAAGACCTACTTCAAGCTAGTGCGGAAGATGATGTTAAAGCTATCAAGGATCTTTTAGATCA
AACTAATAATCTCTTACATAAACTAAAAAATATTTCATATTAGTCAGTCAGAAGATAAGGCAAATATTTTAGACAAGGATGAGCATTTTTATCTAGTATTTGAGGAGTGCTACTTTGAGCTAGCGAATATAGTGCCTCTTTATAACAAAATTAGAAACTATATAACTCAAAAGCCATATAGTGATG
AGAAATTTAAGCTCAATTTTGAGAACTCGACTTTGGCTAATGGTTGGGATAAAAATAAAGAGCCTGACAATACGGCAATTTTATTTATCAAAGATGATAAATATTATCTGGGTGTGATGAATAAGAAAAATAACAAAATATTTGATGATAAAGCTATCAAAGAAAATAAAGGCGAGGGTTATA
AAAAAATTGTTTATAAACTTTTACCTGGCGCAAATAAAATGTTACCTAAGGTTTTCTTTTCTGCTAAATCTATAAAATTTTATAATCCTAGTGAAGATATACTTAGAATAAGAAATCATTCCACACATACAAAAAATGGTAGTCCTCAAAAAGGATATGAAAAATTTGAGTTTAATATTGAAGATT
GCCGAAAATTTATAGATTTTTATAAACAGTCTATAAGTAAGCATCCGGAGTGGAAAGATTTTGGATTTAGATTTTCTGATACTCAAAGATATAATTCTATAGATGAATTTTATAGAGAAGTTGAAAATCAAGGCTACAAACTAACTTTTGAAAATATATCAGAGAGCTATATTGATAGCGTAGTTA
ATCAGGGTAAATTGTACCTATTCCAAATCTATAATAAAGATTTTTCAGCTTATAGCAAAGGGCGACCAAATCTACATACTTTATATTGGAAAGCGCTGTTTGATGAGAGAAATCTTCAAGATGTGGTTTATAAGCTAAATGGTGAGGCAGAGCTTTTTTATCGTAAACAATCAATACCTAAAAAA
ATCACTCACCCAGCTAAAGAGGCAATAGCTAATAAAAACAAAGATAATCCTAAAAAAGAGAGTGTTTTTGAATATGATTTAATCAAAGATAAACGCTTTACTGAAGATAAGTTTTTCTTTCACTGTCCTATTACAATCAATTTTAAATCTAGTGGAGCTAATAAGTTTAATGATGAAATCAATTTA
TTGCTAAAAGAAAAAGCAAATGATGTTCATATATTAAGTATAGATAGAGGTGAAAGACATTTAGCTTACTATACTTTGGTAGATGGTAAAGGCAATATCATCAAACAAGATACTTTCAACATCATTGGTAATGATAGAATGAAAACAAACTACCATGATAAGCTTGCTGCAATAGAGAAAGAT
AGGGATTCAGCTAGGAAAGACTGGAAAAAGATAAATAACATCAAAGAGATGAAAGAGGGCTATCTATCTCAGGTAGTTCATGAAATAGCTAAGCTAGTTATAGAGTATAATGCTATTGTGGTTTTTGAGGATTTAAATTTTGGATTTAAAAGAGGGCGTTTCAAGGTAGAGAAGCAGGTCTAT
CAAAAGTTAGAAAAAATGCTAATTGAGAAACTAAACTATCTAGTTTTCAAAGATAATGAGTTTGATAAAACTGGGGGAGTGCTTAGAGCTTATCAGCTAACAGCACCTTTTGAGACTTTTAAAAAGATGGGTAAACAAACAGGTATTATCTACTATGTACCAGCTGGTTTTACTTCAAAAATTT
GTCCTGTAACTGGTTTTGTAAATCAGTTATATCCTAAGTATGAAAGTGTCAGCAAATCTCAAGAGTTCTTTAGTAAGTTTGACAAGATTTGTTATAACCTTGATAAGGGCTATTTTGAGTTTAGTTTTGATTATAAAAACTTTGGTGACAAGGCTGCCAAAGGCAAGTGGACTATAGCTAGCTTTG
GGAGTAGATTGATTAACTTTAGAAATTCAGATAAAAATCATAATTGGGATACTCGAGAAGTTTATCCAACTAAAGAGTTGGAGAAATTGCTAAAAGATTATTCTATCGAATATGGGCATGGCGAATGTATCAAAGCAGCTATTTGCGGTGAGAGCGACAAAAAGTTTTTTGCTAAGCTAACTA
GTGTCCTAAATACTATCTTACAAATGCGTAACTCAAAAACAGGTACTGAGTTAGATTATCTAATTTCACCAGTAGCAGATGTAAATGGCAATTTCTTTGATTCGCGACAGGCGCCAAAAAATATGCCTCAAGATGCTGATGCCAATGGTGCTTATCATATTGGGCTAAAAGGTCTGATGCTACT
AGGTAGGATCAAAAATAATCAAGAGGGCAAAAAACTCAATTTGGTTATCAAAAATGAAGAGTATTTTGAGTTCGTGCAGAATAGGAATAACCAAGCGGCCGCACTCGAGAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGTCGACACCACCACCACCACCA
CTGAGATCCGGCTGCTAA

Amino acid sequence

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSVKQQIQLRRREVDETADLPAELPPLLRRLYASRGVRSAQELERSVKGMLPWQQLSGVEKAVEILYNAFREGTRIIVVGDFDADGATSTALSVLAMRSLGCSNIDYLVPNRFEDGYGLSPEVVDQAHARGAQLIVTVDNGISSHAGVEHARSL
GIPVIVTDHHLPGETLPAAEAIINPNLRDCNFPSKSLAGVGVAFYLMLALRTFLRDQGWFDERGIAIPNLAELLDLVALGTVADVVPLDANNRILTWQGMSRIRAGKCRPGIKALLEVANRDAQKLAASDLGFALGPRLNAAGRLDDMSVGVALLLCDNIGEARVLANELDALNQTRKEIEQGMQVEALT
LCEKLERSRDTLPGGLAMYHPEWHQGVVGILASRIKERFHRPVIAFAPAGDGTLKGSGRSIQGLHMRDALERLDTLYPGMILKFGGHAMAAGLSLEEDKFELFQQRFGELVTEWLDPSLLQGEVVSDGPLSPAEMTMEVAQLLRDAGPWGQMFPEPLFDGHFRLLQQRLVGERHLKVMVEPVGGGPLL
DGIAFNVDTALWPDNGVREVQLAYKLDINEFRGNRSLQIIDNIWPIGSEFELRRQACGRMSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDL
ILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVI
DKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNL
LHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPE
WKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLV
DGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYF
EFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNNQAAALEKRPAATKK
AGQAKKKKSTPPPPPLRSGC

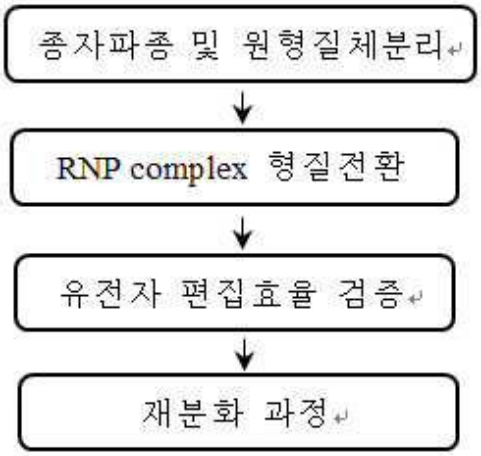
도면50

SSB-FnCpf1-RecJ

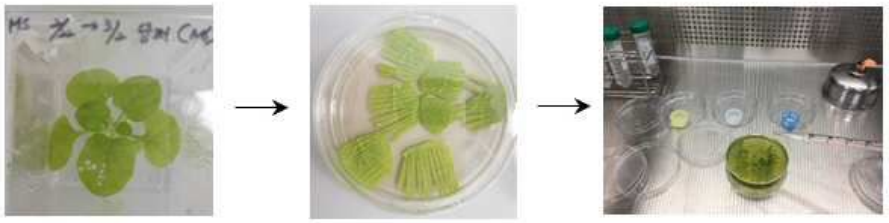
Nucleotide sequence

Amino acid sequence

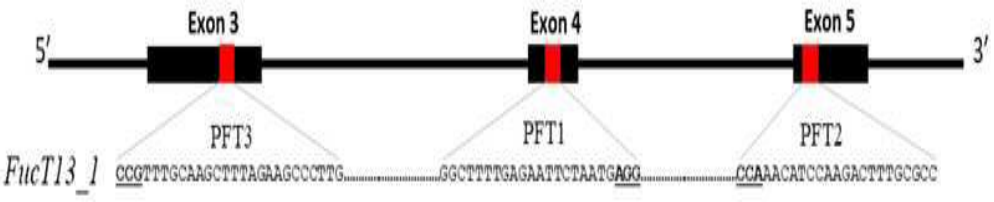
도면51



도면52



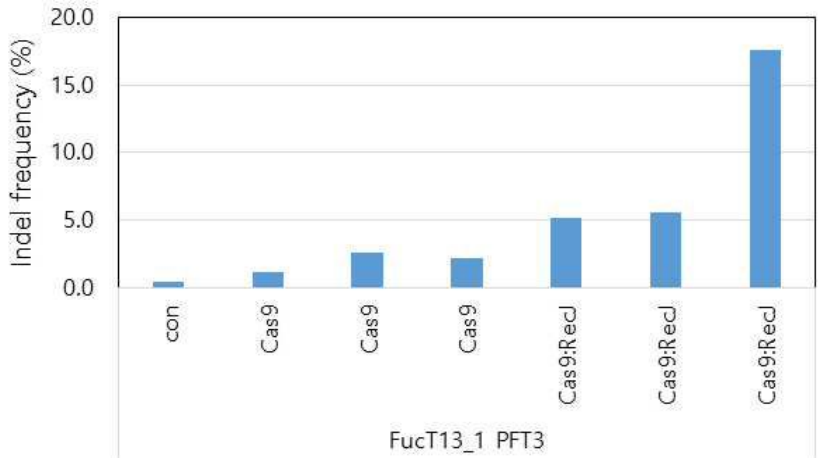
도면53



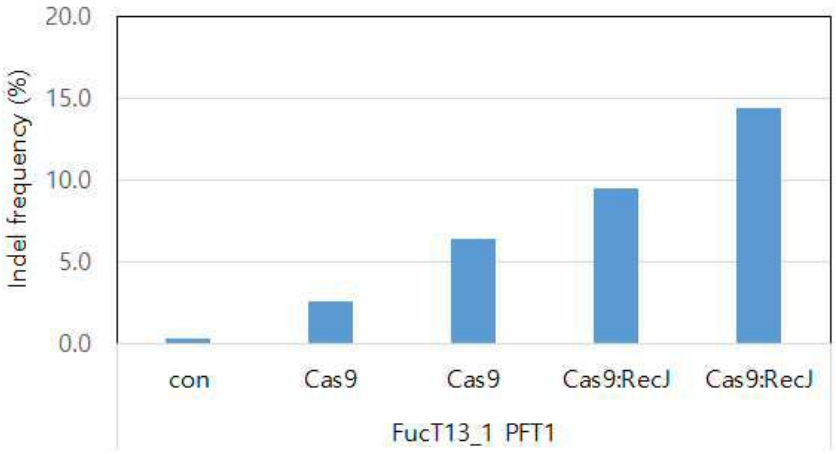
도면54

Conditions	Samples	Total Sequences	With both indicator sequences	More than minimum frequency	Insertions	Deletions	Indel frequency	Indel frequency (%)
13_1 PFT3	con	108344	107309	102874	3	449	452 (0.4%)	0.4
	Cas9	80766	80073	76594	279	574	853 (1.1%)	1.1
	Cas9	73419	72721	69201	879	924	1803 (2.6%)	2.6
	Cas9	95494	94584	90408	629	1353	1982 (2.2%)	2.2
	Cas9:RecJ	102906	101894	97056	2315	2601	4916 (5.1%)	5.1
	Cas9:RecJ	125519	124292	118522	3190	3504	6694 (5.6%)	5.6
	Cas9:RecJ	102660	101069	94971	7330	9381	16711 (17.6%)	17.6
13_1 PFT1	con	196230	144327	138573	7	472	479 (0.3%)	0.3
	Cas9	125076	79885	75976	822	1190	2012 (2.6%)	2.6
	Cas9	38597	21658	20137	711	571	1282 (6.4%)	6.4
	Cas9:RecJ	227861	160703	153332	7741	6873	14614 (9.5%)	9.5
	Cas9:RecJ	198048	152119	144225	11924	8799	20723 (14.4%)	14.4

도면55



도면56



서 열 목 록

<110> g-flas life sciences

<120> MODIFIED CRISPR ASSOCIATED PROTEIN COMPRISING CRISPR ASSOCIATED PROTEIN AND EXONUCLEASE AND USE THEREOF

<130> FPD201811-0042c

<160> 102

<170> KoPatentIn 3.0

<210> 1

<211> 577

<212> PRT

<213> Artificial Sequence

<220><223> Escherichia coli str. K-12 substr. MG1655

<400> 1

Val Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu Thr Ala
1 5 10 15

Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser
20 25 30

Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met
35 40 45

Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu
50 55 60

Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe
65 70 75 80

Asp Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg
85 90 95

Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg Phe Glu
100 105 110

Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg
115 120 125

Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala
130 135 140

Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp
145 150 155 160

His His Leu Pro Gly Asp Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn
165 170 175

Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val

180 185 190

Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp

195 200 205

Gln Gly Trp Phe Asp Glu Arg Asn Ile Ala Ile Pro Asn Leu Ala Glu

210 215 220

Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu

225 230 235 240

Asp Ala Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg

245 250 255

Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn

260 265 270

Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly

275 280 285

Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val

290 295 300

Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn

305 310 315 320

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

325 330 335

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

340 345 350

Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly

355 360 365

Val Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His Arg Pro

370 375 380

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

385 390 395 400

Arg Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg Leu Asp

405 410 415

Thr Leu Tyr Pro Gly Met Met Leu Lys Phe Gly Gly His Ala Met Ala

420 425 430

Ala Gly Leu Ser Leu Glu Glu Asp Lys Phe Lys Leu Phe Gln Gln Arg

435 440 445

Phe Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu Gln Gly

450 455 460

Glu Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr Met Glu
465 470 475 480

Val Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met Phe Pro
485 490 495

Glu Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg Leu Val
500 505 510

Gly Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly Gly Pro
515 520 525

Leu Leu Asp Gly Ile Ala Phe Asn Val Asp Thr Ala Leu Trp Pro Asp
530 535 540

Asn Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile Asn Glu
545 550 555 560

Phe Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile Trp Pro
565 570 575

Ile

<210> 2

<211> 1734

<212> DNA

<213> Artificial Sequence

<220><223> Escherichia coli str. K-12 substr. MG1655

<400> 2

gtgaaacaac agatacaact tcgtcgccgt gaagtcgatg aaacggcaga cttgcccgct 60

gaattgcctc ccttgctgcg ccgtttatac gccagccggg gagtacgcag tgcgcaagaa 120

ctggaacgca gtgttaaagg tatgctgccc tggcagcaac tgagcggcgt cgaaaaggcc 180

gttgagatcc tttacaacgc ttttcgcgaa ggaacgcgga ttattgtggt cggtgatttc 240

gacgccgacg gcgcgaccag cacggctcta agcgtgctgg cgatgcgctc gcttggttgc 300

agcaatatcg actacctggt accaaaccgt ttcgaagacg gttacggctt aagcccggaa 360

gtggtcgatc aggcccatgc ccgtggcgcg cagttaattg tcacggtgga taacggtatt 420

tcctcccatg cgggggttga gcacgctcgc tcgttgggca tcccggttat tgttaccgat 480

caccatttgc caggcgacac attacccgca gcggaagcga tcattaaccc taacttgcgc 540

gactgtaatt tcccgtcgaa atcactggca ggcgtgggtg tggcgtttta tctgatgctg 600

gcgctgcgca ccttttgcg cgatcagggc tggtttgatg agcgtaacat cgcaattcct 660

aacctggcag aactgctgga tctggtcgcg ctggggacag tggcggacgt cgtgccgctg 720

gacgctaata atcgcattct gacctggcag gggatgagtc gcatccgagc cggaaagtgc 780

cgtccgggga ttaaagcgct gcttgaagtg gcaaaccgtg atgcacaaaa actcgccgcc 840

agcgatttag gttttgcgct ggggccacgt ctcaatgctg ccggacgact ggacgatatg 900

tccgtcggtg tggcgctgtt gttgtgcgac aacatcggcg aagcgcgcgt gctggcaaat 960

gaactcgatg cgctaaacca gacgcgaaaa gagatcgaac aaggaatgca aattgaagcc 1020

ctgaccctgt gcgagaaact ggagcgcagc cgtgacacgc tacccggcgg gctggcaatg 1080

tatcaccccg aatggcatca gggcgttgtc ggtattctgg cttcgcgcat caaagagcgt 1140

tttcaccgtc cggttatcgc gtttgcgcca gcaggtgacg gtacgctgaa aggttccggt 1200

cgctccattc aggggctgca tatgcgtgat gcgctggagc gattagacac actctaccct 1260

ggcatgatgc tgaagtttgg cggtcatgcg atggcggcgg gtttgtcgct ggaagaggat 1320

aaattcaaac tctttcaaca acggtttggc gaactggtta ctgagtggct ggacccttcg 1380

ctattgcaag gcgaagtggt atcagacggt ccgttaagcc cggccgaaat gaccatggaa 1440

gtggcgcagc tgctgcgcga tgctggcccg tgggggcaga tgttcccgga gccgctgttt 1500

gacggtcatt tccgtctgct gcaacagcgg ctggtgggcg aacgtcattt gaaggtgatg 1560

gtcgaaccgg tcggcggcgg tccactgctg gatggtattg cttttaatgt cgataccgcc 1620

ctctggccgg ataacggcgt gcgcgaagtg caactggctt ataagctcga tatcaacgag 1680

tttcgcggca accgcagcct gcaaattatc atcgacaata tctggccaat ttag 1734

<210> 3

<211> 744

<212> PRT

<213> Artificial Sequence

<220><223> Streptococcus pneumoniae R6 chromosome

<400> 3

Val Asp Val Phe Leu Ile Thr Pro Thr Tyr Glu Trp Gln Phe Ala Leu

1 5 10 15

```
Gln Val Glu Asp Ala Asp Phe Thr Lys Ile Ala Lys Lys Ala Gly Leu
             20                  25                  30
Gly Pro Glu Val Ala Arg Leu Leu Phe Glu Arg Gly Ile Gln Asp Gln

         35                  40                  45
Glu Ser Leu Lys Lys Phe Leu Glu Pro Ser Leu Glu Asp Leu His Asp
     50                  55                  60
Ala Tyr Leu Leu His Asp Met Asp Lys Ala Val Glu Arg Ile Arg Gln
 65                  70                  75                  80
Ala Ile Glu Glu Gly Glu Asn Ile Leu Val Tyr Gly Asp Tyr Asp Ala
                 85                  90                  95
Asp Gly Met Thr Ser Ala Ser Ile Val Lys Glu Ser Leu Glu Gln Leu
            100                 105                 110

Gly Ala Glu Cys Arg Val Tyr Leu Pro Asn Arg Phe Thr Asp Gly Tyr
        115                 120                 125
Gly Pro Asn Ala Ser Val Tyr Lys Tyr Phe Ile Glu Gln Glu Gly Ile
    130                 135                 140
Ser Leu Ile Val Thr Val Asp Asn Gly Val Ala Gly His Glu Ala Ile
145                 150                 155                 160
Ala Leu Ala Gln Ser Met Gly Val Asp Val Ile Val Thr Asp His His
                165                 170                 175
Ser Met Pro Glu Thr Leu Pro Asp Ala Tyr Ala Ile Val His Pro Glu

            180                 185                 190
His Pro Asp Ala Asp Tyr Pro Phe Lys Tyr Leu Ala Gly Cys Gly Val
        195                 200                 205
Ala Phe Lys Leu Ala Cys Ala Leu Leu Glu Glu Val Gln Val Glu Leu
    210                 215                 220
Leu Asp Leu Val Ala Ile Gly Thr Ile Ala Asp Met Val Ser Leu Thr
225                 230                 235                 240
Asp Glu Asn Arg Ile Leu Val Gln Tyr Gly Leu Glu Met Leu Gly His
                245                 250                 255

Thr Gln Arg Ile Gly Leu Gln Glu Met Leu Asp Met Ala Gly Ile Ala
```

260 265 270

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

275 280 285

Leu Asn Ala Leu Gly Arg Leu Asp Asp Pro Asn Pro Ala Ile Asp Leu

290 295 300

Leu Thr Gly Phe Asp Asp Glu Glu Ala His Glu Ile Ala Leu Met Ile

305 310 315 320

His Gln Lys Asn Glu Glu Arg Lys Glu Ile Val Gln Ser Ile Tyr Glu

325 330 335

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

340 345 350

Lys Glu Gly Trp Asn Pro Gly Val Leu Gly Ile Val Ala Gly Arg Leu

355 360 365

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

370 375 380

Arg Ala Lys Gly Ser Ala Arg Ser Val Glu Ala Val Asp Ile Phe Glu

385 390 395 400

Ala Leu Asp Pro His Arg Asp Leu Phe Ile Ala Phe Gly Gly His Ala

405 410 415

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

420 425 430

Gln Val Leu Glu Asp Tyr Val Arg Glu Lys Gly Ala Asp Ala Gly Gly

435 440 445

Lys Asn Lys Leu Asn Leu Asp Glu Glu Leu Asp Leu Glu Ala Leu Ser

450 455 460

Leu Glu Thr Val Lys Ser Phe Glu Arg Leu Ala Pro Phe Gly Met Asp

465 470 475 480

Asn Gln Lys Pro Ile Phe Tyr Ile Lys Asn Phe Gln Val Glu Ser Ala

485 490 495

Arg Thr Met Gly Ala Gly Asn Ala His Leu Lys Leu Lys Ile Ser Lys

500 505 510

Gly Glu Ala Ser Phe Glu Val Val Ala Phe Gly Gln Gly Arg Trp Ala

515 520 525

Thr Glu Phe Ser Gln Thr Lys Asn Leu Glu Leu Ala Val Lys Leu Ser

530 535 540

Val Asn Gln Trp Asn Gly Gln Thr Ala Leu Gln Leu Met Met Val Asp

545 550 555 560

Ala Arg Val Glu Gly Val Gln Leu Phe Asn Ile Arg Gly Lys Asn Ala

565 570 575

Val Leu Pro Glu Gly Val Pro Val Leu Asp Phe Pro Gly Glu Leu Pro

580 585 590

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

595 600 605

Ile Thr Gln Leu Lys Thr Ile Phe Gln Glu Gln His Phe Ser Ala Val

610 615 620

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

625 630 635 640

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

645 650 655

Glu Phe Asp Ile Arg Tyr Lys Leu Lys Asp Leu Ala Ala Tyr Leu Asn

660 665 670

Ile Gln Gln Ile Leu Leu Val Lys Met Ile Gln Val Phe Glu Glu Leu

675 680 685

Gly Phe Val Thr Ile Lys Asp Gly Val Met Thr Val Asn Lys Glu Ala

690 695 700

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

705 710 715 720

Thr Val Lys Asp Gln Glu Met Met Ala Leu Gly Thr Val Gln Glu Ile

725 730 735

Tyr Asp Phe Leu Met Glu Lys Glu

740

<210> 4

<211> 2235

<212> DNA

<213> Artificial Sequence

<220><223> Streptococcus pneumoniae R6 chromosome

<400> 4

gtggatgtct ttttgataac acctacttat gaatggcagt ttgccctgca ggtagaagat 60

gcggatttta caaagatagc caagaaggct ggactgggtc ctgaggtggc tcggttattg 120

tttgagagag ggattcagga ccaagaaagt ctgaagaagt ttttagaacc ttccttggag 180

gacttacatg atgcttatct gctccatgat atggacaagg cagtggagcg gattcgtcag 240

gctattgaag aaggggaaaa tattctcgtt tatggagact atgatgcgga tggcatgact 300

tcggcttcta ttgtgaagga aagtttggaa caacttggtg ctgagtgccg agtttacctg 360

ccaaatcgtt ttaccgatgg ctatggccct aatgctagtg tttataaata ctttatcgag 420

caagaaggaa tttccttgat tgtgacggtg gacaatgggg ttgctggtca tgaggctatt 480

gcattggctc agtctatggg agtagatgtc attgtgacag accatcattc catgcctgaa 540

accctgccag atgcttatgc tattgtccat cctgaacatc cagatgcgga ttatcctttt 600

aaatatttgg ctggttgtgg agttgctttc aagttggctt gtgccctgtt agaagaagtg 660

caagtggaat tgcttgattt ggtcgctatt ggaactattg cagatatggt gagtctgacg 720

gatgaaaatc gtatcttagt tcaatatggt ctggaaatgt tgggtcatac ccagcgcatt 780

ggtctgcaag aaatgctgga catggctggg attgctgcca acgaagtaac agaagaaacg 840

gttggtttcc agattgctcc tcgtttgaat gccttgggtc gcttggatga tcccaatcct 900

gccattgatt tgttgactgg atttgatgat gaggaagcgc atgagattgc ccttatgatt 960

caccagaaaa acgaagagcg caaggaaatc gttcagtcta tctatgaaga agccaagacc 1020

atggtggatc ctgagaagaa ggttcaggtc ttggccaagg aaggctggaa tcctggggtt 1080

ctaggaatcg tggctggtcg tttattggaa gaattgggac agacagtcat tgttcttaat 1140

atagaagacg gtcgtgccaa gggcagtgct cgtagtgtgg aagcggtcga tatttttgaa 1200

gctctggatc cccatcgaga cctcttcatc gcctttggag gtcatgcagg tgcagcgggt 1260

atgacgctgg aagttgagca actctcagat ttatctcagg ttttggaaga ttatgttcgt 1320

gaaaaaggtg cagatgctgg tggcaagaat aagttaaacc tagatgaaga gttggatttg 1380

gaggcactta gcttggaaac ggtcaaaagt tttgaacgtt tagctccttt tggaatggat 1440

aatcagaaac ctatttttta tatcaagaat tttcaggtcg aaagtgctcg tactatgggg 1500

gcaggtaatg cccatctaaa gctgaaaatt tccaagggtg aggcgagttt tgaagtggta 1560

gcctttggtc aaggcagatg ggcgacagag ttttctcaaa ccaagaatct agagttagcg 1620

gttaaattgt ctgtcaacca atggaatggc caaactgctc tccagttgat gatggtggat 1680
gcgcgagtgg aaggtgttca acttttaac attcgtggaa aaaatgcagt cttgccagaa 1740
ggtgttccag tcttggattt tcctggagaa ctgccaaatc ttgcggctag tgaagctgtt 1800
gtcgtaaaaa acattccaga ggatattact cagctgaaga ccatttttca ggaacagcat 1860
ttctctgctg tctatttcaa aaatgatatt gacaaggctt attatctgac aggttatggg 1920
actagagatc agtttgccaa attgtacaag actatttacc agttcccaga gtttgatatt 1980
cgctacaagc tgaaagattt ggctgcatat cttaatattc aacaaatctt gctggtcaag 2040

atgattcaag tatttgaaga actaggcttt gtgacgataa aagatggtgt gatgacagtc 2100
aataaagagg cgccaaagcg ggagatagga gaaagtcaaa tttaccaaaa tctcaaacaa 2160
accgttaaag accaagaaat gatggcgctg ggtacggtgc aagaaattta tgatttttg 2220
atggaaaaag agtag 2235

<210> 5
<211> 742
<212> PRT
<213> Artificial Sequence
<220><223> Lactococcus lactis subsp. lactis Il1403
<400> 5

Met Ile Lys Ala Lys Tyr Asp Trp Lys Val Ala Asp Thr Ala Ile Ser
1 5 10 15
Glu Asp Phe Leu Lys Ile Ala Lys Lys His Lys Leu Asp Glu Leu Thr
20 25 30
Ser Arg Val Leu Tyr Gln Arg Gly Ile His Ala Glu Ala Glu Ile Glu
35 40 45
Gln Phe Leu Lys Pro Ser Leu Glu Asn Leu His Asp Pro Phe Leu Leu
50 55 60
His Asp Met Glu Lys Ala Thr Gly Arg Ile Leu Ser Ala Ile Glu Gln
65 70 75 80

Gly Glu Asn Ile Leu Ile Tyr Gly Asp Tyr Asp Ala Asp Gly Met Thr
85 90 95
Ala Ser Ser Val Met Lys Ser Ala Leu Asp Glu Leu Gly Ala Glu Ala

100 105 110

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

115 120 125

Asp Val Tyr Gln Tyr Tyr Ile Lys Asn Glu Asn Ile Asn Leu Ile Ile

130 135 140

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

145 150 155 160

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

165 170 175

Gln Leu Pro Pro Ala Tyr Ala Ile Val His Pro Glu His Pro Asp Ser

180 185 190

Gln Tyr Pro Phe Lys Tyr Leu Ala Gly Val Gly Val Ala Phe Lys Val

195 200 205

Ala Cys Ala Leu Leu Glu Tyr Ala Pro Ser Glu Met Leu Asp Leu Val

210 215 220

Ala Ile Gly Thr Ile Ala Asp Met Val Ser Leu Thr Asp Glu Asn Arg

225 230 235 240

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

245 250 255

Gly Leu Gln Glu Leu Met Lys Tyr Ala Gly Val Asp Phe Asp Lys Ile

260 265 270

Thr Glu Glu Thr Val Gly Phe Gln Ile Ala Pro Arg Leu Asn Ala Leu

275 280 285

Gly Arg Leu Asp Asp Pro Asn Pro Ala Ile Glu Leu Leu Thr Gly Trp

290 295 300

Asp Glu Asp Glu Ala His Glu Ile Ala Lys Met Ile Asp Gln Lys Asn

305 310 315 320

Ser Glu Arg Lys Glu Ile Val Glu Lys Ile His Asn Glu Ala Leu Ser

325 330 335

Met Leu Thr Asp Glu Pro Val Gln Ile Leu Tyr His Lys Asp Trp His

340 345 350

Lys Gly Val Leu Gly Ile Val Ala Gly Arg Leu Leu Glu Ala Ile His

355 360 365

Lys Pro Val Ile Met Leu Ala Gln Glu Asp Gly Ile Leu Arg Gly Ser
370 375 380
Ala Arg Ser Ile Glu Asn Phe Asp Ile Phe Lys Ala Leu Asn Ala His
385 390 395 400
Arg Glu Leu Phe Ile Ala Phe Gly Gly His Lys Gln Ala Ala Gly Met
405 410 415
Thr Leu Ser Leu Glu Asn Val Glu Ala Val Lys Lys Ala Met Ile Asp
420 425 430
Tyr Ile Val Asp Asn His Leu Asp Met Ser Lys Lys Ser Pro Leu Glu

435 440 445
Ile Ala Asp Arg Cys His Leu Asp Asp Ile Ser Leu Ser Thr Ile Thr
450 455 460
Asn Leu Ala Lys Leu Ala Pro Phe Gly Met Asp Asn Pro Lys Pro Arg
465 470 475 480
Phe Leu Ile Glu Asp Tyr Lys Val Ile Gln Ser Arg Ser Met Gly Lys
485 490 495
Asp Asn Ala His Leu Lys Leu Lys Ile Gln Glu Glu Lys Gln Gln Ile
500 505 510

Asp Ala Val Tyr Phe Gln His Gly Ser Glu Glu Leu Glu Phe Glu Gln
515 520 525
Ala Gln Thr Lys Leu Val Ala Thr Leu Ser Ser Asn Ser Trp Asn Gly
530 535 540
Asn Thr Ser Leu Gln Leu Met Ile Glu Asp Ala Asp Ser Val Gly Val
545 550 555 560
Glu Leu Leu Asp Ile Arg Ser Lys Gln Ile Pro Ile Pro Lys Glu Ala
565 570 575
Asn Ile Phe Ser Gln Asn Gln Leu Lys His Gly Ile Met Glu Asp Val

580 585 590
Leu Val Ile Glu Glu Ile Pro Gly Asp Leu Ala Gly Leu Ser Val Leu
595 600 605

Lys Glu Ala Val Ser Lys Ala Ala Ser Asp Asp Phe Lys Met Ile Tyr
610 615 620
Phe Lys Asn Lys Ile Thr Asp Ser Tyr Tyr Leu Thr Gly Ser Gly Thr
625 630 635 640
Arg Glu Glu Phe Ala Arg Leu Tyr Lys Ala Ile Tyr Gln Phe Pro Glu
645 650 655

Phe Asp Ile Arg Tyr Lys Leu Lys Ser Leu Ala Asp Tyr Leu Lys Ile
660 665 670
Pro His Leu Leu Leu Val Lys Met Ile Lys Ile Phe Glu Glu Leu Glu
675 680 685
Phe Val Ser Ile Asp Asn Gly Leu Met Thr Val Asn Lys Thr Ala Asp
690 695 700
Lys Arg Glu Ile Ser Glu Ser Thr Ile Tyr Gln Glu Leu Glu Lys Ile
705 710 715 720
Val Lys Met Gln Glu Leu Phe Ala Leu Ala Pro Val Lys Glu Ile Tyr

725 730 735
Gln Asn Leu Leu Glu Lys
740

<210> 6
<211> 2229
<212> DNA
<213> Artificial Sequence
<220><223> Lactococcus lactis subsp. lactis Il1403
<400> 6

atgataaaag caaaatatga ttggaaagtg gctgatacag ccatttccga agattttcta 60
aagatagcga aaaaacataa attagatgaa ttgacaagtc gcgtgcttta tcaaagagga 120
attcatgcag aagctgaaat tgagcaattt ttaaagccaa gtttagaaaa tttacatgac 180
ccttttctac ttcatgatat ggaaaaagca acagggcgta ttttatcagc aattgagcaa 240

ggtgaaaaca tcttaattta tggtgactat gatgctgacg gaatgactgc atcatcagtc 300
atgaaatccg ctctggatga attaggagct gaggctcaag tttatttacc caatcgtttt 360
actgacggtt acgggccaaa tcttgacgtt tatcagtact atattaagaa tgaaaatatc 420
aatcttatta ttactgtaga taatggggtt gctggcttgg aggcaattac ctgggcacag 480

gaaaatggag tggatgtaat tgttaccgac catcactcaa ttcctgacca acttcccct 540

gcttatgcca ttgtccaccc cgaacatcct gacagtcaat atcctttcaa atatttggca 600

ggagttggtg tcgccttcaa agttgcctgt gctcttttag aatatgcacc cagtgaaatg 660

cttgatttag ttgcaattgg tacaattgct gacatggtca gtttgactga tgaaaaccga 720

attttagtgg ctcatggtct aaaagttttg gcccaaacag agcgtgccgg cctacaagaa 780

ttaatgaagt atgccggggt tgattttgat aaaatcactg aggaaacggt tggttttcaa 840

attgctcccc gtttaaacgc tttgggacgt ctcgatgacc caaacccagc cattgaattg 900

ctgacaggtt gggatgaaga tgaagcgcat gaaattgcca aaatgattga ccaaaaaaat 960

tcagagcgaa aagaaattgt ggaaaaaatt cacaatgaag ctctgagtat gctgacagat 1020

gaacctgttc agattttata tcataaggac tggcataaag gtgtcttagg aattgttgcg 1080

ggtcgcttgc tagaagccat acataagcca gtaatcatgc tggcccaaga agatgggatt 1140

ttgcgtgggt ctgctcggtc aatcgaaaat tttgatattt tcaaggcatt aaatgcccat 1200

cgtgaacttt ttattgcttt tggtggtcat aaacaagccg cagggatgac tttaagtctg 1260

gaaaacgtag aagcagttaa aaaggcgatg attgattata tcgtggacaa tcatttagat 1320

atgtctaaaa aaagtccctt agaaattgct gaccgttgcc acttagacga tatttcgctc 1380

tcaacaatca ctaatttagc aaaattggct ccctttggca tggataatcc taagcctcga 1440

tttttaatcg aagattataa ggtgattcaa agccgaagta tggggaaaga caatgctcat 1500

cttaaattaa aaattcagga agaaaagcaa caaattgatg ccgtttattt ccaacatggt 1560

tcagaagaat tggaattcga acaggcacaa acaaaattag tcgcaacact ttcaagtaat 1620

agttggaatg gaaatacgag tcttcaatta atgattgaag atgctgattc agttggtgtc 1680

gaacttcttg atataagaag taaacaaatt cctatcccaa aagaggctaa tatttttct 1740

caaaatcagc taaaacatgg tataatggaa gatgtgcttg tcatcgaaga gatacctgga 1800

gatttagcag gtttgtcagt attaaaagaa gctgtcagta aagcagcgtc ggatgatttt 1860

aaaatgattt attttaaaaa taaaatcact gatagctatt atttaacagg aagtggaaca 1920

cgagaagaat ttgcaagact ttataaagca atttatcaat ttccagaatt tgatattcgc 1980

tataaattaa aatcacttgc ggattatttg aaaatccctc atctcttgct tgtcaagatg 2040

ataaaaatct ttgaggaact tgaatttgtc agcattgaca atggcttaat gaccgttaat 2100

aagacggccg ataaacgcga aatttctgag agtacaattt atcaagaact agaaaaaatt 2160

gtaaaaatgc aagaactttt tgctttagca cccgtcaaag aaatttatca aaatttatta 2220

gaaaaataa 2229

<210> 7

<211> 785

<212> PRT

<213> Unknown

<220><223> Bacillus clausii

<400> 7

Met Leu Gln Ser Lys Thr Arg Trp Arg Ile Leu Glu Gln Asn Glu Leu
1 5 10 15

Leu Ala Gln Glu Leu Ala Asp Glu Leu Arg Val Ser Met Leu Thr Ala
20 25 30

Arg Leu Leu Val Arg Arg Gly Ile Gln Thr Val Gln Ala Ala Lys Arg
35 40 45

Phe Leu His Tyr Glu Glu Pro Thr Phe Tyr Asp Pro Phe Leu Leu Lys
50 55 60

Gly Met Glu Glu Thr Ile Glu Arg Ile Ala Leu Ala Val Lys Arg Lys
65 70 75 80

Glu Arg Ile Leu Val Phe Gly Asp Tyr Asp Ala Asp Gly Val Ser Ser
85 90 95

Thr Ser Val Met Leu Thr Ala Leu Gln Thr Tyr Gly Ala Asp Cys Asp
100 105 110

Tyr Tyr Ile Pro Asn Arg Phe Thr Glu Gly Tyr Gly Pro Asn Cys Pro
115 120 125

Ala Leu Asp Phe Ala Lys Arg Gln Gly Tyr His Leu Val Ile Thr Val

130 135 140

Asp Thr Gly Ile Ser Ala Leu Asn Glu Ala Ala His Ala Lys Asp Ile
145 150 155 160

Gly Leu Asp Phe Ile Ile Thr Asp His His Glu Pro Pro Pro Val Leu
165 170 175

Pro Glu Ala Leu Ala Ile Ile Asn Pro Lys Gln Pro Gly Cys Pro Tyr
180 185 190

Pro Phe Lys Glu Leu Ala Gly Val Gly Val Ala Phe Lys Val Ala His
195 200 205

Ala Leu Leu Gly Arg Leu Pro Glu Glu Leu Leu Asp Tyr Ala Val Ile
210 215 220
Gly Thr Ile Ala Asp Leu Val Pro Leu Ile Asp Glu Asn Arg Leu Leu
225 230 235 240
Ala Lys Lys Gly Leu Arg Ala Ile Glu Ser Ser Gly Arg Pro Gly Ile
245 250 255
Arg Ala Leu Lys Glu Val Cys Gly Met Lys Gln Glu Ala Met Asp Ala
260 265 270
Asp His Ile Gly Phe Ala Ile Gly Pro Arg Leu Asn Ala Ala Gly Arg

275 280 285
Leu Asp Ser Ala Asn Pro Ala Val Glu Leu Leu Leu Ala Asp Asp Glu
290 295 300
Glu Glu Ala Lys Ala Leu Ala Thr Glu Ile Asp Ser Leu Asn Lys Glu
305 310 315 320
Arg Gln Ala Ile Val Ser Lys Met Thr Glu Glu Ala Ile Arg Leu Val
325 330 335
Glu Thr Thr Tyr Gly Thr Lys Ile Pro His Ala Ile Val Val Ala Lys
340 345 350

Glu Gly Trp Asn Pro Gly Val Ile Gly Ile Val Ala Ser Arg Leu Val
355 360 365
Glu Gln Phe Tyr Arg Pro Thr Ile Val Met Ser Ile Asp Glu Ser Ser
370 375 380
Gly Leu Ala Lys Gly Ser Ala Arg Ser Ile Glu Gly Phe Asp Met Tyr
385 390 395 400
Gln Glu Leu Ala Asn Asn Arg Asp Ile Leu Pro His Phe Gly Gly His
405 410 415
Pro Met Ala Ala Gly Met Thr Leu Lys Thr Glu Asp Ile Asp Asp Leu

420 425 430
Arg Ser Arg Leu Ile Lys Gln Ala Lys Glu Thr Leu Thr Asp Asp Met
435 440 445
Leu Thr Pro Ala Thr Asp Ile Asp Leu Val Ala Glu Val Glu Asp Val
450 455 460

Thr Val Gln Val Ile Gly Glu Leu Gln Ala Leu Ala Pro Phe Gly Val
465 470 475 480
Ala Asn Arg Lys Pro Ile Val Leu Val Glu Gly Ala His Ile Ser Asp
485 490 495

Met Arg Arg Ile Gly Ser Asn Gln Asn His Leu Lys Ile Gln Phe Thr
500 505 510
Gly Ala Gln Lys Pro Leu Asp Gly Ile Ala Phe Arg Met Gly His Leu
515 520 525
Phe Glu Glu Ile Thr Pro His Ala Lys Leu Ser Ala Ile Gly Thr Val
530 535 540
Ser Leu Asn Glu Trp Asn Gly Lys Val Lys Pro Gln Leu Ile Ile Asp
545 550 555 560
Asp Leu Ala Val Leu Glu Trp Gln Leu Phe Asp Trp Arg Ser Ile Gln
565 570 575
Pro Asn Arg Leu Asn Ser Arg Leu Leu Asp Leu Pro Arg Glu Lys Leu
580 585 590
Val Ala Ile Ser Phe Gln Glu Gly Thr Lys Glu Arg Leu Gly Leu Asp
595 600 605
Val Pro Val Tyr Asp Tyr Arg Gln Ala Pro Ser Phe Phe Glu Ala Tyr
610 615 620
Val Val Leu Leu Asp Leu Pro Ser Asn Arg Val Glu Leu Glu Ser Leu
625 630 635 640

Phe Ser Lys Lys Gly Glu Pro Ser Arg Val Tyr Val Val Phe Ser Glu
645 650 655
Glu Glu Glu Ser Phe Phe Gln Thr Asn Pro Asn Arg Glu Gln Phe Lys
660 665 670
Trp Tyr Tyr Gly Phe Ile Lys Lys Asn Gln Arg Phe Ser Leu Asn Gln
675 680 685
Leu Gly Ala Lys Leu Glu Lys His Lys Gly Trp Ser Ala Arg Thr Val
690 695 700
Glu Phe Met Thr Thr Val Phe Leu Glu Leu Gly Phe Ile Lys Leu Asp

705 710 715 720

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

725 730 735

Ala Ser Pro Thr Tyr Gln Ala Lys Gln Glu Lys Ala Trp Leu Glu Asn

740 745 750

Glu Phe Val Phe Ala Ser Tyr Gln Gln Leu Lys Glu Trp Phe Gln Ala

755 760 765

Ala Ile Glu Gly Ser Ala Glu Thr Lys Glu Glu Ser Ile Leu Asn Gly

770 775 780

Leu

785

<210> 8

<211> 2358

<212> DNA

<213> Unknown

<220><223> Bacillus clausii

<400> 8

atgttgcaat caaaaacaag atggcgcatt ttagaacaaa acgagctgtt ggcgcaagag 60

ctcgcagacg agctgcgtgt atcgatgtta acggcccgtc ttcttgtacg ccggggaatc 120

caaaccgtcc aggcggccaa gcgcttttta cactacgaag agccgacgtt ttacgatccc 180

tttttgttga aagggatgga ggaaacgatt gaacgaattg ccctcgcagt caaacgaaaa 240

gaacgcattc ttgtttttgg ggactatgat gctgatggcg ttagttcgac ttccgtgatg 300

ctaacagcgt tacaaacgta tggcgccgac tgtgattact atataccaaa ccgttttaca 360

gaaggatacg ggcccaattg cccagcgcta gactttgcaa aaaggcaagg ttatcactta 420

gtgataacag tcgatacagg catttctgct ttgaatgaag ctgcccatgc caaagacatc 480

ggtctggatt ttattatcac tgaccaccat gagccgccgc cggttttgcc agaagcattg 540

gcgatcatta atcctaagca gcctggttgc ccgtatccat ttaaagagct tgccggcgta 600

ggcgttgcct ttaaagttgc acacgctttg ctcgggcgcc tgcctgaaga gctgctcgat 660

tacgctgtga tcggcacgat cgccgacttg gtgcctctca ttgatgaaaa tcggctgctg 720

gccaaaaagg ggcttcgcgc gattgaatca agcggccgcc ctgggatccg tgccttaaag 780

gaagtatgcg gcatgaagca ggaagcaatg gatgccgacc atatcggctt tgcgattggc 840

cctcgtctaa atgcagcagg gcggctcgat tctgctaatc cagctgttga gctgctcctt 900

gctgatgatg aagaagaagc aaaagcgcta gccacagaaa tcgatagcct caataaagaa 960

cggcaagcaa ttgttagcaa aatgacagaa gaagcgattc gccttgtgga aacaacctat 1020

ggcacaaaga tcccccatgc catcgttgta gcgaaggaag gctggaatcc aggggtcatc 1080

gggattgtcg cttctcgcct tgtagagcag ttttaccgtc cgacaattgt catgagcatt 1140

gacgaatctt caggacttgc caaaggctcg gccagaagca tcgaagggtt tgacatgtac 1200

caagaattgg cgaataaccg ggatatactt cctcattttg gcggacaccc gatggcggca 1260

gggatgacac ttaaaaccga ggacattgac gacttacgtt ctcgtttaat caagcaagcc 1320

aaggagacac taacggacga catgcttacg ccagccacgg atatagacct cgttgcagaa 1380

gtggaggatg taaccgttca agtgattggc gagttgcagg cacttgcccc gtttggcgtt 1440

gccaaccgta aaccgattgt acttgttgag ggtgcgcata tttccgatat gcgccggatt 1500

ggcagcaacc aaaaccattt aaaaatccaa tttactggcg cccaaaagcc gcttgacgga 1560

attgccttta gaatggggca tttgtttgaa gagatcacgc cacatgcaaa gttatcggcc 1620

atcggaacgg tatcgttgaa tgaatggaac ggaaaagtaa agccacaact tattattgac 1680

gatcttgccg tattggagtg gcaactattt gattggcgca gcatccagcc aaaccgctta 1740

aacagccggt tgcttgactt gcctcgtgaa aaactcgtcg ctatttcctt tcaagaaggg 1800

acaaaagaaa gacttgggtt ggacgtaccc gtatatgatt accggcaagc gccttccttt 1860

tttgaagctt acgtcgtgtt gctcgatttg cctagcaacc gggttgaact ggaaagtctt 1920

ttttctaaga aaggggagcc aagccgtgtt tatgttgtgt tctctgagga agaagaatcg 1980

ttttttcaaa cgaacccaaa ccgcgagcag ttcaaatggt attatggttt tataaagaaa 2040

aaccaacgct tttcactgaa tcaactcggc gctaagcttg aaaagcacaa aggctggtca 2100

gcccgcacgg tcgaatttat gaccactgta tttttagaac ttggttttat caagcttgac 2160

gacagcatcg ttgaagcagt agaaaaccct gaaaaaaaag cgttgactgc atcgccgacc 2220

taccaagcca agcaagaaaa agcatggctt gagaatgaat ttgtttttgc ttcttatcaa 2280

cagttaaaag aatggttcca ggctgccatc gaagggagcg ctgaaacaaa agaggagtcg 2340

atcctaaatg gattataa 2358

<210> 9

<211> 590

<212> PRT

<213> Unknown

<220><223> Ehrlichia ruminantium strain Welgevonden

<400> 9

Leu Asn Ile Cys Ser Met Leu Asn Cys Asn Leu Glu Asn Ser Glu Tyr
1 5 10 15
Ile Gly Val Thr Gly Ala Leu Trp Lys Pro Tyr Asp Val Asn Leu Arg
20 25 30
Asp Ile Leu Thr Ile Lys Gln Lys Phe Tyr Leu Ser Glu Ile Val Ala
35 40 45
Arg Ile Leu Ser Ala Arg Lys Ile Asn Ile Glu Glu Ile Ser Asn Phe

50 55 60
Leu Tyr Pro Thr Leu Lys Ser Ser Leu Pro Asn Pro Phe His Met Leu
65 70 75 80
Asp Met Asp Lys Ala Val Tyr Arg Ile Cys His Ala Ile His Asn Arg
85 90 95
Glu Asn Ile Val Ile Phe Gly Asp Tyr Asp Val Asp Gly Ala Thr Ser
100 105 110
Ser Ala Leu Ile Lys Gln Tyr Leu Ser Gln Ile Gly Val Pro Thr Thr
115 120 125

Ile Tyr Ile Pro Asp Arg Ile Cys Glu Gly Tyr Gly Pro Asn Thr Gln
130 135 140
Ala Leu Leu Lys Leu Lys Glu Ile Gly Asn Ser Leu Cys Ile Thr Val
145 150 155 160
Asp Cys Gly Thr Ile Ala His Glu Pro Ile Ser Ala Ala Lys Ser Val
165 170 175
Asn Leu Asp Val Ile Val Ile Asp His His Ile Gly Ile Asn Thr Leu
180 185 190
Pro Asp Ala Val Ala Val Ile Asn Pro Asn Arg Leu Asp Glu Thr Ser

195 200 205
Pro Tyr Thr Tyr Leu Ala Gly Val Gly Val Ser Phe Leu Met Leu Val
210 215 220
Ala Leu His Lys Thr Leu Lys Glu Gln Gly Phe Phe Lys Asn Asn Gln
225 230 235 240
Glu Pro Asn Leu Ile Asn Tyr Leu Asp Leu Val Ala Leu Gly Thr Val
245 250 255

Cys Asp Val Met Pro Ile Ile Gly Leu Asn Arg Thr Phe Val Lys Gln
260 265 270

Gly Leu Lys Ile Met Thr Thr Arg Gln Asn Leu Gly Leu Lys Thr Leu
275 280 285
Ser Asp Val Ile Gly Leu Glu Glu Lys Pro Asn Ile Tyr Gln Leu Gly
290 295 300
Phe Asn Ile Gly Pro His Ile Asn Ala Gly Gly Arg Val Gly Asn Ala
305 310 315 320
Ser Leu Gly Ala Arg Leu Leu Ser Thr Asn Asn Glu Glu Glu Ala Leu
325 330 335
Glu Ile Ser Glu Lys Leu Gln Asn Phe Asn Leu Glu Arg Lys Thr Leu

340 345 350
Glu Asn Gln Ser Phe Asn Glu Ala Val Glu Gln Ile Glu Ala Thr Ile
355 360 365
Ile Ser Ser Asn Ile Ile Ile Ala Thr Gly Asn Trp His Pro Gly Ile
370 375 380
Ile Gly Ile Val Ala Gly Arg Leu Lys Asp Lys Phe Phe Leu Pro Ser
385 390 395 400
Ile Val Ile Ser Leu Glu Asn Gly Ile Gly Lys Ala Ser Ala Arg Ser
405 410 415

Ile Pro Asp Val Asp Leu Gly Ala Ala Ile Leu Glu Ala Lys Ala Leu
420 425 430
Gly Ile Ile Ile Glu Gly Gly Gly His Ala Met Ala Ala Gly Phe Ser
435 440 445
Ile Gln Glu Asn Lys Ile Lys Ile Leu His Glu Phe Leu Ser Gln Lys
450 455 460
Phe His Asn Ile Asn Thr Arg Lys Val Phe Lys Val Asp Gly Ile Val
465 470 475 480
Thr Ala Glu Ala Ile Asn Leu Lys Leu Trp Lys Glu Leu Gln Phe Leu

485 490 495
Glu Pro Phe Gly Ile Gly Asn Pro Glu Pro Arg Phe Ile Leu Thr Asn

500 505 510

Ile Arg Ile Lys Asn Ser Glu Val Ile Gly Glu Ser His Ile Arg Cys

515 520 525

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

530 535 540

Ile Gly Thr Glu Leu Gly Thr Thr Leu Leu Glu Cys Thr Thr Thr Thr

545 550 555 560

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

565 570 575

Gln Phe Ile Ile Glu Asp Ala Leu Gln His Asn Gln Ile Arg

580 585 590

<210> 10

<211> 1773

<212> DNA

<213> Unknown

<220><223> Ehrlichia ruminantium strain Welgevonden

<400> 10

ttgaatattt gtagcatgtt aaactgtaat ctagaaaact cggaatatat cggagtaaca 60

ggagcattat ggaaaccata cgatgtaaat ctgagagata ttttaacaat caagcaaaaa 120

ttttaccttt ctgaaattgt tgcaagaata ttatctgcaa gaaaaattaa tatagaagaa 180

attagtaatt ttttatatcc aactctaaaa tcatctctac ctaatccttt tcacatgtta 240

gatatggata aagcagtata cagaatttgt catgcaatac ataatcgaga aaacatagta 300

atctttggag actatgatgt agatggagca acatcatctg cattaataaa acaatatcta 360

tcacagatag gagtaccaac aacgatatat attccagatc gtatatgtga aggatatgga 420

ccaaatacac aggccttact aaaattaaaa gaaataggaa atagtctatg tataaccgta 480

gactgcggta ctatagcaca tgaaccaata tccgcagcaa agtcagtaaa tcttgatgtt 540

atagttatag atcatcatat aggaattaat actttaccag atgcagttgc agtaatcaat 600

cctaatcgct tagacgaaac atccccctac acatatctag caggagtagg cgtatctttt 660

ttaatgttag tagcactaca taaaacttta aaagagcaag gatttttaa aaataatcaa 720

gaaccaaatt taataaatta tttagattta gtagcactag gaactgtatg tgatgttatg 780

cctattatag gactaaatag gacttttgta aaacaaggtc taaagattat gacaacaaga 840

caaaatctag gccttaaaac attatctgat gtaattggat tagaggaaaa gccaaatatt 900

tatcaactgg gatttaatat aggaccacat ataaatgcag gaggaagagt aggtaatgct 960

agcctaggag caagattatt atcaaccaac aatgaagaag aagcattaga aatttcagaa 1020

aaactgcaaa atttcaattt agaaagaaaa acattagaaa atcaaagctt caatgaagca 1080

gtagaacaaa tagaagctac tataatatca agcaatataa ttattgctac aggaaattgg 1140

catccaggaa taataggtat tgttgcagga agattaaaag ataaattttt cttaccaagt 1200

atagtaatat cacttgaaaa tggtattggt aaagcaagtg ctaggtcaat tccagatgta 1260

gatttaggcg cagctatttt agaagcaaaa gccctcggaa taataataga aggaggagga 1320

catgctatgg ctgcaggatt ttcaatacaa gaaaataaaa tcaaaatatt acacgagttt 1380

ctatcacaaa agttccataa tatcaataca cgaaaagtat ttaaagtaga cggtattgtc 1440

actgcagaag ctatcaattt aaaattatgg aaagaattac aatttctaga accatttggc 1500

ataggtaatc ctgaaccaag attcattctc acaaatatca gaataaaaaa ttcagaagtt 1560

ataggagaaa gtcatatcag atgtttaatt tacgataaca aaacatttat aaaaggaatt 1620

tgttttagat gtatcggcac agaattgggt accactttac tagaatgtac tactaccaca 1680

ttactaggaa aaataagtat aaattactgg agaggtaatg agaatataca atttataata 1740

gaagatgctt tacagcataa ccaaataagg taa 1773

<210> 11

<211> 589

<212> PRT

<213> Unknown

<220><223> Chlamydia psittaci 6BC

<400> 11

Met Ala Ser Lys Asp Asn Ser Ser Val Ser Asn Pro Thr Trp Ile Tyr
1 5 10 15

Pro Lys Tyr Asp Pro Ala Leu Leu Ser Ser Ile Ile Lys Glu Leu His
20 25 30

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

Val Asp Glu Val Arg Asp Phe Leu Tyr Val His Leu Ser Asn Leu His

50 55 60

Asp Pro Glu Leu Leu Leu Asp Met Ser Lys Ala Val Gln Arg Leu Leu
65 70 75 80

Leu Ala Lys Glu Arg Gly Glu His Val Met Val Tyr Gly Asp Ser Asp
85 90 95
Val Asp Gly Ile Thr Gly Val Ala Leu Leu Val Glu Phe Leu Arg Ser
100 105 110
Ile Glu Met Lys Val Ser Tyr Cys Phe Leu Gly Ala Phe Leu Lys His
115 120 125

Tyr Gly Glu Pro Ser Leu Leu Ile Ala Lys Met Lys Glu Glu Gly Val
130 135 140
Thr Leu Leu Ile Thr Val Asp Cys Gly Ile Thr Ala Gly Lys Glu Val
145 150 155 160
Ser Asp Ile Asn Lys Gln Gly Ile Asp Val Ile Val Thr Asp His His
165 170 175
Met Pro Thr Gly Lys Ile Pro His Cys Ile Ala Thr Leu Asn Pro Lys
180 185 190
Leu Arg Asp His Thr Tyr Pro Asn Lys Asp Leu Thr Gly Val Gly Val

195 200 205
Ala Phe Lys Leu Ala Arg Gly Val Val Asn Ala Leu Gln Lys Asn Asn
210 215 220
Pro Lys Leu Lys Leu Asp Ile Lys His Leu Leu Asp Leu Val Thr Leu
225 230 235 240
Gly Thr Val Thr Asp Val Gly Thr Leu Leu Gly Glu Asn Arg Thr Met
245 250 255
Val Arg His Gly Ile Lys Glu Ile Ala Lys Gly Ser Arg Leu Gly Leu
260 265 270

Arg Lys Leu Cys Ile Phe Ser Gly Val Lys Pro Ser Glu Val Thr Ser
275 280 285
Thr Asp Ile Val Leu Lys Ile Ser Pro Lys Leu Asn Ser Leu Gly Arg
290 295 300
Leu Ala Asp Ala Ser Lys Gly Val Glu Leu Leu Leu Thr Lys Asp Pro
305 310 315 320
Glu Val Ala Asp Asp Leu Ile Gln Tyr Leu Asp Lys Ile Asn Arg Glu
325 330 335

Arg Gln Lys Ile Glu Ala Asp Val Phe His Asp Val Gln Lys Ile Leu

340 345 350

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

355 360 365

Gln Asp Trp His Ser Arg Val Ile Pro Ile Ile Ser Ala Arg Leu Ala

370 375 380

Lys Ala Tyr Asn Lys Pro Val Ala Ile Ile Ser Asn Gln Gly Gly Ile

385 390 395 400

Gly Lys Gly Ser Leu Arg Thr Ile Gly Ser Phe Pro Leu Leu Gly Ile

405 410 415

Leu Gln Lys Cys Ser Pro Met Phe Ile Ser Tyr Gly Gly His Asp Phe

420 425 430

Ala Ala Gly Ile Ile Ile Asn Glu Asp Arg Ile Glu Ala Phe Arg Lys

435 440 445

Lys Phe Ile His Leu Val Asn Ser Ser Leu Lys Lys Glu Lys Ala Val

450 455 460

Val Thr Leu Pro Leu Asp Ala Arg Ala Asp Phe Asp Glu Ile Asp His

465 470 475 480

Asp Leu Leu Ser Ser Ile Asp Leu Phe Glu Pro Phe Gly Lys Gly Asn

485 490 495

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

500 505 510

Leu Leu Pro Gly Asn His Leu Lys Leu Tyr Leu Asn Tyr Gly Glu Arg

515 520 525

Asn Leu Glu Gly Ile Ala Phe Gly Leu Gly Asp Arg Ile Gly Ala Leu

530 535 540

Lys Ala Ser Trp Asn Gln Pro Leu Glu Leu Ala Tyr Thr Pro Arg Leu

545 550 555 560

Ser Gln Ser Ala Asn Gly Gly Val Ile His Leu Leu Val Arg Asp Phe

565 570 575

Arg Ile Leu Pro Leu Asn Tyr Lys Asp Thr Thr Ala Arg

580 585

<210> 12

<211> 1773

<212> DNA

<213> Unknown

<220><223> Chlamydia psittaci 6BC

<400> 12

atggcaagta aagataattc ttctgtatct aatcctactt ggatatatcc taagtacgat 60

cccgctctac tttcttctat tataaaggag ttacatcttc atccggttgc tgcacagact 120

tttatttctc gaggatttca aacggtagat gaagttcgtg attttcttta tgtacatcta 180

tccaacctgc atgatcccga acttttactc gacatgtcaa aagctgtaca acgtttgctt 240

ctcgcaaaag aacgtggtga gcatgtcatg gtctatggag atagtgatgt tgatgggata 300

acaggcgtgg ctcttcttgt agagtttttg agatctatag aaatgaaggt tagctattgt 360

tttttaggag cattcctgaa acattatggc gaaccttctc tgttgattgc caagatgaag 420

gaagaaggcg tcactttgct gattaccgta gattgcggga ttactgcagg gaaagaagtc 480

agtgatatca ataagcaagg cattgatgtc attgttacag atcatcatat gcctacaggt 540

aaaatccctc attgcatagc aacattaaat ccaaagctca gagatcatac ttatccaaat 600

aaagatctca ctggtgtagg cgtcgctttt aaactcgccc gtggtgttgt taacgcgtta 660

cagaaaaaca atccaaaact caaattagac attaagcatt tattagattt agtaactcta 720

gggacagtga cggatgtggg aacccttta ggagagaatc gcaccatggt tcgccacggc 780

attaaagaaa tcgccaaagg gtcacggtta gggttgcgta aattatgcat tttttcaggg 840

gtaaaacctt cagaggttac ctctacagat attgttttga aaatttctcc aaaactgaat 900

agtttaggaa gacttgccga cgcttctaaa ggcgttgaat tgttgctgac caaagatcca 960

gaagttgctg atgacctcat tcaatatctc gataagatca atagagaacg ccaaaaaata 1020

gaggctgacg tattccatga tgtacaaaaa atcttaaaaa atcagcccga tatcgttaaa 1080

caagcagcta ttgtactatc gtcacaagat tggcattcca gagtcattcc tattatttca 1140

gcacgcctag cgaaagctta taataagccc gtagcgatta tttctaatca gggagggata 1200

gggaaaggat cgttaagaac aatagggtct ttccctctcc ttggaatatt acaaaagtgc 1260

tctcctatgt tcatatccta tggaggacat gattttgcag caggcattat cattaatgaa 1320

gaccgaatag aagcttttag gaaaaagttc attcatcttg tgaattcatc gttaaaaaag 1380

gaaaaagctg tagttaccct tcccttagat gctcgagctg attttgatga gatagatcac 1440

gatttactct cttccatcga tctgttcgaa ccttttggca aaggtaatcc tgtacctatc 1500

ttttatacca tagtacatca agtacgttat ccgaagttat taccagggaa tcatctgaaa 1560

ctctatctta attacggaga aagaaactta gaaggcatag cttttggatt aggggataga 1620

ataggagctc taaaagcgag ttggaatcaa cctttagaac tggcttatac accacgttta 1680

tcccaatctg ctaacggagg agtcattcat ttgttagtgc gtgattttcg tattcttcca 1740

ctaaattaca aggatacaac agcaaggttt taa 1773

<210> 13

<211> 580

<212> PRT

<213> Unknown

<220><223> Finegoldia magna ATCC 29328

<400> 13

Met Asn Lys Trp Leu Ile Asn Asn Arg Gly Asn Asn Tyr Asn Glu Ile
1 5 10 15

Ser Lys Lys Tyr Asn Ile His Pro Leu Ile Ala Lys Ile Leu Leu Asn
20 25 30

Arg Asn Ile Thr Asp Phe Glu Lys Phe Leu Asn Pro Asn Ala Glu Asn
35 40 45

Ser Tyr His Asn Pro Phe Leu Met Lys Asp Met Asp Lys Ala Val Asp

50 55 60

Ile Ile Ile Ser Thr Ile Asp His Asn Glu Lys Ile Arg Ile Val Gly
65 70 75 80

Asp Tyr Asp Gln Asp Gly Asn Ser Ser Thr Met Thr Leu Leu Asp Gly
85 90 95

Leu Gly Tyr Phe Thr Asp Lys Ile Ser Tyr Asp Ile Pro Asn Arg Met
100 105 110

Thr Asp Gly Tyr Gly Ile Ser Phe Ser Ile Ile Asp Lys Cys Ile Glu
115 120 125

Asp Asn Ile Asp Leu Ile Ile Thr Cys Asp Asn Gly Ile Ser Ala Ile
130 135 140

Glu Gln Cys Asp Tyr Ala Arg Lys Asn Gly Ile Lys Ile Ile Val Thr
145 150 155 160

Asp His His Gln Thr Ile Lys His Asp Gly Lys Glu Ile Ile Pro Asn
165 170 175
Ala Asp Ala Val Val Asn Pro Gln Gln Gln Ser Cys Lys Tyr Pro Phe
180 185 190
Lys Ser Leu Cys Gly Ala Gly Val Ala Tyr Lys Leu Ile Gln Ala Ile

195 200 205
Asn Ile Lys Lys Gly Tyr Gly Met Leu Gln Cys Glu Asn Leu Leu Gln
210 215 220
Tyr Val Ala Met Gly Thr Val Cys Asp Ile Val Asp Leu Lys Asp Glu
225 230 235 240
Asn Arg Tyr Phe Val Thr Lys Gly Leu Glu Glu Ile Asn Asn Thr Asp
245 250 255
Asn Tyr Gly Leu Lys Cys Leu Ile Asn Met Thr Gly Ile Lys Asn Ala
260 265 270

Val Asn Val Tyr Ser Leu Gly Phe Ile Ile Gly Pro Cys Ile Asn Ala
275 280 285
Ala Gly Arg Leu Asp Thr Ala Lys Leu Gly Val Glu Leu Phe Arg Asp
290 295 300
Glu Asn Met Asp Asn Val Glu Ala Tyr Ala Lys Ile Leu Val Asp Leu
305 310 315 320
Asn Glu Lys Arg Lys Lys Leu Thr Glu Asp Gly Phe Asn Lys Ala Val
325 330 335
Glu Ile Ile Glu Asn Thr Ala Leu Lys Asn Asp Asp Ile Leu Ile Cys

340 345 350
Asn Val Glu Gly Ile His Glu Ser Val Cys Gly Ile Ile Ala Gly Arg
355 360 365
Ile Lys Glu Lys Tyr Asn Lys Pro Thr Leu Ile Leu Thr Glu Ser Ala
370 375 380
Asn Lys Asn Leu Leu Lys Gly Ser Gly Arg Ser Ile Ser Glu Tyr Asn
385 390 395 400
Ile Phe Glu Glu Phe Asp Glu Phe Arg Glu Met Phe Val Ser Phe Gly

405 410 415

Gly His Pro Met Ala Cys Gly Leu Ser Ile Glu Lys Asn Lys Leu Asp
420 425 430
Glu Phe Arg Thr Lys Val Asn Lys Asn Ser Lys Leu Thr Glu Glu Asp
435 440 445
Phe Val Lys Lys Ile Leu Ile Asp Ser Ser Phe Tyr Ile Asp Lys Ile
450 455 460
Asp Phe Asp Leu Ile Glu Glu Ile Gly Asn Leu Arg Pro Phe Gly Lys
465 470 475 480
Asp Asn Pro Arg Pro Ile Leu Gly Asp Lys Asp Val Glu Ile Ile Phe

485 490 495
Ala Lys Met Ile Gly Lys Asn Gln Asn Val Leu Lys Met Lys Leu Leu
500 505 510
Lys Asn Asn Lys Ala Ile Asp Ala Ile Leu Phe Thr Asp Ala Ile Glu
515 520 525
Lys Tyr Thr Tyr Leu Leu Gly Lys Phe Gly Glu Asn Val Leu Lys Glu
530 535 540
Leu Glu Asn Ser Ile Ser Cys Asn Ala Tyr Ile Asp Ile Leu Tyr Tyr
545 550 555 560

Pro Glu Ile Asn Asp Phe Asn Asn Val Lys Asn Ile Gln Leu Asn Leu
565 570 575
Ile Asp Leu Arg
580

<210> 14
<211> 1743
<212> DNA
<213> Unknown
<220><223> Finegoldia magna ATCC 29328
<400> 14

atgaataaat ggttaattaa taatagagga aataattata acgaaatatc aaaaaaatat 60
aatattcatc cattaatagc aaaaattttg ctaaatagaa atatcactga ttttgaaaaa 120

tttttaaatc ctaatgcaga aaattcatat cataatcctt ttttaatgaa agatatggat 180

aaagcagtag atataattat tagcactata gatcataatg aaaaaataag aatagtagga 240

gattatgatc aagatggaaa ttcttctact atgactctac tagatggact tggatatttt 300

acagataaaa tttcctatga tattcctaat agaatgacag atggatatgg catttctttt 360

agtataatag acaaatgtat cgaagataat atagatttaa taataacttg tgacaatggt 420

attagtgcaa ttgaacagtg tgattatgca aggaaaaatg gaattaaaat aattgtaact 480

gatcatcacc aaacaataaa acatgatgga aaagaaataa ttccaaatgc tgatgcggtc 540

gttaacccac aacaacaatc ttgcaaatat ccttttaaaa gtttatgcgg tgcaggagtt 600

gcttataaat taatacaagc tattaatata aaaaaaggat atggaatgct acagtgtgaa 660

aatttgcttc aatatgtagc aatgggaaca gtatgtgaca tagtcgattt aaaagatgag 720

aataggtatt ttgttacaaa aggtctagaa gaaatcaata atacagataa ttacggttta 780

aaatgcttga tcaacatgac tggaattaag aatgcagtaa atgtctacag tttaggattt 840

ataattggtc cttgtatcaa tgcagcagga agattggata ctgcaaaatt aggtgttgaa 900

ctatttagag atgaaaatat ggataatgtt gaagcatatg caaaaatttt agtcgatctt 960

aatgaaaaaa gaaaaaaact aacagaagat ggatttaata aagctgtcga aattatcgaa 1020

aataccgcac taaaaaacga cgatatttta atttgcaatg tggaaggaat tcatgaaagt 1080

gtttgtggaa tcatagctgg aagaataaaa gaaaaatata ataaaccaac tttaatctta 1140

acggaatctg caaataaaaa tttactcaaa ggatctggaa gaagtatatc agaatataat 1200

atttttgagg agtttgatga gtttagagag atgttcgtat cttttggagg ccatcctatg 1260

gcatgtggat tatcaattga aaaaaataaa ttagatgagt ttagaactaa agttaataaa 1320

aatagcaaac ttacagaaga agattttgtt aaaaaaatct tgatagactc tagcttttat 1380

attgataaaa ttgattttga tttgatagaa gaaattggta atttgcgtcc atttggaaaa 1440

gataatccta ggccaatttt aggagataag gatgttgaaa ttatatttgc caaaatgata 1500

gggaagaatc aaaatgtatt aaaaatgaaa ttactaaaaa acaacaaagc aatagatgca 1560

attttgttta cggatgctat agaaaaatat acctatttat taggaaaatt tggtgaaaat 1620

gtattaaagg aacttgaaaa tagtatttct tgtaatgcat atatagatat tttatactat 1680

ccagaaatta atgattttaa taatgtcaaa aatattcaat taaatttaat tgatttaaga 1740

taa 1743

<210> 15

<211> 739

<212> PRT

<213> Unknown

<220><223> Streptococcus cristatus AS 1.3089

<400> 15

Met Ile Ser Ser Lys Tyr Asp Trp Gln Phe Ala Thr Asn Phe Thr Asp
1 5 10 15

Glu Lys Phe Leu Lys Lys Ala Lys Lys Ala Gly Leu Glu Pro Ala Ala
20 25 30

Ala Ser Leu Leu Tyr Gln Arg Gly Val Gln Thr Glu Glu Ala Leu Gln
35 40 45

Glu Phe Leu Glu Pro Ser Leu Asp Gln Leu His Asp Pro Tyr Asp Leu

50 55 60

His Asp Met Glu Arg Ala Val Glu Arg Ile Arg Ala Ala Ile Glu Asn
65 70 75 80

Tyr Glu Gln Ile Leu Ile Tyr Gly Asp Tyr Asp Ala Asp Gly Met Thr
85 90 95

Ser Ala Ser Ile Val Lys Glu Ala Leu Glu Gln Leu Gly Ala Glu Cys
100 105 110

Gln Val Tyr Leu Pro Asn Arg Phe Thr Asp Gly Tyr Gly Pro Asn Ser
115 120 125

Ser Val Tyr Lys Tyr Phe Ile Glu Asn Gln Gly Ile Ser Leu Ile Ile
130 135 140

Thr Val Asp Asn Gly Val Ala Gly Leu Glu Ala Ile Glu Leu Ala Gln
145 150 155 160

Ser Leu Gly Val Asp Val Ile Val Thr Asp His His Ser Met Pro Glu
165 170 175

Glu Leu Pro Asn Ala Tyr Ala Ile Val His Pro Glu His Ser Gly Ala
180 185 190

Asp Tyr Pro Phe Lys His Leu Ala Gly Cys Gly Val Ala Phe Lys Leu

195 200 205

Ala Thr Ala Leu Leu Glu Glu Val Gln Val Glu Leu Leu Asp Leu Val

210 215 220

Ala Ile Gly Thr Ile Ala Asp Met Val Ser Leu Thr Gly Glu Asn Arg

225 230 235 240

Ile Leu Val Lys Tyr Gly Leu Ser Val Leu Lys Asn Thr Gln Arg Val

245 250 255

Gly Leu Gln Glu Leu Phe Lys Ile Ala Gly Ile Gln Pro Asp Glu Leu

260 265 270

Asp Glu Glu Thr Val Gly Phe Gln Leu Ala Pro Arg Leu Asn Ala Leu

275 280 285

Gly Arg Leu Asp Asp Pro Asn Pro Ala Ile Glu Leu Leu Thr Gly Phe

290 295 300

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

305 310 315 320

Asp Glu Arg Lys Glu Ile Val Gln Gln Ile Tyr Lys Glu Ala Gln Thr

325 330 335

Met Leu Asp Pro Lys Arg Pro Val Gln Val Leu Ala Lys Glu Gly Trp

340 345 350

Asn Pro Gly Val Leu Gly Ile Val Ala Gly Arg Leu Leu Glu Glu Leu

355 360 365

His Gln Pro Val Ile Val Leu Asn Ile Glu Asp Gly Ile Ala Lys Gly

370 375 380

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

385 390 395 400

His Arg Asp Leu Phe Leu Ala Phe Gly Gly His Ala Gly Ala Ala Gly

405 410 415

Met Thr Leu Glu Ala Asp Lys Leu Ala Glu Leu Ala Asp Ile Leu Thr

420 425 430

Glu Tyr Ile Leu Glu Asn Asp Leu Asp Leu Thr Gly Lys Thr Ala Leu

435 440 445

Tyr Leu Asp Glu Glu Leu His Leu Pro Glu Leu Thr Leu Asp Thr Leu

450 455 460

Lys Ser Phe Glu Lys Leu Ala Pro Phe Gly Met Asp Asn Lys Lys Pro

465 470 475 480

Leu Phe Tyr Leu Lys Asp Phe Lys Val Asp Asn Ala Arg Thr Met Gly

485 490 495

Ala Gly Asn Ser His Leu Lys Leu Lys Ile Ser Gln Ala Asp Ala Ser

500 505 510

Phe Glu Val Val Ala Phe Gly Gln Gly Ser Leu Ala Thr Glu Phe Ala

515 520 525

Gln Thr Lys Asn Leu Glu Leu Ala Val Ser Leu Ser Val Asn Lys Trp

530 535 540

Asn Gly Gln Thr Ser Leu Gln Leu Met Leu Val Asp Ala Arg Val Asp

545 550 555 560

Gly Val Gln Leu Phe Asn Ile Arg Gly Lys Asn Ala Thr Leu Pro Asp

565 570 575

Lys Val Pro Val Leu Arg Phe Thr Glu Glu Leu Pro Asp Leu Thr Asn

580 585 590

Ser Arg Ala Val Val Val Tyr Asp Leu Pro Asp Asn Leu Gln Val Leu

595 600 605

Lys Lys Ile Leu Gln Tyr Gln Asp Phe Glu Ala Ile Tyr Phe Lys Asn

610 615 620

Glu Ile Ala Lys Pro Tyr Tyr Leu Thr Gly Tyr Gly Thr Arg Glu Gln

625 630 635 640

Phe Ala Lys Leu Tyr Lys Thr Ile Tyr Gln Phe Pro Glu Phe Asp Val

645 650 655

Arg Tyr Lys Leu Lys Glu Leu Ala Ala Tyr Leu Lys Ile Asp Pro Ile

660 665 670

Leu Leu Val Lys Met Ile Gln Ile Phe Glu Glu Leu Gly Phe Val Ser

675 680 685

Ile Thr Glu Gly Val Met Thr Val Asn Lys Glu Ala Glu Lys Lys Glu

690 695 700

Ile Asp Ser Ser His Ile Tyr Gln Asp Leu Lys Arg Leu Val Lys Glu

705 710 715 720

Gln Glu Leu Met Ala Leu Gly Thr Val Gln Glu Ile Tyr Asp Tyr Leu
725 730 735

Met Asp Cys

<210> 16

<211> 2220

<212> DNA

<213> Unknown

<220><223> Streptococcus cristatus AS 1.3089

<400> 16

atgattagct caaaatatga ctggcagttt gccactaatt ttacagatga aaaattttta 60

aaaaaggcta agaaagctgg gttagaaccc gcagccgcca gtcttcttta ccaaagaggt 120

gtgcagactg aagaggcctt gcaggaattt ttggagccta gtttagacca gcttcacgac 180

ccttatgacc ttcatgatat ggagcgggca gtggagcgca ttcgcgcagc gattgagaac 240

tacgaacaga ttttgatcta tggcgattat gatgcagatg gcatgacctc ggcatctatt 300

gtcaaggaag ccttggagca gctgggggct gagtgtcagg tctatctgcc caatcgcttt 360

acggatggct atggtcctaa tagcagtgtt tacaagtatt ttattgaaaa tcaaggcatc 420

tcgctcatta ttacagtgga taatggagtg gctggacttg aagctatcga gctggcccag 480

tctctcggcg tcgatgtcat cgtgacggat caccactcca tgcccgagga gctgccaaat 540

gcctatgcga tcgtccatcc ggaacatagc ggagcggatt acccttttcaa acatctggct 600

ggctgtgggg tggcttttaa gctggcaaca gccttgttgg aagaagtcca agtcgaactt 660

ttggacttag ttgccattgg caccattgcc gatatggtca gtctgacggg ggaaaatcgg 720

atcttagtca aatatggtct ttcagttctc aaaaataccc agcgagtggg acttcaggag 780

ctcttcaaaa tcgctggtat tcagccagat gaactggatg aagaaacagt tggtttccag 840

cttgccccta gactcaatgc cttgggacgg ttggatgatc ctaatccagc tattgagctt 900

ttgacaggct ttgacgacga agaagctcgt gacattgctc tcatgattaa ccagaaaaat 960

gacgagcgca aggaaatcgt ccagcagatt tataaagaag cacaaaccat gctggaccct 1020

aagagaccag tgcaggtgtt ggccaaggaa ggctggaatc ccggtgtgct agggattgtc 1080

gctgggcgct tgctggaaga gctgcatcag ccagttatag tgctcaatat tgaggatgga 1140

atcgctaagg gcagcgcccg cagtattgaa gcggtcaata tttttgaggc cttggacagt 1200

caccgtgatc tctttctagc ctttggtggg cacgctggag cggctggaat gactcttgaa 1260

gcagacaagc tggcggagct tgctgatatc ttgacagagt atatcttaga aaatgatttg 1320

gatttaactg gcaaaacagc tctgtattta gatgaggagc tgcatctgcc agaactgacc 1380

ctggatacgc tcaaaagctt cgaaaaactg gcgccttttg gtatggataa taagaagcca 1440

ctttttacc tcaaggactt taaagtggac aatgcacgga ctatgggggc tggcaatagc 1500

catctcaaac tgaaaatttc tcaggcagat gcatcttttg aagtagtagc ttttgggcaa 1560

ggaagcctgg cgacagagtt tgcccagact aagaatctgg agctggctgt cagcctctct 1620

gtcaataaat ggaatggaca gaccagtctc cagctcatgc tggtagatgc ccgtgtggac 1680

ggcgttcagc ttttcaatat ccgcggaaaa aatgcgacac tgcctgacaa ggttccagtc 1740

ttgcgtttca cagaggaatt gccggatttg acaaatagca gagcagttgt agtttatgac 1800

ctgccagata atttgcaggt cttgaagaag attcttcagt atcaggattt tgaagctatt 1860

tactttaaga atgaaattgc caagccatac tatctgaccg gttatggtac tcgtgagcaa 1920

tttgctaaac tttacaagac aatctatcag ttccccgagt ttgatgttcg ctataaactc 1980

aaggagctgg cagcttacct gaaaattgat cctattctct tggtcaaaat gattcagatt 2040

tttgaagagc tgggctttgt cagcatcacg gaaggcgtca tgacggtcaa taaggaagct 2100

gaaaagaaag aaattgacag tagtcacatt taccaagacc tcaagcgcct ggtcaaagaa 2160

caagaactta tggctctggg cactgtgcag gagatttacg actatctcat ggattgctag 2220

2220

<210> 17

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Flexible linker

<400> 17

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 18

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Flexible linker

<400> 18

Gly Gly Gly Gly Gly Gly Gly Gly

1 5

<210

> 19

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Flexible linker

<400> 19

Gly Gly Gly Gly Gly Gly

1 5

<210> 20

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Rigid linker

<400> 20

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10 15

<210> 21

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Rigid linker

<400

> 21

Glu Ala Ala Ala Lys

1 5

<210> 22

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Rigid linker

<400> 22

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10

<210> 23

<211> 46

<212> PRT

<213> Artificial Sequence

<220><223> Rigid linker

<400> 23

Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10 15

Glu Ala Ala Ala Lys Ala Leu Glu Ala Glu Ala Ala Ala Lys Glu Ala

20 25 30

Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala

35 40 45

<210> 24

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> Rigid linker

<400> 24

Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala

1 5 10

<210> 25

<211> 1442

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9

<400> 25

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro

1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys

35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala

50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu
85 90 95

Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr
100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His
130 135 140

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg

145 150 155 160

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys
165 170 175

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp
180 185 190

Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys
195 200 205

Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser
210 215 220

Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu
225 230 235 240

Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile
245 250 255

Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala
260 265 270

Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala
275 280 285

Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala

290 295 300

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

305 310 315 320

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

325 330 335

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

340 345 350

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

355 360 365

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

385 390 395 400

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu

405 410 415

Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu

420 425 430

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

435 440 445

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu

450 455 460

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

465 470 475 480

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

485 490 495

Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg

500 505 510

Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

515 520 525

Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp

530 535 540

Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr

545 550 555 560

Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr

565 570 575

Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala

580 585 590

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

595 600 605

Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu

610 615 620

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

625 630 635 640

Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu

645 650 655

Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu

660 665 670

Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe

675 680 685

Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp

690 695 700

Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser

705 710 715 720

Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg

725 730 735

Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp

740 745 750

Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His

755 760 765

Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln

770 775 780

Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys

785 790 795 800

Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln

805 810 815

Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly
820 825 830
Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
835 840 845
Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly
850 855 860
Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp

865 870 875 880
Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser
885 890 895
Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser
900 905 910
Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp
915 920 925
Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn
930 935 940

Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly
945 950 955 960
Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val
965 970 975
Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
980 985 990
Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val
995 1000 1005
Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn

1010 1015 1020
Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr
1025 1030 1035 1040
Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly
1045 1050 1055
Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln
1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met
1075 1080 1085

Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys
1090 1095 1100
Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp
1105 1110 1115 1120
Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln
1125 1130 1135
Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys
1140 1145 1150
Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165
Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val
1170 1175 1180
Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys
1185 1190 1195 1200
Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg
1205 1210 1215
Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr
1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
1235 1240 1245
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu
1250 1255 1260
Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe
1265 1270 1275 1280
Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp
1285 1290 1295
Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp

1300 1305 1310
Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu Lys Arg Pro Ala

1410 1415 1420

Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Glu His His His His

1425 1430 1435 1440

His His

<

210> 26

<211> 4329

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9

<400> 26

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560

tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620

accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680

gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740

aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800

aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860

gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920

gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980

ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040

ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100

tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220

ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280

cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340

ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400

cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460

aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520
ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580
ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640
tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700
gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760
gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820
aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880

ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940
gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000
atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060
cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120
gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180
tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240
tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300

gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360
aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420
aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480
agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540
agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600
aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660
aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780
agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840
ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900
cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960
ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020
aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080
accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctcgag 4260

aaaaggccgg cggccacgaa aaaggccggc caggcaaaaa agaaagagca ccaccaccac 4320

caccactga 4329

<210> 27

<211> 1682

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9-GFP

<400> 27

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu

85 90 95

Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr
100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His
130 135 140

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg
145 150 155 160

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys

165 170 175

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

225 230 235 240
Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile
245 250 255
Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala
260 265 270
Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala
275 280 285
Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala
290 295 300

Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu
305 310 315 320
Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu
325 330 335
Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg
340 345 350
Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys
355 360 365
Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380
Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser
385 390 395 400
Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu
405 410 415
Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu
420 425 430

Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg
435 440 445

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu
450 455 460
His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp
465 470 475 480
Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr
485 490 495
Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg
500 505 510
Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

515 520 525
Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp
530 535 540
Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr
545 550 555 560
Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr
565 570 575
Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala
580 585 590

Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln
595 600 605
Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
610 615 620
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
625 630 635 640
Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
645 650 655
Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu

660 665 670
Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe

675 680 685

Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp

690 695 700

Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser

705 710 715 720

Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg

725 730 735

Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp

740 745 750

Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His

755 760 765

Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln

770 775 780

Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys

785 790 795 800

Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln

805 810 815

Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly

820 825 830

Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn

835 840 845

Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly

850 855 860

Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp

865 870 875 880

Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser

885 890 895

Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser

900 905 910

Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp

915 920 925

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

930 935 940

Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly

945 950 955 960

Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val

965 970 975

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

980 985 990

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly

1045 1050 1055

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

1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met

1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys
1185 1190 1195 1200
Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg
1205 1210 1215
Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr
1220 1225 1230
Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu
1250 1255 1260
Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe
1265 1270 1275 1280
Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp
1285 1290 1295
Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp
1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala
1315 1320 1325
Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp
1330 1335 1340
Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu
1345 1350 1355 1360
Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile
1365 1370 1375
Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu

1380 1385 1390
Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser
1395 1400 1405
Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu Met Ser Lys Gly
1410 1415 1420
Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly
1425 1430 1435 1440

Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp
1445 1450 1455

Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys
1460 1465 1470
Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe Ser Tyr Gly Val
1475 1480 1485
Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg His Asp Phe Phe
1490 1495 1500
Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe
1505 1510 1515 1520
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly

1525 1530 1535
Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu
1540 1545 1550
Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His
1555 1560 1565
Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly Ile Lys Ala Asn
1570 1575 1580
Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp
1585 1590 1595 1600

His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro
1605 1610 1615
Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn
1620 1625 1630
Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly
1635 1640 1645
Ile Thr His Gly Met Asp Glu Leu Tyr Lys Leu Glu Lys Arg Pro Ala
1650 1655 1660
Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His His

1665 1670 1675 1680
His His

<210> 28

<211> 5049

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9-GFP

<400> 28

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560
tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620
accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680
gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740
aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800
aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860
gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920

gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980
ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040
ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100
tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160
ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220
ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280
cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340

ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400
cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460
aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520
ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580
ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640
tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700
gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760

gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820
aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880
ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940
gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000
atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060
cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120
gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180

tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240

tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300

gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360

aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480

agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540

agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600

aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780

agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctcgag 4260

atgagtaaag gagaagaact tttcactgga gttgtcccaa ttcttgttga attagatggt 4320

gatgttaatg ggcacaaatt ttctgtcagt ggagagggtg aaggtgatgc aacatacgga 4380

aaacttaccc ttaaatttat ttgcactact ggaaaactac ctgttccatg gccaacactt 4440

gtcactactt tctcttatgg tgttcaatgc ttttcaagat acccagatca tatgaagcgg 4500

cacgacttct tcaagagcgc catgcctgag ggatacgtgc aggagaggac catctctttc 4560

aaggacgacg ggaactacaa gacacgtgct gaagtcaagt ttgagggaga caccctcgtc 4620

aacaggatcg agcttaaggg aatcgatttc aaggaggacg gaaacatcct cggccacaag 4680

ttggaataca actacaactc ccacaacgta tacatcacgg cagacaaaca aaagaatgga 4740

atcaaagcta acttcaaaat tagacacaac attgaagatg gaagcgttca actagcagac 4800

cattatcaac aaaatactcc aattggcgat ggccctgtcc ttttaccaga caaccattac 4860

ctgtccacac aatctgccct ttcgaaagat cccaacgaaa agagagacca catggtcctt 4920

cttgagtttg taacagctgc tgggattaca catggcatgg atgaactata caaactcgag 4980

aaaaggccgg cggccacgaa aaaggccggc caggcaaaaa agaaaaagca ccaccaccac 5040

caccactga 5049

<210> 29

<211> 1953

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9-hTdT

<400> 29

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu

85 90 95

Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr
100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His
130 135 140

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg
145 150 155 160

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys

165 170 175

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp
180 185 190

Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys

195 200 205

Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser

210 215 220

Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu

225 230 235 240

Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile

245 250 255

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

260 265 270

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

275 280 285

Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala

290 295 300

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

305 310 315 320

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

325 330 335

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

340 345 350

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

355 360 365

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

385 390 395 400

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu

405 410 415

Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu

420 425 430

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

435 440 445

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu
450 455 460
His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp
465 470 475 480
Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr
485 490 495
Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg
500 505 510
Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

515 520 525
Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp
530 535 540
Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr
545 550 555 560
Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr
565 570 575
Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala
580 585 590

Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln
595 600 605
Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
610 615 620
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
625 630 635 640
Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
645 650 655
Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu

660 665 670
Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe
675 680 685
Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp
690 695 700

Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser
705 710 715 720
Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg
725 730 735

Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp
740 745 750
Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His
755 760 765
Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
770 775 780
Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys
785 790 795 800
Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln

805 810 815
Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly
820 825 830
Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
835 840 845
Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly
850 855 860
Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp
865 870 875 880

Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser
885 890 895
Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser
900 905 910
Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp
915 920 925
Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn
930 935 940
Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly

945 950 955 960

Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val

965 970 975

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

980 985 990

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly

1045 1050 1055

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

1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met

1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1185 1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1205 1210 1215

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245

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

1250 1255 1260

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1265 1270 1275 1280

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1285 1290 1295

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

1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Gln Met Asp Pro Pro

1410 1415 1420

Arg Ala Ser His Leu Ser Pro Arg Lys Lys Arg Pro Arg Gln Thr Gly

1425 1430 1435 1440

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

1445 1450 1455

Val Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg Arg Ala Phe Leu
1460 1465 1470
Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu Asn Glu Leu Ser
1475 1480 1485
Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser Gly Ser Asp Val
1490 1495 1500
Leu Glu Trp Leu Gln Ala Gln Lys Val Gln Val Ser Ser Gln Pro Glu
1505 1510 1515 1520
Leu Leu Asp Val Ser Trp Leu Ile Glu Cys Ile Gly Ala Gly Lys Pro

1525 1530 1535
Val Glu Met Thr Gly Lys His Gln Leu Val Val Arg Arg Asp Tyr Ser
1540 1545 1550
Asp Ser Thr Asn Pro Gly Pro Pro Lys Thr Pro Pro Ile Ala Val Gln
1555 1560 1565
Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr Leu Asn Asn Cys
1570 1575 1580
Asn Gln Ile Phe Thr Asp Ala Phe Asp Ile Leu Ala Glu Asn Cys Glu
1585 1590 1595 1600

Phe Arg Glu Asn Glu Asp Ser Cys Val Thr Phe Met Arg Ala Ala Ser
1605 1610 1615
Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met Lys Asp Thr Glu
1620 1625 1630
Gly Ile Pro Cys Leu Gly Ser Lys Val Lys Gly Ile Ile Glu Glu Ile
1635 1640 1645
Ile Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val Leu Asn Asp Glu
1650 1655 1660
Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe Gly Val Gly Leu

1665 1670 1675 1680
Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg Thr Leu Ser Lys
1685 1690 1695
Val Arg Ser Asp Lys Ser Leu Lys Phe Thr Arg Met Gln Lys Ala Gly
1700 1705 1710

Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr Arg Ala Glu Ala
1715 1720 1725
Glu Ala Val Ser Val Leu Val Lys Glu Ala Val Trp Ala Phe Leu Pro
1730 1735 1740

Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg Gly Lys Lys Met
1745 1750 1755 1760
Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly Ser Thr Glu Asp
1765 1770 1775
Glu Glu Gln Leu Leu Gln Lys Val Met Asn Leu Trp Glu Lys Lys Gly
1780 1785 1790
Leu Leu Leu Tyr Tyr Asp Leu Val Glu Ser Thr Phe Glu Lys Leu Arg
1795 1800 1805
Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe Gln Lys Cys Phe

1810 1815 1820
Leu Ile Phe Lys Leu Pro Arg Gln Arg Val Asp Ser Asp Gln Ser Ser
1825 1830 1835 1840
Trp Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val Asp Leu Val Leu
1845 1850 1855
Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu Gly Trp Thr Gly Ser
1860 1865 1870
Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Thr His Glu Arg Lys
1875 1880 1885

Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr Lys Arg Ile Phe
1890 1895 1900
Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His Leu Gly Leu Asp
1905 1910 1915 1920
Tyr Ile Glu Pro Trp Glu Arg Asn Ala Leu Gln Lys Arg Pro Ala Ala
1925 1930 1935
Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His His His
1940 1945 1950
His

<

210> 30

<211> 5862

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9-hTdT

<400> 30

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctaccccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560

tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620
accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680
gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740
aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800
aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860
gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920
gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980

ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040
ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100
tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160
ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220
ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280
cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340
ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400

cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460
aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520
ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580
ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640
tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700
gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760
gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820

aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880
ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940
gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000
atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060
cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120
gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180
tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240

tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300

gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360

aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480

agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540

agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600

aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780

agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctgcag 4260

atggatccac cacgagcgtc ccacttgagc cctcggaaga agagaccccg gcagacgggt 4320

gccttgatgg cctcctctcc tcaagacatc aaatttcaag atttggtcgt cttcattttg 4380

gagaagaaaa tgggaaccac ccgcagagcg ttcctcatgg agctggcccg caggaaaggg 4440

ttcagggttg aaaatgagct cagtgattct gtcacccaca ttgtagcaga gaacaactcg 4500

ggttcggatg ttctggagtg gcttcaagca cagaaagtac aagtcagctc acaaccagag 4560

ctcctcgatg tctcctggct gatcgaatgc ataggagcag ggaaaccggt ggaaatgaca 4620

ggaaaacacc agcttgttgt gagaagagac tattcagata gcaccaaccc aggccccccg 4680

aagactccac caattgctgt acaaaagatc tcccagtatg cgtgtcagag aagaaccact 4740

ttaaacaact gtaaccagat attcacggat gcctttgata tactggctga aaactgtgag 4800

tttagagaaa atgaagactc ctgtgtgaca tttatgagag cagcttctgt attgaaatct 4860

ctgccattca caatcatcag tatgaaggac acagaaggaa ttccctgcct ggggtccaag 4920

gtgaagggta tcatagagga gattattgaa gatggagaaa gttctgaagt taaagctgtg 4980

ttaaatgatg aacgatatca atccttcaaa ctctttactt ctgtatttgg agtggggctg 5040

aagacttctg agaagtggtt caggatgggt ttcagaactc tgagtaaagt aaggtcggac 5100

aaaagcctga aatttacacg aatgcagaaa gcaggatttc tgtattatga agaccttgtc 5160

agctgtgtga ccagggcaga agcagaggcc gtcagtgtgc tggttaaaga ggctgtctgg 5220

gcatttcttc cggatgcttt cgtcaccatg acaggagggt tccggagggg taagaagatg 5280

gggcatgatg tagatttttt aattaccagc ccaggatcaa cagaggatga agagcaactt 5340

ttacagaaag tgatgaactt atgggaaaag aagggattac ttttatatta tgaccttgtg 5400

gagtcaacat ttgaaaagct caggttgcct agcaggaagg ttgatgcttt ggatcatttt 5460

caaaagtgct ttctgatttt caaattgcct cgtcaaagag tggacagtga ccagtccagc 5520

tggcaggaag gaaagacctg gaaggccatc cgtgtggatt tagttctgtg cccctacgag 5580

cgtcgtgcct ttgccctgtt gggatggact ggctcccggc agtttgagag agacctccgg 5640

cgctatgcca cacatgagcg gaagatgatt ctggataacc atgctttata tgacaagacc 5700

aagaggatat tcctcaaagc agaaagtgaa gaagaaattt ttgcgcatct gggattggat 5760

tatattgaac cgtgggaaag aaatgccctg cagaaaaggc cggcggccac gaaaaaggcc 5820

ggccaggcaa aaaagaaaaa gcaccaccac caccaccact ga 5862

<210> 31

<211> 2021

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9-RecJ

<400> 31

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu
85 90 95

Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr

100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln

115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His

130 135 140

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg

145 150 155 160

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys

165 170 175

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp

180 185 190

Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys

195 200 205

Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser

210 215 220

Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu

225 230 235 240

Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile

245 250 255

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

260 265 270

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

275 280 285

Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala

290 295 300

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

305 310 315 320

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

325 330 335

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

340 345 350

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

355 360 365

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

385 390 395 400

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu

405 410 415

Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu

420 425 430

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

435 440 445

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu

450 455 460

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

465 470 475 480

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

485 490 495

Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg

500 505 510

Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

515 520 525

Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp

530 535 540

Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr

545 550 555 560

Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr

565 570 575

Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala

580 585 590

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

595 600 605

Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
610 615 620
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
625 630 635 640
Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
645 650 655
Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu
660 665 670
Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe

675 680 685
Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp
690 695 700
Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser
705 710 715 720
Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg
725 730 735
Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp
740 745 750

Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His
755 760 765
Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
770 775 780
Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys
785 790 795 800
Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln
805 810 815
Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly

820 825 830
Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
835 840 845

Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly
850 855 860
Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp
865 870 875 880
Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser
885 890 895

Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser
900 905 910
Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp
915 920 925
Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn
930 935 940
Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly
945 950 955 960
Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val
965 970 975
Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
980 985 990
Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val
995 1000 1005
Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn
1010 1015 1020
Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr
1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly
1045 1050 1055
Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln
1060 1065 1070
Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met
1075 1080 1085
Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys
1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1185 1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1205 1210 1215

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245

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

1250 1255 1260

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1265 1270 1275 1280

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1285 1290 1295

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

1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Gln Val Lys Gln Gln

1410 1415 1420

Ile Gln Leu Arg Arg Arg Glu Val Asp Glu Thr Ala Asp Leu Pro Ala

1425 1430 1435 1440

Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser Arg Gly Val Arg

1445 1450 1455

Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met Leu Pro Trp Gln

1460 1465 1470

Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu Tyr Asn Ala Phe

1475 1480 1485

Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe Asp Ala Asp Gly

1490 1495 1500

Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg Ser Leu Gly Cys

1505 1510 1515 1520

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

1525 1530 1535

Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg Gly Ala Gln Leu

1540 1545 1550

Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala Gly Val Glu His

1555 1560 1565

Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp His His Leu Pro

1570 1575 1580

Gly Glu Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn Pro Asn Leu Arg

1585 1590 1595 1600

Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val Gly Val Ala Phe

1605 1610 1615

Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp Gln Gly Trp Phe

1620 1625 1630

Asp Glu Arg Gly Ile Ala Ile Pro Asn Leu Ala Glu Leu Leu Asp Leu

1635 1640 1645

Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu Asp Ala Asn Asn

1650 1655 1660

Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg Ala Gly Lys Cys

1665 1670 1675 1680

Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn Arg Asp Ala Gln

1685 1690 1695

Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly Pro Arg Leu Asn

1700 1705 1710

Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val Ala Leu Leu Leu

1715 1720 1725

Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn Glu Leu Asp Ala

1730 1735 1740

Leu Asn Gln Thr Arg Lys Glu Ile Glu Gln Gly Met Gln Val Glu Ala

1745 1750 1755 1760

Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser Arg Asp Thr Leu Pro Gly

1765 1770 1775

Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly Val Val Gly Ile

1780 1785 1790

Leu Ala Ser Arg Ile Lys Glu Arg Phe His Arg Pro Val Ile Ala Phe

1795 1800 1805

Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly Ser Gly Arg Ser Ile Gln

1810 1815 1820

Gly Leu His Met Arg Asp Ala Leu Glu Arg Leu Asp Thr Leu Tyr Pro

1825 1830 1835 1840

Gly Met Ile Leu Lys Phe Gly Gly His Ala Met Ala Ala Gly Leu Ser

1845 1850 1855

Leu Glu Glu Asp Lys Phe Glu Leu Phe Gln Gln Arg Phe Gly Glu Leu
1860 1865 1870
Val Thr Glu Trp Leu Asp Pro Ser Leu Leu Gln Gly Glu Val Val Ser
1875 1880 1885
Asp Gly Pro Leu Ser Pro Ala Glu Met Thr Met Glu Val Ala Gln Leu
1890 1895 1900

Leu Arg Asp Ala Gly Pro Trp Gly Gln Met Phe Pro Glu Pro Leu Phe
1905 1910 1915 1920
Asp Gly His Phe Arg Leu Leu Gln Gln Arg Leu Val Gly Glu Arg His
1925 1930 1935
Leu Lys Val Met Val Glu Pro Val Gly Gly Gly Pro Leu Leu Asp Gly
1940 1945 1950
Ile Ala Phe Asn Val Asp Thr Ala Leu Trp Pro Asp Asn Gly Val Arg
1955 1960 1965
Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile Asn Glu Phe Arg Gly Asn

1970 1975 1980
Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile Trp Pro Ile Leu Gln Lys
1985 1990 1995 2000
Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His
2005 2010 2015
His His His His His
2020

<210> 32
<211> 6066
<212> DNA
<213> Artificial Sequence
<220><223> SpyCas9-RecJ
<400> 32

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120
caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180
gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560

tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620

accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680

gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740

aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800

aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860

gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920

gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980

ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040

ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100

tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220
ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280
cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340
ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400
cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460
aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520
ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580

ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640
tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700
gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760
gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820
aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880
ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940
gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000

atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060
cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120
gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180
tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240
tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300
gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360
aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480
agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540
agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600
aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660
aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720
atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780
agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctgcag 4260

gtgaaacaac agatacaact tcgtcgccgt gaagtcgatg aaacggcaga cttgcccgct 4320

gaattgcctc ccttgctgcg ccgtttatac gccagccggg gcgtgcgcag tgcgcaagaa 4380

ctggaacgca gtgttaaagg tatgttgccc tggcagcaac tgagcggcgt cgaaaaggcc 4440

gttgagatcc tttacaacgc ttttcgcgaa ggaacgcgga ttattgtggt cggcgatttt 4500

gacgccgacg gcgcgaccag cacggctcta agcgtgctgg cgatgcgctc gcttggttgc 4560

agcaatatcg actatctggt accaaaccgt ttcgaagacg gttacggctt aagcccggaa 4620

gtagtcgatc aggcccatgc ccgtggcgcg cagttaattg tcacggtgga taacggtatt 4680

tcctcccatg cgggcgttga acacgctcgc tcgttgggca ttccggttat tgttaccgat 4740

caccatttgc cgggcgaaac attacccgca gcggaagcga tcattaaccc taacttgcgc 4800

gactgtaatt tcccgtcgaa atcactggca ggcgtgggtg tggcgtttta tctgatgctg 4860

gcgctgcgca ccttttgcg cgatcagggc tggtttgatg agcgtggcat cgcaattcct 4920

aacctggcag aactgctgga tctggtcgcg ctgggaacag tggcggacgt cgtgccgctg 4980

gacgctaata atcgcattct gacctggcag gggatgagtc gcatccgtgc cggaaagtgc 5040

cgtccaggga ttaaagcgct gctggaagtg gcaaaccgtg atgcacaaaa actcgccgcc 5100

agcgatttag gttttgcgct ggggccacgt ctcaatgctg ccggacgact ggacgatatg 5160

tccgtcggtg tggcgctctt gctgtgcgac aacatcggcg aagcgcgcgt gctggcaaat 5220

gaactcgatg cgctaaacca gacgcgaaaa gagatcgaac aaggaatgca agttgaagcc 5280

ctgaccctgt gcgagaaact ggagcgaagt cgcgacacgc tacccggcgg gctggcaatg 5340

tatcaccccg aatggcatca gggcgttgtc ggtattctgg cttcgcgcat caaagagcgt 5400

tttcaccgtc cggttatcgc ctttgcgcca gcaggtgatg gtacgctgaa aggttcaggt 5460

cgctccattc aggggctgca tatgcgtgat gcactggagc gattagacac actctaccct 5520

ggcatgatac tgaagtttgg cggtcatgcg atggcggcgg gtttgtcgct ggaagaggat 5580

aaattcgaac tctttcaaca acggtttggc gagctggtta ccgagtggct ggacccttcg 5640

ctattgcaag gcgaagtggt gtcagacggc ccgttaagcc cggccgaaat gaccatggaa 5700

gtggcgcagc tgctgcgcga tgctggcccg tgggggcaga tgttcccgga gccgctgttt 5760

gatggtcatt tccgtctgct gcaacagcgg ctggtgggcg aacgtcattt gaaagtcatg 5820

gtcgaaccgg tcggcggcgg tccgctgctg gatggtattg cttttaatgt cgataccgcc 5880

ctctggccgg ataacggcgt gcgcgaagtg caactggctt acaagctcga tatcaacgag 5940

tttcgcggca accgcagcct gcaaattatc atcgacaata tctggccaat tctgcagaaa 6000

aggccggcgg ccacgaaaaa ggccggccag gcaaaaaaga aaaagcacca ccaccaccac 6060

cactga 6066

<210> 33

<211> 2310

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9-RecE

<400> 33

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu
85 90 95

Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr
100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His
130 135 140

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg

145 150 155 160

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys

165 170 175

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp

180 185 190

Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys

195 200 205

Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser

210 215 220

Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu

225 230 235 240

Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile

245 250 255

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

260 265 270

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

275 280 285

Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala

290 295 300

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

305 310 315 320

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

325 330 335

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

340 345 350

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

355 360 365

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

385 390 395 400

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu
405 410 415
Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu
420 425 430
Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg
435 440 445
Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu

450 455 460
His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp
465 470 475 480
Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr
485 490 495
Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg
500 505 510
Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp
515 520 525

Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp
530 535 540
Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr
545 550 555 560
Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr
565 570 575
Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala
580 585 590
Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln

595 600 605
Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
610 615 620
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
625 630 635 640
Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
645 650 655

Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu
660 665 670

Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe
675 680 685
Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp
690 695 700
Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser
705 710 715 720
Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg
725 730 735
Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp

740 745 750
Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His
755 760 765
Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
770 775 780
Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys
785 790 795 800
Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln
805 810 815

Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly
820 825 830
Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
835 840 845
Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly
850 855 860
Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp
865 870 875 880
Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser

885 890 895
Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser

900 905 910

Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp

915 920 925

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

930 935 940

Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly

945 950 955 960

Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val

965 970 975

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

980 985 990

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly

1045 1050 1055

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

1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met

1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1185 1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1205 1210 1215

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245

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

1250 1255 1260

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1265 1270 1275 1280

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1285 1290 1295

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

1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Gln Met Ser Thr Lys
1410 1415 1420

Pro Leu Phe Leu Leu Arg Lys Ala Lys Lys Ser Ser Gly Glu Pro Asp
1425 1430 1435 1440

Val Val Leu Trp Ala Ser Asn Asp Phe Glu Ser Thr Cys Ala Thr Leu
1445 1450 1455

Asp Tyr Leu Ile Val Lys Ser Gly Lys Lys Leu Ser Ser Tyr Phe Lys

1460 1465 1470

Ala Val Ala Thr Asn Phe Pro Val Val Asn Asp Leu Pro Ala Glu Gly
1475 1480 1485

Glu Ile Asp Phe Thr Trp Ser Glu Arg Tyr Gln Leu Ser Lys Asp Ser
1490 1495 1500

Met Thr Trp Glu Leu Lys Pro Gly Ala Ala Pro Asp Asn Ala His Tyr
1505 1510 1515 1520

Gln Gly Asn Thr Asn Val Asn Gly Glu Asp Met Thr Glu Ile Glu Glu
1525 1530 1535

Asn Met Leu Leu Pro Ile Ser Gly Gln Glu Leu Pro Ile Arg Trp Leu

1540 1545 1550

Ala Gln His Gly Ser Glu Lys Pro Val Thr His Val Ser Arg Asp Gly
1555 1560 1565

Leu Gln Ala Leu His Ile Ala Arg Ala Glu Glu Leu Pro Ala Val Thr
1570 1575 1580

Ala Leu Ala Val Ser His Lys Thr Ser Leu Leu Asp Pro Leu Glu Ile
1585 1590 1595 1600

Arg Glu Leu His Lys Leu Val Arg Asp Thr Asp Lys Val Phe Pro Asn

1605 1610 1615

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

1620 1625 1630

Leu Asn Ala Asp Tyr Thr Asp Arg Gly Leu Leu Thr Lys Glu Trp Met
1635 1640 1645

Lys Gly Asn Arg Val Ser His Ile Thr Arg Thr Ala Ser Gly Ala Asn
1650 1655 1660

Ala Gly Gly Gly Asn Leu Thr Asp Arg Gly Glu Gly Phe Val His Asp
1665 1670 1675 1680

Leu Thr Ser Leu Ala Arg Asp Val Ala Thr Gly Val Leu Ala Arg Ser
1685 1690 1695
Met Asp Leu Asp Ile Tyr Asn Leu His Pro Ala His Ala Lys Arg Ile
1700 1705 1710
Glu Glu Ile Ile Ala Glu Asn Lys Pro Pro Phe Ser Val Phe Arg Asp
1715 1720 1725
Lys Phe Ile Thr Met Pro Gly Gly Leu Asp Tyr Ser Arg Ala Ile Val
1730 1735 1740
Val Ala Ser Val Lys Glu Ala Pro Ile Gly Ile Glu Val Ile Pro Ala

1745 1750 1755 1760
His Val Thr Glu Tyr Leu Asn Lys Val Leu Thr Glu Thr Asp His Ala
1765 1770 1775
Asn Pro Asp Pro Glu Ile Val Asp Ile Ala Cys Gly Arg Ser Ser Ala
1780 1785 1790
Pro Met Pro Gln Arg Val Thr Glu Glu Gly Lys Gln Asp Asp Glu Glu
1795 1800 1805
Lys Pro Gln Pro Ser Gly Thr Thr Ala Val Glu Gln Gly Glu Ala Glu
1810 1815 1820

Thr Met Glu Pro Asp Ala Thr Glu His His Gln Asp Thr Gln Pro Leu
1825 1830 1835 1840
Asp Ala Gln Ser Gln Val Asn Ser Val Asp Ala Lys Tyr Gln Glu Leu
1845 1850 1855
Arg Ala Glu Leu His Glu Ala Arg Lys Asn Ile Pro Ser Lys Asn Pro
1860 1865 1870
Val Asp Asp Asp Lys Leu Leu Ala Ala Ser Arg Gly Glu Phe Val Asp
1875 1880 1885
Gly Ile Ser Asp Pro Asn Asp Pro Lys Trp Val Lys Gly Ile Gln Thr

1890 1895 1900
Arg Asp Cys Val Tyr Gln Asn Gln Pro Glu Thr Glu Lys Thr Ser Pro

1905 1910 1915 1920

Asp Met Asn Gln Pro Glu Pro Val Val Gln Gln Glu Pro Glu Ile Ala

1925 1930 1935

Cys Asn Ala Cys Gly Gln Thr Gly Gly Asp Asn Cys Pro Asp Cys Gly

1940 1945 1950

Ala Val Met Gly Asp Ala Thr Tyr Gln Glu Thr Phe Asp Glu Glu Ser

1955 1960 1965

Gln Val Glu Ala Lys Glu Asn Asp Pro Glu Glu Met Glu Gly Ala Glu

1970 1975 1980

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

1985 1990 1995 2000

Asp Glu Thr Gly Glu Val Ala Asp Pro Val Ile Val Glu Asp Ile Glu

2005 2010 2015

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

2020 2025 2030

Gly Ile Ser Lys Ser Gln Leu Asp Asp Ile Ala Asp Thr Pro Ala Leu

2035 2040 2045

Tyr Leu Trp Arg Lys Asn Ala Pro Val Asp Thr Thr Lys Thr Lys Thr

2050 2055 2060

Leu Asp Leu Gly Thr Ala Phe His Cys Arg Val Leu Glu Pro Glu Glu

2065 2070 2075 2080

Phe Ser Asn Arg Phe Ile Val Ala Pro Glu Phe Asn Arg Arg Thr Asn

2085 2090 2095

Ala Gly Lys Glu Glu Glu Lys Ala Phe Leu Met Glu Cys Ala Ser Thr

2100 2105 2110

Gly Lys Thr Val Ile Thr Ala Glu Glu Gly Arg Lys Ile Glu Leu Met

2115 2120 2125

Tyr Gln Ser Val Met Ala Leu Pro Leu Gly Gln Trp Leu Val Glu Ser

2130 2135 2140

Ala Gly His Ala Glu Ser Ser Ile Tyr Trp Glu Asp Pro Glu Thr Gly

2145 2150 2155 2160

Ile Leu Cys Arg Cys Arg Pro Asp Lys Ile Ile Pro Glu Phe His Trp

2165 2170 2175

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

2180 2185 2190

Tyr Tyr Asp Tyr Arg Tyr His Val Gln Asp Ala Phe Tyr Ser Asp Gly

2195 2200 2205

Tyr Glu Ala Gln Phe Gly Val Gln Pro Thr Phe Val Phe Leu Val Ala

2210 2215 2220

Ser Thr Thr Ile Glu Cys Gly Arg Tyr Pro Val Glu Ile Phe Met Met

2225 2230 2235 2240

Gly Glu Glu Ala Lys Leu Ala Gly Gln Gln Glu Tyr His Arg Asn Leu

2245 2250 2255

Arg Thr Leu Ser Asp Cys Leu Asn Thr Asp Glu Trp Pro Ala Ile Lys

2260 2265 2270

Thr Leu Ser Leu Pro Arg Trp Ala Lys Glu Tyr Ala Asn Asp Leu Gln

2275 2280 2285

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

2290 2295 2300

His His His His His His

2305 2310

<210> 34

<211> 6933

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9-RecE

<400> 34

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480
caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540
taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600
ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660
cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720
ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780
ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900
ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960
gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020
ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080
cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140
aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200
aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320
gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380
ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440
aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggccccctg 1500
gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560
tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620
accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680

gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740
aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800
aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860
gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920
gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980
ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040
ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100

tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220

ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280

cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340

ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400

cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460

aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520

ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580

ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640

tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700

gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760

gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820

aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880

ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940

gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000

atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060

cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120

gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180

tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240

tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300

gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360

aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480

agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540

agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600

aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780

agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgccccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctgcag 4260

atgagcacaa aaccactctt cctgttacgg aaagcgaaaa aatcatccgg tgaacctgac 4320

gtcgtcctgt gggcaagcaa cgattttgaa tcgacctgtg ccactctgga ctacctgatc 4380

gttaagtcag gtaaaaaact gagcagctat tttaaagctg ttgccacgaa ttttcctgtc 4440

gttaatgacc tgcccgctga aggtgagatc gattttacct ggagtgaacg ctatcaactc 4500

agcaaagact ccatgacatg ggaactaaaa ccgggagcag caccagacaa cgctcactat 4560

caaggcaata ccaacgtcaa cggcgaagac atgactgaga ttgaggagaa tatgctactc 4620

ccaatttctg gccaggaact gcccattcgt tggcttgctc aacacggcag cgaaaaaccg 4680

gtaacgcacg tttcacgcga cggactccag gcattacaca ttgctcgggc tgaagaacta 4740

ccggctgtta ctgccctggc tgtttcccac aaaaccagcc tgctcgaccc gctggaaatt 4800

cgcgaactcc acaaactggt tcgtgacact gacaaagttt tccctaatcc tggtaattca 4860

aacctgggac tgataactgc ttttttcgaa gcatacctga acgctgacta caccgatcga 4920

ggactgctga caaaagagtg gatgaagggt aatcgtgttt cacacatcac tcgcacggct 4980

tccggtgcta atgctggcgg cggaaacctc accgatcgcg gcgaaggttt cgtacacgat 5040

ctgacgtcac tggcgcgcga cgtagccact ggcgtactgg cccgttcaat ggatctggac 5100

atctataacc ttcatccggc acacgctaaa cgcattgagg aaattatcgc tgaaaataaa 5160

ccgccctttt ctgttttccg cgacaaattc atcaccatgc ctggcgggct ggattattcc 5220

cgcgccatcg tggttgcgtc cgtaaaagaa gcaccaattg ggatcgaggt catccccgcg 5280

cacgtcactg aatatctgaa caaagtactg actgaaaccg atcatgccaa ccctgatccg 5340

gaaatcgtgg atattgcctg cggtcgctcc tctgccccga tgccgcagcg agtaacagaa 5400

gaaggaaaac aggatgatga agaaaaaccg caaccatctg gaacaacggc agttgaacag 5460

ggagaggctg aaacaatgga accggacgca actgaacatc atcaggacac gcagccgctg 5520

gatgctcagt cacaggtaaa ttctgttgat gcgaaatatc aggaactgcg ggcagaactc 5580

catgaagccc ggaaaaacat tccatcaaaa aatcctgtcg atgacgataa attgcttgct 5640

gcatcacgtg gtgaatttgt tgacggaatt agcgacccga acgatccgaa atgggtaaag 5700

gggatccaga ctcgcgattg tgtgtaccag aaccagccag aaacggaaaa aaccagccca 5760

gatatgaatc aacctgagcc agtagtgcaa caggaaccgg aaatagcctg caatgcctgc 5820

ggccagactg gcggggataa ctgccctgac tgtggtgcgg tgatgggcga cgcaacatac 5880

caggaaacat tcgatgaaga gagtcaggtt gaagctaagg aaaatgatcc ggaggaaatg 5940
gaaggcgctg aacatccgca caatgagaat gctggcagcg atccgcatcg cgattgcagt 6000
gatgaaactg gcgaagtcgc agatcccgta atcgtagaag acatagagcc aggtatttat 6060
tacggaattt cgaatgagaa ttaccacgcg ggtcccggta tcagtaagtc tcagctcgat 6120
gacattgctg atactccggc actatatttg tggcgtaaaa atgcccccgt ggacaccaca 6180
aagacaaaaa cgctcgattt aggaactgct ttccactgcc gggtacttga accggaagaa 6240
ttcagtaacc gctttatcgt agcacctgaa tttaaccgcc gtacaaacgc cggaaaagaa 6300

gaagagaaag cgtttctgat ggaatgcgca agcacaggaa aaacggttat cactgcggaa 6360
gaaggccgga aaattgaact catgtatcaa agcgttatgg ctttgccgct ggggcaatgg 6420
cttgttgaaa gcgccggaca cgctgaatca tcaatttact gggaagatcc tgaaacagga 6480
attttgtgtc ggtgccgtcc ggacaaaatt atccctgaat ttcactggat catggacgtg 6540
aaaactacgg cggatattca acgattcaaa accgcttatt acgactaccg ctatcacgtt 6600
caggatgcat tctacagtga cggttatgaa gcacagtttg gagtgcagcc aactttcgtt 6660
tttctggttg ccagcacaac tattgaatgc ggacgttatc cggttgaaat tttcatgatg 6720

ggcgaagaag caaaactggc aggtcaacag gaatatcacc gcaatctgcg aaccctgtct 6780
gactgcctga ataccgatga atggccagct attaagacat tatcactgcc ccgctgggct 6840
aaggaatatg caaatgacct gcagaaaagg ccggcggcca cgaaaaaggc cggccaggca 6900
aaaaagaaaa agcaccacca ccaccaccac tga 6933

<210> 35
<211> 1670
<212> PRT
<213> Artificial Sequence
<220><223> SpyCas9-lambda
<400> 35

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30
Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60
Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80
Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu

85 90 95
Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr
100 105 110
Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125
Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His
130 135 140
Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg
145 150 155 160

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

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

290 295 300

Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu
305 310 315 320
Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu
325 330 335
Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg
340 345 350
Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys
355 360 365
Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380
Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser
385 390 395 400
Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu
405 410 415
Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu
420 425 430
Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg
435 440 445

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu
450 455 460
His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp
465 470 475 480
Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr
485 490 495
Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg
500 505 510
Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

515 520 525
Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp
530 535 540

Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr
545 550 555 560
Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr
565 570 575
Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala
580 585 590

Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln
595 600 605
Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
610 615 620
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
625 630 635 640
Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
645 650 655
Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu

660 665 670
Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe
675 680 685
Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp
690 695 700
Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser
705 710 715 720
Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg
725 730 735

Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp
740 745 750
Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His
755 760 765
Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
770 775 780
Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys
785 790 795 800

Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln

805 810 815

Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly

820 825 830

Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn

835 840 845

Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly

850 855 860

Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp

865 870 875 880

Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser

885 890 895

Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser

900 905 910

Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp

915 920 925

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

930 935 940

Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly

945 950 955 960

Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val

965 970 975

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

980 985 990

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly

1045 1050 1055

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

1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met

1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1185 1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1205 1210 1215

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245

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

1250 1255 1260

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1265 1270 1275 1280

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1285 1290 1295

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

1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu Met Thr Pro Asp

1410 1415 1420

Ile Ile Leu Gln Arg Thr Gly Ile Asp Val Arg Ala Val Glu Gln Gly

1425 1430 1435 1440

Asp Asp Ala Trp His Lys Leu Arg Leu Gly Val Ile Thr Ala Ser Glu

1445 1450 1455

Val His Asn Val Ile Ala Lys Pro Arg Ser Gly Lys Lys Trp Pro Asp

1460 1465 1470

Met Lys Met Ser Tyr Phe His Thr Leu Leu Ala Glu Val Cys Thr Gly

1475 1480 1485

Val Ala Pro Glu Val Asn Ala Lys Ala Leu Ala Trp Gly Lys Gln Tyr

1490 1495 1500

Glu Asn Asp Ala Arg Thr Leu Phe Glu Phe Thr Ser Gly Val Asn Val

1505 1510 1515 1520

Thr Glu Ser Pro Ile Ile Tyr Arg Asp Glu Ser Met Arg Thr Ala Cys

1525 1530 1535

Ser Pro Asp Gly Leu Cys Ser Asp Gly Asn Gly Leu Glu Leu Lys Cys

1540 1545 1550

Pro Phe Thr Ser Arg Asp Phe Met Lys Phe Arg Leu Gly Gly Phe Glu
1555 1560 1565

Ala Ile Lys Ser Ala Tyr Met Ala Gln Val Gln Tyr Ser Met Trp Val
1570 1575 1580

Thr Arg Lys Asn Ala Trp Tyr Phe Ala Asn Tyr Asp Pro Arg Met Lys
1585 1590 1595 1600

Arg Glu Gly Leu His Tyr Val Val Ile Glu Arg Asp Glu Lys Tyr Met
1605 1610 1615

Ala Ser Phe Asp Glu Ile Val Pro Glu Phe Ile Glu Lys Met Asp Glu
1620 1625 1630

Ala Leu Ala Glu Ile Gly Phe Val Phe Gly Glu Gln Trp Arg Leu Glu
1635 1640 1645

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1650 1655 1660

His His His His His His

1665 1670

<210> 36

<211> 5013

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9-lambda

<400> 36

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggcccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560

tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620

accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680

gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740

aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800

aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860

gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920

gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980

ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040

ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100

tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220

ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280

cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340

ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400

cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460
aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520
ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580
ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640
tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700
gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760
gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820

aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880
ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940
gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000
atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060
cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120
gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180
tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240

tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300
gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360
aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420
aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480
agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540
agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600
aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720
atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780
agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840
ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900
cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960
ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020
aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctcgag 4260

atgacaccgg acattatcct gcagcgtacc gggatcgatg tgagagctgt cgaacagggg 4320

gatgatgcgt ggcacaaatt acggctcggc gtcatcaccg cttcagaagt tcacaacgtg 4380

atagcaaaac cccgctccgg aaagaagtgg cctgacatga aaatgtccta cttccacacc 4440

ctgcttgctg aggtttgcac cggtgtggct ccggaagtta acgctaaagc actggcctgg 4500

ggaaaacagt acgagaacga cgccagaacc ctgtttgaat tcacttccgg cgtgaatgtt 4560

actgaatccc cgatcatcta tcgcgacgaa agtatgcgta ccgcctgctc tcccgatggt 4620

ttatgcagtg acggcaacgg ccttgaactg aaatgcccgt ttacctcccg ggatttcatg 4680

aagttccggc tcggtggttt cgaggccata aagtcagctt acatggccca ggtgcagtac 4740

agcatgtggg tgacgcgaaa aaatgcctgg tactttgcca actatgaccc gcgtatgaag 4800

cgtgaaggcc tgcattatgt cgtgattgag cgggatgaaa agtacatggc gagttttgac 4860

gagatcgtgc cggagttcat cgaaaaaatg gacgaggcac tggctgaaat tggttttgta 4920

tttggggagc aatggcgact cgagaaaagg ccggcggcca cgaaaaaggc cggccaggca 4980

aaaaagaaaa agcaccacca ccaccaccac tga 5013

<210> 37

<211> 1799

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9-mungbean

<400> 37

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu
85 90 95
Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr

100 105 110
Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125
Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His
130 135 140
Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg
145 150 155 160
His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys
165 170 175

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp
180 185 190
Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys
195 200 205
Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser
210 215 220
Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu
225 230 235 240
Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile

245 250 255
Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala
260 265 270
Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala
275 280 285
Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala
290 295 300
Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu
305 310 315 320

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

325 330 335

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

340 345 350

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

355 360 365

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

385 390 395 400

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu

405 410 415

Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu

420 425 430

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

435 440 445

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu

450 455 460

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

465 470 475 480

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

485 490 495

Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg

500 505 510

Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

515 520 525

Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp

530 535 540

Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr

545 550 555 560

Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr

565 570 575

Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala

580 585 590

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

595 600 605

Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu

610 615 620

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

625 630 635 640

Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu

645 650 655

Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu

660 665 670

Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe

675 680 685

Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp

690 695 700

Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser

705 710 715 720

Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg

725 730 735

Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp

740 745 750

Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His

755 760 765

Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln

770 775 780

Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys

785 790 795 800

Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln

805 810 815

Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly

820 825 830

Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
835 840 845

Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly
850 855 860

Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp
865 870 875 880

Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser
885 890 895

Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser
900 905 910

Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp
915 920 925

Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn
930 935 940

Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly
945 950 955 960

Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val
965 970 975

Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
980 985 990

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val
995 1000 1005

Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn
1010 1015 1020

Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr
1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly
1045 1050 1055

Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln
1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met
1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1185 1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1205 1210 1215

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245

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

1250 1255 1260

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1265 1270 1275 1280

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1285 1290 1295

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

1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Gln Met Gln Thr Leu

1410 1415 1420

Gln Met Ser Leu Leu Thr Gln Pro Tyr Val Gln Pro Arg Phe Pro Cys

1425 1430 1435 1440

Lys Arg Tyr Pro Thr Phe Ser Ala Ser Cys Arg Thr Gln Lys Thr Ala

1445 1450 1455

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

1460 1465 1470

Arg Cys Thr Gln Pro Leu Ser Glu Lys Lys Lys Arg Val Phe Phe Leu

1475 1480 1485

Asp Val Asn Pro Leu Cys Tyr Glu Gly Ser Lys Pro Ser Leu Arg Ser

1490 1495 1500

Phe Gly Arg Trp Leu Ser Leu Phe Leu His Gln Val Ser Leu Thr Asp

1505 1510 1515 1520

Pro Val Ile Ala Val Ile Asp Gly Glu Gly Gly Ser Glu His Arg Arg

1525 1530 1535

Lys Leu Leu Pro Ser Tyr Lys Ala His Arg Lys Lys Phe Met Arg His

1540 1545 1550

Met Ser Ser Gly His Val Gly Arg Ser His Gln Val Ile Asn Asp Val

1555 1560 1565

Leu Gly Lys Cys Asn Val Pro Val Ile Lys Val Ala Gly His Glu Ala

1570 1575 1580

Asp Asp Val Val Ala Thr Leu Ala Gly Gln Val Val Asn Lys Gly Phe

1585 1590 1595 1600

Arg Val Val Ile Gly Ser Pro Asp Lys Asp Phe Lys Gln Leu Ile Ser

1605 1610 1615

Glu Asp Val Gln Ile Val Met Pro Leu Pro Glu Leu Gln Arg Trp Ser

1620 1625 1630

Phe Tyr Thr Leu Arg His Tyr Arg Asp Gln Tyr Asn Cys Asp Pro Glu

1635 1640 1645

Ser Asp Leu Ser Phe Arg Cys Ile Val Gly Asp Glu Val Asp Gly Val

1650 1655 1660

Pro Gly Ile Gln His Leu Val Pro Ser Phe Gly Arg Lys Thr Ala Met

1665 1670 1675 1680

Lys Leu Ile Lys Lys His Gly Ser Leu Glu Thr Leu Leu Asn Ala Ala

1685 1690 1695

Ala Ile Arg Thr Val Gly Arg Pro Tyr Ala Gln Asp Ala Leu Lys Asn

1700 1705 1710

His Ala Asp Tyr Leu Arg Arg Asn Tyr Glu Val Leu Ala Leu Lys Arg

1715 1720 1725

Asp Val Asn Ile Gln Leu Tyr Asp Glu Trp Leu Val Lys Arg Asp Asn

1730 1735 1740

His Asn Asp Lys Thr Ala Leu Ser Ser Phe Phe Lys Tyr Leu Gly Glu

1745 1750 1755 1760

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

1765 1770 1775

Gln Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys

1780 1785 1790

Lys His His His His His His

1795

<210> 38

<211> 5400

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9-mungbean

<400> 38

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560

tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620

accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680

gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740

aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800
aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860
gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920
gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980
ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040
ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100
tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220
ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280
cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340
ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400
cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460
aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520
ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580

ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640
tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700
gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760
gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820
aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880
ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940
gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000

atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060
cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120
gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180
tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240
tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300
gagatccgca agcgccccct gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360
aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480

agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540

agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600

aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780

agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctgcag 4260

atgcaaacgt tacagatgag tctgttgaca caaccttacg ttcagcctcg tttcccttgc 4320

aagcgttacc cgaccttctc cgcatcctgc agaactcaaa agacagcgat cacgaaaaca 4380

gagaaggtgt ttttcagtga gtcatttgat caaacacgtt gcacgcagcc tctctcggaa 4440

aagaagaaga gggtgttctt tttggacgtt aacccgctct gttacgaagg aagcaagccc 4500

agcttgcgct ccttcgggcg gtggctctct ctgtttctcc atcaagtcag cctcactgac 4560

cccgtcattg ctgttattga tggagaagga ggcagcgagc atcgcagaaa gttgctacct 4620

tcatataaag cacataggaa aaagttcatg agacacatgt caagtggcca tgttgggagg 4680

tctcatcaag ttataaatga tgttcttgga aaatgcaacg tgccagttat aaaggttgct 4740

ggtcatgaag ctgatgatgt tgtagctact ctagctggac aagttgtcaa taaagggttt 4800

cgagtggtca ttggctcccc tgataaggat tttaagcagc ttatatctga agatgtgcaa 4860

atagttatgc ctttgccaga gttacaaagg tggtccttct acactctgag gcactacagg 4920

gatcagtata attgtgatcc agaatctgat ctgagcttta gatgcattgt aggtgatgaa 4980

gtagacggcg ttcctggtat ccagcatttg gtccctagtt ttggtcggaa gactgctatg 5040

aaacttatta aaaaacatgg ttccttggaa actttattaa atgcggctgc aataaggact 5100

gtaggcagac catatgcaca ggatgccctc aaaaaccatg ctgattacct tcggagaaac 5160

tatgaagttc ttgccttgaa aagggatgta aatatccaac tttatgatga gtggttggtt 5220

aagagagaca atcacaatga taaaactgca ctatcttcct tcttcaaata tttgggagaa 5280

agtaaggagc tcagttacaa tggcagacct atctcttaca atggtctgca gaaaaggccg 5340

gcggccacga aaaaggccgg ccaggcaaaa aagaaaaagc accaccacca ccaccactga 5400

5400

<210> 39

<211> 1720

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9-T5

<400> 39

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
35 40 45

Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
50 55 60

Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
65 70 75 80

Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu
85 90 95

Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr
100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His

130 135 140

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg
145 150 155 160

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys
165 170 175

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp

180 185 190

Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys

195 200 205

Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser

210 215 220

Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu

225 230 235 240

Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile

245 250 255

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

260 265 270

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

275 280 285

Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala

290 295 300

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

305 310 315 320

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

325 330 335

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

340 345 350

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

355 360 365

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

370 375 380

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

385 390 395 400

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu

405 410 415

Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu

420 425 430

Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg
435 440 445
Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu
450 455 460
His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp
465 470 475 480
Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr
485 490 495

Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg
500 505 510
Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp
515 520 525
Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp
530 535 540
Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr
545 550 555 560
Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr
565 570 575
Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala
580 585 590
Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln
595 600 605
Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
610 615 620
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
625 630 635 640

Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
645 650 655
Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu
660 665 670
Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe
675 680 685

Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp
690 695 700
Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser

705 710 715 720
Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg
725 730 735
Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp
740 745 750
Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His
755 760 765
Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
770 775 780

Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys
785 790 795 800
Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln
805 810 815
Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly
820 825 830
Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
835 840 845
Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly

850 855 860
Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp
865 870 875 880
Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser
885 890 895
Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser
900 905 910
Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp
915 920 925

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

930 935 940

Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly

945 950 955 960

Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val

965 970 975

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

980 985 990

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035 1040

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly

1045 1050 1055

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

1060 1065 1070

Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met

1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1125 1130 1135

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

1140 1145 1150

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1155 1160 1165

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1170 1175 1180

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1185 1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1205 1210 1215

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1220 1225 1230

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1235 1240 1245

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

1250 1255 1260

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1265 1270 1275 1280

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1285 1290 1295

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

1300 1305 1310

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1315 1320 1325

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1330 1335 1340

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1395 1400 1405

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Gln Met Ala Ser Arg

1410 1415 1420

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

1425 1430 1435 1440

His Asn Asn Ser Lys Lys Pro Phe Ala Ser Ser Tyr Val Ser Thr Ile
1445 1450 1455
Gln Ser Leu Ala Lys Ser Tyr Ser Ala Arg Thr Thr Ile Val Leu Gly
1460 1465 1470
Asp Lys Gly Lys Ser Val Phe Arg Leu Glu His Leu Pro Glu Tyr Lys
1475 1480 1485
Gly Asn Arg Asp Glu Lys Tyr Ala Gln Arg Thr Glu Glu Glu Lys Ala
1490 1495 1500

Leu Asp Glu Gln Phe Phe Glu Tyr Leu Lys Asp Ala Phe Glu Leu Cys
1505 1510 1515 1520
Lys Thr Thr Phe Pro Thr Phe Thr Ile Arg Gly Val Glu Ala Asp Asp
1525 1530 1535
Met Ala Ala Tyr Ile Val Lys Leu Ile Gly His Leu Tyr Asp His Val
1540 1545 1550
Trp Leu Ile Ser Thr Asp Gly Asp Trp Asp Thr Leu Leu Thr Asp Lys
1555 1560 1565
Val Ser Arg Phe Ser Phe Thr Thr Arg Arg Glu Tyr His Leu Arg Asp

1570 1575 1580
Met Tyr Glu His His Asn Val Asp Asp Val Glu Gln Phe Ile Ser Leu
1585 1590 1595 1600
Lys Ala Ile Met Gly Asp Leu Gly Asp Asn Ile Arg Gly Val Glu Gly
1605 1610 1615
Ile Gly Ala Lys Arg Gly Tyr Asn Ile Ile Arg Glu Phe Gly Asn Val
1620 1625 1630
Leu Asp Ile Ile Asp Gln Leu Pro Leu Pro Gly Lys Gln Lys Tyr Ile
1635 1640 1645

Gln Asn Leu Asn Ala Ser Glu Glu Leu Leu Phe Arg Asn Leu Ile Leu
1650 1655 1660
Val Asp Leu Pro Thr Tyr Cys Val Asp Ala Ile Ala Ala Val Gly Gln
1665 1670 1675 1680
Asp Val Leu Asp Lys Phe Thr Lys Asp Ile Leu Glu Ile Ala Glu Gln
1685 1690 1695

Leu Gln Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys

1700 1705 1710

Lys Lys His His His His His His

1715 1720

<210> 40

<211> 5163

<212> DNA

<213> Artificial Sequence

<220><223> SpyCas9-T5

<400> 40

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 180

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 240

ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 300

ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 360

aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 420

agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 480

caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 540

taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 600

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 660

cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 720

ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 780

ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 840

ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 900

ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 960

gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1020

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1080

cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1140

aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1200

aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1260

atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 1320

gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 1380

ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 1440

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggcccccctg 1500

gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 1560

tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 1620

accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 1680

gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 1740

aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 1800

aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 1860

gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 1920

gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 1980

ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2040

ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2100

tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2160

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2220

ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 2280

cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 2340

ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 2400

cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 2460

aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 2520

ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 2580

ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 2640

tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 2700

gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 2760

gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 2820

aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 2880

ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 2940

gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3000

atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3060

cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3120

gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3180

tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3240

tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 3300

gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 3360

aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 3420

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 3480

agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 3540

agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 3600

aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 3660

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 3720

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 3780

agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 3840

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 3900

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3960

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4020

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4080

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4140

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4200

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctgcag 4260

atggcttccc gtcgtaatct aatgattgtc gatggaacta acttaggctt tcgcttcaaa 4320

cataacaata gtaaaaaacc atttgcctca agttatgttt caactattca atctctggca 4380

aaatcctact ctgccagaac tacgattgtt ctaggtgata agggaaaatc tgtatttcgt 4440

ctagaacatc taccagagta taaaggtaat cgtgatgaaa agtacgcaca acgtacggaa 4500

gaggagaaag cgctagatga gcagttcttt gagtatttga aggatgcttt cgagttgtgt 4560

aaaactacat tcccaacttt taccattcgt ggtgtagaag cagacgatat ggcagcttat 4620

attgttaagc tcatcgggca tctttatgat cacgtttggc taatatctac agatggtgac 4680

tgggatactt tattaacgga taaagtttct cgtttttctt tcacaacacg tcgtgagtat 4740

catcttcgtg atatgtatga acatcataat gttgatgatg ttgagcagtt tatctccctg 4800

aaagcaatta tgggagatct aggagataat attcgtggtg ttgaaggaat aggagcaaaa 4860

cgcggatata atattattcg tgagtttggt aacgtactgg atattattga tcagcttcca 4920

ctgcctggaa agcagaaata tatacagaac ctgaatgcat cggaagaact gcttttccga 4980
aacttgattc tggttgattt acctacctac tgtgtggatg ctattgctgc tgtaggtcaa 5040
gatgtgttag ataagtttac aaaagatatt ttggagattg cagaacaact gcagaaaagg 5100
ccggcggcca cgaaaaaggc cggccaggca aaaaagaaaa agcaccacca ccaccaccac 5160
tga 5163

<210> 41
<211> 2021
<212> PRT
<213> Artificial Sequence
<220><223> RecJ-SpyCas9
<400> 41

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30
Gly Ser Val Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu
35 40 45
Thr Ala Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr
50 55 60
Ala Ser Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys

65 70 75 80
Gly Met Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu
85 90 95
Ile Leu Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly
100 105 110
Asp Phe Asp Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala
115 120 125
Met Arg Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg
130 135 140

Phe Glu Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His

145 150 155 160

Ala Arg Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser

165 170 175

His Ala Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val

180 185 190

Thr Asp His His Leu Pro Gly Glu Thr Leu Pro Ala Ala Glu Ala Ile

195 200 205

Ile Asn Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala

210 215 220

Gly Val Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu

225 230 235 240

Arg Asp Gln Gly Trp Phe Asp Glu Arg Gly Ile Ala Ile Pro Asn Leu

245 250 255

Ala Glu Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val

260 265 270

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

275 280 285

Ile Arg Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val

290 295 300

Ala Asn Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala

305 310 315 320

Leu Gly Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val

325 330 335

Gly Val Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu

340 345 350

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

355 360 365

Gly Met Gln Val Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser

370 375 380

Arg Asp Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His

385 390 395 400

Gln Gly Val Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His

405 410 415

Arg Pro Val Ile Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly

420 425 430

Ser Gly Arg Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg

435 440 445

Leu Asp Thr Leu Tyr Pro Gly Met Ile Leu Lys Phe Gly Gly His Ala

450 455 460

Met Ala Ala Gly Leu Ser Leu Glu Glu Asp Lys Phe Glu Leu Phe Gln

465 470 475 480

Gln Arg Phe Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu

485 490 495

Gln Gly Glu Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr

500 505 510

Met Glu Val Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met

515 520 525

Phe Pro Glu Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg

530 535 540

Leu Val Gly Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly

545 550 555 560

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

565 570 575

Pro Asp Asn Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile

580 585 590

Asn Glu Phe Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile

595 600 605

Trp Pro Ile Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg

610 615 620

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

625 630 635 640

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

645 650 655

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
660 665 670
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
675 680 685
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
690 695 700
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
705 710 715 720

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
725 730 735
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
740 745 750
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
755 760 765
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
770 775 780
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

785 790 795 800
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
805 810 815
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
820 825 830
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
835 840 845
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
850 855 860

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
865 870 875 880
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
885 890 895
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
900 905 910

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
915 920 925
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser

930 935 940
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
945 950 955 960
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
965 970 975
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
980 985 990
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
995 1000 1005

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
1010 1015 1020
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
1025 1030 1035 1040
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
1045 1050 1055
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
1060 1065 1070
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp

1075 1080 1085
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
1090 1095 1100
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
1105 1110 1115 1120
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
1125 1130 1135
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
1140 1145 1150

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln

1155 1160 1165

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

1170 1175 1180

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp

1185 1190 1195 1200

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

1205 1210 1215

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

1220 1225 1230

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr

1235 1240 1245

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala

1250 1255 1260

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr

1265 1270 1275 1280

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp

1285 1290 1295

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe

1300 1305 1310

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

1315 1320 1325

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

1330 1335 1340

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

1345 1350 1355 1360

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly

1365 1370 1375

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

1380 1385 1390

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile

1395 1400 1405

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro

1410 1415 1420

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu

1425 1430 1435 1440

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

1445 1450 1455

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys

1460 1465 1470

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

1475 1480 1485

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

1490 1495 1500

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

1505 1510 1515 1520

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp

1525 1530 1535

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr

1540 1545 1550

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp

1555 1560 1565

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

1570 1575 1580

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

1585 1590 1595 1600

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

1605 1610 1615

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe

1620 1625 1630

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

1635 1640 1645

Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser

1650 1655 1660

Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu
1665 1670 1675 1680
Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile
1685 1690 1695
Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser
1700 1705 1710
Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly
1715 1720 1725

Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
1730 1735 1740
Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser
1745 1750 1755 1760
Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly
1765 1770 1775
Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile
1780 1785 1790
Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala

1795 1800 1805
Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys
1810 1815 1820
Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser
1825 1830 1835 1840
Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr
1845 1850 1855
Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
1860 1865 1870

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His
1875 1880 1885
Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val
1890 1895 1900
Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys
1905 1910 1915 1920

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu
1925 1930 1935
Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp

1940 1945 1950
Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp
1955 1960 1965
Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1970 1975 1980
Asp Leu Ser Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu Lys
1985 1990 1995 2000
Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His
2005 2010 2015

His His His His His
2020

<210> 42
<211> 6066
<212> DNA
<213> Artificial Sequence
<220><223> RecJ-SpyCas9
<400> 42

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60
atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgtgaaaca acagatacaa 120
cttcgtcgcc gtgaagtcga tgaaacggca gacttgcccg ctgaattgcc tcccttgctg 180
cgccgtttat acgccagccg gggcgtgcgc agtgcgcaag aactggaacg cagtgttaaa 240
ggtatgttgc cctggcagca actgagcggc gtcgaaaagg ccgttgagat cctttacaac 300

gcttttcgcg aaggaacgcg gattattgtg gtcggcgatt ttgacgccga cggcgcgacc 360
agcacggctc taagcgtgct ggcgatgcgc tcgcttggtt gcagcaatat cgactatctg 420
gtaccaaacc gtttcgaaga cggttacggc ttaagcccgg aagtagtcga tcaggcccat 480
gcccgtggcg cgcagttaat tgtcacggtg gataacggta tttcctccca tgcgggcgtt 540
gaacacgctc gctcgttggg cattccggtt attgttaccg atcaccattt gccgggcgaa 600
acattacccg cagcggaagc gatcattaac cctaacttgc gcgactgtaa tttcccgtcg 660

aaatcactgg caggcgtggg tgtggcgttt tatctgatgc tggcgctgcg cacctttttg 720

cgcgatcagg gctggtttga tgagcgtggc atcgcaattc ctaacctggc agaactgctg 780

gatctggtcg cgctgggaac agtggcggac gtcgtgccgc tggacgctaa taatcgcatt 840

ctgacctggc aggggatgag tcgcatccgt gccggaaagt gccgtccagg gattaaagcg 900

ctgctggaag tggcaaaccg tgatgcacaa aaactcgccg ccagcgattt aggttttgcg 960

ctggggccac gtctcaatgc tgccggacga ctggacgata tgtccgtcgg tgtggcgctc 1020

ttgctgtgcg acaacatcgg cgaagcgcgc gtgctggcaa atgaactcga tgcgctaaac 1080

cagacgcgaa aagagatcga acaaggaatg caagttgaag ccctgaccct gtgcgagaaa 1140

ctggagcgaa gtcgcgacac gctacccggc gggctggcaa tgtatcaccc cgaatggcat 1200

cagggcgttg tcggtattct ggcttcgcgc atcaaagagc gttttcaccg tccggttatc 1260

gcctttgcgc cagcaggtga tggtacgctg aaaggttcag gtcgctccat tcaggggctg 1320

catatgcgtg atgcactgga gcgattagac acactctacc ctggcatgat actgaagttt 1380

ggcggtcatg cgatggcggc gggtttgtcg ctggaagagg ataaattcga actctttcaa 1440

caacggtttg gcgagctggt taccgagtgg ctggaccctt cgctattgca aggcgaagtg 1500

gtgtcagacg gcccgttaag cccggccgaa atgaccatgg aagtggcgca gctgctgcgc 1560

gatgctggcc cgtgggggca gatgttcccg gagccgctgt ttgatggtca tttccgtctg 1620

ctgcaacagc ggctggtggg cgaacgtcat ttgaaagtca tggtcgaacc ggtcggcggc 1680

ggtccgctgc tggatggtat tgcttttaat gtcgataccg ccctctggcc ggataacggc 1740

gtgcgcgaag tgcaactggc ttacaagctc gatatcaacg agtttcgcgg caaccgcagc 1800

ctgcaaatta tcatcgacaa tatctggcca attggatccg aattcgagct ccgtcgacaa 1860

gcttgcggcc gcatggacaa gaagtacagc atcggcctgg acatcggtac caacagcgtg 1920

ggctgggccg tgatcaccga cgagtacaag gtgcccagca agaagttcaa ggtgctgggc 1980

aacaccgacc gccacagcat caagaagaac ctgatcggcg ccctgctgtt cgacagcggc 2040

gagaccgccg aggccacccg cctgaagcgc accgcccgcc gccgctacac ccgccgcaag 2100

aaccgcatct gctacctgca ggagatcttc agcaacgaga tggccaaggt ggacgacagc 2160

ttcttccacc gcctggagga gagcttcctg gtggaggagg acaagaagca cgagcgccac 2220

cccatcttcg gcaacatcgt ggacgaggtg gcctaccacg agaagtaccc caccatctac 2280

cacctgcgca agaagctggt ggacagcacc gacaaggccg acctgcgcct gatctacctg 2340

gccctggccc acatgatcaa gttccgcggc cacttcctga tcgagggcga cctgaacccc 2400

gacaacagcg acgtggacaa gctgttcatc cagctggtgc agacctacaa ccagctgttc 2460

gaggagaacc ccatcaacgc cagcggcgtg gacgccaagg ccatcctgag cgcccgcctg 2520

agcaagagcc gccgcctgga gaacctgatc gcccagctgc ccggcgagaa gaagaacggc 2580

ctgttcggca acctgatcgc cctgagcctg ggcctgaccc ccaacttcaa gagcaacttc 2640

gacctggccg aggacgccaa gctgcagctg agcaaggaca cctacgacga cgacctggac 2700

aacctgctgg cccagatcgg cgaccagtac gccgacctgt tcctggccgc caagaacctg 2760

agcgacgcca tcctgctgag cgacatcctg cgcgtgaaca ccgagatcac caaggccccc 2820

ctgagcgcca gcatgatcaa gcgctacgac gagcaccacc aggacctgac cctgctgaag 2880

gccctggtgc gccagcagct gcccgagaag tacaaggaga tcttcttcga ccagagcaag 2940

aacggctacg ccggctacat cgacggcggc gccagccagg aggagttcta caagttcatc 3000

aagcccatcc tggagaagat ggacggcacc gaggagctgc tggtgaagct gaaccgcgag 3060

gacctgctgc gcaagcagcg caccttcgac aacggcagca tcccccacca gatccacctg 3120

ggcgagctgc acgccatcct gcgccgccag gaggacttct acccccttcct gaaggacaac 3180

cgcgagaaga tcgagaagat cctgaccttc cgcatccccct actacgtggg cccccctggcc 3240

cgcggcaaca gccgcttcgc ctggatgacc cgcaagagcg aggagaccat caccccctgg 3300

aacttcgagg aggtggtgga caagggcgcc agcgcccaga gcttcatcga gcgcatgacc 3360

aacttcgaca agaacctgcc caacgagaag gtgctgccca agcacagcct gctgtacgag 3420

tacttcaccg tgtacaacga gctgaccaag gtgaagtacg tgaccgaggg catgcgcaag 3480

cccgccttcc tgagcggcga gcagaagaag gccatcgtgg acctgctgtt caagaccaac 3540

cgcaaggtga ccgtgaagca gctgaaggag gactacttca agaagatcga gtgcttcgac 3600

agcgtggaga tcagcggcgt ggaggaccgc ttcaacgcca gcctgggcac ctaccacgac 3660

ctgctgaaga tcatcaagga caaggacttc ctggacaacg aggagaacga ggacatcctg 3720

gaggacatcg tgctgaccct gaccctgttc gaggaccgcg agatgatcga ggagcgcctg 3780

aagacctacg cccacctgtt cgacgacaag gtgatgaagc agctgaagcg ccgccgctac 3840

accggctggg gccgcctgag ccgcaagctt atcaacggca tccgcgacaa gcagagcggc 3900

aagaccatcc tggacttcct gaagagcgac ggcttcgcca accgcaactt catgcagctg 3960

atccacgacg acagcctgac cttcaaggag gacatccaga aggcccaggt gagcggccag 4020

ggcgacagcc tgcacgagca catcgccaac ctggccggca gccccgccat caagaagggc 4080

atcctgcaga ccgtgaaggt ggtggacgag ctggtgaagg tgatgggccg ccacaagccc 4140

gagaacatcg tgatcgagat ggcccgcgag aaccagacca cccagaaggg ccagaagaac 4200

agccgcgagc gcatgaagcg catcgaggag ggcatcaagg agctgggcag ccagatcctg 4260

aaggagcacc ccgtggagaa cacccagctg cagaacgaga agctgtacct gtactacctg 4320

cagaacggcc gcgacatgta cgtggaccag gagctggaca tcaaccgcct gagcgactac 4380

gacgtggacc acatcgtgcc ccagagcttc ctgaaggacg acagcatcga caacaaggtg 4440

ctgacccgca gcgacaagaa ccgcggcaag agcgacaacg tgcccagcga ggaggtggtg 4500

aagaagatga agaactactg gcgccagctg ctgaacgcca agctgatcac ccagcgcaag 4560

ttcgacaacc tgaccaaggc cgagcgcggc ggcctgagcg agctggacaa ggccggcttc 4620

atcaagcgcc agctggtgga gacccgccag atcaccaagc acgtggccca gatcctggac 4680

agccgcatga acaccaagta cgacgagaac gacaagctga tccgcgaggt gaaggtgatc 4740

accctgaaga gcaagctggt gagcgacttc cgcaaggact tccagttcta caaggtgcgc 4800

gagatcaaca actaccacca cgcccacgac gcctacctga acgccgtggt gggcaccgcc 4860

ctgatcaaga agtaccccaa gctggagagc gagttcgtgt acggcgacta caaggtgtac 4920

gacgtgcgca agatgatcgc caagagcgag caggagatcg gcaaggccac cgccaagtac 4980

ttcttctaca gcaacatcat gaacttcttc aagaccgaga tcaccctggc caacggcgag 5040

atccgcaagc gccccctgat cgagaccaac ggcgagaccg gcgagatcgt gtgggacaag 5100

ggccgcgact tcgccaccgt gcgcaaggtg ctgagcatgc cccaggtgaa catcgtgaag 5160

aagaccgagg tgcagaccgg cggcttcagc aaggagagca tcctgcccaa gcgcaacagc 5220

gacaagctga tcgcccgcaa gaaggactgg gaccccaaga agtacggcgg cttcgacagc 5280

cccaccgtgg cctacagcgt gctggtggtg gccaaggtgg agaagggcaa gagcaagaag 5340

ctgaagagcg tgaaggagct gctgggcatc accatcatgg agcgcagcag cttcgagaag 5400

aaccccatcg acttcctgga ggccaagggc tacaaggagg tgaagaagga cctgatcatc 5460

aagctgccca agtacagcct gttcgagctg gagaacggcc gcaagcgcat gctggccagc 5520

gccggcgagc tgcagaaggg caacgagctg gccctgccca gcaagtacgt gaacttcctg 5580

tacctggcca gccactacga gaagctgaag ggcagccccg aggacaacga gcagaagcag 5640

ctgttcgtgg agcagcacaa gcactacctg gacgagatca tcgagcagat cagcgagttc 5700

agcaagcgcg tgatcctggc cgacgccaac ctggacaagg tgctgagcgc ctacaacaag 5760

caccgcgaca agcccatccg cgagcaggcc gagaacatca tccacctgtt caccctgacc 5820

aacctgggcg cccccgccgc cttcaagtac ttcgacacca ccatcgaccg caagcgctac 5880

accagcacca aggaggtgct ggacgccacc ctgatccacc agagcatcac cggtctgtac 5940

gagacccgca tcgacctgag ccagctgggc ggcgacgcgg ccgcactcga cctcgagaaa 6000

aggccggcgg ccacgaaaaa ggccggccag gcaaaaaaga aaaagcacca ccaccaccac 6060

cactga 6066

<210> 43

<211> 2310

<212> PRT

<213> Artificial Sequence

<220><223> RecE-SpyCas9

<400> 43

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ser Thr Lys Pro Leu Phe Leu Leu Arg Lys Ala Lys Lys
35 40 45

Ser Ser Gly Glu Pro Asp Val Val Leu Trp Ala Ser Asn Asp Phe Glu
50 55 60

Ser Thr Cys Ala Thr Leu Asp Tyr Leu Ile Val Lys Ser Gly Lys Lys
65 70 75 80

Leu Ser Ser Tyr Phe Lys Ala Val Ala Thr Asn Phe Pro Val Val Asn
85 90 95

Asp Leu Pro Ala Glu Gly Glu Ile Asp Phe Thr Trp Ser Glu Arg Tyr
100 105 110

Gln Leu Ser Lys Asp Ser Met Thr Trp Glu Leu Lys Pro Gly Ala Ala
115 120 125

Pro Asp Asn Ala His Tyr Gln Gly Asn Thr Asn Val Asn Gly Glu Asp

130 135 140

Met Thr Glu Ile Glu Glu Asn Met Leu Leu Pro Ile Ser Gly Gln Glu
145 150 155 160

Leu Pro Ile Arg Trp Leu Ala Gln His Gly Ser Glu Lys Pro Val Thr
165 170 175

His Val Ser Arg Asp Gly Leu Gln Ala Leu His Ile Ala Arg Ala Glu
180 185 190

Glu Leu Pro Ala Val Thr Ala Leu Ala Val Ser His Lys Thr Ser Leu

195 200 205

Leu Asp Pro Leu Glu Ile Arg Glu Leu His Lys Leu Val Arg Asp Thr
210 215 220
Asp Lys Val Phe Pro Asn Pro Gly Asn Ser Asn Leu Gly Leu Ile Thr
225 230 235 240
Ala Phe Phe Glu Ala Tyr Leu Asn Ala Asp Tyr Thr Asp Arg Gly Leu
245 250 255
Leu Thr Lys Glu Trp Met Lys Gly Asn Arg Val Ser His Ile Thr Arg
260 265 270
Thr Ala Ser Gly Ala Asn Ala Gly Gly Gly Asn Leu Thr Asp Arg Gly

275 280 285
Glu Gly Phe Val His Asp Leu Thr Ser Leu Ala Arg Asp Val Ala Thr
290 295 300
Gly Val Leu Ala Arg Ser Met Asp Leu Asp Ile Tyr Asn Leu His Pro
305 310 315 320
Ala His Ala Lys Arg Ile Glu Glu Ile Ile Ala Glu Asn Lys Pro Pro
325 330 335
Phe Ser Val Phe Arg Asp Lys Phe Ile Thr Met Pro Gly Gly Leu Asp
340 345 350

Tyr Ser Arg Ala Ile Val Val Ala Ser Val Lys Glu Ala Pro Ile Gly
355 360 365
Ile Glu Val Ile Pro Ala His Val Thr Glu Tyr Leu Asn Lys Val Leu
370 375 380
Thr Glu Thr Asp His Ala Asn Pro Asp Pro Glu Ile Val Asp Ile Ala
385 390 395 400
Cys Gly Arg Ser Ser Ala Pro Met Pro Gln Arg Val Thr Glu Glu Gly
405 410 415
Lys Gln Asp Asp Glu Glu Lys Pro Gln Pro Ser Gly Thr Thr Ala Val

420 425 430
Glu Gln Gly Glu Ala Glu Thr Met Glu Pro Asp Ala Thr Glu His His
435 440 445

Gln Asp Thr Gln Pro Leu Asp Ala Gln Ser Gln Val Asn Ser Val Asp
450 455 460
Ala Lys Tyr Gln Glu Leu Arg Ala Glu Leu His Glu Ala Arg Lys Asn
465 470 475 480
Ile Pro Ser Lys Asn Pro Val Asp Asp Asp Lys Leu Leu Ala Ala Ser
485 490 495

Arg Gly Glu Phe Val Asp Gly Ile Ser Asp Pro Asn Asp Pro Lys Trp
500 505 510
Val Lys Gly Ile Gln Thr Arg Asp Cys Val Tyr Gln Asn Gln Pro Glu
515 520 525
Thr Glu Lys Thr Ser Pro Asp Met Asn Gln Pro Glu Pro Val Val Gln
530 535 540
Gln Glu Pro Glu Ile Ala Cys Asn Ala Cys Gly Gln Thr Gly Gly Asp
545 550 555 560
Asn Cys Pro Asp Cys Gly Ala Val Met Gly Asp Ala Thr Tyr Gln Glu

565 570 575
Thr Phe Asp Glu Glu Ser Gln Val Glu Ala Lys Glu Asn Asp Pro Glu
580 585 590
Glu Met Glu Gly Ala Glu His Pro His Asn Glu Asn Ala Gly Ser Asp
595 600 605
Pro His Arg Asp Cys Ser Asp Glu Thr Gly Glu Val Ala Asp Pro Val
610 615 620
Ile Val Glu Asp Ile Glu Pro Gly Ile Tyr Tyr Gly Ile Ser Asn Glu
625 630 635 640

Asn Tyr His Ala Gly Pro Gly Ile Ser Lys Ser Gln Leu Asp Asp Ile
645 650 655
Ala Asp Thr Pro Ala Leu Tyr Leu Trp Arg Lys Asn Ala Pro Val Asp
660 665 670
Thr Thr Lys Thr Lys Thr Leu Asp Leu Gly Thr Ala Phe His Cys Arg
675 680 685
Val Leu Glu Pro Glu Glu Phe Ser Asn Arg Phe Ile Val Ala Pro Glu
690 695 700

Phe Asn Arg Arg Thr Asn Ala Gly Lys Glu Glu Glu Lys Ala Phe Leu

705 710 715 720

Met Glu Cys Ala Ser Thr Gly Lys Thr Val Ile Thr Ala Glu Glu Gly

725 730 735

Arg Lys Ile Glu Leu Met Tyr Gln Ser Val Met Ala Leu Pro Leu Gly

740 745 750

Gln Trp Leu Val Glu Ser Ala Gly His Ala Glu Ser Ser Ile Tyr Trp

755 760 765

Glu Asp Pro Glu Thr Gly Ile Leu Cys Arg Cys Arg Pro Asp Lys Ile

770 775 780

Ile Pro Glu Phe His Trp Ile Met Asp Val Lys Thr Thr Ala Asp Ile

785 790 795 800

Gln Arg Phe Lys Thr Ala Tyr Tyr Asp Tyr Arg Tyr His Val Gln Asp

805 810 815

Ala Phe Tyr Ser Asp Gly Tyr Glu Ala Gln Phe Gly Val Gln Pro Thr

820 825 830

Phe Val Phe Leu Val Ala Ser Thr Thr Ile Glu Cys Gly Arg Tyr Pro

835 840 845

Val Glu Ile Phe Met Met Gly Glu Glu Ala Lys Leu Ala Gly Gln Gln

850 855 860

Glu Tyr His Arg Asn Leu Arg Thr Leu Ser Asp Cys Leu Asn Thr Asp

865 870 875 880

Glu Trp Pro Ala Ile Lys Thr Leu Ser Leu Pro Arg Trp Ala Lys Glu

885 890 895

Tyr Ala Asn Asp Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly

900 905 910

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

915 920 925

Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys

930 935 940

Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu

945 950 955 960

Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg

965 970 975

Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile

980 985 990

Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp

995 1000 1005

Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys

1010 1015 1020

Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala

1025 1030 1035 1040

Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val

1045 1050 1055

Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala

1060 1065 1070

His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn

1075 1080 1085

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

1090 1095 1100

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

1105 1110 1115 1120

Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu

1125 1130 1135

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

1140 1145 1150

Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn

1155 1160 1165

Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr

1170 1175 1180

Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala

1185 1190 1195 1200

Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser

1205 1210 1215

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

1220 1225 1230

Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu

1235 1240 1245

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

1250 1255 1260

Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala

1265 1270 1275 1280

Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met

1285 1290 1295

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

1300 1305 1310

Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His

1315 1320 1325

Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro

1330 1335 1340

Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg

1345 1350 1355 1360

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

1365 1370 1375

Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu

1380 1385 1390

Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met

1395 1400 1405

Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His

1410 1415 1420

Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val

1425 1430 1435 1440

Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu

1445 1450 1455

Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val
1460 1465 1470
Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe
1475 1480 1485
Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu
1490 1495 1500

Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu
1505 1510 1515 1520
Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu
1525 1530 1535
Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr
1540 1545 1550
Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg
1555 1560 1565
Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg

1570 1575 1580
Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly
1585 1590 1595 1600
Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr
1605 1610 1615
Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser
1620 1625 1630
Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys
1635 1640 1645

Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met
1650 1655 1660
Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn
1665 1670 1675 1680
Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg
1685 1690 1695
Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His
1700 1705 1710

Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr

1715 1720 1725

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

1730 1735 1740

Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu

1745 1750 1755 1760

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

1765 1770 1775

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

1780 1785 1790

Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg

1795 1800 1805

Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu

1810 1815 1820

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

1825 1830 1835 1840

Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr

1845 1850 1855

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

1860 1865 1870

Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val

1875 1880 1885

Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala

1890 1895 1900

Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu

1905 1910 1915 1920

Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala

1925 1930 1935

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr

1940 1945 1950

Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly

1955 1960 1965

Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu

1970 1975 1980

Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu

1985 1990 1995 2000

Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly

2005 2010 2015

Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu

2020 2025 2030

Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp

2035 2040 2045

Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys

2050 2055 2060

Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr

2065 2070 2075 2080

Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu

2085 2090 2095

Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro

2100 2105 2110

Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala

2115 2120 2125

Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys

2130 2135 2140

Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly

2145 2150 2155 2160

Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys

2165 2170 2175

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg

2180 2185 2190

Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn

2195 2200 2205

Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His

2210 2215 2220

Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe
2225 2230 2235 2240
Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu
2245 2250 2255
Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg
2260 2265 2270
Ile Asp Leu Ser Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu
2275 2280 2285
Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

2290 2295 2300
His His His His His His
2305 2310

<210> 44
<211> 6933
<212> DNA
<213> Artificial Sequence
<220><223> RecE-SpyCas9
<400> 44

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60
atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatgagcac aaaaccactc 120
ttcctgttac ggaaagcgaa aaaatcatcc ggtgaacctg acgtcgtcct gtgggcaagc 180
aacgattttg aatcgacctg tgccactctg gactacctga tcgttaagtc aggtaaaaaa 240

ctgagcagct attttaaagc tgttgccacg aattttcctg tcgttaatga cctgcccgct 300
gaaggtgaga tcgattttac ctggagtgaa cgctatcaac tcagcaaaga ctccatgaca 360
tgggaactaa aaccgggagc agcaccagac aacgctcact atcaaggcaa taccaacgtc 420
aacggcgaag acatgactga gattgaggag aatatgctac tcccaatttc tggccaggaa 480
ctgcccattc gttggcttgc tcaacacggc agcgaaaaac cggtaacgca cgtttcacgc 540
gacggactcc aggcattaca cattgctcgg gctgaagaac taccggctgt tactgccctg 600
gctgtttccc acaaaaccag cctgctcgac ccgctggaaa ttcgcgaact ccacaaactg 660

gttcgtgaca ctgacaaagt tttccctaat cctggtaatt caaacctggg actgataact 720

gctttttcg aagcatacct gaacgctgac tacaccgatc gaggactgct gacaaaagag 780

tggatgaagg gtaatcgtgt ttcacacatc actcgcacgg cttccggtgc taatgctggc 840

ggcggaaacc tcaccgatcg cggcgaaggt ttcgtacacg atctgacgtc actggcgcgc 900

gacgtagcca ctggcgtact ggcccgttca atggatctgg acatctataa ccttcatccg 960

gcacacgcta aacgcattga ggaaattatc gctgaaaata aaccgccctt ttctgttttc 1020

cgcgacaaat tcatcaccat gcctggcggg ctggattatt cccgcgccat cgtggttgcg 1080

tccgtaaaag aagcaccaat tgggatcgag gtcatccccg cgcacgtcac tgaatatctg 1140

aacaaagtac tgactgaaac cgatcatgcc aaccctgatc cggaaatcgt ggatattgcc 1200

tgcggtcgct cctctgcccc gatgccgcag cgagtaacag aagaaggaaa acaggatgat 1260

gaagaaaaac cgcaaccatc tggaacaacg gcagttgaac agggagaggc tgaaacaatg 1320

gaaccggacg caactgaaca tcatcaggac acgcagccgc tggatgctca gtcacaggta 1380

aattctgttg atgcgaaata tcaggaactg cgggcagaac tccatgaagc ccggaaaaac 1440

attccatcaa aaaatcctgt cgatgacgat aaattgcttg ctgcatcacg tggtgaattt 1500

gttgacggaa ttagcgaccc gaacgatccg aaatgggtaa aggggatcca gactcgcgat 1560

tgtgtgtacc agaaccagcc agaaacggaa aaaaccagcc cagatatgaa tcaacctgag 1620

ccagtagtgc aacaggaacc ggaaatagcc tgcaatgcct gcggccagac tggcggggat 1680

aactgccctg actgtggtgc ggtgatgggc gacgcaacat accaggaaac attcgatgaa 1740

gagagtcagg ttgaagctaa ggaaaatgat ccggaggaaa tggaaggcgc tgaacatccg 1800

cacaatgaga atgctggcag cgatccgcat cgcgattgca gtgatgaaac tggcgaagtc 1860

gcagatcccg taatcgtaga agacatagag ccaggtattt attacggaat ttcgaatgag 1920

aattaccacg cgggtcccgg tatcagtaag tctcagctcg atgacattgc tgatactccg 1980

gcactatatt tgtggcgtaa aaatgccccc gtggacacca caaagacaaa aacgctcgat 2040

ttaggaactg ctttccactg ccgggtactt gaaccggaag aattcagtaa ccgctttatc 2100

gtagcacctg aatttaaccg ccgtacaaac gccggaaaag aagaagagaa agcgtttctg 2160

atggaatgcg caagcacagg aaaaacggtt atcactgcgg aagaaggccg gaaaattgaa 2220

ctcatgtatc aaagcgttat ggctttgccg ctggggcaat ggcttgttga aagcgccgga 2280

cacgctgaat catcaattta ctgggaagat cctgaaacag gaattttgtg tcggtgccgt 2340

ccggacaaaa ttatccctga atttcactgg atcatggacg tgaaaactac ggcggatatt 2400

caacgattca aaaccgctta ttacgactac cgctatcacg ttcaggatgc attctacagt 2460

gacggttatg aagcacagtt tggagtgcag ccaactttcg tttttctggt tgccagcaca 2520

actattgaat gcggacgtta tccggttgaa attttcatga tgggcgaaga agcaaaactg 2580

gcaggtcaac aggaatatca ccgcaatctg cgaaccctgt ctgactgcct gaataccgat 2640

gaatggccag ctattaagac attatcactg ccccgctggg ctaaggaata tgcaaatgac 2700

ggatccgaat tcgagctccg tcgacaagct tgcggccgca tggacaagaa gtacagcatc 2760

ggcctggaca tcggtaccaa cagcgtgggc tgggccgtga tcaccgacga gtacaaggtg 2820

cccagcaaga agttcaaggt gctgggcaac accgaccgcc acagcatcaa gaagaacctg 2880

atcggcgccc tgctgttcga cagcggcgag accgccgagg ccacccgcct gaagcgcacc 2940

gcccgccgcc gctacacccg ccgcaagaac cgcatctgct acctgcagga gatcttcagc 3000

aacgagatgg ccaaggtgga cgacagcttc ttccaccgcc tggaggagag cttcctggtg 3060

gaggaggaca agaagcacga gcgccacccc atcttcggca acatcgtgga cgaggtggcc 3120

taccacgaga agtaccccac catctaccac ctgcgcaaga agctggtgga cagcaccgac 3180

aaggccgacc tgcgcctgat ctacctggcc ctggcccaca tgatcaagtt ccgcggccac 3240

ttcctgatcg agggcgacct gaaccccgac aacagcgacg tggacaagct gttcatccag 3300

ctggtgcaga cctacaacca gctgttcgag gagaacccca tcaacgccag cggcgtggac 3360

gccaaggcca tcctgagcgc ccgcctgagc aagagccgcc gcctggagaa cctgatcgcc 3420

cagctgcccg gcgagaagaa gaacggcctg ttcggcaacc tgatcgccct gagcctgggc 3480

ctgaccccca acttcaagag caacttcgac ctggccgagg acgccaagct gcagctgagc 3540

aaggacacct acgacgacga cctggacaac ctgctggccc agatcggcga ccagtacgcc 3600

gacctgttcc tggccgccaa gaacctgagc gacgccatcc tgctgagcga catcctgcgc 3660

gtgaacaccg agatcaccaa ggccccctg agcgccagca tgatcaagcg ctacgacgag 3720

caccaccagg acctgaccct gctgaaggcc ctggtgcgcc agcagctgcc cgagaagtac 3780

aaggagatct tcttcgacca gagcaagaac ggctacgccg gctacatcga cggcggcgcc 3840

agccaggagg agttctacaa gttcatcaag cccatcctgg agaagatgga cggcaccgag 3900

gagctgctgg tgaagctgaa ccgcgaggac ctgctgcgca agcagcgcac cttcgacaac 3960

ggcagcatcc cccaccagat ccacctgggc gagctgcacg ccatcctgcg ccgccaggag 4020

gacttctacc ccttcctgaa ggacaaccgc gagaagatcg agaagatcct gaccttccgc 4080

atcccctact acgtgggccc cctggcccgc ggcaacagcc gcttcgcctg gatgacccgc 4140

aagagcgagg agaccatcac cccctggaac ttcgaggagg tggtggacaa gggcgccagc 4200

gcccagagct tcatcgagcg catgaccaac ttcgacaaga acctgcccaa cgagaaggtg 4260

ctgcccaagc acagcctgct gtacgagtac ttcaccgtgt acaacgagct gaccaaggtg 4320

aagtacgtga ccgagggcat gcgcaagccc gccttcctga gcggcgagca gaagaaggcc 4380

atcgtggacc tgctgttcaa gaccaaccgc aaggtgaccg tgaagcagct gaaggaggac 4440

tacttcaaga agatcgagtg cttcgacagc gtggagatca gcggcgtgga ggaccgcttc 4500
aacgccagcc tgggcaccta ccacgacctg ctgaagatca tcaaggacaa ggacttcctg 4560
gacaacgagg agaacgagga catcctggag gacatcgtgc tgaccctgac cctgttcgag 4620
gaccgcgaga tgatcgagga gcgcctgaag acctacgccc acctgttcga cgacaaggtg 4680
atgaagcagc tgaagcgccg ccgctacacc ggctggggcc gcctgagccg caagcttatc 4740
aacggcatcc gcgacaagca gagcggcaag accatcctgg acttcctgaa gagcgacggc 4800
ttcgccaacc gcaacttcat gcagctgatc cacgacgaca gcctgacctt caaggaggac 4860

atccagaagg cccaggtgag cggccagggc gacagcctgc acgagcacat cgccaacctg 4920
gccggcagcc ccgccatcaa gaagggcatc ctgcagaccg tgaaggtggt ggacgagctg 4980
gtgaaggtga tgggccgcca caagcccgag aacatcgtga tcgagatggc ccgcgagaac 5040
cagaccaccc agaagggcca gaagaacagc cgcgagcgca tgaagcgcat cgaggagggc 5100
atcaaggagc tgggcagcca gatcctgaag gagcaccccg tggagaacac ccagctgcag 5160
aacgagaagc tgtacctgta ctacctgcag aacggccgcg acatgtacgt ggaccaggag 5220
ctggacatca accgcctgag cgactacgac gtggaccaca tcgtgcccca gagcttcctg 5280

aaggacgaca gcatcgacaa caaggtgctg acccgcagcg acaagaaccg cggcaagagc 5340
gacaacgtgc ccagcgagga ggtggtgaag aagatgaaga actactggcg ccagctgctg 5400
aacgccaagc tgatcaccca gcgcaagttc gacaacctga ccaaggccga gcgcggcggc 5460
ctgagcgagc tggacaaggc cggcttcatc aagcgccagc tggtggagac ccgccagatc 5520
accaagcacg tggcccagat cctggacagc cgcatgaaca ccaagtacga cgagaacgac 5580
aagctgatcc gcgaggtgaa ggtgatcacc ctgaagagca agctggtgag cgacttccgc 5640
aaggacttcc agttctacaa ggtgcgcgag atcaacaact accaccacgc ccacgacgcc 5700

tacctgaacg ccgtggtggg caccgccctg atcaagaagt accccaagct ggagagcgag 5760
ttcgtgtacg gcgactacaa ggtgtacgac gtgcgcaaga tgatcgccaa gagcgagcag 5820
gagatcggca aggccaccgc caagtacttc ttctacagca acatcatgaa cttcttcaag 5880
accgagatca ccctggccaa cggcgagatc cgcaagcgcc ccctgatcga gaccaacggc 5940
gagaccggcg agatcgtgtg ggacaagggc cgcgacttcg ccaccgtgcg caaggtgctg 6000
agcatgcccc aggtgaacat cgtgaagaag accgaggtgc agaccggcgg cttcagcaag 6060
gagagcatcc tgcccaagcg caacagcgac aagctgatcg cccgcaagaa ggactgggac 6120

cccaagaagt acggcggctt cgacagcccc accgtggcct acagcgtgct ggtggtggcc 6180

aaggtggaga agggcaagag caagaagctg aagagcgtga aggagctgct gggcatcacc 6240

atcatggagc gcagcagctt cgagaagaac cccatcgact tcctggaggc caagggctac 6300

aaggaggtga agaaggacct gatcatcaag ctgcccaagt acagcctgtt cgagctggag 6360

aacggccgca agcgcatgct ggccagcgcc ggcgagctgc agaagggcaa cgagctggcc 6420

ctgcccagca agtacgtgaa cttcctgtac ctggccagcc actacgagaa gctgaagggc 6480

agccccgagg acaacgagca gaagcagctg ttcgtggagc agcacaagca ctacctggac 6540

gagatcatcg agcagatcag cgagttcagc aagcgcgtga tcctggccga cgccaacctg 6600

gacaaggtgc tgagcgccta caacaagcac cgcgacaagc ccatccgcga gcaggccgag 6660

aacatcatcc acctgttcac cctgaccaac ctgggcgccc ccgccgcctt caagtacttc 6720

gacaccacca tcgaccgcaa gcgctacacc agcaccaagg aggtgctgga cgccaccctg 6780

atccaccaga gcatcaccgg tctgtacgag acccgcatcg acctgagcca gctgggcggc 6840

gacgcggccg cactcgacct cgagaaaagg ccggcggcca cgaaaaaggc cggccaggca 6900

aaaaagaaaa agcaccacca ccaccaccac tga 6933

<210> 45

<211> 1682

<212> PRT

<213> Artificial Sequence

<220><223> GFP-SpyCas9

<400> 45

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
35 40 45

Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser

50 55 60

Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe
65 70 75 80

Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr
85 90 95

Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met
100 105 110
Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln
115 120 125

Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala
130 135 140
Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys
145 150 155 160
Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu
165 170 175
Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys
180 185 190
Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly
195 200 205
Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp
210 215 220
Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala
225 230 235 240
Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu
245 250 255
Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
260 265 270

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys
275 280 285
Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala
290 295 300
Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu
305 310 315 320
Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu
325 330 335
Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr

340 345 350

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln

355 360 365

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His

370 375 380

Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg

385 390 395 400

His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys

405 410 415

Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp

420 425 430

Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys

435 440 445

Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser

450 455 460

Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu

465 470 475 480

Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile

485 490 495

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

500 505 510

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

515 520 525

Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala

530 535 540

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

545 550 555 560

Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu

565 570 575

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg

580 585 590

Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys

595 600 605

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val

610 615 620

Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser

625 630 635 640

Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu

645 650 655

Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu

660 665 670

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

675 680 685

Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu

690 695 700

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

705 710 715 720

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

725 730 735

Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg

740 745 750

Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp

755 760 765

Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp

770 775 780

Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr

785 790 795 800

Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr

805 810 815

Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala

820 825 830

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

835 840 845

Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu
850 855 860
Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His
865 870 875 880
Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu
885 890 895
Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu
900 905 910
Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe

915 920 925
Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp
930 935 940
Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser
945 950 955 960
Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg
965 970 975
Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp
980 985 990

Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His
995 1000 1005
Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
1010 1015 1020
Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys
1025 1030 1035 1040
Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln
1045 1050 1055
Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly

1060 1065 1070
Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
1075 1080 1085
Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly
1090 1095 1100

Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp
1105 1110 1115 1120
Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser
1125 1130 1135

Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser
1140 1145 1150
Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp
1155 1160 1165
Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn
1170 1175 1180
Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly
1185 1190 1195 1200
Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val

1205 1210 1215
Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
1220 1225 1230
Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val
1235 1240 1245
Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn
1250 1255 1260
Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr
1265 1270 1275 1280

Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly
1285 1290 1295
Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln
1300 1305 1310
Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met
1315 1320 1325
Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys
1330 1335 1340
Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp

1345 1350 1355 1360

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln

1365 1370 1375

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

1380 1385 1390

Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys

1395 1400 1405

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val

1410 1415 1420

Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys

1425 1430 1435 1440

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg

1445 1450 1455

Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1460 1465 1470

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1475 1480 1485

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

1490 1495 1500

Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe

1505 1510 1515 1520

Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp

1525 1530 1535

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

1540 1545 1550

Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala

1555 1560 1565

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp

1570 1575 1580

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu

1585 1590 1595 1600

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

1605 1610 1615

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

1620 1625 1630

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser

1635 1640 1645

Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu Lys Arg Pro Ala

1650 1655 1660

Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His His

1665 1670 1675 1680

His His

<210> 46

<211> 5049

<212> DNA

<213> Artificial Sequence

<220><223> GFP-SpyCas9

<400> 46

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatgagtaa aggagaagaa 120

cttttcactg gagttgtccc aattcttgtt gaattagatg gtgatgttaa tgggcacaaa 180

ttttctgtca gtggagaggg tgaaggtgat gcaacatacg gaaaacttac ccttaaattt 240

atttgcacta ctggaaaact acctgttcca tggccaacac ttgtcactac tttctcttat 300

ggtgttcaat gcttttcaag atacccagat catatgaagc ggcacgactt cttcaagagc 360

gccatgcctg agggatacgt gcaggagagg accatctctt tcaaggacga cgggaactac 420

aagacacgtg ctgaagtcaa gtttgaggga gacaccctcg tcaacaggat cgagcttaag 480

ggaatcgatt tcaaggagga cggaaacatc ctcggccaca agttggaata caactacaac 540

tcccacaacg tatacatcac ggcagacaaa caaaagaatg gaatcaaagc taacttcaaa 600

attagacaca acattgaaga tggaagcgtt caactagcag accattatca acaaaatact 660

ccaattggcg atggccctgt ccttttacca gacaaccatt acctgtccac acaatctgcc 720

ctttcgaaag atcccaacga aaagagagac cacatggtcc ttcttgagtt tgtaacagct 780

gctgggatta cacatggcat ggatgaacta tacaaaggat ccgaattcga gctccgtcga 840

caagcttgcg gccgcatgga caagaagtac agcatcggcc tggacatcgg taccaacagc 900

gtgggctggg ccgtgatcac cgacgagtac aaggtgccca gcaagaagtt caaggtgctg 960
ggcaacaccg accgccacag catcaagaag aacctgatcg gcgccctgct gttcgacagc 1020
ggcgagaccg ccgaggccac ccgcctgaag cgcaccgccc gccgccgcta cacccgccgc 1080
aagaaccgca tctgctacct gcaggagatc ttcagcaacg agatggccaa ggtggacgac 1140
agcttcttcc accgcctgga ggagagcttc ctggtggagg aggacaagaa gcacgagcgc 1200
caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta ccccaccatc 1260
taccacctgc gcaagaagct ggtggacagc accgacaagg ccgacctgcg cctgatctac 1320

ctggccctgg cccacatgat caagttccgc ggccacttcc tgatcgaggg cgacctgaac 1380
cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg 1440
ttcgaggaga accccatcaa cgccagcggc gtggacgcca aggccatcct gagcgcccgc 1500
ctgagcaaga gccgccgcct ggagaacctg atcgcccagc tgcccggcga gaagaagaac 1560
ggcctgttcg gcaacctgat cgccctgagc ctgggcctga cccccaactt caagagcaac 1620
ttcgacctgg ccgaggacgc caagctgcag ctgagcaagg acacctacga cgacgacctg 1680
gacaacctgc tggcccagat cggcgaccag tacgccgacc tgttcctggc cgccaagaac 1740

ctgagcgacg ccatcctgct gagcgacatc ctgcgcgtga acaccgagat caccaaggcc 1800
cccctgagcg ccagcatgat caagcgctac gacgagcacc accaggacct gaccctgctg 1860
aaggccctgg tgcgccagca gctgcccgag aagtacaagg agatcttctt cgaccagagc 1920
aagaacggct acgccggcta catcgacggc ggcgccagcc aggaggagtt ctacaagttc 1980
atcaagccca tcctggagaa gatggacggc accgaggagc tgctggtgaa gctgaaccgc 2040
gaggacctgc tgcgcaagca gcgcaccttc gacaacggca gcatccccca ccagatccac 2100
ctgggcgagc tgcacgccat cctgcgccgc caggaggact tctacccctt cctgaaggac 2160

aaccgcgaga agatcgagaa gatcctgacc ttccgcatcc cctactacgt gggcccccctg 2220
gcccgcggca acagccgctt cgcctggatg acccgcaaga gcgaggagac catcaccccc 2280
tggaacttcg aggaggtggt ggacaagggc gccagcgccc agagcttcat cgagcgcatg 2340
accaacttcg acaagaacct gcccaacgag aaggtgctgc ccaagcacag cctgctgtac 2400
gagtacttca ccgtgtacaa cgagctgacc aaggtgaagt acgtgaccga gggcatgcgc 2460
aagcccgcct tcctgagcgg cgagcagaag aaggccatcg tggacctgct gttcaagacc 2520
aaccgcaagg tgaccgtgaa gcagctgaag gaggactact tcaagaagat cgagtgcttc 2580

gacagcgtgg agatcagcgg cgtggaggac cgcttcaacg ccagcctggg cacctaccac 2640

gacctgctga agatcatcaa ggacaaggac ttcctggaca acgaggagaa cgaggacatc 2700

ctggaggaca tcgtgctgac cctgaccctg ttcgaggacc gcgagatgat cgaggagcgc 2760

ctgaagacct acgcccacct gttcgacgac aaggtgatga agcagctgaa gcgccgccgc 2820

tacaccggct ggggccgcct gagccgcaag cttatcaacg gcatccgcga caagcagagc 2880

ggcaagacca tcctggactt cctgaagagc gacggcttcg ccaaccgcaa cttcatgcag 2940

ctgatccacg acgacagcct gaccttcaag gaggacatcc agaaggccca ggtgagcggc 3000

cagggcgaca gcctgcacga gcacatcgcc aacctggccg gcagccccgc catcaagaag 3060

ggcatcctgc agaccgtgaa ggtggtggac gagctggtga aggtgatggg ccgccacaag 3120

cccgagaaca tcgtgatcga gatggcccgc gagaaccaga ccacccagaa gggccagaag 3180

aacagccgcg agcgcatgaa gcgcatcgag gagggcatca aggagctggg cagccagatc 3240

ctgaaggagc accccgtgga gaacacccag ctgcagaacg agaagctgta cctgtactac 3300

ctgcagaacg gccgcgacat gtacgtggac caggagctgg acatcaaccg cctgagcgac 3360

tacgacgtgg accacatcgt gccccagagc ttcctgaagg acgacagcat cgacaacaag 3420

gtgctgaccc gcagcgacaa gaaccgcggc aagagcgaca acgtgcccag cgaggaggtg 3480

gtgaagaaga tgaagaacta ctggcgccag ctgctgaacg ccaagctgat cacccagcgc 3540

aagttcgaca acctgaccaa ggccgagcgc ggcggcctga gcgagctgga caaggccggc 3600

ttcatcaagc gccagctggt ggagacccgc cagatcacca agcacgtggc ccagatcctg 3660

gacagccgca tgaacaccaa gtacgacgag aacgacaagc tgatccgcga ggtgaaggtg 3720

atcaccctga agagcaagct ggtgagcgac ttccgcaagg acttccagtt ctacaaggtg 3780

cgcgagatca acaactacca ccacgcccac gacgcctacc tgaacgccgt ggtgggcacc 3840

gccctgatca agaagtaccc caagctggag agcgagttcg tgtacggcga ctacaaggtg 3900

tacgacgtgc gcaagatgat cgccaagagc gagcaggaga tcggcaaggc caccgccaag 3960

tacttcttct acagcaacat catgaacttc ttcaagaccg agatcaccct ggccaacggc 4020

gagatccgca agcgcccccc gatcgagacc aacggcgaga ccggcgagat cgtgtgggac 4080

aagggccgcg acttcgccac cgtgcgcaag gtgctgagca tgccccaggt gaacatcgtg 4140

aagaagaccg aggtgcagac cggcggcttc agcaaggaga gcatcctgcc caagcgcaac 4200

agcgacaagc tgatcgcccg caagaaggac tgggacccca agaagtacgg cggcttcgac 4260

agccccaccg tggcctacag cgtgctggtg gtggccaagg tggagaaggg caagagcaag 4320

aagctgaaga gcgtgaagga gctgctgggc atcaccatca tggagcgcag cagcttcgag 4380

aagaacccca tcgacttcct ggaggccaag ggctacaagg aggtgaagaa ggacctgatc 4440

atcaagctgc ccaagtacag cctgttcgag ctggagaacg gccgcaagcg catgctggcc 4500

agcgccggcg agctgcagaa gggcaacgag ctggccctgc ccagcaagta cgtgaacttc 4560

ctgtacctgg ccagccacta cgagaagctg aagggcagcc ccgaggacaa cgagcagaag 4620

cagctgttcg tggagcagca caagcactac ctggacgaga tcatcgagca gatcagcgag 4680

ttcagcaagc gcgtgatcct ggccgacgcc aacctggaca aggtgctgag cgcctacaac 4740

aagcaccgcg acaagcccat ccgcgagcag gccgagaaca tcatccacct gttcaccctg 4800

accaacctgg gcgcccccgc cgccttcaag tacttcgaca ccaccatcga ccgcaagcgc 4860

tacaccagca ccaaggaggt gctggacgcc accctgatcc accagagcat caccggtctg 4920

tacgagaccc gcatcgacct gagccagctg ggcggcgacg cggccgcact cgacctcgag 4980

aaaaggccgg cggccacgaa aaaggccggc caggcaaaaa agaaaaagca ccaccaccac 5040

caccactga 5049

<210> 47

<211> 1560

<212> PRT

<213> Artificial Sequence

<220><223> SSB-SpyCas9

<400> 47

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ala Ser Arg Gly Val Asn Lys Val Ile Leu Val Gly Asn
35 40 45

Leu Gly Gln Asp Pro Glu Val Arg Tyr Met Pro Asn Gly Gly Ala Val
50 55 60

Ala Asn Ile Thr Leu Ala Thr Ser Glu Ser Trp Arg Asp Lys Ala Thr
65 70 75 80

Gly Glu Met Lys Glu Gln Thr Glu Trp His Arg Val Val Leu Phe Gly
85 90 95

Lys Leu Ala Glu Val Ala Ser Glu Tyr Leu Arg Lys Gly Ser Gln Val
100 105 110

Tyr Ile Glu Gly Gln Leu Arg Thr Arg Lys Trp Thr Asp Gln Ser Gly

115 120 125

Gln Asp Arg Tyr Thr Thr Glu Val Val Val Asn Val Gly Gly Thr Met
130 135 140
Gln Met Leu Gly Gly Arg Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala
145 150 155 160
Cys Gly Arg Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr
165 170 175
Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser
180 185 190
Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys

195 200 205
Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala
210 215 220
Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn
225 230 235 240
Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val
245 250 255
Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu
260 265 270

Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu
275 280 285
Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys
290 295 300
Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala
305 310 315 320
Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp
325 330 335
Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val

340 345 350
Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly
355 360 365

Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg
370 375 380
Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu
385 390 395 400
Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys
405 410 415

Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp
420 425 430
Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln
435 440 445
Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu
450 455 460
Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu
465 470 475 480
Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr
485 490 495
Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu
500 505 510
Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly
515 520 525
Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu
530 535 540
Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp
545 550 555 560

Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln
565 570 575
Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe
580 585 590
Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr
595 600 605
Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg
610 615 620

Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn

625 630 635 640

Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu

645 650 655

Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro

660 665 670

Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr

675 680 685

Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser

690 695 700

Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg

705 710 715 720

Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu

725 730 735

Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala

740 745 750

Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp

755 760 765

Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu

770 775 780

Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys

785 790 795 800

Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg

805 810 815

Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly

820 825 830

Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser

835 840 845

Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser

850 855 860

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

865 870 875 880

Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile

885 890 895

Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys

900 905 910

Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg

915 920 925

Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met

930 935 940

Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys

945 950 955 960

Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu

965 970 975

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

980 985 990

Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser

995 1000 1005

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

1010 1015 1020

Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys

1025 1030 1035 1040

Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr

1045 1050 1055

Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser

1060 1065 1070

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

1075 1080 1085

Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr

1090 1095 1100

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

1105 1110 1115 1120

Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr

1125 1130 1135

Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu
1140 1145 1150
Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu
1155 1160 1165
Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met
1170 1175 1180
Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe
1185 1190 1195 1200
Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala

1205 1210 1215
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr
1220 1225 1230
Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys
1235 1240 1245
Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln
1250 1255 1260
Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp
1265 1270 1275 1280

Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly
1285 1290 1295
Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val
1300 1305 1310
Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly
1315 1320 1325
Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe
1330 1335 1340
Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys

1345 1350 1355 1360
Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met
1365 1370 1375

Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro

1380 1385 1390

Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu

1395 1400 1405

Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln

1410 1415 1420

His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser

1425 1430 1435 1440

Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala

1445 1450 1455

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile

1460 1465 1470

Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys

1475 1480 1485

Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu

1490 1495 1500

Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu

1505 1510 1515 1520

Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp

1525 1530 1535

Leu Glu Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys

1540 1545 1550

Lys Lys His His His His His His

1555 1560

<210> 48

<211> 4683

<212> DNA

<213> Artificial Sequence

<220><223> SSB-SpyCas9

<400> 48

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggccag cagaggcgta 120

aacaaggtta ttctcgttgg taatctgggt caggacccgg aagtacgcta catgccaaat 180

ggtggcgcag ttgccaacat tacgctggct acttccgaat cctggcgtga taaagcgacc 240

ggcgagatga aagaacagac tgaatggcac cgcgttgtgc tgttcggcaa actggcagaa 300

gtggcgagcg aatatctgcg taaaggttct caggtttata tcgaaggtca gctgcgtacc 360

cgtaaatgga ccgatcaatc cggtcaggat cgctacacca cagaagtcgt ggtgaacgtt 420

ggcggcacca tgcagatgct gggtggtcgt ggatccgaat tcgagctccg tcgacaagct 480

tgcggccgca tggacaagaa gtacagcatc ggcctggaca tcggtaccaa cagcgtgggc 540

tgggccgtga tcaccgacga gtacaaggtg cccagcaaga agttcaaggt gctgggcaac 600

accgaccgcc acagcatcaa gaagaacctg atcggcgccc tgctgttcga cagcggcgag 660

accgccgagg ccacccgcct gaagcgcacc gcccgccgcc gctacacccg ccgcaagaac 720

cgcatctgct acctgcagga gatcttcagc aacgagatgg ccaaggtgga cgacagcttc 780

ttccaccgcc tggaggagag cttcctggtg gaggaggaca agaagcacga gcgccacccc 840

atcttcggca acatcgtgga cgaggtggcc taccacgaga agtaccccac catctaccac 900

ctgcgcaaga agctggtgga cagcaccgac aaggccgacc tgcgcctgat ctacctggcc 960

ctggcccaca tgatcaagtt ccgcggccac ttcctgatcg agggcgacct gaaccccgac 1020

aacagcgacg tggacaagct gttcatccag ctggtgcaga cctacaacca gctgttcgag 1080

gagaacccca tcaacgccag cggcgtggac gccaaggcca tcctgagcgc ccgcctgagc 1140

aagagccgcc gcctggagaa cctgatcgcc cagctgcccg gcgagaagaa gaacggcctg 1200

ttcggcaacc tgatcgccct gagcctgggc ctgaccccca acttcaagag caacttcgac 1260

ctggccgagg acgccaagct gcagctgagc aaggacacct acgacgacga cctggacaac 1320

ctgctggccc agatcggcga ccagtacgcc gacctgttcc tggccgccaa gaacctgagc 1380

gacgccatcc tgctgagcga catcctgcgc gtgaacaccg agatcaccaa ggcccccctg 1440

agcgccagca tgatcaagcg ctacgacgag caccaccagg acctgaccct gctgaaggcc 1500

ctggtgcgcc agcagctgcc cgagaagtac aaggagatct tcttcgacca gagcaagaac 1560

ggctacgccg gctacatcga cggcggcgcc agccaggagg agttctacaa gttcatcaag 1620

cccatcctgg agaagatgga cggcaccgag gagctgctgg tgaagctgaa ccgcgaggac 1680

ctgctgcgca agcagcgcac cttcgacaac ggcagcatcc cccaccagat ccacctgggc 1740

gagctgcacg ccatcctgcg ccgccaggag gacttctacc ccttcctgaa ggacaaccgc 1800

gagaagatcg agaagatcct gaccttccgc atcccctact acgtgggccc cctggcccgc 1860

ggcaacagcc gcttcgcctg gatgacccgc aagagcgagg agaccatcac cccctggaac 1920

ttcgaggagg tggtggacaa gggcgccagc gcccagagct tcatcgagcg catgaccaac 1980

ttcgacaaga acctgcccaa cgagaaggtg ctgcccaagc acagcctgct gtacgagtac 2040

ttcaccgtgt acaacgagct gaccaaggtg aagtacgtga ccgagggcat gcgcaagccc 2100
gccttcctga gcggcgagca gaagaaggcc atcgtggacc tgctgttcaa gaccaaccgc 2160
aaggtgaccg tgaagcagct gaaggaggac tacttcaaga agatcgagtg cttcgacagc 2220
gtggagatca gcggcgtgga ggaccgcttc aacgccagcc tgggcaccta ccacgacctg 2280
ctgaagatca tcaaggacaa ggacttcctg gacaacgagg agaacgagga catcctggag 2340
gacatcgtgc tgaccctgac cctgttcgag gaccgcgaga tgatcgagga gcgcctgaag 2400
acctacgccc acctgttcga cgacaaggtg atgaagcagc tgaagcgccg ccgctacacc 2460

ggctggggcc gcctgagccg caagcttatc aacggcatcc gcgacaagca gagcggcaag 2520
accatcctgg acttcctgaa gagcgacggc ttcgccaacc gcaacttcat gcagctgatc 2580
cacgacgaca gcctgacctt caaggaggac atccagaagg cccaggtgag cggccagggc 2640
gacagcctgc acgagcacat cgccaacctg gccggcagcc ccgccatcaa gaagggcatc 2700
ctgcagaccg tgaaggtggt ggacgagctg gtgaaggtga tgggccgcca caagcccgag 2760
aacatcgtga tcgagatggc ccgcgagaac cagaccaccc agaagggcca gaagaacagc 2820
cgcgagcgca tgaagcgcat cgaggagggc atcaaggagc tgggcagcca gatcctgaag 2880

gagcaccccg tggagaacac ccagctgcag aacgagaagc tgtacctgta ctacctgcag 2940
aacggccgcg acatgtacgt ggaccaggag ctggacatca accgcctgag cgactacgac 3000
gtggaccaca tcgtgcccca gagcttcctg aaggacgaca gcatcgacaa caaggtgctg 3060
acccgcagcg acaagaaccg cggcaagagc gacaacgtgc ccagcgagga ggtggtgaag 3120
aagatgaaga actactggcg ccagctgctg aacgccaagc tgatcaccca gcgcaagttc 3180
gacaacctga ccaaggccga gcgcggcggc ctgagcgagc tggacaaggc cggcttcatc 3240
aagcgccagc tggtggagac ccgccagatc accaagcacg tggcccagat cctggacagc 3300

cgcatgaaca ccaagtacga cgagaacgac aagctgatcc gcgaggtgaa ggtgatcacc 3360
ctgaagagca agctggtgag cgacttccgc aaggacttcc agttctacaa ggtgcgcgag 3420
atcaacaact accaccacgc ccacgacgcc tacctgaacg ccgtggtggg caccgccctg 3480
atcaagaagt accccaagct ggagagcgag ttcgtgtacg gcgactacaa ggtgtacgac 3540
gtgcgcaaga tgatcgccaa gagcgagcag gagatcggca aggccaccgc caagtacttc 3600
ttctacagca acatcatgaa cttcttcaag accgagatca ccctggccaa cggcgagatc 3660
cgcaagcgcc ccctgatcga gaccaacggc gagaccggcg agatcgtgtg ggacaagggc 3720

cgcgacttcg ccaccgtgcg caaggtgctg agcatgcccc aggtgaacat cgtgaagaag 3780

accgaggtgc agaccggcgg cttcagcaag gagagcatcc tgcccaagcg caacagcgac 3840

aagctgatcg cccgcaagaa ggactgggac cccaagaagt acggcggctt cgacagcccc 3900

accgtggcct acagcgtgct ggtggtggcc aaggtggaga agggcaagag caagaagctg 3960

aagagcgtga aggagctgct gggcatcacc atcatggagc gcagcagctt cgagaagaac 4020

cccatcgact tcctggaggc caagggctac aaggaggtga agaaggacct gatcatcaag 4080

ctgcccaagt acagcctgtt cgagctggag aacggccgca agcgcatgct ggccagcgcc 4140

ggcgagctgc agaagggcaa cgagctggcc ctgcccagca agtacgtgaa cttcctgtac 4200

ctggccagcc actacgagaa gctgaagggc agccccgagg acaacgagca gaagcagctg 4260

ttcgtggagc agcacaagca ctacctggac gagatcatcg agcagatcag cgagttcagc 4320

aagcgcgtga tcctggccga cgccaacctg gacaaggtgc tgagcgccta caacaagcac 4380

cgcgacaagc ccatccgcga gcaggccgag aacatcatcc acctgttcac cctgaccaac 4440

ctgggcgccc ccgccgcctt caagtacttc gacaccacca tcgaccgcaa gcgctacacc 4500

agcaccaagg aggtgctgga cgccaccctg atccaccaga gcatcaccgg tctgtacgag 4560

acccgcatcg acctgagcca gctgggcggc gacgcggccg cactcgacct cgagaaaagg 4620

ccggcggcca cgaaaaaggc cggccaggca aaaaagaaaa agcaccacca ccaccaccac 4680

tga 4683

<210> 49

<211> 2139

<212> PRT

<213> Artificial Sequence

<220><223> SSB-SpyCas9-RecJ

<400> 49

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Met Ala Ser Arg Gly Val Asn Lys Val Ile Leu Val Gly Asn
35 40 45

Leu Gly Gln Asp Pro Glu Val Arg Tyr Met Pro Asn Gly Gly Ala Val
50 55 60

Ala Asn Ile Thr Leu Ala Thr Ser Glu Ser Trp Arg Asp Lys Ala Thr

65 70 75 80

Gly Glu Met Lys Glu Gln Thr Glu Trp His Arg Val Val Leu Phe Gly

85 90 95

Lys Leu Ala Glu Val Ala Ser Glu Tyr Leu Arg Lys Gly Ser Gln Val

100 105 110

Tyr Ile Glu Gly Gln Leu Arg Thr Arg Lys Trp Thr Asp Gln Ser Gly

115 120 125

Gln Asp Arg Tyr Thr Thr Glu Val Val Val Asn Val Gly Gly Thr Met

130 135 140

Gln Met Leu Gly Gly Arg Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala

145 150 155 160

Cys Gly Arg Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr

165 170 175

Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser

180 185 190

Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys

195 200 205

Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala

210 215 220

Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn

225 230 235 240

Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val

245 250 255

Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu

260 265 270

Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu

275 280 285

Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys

290 295 300

Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala

305 310 315 320

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

Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu
385 390 395 400
Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys
405 410 415
Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp
420 425 430
Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln
435 440 445
Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu

450 455 460
Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu
465 470 475 480
Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr
485 490 495
Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu
500 505 510
Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly
515 520 525

Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu
530 535 540
Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp
545 550 555 560
Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln
565 570 575

Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe
580 585 590
Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr

595 600 605
Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg
610 615 620
Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn
625 630 635 640
Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu
645 650 655
Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro
660 665 670

Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr
675 680 685
Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser
690 695 700
Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg
705 710 715 720
Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu
725 730 735
Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala

740 745 750
Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp
755 760 765
Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu
770 775 780
Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys
785 790 795 800
Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg
805 810 815

Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly

820 825 830

Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser

835 840 845

Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser

850 855 860

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

865 870 875 880

Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile

885 890 895

Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys

900 905 910

Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg

915 920 925

Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met

930 935 940

Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys

945 950 955 960

Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu

965 970 975

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

980 985 990

Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser

995 1000 1005

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

1010 1015 1020

Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys

1025 1030 1035 1040

Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr

1045 1050 1055

Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser

1060 1065 1070

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

1075 1080 1085

Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr

1090 1095 1100

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

1105 1110 1115 1120

Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr

1125 1130 1135

Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu

1140 1145 1150

Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu

1155 1160 1165

Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met

1170 1175 1180

Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe

1185 1190 1195 1200

Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala

1205 1210 1215

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr

1220 1225 1230

Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys

1235 1240 1245

Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln

1250 1255 1260

Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp

1265 1270 1275 1280

Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly

1285 1290 1295

Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val

1300 1305 1310

Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly

1315 1320 1325

Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe
1330 1335 1340
Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys
1345 1350 1355 1360
Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met
1365 1370 1375
Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro
1380 1385 1390

Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu
1395 1400 1405
Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln
1410 1415 1420
His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser
1425 1430 1435 1440
Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
1445 1450 1455
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile

1460 1465 1470
Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys
1475 1480 1485
Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu
1490 1495 1500
Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu
1505 1510 1515 1520
Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp
1525 1530 1535

Leu Gln Val Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu
1540 1545 1550
Thr Ala Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr
1555 1560 1565
Ala Ser Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys
1570 1575 1580

Gly Met Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu
1585 1590 1595 1600
Ile Leu Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly

1605 1610 1615
Asp Phe Asp Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala
1620 1625 1630
Met Arg Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg
1635 1640 1645
Phe Glu Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His
1650 1655 1660
Ala Arg Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser
1665 1670 1675 1680

His Ala Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val
1685 1690 1695
Thr Asp His His Leu Pro Gly Asp Thr Leu Pro Ala Ala Glu Ala Ile
1700 1705 1710
Ile Asn Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala
1715 1720 1725
Gly Val Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu
1730 1735 1740
Arg Asp Gln Gly Trp Phe Asp Glu Arg Asn Ile Ala Ile Pro Asn Leu

1745 1750 1755 1760
Ala Glu Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val
1765 1770 1775
Pro Leu Asp Ala Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg
1780 1785 1790
Ile Arg Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val
1795 1800 1805
Ala Asn Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala
1810 1815 1820

Leu Gly Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val

1825 1830 1835 1840

Gly Val Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu

1845 1850 1855

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

1860 1865 1870

Gly Met Gln Ile Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser

1875 1880 1885

Arg Asp Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His

1890 1895 1900

Gln Gly Val Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His

1905 1910 1915 1920

Arg Pro Val Ile Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly

1925 1930 1935

Ser Gly Arg Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg

1940 1945 1950

Leu Asp Thr Leu Tyr Pro Gly Met Met Leu Lys Phe Gly Gly His Ala

1955 1960 1965

Met Ala Ala Gly Leu Ser Leu Glu Glu Asp Lys Phe Lys Leu Phe Gln

1970 1975 1980

Gln Arg Phe Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu

1985 1990 1995 2000

Gln Gly Glu Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr

2005 2010 2015

Met Glu Val Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met

2020 2025 2030

Phe Pro Glu Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg

2035 2040 2045

Leu Val Gly Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly

2050 2055 2060

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

2065 2070 2075 2080

Pro Asp Asn Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile

2085 2090 2095

Asn Glu Phe Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile

2100 2105 2110

Trp Pro Ile Leu Gln Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln

2115 2120 2125

Ala Lys Lys Lys Lys His His His His His His

2130 2135

<210> 50

<211> 6420

<212> DNA

<213> Artificial Sequence

<220><223> SSB-SpyCas9-RecJ

<400> 50

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggccag cagaggcgta 120

aacaaggtta ttctcgttgg taatctgggt caggacccgg aagtacgcta catgccaaat 180

ggtggcgcag ttgccaacat tacgctggct acttccgaat cctggcgtga taaagcgacc 240

ggcgagatga aagaacagac tgaatggcac cgcgttgtgc tgttcggcaa actggcagaa 300

gtggcgagcg aatatctgcg taaaggttct caggtttata tcgaaggtca gctgcgtacc 360

cgtaaatgga ccgatcaatc cggtcaggat cgctacacca cagaagtcgt ggtgaacgtt 420

ggcggcacca tgcagatgct gggtggtcgt ggatccgaat tcgagctccg tcgacaagct 480

tgcggccgca tggacaagaa gtacagcatc ggcctggaca tcggtaccaa cagcgtgggc 540

tgggccgtga tcaccgacga gtacaaggtg cccagcaaga agttcaaggt gctgggcaac 600

accgaccgcc acagcatcaa gaagaacctg atcggcgccc tgctgttcga cagcggcgag 660

accgccgagg ccacccgcct gaagcgcacc gcccgccgcc gctacacccg ccgcaagaac 720

cgcatctgct acctgcagga gatcttcagc aacgagatgg ccaaggtgga cgacagcttc 780

ttccaccgcc tggaggagag cttcctggtg gaggaggaca agaagcacga gcgccacccc 840

atcttcggca acatcgtgga cgaggtggcc taccacgaga agtaccccac catctaccac 900

ctgcgcaaga agctggtgga cagcaccgac aaggccgacc tgcgcctgat ctacctggcc 960

ctggcccaca tgatcaagtt ccgcggccac ttcctgatcg agggcgacct gaaccccgac 1020

aacagcgacg tggacaagct gttcatccag ctggtgcaga cctacaacca gctgttcgag 1080

gagaacccca tcaacgccag cggcgtggac gccaaggcca tcctgagcgc ccgcctgagc 1140

aagagccgcc gcctggagaa cctgatcgcc cagctgcccg gcgagaagaa gaacggcctg 1200

ttcggcaacc tgatcgccct gagcctgggc ctgaccccca acttcaagag caacttcgac 1260

ctggccgagg acgccaagct gcagctgagc aaggacacct acgacgacga cctggacaac 1320

ctgctggccc agatcggcga ccagtacgcc gacctgttcc tggccgccaa gaacctgagc 1380

gacgccatcc tgctgagcga catcctgcgc gtgaacaccg agatcaccaa ggccccccctg 1440

agcgccagca tgatcaagcg ctacgacgag caccaccagg acctgaccct gctgaaggcc 1500

ctggtgcgcc agcagctgcc cgagaagtac aaggagatct tcttcgacca gagcaagaac 1560

ggctacgccg gctacatcga cggcggcgcc agccaggagg agttctacaa gttcatcaag 1620

cccatcctgg agaagatgga cggcaccgag gagctgctgg tgaagctgaa ccgcgaggac 1680

ctgctgcgca agcagcgcac cttcgacaac ggcagcatcc cccaccagat ccacctgggc 1740

gagctgcacg ccatcctgcg ccgccaggag gacttctacc ccttcctgaa ggacaaccgc 1800

gagaagatcg agaagatcct gaccttccgc atcccctact acgtgggccc cctggcccgc 1860

ggcaacagcc gcttcgcctg gatgacccgc aagagcgagg agaccatcac cccctggaac 1920

ttcgaggagg tggtggacaa gggcgccagc gcccagagct tcatcgagcg catgaccaac 1980

ttcgacaaga acctgcccaa cgagaaggtg ctgcccaagc acagcctgct gtacgagtac 2040

ttcaccgtgt acaacgagct gaccaaggtg aagtacgtga ccgagggcat gcgcaagccc 2100

gccttcctga gcggcgagca gaagaaggcc atcgtggacc tgctgttcaa gaccaaccgc 2160

aaggtgaccg tgaagcagct gaaggaggac tacttcaaga agatcgagtg cttcgacagc 2220

gtggagatca gcggcgtgga ggaccgcttc aacgccagcc tgggcaccta ccacgacctg 2280

ctgaagatca tcaaggacaa ggacttcctg gacaacgagg agaacgagga catcctggag 2340

gacatcgtgc tgaccctgac cctgttcgag gaccgcgaga tgatcgagga gcgcctgaag 2400

acctacgccc acctgttcga cgacaaggtg atgaagcagc tgaagcgccg ccgctacacc 2460

ggctggggcc gcctgagccg caagcttatc aacggcatcc gcgacaagca gagcggcaag 2520

accatcctgg acttcctgaa gagcgacggc ttcgccaacc gcaacttcat gcagctgatc 2580

cacgacgaca gcctgacctt caaggaggac atccagaagg cccaggtgag cggccagggc 2640

gacagcctgc acgagcacat cgccaacctg gccggcagcc ccgccatcaa gaagggcatc 2700

ctgcagaccg tgaaggtggt ggacgagctg gtgaaggtga tgggccgcca caagcccgag 2760

aacatcgtga tcgagatggc ccgcgagaac cagaccaccc agaagggcca gaagaacagc 2820

cgcgagcgca tgaagcgcat cgaggagggc atcaaggagc tgggcagcca gatcctgaag 2880

gagcaccccg tggagaacac ccagctgcag aacgagaagc tgtacctgta ctacctgcag 2940

aacggccgcg acatgtacgt ggaccaggag ctggacatca accgcctgag cgactacgac 3000

gtggaccaca tcgtgcccca gagcttcctg aaggacgaca gcatcgacaa caaggtgctg 3060

acccgcagcg acaagaaccg cggcaagagc gacaacgtgc ccagcgagga ggtggtgaag 3120

aagatgaaga actactggcg ccagctgctg aacgccaagc tgatcaccca gcgcaagttc 3180

gacaacctga ccaaggccga gcgcggcggc ctgagcgagc tggacaaggc cggcttcatc 3240

aagcgccagc tggtggagac ccgccagatc accaagcacg tggcccagat cctggacagc 3300

cgcatgaaca ccaagtacga cgagaacgac aagctgatcc gcgaggtgaa ggtgatcacc 3360

ctgaagagca agctggtgag cgacttccgc aaggacttcc agttctacaa ggtgcgcgag 3420

atcaacaact accaccacgc ccacgacgcc tacctgaacg ccgtggtggg caccgccctg 3480

atcaagaagt accccaagct ggagagcgag ttcgtgtacg gcgactacaa ggtgtacgac 3540

gtgcgcaaga tgatcgccaa gagcgagcag gagatcggca aggccaccgc caagtacttc 3600

ttctacagca acatcatgaa cttcttcaag accgagatca ccctggccaa cggcgagatc 3660

cgcaagcgcc ccctgatcga gaccaacggc gagaccggcg agatcgtgtg ggacaagggc 3720

cgcgacttcg ccaccgtgcg caaggtgctg agcatgcccc aggtgaacat cgtgaagaag 3780

accgaggtgc agaccggcgg cttcagcaag gagagcatcc tgcccaagcg caacagcgac 3840

aagctgatcg cccgcaagaa ggactgggac cccaagaagt acggcggctt cgacagcccc 3900

accgtggcct acagcgtgct ggtggtggcc aaggtggaga agggcaagag caagaagctg 3960

aagagcgtga aggagctgct gggcatcacc atcatggagc gcagcagctt cgagaagaac 4020

cccatcgact tcctggaggc caagggctac aaggaggtga agaaggacct gatcatcaag 4080

ctgcccaagt acagcctgtt cgagctggag aacggccgca agcgcatgct ggccagcgcc 4140

ggcgagctgc agaagggcaa cgagctggcc ctgcccagca agtacgtgaa cttcctgtac 4200

ctggccagcc actacgagaa gctgaagggc agccccgagg acaacgagca gaagcagctg 4260

ttcgtggagc agcacaagca ctacctggac gagatcatcg agcagatcag cgagttcagc 4320

aagcgcgtga tcctggccga cgccaacctg gacaaggtgc tgagcgccta caacaagcac 4380

cgcgacaagc ccatccgcga gcaggccgag aacatcatcc acctgttcac cctgaccaac 4440

ctgggcgccc ccgccgcctt caagtacttc gacaccacca tcgaccgcaa gcgctacacc 4500

agcaccaagg aggtgctgga cgccaccctg atccaccaga gcatcaccgg tctgtacgag 4560

acccgcatcg acctgagcca gctgggcggc gacgcggccg cactcgacct gcaggtgaaa 4620

caacagatac aacttcgtcg ccgtgaagtc gatgaaacgg cagacttgcc cgctgaattg 4680

cctcccttgc tgcgccgttt atacgccagc cggggagtac gcagtgcgca agaactggaa 4740

cgcagtgtta aaggtatgct gccctggcag caactgagcg gcgtcgaaaa ggccgttgag 4800

atcctttaca acgcttttcg cgaaggaacg cggattattg tggtcggtga tttcgacgcc 4860
gacggcgcga ccagcacggc tctaagcgtg ctggcgatgc gctcgcttgg ttgcagcaat 4920
atcgactacc tggtaccaaa ccgtttcgaa gacggttacg gcttaagccc ggaagtggtc 4980
gatcaggccc atgcccgtgg cgcgcagtta attgtcacgg tggataacgg tatttcctcc 5040
catgcggggg ttgagcacgc tcgctcgttg ggcatcccgg ttattgttac cgatcaccat 5100
ttgccaggcg acacattacc cgcagcggaa gcgatcatta accctaactt gcgcgactgt 5160
aatttcccgt cgaaatcact ggcaggcgtg ggtgtggcgt tttatctgat gctggcgctg 5220

cgcacctttt tgcgcgatca gggctggttt gatgagcgta acatcgcaat tcctaacctg 5280
gcagaactgc tggatctggt cgcgctgggg acagtggcgg acgtcgtgcc gctggacgct 5340
aataatcgca ttctgacctg gcaggggatg agtcgcatcc gagccggaaa gtgccgtccg 5400
gggattaaag cgctgcttga agtggcaaac cgtgatgcac aaaaactcgc cgccagcgat 5460
ttaggttttg cgctggggcc acgtctcaat gctgccggac gactggacga tatgtccgtc 5520
ggtgtggcgc tgttgttgtg cgacaacatc ggcgaagcgc gcgtgctggc aaatgaactc 5580
gatgcgctaa accagacgcg aaaagagatc gaacaaggaa tgcaaattga agccctgacc 5640

ctgtgcgaga aactggagcg cagccgtgac acgctacccg gcgggctggc aatgtatcac 5700
cccgaatggc atcagggcgt tgtcggtatt ctggcttcgc gcatcaaaga gcgttttcac 5760
cgtccggtta tcgcgtttgc gccagcaggt gacggtacgc tgaaaggttc cggtcgctcc 5820
attcaggggc tgcatatgcg tgatgcgctg gagcgattag acacactcta ccctggcatg 5880
atgctgaagt ttggcggtca tgcgatggcg gcgggtttgt cgctggaaga ggataaattc 5940
aaactctttc aacaacggtt tggcgaactg gttactgagt ggctggaccc ttcgctattg 6000
caaggcgaag tggtatcaga cggtccgtta agcccggccg aaatgaccat ggaagtggcg 6060

cagctgctgc gcgatgctgg cccgtggggg cagatgttcc cggagccgct gtttgacggt 6120
catttccgtc tgctgcaaca gcggctggtg ggcgaacgtc atttgaaggt gatggtcgaa 6180
ccggtcggcg gcggtccact gctggatggt attgctttta atgtcgatac cgccctctgg 6240
ccggataacg gcgtgcgcga agtgcaactg gcttataagc tcgatatcaa cgagtttcgc 6300
ggcaaccgca gcctgcaaat tatcatcgac aatatctggc caattctgca gaaaaggccg 6360
gcggccacga aaaaggccgg ccaggcaaaa aagaaaaagc accaccacca ccaccactga 6420
6420

<210> 51

<211> 1533

<212> PRT

<213> Artificial Sequence

<220><223> DSB-SpyCas9

<400> 51

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ala Lys Lys Glu Met Val Glu Phe Asp Glu Ala Ile His
35 40 45

Gly Glu Asp Leu Ala Lys Phe Ile Lys Glu Ala Ser Asp His Lys Leu
50 55 60

Lys Ile Ser Gly Tyr Asn Glu Leu Ile Lys Asp Ile Arg Ile Arg Ala
65 70 75 80

Lys Asp Glu Leu Gly Val Asp Gly Lys Met Phe Asn Arg Leu Leu Ala
85 90 95

Leu Tyr His Lys Asp Asn Arg Asp Val Phe Glu Ala Glu Thr Glu Glu
100 105 110

Val Val Glu Leu Tyr Asp Thr Val Phe Ser Lys Gly Ser Glu Phe Glu
115 120 125

Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys Lys Tyr Ser Ile Gly
130 135 140

Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu
145 150 155 160

Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg
165 170 175

His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly
180 185 190

Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr
195 200 205

Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn
210 215 220
Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser
225 230 235 240
Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly
245 250 255
Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr
260 265 270

His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg
275 280 285
Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe
290 295 300
Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu
305 310 315 320
Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro
325 330 335
Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu

340 345 350
Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu
355 360 365
Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu
370 375 380
Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu
385 390 395 400
Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala
405 410 415

Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu
420 425 430
Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile
435 440 445
Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His
450 455 460

His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro
465 470 475 480
Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala

485 490 495
Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile
500 505 510
Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys
515 520 525
Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
530 535 540
Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg
545 550 555 560

Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile
565 570 575
Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala
580 585 590
Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr
595 600 605
Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala
610 615 620
Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn

625 630 635 640
Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val
645 650 655
Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys
660 665 670
Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu
675 680 685
Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr
690 695 700

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu

705 710 715 720

Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile

725 730 735

Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu

740 745 750

Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile

755 760 765

Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met

770 775 780

Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg

785 790 795 800

Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu

805 810 815

Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu

820 825 830

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

835 840 845

Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala

850 855 860

Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val

865 870 875 880

Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val

885 890 895

Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn

900 905 910

Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly

915 920 925

Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn

930 935 940

Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val

945 950 955 960

Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His

965 970 975

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

980 985 990

Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser

995 1000 1005

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn

1010 1015 1020

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

1025 1030 1035 1040

Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln

1045 1050 1055

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

1060 1065 1070

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

1075 1080 1085

Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys

1090 1095 1100

Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala

1105 1110 1115 1120

His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys

1125 1130 1135

Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr

1140 1145 1150

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

1155 1160 1165

Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr

1170 1175 1180

Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu

1185 1190 1195 1200

Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe

1205 1210 1215

Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys
1220 1225 1230
Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro
1235 1240 1245
Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1250 1255 1260
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu
1265 1270 1275 1280

Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val
1285 1290 1295
Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys
1300 1305 1310
Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys
1315 1320 1325
Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn
1330 1335 1340
Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn

1345 1350 1355 1360
Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser
1365 1370 1375
His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln
1380 1385 1390
Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln
1395 1400 1405
Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp
1410 1415 1420

Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu
1425 1430 1435 1440
Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala
1445 1450 1455
Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr
1460 1465 1470

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

1475 1480 1485

Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1490 1495 1500

Ala Ala Ala Leu Asp Leu Glu Lys Arg Pro Ala Ala Thr Lys Lys Ala

1505 1510 1515 1520

Gly Gln Ala Lys Lys Lys Lys His His His His His His

1525 1530

<210> 52

<211> 4602

<212> DNA

<213> Artificial Sequence

<220><223> DSB-SpyCas9

<400> 52

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggctaa aaaagaaatg 120

gttgaatttg atgaagctat ccatggcgaa gacttggcta aatttattaa agaagcatct 180

gatcataaac tgaaaatttc cggttataat gaactgatta aagatattcg aattcgtgct 240

aaagatgaac ttggcgttga tggtaagatg tttaatcgtc tattagcttt gtatcataaa 300

gataaccgtg atgtgtttga agctgaaact gaagaggtag ttgaacttta tgacacagtt 360

ttctctaaag gatccgaatt cgagctccgt cgacaagctt gcggccgcat ggacaagaag 420

tacagcatcg gcctggacat cggtaccaac agcgtgggct gggccgtgat caccgacgag 480

tacaaggtgc ccagcaagaa gttcaaggtg ctgggcaaca ccgaccgcca cagcatcaag 540

aagaacctga tcggcgccct gctgttcgac agcggcgaga ccgccgaggc cacccgcctg 600

aagcgcaccg cccgccgccg ctacacccgc cgcaagaacc gcatctgcta cctgcaggag 660

atcttcagca acgagatggc caaggtggac gacagcttct tccaccgcct ggaggagagc 720

ttcctggtgg aggaggacaa gaagcacgag cgccacccca tcttcggcaa catcgtggac 780

gaggtggcct accacgagaa gtaccccacc atctaccacc tgcgcaagaa gctggtggac 840

agcaccgaca aggccgacct gcgcctgatc tacctggccc tggcccacat gatcaagttc 900

cgcggccact tcctgatcga gggcgacctg aaccccgaca acagcgacgt ggacaagctg 960

ttcatccagc tggtgcagac ctacaaccag ctgttcgagg agaaccccat caacgccagc 1020

ggcgtggacg ccaaggccat cctgagcgcc cgcctgagca agagccgccg cctggagaac 1080

ctgatcgccc agctgcccgg cgagaagaag aacggcctgt tcggcaacct gatcgccctg 1140

agcctgggcc tgaccccaa cttcaagagc aacttcgacc tggccgagga cgccaagctg 1200

cagctgagca aggacaccta cgacgacgac ctggacaacc tgctggccca gatcggcgac 1260

cagtacgccg acctgttcct ggccgccaag aacctgagcg acgccatcct gctgagcgac 1320

atcctgcgcg tgaacaccga gatcaccaag gccccctga gcgccagcat gatcaagcgc 1380

tacgacgagc accaccagga cctgaccctg ctgaaggccc tggtgcgcca gcagctgccc 1440

gagaagtaca aggagatctt cttcgaccag agcaagaacg gctacgccgg ctacatcgac 1500

ggcggcgcca gccaggagga gttctacaag ttcatcaagc ccatcctgga gaagatggac 1560

ggcaccgagg agctgctggt gaagctgaac cgcgaggacc tgctgcgcaa gcagcgcacc 1620

ttcgacaacg gcagcatccc ccaccagatc cacctgggcg agctgcacgc catcctgcgc 1680

cgccaggagg acttctaccc cttcctgaag gacaaccgcg agaagatcga gaagatcctg 1740

accttccgca tcccctacta cgtgggcccc ctggcccgcg gcaacagccg cttcgcctgg 1800

atgacccgca agagcgagga gaccatcacc ccctggaact tcgaggaggt ggtggacaag 1860

ggcgccagcg cccagagctt catcgagcgc atgaccaact tcgacaagaa cctgcccaac 1920

gagaaggtgc tgcccaagca cagcctgctg tacgagtact tcaccgtgta caacgagctg 1980

accaaggtga agtacgtgac cgagggcatg cgcaagcccg ccttcctgag cggcgagcag 2040

aagaaggcca tcgtggacct gctgttcaag accaaccgca aggtgaccgt gaagcagctg 2100

aaggaggact acttcaagaa gatcgagtgc ttcgacagcg tggagatcag cggcgtggag 2160

gaccgcttca acgccagcct gggcacctac cacgacctgc tgaagatcat caaggacaag 2220

gacttcctgg acaacgagga gaacgaggac atcctggagg acatcgtgct gaccctgacc 2280

ctgttcgagg accgcgagat gatcgaggag cgcctgaaga cctacgccca cctgttcgac 2340

gacaaggtga tgaagcagct gaagcgccgc cgctacaccg gctggggccg cctgagccgc 2400

aagcttatca acggcatccg cgacaagcag agcggcaaga ccatcctgga cttcctgaag 2460

agcgacggct tcgccaaccg caacttcatg cagctgatcc acgacgacag cctgaccttc 2520

aaggaggaca tccagaaggc ccaggtgagc ggccagggcg acagcctgca cgagcacatc 2580

gccaacctgg ccggcagccc cgccatcaag aagggcatcc tgcagaccgt gaaggtggtg 2640

gacgagctgg tgaaggtgat gggccgccac aagcccgaga acatcgtgat cgagatggcc 2700

cgcgagaacc agaccaccca gaagggccag aagaacagcc gcgagcgcat gaagcgcatc 2760

gaggagggca tcaaggagct gggcagccag atcctgaagg agcaccccgt ggagaacacc 2820

cagctgcaga acgagaagct gtacctgtac tacctgcaga acggccgcga catgtacgtg 2880

gaccaggagc tggacatcaa ccgcctgagc gactacgacg tggaccacat cgtgccccag 2940

agcttcctga aggacgacag catcgacaac aaggtgctga cccgcagcga caagaaccgc 3000

ggcaagagcg acaacgtgcc cagcgaggag gtggtgaaga agatgaagaa ctactggcgc 3060

cagctgctga acgccaagct gatcacccag cgcaagttcg acaacctgac caaggccgag 3120

cgcggcggcc tgagcgagct ggacaaggcc ggcttcatca agcgccagct ggtggagacc 3180

cgccagatca ccaagcacgt ggcccagatc ctggacagcc gcatgaacac caagtacgac 3240

gagaacgaca agctgatccg cgaggtgaag gtgatcaccc tgaagagcaa gctggtgagc 3300

gacttccgca aggacttcca gttctacaag gtgcgcgaga tcaacaacta ccaccacgcc 3360

cacgacgcct acctgaacgc cgtggtgggc accgccctga tcaagaagta ccccaagctg 3420

gagagcgagt tcgtgtacgg cgactacaag gtgtacgacg tgcgcaagat gatcgccaag 3480

agcgagcagg agatcggcaa ggccaccgcc aagtacttct tctacagcaa catcatgaac 3540

ttcttcaaga ccgagatcac cctggccaac ggcgagatcc gcaagcgccc cctgatcgag 3600

accaacggcg agaccggcga gatcgtgtgg gacaagggcc gcgacttcgc caccgtgcgc 3660

aaggtgctga gcatgcccca ggtgaacatc gtgaagaaga ccgaggtgca gaccggcggc 3720

ttcagcaagg agagcatcct gcccaagcgc aacagcgaca agctgatcgc ccgcaagaag 3780

gactgggacc ccaagaagta cggcggcttc gacagcccca ccgtggccta cagcgtgctg 3840

gtggtggcca aggtggagaa gggcaagagc aagaagctga agagcgtgaa ggagctgctg 3900

ggcatcacca tcatggagcg cagcagcttc gagaagaacc ccatcgactt cctggaggcc 3960

aagggctaca aggaggtgaa gaaggacctg atcatcaagc tgcccaagta cagcctgttc 4020

gagctggaga acggccgcaa gcgcatgctg gccagcgccg gcgagctgca gaagggcaac 4080

gagctggccc tgcccagcaa gtacgtgaac ttcctgtacc tggccagcca ctacgagaag 4140

ctgaagggca gccccgagga caacgagcag aagcagctgt tcgtggagca gcacaagcac 4200

tacctggacg agatcatcga gcagatcagc gagttcagca agcgcgtgat cctggccgac 4260

gccaacctgg acaaggtgct gagcgcctac aacaagcacc gcgacaagcc catccgcgag 4320

caggccgaga acatcatcca cctgttcacc ctgaccaacc tgggcgcccc cgccgccttc 4380

aagtacttcg acaccaccat cgaccgcaag cgctacacca gcaccaagga ggtgctggac 4440

gccaccctga tccaccagag catcaccggt ctgtacgaga cccgcatcga cctgagccag 4500

ctgggcggcg acgcggccgc actcgacctc gagaaaaggc cggcggccac gaaaaaggcc 4560

ggccaggcaa aaaagaaaaa gcaccaccac caccaccact ga 4602

<210> 53

<211> 2112

<212> PRT

<213> Artificial Sequence

<220><223> DSB-SpyCas9-RecJ

<400> 53

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ala Lys Lys Glu Met Val Glu Phe Asp Glu Ala Ile His
35 40 45

Gly Glu Asp Leu Ala Lys Phe Ile Lys Glu Ala Ser Asp His Lys Leu
50 55 60

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

65 70 75 80

Lys Asp Glu Leu Gly Val Asp Gly Lys Met Phe Asn Arg Leu Leu Ala
85 90 95

Leu Tyr His Lys Asp Asn Arg Asp Val Phe Glu Ala Glu Thr Glu Glu
100 105 110

Val Val Glu Leu Tyr Asp Thr Val Phe Ser Lys Gly Ser Glu Phe Glu
115 120 125

Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys Lys Tyr Ser Ile Gly
130 135 140

Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu
145 150 155 160

Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg
165 170 175

His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly
180 185 190

Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr
195 200 205

Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn

210 215 220

Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser

225 230 235 240

Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly

245 250 255

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

260 265 270

His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg

275 280 285

Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe

290 295 300

Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu

305 310 315 320

Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro

325 330 335

Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu

340 345 350

Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu

355 360 365

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

370 375 380

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

385 390 395 400

Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala

405 410 415

Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu

420 425 430

Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile

435 440 445

Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His

450 455 460

His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro

465 470 475 480

Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala

485 490 495

Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile

500 505 510

Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys

515 520 525

Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly

530 535 540

Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg

545 550 555 560

Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile

565 570 575

Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala

580 585 590

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr

595 600 605

Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala

610 615 620

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

625 630 635 640

Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val

645 650 655

Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys

660 665 670

Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu

675 680 685

Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr

690 695 700

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu

705 710 715 720

Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile
725 730 735
Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu
740 745 750
Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile
755 760 765
Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met
770 775 780
Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg

785 790 795 800
Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu
805 810 815
Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu
820 825 830
Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln
835 840 845
Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala
850 855 860

Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val
865 870 875 880
Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val
885 890 895
Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn
900 905 910
Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly
915 920 925
Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn

930 935 940
Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val
945 950 955 960
Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His
965 970 975

Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val
980 985 990
Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser
995 1000 1005

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn
1010 1015 1020
Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu
1025 1030 1035 1040
Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln
1045 1050 1055
Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp
1060 1065 1070
Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu

1075 1080 1085
Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys
1090 1095 1100
Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala
1105 1110 1115 1120
His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys
1125 1130 1135
Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr
1140 1145 1150

Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala
1155 1160 1165
Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr
1170 1175 1180
Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu
1185 1190 1195 1200
Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe
1205 1210 1215
Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys

1220 1225 1230

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

1235 1240 1245

Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro

1250 1255 1260

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu

1265 1270 1275 1280

Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val

1285 1290 1295

Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys

1300 1305 1310

Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys

1315 1320 1325

Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn

1330 1335 1340

Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn

1345 1350 1355 1360

Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser

1365 1370 1375

His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln

1380 1385 1390

Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln

1395 1400 1405

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

1410 1415 1420

Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu

1425 1430 1435 1440

Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala

1445 1450 1455

Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr

1460 1465 1470

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

1475 1480 1485

Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1490 1495 1500

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

1505 1510 1515 1520

Arg Glu Val Asp Glu Thr Ala Asp Leu Pro Ala Glu Leu Pro Pro Leu

1525 1530 1535

Leu Arg Arg Leu Tyr Ala Ser Arg Gly Val Arg Ser Ala Gln Glu Leu

1540 1545 1550

Glu Arg Ser Val Lys Gly Met Leu Pro Trp Gln Gln Leu Ser Gly Val

1555 1560 1565

Glu Lys Ala Val Glu Ile Leu Tyr Asn Ala Phe Arg Glu Gly Thr Arg

1570 1575 1580

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

1585 1590 1595 1600

Leu Ser Val Leu Ala Met Arg Ser Leu Gly Cys Ser Asn Ile Asp Tyr

1605 1610 1615

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

1620 1625 1630

Val Asp Gln Ala His Ala Arg Gly Ala Gln Leu Ile Val Thr Val Asp

1635 1640 1645

Asn Gly Ile Ser Ser His Ala Gly Val Glu His Ala Arg Ser Leu Gly

1650 1655 1660

Ile Pro Val Ile Val Thr Asp His His Leu Pro Gly Asp Thr Leu Pro

1665 1670 1675 1680

Ala Ala Glu Ala Ile Ile Asn Pro Asn Leu Arg Asp Cys Asn Phe Pro

1685 1690 1695

Ser Lys Ser Leu Ala Gly Val Gly Val Ala Phe Tyr Leu Met Leu Ala

1700 1705 1710

Leu Arg Thr Phe Leu Arg Asp Gln Gly Trp Phe Asp Glu Arg Asn Ile

1715 1720 1725

Ala Ile Pro Asn Leu Ala Glu Leu Leu Asp Leu Val Ala Leu Gly Thr
1730 1735 1740

Val Ala Asp Val Val Pro Leu Asp Ala Asn Asn Arg Ile Leu Thr Trp
1745 1750 1755 1760

Gln Gly Met Ser Arg Ile Arg Ala Gly Lys Cys Arg Pro Gly Ile Lys
1765 1770 1775

Ala Leu Leu Glu Val Ala Asn Arg Asp Ala Gln Lys Leu Ala Ala Ser
1780 1785 1790

Asp Leu Gly Phe Ala Leu Gly Pro Arg Leu Asn Ala Ala Gly Arg Leu

1795 1800 1805

Asp Asp Met Ser Val Gly Val Ala Leu Leu Leu Cys Asp Asn Ile Gly
1810 1815 1820

Glu Ala Arg Val Leu Ala Asn Glu Leu Asp Ala Leu Asn Gln Thr Arg
1825 1830 1835 1840

Lys Glu Ile Glu Gln Gly Met Gln Ile Glu Ala Leu Thr Leu Cys Glu
1845 1850 1855

Lys Leu Glu Arg Ser Arg Asp Thr Leu Pro Gly Gly Leu Ala Met Tyr
1860 1865 1870

His Pro Glu Trp His Gln Gly Val Val Gly Ile Leu Ala Ser Arg Ile
1875 1880 1885

Lys Glu Arg Phe His Arg Pro Val Ile Ala Phe Ala Pro Ala Gly Asp
1890 1895 1900

Gly Thr Leu Lys Gly Ser Gly Arg Ser Ile Gln Gly Leu His Met Arg
1905 1910 1915 1920

Asp Ala Leu Glu Arg Leu Asp Thr Leu Tyr Pro Gly Met Met Leu Lys
1925 1930 1935

Phe Gly Gly His Ala Met Ala Ala Gly Leu Ser Leu Glu Glu Asp Lys

1940 1945 1950

Phe Lys Leu Phe Gln Gln Arg Phe Gly Glu Leu Val Thr Glu Trp Leu
1955 1960 1965

Asp Pro Ser Leu Leu Gln Gly Glu Val Val Ser Asp Gly Pro Leu Ser
1970 1975 1980

Pro Ala Glu Met Thr Met Glu Val Ala Gln Leu Leu Arg Asp Ala Gly
1985 1990 1995 2000
Pro Trp Gly Gln Met Phe Pro Glu Pro Leu Phe Asp Gly His Phe Arg
2005 2010 2015

Leu Leu Gln Gln Arg Leu Val Gly Glu Arg His Leu Lys Val Met Val
2020 2025 2030
Glu Pro Val Gly Gly Gly Pro Leu Leu Asp Gly Ile Ala Phe Asn Val
2035 2040 2045
Asp Thr Ala Leu Trp Pro Asp Asn Gly Val Arg Glu Val Gln Leu Ala
2050 2055 2060
Tyr Lys Leu Asp Ile Asn Glu Phe Arg Gly Asn Arg Ser Leu Gln Ile
2065 2070 2075 2080
Ile Ile Asp Asn Ile Trp Pro Ile Leu Gln Lys Arg Pro Ala Ala Thr

2085 2090 2095
Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His His His His
2100 2105 2110

<210> 54
<211> 6339
<212> DNA
<213> Artificial Sequence
<220><223> DSB-SpyCas9-RecJ
<400> 54

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60
atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggctaa aaaagaaatg 120
gttgaatttg atgaagctat ccatggcgaa gacttggcta aatttattaa agaagcatct 180

gatcataaac tgaaaatttc cggttataat gaactgatta aagatattcg aattcgtgct 240
aaagatgaac ttggcgttga tggtaagatg tttaatcgtc tattagcttt gtatcataaa 300
gataaccgtg atgtgtttga agctgaaact gaagaggtag ttgaacttta tgacacagtt 360
ttctctaaag gatccgaatt cgagctccgt cgacaagctt gcggccgcat ggacaagaag 420
tacagcatcg gcctggacat cggtaccaac agcgtgggct gggccgtgat caccgacgag 480
tacaaggtgc ccagcaagaa gttcaaggtg ctgggcaaca ccgaccgcca cagcatcaag 540

aagaacctga tcggcgccct gctgttcgac agcggcgaga ccgccgaggc cacccgcctg 600

aagcgcaccg cccgccgccg ctacacccgc cgcaagaacc gcatctgcta cctgcaggag 660

atcttcagca acgagatggc caaggtggac gacagcttct tccaccgcct ggaggagagc 720

ttcctggtgg aggaggacaa gaagcacgag cgccacccca tcttcggcaa catcgtggac 780

gaggtggcct accacgagaa gtaccccacc atctaccacc tgcgcaagaa gctggtggac 840

agcaccgaca aggccgacct gcgcctgatc tacctggccc tggcccacat gatcaagttc 900

cgcggccact tcctgatcga gggcgacctg aaccccgaca acagcgacgt ggacaagctg 960

ttcatccagc tggtgcagac ctacaaccag ctgttcgagg agaaccccat caacgccagc 1020

ggcgtggacg ccaaggccat cctgagcgcc cgcctgagca agagccgccg cctggagaac 1080

ctgatcgccc agctgcccgg cgagaagaag aacggcctgt tcggcaacct gatcgccctg 1140

agcctgggcc tgaccccaa cttcaagagc aacttcgacc tggccgagga cgccaagctg 1200

cagctgagca aggacaccta cgacgacgac ctggacaacc tgctggccca gatcggcgac 1260

cagtacgccg acctgttcct ggccgccaag aacctgagcg acgccatcct gctgagcgac 1320

atcctgcgcg tgaacaccga gatcaccaag gccccctga gcgccagcat gatcaagcgc 1380

tacgacgagc accaccagga cctgaccctg ctgaaggccc tggtgcgcca gcagctgccc 1440

gagaagtaca aggagatctt cttcgaccag agcaagaacg gctacgccgg ctacatcgac 1500

ggcggcgcca gccaggagga gttctacaag ttcatcaagc ccatcctgga gaagatggac 1560

ggcaccgagg agctgctggt gaagctgaac cgcgaggacc tgctgcgcaa gcagcgcacc 1620

ttcgacaacg gcagcatccc ccaccagatc cacctgggcg agctgcacgc catcctgcgc 1680

cgccaggagg acttctaccc cttcctgaag gacaaccgcg agaagatcga gaagatcctg 1740

accttccgca tcccctacta cgtgggcccc ctggcccgcg gcaacagccg cttcgcctgg 1800

atgacccgca agagcgagga gaccatcacc ccctggaact tcgaggaggt ggtggacaag 1860

ggcgccagcg cccagagctt catcgagcgc atgaccaact tcgacaagaa cctgcccaac 1920

gagaaggtgc tgcccaagca cagcctgctg tacgagtact tcaccgtgta caacgagctg 1980

accaaggtga agtacgtgac cgagggcatg cgcaagcccg ccttcctgag cggcgagcag 2040

aagaaggcca tcgtggacct gctgttcaag accaaccgca aggtgaccgt gaagcagctg 2100

aaggaggact acttcaagaa gatcgagtgc ttcgacagcg tggagatcag cggcgtggag 2160

gaccgcttca acgccagcct gggcacctac cacgacctgc tgaagatcat caaggacaag 2220

gacttcctgg acaacgagga gaacgaggac atcctggagg acatcgtgct gaccctgacc 2280

ctgttcgagg accgcgagat gatcgaggag cgcctgaaga cctacgccca cctgttcgac 2340

gacaaggtga tgaagcagct gaagcgccgc cgctacaccg gctggggccg cctgagccgc 2400

aagcttatca acggcatccg cgacaagcag agcggcaaga ccatcctgga cttcctgaag 2460

agcgacggct tcgccaaccg caacttcatg cagctgatcc acgacgacag cctgaccttc 2520

aaggaggaca tccagaaggc ccaggtgagc ggccagggcg acagcctgca cgagcacatc 2580

gccaacctgg ccggcagccc cgccatcaag aagggcatcc tgcagaccgt gaaggtggtg 2640

gacgagctgg tgaaggtgat gggccgccac aagcccgaga acatcgtgat cgagatggcc 2700

cgcgagaacc agaccaccca gaagggccag aagaacagcc gcgagcgcat gaagcgcatc 2760

gaggagggca tcaaggagct gggcagccag atcctgaagg agcaccccgt ggagaacacc 2820

cagctgcaga acgagaagct gtacctgtac tacctgcaga acggccgcga catgtacgtg 2880

gaccaggagc tggacatcaa ccgcctgagc gactacgacg tggaccacat cgtgccccag 2940

agcttcctga aggacgacag catcgacaac aaggtgctga cccgcagcga caagaaccgc 3000

ggcaagagcg acaacgtgcc cagcgaggag gtggtgaaga agatgaagaa ctactggcgc 3060

cagctgctga acgccaagct gatcacccag cgcaagttcg acaacctgac caaggccgag 3120

cgcggcggcc tgagcgagct ggacaaggcc ggcttcatca agcgccagct ggtggagacc 3180

cgccagatca ccaagcacgt ggcccagatc ctggacagcc gcatgaacac caagtacgac 3240

gagaacgaca agctgatccg cgaggtgaag gtgatcaccc tgaagagcaa gctggtgagc 3300

gacttccgca aggacttcca gttctacaag gtgcgcgaga tcaacaacta ccaccacgcc 3360

cacgacgcct acctgaacgc cgtggtgggc accgccctga tcaagaagta ccccaagctg 3420

gagagcgagt tcgtgtacgg cgactacaag gtgtacgacg tgcgcaagat gatcgccaag 3480

agcgagcagg agatcggcaa ggccaccgcc aagtacttct tctacagcaa catcatgaac 3540

ttcttcaaga ccgagatcac cctggccaac ggcgagatcc gcaagcgccc cctgatcgag 3600

accaacggcg agaccggcga gatcgtgtgg gacaagggcc gcgacttcgc caccgtgcgc 3660

aaggtgctga gcatgcccca ggtgaacatc gtgaagaaga ccgaggtgca gaccggcggc 3720

ttcagcaagg agagcatcct gcccaagcgc aacagcgaca agctgatcgc ccgcaagaag 3780

gactgggacc ccaagaagta cggcggcttc gacagcccca ccgtggccta cagcgtgctg 3840

gtggtggcca aggtggagaa gggcaagagc aagaagctga agagcgtgaa ggagctgctg 3900

ggcatcacca tcatggagcg cagcagcttc gagaagaacc ccatcgactt cctggaggcc 3960

aagggctaca aggaggtgaa gaaggacctg atcatcaagc tgcccaagta cagcctgttc 4020

gagctggaga acggccgcaa gcgcatgctg gccagcgccg gcgagctgca gaagggcaac 4080

gagctggccc tgcccagcaa gtacgtgaac ttcctgtacc tggccagcca ctacgagaag 4140

ctgaagggca gccccgagga caacgagcag aagcagctgt tcgtggagca gcacaagcac 4200

tacctggacg agatcatcga gcagatcagc gagttcagca agcgcgtgat cctggccgac 4260

gccaacctgg acaaggtgct gagcgcctac aacaagcacc gcgacaagcc catccgcgag 4320

caggccgaga acatcatcca cctgttcacc ctgaccaacc tgggcgcccc cgccgccttc 4380

aagtacttcg acaccaccat cgaccgcaag cgctacacca gcaccaagga ggtgctggac 4440

gccaccctga tccaccagag catcaccggt ctgtacgaga cccgcatcga cctgagccag 4500

ctgggcggcg acgcggccgc actcgacctg caggtgaaac aacagataca acttcgtcgc 4560

cgtgaagtcg atgaaacggc agacttgccc gctgaattgc ctcccttgct gcgccgttta 4620

tacgccagcc ggggagtacg cagtgcgcaa gaactggaac gcagtgttaa aggtatgctg 4680

ccctggcagc aactgagcgg cgtcgaaaag gccgttgaga tcctttacaa cgcttttcgc 4740

gaaggaacgc ggattattgt ggtcggtgat ttcgacgccg acggcgcgac cagcacggct 4800

ctaagcgtgc tggcgatgcg ctcgcttggt tgcagcaata tcgactacct ggtaccaaac 4860

cgtttcgaag acggttacgg cttaagcccg gaagtggtcg atcaggccca tgcccgtggc 4920

gcgcagttaa ttgtcacggt ggataacggt atttcctccc atgcgggggt tgagcacgct 4980

cgctcgttgg gcatcccggt tattgttacc gatcaccatt tgccaggcga cacattaccc 5040

gcagcggaag cgatcattaa ccctaacttg cgcgactgta atttcccgtc gaaatcactg 5100

gcaggcgtgg gtgtggcgtt ttatctgatg ctggcgctgc gcacctttt gcgcgatcag 5160

ggctggtttg atgagcgtaa catcgcaatt cctaacctgg cagaactgct ggatctggtc 5220

gcgctgggga cagtggcgga cgtcgtgccg ctggacgcta ataatcgcat tctgacctgg 5280

caggggatga gtcgcatccg agccggaaag tgccgtccgg ggattaaagc gctgcttgaa 5340

gtggcaaacc gtgatgcaca aaaactcgcc gccagcgatt taggttttgc gctggggcca 5400

cgtctcaatg ctgccggacg actggacgat atgtccgtcg gtgtggcgct gttgttgtgc 5460

gacaacatcg gcgaagcgcg cgtgctggca aatgaactcg atgcgctaaa ccagacgcga 5520

aaagagatcg aacaaggaat gcaaattgaa gccctgaccc tgtgcgagaa actggagcgc 5580

agccgtgaca cgctacccgg cgggctggca atgtatcacc ccgaatggca tcagggcgtt 5640

gtcggtattc tggcttcgcg catcaaagag cgttttcacc gtccggttat cgcgtttgcg 5700

ccagcaggtg acggtacgct gaaaggttcc ggtcgctcca ttcaggggct gcatatgcgt 5760

gatgcgctgg agcgattaga cacactctac cctggcatga tgctgaagtt tggcggtcat 5820

gcgatggcgg cgggtttgtc gctggaagag gataaattca aactctttca acaacggttt 5880

ggcgaactgg ttactgagtg gctggaccct tcgctattgc aaggcgaagt ggtatcagac 5940

ggtccgttaa gcccggccga aatgaccatg gaagtggcgc agctgctgcg cgatgctggc 6000

ccgtgggggc agatgttccc ggagccgctg tttgacggtc atttccgtct gctgcaacag 6060

cggctggtgg gcgaacgtca tttgaaggtg atggtcgaac cggtcggcgg cggtccactg 6120

ctggatggta ttgcttttaa tgtcgatacc gccctctggc cggataacgg cgtgcgcgaa 6180

gtgcaactgg cttataagct cgatatcaac gagtttcgcg gcaaccgcag cctgcaaatt 6240

atcatcgaca atatctggcc aattctgcag aaaaggccgg cggccacgaa aaaggccggc 6300

caggcaaaaa agaaaaagca ccaccaccac caccactga 6339

<210> 55

<211> 1379

<212> PRT

<213> Artificial Sequence

<220><223> FnCpf1

<400> 55

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro

1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile

35 40 45

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe

50 55 60

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

65 70 75 80

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys

85 90 95

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

100 105 110

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr

115 120 125

Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys

130 135 140

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

145 150 155 160

Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys

165 170 175

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp

180 185 190

Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp

195 200 205

Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

210 215 220

Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile

225 230 235 240

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

245 250 255

Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu

260 265 270

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr

275 280 285

Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser

290 295 300

Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln

305 310 315 320

Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn

325 330 335

Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr

340 345 350

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

355 360 365

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

370 375 380

Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe

385 390 395 400

Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys

405 410 415

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

420 425 430

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

435 440 445

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

450 455 460

Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys

465 470 475 480

Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu

485 490 495

Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg

500 505 510

Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala

515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu

530 535 540

Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu

545 550 555 560

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

565 570 575

Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln

580 585 590

Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu

595 600 605

Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr

610 615 620

Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640

Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys

645 650 655

Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys Asp Asp Lys
660 665 670

Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp
675 680 685

Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr
690 695 700

Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val Phe Phe Ser
705 710 715 720

Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile

725 730 735

Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr
740 745 750

Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe
755 760 765

Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe
770 775 780

Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu Phe Tyr Arg
785 790 795 800

Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn Ile Ser Glu
805 810 815

Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln
820 825 830

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu
835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp
850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln

865 870 875 880

Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn
885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu
900 905 910

Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro
915 920 925
Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu
930 935 940

Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser
945 950 955 960
Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly
965 970 975
Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp
980 985 990
Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp
995 1000 1005
Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1010 1015 1020
Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu
1025 1030 1035 1040
Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly
1045 1050 1055
Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu
1060 1065 1070
Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn
1075 1080 1085

Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala
1090 1095 1100
Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr
1105 1110 1115 1120
Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe
1125 1130 1135
Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu
1140 1145 1150
Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr

1155 1160 1165

Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys

1170 1175 1180

Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg

1185 1190 1195 1200

Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr

1205 1210 1215

Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1220 1225 1230

Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe

1235 1240 1245

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

1250 1255 1260

Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp

1265 1270 1275 1280

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro

1285 1290 1295

Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu

1300 1305 1310

Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn

1315 1320 1325

Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn

1330 1335 1340

Asn Gln Ala Ala Ala Leu Glu Lys Arg Pro Ala Ala Thr Lys Lys Ala

1345 1350 1355 1360

Gly Gln Ala Lys Lys Lys Lys Ser Thr Pro Pro Pro Pro Pro Leu Arg

1365 1370 1375

Ser Gly Cys

<210> 56

<211> 4140

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1

<400> 56

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agttttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttatttttaaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800

gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040

aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100

aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160

gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220

acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580

aatcttcaag atgtggttta taagctaaat ggtgaggcag agctttttta tcgtaaacaa 2640

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag caccttttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020

cagaatagga ataaccaagc ggccgcactc gagaaaaggc cggcggccac gaaaaaggcc 4080

ggccaggcaa aaaagaaaaa gtcgacacca ccaccaccac cactgagatc cggctgctaa 4140

4140

<210> 57

<211> 1954

<212> PRT

<213> Artificial Sequence

<220><223> FnCpf1-RecJ

<400> 57

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile
35 40 45

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe
50 55 60

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly
65 70 75 80

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys
85 90 95

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

100 105 110

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr
115 120 125
Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys
130 135 140
Ser Ala Lys Asp Thr Ile Lys Lys Gln Ile Ser Glu Tyr Ile Lys Asp
145 150 155 160
Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys
165 170 175

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp
180 185 190
Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp
195 200 205
Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe
210 215 220
Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile
225 230 235 240
Pro Thr Ser Ile Ile Tyr Arg Ile Val Asp Asp Asn Leu Pro Lys Phe

245 250 255
Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu
260 265 270
Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr
275 280 285
Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser
290 295 300
Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln
305 310 315 320

Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn
325 330 335
Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr
340 345 350
Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys Met Ser Val
355 360 365

Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser Phe Val Ile
370 375 380
Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe

385 390 395 400
Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys
405 410 415
Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln Lys Leu Asp
420 425 430
Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr Asp Leu Ser
435 440 445
Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala Val Leu Glu
450 455 460

Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys
465 470 475 480
Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu
485 490 495
Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg
500 505 510
Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala
515 520 525
Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu

530 535 540
Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu
545 550 555 560
Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp Leu Leu Asp
565 570 575
Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln
580 585 590
Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu
595 600 605

Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr

610 615 620

Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640

Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys

645 650 655

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

660 665 670

Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp

675 680 685

Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr

690 695 700

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

705 710 715 720

Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile

725 730 735

Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr

740 745 750

Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe

755 760 765

Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe

770 775 780

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

785 790 795 800

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

805 810 815

Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln

820 825 830

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu

835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp

850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln

865 870 875 880

Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn

885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu

900 905 910

Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro

915 920 925

Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu

930 935 940

Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser

945 950 955 960

Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly

965 970 975

Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp

980 985 990

Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp

995 1000 1005

Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1010 1015 1020

Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu

1025 1030 1035 1040

Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly

1045 1050 1055

Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu

1060 1065 1070

Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn

1075 1080 1085

Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala

1090 1095 1100

Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr

1105 1110 1115 1120

Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe
1125 1130 1135
Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu
1140 1145 1150
Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr
1155 1160 1165
Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys
1170 1175 1180

Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg
1185 1190 1195 1200
Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr
1205 1210 1215
Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His
1220 1225 1230
Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe
1235 1240 1245
Phe Ala Lys Leu Thr Ser Val Leu Asn Thr Ile Leu Gln Met Arg Asn

1250 1255 1260
Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp
1265 1270 1275 1280
Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro
1285 1290 1295
Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu
1300 1305 1310
Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn
1315 1320 1325

Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn
1330 1335 1340
Asn Gln Ala Ala Ala Leu Asp Leu Gln Val Lys Gln Gln Ile Gln Leu
1345 1350 1355 1360
Arg Arg Arg Glu Val Asp Glu Thr Ala Asp Leu Pro Ala Glu Leu Pro
1365 1370 1375

Pro Leu Leu Arg Arg Leu Tyr Ala Ser Arg Gly Val Arg Ser Ala Gln
1380 1385 1390
Glu Leu Glu Arg Ser Val Lys Gly Met Leu Pro Trp Gln Gln Leu Ser

1395 1400 1405
Gly Val Glu Lys Ala Val Glu Ile Leu Tyr Asn Ala Phe Arg Glu Gly
1410 1415 1420
Thr Arg Ile Ile Val Val Gly Asp Phe Asp Ala Asp Gly Ala Thr Ser
1425 1430 1435 1440
Thr Ala Leu Ser Val Leu Ala Met Arg Ser Leu Gly Cys Ser Asn Ile
1445 1450 1455
Asp Tyr Leu Val Pro Asn Arg Phe Glu Asp Gly Tyr Gly Leu Ser Pro
1460 1465 1470

Glu Val Val Asp Gln Ala His Ala Arg Gly Ala Gln Leu Ile Val Thr
1475 1480 1485
Val Asp Asn Gly Ile Ser Ser His Ala Gly Val Glu His Ala Arg Ser
1490 1495 1500
Leu Gly Ile Pro Val Ile Val Thr Asp His His Leu Pro Gly Glu Thr
1505 1510 1515 1520
Leu Pro Ala Ala Glu Ala Ile Ile Asn Pro Asn Leu Arg Asp Cys Asn
1525 1530 1535
Phe Pro Ser Lys Ser Leu Ala Gly Val Gly Val Ala Phe Tyr Leu Met

1540 1545 1550
Leu Ala Leu Arg Thr Phe Leu Arg Asp Gln Gly Trp Phe Asp Glu Arg
1555 1560 1565
Gly Ile Ala Ile Pro Asn Leu Ala Glu Leu Leu Asp Leu Val Ala Leu
1570 1575 1580
Gly Thr Val Ala Asp Val Val Pro Leu Asp Ala Asn Asn Arg Ile Leu
1585 1590 1595 1600
Thr Trp Gln Gly Met Ser Arg Ile Arg Ala Gly Lys Cys Arg Pro Gly
1605 1610 1615

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

1620 1625 1630

Ala Ser Asp Leu Gly Phe Ala Leu Gly Pro Arg Leu Asn Ala Ala Gly

1635 1640 1645

Arg Leu Asp Asp Met Ser Val Gly Val Ala Leu Leu Leu Cys Asp Asn

1650 1655 1660

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

1665 1670 1675 1680

Thr Arg Lys Glu Ile Glu Gln Gly Met Gln Val Glu Ala Leu Thr Leu

1685 1690 1695

Cys Glu Lys Leu Glu Arg Ser Arg Asp Thr Leu Pro Gly Gly Leu Ala

1700 1705 1710

Met Tyr His Pro Glu Trp His Gln Gly Val Val Gly Ile Leu Ala Ser

1715 1720 1725

Arg Ile Lys Glu Arg Phe His Arg Pro Val Ile Ala Phe Ala Pro Ala

1730 1735 1740

Gly Asp Gly Thr Leu Lys Gly Ser Gly Arg Ser Ile Gln Gly Leu His

1745 1750 1755 1760

Met Arg Asp Ala Leu Glu Arg Leu Asp Thr Leu Tyr Pro Gly Met Ile

1765 1770 1775

Leu Lys Phe Gly Gly His Ala Met Ala Ala Gly Leu Ser Leu Glu Glu

1780 1785 1790

Asp Lys Phe Glu Leu Phe Gln Gln Arg Phe Gly Glu Leu Val Thr Glu

1795 1800 1805

Trp Leu Asp Pro Ser Leu Leu Gln Gly Glu Val Val Ser Asp Gly Pro

1810 1815 1820

Leu Ser Pro Ala Glu Met Thr Met Glu Val Ala Gln Leu Leu Arg Asp

1825 1830 1835 1840

Ala Gly Pro Trp Gly Gln Met Phe Pro Glu Pro Leu Phe Asp Gly His

1845 1850 1855

Phe Arg Leu Leu Gln Gln Arg Leu Val Gly Glu Arg His Leu Lys Val

1860 1865 1870

Met Val Glu Pro Val Gly Gly Gly Pro Leu Leu Asp Gly Ile Ala Phe

1875 1880 1885

Asn Val Asp Thr Ala Leu Trp Pro Asp Asn Gly Val Arg Glu Val Gln

1890 1895 1900

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

1905 1910 1915 1920

Gln Ile Ile Ile Asp Asn Ile Trp Pro Ile Leu Gln Lys Arg Pro Ala

1925 1930 1935

Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His His

1940 1945 1950

His His

<210> 58

<211> 5865

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1-RecJ

<400>

58

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agtttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttatttttaaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800

gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040

aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100

aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160

gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220

acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580

aatcttcaag atgtggttta taagctaaat ggtgaggcag agcttttta tcgtaaacaa 2640

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag caccttttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gtttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020

cagaatagga ataaccaagc ggccgcactc gacctgcagg tgaaacaaca gatacaactt 4080

cgtcgccgtg aagtcgatga aacggcagac ttgcccgctg aattgcctcc cttgctgcgc 4140

cgtttatacg ccagccgggg cgtgcgcagt gcgcaagaac tggaacgcag tgttaaaggt 4200

atgttgccct ggcagcaact gagcggcgtc gaaaaggccg ttgagatcct ttacaacgct 4260

tttcgcgaag gaacgcggat tattgtggtc ggcgattttg acgccgacgg cgcgaccagc 4320

acggctctaa gcgtgctggc gatgcgctcg cttggttgca gcaatatcga ctatctggta 4380

ccaaaccgtt tcgaagacgg ttacggctta agcccggaag tagtcgatca ggcccatgcc 4440

cgtggcgcgc agttaattgt cacggtggat aacggtattt cctcccatgc gggcgttgaa 4500

cacgctcgct cgttgggcat tccggttatt gttaccgatc accatttgcc gggcgaaaca 4560

ttacccgcag cggaagcgat cattaaccct aacttgcgcg actgtaattt cccgtcgaaa 4620

tcactggcag gcgtgggtgt ggcgttttat ctgatgctgg cgctgcgcac ctttttgcgc 4680
gatcagggct ggtttgatga gcgtggcatc gcaattccta acctggcaga actgctggat 4740
ctggtcgcgc tgggaacagt ggcggacgtc gtgccgctgg acgctaataa tcgcattctg 4800
acctggcagg ggatgagtcg catccgtgcc ggaaagtgcc gtccagggat taaagcgctg 4860
ctggaagtgg caaaccgtga tgcacaaaaa ctcgccgcca gcgatttagg ttttgcgctg 4920
gggccacgtc tcaatgctgc cggacgactg gacgatatgt ccgtcggtgt ggcgctcttg 4980
ctgtgcgaca acatcggcga agcgcgcgtg ctggcaaatg aactcgatgc gctaaaccag 5040

acgcgaaaag agatcgaaca aggaatgcaa gttgaagccc tgaccctgtg cgagaaactg 5100
gagcgaagtc gcgacacgct acccggcggg ctggcaatgt atcaccccga atggcatcag 5160
ggcgttgtcg gtattctggc ttcgcgcatc aaagagcgtt ttcaccgtcc ggttatcgcc 5220
tttgcgccag caggtgatgg tacgctgaaa ggttcaggtc gctccattca ggggctgcat 5280
atgcgtgatg cactggagcg attagacaca ctctaccctg gcatgatact gaagtttggc 5340
ggtcatgcga tggcggcggg tttgtcgctg gaagaggata aattcgaact ctttcaacaa 5400
cggtttggcg agctggttac cgagtggctg gacccttcgc tattgcaagg cgaagtggtg 5460

tcagacggcc cgttaagccc ggccgaaatg accatggaag tggcgcagct gctgcgcgat 5520
gctggcccgt gggggcagat gttcccggag ccgctgtttg atggtcattt ccgtctgctg 5580
caacagcggc tggtgggcga acgtcatttg aaagtcatgg tcgaaccggt cggcggcggt 5640
ccgctgctgg atggtattgc ttttaatgtc gataccgccc tctggccgga taacggcgtg 5700
cgcgaagtgc aactggctta caagctcgat atcaacgagt ttcgcggcaa ccgcagcctg 5760
caaattatca tcgacaatat ctggccaatt ctgcagaaaa ggccggcggc cacgaaaaag 5820
gccggccagg caaaaaagaa aaagcaccac caccaccacc actga 5865

<210> 59
<211> 2243
<212> PRT
<213> Artificial Sequence
<220><223> FnCpf1-RecE
<400> 59

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile

35 40 45

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe

50 55 60

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

65 70 75 80

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys

85 90 95

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

100 105 110

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr

115 120 125

Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys

130 135 140

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

145 150 155 160

Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys

165 170 175

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp

180 185 190

Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp

195 200 205

Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

210 215 220

Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile

225 230 235 240

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

245 250 255

Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu

260 265 270

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr
275 280 285
Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser
290 295 300
Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln
305 310 315 320
Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn
325 330 335
Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr

340 345 350
Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys Met Ser Val
355 360 365
Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser Phe Val Ile
370 375 380
Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe
385 390 395 400
Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys
405 410 415

Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln Lys Leu Asp
420 425 430
Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr Asp Leu Ser
435 440 445
Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala Val Leu Glu
450 455 460
Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys
465 470 475 480
Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu

485 490 495
Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg
500 505 510
Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala
515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu
530 535 540
Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu
545 550 555 560

Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp Leu Leu Asp
565 570 575
Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln
580 585 590
Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu
595 600 605
Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr
610 615 620
Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640
Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys
645 650 655
Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys Asp Asp Lys
660 665 670
Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp
675 680 685
Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr
690 695 700

Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val Phe Phe Ser
705 710 715 720
Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile
725 730 735
Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr
740 745 750
Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe
755 760 765
Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe

770 775 780

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

785 790 795 800

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

805 810 815

Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln

820 825 830

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu

835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp

850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln

865 870 875 880

Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn

885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu

900 905 910

Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro

915 920 925

Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu

930 935 940

Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser

945 950 955 960

Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly

965 970 975

Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp

980 985 990

Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp

995 1000 1005

Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1010 1015 1020

Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu

1025 1030 1035 1040

Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly

1045 1050 1055

Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu

1060 1065 1070

Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn

1075 1080 1085

Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala

1090 1095 1100

Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr

1105 1110 1115 1120

Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe

1125 1130 1135

Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu

1140 1145 1150

Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr

1155 1160 1165

Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys

1170 1175 1180

Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg

1185 1190 1195 1200

Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr

1205 1210 1215

Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1220 1225 1230

Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe

1235 1240 1245

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

1250 1255 1260

Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp

1265 1270 1275 1280

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro
1285 1290 1295
Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu
1300 1305 1310
Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn
1315 1320 1325
Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn
1330 1335 1340
Asn Gln Ala Ala Ala Leu Asp Leu Gln Met Ser Thr Lys Pro Leu Phe

1345 1350 1355 1360
Leu Leu Arg Lys Ala Lys Lys Ser Ser Gly Glu Pro Asp Val Val Leu
1365 1370 1375
Trp Ala Ser Asn Asp Phe Glu Ser Thr Cys Ala Thr Leu Asp Tyr Leu
1380 1385 1390
Ile Val Lys Ser Gly Lys Lys Leu Ser Ser Tyr Phe Lys Ala Val Ala
1395 1400 1405
Thr Asn Phe Pro Val Val Asn Asp Leu Pro Ala Glu Gly Glu Ile Asp
1410 1415 1420

Phe Thr Trp Ser Glu Arg Tyr Gln Leu Ser Lys Asp Ser Met Thr Trp
1425 1430 1435 1440
Glu Leu Lys Pro Gly Ala Ala Pro Asp Asn Ala His Tyr Gln Gly Asn
1445 1450 1455
Thr Asn Val Asn Gly Glu Asp Met Thr Glu Ile Glu Glu Asn Met Leu
1460 1465 1470
Leu Pro Ile Ser Gly Gln Glu Leu Pro Ile Arg Trp Leu Ala Gln His
1475 1480 1485
Gly Ser Glu Lys Pro Val Thr His Val Ser Arg Asp Gly Leu Gln Ala

1490 1495 1500
Leu His Ile Ala Arg Ala Glu Glu Leu Pro Ala Val Thr Ala Leu Ala
1505 1510 1515 1520
Val Ser His Lys Thr Ser Leu Leu Asp Pro Leu Glu Ile Arg Glu Leu
1525 1530 1535

His Lys Leu Val Arg Asp Thr Asp Lys Val Phe Pro Asn Pro Gly Asn
1540 1545 1550
Ser Asn Leu Gly Leu Ile Thr Ala Phe Phe Glu Ala Tyr Leu Asn Ala
1555 1560 1565

Asp Tyr Thr Asp Arg Gly Leu Leu Thr Lys Glu Trp Met Lys Gly Asn
1570 1575 1580
Arg Val Ser His Ile Thr Arg Thr Ala Ser Gly Ala Asn Ala Gly Gly
1585 1590 1595 1600
Gly Asn Leu Thr Asp Arg Gly Glu Gly Phe Val His Asp Leu Thr Ser
1605 1610 1615
Leu Ala Arg Asp Val Ala Thr Gly Val Leu Ala Arg Ser Met Asp Leu
1620 1625 1630
Asp Ile Tyr Asn Leu His Pro Ala His Ala Lys Arg Ile Glu Glu Ile

1635 1640 1645
Ile Ala Glu Asn Lys Pro Pro Phe Ser Val Phe Arg Asp Lys Phe Ile
1650 1655 1660
Thr Met Pro Gly Gly Leu Asp Tyr Ser Arg Ala Ile Val Val Ala Ser
1665 1670 1675 1680
Val Lys Glu Ala Pro Ile Gly Ile Glu Val Ile Pro Ala His Val Thr
1685 1690 1695
Glu Tyr Leu Asn Lys Val Leu Thr Glu Thr Asp His Ala Asn Pro Asp
1700 1705 1710

Pro Glu Ile Val Asp Ile Ala Cys Gly Arg Ser Ser Ala Pro Met Pro
1715 1720 1725
Gln Arg Val Thr Glu Glu Gly Lys Gln Asp Asp Glu Glu Lys Pro Gln
1730 1735 1740
Pro Ser Gly Thr Thr Ala Val Glu Gln Gly Glu Ala Glu Thr Met Glu
1745 1750 1755 1760
Pro Asp Ala Thr Glu His His Gln Asp Thr Gln Pro Leu Asp Ala Gln
1765 1770 1775
Ser Gln Val Asn Ser Val Asp Ala Lys Tyr Gln Glu Leu Arg Ala Glu

1780 1785 1790

Leu His Glu Ala Arg Lys Asn Ile Pro Ser Lys Asn Pro Val Asp Asp

1795 1800 1805

Asp Lys Leu Leu Ala Ala Ser Arg Gly Glu Phe Val Asp Gly Ile Ser

1810 1815 1820

Asp Pro Asn Asp Pro Lys Trp Val Lys Gly Ile Gln Thr Arg Asp Cys

1825 1830 1835 1840

Val Tyr Gln Asn Gln Pro Glu Thr Glu Lys Thr Ser Pro Asp Met Asn

1845 1850 1855

Gln Pro Glu Pro Val Val Gln Gln Glu Pro Glu Ile Ala Cys Asn Ala

1860 1865 1870

Cys Gly Gln Thr Gly Gly Asp Asn Cys Pro Asp Cys Gly Ala Val Met

1875 1880 1885

Gly Asp Ala Thr Tyr Gln Glu Thr Phe Asp Glu Glu Ser Gln Val Glu

1890 1895 1900

Ala Lys Glu Asn Asp Pro Glu Glu Met Glu Gly Ala Glu His Pro His

1905 1910 1915 1920

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

1925 1930 1935

Gly Glu Val Ala Asp Pro Val Ile Val Glu Asp Ile Glu Pro Gly Ile

1940 1945 1950

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

1955 1960 1965

Lys Ser Gln Leu Asp Asp Ile Ala Asp Thr Pro Ala Leu Tyr Leu Trp

1970 1975 1980

Arg Lys Asn Ala Pro Val Asp Thr Thr Lys Thr Lys Thr Leu Asp Leu

1985 1990 1995 2000

Gly Thr Ala Phe His Cys Arg Val Leu Glu Pro Glu Glu Phe Ser Asn

2005 2010 2015

Arg Phe Ile Val Ala Pro Glu Phe Asn Arg Arg Thr Asn Ala Gly Lys

2020 2025 2030

Glu Glu Glu Lys Ala Phe Leu Met Glu Cys Ala Ser Thr Gly Lys Thr

2035 2040 2045

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

2050 2055 2060

Val Met Ala Leu Pro Leu Gly Gln Trp Leu Val Glu Ser Ala Gly His

2065 2070 2075 2080

Ala Glu Ser Ser Ile Tyr Trp Glu Asp Pro Glu Thr Gly Ile Leu Cys

2085 2090 2095

Arg Cys Arg Pro Asp Lys Ile Ile Pro Glu Phe His Trp Ile Met Asp

2100 2105 2110

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

2115 2120 2125

Tyr Arg Tyr His Val Gln Asp Ala Phe Tyr Ser Asp Gly Tyr Glu Ala

2130 2135 2140

Gln Phe Gly Val Gln Pro Thr Phe Val Phe Leu Val Ala Ser Thr Thr

2145 2150 2155 2160

Ile Glu Cys Gly Arg Tyr Pro Val Glu Ile Phe Met Met Gly Glu Glu

2165 2170 2175

Ala Lys Leu Ala Gly Gln Gln Glu Tyr His Arg Asn Leu Arg Thr Leu

2180 2185 2190

Ser Asp Cys Leu Asn Thr Asp Glu Trp Pro Ala Ile Lys Thr Leu Ser

2195 2200 2205

Leu Pro Arg Trp Ala Lys Glu Tyr Ala Asn Asp Leu Gln Lys Arg Pro

2210 2215 2220

Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His

2225 2230 2235 2240

His His His

<210> 60

<211> 6732

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1-RecE

<400> 60

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agttttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttatttttaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800

gacaaggatg agcatttta tctagtattt gaggagtgct actttgagct agcgaatata 1860

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920
tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980
gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040
aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100
aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160
gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220
acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340
gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400
gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460
agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520
tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580
aatcttcaag atgtggttta taagctaaat ggtgaggcag agcttttta tcgtaaacaa 2640
tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760
gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820
tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880
atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940
atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000
aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060
aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180
cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240
aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300
cagctaacag caccttttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360
tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420
tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480
tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020

cagaatagga ataaccaagc ggccgcactc gacctgcaga tgagcacaaa accactcttc 4080

ctgttacgga aagcgaaaaa atcatccggt gaacctgacg tcgtcctgtg ggcaagcaac 4140

gattttgaat cgacctgtgc cactctggac tacctgatcg ttaagtcagg taaaaaactg 4200

agcagctatt ttaaagctgt tgccacgaat tttcctgtcg ttaatgacct gcccgctgaa 4260

ggtgagatcg attttacctg gagtgaacgc tatcaactca gcaaagactc catgacatgg 4320

gaactaaaac cgggagcagc accagacaac gctcactatc aaggcaatac caacgtcaac 4380

ggcgaagaca tgactgagat tgaggagaat atgctactcc caatttctgg ccaggaactg 4440

cccattcgtt ggcttgctca acacggcagc gaaaaaccgg taacgcacgt ttcacgcgac 4500

ggactccagg cattacacat tgctcgggct gaagaactac cggctgttac tgccctggct 4560

gtttcccaca aaaccagcct gctcgacccg ctggaaattc gcgaactcca caaactggtt 4620

cgtgacactg acaaagtttt ccctaatcct ggtaattcaa acctgggact gataactgct 4680

tttttcgaag catacctgaa cgctgactac accgatcgag gactgctgac aaaagagtgg 4740

atgaagggta atcgtgtttc acacatcact cgcacggctt ccggtgctaa tgctggcggc 4800

ggaaacctca ccgatcgcgg cgaaggtttc gtacacgatc tgacgtcact ggcgcgcgac 4860

gtagccactg gcgtactggc ccgttcaatg gatctggaca tctataacct tcatccggca 4920

cacgctaaac gcattgagga aattatcgct gaaaataaac cgccctttc tgttttccgc 4980

gacaaattca tcaccatgcc tggcgggctg gattattccc gcgccatcgt ggttgcgtcc 5040

gtaaaagaag caccaattgg gatcgaggtc atccccgcgc acgtcactga atatctgaac 5100

aaagtactga ctgaaaccga tcatgccaac cctgatccgg aaatcgtgga tattgcctgc 5160

ggtcgctcct ctgccccgat gccgcagcga gtaacagaag aaggaaaaca ggatgatgaa 5220

gaaaaaccgc aaccatctgg aacaacggca gttgaacagg gagaggctga aacaatggaa 5280

ccggacgcaa ctgaacatca tcaggacacg cagccgctgg atgctcagtc acaggtaaat 5340

tctgttgatg cgaaatatca ggaactgcgg gcagaactcc atgaagcccg gaaaaacatt 5400

ccatcaaaaa atcctgtcga tgacgataaa ttgcttgctg catcacgtgg tgaatttgtt 5460

gacggaatta gcgacccgaa cgatccgaaa tgggtaaagg ggatccagac tcgcgattgt 5520

gtgtaccaga accagccaga aacggaaaaa accagcccag atatgaatca acctgagcca 5580

gtagtgcaac aggaaccgga aatagcctgc aatgcctgcg gccagactgg cggggataac 5640

tgccctgact gtggtgcggt gatgggcgac gcaacatacc aggaaacatt cgatgaagag 5700

agtcaggttg aagctaagga aaatgatccg gaggaaatgg aaggcgctga acatccgcac 5760

aatgagaatg ctggcagcga tccgcatcgc gattgcagtg atgaaactgg cgaagtcgca 5820

gatcccgtaa tcgtagaaga catagagcca ggtatttatt acggaatttc gaatgagaat 5880

taccacgcgg gtcccggtat cagtaagtct cagctcgatg acattgctga tactccggca 5940

ctatatttgt ggcgtaaaaa tgcccccgtg gacaccacaa agacaaaaac gctcgattta 6000

ggaactgctt tccactgccg ggtacttgaa ccggaagaat tcagtaaccg ctttatcgta 6060

gcacctgaat ttaaccgccg tacaaacgcc ggaaaagaag aagagaaagc gtttctgatg 6120

gaatgcgcaa gcacaggaaa aacggttatc actgcggaag aaggccggaa aattgaactc 6180

atgtatcaaa gcgttatggc tttgccgctg gggcaatggc ttgttgaaag cgccggacac 6240

gctgaatcat caatttactg ggaagatcct gaaacaggaa ttttgtgtcg gtgccgtccg 6300

gacaaaatta tccctgaatt tcactggatc atggacgtga aaactacggc ggatattcaa 6360

cgattcaaaa ccgcttatta cgactaccgc tatcacgttc aggatgcatt ctacagtgac 6420

ggttatgaag cacagtttgg agtgcagcca actttcgttt ttctggttgc cagcacaact 6480

attgaatgcg gacgttatcc ggttgaaatt ttcatgatgg gcgaagaagc aaaactggca 6540

ggtcaacagg aatatcaccg caatctgcga accctgtctg actgcctgaa taccgatgaa 6600

tggccagcta ttaagacatt atcactgccc cgctgggcta aggaatatgc aaatgacctg 6660

cagaaaaggc cggcggccac gaaaaaggcc ggccaggcaa aaaagaaaaa gcaccaccac 6720

caccaccact ga 6732

<210> 61

<211> 1885

<212> PRT

<213> Artificial Sequence

<220><223> FnCpf1-hTdT

<400> 61

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro

1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30
Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile
35 40 45
Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe
50 55 60
Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

65 70 75 80
Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys
85 90 95
Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser
100 105 110
Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr
115 120 125
Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys
130 135 140

Ser Ala Lys Asp Thr Ile Lys Lys Gln Ile Ser Glu Tyr Ile Lys Asp
145 150 155 160
Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys
165 170 175
Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp
180 185 190
Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp
195 200 205
Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

210 215 220
Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile
225 230 235 240
Pro Thr Ser Ile Ile Tyr Arg Ile Val Asp Asp Asn Leu Pro Lys Phe
245 250 255
Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu
260 265 270

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr
275 280 285

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

355 360 365
Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser Phe Val Ile
370 375 380
Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe
385 390 395 400
Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys
405 410 415
Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln Lys Leu Asp
420 425 430

Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr Asp Leu Ser
435 440 445
Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala Val Leu Glu
450 455 460
Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys
465 470 475 480
Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu
485 490 495
Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg

500 505 510
Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala

515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu

530 535 540

Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu

545 550 555 560

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

565 570 575

Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln

580 585 590

Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu

595 600 605

Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr

610 615 620

Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640

Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys

645 650 655

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

660 665 670

Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp

675 680 685

Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr

690 695 700

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

705 710 715 720

Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile

725 730 735

Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr

740 745 750

Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe

755 760 765

Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe

770 775 780

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

785 790 795 800

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

805 810 815

Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln

820 825 830

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu

835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp

850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln

865 870 875 880

Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn

885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu

900 905 910

Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro

915 920 925

Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu

930 935 940

Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser

945 950 955 960

Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly

965 970 975

Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp

980 985 990

Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp

995 1000 1005

Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1010 1015 1020

Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu
1025 1030 1035 1040
Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly
1045 1050 1055
Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu
1060 1065 1070
Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn

1075 1080 1085
Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala
1090 1095 1100
Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr
1105 1110 1115 1120
Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe
1125 1130 1135
Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu
1140 1145 1150

Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr
1155 1160 1165
Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys
1170 1175 1180
Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg
1185 1190 1195 1200
Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr
1205 1210 1215
Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1220 1225 1230
Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe
1235 1240 1245
Phe Ala Lys Leu Thr Ser Val Leu Asn Thr Ile Leu Gln Met Arg Asn
1250 1255 1260
Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp
1265 1270 1275 1280

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro
1285 1290 1295

Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu
1300 1305 1310
Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn
1315 1320 1325
Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn
1330 1335 1340
Asn Gln Ala Ala Ala Leu Asp Leu Glu Met Asp Pro Pro Arg Ala Ser
1345 1350 1355 1360
His Leu Ser Pro Arg Lys Lys Arg Pro Arg Gln Thr Gly Ala Leu Met

1365 1370 1375
Ala Ser Ser Pro Gln Asp Ile Lys Phe Gln Asp Leu Val Val Phe Ile
1380 1385 1390
Leu Glu Lys Lys Met Gly Thr Thr Arg Arg Ala Phe Leu Met Glu Leu
1395 1400 1405
Ala Arg Arg Lys Gly Phe Arg Val Glu Asn Glu Leu Ser Asp Ser Val
1410 1415 1420
Thr His Ile Val Ala Glu Asn Asn Ser Gly Ser Asp Val Leu Glu Trp
1425 1430 1435 1440

Leu Gln Ala Gln Lys Val Gln Val Ser Ser Gln Pro Glu Leu Leu Asp
1445 1450 1455
Val Ser Trp Leu Ile Glu Cys Ile Gly Ala Gly Lys Pro Val Glu Met
1460 1465 1470
Thr Gly Lys His Gln Leu Val Val Arg Arg Asp Tyr Ser Asp Ser Thr
1475 1480 1485
Asn Pro Gly Pro Pro Lys Thr Pro Pro Ile Ala Val Gln Lys Ile Ser
1490 1495 1500
Gln Tyr Ala Cys Gln Arg Arg Thr Thr Leu Asn Asn Cys Asn Gln Ile

1505 1510 1515 1520
Phe Thr Asp Ala Phe Asp Ile Leu Ala Glu Asn Cys Glu Phe Arg Glu

1525 1530 1535

Asn Glu Asp Ser Cys Val Thr Phe Met Arg Ala Ala Ser Val Leu Lys

1540 1545 1550

Ser Leu Pro Phe Thr Ile Ile Ser Met Lys Asp Thr Glu Gly Ile Pro

1555 1560 1565

Cys Leu Gly Ser Lys Val Lys Gly Ile Ile Glu Glu Ile Ile Glu Asp

1570 1575 1580

Gly Glu Ser Ser Glu Val Lys Ala Val Leu Asn Asp Glu Arg Tyr Gln

1585 1590 1595 1600

Ser Phe Lys Leu Phe Thr Ser Val Phe Gly Val Gly Leu Lys Thr Ser

1605 1610 1615

Glu Lys Trp Phe Arg Met Gly Phe Arg Thr Leu Ser Lys Val Arg Ser

1620 1625 1630

Asp Lys Ser Leu Lys Phe Thr Arg Met Gln Lys Ala Gly Phe Leu Tyr

1635 1640 1645

Tyr Glu Asp Leu Val Ser Cys Val Thr Arg Ala Glu Ala Glu Ala Val

1650 1655 1660

Ser Val Leu Val Lys Glu Ala Val Trp Ala Phe Leu Pro Asp Ala Phe

1665 1670 1675 1680

Val Thr Met Thr Gly Gly Phe Arg Arg Gly Lys Lys Met Gly His Asp

1685 1690 1695

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

1700 1705 1710

Leu Leu Gln Lys Val Met Asn Leu Trp Glu Lys Lys Gly Leu Leu Leu

1715 1720 1725

Tyr Tyr Asp Leu Val Glu Ser Thr Phe Glu Lys Leu Arg Leu Pro Ser

1730 1735 1740

Arg Lys Val Asp Ala Leu Asp His Phe Gln Lys Cys Phe Leu Ile Phe

1745 1750 1755 1760

Lys Leu Pro Arg Gln Arg Val Asp Ser Asp Gln Ser Ser Trp Gln Glu

1765 1770 1775

Gly Lys Thr Trp Lys Ala Ile Arg Val Asp Leu Val Leu Cys Pro Tyr

1780 1785 1790

Glu Arg Arg Ala Phe Ala Leu Leu Gly Trp Thr Gly Ser Arg Gln Phe

1795 1800 1805

Glu Arg Asp Leu Arg Arg Tyr Ala Thr His Glu Arg Lys Met Ile Leu

1810 1815 1820

Asp Asn His Ala Leu Tyr Asp Lys Thr Lys Arg Ile Phe Leu Lys Ala

1825 1830 1835 1840

Glu Ser Glu Glu Glu Ile Phe Ala His Leu Gly Leu Asp Tyr Ile Glu

1845 1850 1855

Pro Trp Glu Arg Asn Ala Leu Glu Lys Arg Pro Ala Ala Thr Lys Lys

1860 1865 1870

Ala Gly Gln Ala Lys Lys Lys Lys His His His His His His

1875 1880 1885

<210> 62

<211> 5658

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1-hTdT

<400> 62

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agttttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttattttaaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800

gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040

aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100

aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160

gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220

acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580

aatcttcaag atgtggttta taagctaaat ggtgaggcag agctttttta tcgtaaacaa 2640

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag caccttttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020

cagaatagga ataaccaagc ggccgcactc gacctcgaga tggatccacc acgagcgtcc 4080

cacttgagcc ctcggaagaa gagaccccgg cagacgggtg ccttgatggc ctcctctcct 4140

caagacatca aatttcaaga tttggtcgtc ttcattttgg agaagaaaat gggaaccacc 4200

cgcagagcgt tcctcatgga gctggcccgc aggaaagggt tcagggttga aaatgagctc 4260

agtgattctg tcacccacat tgtagcagag aacaactcgg gttcggatgt tctggagtgg 4320

cttcaagcac agaaagtaca agtcagctca caaccagagc tcctcgatgt ctcctggctg 4380

atcgaatgca taggagcagg gaaaccggtg gaaatgacag gaaaacacca gcttgttgtg 4440

agaagagact attcagatag caccaaccca ggccccccga agactccacc aattgctgta 4500
caaaagatct cccagtatgc gtgtcagaga agaaccactt taaacaactg taaccagata 4560
ttcacggatg cctttgatat actggctgaa aactgtgagt ttagagaaaa tgaagactcc 4620
tgtgtgacat ttatgagagc agcttctgta ttgaaatctc tgccattcac aatcatcagt 4680
atgaaggaca cagaaggaat tccctgcctg gggtccaagg tgaagggtat catagaggag 4740
attattgaag atggagaaag ttctgaagtt aaagctgtgt taaatgatga acgatatcaa 4800
tccttcaaac tctttacttc tgtatttgga gtggggctga agacttctga gaagtggttc 4860

aggatgggtt tcagaactct gagtaaagta aggtcggaca aaagcctgaa atttacacga 4920
atgcagaaag caggatttct gtattatgaa gaccttgtca gctgtgtgac cagggcagaa 4980
gcagaggccg tcagtgtgct ggttaaagag gctgtctggg catttcttcc ggatgctttc 5040
gtcaccatga caggagggtt ccggaggggt aagaagatgg ggcatgatgt agatttttta 5100
attaccagcc caggatcaac agaggatgaa gagcaacttt tacagaaagt gatgaactta 5160
tgggaaaaga agggattact tttatattat gaccttgtgg agtcaacatt tgaaaagctc 5220
aggttgccta gcaggaaggt tgatgctttg gatcattttc aaaagtgctt tctgattttc 5280

aaattgcctc gtcaaagagt ggacagtgac cagtccagct ggcaggaagg aaagacctgg 5340
aaggccatcc gtgtggattt agttctgtgc ccctacgagc gtcgtgcctt tgccctgttg 5400
ggatggactg gctcccggca gtttgagaga gacctccggc gctatgccac acatgagcgg 5460
aagatgattc tggataacca tgctttatat gacaagacca agaggatatt cctcaaagca 5520
gaaagtgaag aagaaatttt tgcgcatctg ggattggatt atattgaacc gtgggaaaga 5580
aatgccctcg agaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag 5640
caccaccacc accactga 5658

<210> 63
<211> 1652
<212> PRT
<213> Artificial Sequence
<220><223> FnCpf1-T5
<400> 63
Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile

35 40 45

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe

50 55 60

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

65 70 75 80

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys

85 90 95

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

100 105 110

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr

115 120 125

Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys

130 135 140

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

145 150 155 160

Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys

165 170 175

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp

180 185 190

Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp

195 200 205

Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

210 215 220

Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile

225 230 235 240

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

245 250 255

Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu

260 265 270

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr
275 280 285
Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser
290 295 300
Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln
305 310 315 320
Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn
325 330 335
Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr

340 345 350
Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys Met Ser Val
355 360 365
Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser Phe Val Ile
370 375 380
Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe
385 390 395 400
Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys
405 410 415

Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln Lys Leu Asp
420 425 430
Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr Asp Leu Ser
435 440 445
Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala Val Leu Glu
450 455 460
Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys
465 470 475 480
Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu

485 490 495
Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg
500 505 510
Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala
515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu
530 535 540
Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu
545 550 555 560

Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp Leu Leu Asp
565 570 575
Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln
580 585 590
Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu
595 600 605
Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr
610 615 620
Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640
Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys
645 650 655
Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys Asp Asp Lys
660 665 670
Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp
675 680 685
Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr
690 695 700

Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val Phe Phe Ser
705 710 715 720
Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile
725 730 735
Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr
740 745 750
Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe
755 760 765
Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe

770 775 780

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

785 790 795 800

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

805 810 815

Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln

820 825 830

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu

835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp

850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln

865 870 875 880

Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn

885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu

900 905 910

Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro

915 920 925

Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu

930 935 940

Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser

945 950 955 960

Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly

965 970 975

Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp

980 985 990

Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp

995 1000 1005

Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1010 1015 1020

Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu

1025 1030 1035 1040

Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly

1045 1050 1055

Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu

1060 1065 1070

Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn

1075 1080 1085

Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala

1090 1095 1100

Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr

1105 1110 1115 1120

Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe

1125 1130 1135

Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu

1140 1145 1150

Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr

1155 1160 1165

Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys

1170 1175 1180

Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg

1185 1190 1195 1200

Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr

1205 1210 1215

Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1220 1225 1230

Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe

1235 1240 1245

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

1250 1255 1260

Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp

1265 1270 1275 1280

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro
1285 1290 1295
Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu
1300 1305 1310
Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn
1315 1320 1325
Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn
1330 1335 1340
Asn Gln Ala Ala Ala Leu Asp Leu Gln Met Ala Ser Arg Arg Asn Leu

1345 1350 1355 1360
Met Ile Val Asp Gly Thr Asn Leu Gly Phe Arg Phe Lys His Asn Asn
1365 1370 1375
Ser Lys Lys Pro Phe Ala Ser Ser Tyr Val Ser Thr Ile Gln Ser Leu
1380 1385 1390
Ala Lys Ser Tyr Ser Ala Arg Thr Thr Ile Val Leu Gly Asp Lys Gly
1395 1400 1405
Lys Ser Val Phe Arg Leu Glu His Leu Pro Glu Tyr Lys Gly Asn Arg
1410 1415 1420

Asp Glu Lys Tyr Ala Gln Arg Thr Glu Glu Glu Lys Ala Leu Asp Glu
1425 1430 1435 1440
Gln Phe Phe Glu Tyr Leu Lys Asp Ala Phe Glu Leu Cys Lys Thr Thr
1445 1450 1455
Phe Pro Thr Phe Thr Ile Arg Gly Val Glu Ala Asp Asp Met Ala Ala
1460 1465 1470
Tyr Ile Val Lys Leu Ile Gly His Leu Tyr Asp His Val Trp Leu Ile
1475 1480 1485
Ser Thr Asp Gly Asp Trp Asp Thr Leu Leu Thr Asp Lys Val Ser Arg

1490 1495 1500
Phe Ser Phe Thr Thr Arg Arg Glu Tyr His Leu Arg Asp Met Tyr Glu
1505 1510 1515 1520
His His Asn Val Asp Asp Val Glu Gln Phe Ile Ser Leu Lys Ala Ile
1525 1530 1535

Met Gly Asp Leu Gly Asp Asn Ile Arg Gly Val Glu Gly Ile Gly Ala
1540 1545 1550
Lys Arg Gly Tyr Asn Ile Ile Arg Glu Phe Gly Asn Val Leu Asp Ile
1555 1560 1565

Ile Asp Gln Leu Pro Leu Pro Gly Lys Gln Lys Tyr Ile Gln Asn Leu
1570 1575 1580
Asn Ala Ser Glu Glu Leu Leu Phe Arg Asn Leu Ile Leu Val Asp Leu
1585 1590 1595 1600
Pro Thr Tyr Cys Val Asp Ala Ile Ala Ala Val Gly Gln Asp Val Leu
1605 1610 1615
Asp Lys Phe Thr Lys Asp Ile Leu Glu Ile Ala Glu Gln Leu Gln Lys
1620 1625 1630
Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His

1635 1640 1645
His His His His
1650

<210> 64
<211> 4959
<212> DNA
<213> Artificial Sequence
<220><223> FnCpf1-T5
<400> 64

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60
atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120
caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180
actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300
aaatatcatc agttttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360
ttacaaaact attctgatgt ttatttttaa cttaaaaaga gtgatgatga taatctacaa 420
aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480
tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540
tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720
cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780
gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840
aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900
agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960
agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020
aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140
tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200
tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260
ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320
gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380
gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440
aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560
gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620
aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680
caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740
ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800
gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860
gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980
gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040
aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100
aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160
gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220
acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280
gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580

aatcttcaag atgtggttta taagctaaat ggtgaggcag agctttttta tcgtaaacaa 2640

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag cacctttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020

cagaatagga ataaccaagc ggccgcactc gacctgcaga tggcttcccg tcgtaatcta 4080

atgattgtcg atggaactaa cttaggcttt cgcttcaaac ataacaatag taaaaaacca 4140

tttgcctcaa gttatgtttc aactattcaa tctctggcaa aatcctactc tgccagaact 4200

acgattgttc taggtgataa gggaaaatct gtatttcgtc tagaacatct accagagtat 4260

aaaggtaatc gtgatgaaaa gtacgcacaa cgtacggaag aggagaaagc gctagatgag 4320

cagttctttg agtatttgaa ggatgctttc gagttgtgta aaactacatt cccaactttt 4380

accattcgtg gtgtagaagc agacgatatg gcagcttata ttgttaagct catcgggcat 4440

ctttatgatc acgtttggct aatatctaca gatggtgact gggatacttt attaacggat 4500

aaagtttctc gtttttcttt cacaacacgt cgtgagtatc atcttcgtga tatgtatgaa 4560

catcataatg ttgatgatgt tgagcagttt atctccctga aagcaattat gggagatcta 4620

ggagataata ttcgtggtgt tgaaggaata ggagcaaaac gcggatataa tattattcgt 4680

gagtttggta acgtactgga tattattgat cagcttccac tgcctggaaa gcagaaatat 4740

atacagaacc tgaatgcatc ggaagaactg cttttccgaa acttgattct ggttgattta 4800

cctacctact gtgtggatgc tattgctgct gtaggtcaag atgtgttaga taagtttaca 4860

aaagatattt tggagattgc agaacaactg cagaaaaggc cggcggccac gaaaaaggcc 4920

ggccaggcaa aaaagaaaaa gcaccaccac caccactga 4959

<210> 65

<211> 1602

<212> PRT

<213> Artificial Sequence

<220><223> FnCpf1-lambda

<400> 65

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile
35 40 45

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe
50 55 60

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly
65 70 75 80

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys
85 90 95

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

100 105 110

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr

115 120 125

Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys

130 135 140

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

145 150 155 160

Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys

165 170 175

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp

180 185 190

Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp

195 200 205

Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

210 215 220

Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile

225 230 235 240

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

245 250 255

Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu

260 265 270

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr

275 280 285

Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser

290 295 300

Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln

305 310 315 320

Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn

325 330 335

Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr

340 345 350

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

355 360 365

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

370 375 380

Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe

385 390 395 400

Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys

405 410 415

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

420 425 430

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

435 440 445

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

450 455 460

Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys

465 470 475 480

Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu

485 490 495

Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg

500 505 510

Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala

515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu

530 535 540

Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu

545 550 555 560

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

565 570 575

Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln

580 585 590

Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu

595 600 605

Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr
610 615 620
Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys
625 630 635 640
Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys
645 650 655
Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys Asp Asp Lys
660 665 670
Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp

675 680 685
Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr
690 695 700
Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val Phe Phe Ser
705 710 715 720
Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile
725 730 735
Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr
740 745 750

Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe
755 760 765
Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe
770 775 780
Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu Phe Tyr Arg
785 790 795 800
Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn Ile Ser Glu
805 810 815
Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln

820 825 830
Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu
835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp
850 855 860
Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln
865 870 875 880
Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn
885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu
900 905 910
Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro
915 920 925
Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu
930 935 940
Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser
945 950 955 960
Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly

965 970 975
Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp
980 985 990
Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp
995 1000 1005
Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu
1010 1015 1020
Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu
1025 1030 1035 1040

Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly
1045 1050 1055
Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu
1060 1065 1070
Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn
1075 1080 1085
Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala
1090 1095 1100

Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr

1105 1110 1115 1120

Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe

1125 1130 1135

Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu

1140 1145 1150

Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr

1155 1160 1165

Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys

1170 1175 1180

Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg

1185 1190 1195 1200

Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr

1205 1210 1215

Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1220 1225 1230

Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe

1235 1240 1245

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

1250 1255 1260

Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp

1265 1270 1275 1280

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro

1285 1290 1295

Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu

1300 1305 1310

Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn

1315 1320 1325

Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn

1330 1335 1340

Asn Gln Ala Ala Ala Leu Asp Leu Glu Met Thr Pro Asp Ile Ile Leu

1345 1350 1355 1360

Gln Arg Thr Gly Ile Asp Val Arg Ala Val Glu Gln Gly Asp Asp Ala

1365 1370 1375

Trp His Lys Leu Arg Leu Gly Val Ile Thr Ala Ser Glu Val His Asn

1380 1385 1390

Val Ile Ala Lys Pro Arg Ser Gly Lys Lys Trp Pro Asp Met Lys Met

1395 1400 1405

Ser Tyr Phe His Thr Leu Leu Ala Glu Val Cys Thr Gly Val Ala Pro

1410 1415 1420

Glu Val Asn Ala Lys Ala Leu Ala Trp Gly Lys Gln Tyr Glu Asn Asp

1425 1430 1435 1440

Ala Arg Thr Leu Phe Glu Phe Thr Ser Gly Val Asn Val Thr Glu Ser

1445 1450 1455

Pro Ile Ile Tyr Arg Asp Glu Ser Met Arg Thr Ala Cys Ser Pro Asp

1460 1465 1470

Gly Leu Cys Ser Asp Gly Asn Gly Leu Glu Leu Lys Cys Pro Phe Thr

1475 1480 1485

Ser Arg Asp Phe Met Lys Phe Arg Leu Gly Gly Phe Glu Ala Ile Lys

1490 1495 1500

Ser Ala Tyr Met Ala Gln Val Gln Tyr Ser Met Trp Val Thr Arg Lys

1505 1510 1515 1520

Asn Ala Trp Tyr Phe Ala Asn Tyr Asp Pro Arg Met Lys Arg Glu Gly

1525 1530 1535

Leu His Tyr Val Val Ile Glu Arg Asp Glu Lys Tyr Met Ala Ser Phe

1540 1545 1550

Asp Glu Ile Val Pro Glu Phe Ile Glu Lys Met Asp Glu Ala Leu Ala

1555 1560 1565

Glu Ile Gly Phe Val Phe Gly Glu Gln Trp Arg Leu Glu Lys Arg Pro

1570 1575 1580

Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His

1585 1590 1595 1600

His His

<210> 66

<211> 4809

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1-lambda

<400> 66

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agttttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttattttaaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800

gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040

aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100

aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160

gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220

acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580

aatcttcaag atgtggttta taagctaaat ggtgaggcag agcttttta tcgtaaacaa 2640

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag cacctttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360
tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420
tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480
tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540
aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600
aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660
aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780
caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840
gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900
gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960
aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020
cagaatagga ataaccaagc ggccgcactc gacctcgaga tgacaccgga cattatcctg 4080
cagcgtaccg ggatcgatgt gagagctgtc gaacaggggg atgatgcgtg gcacaaatta 4140

cggctcggcg tcatcaccgc ttcagaagtt cacaacgtga tagcaaaacc ccgctccgga 4200
aagaagtggc ctgacatgaa aatgtcctac ttccacaccc tgcttgctga ggtttgcacc 4260
ggtgtggctc cggaagttaa cgctaaagca ctggcctggg gaaaacagta cgagaacgac 4320
gccagaaccc tgtttgaatt cacttccggc gtgaatgtta ctgaatcccc gatcatctat 4380
cgcgacgaaa gtatgcgtac cgcctgctct cccgatggtt tatgcagtga cggcaacggc 4440
cttgaactga aatgcccgtt tacctcccgg gatttcatga agttccggct cggtggtttc 4500
gaggccataa agtcagctta catggcccag gtgcagtaca gcatgtgggt gacgcgaaaa 4560

aatgcctggt actttgccaa ctatgacccg cgtatgaagc gtgaaggcct gcattatgtc 4620
gtgattgagc gggatgaaaa gtacatggcg agttttgacg agatcgtgcc ggagttcatc 4680
gaaaaaatgg acgaggcact ggctgaaatt ggttttgtat ttggggagca atggcgactc 4740
gagaaaaggc cggcggccac gaaaaaggcc ggccaggcaa aaaagaaaaa gcaccaccac 4800
caccactga 4809

<210> 67

<211> 1731

<212> PRT

<213> Artificial Sequence

<220><223> FnCpf1-mungbean

<400> 67

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30
Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile
35 40 45
Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe
50 55 60
Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

65 70 75 80
Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys
85 90 95
Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser
100 105 110
Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr
115 120 125
Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys
130 135 140

Ser Ala Lys Asp Thr Ile Lys Lys Gln Ile Ser Glu Tyr Ile Lys Asp
145 150 155 160
Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys
165 170 175
Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp
180 185 190
Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp
195 200 205
Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

210 215 220
Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile

225 230 235 240

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

245 250 255

Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu

260 265 270

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr

275 280 285

Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser

290 295 300

Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln

305 310 315 320

Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn

325 330 335

Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr

340 345 350

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

355 360 365

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

370 375 380

Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe

385 390 395 400

Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys

405 410 415

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

420 425 430

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

435 440 445

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

450 455 460

Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys

465 470 475 480

Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu

485 490 495

Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg

500 505 510

Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala

515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu

530 535 540

Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu

545 550 555 560

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

565 570 575

Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln

580 585 590

Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu

595 600 605

Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr

610 615 620

Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640

Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys

645 650 655

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

660 665 670

Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp

675 680 685

Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr

690 695 700

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

705 710 715 720

Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile

725 730 735

Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr
740 745 750
Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe
755 760 765
Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe
770 775 780
Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu Phe Tyr Arg

785 790 795 800
Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn Ile Ser Glu
805 810 815
Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln
820 825 830
Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu
835 840 845
His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp
850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln
865 870 875 880
Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn
885 890 895
Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu
900 905 910
Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro
915 920 925
Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu

930 935 940
Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser
945 950 955 960
Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly
965 970 975
Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp
980 985 990

Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp
995 1000 1005

Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu
1010 1015 1020
Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu
1025 1030 1035 1040
Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly
1045 1050 1055
Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu
1060 1065 1070
Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn

1075 1080 1085
Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala
1090 1095 1100
Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr
1105 1110 1115 1120
Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe
1125 1130 1135
Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu
1140 1145 1150

Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr
1155 1160 1165
Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys
1170 1175 1180
Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg
1185 1190 1195 1200
Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr
1205 1210 1215
Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1220 1225 1230
Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe

1235 1240 1245

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

1250 1255 1260

Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp

1265 1270 1275 1280

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro

1285 1290 1295

Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu

1300 1305 1310

Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn

1315 1320 1325

Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn

1330 1335 1340

Asn Gln Ala Ala Ala Leu Asp Leu Gln Met Gln Thr Leu Gln Met Ser

1345 1350 1355 1360

Leu Leu Thr Gln Pro Tyr Val Gln Pro Arg Phe Pro Cys Lys Arg Tyr

1365 1370 1375

Pro Thr Phe Ser Ala Ser Cys Arg Thr Gln Lys Thr Ala Ile Thr Lys

1380 1385 1390

Thr Glu Lys Val Phe Phe Ser Glu Ser Phe Asp Gln Thr Arg Cys Thr

1395 1400 1405

Gln Pro Leu Ser Glu Lys Lys Lys Arg Val Phe Phe Leu Asp Val Asn

1410 1415 1420

Pro Leu Cys Tyr Glu Gly Ser Lys Pro Ser Leu Arg Ser Phe Gly Arg

1425 1430 1435 1440

Trp Leu Ser Leu Phe Leu His Gln Val Ser Leu Thr Asp Pro Val Ile

1445 1450 1455

Ala Val Ile Asp Gly Glu Gly Gly Ser Glu His Arg Arg Lys Leu Leu

1460 1465 1470

Pro Ser Tyr Lys Ala His Arg Lys Lys Phe Met Arg His Met Ser Ser

1475 1480 1485

Gly His Val Gly Arg Ser His Gln Val Ile Asn Asp Val Leu Gly Lys

1490 1495 1500

Cys Asn Val Pro Val Ile Lys Val Ala Gly His Glu Ala Asp Asp Val

1505 1510 1515 1520

Val Ala Thr Leu Ala Gly Gln Val Val Asn Lys Gly Phe Arg Val Val

1525 1530 1535

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

1540 1545 1550

Gln Ile Val Met Pro Leu Pro Glu Leu Gln Arg Trp Ser Phe Tyr Thr

1555 1560 1565

Leu Arg His Tyr Arg Asp Gln Tyr Asn Cys Asp Pro Glu Ser Asp Leu

1570 1575 1580

Ser Phe Arg Cys Ile Val Gly Asp Glu Val Asp Gly Val Pro Gly Ile

1585 1590 1595 1600

Gln His Leu Val Pro Ser Phe Gly Arg Lys Thr Ala Met Lys Leu Ile

1605 1610 1615

Lys Lys His Gly Ser Leu Glu Thr Leu Leu Asn Ala Ala Ala Ile Arg

1620 1625 1630

Thr Val Gly Arg Pro Tyr Ala Gln Asp Ala Leu Lys Asn His Ala Asp

1635 1640 1645

Tyr Leu Arg Arg Asn Tyr Glu Val Leu Ala Leu Lys Arg Asp Val Asn

1650 1655 1660

Ile Gln Leu Tyr Asp Glu Trp Leu Val Lys Arg Asp Asn His Asn Asp

1665 1670 1675 1680

Lys Thr Ala Leu Ser Ser Phe Phe Lys Tyr Leu Gly Glu Ser Lys Glu

1685 1690 1695

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

1700 1705 1710

Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys His His

1715 1720 1725

His His His

1730

<210> 68

<211> 5196

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1-mungbean

<400> 68

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agttttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttatttttaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620
aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680
caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740
ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800
gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860
gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920
tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040
aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100
aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160
gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220
acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280
gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340
gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460
agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520
tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580
aatcttcaag atgtggttta taagctaaat ggtgaggcag agcttttta tcgtaaacaa 2640
tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700
aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760
gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880
atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940
atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000
aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060
aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120
gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180
cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag caccttttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020

cagaatagga ataaccaagc ggccgcactc gacctgcaga tgcaaacgtt acagatgagt 4080

ctgttgacac aaccttacgt tcagcctcgt ttcccttgca agcgttaccc gaccttctcc 4140

gcatcctgca gaactcaaaa gacagcgatc acgaaaacag agaaggtgtt tttcagtgag 4200

tcatttgatc aaacacgttg cacgcagcct ctctcggaaa agaagaagag ggtgttcttt 4260

ttggacgtta acccgctctg ttacgaagga agcaagccca gcttgcgctc cttcgggcgg 4320

tggctctctc tgtttctcca tcaagtcagc ctcactgacc ccgtcattgc tgttattgat 4380

ggagaaggag gcagcgagca tcgcagaaag ttgctacctt catataaagc acataggaaa 4440

aagttcatga gacacatgtc aagtggccat gttgggaggt ctcatcaagt tataaatgat 4500

gttcttggaa aatgcaacgt gccagttata aaggttgctg gtcatgaagc tgatgatgtt 4560

gtagctactc tagctggaca agttgtcaat aaagggtttc gagtggtcat tggctcccct 4620

gataaggatt ttaagcagct tatatctgaa gatgtgcaaa tagttatgcc tttgccagag 4680

ttacaaaggt ggtccttcta cactctgagg cactacaggg atcagtataa ttgtgatcca 4740

gaatctgatc tgagctttag atgcattgta ggtgatgaag tagacggcgt tcctggtatc 4800

cagcatttgg tccctagttt tggtcggaag actgctatga aacttattaa aaaacatggt 4860

tccttggaaa ctttattaaa tgcggctgca ataaggactg taggcagacc atatgcacag 4920

gatgccctca aaaaccatgc tgattacctt cggagaaact atgaagttct tgccttgaaa 4980

agggatgtaa atatccaact ttatgatgag tggttggtta agagagacaa tcacaatgat 5040

aaaactgcac tatcttcctt cttcaaatat ttgggagaaa gtaaggagct cagttacaat 5100

ggcagaccta tctcttacaa tggtctgcag aaaaggccgg cggccacgaa aaaggccggc 5160

caggcaaaaa agaaaaagca ccaccaccac cactga 5196

<210> 69

<211> 1614

<212> PRT

<213> Artificial Sequence

<220><223> FnCpf1-GFP

<400> 69

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro

1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile

35 40 45

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe

50 55 60

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

65 70 75 80

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys

85 90 95

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

100 105 110

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr

115 120 125

Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys

130 135 140

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

145 150 155 160

Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys

165 170 175

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp

180 185 190

Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp
195 200 205
Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe
210 215 220

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

290 295 300
Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln
305 310 315 320
Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn
325 330 335
Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr
340 345 350
Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys Met Ser Val
355 360 365

Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser Phe Val Ile
370 375 380
Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe
385 390 395 400
Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys
405 410 415
Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln Lys Leu Asp
420 425 430
Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr Asp Leu Ser

435 440 445

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

450 455 460

Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys

465 470 475 480

Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu

485 490 495

Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg

500 505 510

Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala

515 520 525

Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu

530 535 540

Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu

545 550 555 560

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

565 570 575

Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln

580 585 590

Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu

595 600 605

Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr

610 615 620

Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys

625 630 635 640

Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys

645 650 655

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

660 665 670

Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp

675 680 685

Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr

690 695 700

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

705 710 715 720

Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile

725 730 735

Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr

740 745 750

Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe

755 760 765

Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe

770 775 780

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

785 790 795 800

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

805 810 815

Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln

820 825 830

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu

835 840 845

His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp

850 855 860

Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln

865 870 875 880

Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn

885 890 895

Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu

900 905 910

Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro

915 920 925

Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu

930 935 940

Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser
945 950 955 960
Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly
965 970 975
Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp
980 985 990
Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp
995 1000 1005
Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1010 1015 1020
Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu
1025 1030 1035 1040
Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly
1045 1050 1055
Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu
1060 1065 1070
Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn
1075 1080 1085

Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala
1090 1095 1100
Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr
1105 1110 1115 1120
Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe
1125 1130 1135
Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu
1140 1145 1150
Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr

1155 1160 1165
Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys
1170 1175 1180
Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg
1185 1190 1195 1200

Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr
1205 1210 1215
Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His
1220 1225 1230

Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe
1235 1240 1245
Phe Ala Lys Leu Thr Ser Val Leu Asn Thr Ile Leu Gln Met Arg Asn
1250 1255 1260
Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp
1265 1270 1275 1280
Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro
1285 1290 1295
Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu

1300 1305 1310
Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn
1315 1320 1325
Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn
1330 1335 1340
Asn Gln Ala Ala Ala Leu Asp Leu Glu Met Ser Lys Gly Glu Glu Leu
1345 1350 1355 1360
Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn
1365 1370 1375

Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr
1380 1385 1390
Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val
1395 1400 1405
Pro Trp Pro Thr Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe
1410 1415 1420
Ser Arg Tyr Pro Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala
1425 1430 1435 1440
Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp

1445 1450 1455

Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu

1460 1465 1470

Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn

1475 1480 1485

Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr

1490 1495 1500

Ile Thr Ala Asp Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Ile

1505 1510 1515 1520

Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln

1525 1530 1535

Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His

1540 1545 1550

Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg

1555 1560 1565

Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His

1570 1575 1580

Gly Met Asp Glu Leu Tyr Lys Leu Glu Lys Arg Pro Ala Ala Thr Lys

1585 1590 1595 1600

Lys Ala Gly Gln Ala Lys Lys Lys Lys His His His His His His

1605 1610

<210> 70

<211> 4845

<212> DNA

<213> Artificial Sequence

<220><223> FnCpf1-GFP

<400> 70

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgaattcga gctccgtcga 120

caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 180

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 240

ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 300

aaatatcatc agtttttat agaggagata ttaagttcgg tttgtattag cgaagattta 360

ttacaaaact attctgatgt ttatttaaa cttaaaaaga gtgatgatga taatctacaa 420

aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 480

tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 540

tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 600

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 660

acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 720

cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 780

gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 840

aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 900

agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 960

agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1020

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1080

ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1140

tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1200

tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1260

ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 1320

gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 1380

gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 1440

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 1500

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 1560

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 1620

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 1680

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 1740

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 1800

gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 1860

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 1920

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 1980

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2040

aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2100

aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2160

gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2220

acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 2280

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 2340

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 2400

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 2460

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 2520

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 2580

aatcttcaag atgtggttta taagctaaat ggtgaggcag agctttttta tcgtaaacaa 2640

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 2700

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 2760

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 2820

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 2880

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 2940

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3000

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3060

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3120

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3180

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3240

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 3300

cagctaacag cacctttga gacttttaaa aagatgggta aacaaacagg tattatctac 3360

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 3420

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 3480

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 3540

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 3600

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 3660

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 3720

tgcggtgaga gcgacaaaaa gttttttgct aagctaacta gtgtcctaaa tactatctta 3780

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 3840

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 3900

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 3960

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4020
cagaatagga ataaccaagc ggccgcactc gacctcgaga tgagtaaagg agaagaactt 4080
ttcactggag ttgtcccaat tcttgttgaa ttagatggtg atgttaatgg gcacaaattt 4140
tctgtcagtg gagagggtga aggtgatgca acatacggaa aacttaccct taaatttatt 4200
tgcactactg gaaaactacc tgttccatgg ccaacacttg tcactacttt ctcttatggt 4260
gttcaatgct tttcaagata cccagatcat atgaagcggc acgacttctt caagagcgcc 4320
atgcctgagg gatacgtgca ggagaggacc atctctttca aggacgacgg gaactacaag 4380

acacgtgctg aagtcaagtt tgagggagac accctcgtca acaggatcga gcttaaggga 4440
atcgatttca aggaggacgg aaacatcctc ggccacaagt tggaatacaa ctacaactcc 4500
cacaacgtat acatcacggc agacaaacaa aagaatggaa tcaaagctaa cttcaaaatt 4560
agacacaaca ttgaagatgg aagcgttcaa ctagcagacc attatcaaca aaatactcca 4620
attggcgatg gccctgtcct tttaccagac aaccattacc tgtccacaca atctgccctt 4680
tcgaaagatc ccaacgaaaa gagagaccac atggtccttc ttgagtttgt aacagctgct 4740
gggattacac atggcatgga tgaactatac aaactcgaga aaaggccggc ggccacgaaa 4800

aaggccggcc aggcaaaaaa gaaaaagcac caccaccacc actga 4845

<210> 71
<211> 1958
<212> PRT
<213> Artificial Sequence
<220><223> RecJ-FnCpf1
<400> 71

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15
Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30
Gly Ser Val Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu

35 40 45
Thr Ala Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr
50 55 60
Ala Ser Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys

65 70 75 80

Gly Met Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu

85 90 95

Ile Leu Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly

100 105 110

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

115 120 125

Met Arg Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg

130 135 140

Phe Glu Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His

145 150 155 160

Ala Arg Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser

165 170 175

His Ala Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val

180 185 190

Thr Asp His His Leu Pro Gly Glu Thr Leu Pro Ala Ala Glu Ala Ile

195 200 205

Ile Asn Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala

210 215 220

Gly Val Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu

225 230 235 240

Arg Asp Gln Gly Trp Phe Asp Glu Arg Gly Ile Ala Ile Pro Asn Leu

245 250 255

Ala Glu Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val

260 265 270

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

275 280 285

Ile Arg Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val

290 295 300

Ala Asn Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala

305 310 315 320

Leu Gly Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val

325 330 335

Gly Val Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu

340 345 350

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

355 360 365

Gly Met Gln Val Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser

370 375 380

Arg Asp Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His

385 390 395 400

Gln Gly Val Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His

405 410 415

Arg Pro Val Ile Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly

420 425 430

Ser Gly Arg Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg

435 440 445

Leu Asp Thr Leu Tyr Pro Gly Met Ile Leu Lys Phe Gly Gly His Ala

450 455 460

Met Ala Ala Gly Leu Ser Leu Glu Glu Asp Lys Phe Glu Leu Phe Gln

465 470 475 480

Gln Arg Phe Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu

485 490 495

Gln Gly Glu Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr

500 505 510

Met Glu Val Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met

515 520 525

Phe Pro Glu Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg

530 535 540

Leu Val Gly Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly

545 550 555 560

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

565 570 575

Pro Asp Asn Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile

580 585 590

Asn Glu Phe Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile

595 600 605

Trp Pro Ile Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg

610 615 620

Met Ser Ile Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr

625 630 635 640

Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys

645 650 655

Ala Arg Gly Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys

660 665 670

Lys Ala Lys Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu

675 680 685

Ile Leu Ser Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser

690 695 700

Asp Val Tyr Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys

705 710 715 720

Asp Phe Lys Ser Ala Lys Asp Thr Ile Lys Lys Gln Ile Ser Glu Tyr

725 730 735

Ile Lys Asp Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile

740 745 750

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

755 760 765

Ser Lys Asp Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr

770 775 780

Asp Ile Asp Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr

785 790 795 800

Thr Tyr Phe Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser

805 810 815

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

820 825 830

Pro Lys Phe Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys
835 840 845
Ala Pro Glu Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu
850 855 860
Glu Leu Thr Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg
865 870 875 880
Val Phe Ser Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr
885 890 895
Leu Asn Gln Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys

900 905 910
Phe Val Asn Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile
915 920 925
Asn Leu Tyr Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys
930 935 940
Met Ser Val Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser
945 950 955 960
Phe Val Ile Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met
965 970 975

Gln Ser Phe Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys
980 985 990
Ser Ile Lys Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln
995 1000 1005
Lys Leu Asp Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr
1010 1015 1020
Asp Leu Ser Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala
1025 1030 1035 1040
Val Leu Glu Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn

1045 1050 1055
Pro Ser Lys Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala
1060 1065 1070

Lys Tyr Leu Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn
1075 1080 1085
Lys His Arg Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala
1090 1095 1100
Asn Phe Ala Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys
1105 1110 1115 1120

Asp Asn Leu Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys
1125 1130 1135
Asp Leu Leu Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp
1140 1145 1150
Leu Leu Asp Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His
1155 1160 1165
Ile Ser Gln Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His
1170 1175 1180
Phe Tyr Leu Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val

1185 1190 1195 1200
Pro Leu Tyr Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser
1205 1210 1215
Asp Glu Lys Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly
1220 1225 1230
Trp Asp Lys Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys
1235 1240 1245
Asp Asp Lys Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile
1250 1255 1260

Phe Asp Asp Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys
1265 1270 1275 1280
Ile Val Tyr Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val
1285 1290 1295
Phe Phe Ser Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile
1300 1305 1310
Leu Arg Ile Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln
1315 1320 1325

Lys Gly Tyr Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe

1330 1335 1340

Ile Asp Phe Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp

1345 1350 1355 1360

Phe Gly Phe Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu

1365 1370 1375

Phe Tyr Arg Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn

1380 1385 1390

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

1395 1400 1405

Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg

1410 1415 1420

Pro Asn Leu His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn

1425 1430 1435 1440

Leu Gln Asp Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr

1445 1450 1455

Arg Lys Gln Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala

1460 1465 1470

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

1475 1480 1485

Tyr Asp Leu Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe

1490 1495 1500

His Cys Pro Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe

1505 1510 1515 1520

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

1525 1530 1535

Ile Leu Ser Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu

1540 1545 1550

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

1555 1560 1565

Gly Asn Asp Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile

1570 1575 1580

Glu Lys Asp Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn

1585 1590 1595 1600

Ile Lys Glu Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile

1605 1610 1615

Ala Lys Leu Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu

1620 1625 1630

Asn Phe Gly Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr

1635 1640 1645

Gln Lys Leu Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe

1650 1655 1660

Lys Asp Asn Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln

1665 1670 1675 1680

Leu Thr Ala Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly

1685 1690 1695

Ile Ile Tyr Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val

1700 1705 1710

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

1715 1720 1725

Ser Gln Glu Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp

1730 1735 1740

Lys Gly Tyr Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys

1745 1750 1755 1760

Ala Ala Lys Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile

1765 1770 1775

Asn Phe Arg Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val

1780 1785 1790

Tyr Pro Thr Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu

1795 1800 1805

Tyr Gly His Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp

1810 1815 1820

Lys Lys Phe Phe Ala Lys Leu Thr Ser Val Leu Asn Thr Ile Leu Gln

1825 1830 1835 1840

Met Arg Asn Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro
1845 1850 1855
Val Ala Asp Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys
1860 1865 1870
Asn Met Pro Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu
1875 1880 1885
Lys Gly Leu Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys
1890 1895 1900
Lys Leu Asn Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln

1905 1910 1915 1920
Asn Arg Asn Asn Gln Ala Ala Ala Leu Glu Lys Arg Pro Ala Ala Thr
1925 1930 1935
Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Ser Thr Pro Pro Pro Pro
1940 1945 1950
Pro Leu Arg Ser Gly Cys
1955

<210> 72
<211> 5877
<212> DNA
<213> Artificial Sequence
<220><223> RecJ-FnCpf1
<400> 72
atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccgtgaaaca acagatacaa 120
cttcgtcgcc gtgaagtcga tgaaacggca gacttgcccg ctgaattgcc tcccttgctg 180
cgccgtttat acgccagccg gggcgtgcgc agtgcgcaag aactggaacg cagtgttaaa 240
ggtatgttgc cctggcagca actgagcggc gtcgaaaagg ccgttgagat cctttacaac 300
gcttttcgcg aaggaacgcg gattattgtg gtcggcgatt ttgacgccga cggcgcgacc 360
agcacggctc taagcgtgct ggcgatgcgc tcgcttggtt gcagcaatat cgactatctg 420
gtaccaaacc gtttcgaaga cggttacggc ttaagcccgg aagtagtcga tcaggcccat 480

gcccgtggcg cgcagttaat tgtcacggtg gataacggta tttcctccca tgcgggcgtt 540

gaacacgctc gctcgttggg cattccggtt attgttaccg atcaccattt gccgggcgaa 600

acattacccg cagcggaagc gatcattaac cctaacttgc gcgactgtaa tttcccgtcg 660

aaatcactgg caggcgtggg tgtggcgttt tatctgatgc tggcgctgcg caccttttg 720

cgcgatcagg gctggtttga tgagcgtggc atcgcaattc ctaacctggc agaactgctg 780

gatctggtcg cgctgggaac agtggcggac gtcgtgccgc tggacgctaa taatcgcatt 840

ctgacctggc aggggatgag tcgcatccgt gccggaaagt gccgtccagg gattaaagcg 900

ctgctggaag tggcaaaccg tgatgcacaa aaactcgccg ccagcgattt aggttttgcg 960

ctggggccac gtctcaatgc tgccggacga ctggacgata tgtccgtcgg tgtggcgctc 1020

ttgctgtgcg acaacatcgg cgaagcgcgc gtgctggcaa atgaactcga tgcgctaaac 1080

cagacgcgaa aagagatcga acaaggaatg caagttgaag ccctgaccct gtgcgagaaa 1140

ctggagcgaa gtcgcgacac gctacccggc gggctggcaa tgtatcaccc cgaatggcat 1200

cagggcgttg tcggtattct ggcttcgcgc atcaaagagc gttttcaccg tccggttatc 1260

gcctttgcgc cagcaggtga tggtacgctg aaaggttcag gtcgctccat tcaggggctg 1320

catatgcgtg atgcactgga gcgattagac acactctacc ctggcatgat actgaagttt 1380

ggcggtcatg cgatggcggc gggtttgtcg ctggaagagg ataaattcga actctttcaa 1440

caacggtttg gcgagctggt taccgagtgg ctggaccctt cgctattgca aggcgaagtg 1500

gtgtcagacg gcccgttaag cccggccgaa atgaccatgg aagtggcgca gctgctgcgc 1560

gatgctggcc cgtgggggca gatgttcccg gagccgctgt ttgatggtca tttccgtctg 1620

ctgcaacagc ggctggtggg cgaacgtcat ttgaaagtca tggtcgaacc ggtcggcggc 1680

ggtccgctgc tggatggtat tgcttttaat gtcgataccg ccctctggcc ggataacggc 1740

gtgcgcgaag tgcaactggc ttacaagctc gatatcaacg agtttcgcgg caaccgcagc 1800

ctgcaaatta tcatcgacaa tatctggcca attggatccg aattcgagct ccgtcgacaa 1860

gcttgcggcc gcatgtcaat ttatcaagaa tttgttaata aatatagttt aagtaaaact 1920

ctaagatttg agttaatccc acagggtaaa acacttgaaa acataaaagc aagaggtttg 1980

attttagatg atgagaaaag agctaaagac tacaaaaagg ctaaacaaat aattgataaa 2040

tatcatcagt tttttataga ggagatatta agttcggttt gtattagcga agatttatta 2100

caaaactatt ctgatgttta ttttaaactt aaaaagagtg atgatgataa tctacaaaaa 2160

gattttaaaa gtgcaaaaga tacgataaag aaacaaatat ctgaatatat aaaggactca 2220

gagaaattta agaatttgtt taatcaaaac cttatcgatg ctaaaaaagg gcaagagtca 2280

gatttaattc tatggctaaa gcaatctaag gataatggta tagaactatt taaagccaat 2340

agtgatatca cagatataga tgaggcgtta gaaataatca aatcttttaa aggttggaca 2400

acttatttta agggttttca tgaaaataga aaaaatgttt atagtagcaa tgatattcct 2460

acatctatta tttataggat agtagatgat aatttgccta aatttctaga aaataaagct 2520

aagtatgaga gtttaaaaga caaagctcca gaagctataa actatgaaca aattaaaaaa 2580

gatttggcag aagagctaac ctttgatatt gactacaaaa catctgaagt taatcaaaga 2640

gtttttcac ttgatgaagt ttttgagata gcaaacttta ataattatct aaatcaaagt 2700

ggtattacta aatttaatac tattattggt ggtaaatttg taaatggtga aaatacaaag 2760

agaaaaggta taaatgaata tataaatcta tactcacagc aaataaatga taaaacactc 2820

aaaaaatata aaatgagtgt tttatttaag caaattttaa gtgatacaga atctaaatct 2880

tttgtaattg ataagttaga agatgatagt gatgtagtta caacgatgca aagtttttat 2940

gagcaaatag cagcttttaa aacagtagaa gaaaaatcta ttaaagaaac actatcttta 3000

ttatttgatg atttaaaagc tcaaaaactt gatttgagta aaatttattt taaaaatgat 3060

aaatctctta ctgatctatc acaacaagtt tttgatgatt atagtgttat tggtacagcg 3120

gtactagaat atataactca acaaatagca cctaaaaatc ttgataaccc tagtaagaaa 3180

gagcaagaat taatagccaa aaaaactgaa aaagcaaaat acttatctct agaaactata 3240

aagcttgcct tagaagaatt taataagcat agagatatag ataaacagtg taggtttgaa 3300

gaaatacttg caaactttgc ggctattccg atgatatttg atgaaatagc tcaaaacaaa 3360

gacaatttgg cacagatatc tatcaaatat caaaatcaag gtaaaaaaga cctacttcaa 3420

gctagtgcgg aagatgatgt taaagctatc aaggatcttt tagatcaaac taataatctc 3480

ttacataaac taaaaatatt tcatattagt cagtcagaag ataaggcaaa tattttagac 3540

aaggatgagc atttttatct agtatttgag gagtgctact ttgagctagc gaatatagtg 3600

cctctttata acaaaattag aaactatata actcaaaagc catatagtga tgagaaattt 3660

aagctcaatt ttgagaactc gactttggct aatggttggg ataaaaataa agagcctgac 3720

aatacggcaa ttttatttat caaagatgat aaatattatc tgggtgtgat gaataagaaa 3780

aataacaaaa tatttgatga taaagctatc aaagaaaata aaggcgaggg ttataaaaaa 3840

attgtttata aacttttacc tggcgcaaat aaaatgttac ctaaggtttt cttttctgct 3900

aaatctataa aattttataa tcctagtgaa gatatactta gaataagaaa tcattccaca 3960

catacaaaaa atggtagtcc tcaaaaagga tatgaaaaat ttgagtttaa tattgaagat 4020

tgccgaaaat ttatagattt ttataaacag tctataagta agcatccgga gtggaaagat 4080

tttggattta gattttctga tactcaaaga tataattcta tagatgaatt ttatagagaa 4140

gttgaaaatc aaggctacaa actaactttt gaaaatatat cagagagcta tattgatagc 4200

gtagttaatc agggtaaatt gtacctattc caaatctata ataaagattt ttcagcttat 4260

agcaaagggc gaccaaatct acatacttta tattggaaag cgctgtttga tgagagaaat 4320
cttcaagatg tggtttataa gctaaatggt gaggcagagc ttttttatcg taaacaatca 4380
atacctaaaa aaatcactca cccagctaaa gaggcaatag ctaataaaaa caaagataat 4440
cctaaaaaag agagtgtttt tgaatatgat ttaatcaaag ataaacgctt tactgaagat 4500
aagtttttct ttcactgtcc tattacaatc aattttaaat ctagtggagc taataagttt 4560
aatgatgaaa tcaatttatt gctaaaagaa aaagcaaatg atgttcatat attaagtata 4620
gatagaggtg aaagacattt agcttactat actttggtag atggtaaagg caatatcatc 4680

aaacaagata ctttcaacat cattggtaat gatagaatga aaacaaacta ccatgataag 4740
cttgctgcaa tagagaaaga tagggattca gctaggaaag actggaaaaa gataaataac 4800
atcaaagaga tgaaagaggg ctatctatct caggtagttc atgaaatagc taagctagtt 4860
atagagtata atgctattgt ggtttttgag gatttaaatt ttggatttaa aagagggcgt 4920
ttcaaggtag agaagcaggt ctatcaaaag ttagaaaaaa tgctaattga gaaactaaac 4980
tatctagttt tcaaagataa tgagtttgat aaaactgggg gagtgcttag agcttatcag 5040
ctaacagcac cttttgagac ttttaaaaag atgggtaaac aaacaggtat tatctactat 5100

gtaccagctg gttttacttc aaaaatttgt cctgtaactg gttttgtaaa tcagttatat 5160
cctaagtatg aaagtgtcag caaatctcaa gagttcttta gtaagtttga caagatttgt 5220
tataaccttg ataagggcta ttttgagttt agttttgatt ataaaaactt tggtgacaag 5280
gctgccaaag gcaagtggac tatagctagc tttgggagta gattgattaa ctttagaaat 5340
tcagataaaa atcataattg ggatactcga gaagtttatc caactaaaga gttggagaaa 5400
ttgctaaaag attattctat cgaatatggg catggcgaat gtatcaaagc agctatttgc 5460
ggtgagagcg acaaaaagtt ttttgctaag ctaactagtg tcctaaatac tatcttacaa 5520

atgcgtaact caaaaacagg tactgagtta gattatctaa tttcaccagt agcagatgta 5580
aatggcaatt tctttgattc gcgacaggcg ccaaaaaata tgcctcaaga tgctgatgcc 5640
aatggtgctt atcatattgg gctaaaaggt ctgatgctac taggtaggat caaaaataat 5700
caagagggca aaaaactcaa tttggttatc aaaaatgaag agtattttga gttcgtgcag 5760
aataggaata accaagcggc cgcactcgag aaaaggccgg cggccacgaa aaaggccggc 5820
caggcaaaaa agaaaaagtc gacaccacca ccaccaccac tgagatccgg ctgctaa 5877

<210> 73

<211> 1619

<212> PRT

<213> Artificial Sequence

<220><223> GFP-FnCpf1

<400> 73

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
35 40 45

Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser
50 55 60

Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe
65 70 75 80

Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr
85 90 95

Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met
100 105 110

Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln
115 120 125

Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala

130 135 140

Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys
145 150 155 160

Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu
165 170 175

Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys
180 185 190

Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly
195 200 205

Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp

210 215 220

Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala

225 230 235 240

Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu

245 250 255

Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys

260 265 270

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile

275 280 285

Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe

290 295 300

Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly

305 310 315 320

Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys

325 330 335

Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser

340 345 350

Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser Asp Val Tyr

355 360 365

Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys

370 375 380

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

385 390 395 400

Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys

405 410 415

Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp

420 425 430

Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr Asp Ile Asp

435 440 445

Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe

450 455 460

Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser Asn Asp Ile

465 470 475 480

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

485 490 495

Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu

500 505 510

Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr

515 520 525

Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser

530 535 540

Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln

545 550 555 560

Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn

565 570 575

Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile Asn Leu Tyr

580 585 590

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

595 600 605

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

610 615 620

Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe

625 630 635 640

Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys

645 650 655

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

660 665 670

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

675 680 685

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

690 695 700

Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys

705 710 715 720

Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu
725 730 735
Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg
740 745 750
Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala
755 760 765
Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu
770 775 780

Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu
785 790 795 800
Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp Leu Leu Asp
805 810 815
Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His Ile Ser Gln
820 825 830
Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu
835 840 845
Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr

850 855 860
Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys
865 870 875 880
Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys
885 890 895
Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys Asp Asp Lys
900 905 910
Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp
915 920 925

Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr
930 935 940
Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val Phe Phe Ser
945 950 955 960
Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile
965 970 975

Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr
980 985 990
Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe

995 1000 1005
Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe
1010 1015 1020
Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu Phe Tyr Arg
1025 1030 1035 1040
Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn Ile Ser Glu
1045 1050 1055
Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln
1060 1065 1070

Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu
1075 1080 1085
His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp
1090 1095 1100
Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln
1105 1110 1115 1120
Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn
1125 1130 1135
Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu

1140 1145 1150
Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro
1155 1160 1165
Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu
1170 1175 1180
Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser
1185 1190 1195 1200
Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly
1205 1210 1215

Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp

1220 1225 1230

Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp

1235 1240 1245

Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu

1250 1255 1260

Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile Ala Lys Leu

1265 1270 1275 1280

Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu Asn Phe Gly

1285 1290 1295

Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu

1300 1305 1310

Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn

1315 1320 1325

Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala

1330 1335 1340

Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr

1345 1350 1355 1360

Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val Thr Gly Phe

1365 1370 1375

Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu

1380 1385 1390

Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr

1395 1400 1405

Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys

1410 1415 1420

Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg

1425 1430 1435 1440

Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr

1445 1450 1455

Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His

1460 1465 1470

Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe

1475 1480 1485

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

1490 1495 1500

Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp

1505 1510 1515 1520

Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro

1525 1530 1535

Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu

1540 1545 1550

Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn

1555 1560 1565

Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn

1570 1575 1580

Asn Gln Ala Ala Ala Leu Glu Lys Arg Pro Ala Ala Thr Lys Lys Ala

1585 1590 1595 1600

Gly Gln Ala Lys Lys Lys Lys Ser Thr Pro Pro Pro Pro Pro Leu Arg

1605 1610 1615

Ser Gly Cys

<210> 74

<211> 4860

<212> DNA

<213> Artificial Sequence

<220><223> GFP-FnCpf1

<400> 74

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatgagtaa aggagaagaa 120

cttttcactg gagttgtccc aattcttgtt gaattagatg gtgatgttaa tgggcacaaa 180

ttttctgtca gtggagaggg tgaaggtgat gcaacatacg gaaaacttac ccttaaattt 240

atttgcacta ctggaaaact acctgttcca tggccaacac ttgtcactac tttctcttat 300

ggtgttcaat gcttttcaag atacccagat catatgaagc ggcacgactt cttcaagagc 360

gccatgcctg agggatacgt gcaggagagg accatctctt tcaaggacga cgggaactac 420

aagacacgtg ctgaagtcaa gtttgaggga gacaccctcg tcaacaggat cgagcttaag 480

ggaatcgatt tcaaggagga cggaaacatc ctcggccaca agttggaata caactacaac 540
tcccacaacg tatacatcac ggcagacaaa caaaagaatg gaatcaaagc taacttcaaa 600
attagacaca acattgaaga tggaagcgtt caactagcag accattatca acaaaatact 660
ccaattggcg atggccctgt ccttttacca gacaaccatt acctgtccac acaatctgcc 720
ctttcgaaag atcccaacga aaagagagac cacatggtcc ttcttgagtt tgtaacagct 780
gctgggatta cacatggcat ggatgaacta tacaaaggat ccgaattcga gctccgtcga 840
caagcttgcg gccgcatgtc aatttatcaa gaatttgtta ataaatatag tttaagtaaa 900

actctaagat ttgagttaat cccacagggt aaaacacttg aaaacataaa agcaagaggt 960
ttgattttag atgatgagaa aagagctaaa gactacaaaa aggctaaaca aataattgat 1020
aaatatcatc agtttttat agaggagata ttaagttcgg tttgtattag cgaagattta 1080
ttacaaaact attctgatgt ttatttttaaa cttaaaaaga gtgatgatga taatctacaa 1140
aaagatttta aaagtgcaaa agatacgata aagaaacaaa tatctgaata tataaaggac 1200
tcagagaaat ttaagaattt gtttaatcaa aaccttatcg atgctaaaaa agggcaagag 1260
tcagatttaa ttctatggct aaagcaatct aaggataatg gtatagaact atttaaagcc 1320

aatagtgata tcacagatat agatgaggcg ttagaaataa tcaaatcttt taaaggttgg 1380
acaacttatt ttaagggttt tcatgaaaat agaaaaaatg tttatagtag caatgatatt 1440
cctacatcta ttatttatag gatagtagat gataatttgc ctaaatttct agaaaataaa 1500
gctaagtatg agagtttaaa agacaaagct ccagaagcta taaactatga acaaattaaa 1560
aaagatttgg cagaagagct aacctttgat attgactaca aaacatctga agttaatcaa 1620
agagtttttt cacttgatga agtttttgag atagcaaact ttaataatta tctaaatcaa 1680
agtggtatta ctaaatttaa tactattatt ggtggtaaat ttgtaaatgg tgaaaataca 1740

aagagaaaag gtataaatga atatataaat ctatactcac agcaaataaa tgataaaaca 1800
ctcaaaaaat ataaaatgag tgttttattt aagcaaattt taagtgatac agaatctaaa 1860
tcttttgtaa ttgataagtt agaagatgat agtgatgtag ttacaacgat gcaaagtttt 1920
tatgagcaaa tagcagcttt taaaacagta gaagaaaaat ctattaaaga aacactatct 1980
ttattatttg atgatttaaa agctcaaaaa cttgatttga gtaaaattta ttttaaaaat 2040
gataaatctc ttactgatct atcacaacaa gtttttgatg attatagtgt tattggtaca 2100
gcggtactag aatatataac tcaacaaata gcacctaaaa atcttgataa ccctagtaag 2160

aaagagcaag aattaatagc caaaaaaact gaaaaagcaa aatacttatc tctagaaact 2220

ataaagcttg ccttagaaga atttaataag catagagata tagataaaca gtgtaggttt 2280

gaagaaatac ttgcaaactt tgcggctatt ccgatgatat ttgatgaaat agctcaaaac 2340

aaagacaatt tggcacagat atctatcaaa tatcaaaatc aaggtaaaaa agacctactt 2400

caagctagtg cggaagatga tgttaaagct atcaaggatc ttttagatca aactaataat 2460

ctcttacata aactaaaaat atttcatatt agtcagtcag aagataaggc aaatatttta 2520

gacaaggatg agcattttta tctagtattt gaggagtgct actttgagct agcgaatata 2580

gtgcctcttt ataacaaaat tagaaactat ataactcaaa agccatatag tgatgagaaa 2640

tttaagctca attttgagaa ctcgactttg gctaatggtt gggataaaaa taaagagcct 2700

gacaatacgg caattttatt tatcaaagat gataaatatt atctgggtgt gatgaataag 2760

aaaaataaca aaatatttga tgataaagct atcaaagaaa ataaaggcga gggttataaa 2820

aaaattgttt ataaactttt acctggcgca aataaaatgt tacctaaggt tttcttttct 2880

gctaaatcta taaaatttta taatcctagt gaagatatac ttagaataag aaatcattcc 2940

acacatacaa aaaatggtag tcctcaaaaa ggatatgaaa aatttgagtt taatattgaa 3000

gattgccgaa aatttataga tttttataaa cagtctataa gtaagcatcc ggagtggaaa 3060

gattttggat ttagattttc tgatactcaa agatataatt ctatagatga attttataga 3120

gaagttgaaa atcaaggcta caaactaact tttgaaaata tatcagagag ctatattgat 3180

agcgtagtta atcagggtaa attgtaccta ttccaaatct ataataaaga tttttcagct 3240

tatagcaaag ggcgaccaaa tctacatact ttatattgga aagcgctgtt tgatgagaga 3300

aatcttcaag atgtggttta taagctaaat ggtgaggcag agcttttta tcgtaaacaa 3360

tcaataccta aaaaaatcac tcacccagct aaagaggcaa tagctaataa aaacaaagat 3420

aatcctaaaa aagagagtgt ttttgaatat gatttaatca aagataaacg ctttactgaa 3480

gataagtttt tctttcactg tcctattaca atcaatttta aatctagtgg agctaataag 3540

tttaatgatg aaatcaattt attgctaaaa gaaaaagcaa atgatgttca tatattaagt 3600

atagatagag gtgaaagaca tttagcttac tatactttgg tagatggtaa aggcaatatc 3660

atcaaacaag atactttcaa catcattggt aatgatagaa tgaaaacaaa ctaccatgat 3720

aagcttgctg caatagagaa agatagggat tcagctagga aagactggaa aaagataaat 3780

aacatcaaag agatgaaaga gggctatcta tctcaggtag ttcatgaaat agctaagcta 3840

gttatagagt ataatgctat tgtggttttt gaggatttaa attttggatt taaaagaggg 3900

cgtttcaagg tagagaagca ggtctatcaa aagttagaaa aaatgctaat tgagaaacta 3960

aactatctag ttttcaaaga taatgagttt gataaaactg ggggagtgct tagagcttat 4020

cagctaacag caccttttga gacttttaaa aagatgggta aacaaacagg tattatctac 4080

tatgtaccag ctggttttac ttcaaaaatt tgtcctgtaa ctggttttgt aaatcagtta 4140

tatcctaagt atgaaagtgt cagcaaatct caagagttct ttagtaagtt tgacaagatt 4200

tgttataacc ttgataaggg ctattttgag tttagttttg attataaaaa ctttggtgac 4260

aaggctgcca aaggcaagtg gactatagct agctttggga gtagattgat taactttaga 4320

aattcagata aaaatcataa ttgggatact cgagaagttt atccaactaa agagttggag 4380

aaattgctaa aagattattc tatcgaatat gggcatggcg aatgtatcaa agcagctatt 4440

tgcggtgaga gcgacaaaaa gtttttgct aagctaacta gtgtcctaaa tactatctta 4500

caaatgcgta actcaaaaac aggtactgag ttagattatc taatttcacc agtagcagat 4560

gtaaatggca atttctttga ttcgcgacag gcgccaaaaa atatgcctca agatgctgat 4620

gccaatggtg cttatcatat tgggctaaaa ggtctgatgc tactaggtag gatcaaaaat 4680

aatcaagagg gcaaaaaact caatttggtt atcaaaaatg aagagtattt tgagttcgtg 4740

cagaatagga ataaccaagc ggccgcactc gagaaaaggc cggcggccac gaaaaaggcc 4800

ggccaggcaa aaaagaaaaa gtcgacacca ccaccaccac cactgagatc cggctgctaa 4860

4860

<210> 75

<211> 1497

<212> PRT

<213> Artificial Sequence

<220><223> SSB-FnCpf1

<400> 75

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ala Ser Arg Gly Val Asn Lys Val Ile Leu Val Gly Asn
35 40 45

Leu Gly Gln Asp Pro Glu Val Arg Tyr Met Pro Asn Gly Gly Ala Val
50 55 60

Ala Asn Ile Thr Leu Ala Thr Ser Glu Ser Trp Arg Asp Lys Ala Thr
65 70 75 80

Gly Glu Met Lys Glu Gln Thr Glu Trp His Arg Val Val Leu Phe Gly

85 90 95

Lys Leu Ala Glu Val Ala Ser Glu Tyr Leu Arg Lys Gly Ser Gln Val

100 105 110

Tyr Ile Glu Gly Gln Leu Arg Thr Arg Lys Trp Thr Asp Gln Ser Gly

115 120 125

Gln Asp Arg Tyr Thr Thr Glu Val Val Val Asn Val Gly Gly Thr Met

130 135 140

Gln Met Leu Gly Gly Arg Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala

145 150 155 160

Cys Gly Arg Met Ser Ile Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu

165 170 175

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

180 185 190

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

195 200 205

Asp Tyr Lys Lys Ala Lys Gln Ile Ile Asp Lys Tyr His Gln Phe Phe

210 215 220

Ile Glu Glu Ile Leu Ser Ser Val Cys Ile Ser Glu Asp Leu Leu Gln

225 230 235 240

Asn Tyr Ser Asp Val Tyr Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn

245 250 255

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

260 265 270

Ser Glu Tyr Ile Lys Asp Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln

275 280 285

Asn Leu Ile Asp Ala Lys Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp

290 295 300

Leu Lys Gln Ser Lys Asp Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser

305 310 315 320

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

325 330 335

Gly Trp Thr Thr Tyr Phe Lys Gly Phe His Glu Asn Arg Lys Asn Val

340 345 350

Tyr Ser Ser Asn Asp Ile Pro Thr Ser Ile Ile Tyr Arg Ile Val Asp

355 360 365

Asp Asn Leu Pro Lys Phe Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu

370 375 380

Lys Asp Lys Ala Pro Glu Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp

385 390 395 400

Leu Ala Glu Glu Leu Thr Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val

405 410 415

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

420 425 430

Asn Asn Tyr Leu Asn Gln Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile

435 440 445

Gly Gly Lys Phe Val Asn Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn

450 455 460

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

465 470 475 480

Lys Tyr Lys Met Ser Val Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu

485 490 495

Ser Lys Ser Phe Val Ile Asp Lys Leu Glu Asp Asp Ser Asp Val Val

500 505 510

Thr Thr Met Gln Ser Phe Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val

515 520 525

Glu Glu Lys Ser Ile Lys Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu

530 535 540

Lys Ala Gln Lys Leu Asp Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys

545 550 555 560

Ser Leu Thr Asp Leu Ser Gln Gln Val Phe Asp Asp Tyr Ser Val Ile

565 570 575

Gly Thr Ala Val Leu Glu Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn

580 585 590

Leu Asp Asn Pro Ser Lys Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr
595 600 605
Glu Lys Ala Lys Tyr Leu Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu
610 615 620
Glu Phe Asn Lys His Arg Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu
625 630 635 640
Ile Leu Ala Asn Phe Ala Ala Ile Pro Met Ile Phe Asp Glu Ile Ala
645 650 655
Gln Asn Lys Asp Asn Leu Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln

660 665 670
Gly Lys Lys Asp Leu Leu Gln Ala Ser Ala Glu Asp Asp Val Lys Ala
675 680 685
Ile Lys Asp Leu Leu Asp Gln Thr Asn Asn Leu Leu His Lys Leu Lys
690 695 700
Ile Phe His Ile Ser Gln Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys
705 710 715 720
Asp Glu His Phe Tyr Leu Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala
725 730 735

Asn Ile Val Pro Leu Tyr Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys
740 745 750
Pro Tyr Ser Asp Glu Lys Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu
755 760 765
Ala Asn Gly Trp Asp Lys Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu
770 775 780
Phe Ile Lys Asp Asp Lys Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn
785 790 795 800
Asn Lys Ile Phe Asp Asp Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly

805 810 815
Tyr Lys Lys Ile Val Tyr Lys Leu Leu Pro Gly Ala Asn Lys Met Leu
820 825 830

Pro Lys Val Phe Phe Ser Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser
835 840 845

Glu Asp Ile Leu Arg Ile Arg Asn His Ser Thr His Thr Lys Asn Gly
850 855 860

Ser Pro Gln Lys Gly Tyr Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys
865 870 875 880

Arg Lys Phe Ile Asp Phe Tyr Lys Gln Ser Ile Ser Lys His Pro Glu
885 890 895

Trp Lys Asp Phe Gly Phe Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser
900 905 910

Ile Asp Glu Phe Tyr Arg Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr
915 920 925

Phe Glu Asn Ile Ser Glu Ser Tyr Ile Asp Ser Val Val Asn Gln Gly
930 935 940

Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser

945 950 955 960

Lys Gly Arg Pro Asn Leu His Thr Leu Tyr Trp Lys Ala Leu Phe Asp
965 970 975

Glu Arg Asn Leu Gln Asp Val Val Tyr Lys Leu Asn Gly Glu Ala Glu
980 985 990

Leu Phe Tyr Arg Lys Gln Ser Ile Pro Lys Lys Ile Thr His Pro Ala
995 1000 1005

Lys Glu Ala Ile Ala Asn Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser
1010 1015 1020

Val Phe Glu Tyr Asp Leu Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys
1025 1030 1035 1040

Phe Phe Phe His Cys Pro Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala
1045 1050 1055

Asn Lys Phe Asn Asp Glu Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn
1060 1065 1070

Asp Val His Ile Leu Ser Ile Asp Arg Gly Glu Arg His Leu Ala Tyr
1075 1080 1085

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

1090 1095 1100

Asn Ile Ile Gly Asn Asp Arg Met Lys Thr Asn Tyr His Asp Lys Leu
1105 1110 1115 1120

Ala Ala Ile Glu Lys Asp Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys
1125 1130 1135

Ile Asn Asn Ile Lys Glu Met Lys Glu Gly Tyr Leu Ser Gln Val Val
1140 1145 1150

His Glu Ile Ala Lys Leu Val Ile Glu Tyr Asn Ala Ile Val Val Phe
1155 1160 1165

Glu Asp Leu Asn Phe Gly Phe Lys Arg Gly Arg Phe Lys Val Glu Lys
1170 1175 1180

Gln Val Tyr Gln Lys Leu Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr
1185 1190 1195 1200

Leu Val Phe Lys Asp Asn Glu Phe Asp Lys Thr Gly Gly Val Leu Arg
1205 1210 1215

Ala Tyr Gln Leu Thr Ala Pro Phe Glu Thr Phe Lys Lys Met Gly Lys
1220 1225 1230

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

1235 1240 1245

Cys Pro Val Thr Gly Phe Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser
1250 1255 1260

Val Ser Lys Ser Gln Glu Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr
1265 1270 1275 1280

Asn Leu Asp Lys Gly Tyr Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe
1285 1290 1295

Gly Asp Lys Ala Ala Lys Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser
1300 1305 1310

Arg Leu Ile Asn Phe Arg Asn Ser Asp Lys Asn His Asn Trp Asp Thr
1315 1320 1325

Arg Glu Val Tyr Pro Thr Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr

1330 1335 1340

Ser Ile Glu Tyr Gly His Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

Ile Ser Pro Val Ala Asp Val Asn Gly Asn Phe Phe Asp Ser Arg Gln

1395 1400 1405

Ala Pro Lys Asn Met Pro Gln Asp Ala Asp Ala Asn Gly Ala Tyr His

1410 1415 1420

Ile Gly Leu Lys Gly Leu Met Leu Leu Gly Arg Ile Lys Asn Asn Gln

1425 1430 1435 1440

Glu Gly Lys Lys Leu Asn Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu

1445 1450 1455

Phe Val Gln Asn Arg Asn Asn Gln Ala Ala Ala Leu Glu Lys Arg Pro

1460 1465 1470

Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Ser Thr Pro

1475 1480 1485

Pro Pro Pro Pro Leu Arg Ser Gly Cys

1490 1495

<210> 76

<211> 4494

<212> DNA

<213> Artificial Sequence

<220><223> SSB-FnCpf1

<400> 76

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggccag cagaggcgta 120

aacaaggtta ttctcgttgg taatctgggt caggacccgg aagtacgcta catgccaaat 180

ggtggcgcag ttgccaacat tacgctggct acttccgaat cctggcgtga taaagcgacc 240

ggcgagatga aagaacagac tgaatggcac cgcgttgtgc tgttcggcaa actggcagaa 300

gtggcgagcg aatatctgcg taaaggttct caggtttata tcgaaggtca gctgcgtacc 360

cgtaaatgga ccgatcaatc cggtcaggat cgctacacca cagaagtcgt ggtgaacgtt 420

ggcggcacca tgcagatgct gggtggtcgt ggatccgaat tcgagctccg tcgacaagct 480

tgcggccgca tgtcaattta tcaagaattt gttaataaat atagtttaag taaaactcta 540

agatttgagt taatcccaca gggtaaaaca cttgaaaaca taaaagcaag aggtttgatt 600

ttagatgatg agaaaagagc taaagactac aaaaaggcta aacaaataat tgataaatat 660

catcagtttt ttatagagga gatattaagt tcggtttgta ttagcgaaga tttattacaa 720

aactattctg atgtttattt taaacttaaa aagagtgatg atgataatct acaaaaagat 780

tttaaaagtg caaaagatac gataaagaaa caaatatctg aatatataaa ggactcagag 840

aaatttaaga atttgtttaa tcaaaacctt atcgatgcta aaaaagggca agagtcagat 900

ttaattctat ggctaaagca atctaaggat aatggtatag aactatttaa agccaatagt 960

gatatcacag atatagatga ggcgttagaa ataatcaaat cttttaaagg ttggacaact 1020

tattttaagg gttttcatga aaatagaaaa aatgtttata gtagcaatga tattcctaca 1080

tctattattt ataggatagt agatgataat ttgcctaaat ttctagaaaa taaagctaag 1140

tatgagagtt taaaagacaa agctccagaa gctataaact atgaacaaat taaaaaagat 1200

ttggcagaag agctaacctt tgatattgac tacaaaacat ctgaagttaa tcaaagagtt 1260

ttttcacttg atgaagtttt tgagatagca aactttaata attatctaaa tcaaagtggt 1320

attactaaat ttaatactat tattggtggt aaatttgtaa atggtgaaaa tacaaagaga 1380

aaaggtataa atgaatatat aaatctatac tcacagcaaa taaatgataa aacactcaaa 1440

aaatataaaa tgagtgtttt atttaagcaa attttaagtg atacagaatc taaatctttt 1500

gtaattgata agttagaaga tgatagtgat gtagttacaa cgatgcaaag ttttttatgag 1560

caaatagcag cttttaaaac agtagaagaa aaatctatta aagaaacact atctttatta 1620

tttgatgatt taaaagctca aaaacttgat ttgagtaaaa tttattttaa aaatgataaa 1680

tctcttactg atctatcaca acaagttttt gatgattata gtgttattgg tacagcggta 1740

ctagaatata taactcaaca aatagcacct aaaaatcttg ataaccctag taagaaagag 1800

caagaattaa tagccaaaaa aactgaaaaa gcaaaatact tatctctaga aactataaag 1860

cttgccttag aagaatttaa taagcataga gatatagata aacagtgtag gtttgaagaa 1920

atacttgcaa actttgcggc tattccgatg atatttgatg aaatagctca aaacaaagac 1980

aatttggcac agatatctat caaatatcaa aatcaaggta aaaaagacct acttcaagct 2040

agtgcggaag atgatgttaa agctatcaag gatcttttag atcaaactaa taatctctta 2100

cataaactaa aaatatttca tattagtcag tcagaagata aggcaaatat tttagacaag 2160

gatgagcatt tttatctagt atttgaggag tgctactttg agctagcgaa tatagtgcct 2220
ctttataaca aaattagaaa ctatataact caaaagccat atagtgatga gaaatttaag 2280
ctcaattttg agaactcgac tttggctaat ggttgggata aaaataaaga gcctgacaat 2340
acggcaattt tatttatcaa agatgataaa tattatctgg gtgtgatgaa taagaaaaat 2400
aacaaaatat ttgatgataa agctatcaaa gaaaataaag gcgagggtta taaaaaaatt 2460
gtttataaac ttttacctgg cgcaaataaa atgttaccta aggttttctt ttctgctaaa 2520
tctataaaat tttataatcc tagtgaagat atacttagaa taagaaatca ttccacacat 2580

acaaaaaatg gtagtcctca aaaaggatat gaaaaatttg agtttaatat tgaagattgc 2640
cgaaaattta tagattttta taaacagtct ataagtaagc atccggagtg gaaagatttt 2700
ggatttagat tttctgatac tcaaagatat aattctatag atgaatttta tagagaagtt 2760
gaaaatcaag gctacaaact aacttttgaa aatatatcag agagctatat tgatagcgta 2820
gttaatcagg gtaaattgta cctattccaa atctataata aagatttttc agcttatagc 2880
aaagggcgac caaatctaca tactttatat tggaaagcgc tgtttgatga gagaaatctt 2940
caagatgtgg tttataagct aaatggtgag gcagagcttt tttatcgtaa acaatcaata 3000

cctaaaaaaa tcactcaccc agctaaagag gcaatagcta ataaaaacaa agataatcct 3060
aaaaaagaga gtgtttttga atatgattta atcaaagata aacgctttac tgaagataag 3120
tttttctttc actgtcctat tacaatcaat tttaaatcta gtggagctaa taagtttaat 3180
gatgaaatca atttattgct aaaagaaaaa gcaaatgatg ttcatatatt aagtatagat 3240
agaggtgaaa gacatttagc ttactatact ttggtagatg gtaaaggcaa tatcatcaaa 3300
caagatactt tcaacatcat tggtaatgat agaatgaaaa caaactacca tgataagctt 3360
gctgcaatag agaaagatag ggattcagct aggaaagact ggaaaaagat aaataacatc 3420

aaagagatga aagagggcta tctatctcag gtagttcatg aaatagctaa gctagttata 3480
gagtataatg ctattgtggt ttttgaggat ttaaattttg gatttaaaag agggcgtttc 3540
aaggtagaga agcaggtcta tcaaaagtta gaaaaaatgc taattgagaa actaaactat 3600
ctagttttca aagataatga gtttgataaa actgggggag tgcttagagc ttatcagcta 3660
acagcacctt ttgagacttt taaaaagatg ggtaaacaaa caggtattat ctactatgta 3720
ccagctggtt ttacttcaaa aatttgtcct gtaactggtt ttgtaaatca gttatatcct 3780
aagtatgaaa gtgtcagcaa atctcaagag ttctttagta agtttgacaa gatttgttat 3840

aaccttgata agggctattt tgagtttagt tttgattata aaaactttgg tgacaaggct 3900

gccaaaggca agtggactat agctagcttt gggagtagat tgattaactt tagaaattca 3960

gataaaaatc ataattggga tactcgagaa gtttatccaa ctaaagagtt ggagaaattg 4020

ctaaaagatt attctatcga atatgggcat ggcgaatgta tcaaagcagc tatttgcggt 4080

gagagcgaca aaaagttttt tgctaagcta actagtgtcc taaatactat cttacaaatg 4140

cgtaactcaa aaacaggtac tgagttagat tatctaattt caccagtagc agatgtaaat 4200

ggcaatttct ttgattcgcg acaggcgcca aaaaatatgc ctcaagatgc tgatgccaat 4260

ggtgcttatc atattgggct aaaaggtctg atgctactag gtaggatcaa aaataatcaa 4320

gagggcaaaa aactcaattt ggttatcaaa aatgaagagt attttgagtt cgtgcagaat 4380

aggaataacc aagcggccgc actcgagaaa aggccggcgg ccacgaaaaa ggccggccag 4440

gcaaaaaaga aaaagtcgac accaccacca ccaccactga gatccggctg ctaa 4494

<210> 77

<211> 2072

<212> PRT

<213> Artificial Sequence

<220><223> SSB-FnCpf1-RecJ

<400> 77

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30

Gly Ser Met Ala Ser Arg Gly Val Asn Lys Val Ile Leu Val Gly Asn
35 40 45

Leu Gly Gln Asp Pro Glu Val Arg Tyr Met Pro Asn Gly Gly Ala Val
50 55 60

Ala Asn Ile Thr Leu Ala Thr Ser Glu Ser Trp Arg Asp Lys Ala Thr
65 70 75 80

Gly Glu Met Lys Glu Gln Thr Glu Trp His Arg Val Val Leu Phe Gly

85 90 95

Lys Leu Ala Glu Val Ala Ser Glu Tyr Leu Arg Lys Gly Ser Gln Val
100 105 110

Tyr Ile Glu Gly Gln Leu Arg Thr Arg Lys Trp Thr Asp Gln Ser Gly

115 120 125

Gln Asp Arg Tyr Thr Thr Glu Val Val Val Asn Val Gly Gly Thr Met

130 135 140

Gln Met Leu Gly Gly Arg Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala

145 150 155 160

Cys Gly Arg Met Ser Ile Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu

165 170 175

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

180 185 190

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

195 200 205

Asp Tyr Lys Lys Ala Lys Gln Ile Ile Asp Lys Tyr His Gln Phe Phe

210 215 220

Ile Glu Glu Ile Leu Ser Ser Val Cys Ile Ser Glu Asp Leu Leu Gln

225 230 235 240

Asn Tyr Ser Asp Val Tyr Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn

245 250 255

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

260 265 270

Ser Glu Tyr Ile Lys Asp Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln

275 280 285

Asn Leu Ile Asp Ala Lys Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp

290 295 300

Leu Lys Gln Ser Lys Asp Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser

305 310 315 320

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

325 330 335

Gly Trp Thr Thr Tyr Phe Lys Gly Phe His Glu Asn Arg Lys Asn Val

340 345 350

Tyr Ser Ser Asn Asp Ile Pro Thr Ser Ile Ile Tyr Arg Ile Val Asp

355 360 365

Asp Asn Leu Pro Lys Phe Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu

370 375 380

Lys Asp Lys Ala Pro Glu Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp

385 390 395 400

Leu Ala Glu Glu Leu Thr Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val

405 410 415

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

420 425 430

Asn Asn Tyr Leu Asn Gln Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile

435 440 445

Gly Gly Lys Phe Val Asn Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn

450 455 460

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

465 470 475 480

Lys Tyr Lys Met Ser Val Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu

485 490 495

Ser Lys Ser Phe Val Ile Asp Lys Leu Glu Asp Asp Ser Asp Val Val

500 505 510

Thr Thr Met Gln Ser Phe Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val

515 520 525

Glu Glu Lys Ser Ile Lys Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu

530 535 540

Lys Ala Gln Lys Leu Asp Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys

545 550 555 560

Ser Leu Thr Asp Leu Ser Gln Gln Val Phe Asp Asp Tyr Ser Val Ile

565 570 575

Gly Thr Ala Val Leu Glu Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn

580 585 590

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

595 600 605

Glu Lys Ala Lys Tyr Leu Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu

610 615 620

Glu Phe Asn Lys His Arg Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu

625 630 635 640

Ile Leu Ala Asn Phe Ala Ala Ile Pro Met Ile Phe Asp Glu Ile Ala

645 650 655

Gln Asn Lys Asp Asn Leu Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln

660 665 670

Gly Lys Lys Asp Leu Leu Gln Ala Ser Ala Glu Asp Asp Val Lys Ala

675 680 685

Ile Lys Asp Leu Leu Asp Gln Thr Asn Asn Leu Leu His Lys Leu Lys

690 695 700

Ile Phe His Ile Ser Gln Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys

705 710 715 720

Asp Glu His Phe Tyr Leu Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala

725 730 735

Asn Ile Val Pro Leu Tyr Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys

740 745 750

Pro Tyr Ser Asp Glu Lys Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu

755 760 765

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

770 775 780

Phe Ile Lys Asp Asp Lys Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn

785 790 795 800

Asn Lys Ile Phe Asp Asp Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly

805 810 815

Tyr Lys Lys Ile Val Tyr Lys Leu Leu Pro Gly Ala Asn Lys Met Leu

820 825 830

Pro Lys Val Phe Phe Ser Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser

835 840 845

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

850 855 860

Ser Pro Gln Lys Gly Tyr Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys

865 870 875 880

Arg Lys Phe Ile Asp Phe Tyr Lys Gln Ser Ile Ser Lys His Pro Glu
885 890 895
Trp Lys Asp Phe Gly Phe Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser
900 905 910
Ile Asp Glu Phe Tyr Arg Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr
915 920 925
Phe Glu Asn Ile Ser Glu Ser Tyr Ile Asp Ser Val Val Asn Gln Gly
930 935 940
Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser

945 950 955 960
Lys Gly Arg Pro Asn Leu His Thr Leu Tyr Trp Lys Ala Leu Phe Asp
965 970 975
Glu Arg Asn Leu Gln Asp Val Val Tyr Lys Leu Asn Gly Glu Ala Glu
980 985 990
Leu Phe Tyr Arg Lys Gln Ser Ile Pro Lys Lys Ile Thr His Pro Ala
995 1000 1005
Lys Glu Ala Ile Ala Asn Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser
1010 1015 1020

Val Phe Glu Tyr Asp Leu Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys
1025 1030 1035 1040
Phe Phe Phe His Cys Pro Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala
1045 1050 1055
Asn Lys Phe Asn Asp Glu Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn
1060 1065 1070
Asp Val His Ile Leu Ser Ile Asp Arg Gly Glu Arg His Leu Ala Tyr
1075 1080 1085
Tyr Thr Leu Val Asp Gly Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe

1090 1095 1100
Asn Ile Ile Gly Asn Asp Arg Met Lys Thr Asn Tyr His Asp Lys Leu
1105 1110 1115 1120

Ala Ala Ile Glu Lys Asp Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys
1125 1130 1135
Ile Asn Asn Ile Lys Glu Met Lys Glu Gly Tyr Leu Ser Gln Val Val
1140 1145 1150
His Glu Ile Ala Lys Leu Val Ile Glu Tyr Asn Ala Ile Val Val Phe
1155 1160 1165

Glu Asp Leu Asn Phe Gly Phe Lys Arg Gly Arg Phe Lys Val Glu Lys
1170 1175 1180
Gln Val Tyr Gln Lys Leu Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr
1185 1190 1195 1200
Leu Val Phe Lys Asp Asn Glu Phe Asp Lys Thr Gly Gly Val Leu Arg
1205 1210 1215
Ala Tyr Gln Leu Thr Ala Pro Phe Glu Thr Phe Lys Lys Met Gly Lys
1220 1225 1230
Gln Thr Gly Ile Ile Tyr Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile

1235 1240 1245
Cys Pro Val Thr Gly Phe Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser
1250 1255 1260
Val Ser Lys Ser Gln Glu Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr
1265 1270 1275 1280
Asn Leu Asp Lys Gly Tyr Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe
1285 1290 1295
Gly Asp Lys Ala Ala Lys Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser
1300 1305 1310

Arg Leu Ile Asn Phe Arg Asn Ser Asp Lys Asn His Asn Trp Asp Thr
1315 1320 1325
Arg Glu Val Tyr Pro Thr Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr
1330 1335 1340
Ser Ile Glu Tyr Gly His Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly
1345 1350 1355 1360
Glu Ser Asp Lys Lys Phe Phe Ala Lys Leu Thr Ser Val Leu Asn Thr
1365 1370 1375

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

1380 1385 1390

Ile Ser Pro Val Ala Asp Val Asn Gly Asn Phe Phe Asp Ser Arg Gln

1395 1400 1405

Ala Pro Lys Asn Met Pro Gln Asp Ala Asp Ala Asn Gly Ala Tyr His

1410 1415 1420

Ile Gly Leu Lys Gly Leu Met Leu Leu Gly Arg Ile Lys Asn Asn Gln

1425 1430 1435 1440

Glu Gly Lys Lys Leu Asn Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu

1445 1450 1455

Phe Val Gln Asn Arg Asn Asn Gln Ala Ala Ala Leu Asp Leu Gln Val

1460 1465 1470

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

1475 1480 1485

Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser Arg

1490 1495 1500

Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met Leu

1505 1510 1515 1520

Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu Tyr

1525 1530 1535

Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe Asp

1540 1545 1550

Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg Ser

1555 1560 1565

Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg Phe Glu Asp

1570 1575 1580

Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg Gly

1585 1590 1595 1600

Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala Gly

1605 1610 1615

Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp His

1620 1625 1630

His Leu Pro Gly Asp Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn Pro

1635 1640 1645

Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val Gly

1650 1655 1660

Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp Gln

1665 1670 1675 1680

Gly Trp Phe Asp Glu Arg Asn Ile Ala Ile Pro Asn Leu Ala Glu Leu

1685 1690 1695

Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu Asp

1700 1705 1710

Ala Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg Ala

1715 1720 1725

Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn Arg

1730 1735 1740

Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly Pro

1745 1750 1755 1760

Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val Ala

1765 1770 1775

Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn Glu

1780 1785 1790

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

1795 1800 1805

Ile Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser Arg Asp Thr

1810 1815 1820

Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly Val

1825 1830 1835 1840

Val Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His Arg Pro Val

1845 1850 1855

Ile Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly Ser Gly Arg

1860 1865 1870

Ser Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg Leu Asp Thr

1875 1880 1885

Leu Tyr Pro Gly Met Met Leu Lys Phe Gly Gly His Ala Met Ala Ala
1890 1895 1900
Gly Leu Ser Leu Glu Glu Asp Lys Phe Lys Leu Phe Gln Gln Arg Phe
1905 1910 1915 1920
Gly Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu Gln Gly Glu
1925 1930 1935
Val Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr Met Glu Val
1940 1945 1950
Ala Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met Phe Pro Glu

1955 1960 1965
Pro Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg Leu Val Gly
1970 1975 1980
Glu Arg His Leu Lys Val Met Val Glu Pro Val Gly Gly Gly Pro Leu
1985 1990 1995 2000
Leu Asp Gly Ile Ala Phe Asn Val Asp Thr Ala Leu Trp Pro Asp Asn
2005 2010 2015
Gly Val Arg Glu Val Gln Leu Ala Tyr Lys Leu Asp Ile Asn Glu Phe
2020 2025 2030

Arg Gly Asn Arg Ser Leu Gln Ile Ile Ile Asp Asn Ile Trp Pro Ile
2035 2040 2045
Leu Gln Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys
2050 2055 2060
Lys Lys His His His His His His
2065 2070

<210> 78
<211> 6219
<212> DNA
<213> Artificial Sequence
<220><223> SSB-FnCpf1-RecJ
<400> 78
atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggccag cagaggcgta 120

aacaaggtta ttctcgttgg taatctgggt caggacccgg aagtacgcta catgccaaat 180
ggtggcgcag ttgccaacat tacgctggct acttccgaat cctggcgtga taaagcgacc 240
ggcgagatga aagaacagac tgaatggcac cgcgttgtgc tgttcggcaa actggcagaa 300
gtggcgagcg aatatctgcg taaaggttct caggtttata tcgaaggtca gctgcgtacc 360
cgtaaatgga ccgatcaatc cggtcaggat cgctacacca cagaagtcgt ggtgaacgtt 420
ggcggcacca tgcagatgct gggtggtcgt ggatccgaat tcgagctccg tcgacaagct 480
tgcggccgca tgtcaattta tcaagaattt gttaataaat atagtttaag taaaactcta 540

agatttgagt taatcccaca gggtaaaaca cttgaaaaca taaaagcaag aggtttgatt 600
ttagatgatg agaaaagagc taaagactac aaaaaggcta aacaaataat tgataaatat 660
catcagtttt ttatagagga gatattaagt tcggtttgta ttagcgaaga tttattacaa 720
aactattctg atgtttattt taaacttaaa aagagtgatg atgataatct acaaaaagat 780
tttaaaagtg caaaagatac gataaagaaa caaatatctg aatatataaa ggactcagag 840
aaatttaaga atttgtttaa tcaaaacctt atcgatgcta aaaaagggca agagtcagat 900
ttaattctat ggctaaagca atctaaggat aatggtatag aactatttaa agccaatagt 960

gatatcacag atatagatga ggcgttagaa ataatcaaat cttttaaagg ttggacaact 1020
tattttaagg gttttcatga aaatagaaaa aatgtttata gtagcaatga tattcctaca 1080
tctattattt ataggatagt agatgataat ttgcctaaat ttctagaaaa taaagctaag 1140
tatgagagtt taaaagacaa agctccagaa gctataaact atgaacaaat taaaaaagat 1200
ttggcagaag agctaacctt tgatattgac tacaaaacat ctgaagttaa tcaaagagtt 1260
ttttcacttg atgaagtttt tgagatagca aactttaata attatctaaa tcaaagtggt 1320
attactaaat ttaatactat tattggtggt aaatttgtaa atggtgaaaa tacaaagaga 1380

aaaggtataa atgaatatat aaatctatac tcacagcaaa taaatgataa aacactcaaa 1440
aaatataaaa tgagtgtttt atttaagcaa attttaagtg atacagaatc taaatctttt 1500
gtaattgata agttagaaga tgatagtgat gtagttacaa cgatgcaaag ttttttatgag 1560
caaatagcag cttttaaaac agtagaagaa aaatctatta aagaaacact atctttatta 1620
tttgatgatt taaaagctca aaaacttgat ttgagtaaaa tttattttaa aaatgataaa 1680
tctcttactg atctatcaca acaagttttt gatgattata gtgttattgg tacagcggta 1740
ctagaatata taactcaaca aatagcacct aaaaatcttg ataaccctag taagaaagag 1800

caagaattaa tagccaaaaa aactgaaaaa gcaaaatact tatctctaga aactataaag 1860

cttgccttag aagaatttaa taagcataga gatatagata aacagtgtag gtttgaagaa 1920

atacttgcaa actttgcggc tattccgatg atatttgatg aaatagctca aaacaaagac 1980

aatttggcac agatatctat caaatatcaa aatcaaggta aaaaagacct acttcaagct 2040

agtgcggaag atgatgttaa agctatcaag gatcttttag atcaaactaa taatctctta 2100

cataaactaa aaatatttca tattagtcag tcagaagata aggcaaatat tttagacaag 2160

gatgagcatt tttatctagt atttgaggag tgctactttg agctagcgaa tatagtgcct 2220

ctttataaca aaattagaaa ctatataact caaaagccat atagtgatga gaaatttaag 2280

ctcaattttg agaactcgac tttggctaat ggttgggata aaaataaaga gcctgacaat 2340

acggcaattt tatttatcaa agatgataaa tattatctgg gtgtgatgaa taagaaaaat 2400

aacaaaatat ttgatgataa agctatcaaa gaaaataaag gcgagggtta taaaaaaatt 2460

gtttataaac ttttacctgg cgcaaataaa atgttaccta aggttttctt ttctgctaaa 2520

tctataaaat tttataatcc tagtgaagat atacttagaa taagaaatca ttccacacat 2580

acaaaaaatg gtagtcctca aaaaggatat gaaaaatttg agtttaatat tgaagattgc 2640

cgaaaattta tagattttta taaacagtct ataagtaagc atccggagtg gaaagatttt 2700

ggatttagat tttctgatac tcaaagatat aattctatag atgaatttta tagagaagtt 2760

gaaaatcaag gctacaaact aacttttgaa aatatatcag agagctatat tgatagcgta 2820

gttaatcagg gtaaattgta cctattccaa atctataata aagatttttc agcttatagc 2880

aaagggcgac caaatctaca tactttatat tggaaagcgc tgtttgatga gagaaatctt 2940

caagatgtgg tttataagct aaatggtgag gcagagcttt tttatcgtaa acaatcaata 3000

cctaaaaaaa tcactcaccc agctaaagag gcaatagcta ataaaaacaa agataatcct 3060

aaaaaagaga gtgtttttga atatgattta atcaaagata aacgctttac tgaagataag 3120

ttttttctttc actgtcctat tacaatcaat tttaaatcta gtggagctaa taagtttaat 3180

gatgaaatca atttattgct aaaagaaaaa gcaaatgatg ttcatatatt aagtatagat 3240

agaggtgaaa gacatttagc ttactatact ttggtagatg gtaaaggcaa tatcatcaaa 3300

caagatactt tcaacatcat tggtaatgat agaatgaaaa caaactacca tgataagctt 3360

gctgcaatag agaaagatag ggattcagct aggaaagact ggaaaaagat aaataacatc 3420

aaagagatga aagagggcta tctatctcag gtagttcatg aaatagctaa gctagttata 3480

gagtataatg ctattgtggt ttttgaggat ttaaattttg gatttaaaag agggcgtttc 3540

aaggtagaga agcaggtcta tcaaaagtta gaaaaaatgc taattgagaa actaaactat 3600

ctagttttca aagataatga gtttgataaa actgggggag tgcttagagc ttatcagcta 3660

acagcacctt ttgagacttt taaaaagatg ggtaaacaaa caggtattat ctactatgta 3720

ccagctggtt ttacttcaaa aatttgtcct gtaactggtt ttgtaaatca gttatatcct 3780

aagtatgaaa gtgtcagcaa atctcaagag ttctttagta agtttgacaa gatttgttat 3840

aaccttgata agggctattt tgagtttagt tttgattata aaaactttgg tgacaaggct 3900

gccaaaggca agtggactat agctagcttt gggagtagat tgattaactt tagaaattca 3960

gataaaaatc ataattggga tactcgagaa gtttatccaa ctaaagagtt ggagaaattg 4020

ctaaaagatt attctatcga atatgggcat ggcgaatgta tcaaagcagc tatttgcggt 4080

gagagcgaca aaaagttttt tgctaagcta actagtgtcc taaatactat cttacaaatg 4140

cgtaactcaa aaacaggtac tgagttagat tatctaattt caccagtagc agatgtaaat 4200

ggcaatttct ttgattcgcg acaggcgcca aaaaatatgc ctcaagatgc tgatgccaat 4260

ggtgcttatc atattgggct aaaaggtctg atgctactag gtaggatcaa aaataatcaa 4320

gagggcaaaa aactcaattt ggttatcaaa aatgaagagt attttgagtt cgtgcagaat 4380

aggaataacc aagcggccgc actcgacctg caggtgaaac aacagataca acttcgtcgc 4440

cgtgaagtcg atgaaacggc agacttgccc gctgaattgc ctcccttgct gcgccgttta 4500

tacgccagcc ggggagtacg cagtgcgcaa gaactggaac gcagtgttaa aggtatgctg 4560

ccctggcagc aactgagcgg cgtcgaaaag gccgttgaga tcctttacaa cgcttttcgc 4620

gaaggaacgc ggattattgt ggtcggtgat ttcgacgccg acggcgcgac cagcacggct 4680

ctaagcgtgc tggcgatgcg ctcgcttggt tgcagcaata tcgactacct ggtaccaaac 4740

cgtttcgaag acggttacgg cttaagcccg gaagtggtcg atcaggccca tgcccgtggc 4800

gcgcagttaa ttgtcacggt ggataacggt atttcctccc atgcgggggt tgagcacgct 4860

cgctcgttgg gcatcccggt tattgttacc gatcaccatt tgccaggcga cacattaccc 4920

gcagcggaag cgatcattaa ccctaacttg cgcgactgta atttcccgtc gaaatcactg 4980

gcaggcgtgg gtgtggcgtt ttatctgatg ctggcgctgc gcacctttt gcgcgatcag 5040

ggctggtttg atgagcgtaa catcgcaatt cctaacctgg cagaactgct ggatctggtc 5100

gcgctgggga cagtggcgga cgtcgtgccg ctggacgcta ataatcgcat tctgacctgg 5160

caggggatga gtcgcatccg agccggaaag tgccgtccgg ggattaaagc gctgcttgaa 5220

gtggcaaacc gtgatgcaca aaaactcgcc gccagcgatt taggttttgc gctggggcca 5280

cgtctcaatg ctgccggacg actggacgat atgtccgtcg gtgtggcgct gttgttgtgc 5340

gacaacatcg gcgaagcgcg cgtgctggca aatgaactcg atgcgctaaa ccagacgcga 5400

aaagagatcg aacaaggaat gcaaattgaa gccctgaccc tgtgcgagaa actggagcgc 5460

agccgtgaca cgctacccgg cgggctggca atgtatcacc ccgaatggca tcagggcgtt 5520

gtcggtattc tggcttcgcg catcaaagag cgttttcacc gtccggttat cgcgtttgcg 5580

ccagcaggtg acggtacgct gaaaggttcc ggtcgctcca ttcaggggct gcatatgcgt 5640
gatgcgctgg agcgattaga cacactctac cctggcatga tgctgaagtt tggcggtcat 5700
gcgatggcgg cgggtttgtc gctggaagag gataaattca aactctttca acaacggttt 5760
ggcgaactgg ttactgagtg gctggaccct tcgctattgc aaggcgaagt ggtatcagac 5820
ggtccgttaa gcccggccga aatgaccatg gaagtggcgc agctgctgcg cgatgctggc 5880
ccgtgggggc agatgttccc ggagccgctg tttgacggtc atttccgtct gctgcaacag 5940
cggctggtgg gcgaacgtca tttgaaggtg atggtcgaac cggtcggcgg cggtccactg 6000

ctggatggta ttgcttttaa tgtcgatacc gccctctggc cggataacgg cgtgcgcgaa 6060
gtgcaactgg cttataagct cgatatcaac gagtttcgcg gcaaccgcag cctgcaaatt 6120
atcatcgaca atatctggcc aattctgcag aaaaggccgg cggccacgaa aaaggccggc 6180
caggcaaaaa agaaaaagca ccaccaccac caccactga 6219

<210> 79
<211> 1470
<212> PRT
<213> Artificial Sequence
<220><223> DSB-FnCpf1
<400> 79

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
20 25 30
Gly Ser Met Ala Lys Lys Glu Met Val Glu Phe Asp Glu Ala Ile His
35 40 45
Gly Glu Asp Leu Ala Lys Phe Ile Lys Glu Ala Ser Asp His Lys Leu
50 55 60
Lys Ile Ser Gly Tyr Asn Glu Leu Ile Lys Asp Ile Arg Ile Arg Ala
65 70 75 80
Lys Asp Glu Leu Gly Val Asp Gly Lys Met Phe Asn Arg Leu Leu Ala

85 90 95
Leu Tyr His Lys Asp Asn Arg Asp Val Phe Glu Ala Glu Thr Glu Glu

100 105 110

Val Val Glu Leu Tyr Asp Thr Val Phe Ser Lys Gly Ser Glu Phe Glu

115 120 125

Leu Arg Arg Gln Ala Cys Gly Arg Met Ser Ile Tyr Gln Glu Phe Val

130 135 140

Asn Lys Tyr Ser Leu Ser Lys Thr Leu Arg Phe Glu Leu Ile Pro Gln

145 150 155 160

Gly Lys Thr Leu Glu Asn Ile Lys Ala Arg Gly Leu Ile Leu Asp Asp

165 170 175

Glu Lys Arg Ala Lys Asp Tyr Lys Lys Ala Lys Gln Ile Ile Asp Lys

180 185 190

Tyr His Gln Phe Phe Ile Glu Glu Ile Leu Ser Ser Val Cys Ile Ser

195 200 205

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

210 215 220

Ser Asp Asp Asp Asn Leu Gln Lys Asp Phe Lys Ser Ala Lys Asp Thr

225 230 235 240

Ile Lys Lys Gln Ile Ser Glu Tyr Ile Lys Asp Ser Glu Lys Phe Lys

245 250 255

Asn Leu Phe Asn Gln Asn Leu Ile Asp Ala Lys Lys Gly Gln Glu Ser

260 265 270

Asp Leu Ile Leu Trp Leu Lys Gln Ser Lys Asp Asn Gly Ile Glu Leu

275 280 285

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

290 295 300

Ile Lys Ser Phe Lys Gly Trp Thr Thr Tyr Phe Lys Gly Phe His Glu

305 310 315 320

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

325 330 335

Tyr Arg Ile Val Asp Asp Asn Leu Pro Lys Phe Leu Glu Asn Lys Ala

340 345 350

Lys Tyr Glu Ser Leu Lys Asp Lys Ala Pro Glu Ala Ile Asn Tyr Glu

355 360 365

Gln Ile Lys Lys Asp Leu Ala Glu Glu Leu Thr Phe Asp Ile Asp Tyr

370 375 380

Lys Thr Ser Glu Val Asn Gln Arg Val Phe Ser Leu Asp Glu Val Phe

385 390 395 400

Glu Ile Ala Asn Phe Asn Asn Tyr Leu Asn Gln Ser Gly Ile Thr Lys

405 410 415

Phe Asn Thr Ile Ile Gly Gly Lys Phe Val Asn Gly Glu Asn Thr Lys

420 425 430

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

435 440 445

Asp Lys Thr Leu Lys Lys Tyr Lys Met Ser Val Leu Phe Lys Gln Ile

450 455 460

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

465 470 475 480

Asp Ser Asp Val Val Thr Thr Met Gln Ser Phe Tyr Glu Gln Ile Ala

485 490 495

Ala Phe Lys Thr Val Glu Glu Lys Ser Ile Lys Glu Thr Leu Ser Leu

500 505 510

Leu Phe Asp Asp Leu Lys Ala Gln Lys Leu Asp Leu Ser Lys Ile Tyr

515 520 525

Phe Lys Asn Asp Lys Ser Leu Thr Asp Leu Ser Gln Gln Val Phe Asp

530 535 540

Asp Tyr Ser Val Ile Gly Thr Ala Val Leu Glu Tyr Ile Thr Gln Gln

545 550 555 560

Ile Ala Pro Lys Asn Leu Asp Asn Pro Ser Lys Lys Glu Gln Glu Leu

565 570 575

Ile Ala Lys Lys Thr Glu Lys Ala Lys Tyr Leu Ser Leu Glu Thr Ile

580 585 590

Lys Leu Ala Leu Glu Glu Phe Asn Lys His Arg Asp Ile Asp Lys Gln

595 600 605

Cys Arg Phe Glu Glu Ile Leu Ala Asn Phe Ala Ala Ile Pro Met Ile
610 615 620
Phe Asp Glu Ile Ala Gln Asn Lys Asp Asn Leu Ala Gln Ile Ser Ile
625 630 635 640
Lys Tyr Gln Asn Gln Gly Lys Lys Asp Leu Leu Gln Ala Ser Ala Glu
645 650 655
Asp Asp Val Lys Ala Ile Lys Asp Leu Leu Asp Gln Thr Asn Asn Leu

660 665 670
Leu His Lys Leu Lys Ile Phe His Ile Ser Gln Ser Glu Asp Lys Ala
675 680 685
Asn Ile Leu Asp Lys Asp Glu His Phe Tyr Leu Val Phe Glu Glu Cys
690 695 700
Tyr Phe Glu Leu Ala Asn Ile Val Pro Leu Tyr Asn Lys Ile Arg Asn
705 710 715 720
Tyr Ile Thr Gln Lys Pro Tyr Ser Asp Glu Lys Phe Lys Leu Asn Phe
725 730 735

Glu Asn Ser Thr Leu Ala Asn Gly Trp Asp Lys Asn Lys Glu Pro Asp
740 745 750
Asn Thr Ala Ile Leu Phe Ile Lys Asp Asp Lys Tyr Tyr Leu Gly Val
755 760 765
Met Asn Lys Lys Asn Asn Lys Ile Phe Asp Asp Lys Ala Ile Lys Glu
770 775 780
Asn Lys Gly Glu Gly Tyr Lys Lys Ile Val Tyr Lys Leu Leu Pro Gly
785 790 795 800
Ala Asn Lys Met Leu Pro Lys Val Phe Phe Ser Ala Lys Ser Ile Lys

805 810 815
Phe Tyr Asn Pro Ser Glu Asp Ile Leu Arg Ile Arg Asn His Ser Thr
820 825 830
His Thr Lys Asn Gly Ser Pro Gln Lys Gly Tyr Glu Lys Phe Glu Phe
835 840 845
Asn Ile Glu Asp Cys Arg Lys Phe Ile Asp Phe Tyr Lys Gln Ser Ile
850 855 860

Ser Lys His Pro Glu Trp Lys Asp Phe Gly Phe Arg Phe Ser Asp Thr
865 870 875 880

Gln Arg Tyr Asn Ser Ile Asp Glu Phe Tyr Arg Glu Val Glu Asn Gln
885 890 895
Gly Tyr Lys Leu Thr Phe Glu Asn Ile Ser Glu Ser Tyr Ile Asp Ser
900 905 910
Val Val Asn Gln Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp
915 920 925
Phe Ser Ala Tyr Ser Lys Gly Arg Pro Asn Leu His Thr Leu Tyr Trp
930 935 940
Lys Ala Leu Phe Asp Glu Arg Asn Leu Gln Asp Val Val Tyr Lys Leu

945 950 955 960
Asn Gly Glu Ala Glu Leu Phe Tyr Arg Lys Gln Ser Ile Pro Lys Lys
965 970 975
Ile Thr His Pro Ala Lys Glu Ala Ile Ala Asn Lys Asn Lys Asp Asn
980 985 990
Pro Lys Lys Glu Ser Val Phe Glu Tyr Asp Leu Ile Lys Asp Lys Arg
995 1000 1005
Phe Thr Glu Asp Lys Phe Phe Phe His Cys Pro Ile Thr Ile Asn Phe
1010 1015 1020

Lys Ser Ser Gly Ala Asn Lys Phe Asn Asp Glu Ile Asn Leu Leu Leu
1025 1030 1035 1040
Lys Glu Lys Ala Asn Asp Val His Ile Leu Ser Ile Asp Arg Gly Glu
1045 1050 1055
Arg His Leu Ala Tyr Tyr Thr Leu Val Asp Gly Lys Gly Asn Ile Ile
1060 1065 1070
Lys Gln Asp Thr Phe Asn Ile Ile Gly Asn Asp Arg Met Lys Thr Asn
1075 1080 1085
Tyr His Asp Lys Leu Ala Ala Ile Glu Lys Asp Arg Asp Ser Ala Arg

1090 1095 1100
Lys Asp Trp Lys Lys Ile Asn Asn Ile Lys Glu Met Lys Glu Gly Tyr

1105 1110 1115 1120

Leu Ser Gln Val Val His Glu Ile Ala Lys Leu Val Ile Glu Tyr Asn

1125 1130 1135

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

1140 1145 1150

Phe Lys Val Glu Lys Gln Val Tyr Gln Lys Leu Glu Lys Met Leu Ile

1155 1160 1165

Glu Lys Leu Asn Tyr Leu Val Phe Lys Asp Asn Glu Phe Asp Lys Thr

1170 1175 1180

Gly Gly Val Leu Arg Ala Tyr Gln Leu Thr Ala Pro Phe Glu Thr Phe

1185 1190 1195 1200

Lys Lys Met Gly Lys Gln Thr Gly Ile Ile Tyr Tyr Val Pro Ala Gly

1205 1210 1215

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

1220 1225 1230

Pro Lys Tyr Glu Ser Val Ser Lys Ser Gln Glu Phe Phe Ser Lys Phe

1235 1240 1245

Asp Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr Phe Glu Phe Ser Phe

1250 1255 1260

Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys Gly Lys Trp Thr Ile

1265 1270 1275 1280

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

1285 1290 1295

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

1300 1305 1310

Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His Gly Glu Cys Ile Lys

1315 1320 1325

Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe Phe Ala Lys Leu Thr

1330 1335 1340

Ser Val Leu Asn Thr Ile Leu Gln Met Arg Asn Ser Lys Thr Gly Thr

1345 1350 1355 1360

Glu Leu Asp Tyr Leu Ile Ser Pro Val Ala Asp Val Asn Gly Asn Phe

1365 1370 1375

Phe Asp Ser Arg Gln Ala Pro Lys Asn Met Pro Gln Asp Ala Asp Ala

1380 1385 1390

Asn Gly Ala Tyr His Ile Gly Leu Lys Gly Leu Met Leu Leu Gly Arg

1395 1400 1405

Ile Lys Asn Asn Gln Glu Gly Lys Lys Leu Asn Leu Val Ile Lys Asn

1410 1415 1420

Glu Glu Tyr Phe Glu Phe Val Gln Asn Arg Asn Asn Gln Ala Ala Ala

1425 1430 1435 1440

Leu Glu Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys

1445 1450 1455

Lys Lys Ser Thr Pro Pro Pro Pro Pro Leu Arg Ser Gly Cys

1460 1465 1470

<210> 80

<211> 4413

<212> DNA

<213> Artificial Sequence

<220><223> DSB-FnCpf1

<400> 80

atgggcagca gccatcatca tcatcatcac agcagcggcc tggtgccgcg cggcagccat 60

atggctagca tgactggtgg acagcaaatg ggtcgcggat ccatggctaa aaaagaaatg 120

gttgaatttg atgaagctat ccatggcgaa gacttggcta aatttattaa agaagcatct 180

gatcataaac tgaaaatttc cggttataat gaactgatta aagatattcg aattcgtgct 240

aaagatgaac ttggcgttga tggtaagatg tttaatcgtc tattagcttt gtatcataaa 300

gataaccgtg atgtgtttga agctgaaact gaagaggtag ttgaacttta tgacacagtt 360

ttctctaaag gatccgaatt cgagctccgt cgacaagctt gcggccgcat gtcaatttat 420

caagaatttg ttaataaata tagtttaagt aaaactctaa gatttgagtt aatcccacag 480

ggtaaaacac ttgaaaacat aaaagcaaga ggtttgattt tagatgatga gaaaagagct 540

aaagactaca aaaaggctaa acaaataatt gataaatatc atcagttttt tatagaggag 600

atattaagtt cggtttgtat tagcgaagat ttattacaaa actattctga tgtttatttt 660

aaacttaaaa agagtgatga tgataatcta caaaaagatt ttaaaagtgc aaaagatacg 720

ataaagaaac aaatatctga atatataaag gactcagaga aatttaagaa tttgtttaat 780

caaaacctta tcgatgctaa aaaagggcaa gagtcagatt taattctatg gctaaagcaa 840

tctaaggata atggtataga actatttaaa gccaatagtg atatcacaga tatagatgag 900

gcgttagaaa taatcaaatc ttttaaaggt tggacaactt attttaaggg ttttcatgaa 960

aatagaaaaa atgtttatag tagcaatgat attcctacat ctattattta taggatagta 1020

gatgataatt tgcctaaatt tctagaaaat aaagctaagt atgagagttt aaaagacaaa 1080

gctccagaag ctataaacta tgaacaaatt aaaaaagatt tggcagaaga gctaaccttt 1140

gatattgact acaaaacatc tgaagttaat caaagagttt tttcacttga tgaagttttt 1200

gagatagcaa actttaataa ttatctaaat caaagtggta ttactaaatt taatactatt 1260

attggtggta aatttgtaaa tggtgaaaat acaaagagaa aaggtataaa tgaatatata 1320

aatctatact cacagcaaat aaatgataaa acactcaaaa aatataaaat gagtgtttta 1380

tttaagcaaa ttttaagtga tacagaatct aaatcttttg taattgataa gttagaagat 1440

gatagtgatg tagttacaac gatgcaaagt ttttatgagc aaatagcagc ttttaaaaca 1500

gtagaagaaa aatctattaa agaaacacta tctttattat ttgatgattt aaaagctcaa 1560

aaacttgatt tgagtaaaat ttatttttaaa aatgataaat ctcttactga tctatcacaa 1620

caagtttttg atgattatag tgttattggt acagcggtac tagaatatat aactcaacaa 1680

atagcaccta aaaatcttga taaccctagt aagaaagagc aagaattaat agccaaaaaa 1740

actgaaaaag caaaatactt atctctagaa actataaagc ttgccttaga agaatttaat 1800

aagcatagag atatagataa acagtgtagg tttgaagaaa tacttgcaaa ctttgcggct 1860

attccgatga tatttgatga aatagctcaa aacaaagaca atttggcaca gatatctatc 1920

aaatatcaaa atcaaggtaa aaaagaccta cttcaagcta gtgcggaaga tgatgttaaa 1980

gctatcaagg atcttttaga tcaaactaat aatctcttac ataaactaaa aatatttcat 2040

attagtcagt cagaagataa ggcaaatatt ttagacaagg atgagcattt ttatctagta 2100

tttgaggagt gctactttga gctagcgaat atagtgcctc tttataacaa aattagaaac 2160

tatataactc aaaagccata tagtgatgag aaatttaagc tcaattttga gaactcgact 2220

ttggctaatg gttgggataa aaataaagag cctgacaata cggcaatttt atttatcaaa 2280

gatgataaat attatctggg tgtgatgaat aagaaaaata acaaaatatt tgatgataaa 2340

gctatcaaag aaaataaagg cgagggttat aaaaaaattg tttataaact tttacctggc 2400

gcaaataaaa tgttacctaa ggttttcttt tctgctaaat ctataaaatt ttataatcct 2460

agtgaagata tacttagaat aagaaatcat tccacacata caaaaaatgg tagtcctcaa 2520

aaaggatatg aaaaatttga gtttaatatt gaagattgcc gaaaatttat agatttttat 2580

aaacagtcta taagtaagca tccggagtgg aaagattttg gatttagatt ttctgatact 2640

caaagatata attctataga tgaattttat agagaagttg aaaatcaagg ctacaaacta 2700

acttttgaaa atatatcaga gagctatatt gatagcgtag ttaatcaggg taaattgtac 2760

ctattccaaa tctataataa agatttttca gcttatagca aagggcgacc aaatctacat 2820

actttatatt ggaaagcgct gtttgatgag agaaatcttc aagatgtggt ttataagcta 2880

aatggtgagg cagagctttt ttatcgtaaa caatcaatac ctaaaaaaat cactcaccca 2940

gctaaagagg caatagctaa taaaaacaaa gataatccta aaaaagagag tgtttttgaa 3000

tatgatttaa tcaaagataa acgctttact gaagataagt ttttctttca ctgtcctatt 3060

acaatcaatt ttaaatctag tggagctaat aagtttaatg atgaaatcaa tttattgcta 3120

aaagaaaaag caaatgatgt tcatatatta agtatagata gaggtgaaag acatttagct 3180

tactatactt tggtagatgg taaaggcaat atcatcaaac aagatacttt caacatcatt 3240

ggtaatgata gaatgaaaac aaactaccat gataagcttg ctgcaataga gaaagatagg 3300

gattcagcta ggaaagactg gaaaaagata aataacatca aagagatgaa agagggctat 3360

ctatctcagg tagttcatga aatagctaag ctagttatag agtataatgc tattgtggtt 3420

tttgaggatt taaattttgg atttaaaaga gggcgtttca aggtagagaa gcaggtctat 3480

caaaagttag aaaaaatgct aattgagaaa ctaaactatc tagttttcaa agataatgag 3540

tttgataaaa ctgggggagt gcttagagct tatcagctaa cagcaccttt tgagactttt 3600

aaaaagatgg gtaaacaaac aggtattatc tactatgtac cagctggttt tacttcaaaa 3660

atttgtcctg taactggttt tgtaaatcag ttatatccta agtatgaaag tgtcagcaaa 3720

tctcaagagt tctttagtaa gtttgacaag atttgttata accttgataa gggctatttt 3780

gagtttagtt ttgattataa aaactttggt gacaaggctg ccaaaggcaa gtggactata 3840

gctagctttg ggagtagatt gattaacttt agaaattcag ataaaaatca taattgggat 3900

actcgagaag tttatccaac taaagagttg gagaaattgc taaaagatta ttctatcgaa 3960

tatgggcatg gcgaatgtat caaagcagct atttgcggtg agagcgacaa aaagtttttt 4020

gctaagctaa ctagtgtcct aaatactatc ttacaaatgc gtaactcaaa aacaggtact 4080

gagttagatt atctaatttc accagtagca gatgtaaatg gcaatttctt tgattcgcga 4140

caggcgccaa aaaatatgcc tcaagatgct gatgccaatg gtgcttatca tattgggcta 4200

aaaggtctga tgctactagg taggatcaaa aataatcaag agggcaaaaa actcaatttg 4260

gttatcaaaa atgaagagta ttttgagttc gtgcagaata ggaataacca agcggccgca 4320

ctcgagaaaa ggccggcggc cacgaaaaag gccggccagg caaaaaagaa aaagtcgaca 4380

ccaccaccac caccactgag atccggctgc taa 4413

<210> 81

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> Adapter primer sequences of CCR5

<400> 81

tcgtcggcag cgtcagatgt gtataagaga cagggagggc aactaaatac attct 55

<210> 82

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> Adapter primer sequences of CCR5

<400> 82

gtctcgtggg ctcggagatg tgtataagag acaggaacac cagtgagtag agcgg 55

<210> 83

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223

> Adapter primer sequences of DHCR7

<400> 83

tcgtcggcag cgtcagatgt gtataagaga cagggtttga gcaacagttc tcc 53

<210> 84

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> Adapter primer sequences of DHCR7

<400> 84

gtctcgtggg ctcggagatg tgtataagag acagtgacga tgtccaccac ag 52

<210> 85

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> Adapter primer sequences of CCR5

<400> 85

tcgtcggcag cgtcagatgt gtataagaga cagggtattt ctgttcagat cac 53

<210> 86

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> Adapter primer sequences of CCR5

<400> 86

gtctcgtggg ctcggagatg tgtataagag acaggcccat caattataga aagcc 55

<210> 87

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> Adapter primer sequences of DNMT1

<400> 87

tcgtcggcag cgtcagatgt gtataagaga cagctgcaca cagcaggcct ttg 53

<210> 88

<211> 54

<212> DNA

<213> Artificial Sequence

<220><

223> Adapter primer sequences of DNMT1

<400> 88

gtctcgtggg ctcggagatg tgtataagag acagcccaat aagtggcaga gtgc 54

<210> 89

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Spacer sequence of CCR5

<400> 89

tgacatcaat tattatacat 20

<210> 90

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Spacer sequence of ADCY5

<400> 90

tgacatcaat tattatagat 20

<210> 91

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Spacer sequence of KCNJ6

<400> 91

tgacatcact tattatgcat 20

<210> 92

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Spacer sequence of CNTNAP2

<400> 92

tgacataaat tattctacat 20

<210> 93

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Spacer sequence of Chr. 5N/A

<400> 93

tgaaatcaat tatcatagat 20

<210> 94

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for CCR5

<400> 94

tggaacaaga tggattatca agtgtcaagt ccaatctatg acatcaatta ttat 54

<210> 95

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for CCR5

<400> 95

acatatgcat cggagccctg ccaaaaaatc aatgtgaagc aaatcgcagc cc 52

<210> 96

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for DHCR7

<400> 96

atgggcccca gtgtgactgc ctgcatccgt cctcgcaggg aggtggactg gt 52

<210> 97

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for DHCR7

<400> 97

tttgaattcc actggcgagc gtcatcttcc tactgctgtt cgccccccttc atcg 54

<210> 98

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223

> ssODN donor sequences for CCR5

<400> 98

aattctctga ggctttcttt taaatataca taaggaactt tcggagtgaa ggg 53

<210> 99

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for CCR5

<400> 99

agagtttcat atggtcaata acttgatgca tgtgaagggg agataaaaag gtt 53

<210> 100

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for DNMT1

<400> 100

tggccctggg gccgtttccc tcactcctgc tcggtgaatt tggctcagca ggc 53

<210> 101

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ssODN donor sequences for DNMT1

<400> 101

acctgccgaa ttctcagctg ctcacttgag cctctgggtc tagaaccctc tgg 53

<210> 102

<211> 3358

<212> DNA

<213> Artificial Sequence

<220><223> pGEM _447

<400> 102

gggcgaattg ggcccgacgt cgcatgctcc cggccgccat ggccgcgggt ttgtcgctgg 60

aagaggataa attcgaactc tttcaacaac ggtttggcga gctggttacc gagtggctgg 120

acccttcgct attgcaaggc gaagtggtgt cagacggccc gttaagcccg gccgaaatga 180

ccatggaagt ggcgcagctg ctgcgcgatg ctggcccgtg ggggcagatg ttcccggagc 240

cgctgtttga tggtcatttc cgtctgctgc aacagcggct ggtgggcgaa cgtcatttga 300

aagtcatggt cgaaccggtc ggcggcggtc cgctgctgga tggtattgct tttaatgtcg 360

ataccgccct ctggccggat aacggcgtgc gcgaagtgca actggcttaa tcactagtgc 420

ggccgcctgc aggtcgacca tatgggagag ctcccaacgc gttggatgca tagcttgagt 480

attctatagt gtcacctaaa tagcttggcg taatcatggt catagctgtt tcctgtgtga 540

aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc 600

tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc 660

cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc 720

ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 780

cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 840

ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 900

aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 960

cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 1020

cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 1080

gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 1140

tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccccgt tcagcccgac 1200

cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 1260

ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 1320

gagttcttga agtggtggcc taactacggc tacactagaa gaacagtatt tggtatctgc 1380

gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 1440

accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 1500

ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 1560

tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 1620

aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 1680

taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 1740

gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 1800

agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac 1860

cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 1920

tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac 1980

gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc 2040

agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg 2100

gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc 2160

atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct 2220

gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc 2280

tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc 2340

atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc 2400

agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc 2460

gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca 2520

cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt 2580

tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt 2640

ccgcgcacat ttccccgaaa agtgccacct gatgcggtgt gaaataccgc acagatgcgt 2700

aaggagaaaa taccgcatca ggaaattgta agcgttaata ttttgttaaa attcgcgtta 2760

aatttttgtt aaatcagctc atttttttaac caataggccg aaatcggcaa aatcccttat 2820

aaatcaaaag aatagaccga gatagggttg agtgttgttc cagtttggaa caagagtcca 2880

ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 2940

ccactacgtg aaccatcacc ctaatcaagt tttttggggt cgaggtgccg taaagcacta 3000

aatcggaacc ctaaagggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg 3060

gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg 3120

gtcacgctgc gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc 3180

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 3240

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 3300

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gtaatacgac tcactata 3358