



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0123492
(43) 공개일자 2023년08월23일

(51) 국제특허분류(Int. Cl.)
C12N 15/62 (2006.01) C12N 15/10 (2017.01)
C12N 15/113 (2010.01) C12N 15/90 (2006.01)
C12N 9/12 (2006.01) C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/62 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2023-7023944
(22) 출원일자(국제) 2021년12월16일
심사청구일자 없음
(85) 번역문제출일자 2023년07월13일
(86) 국제출원번호 PCT/EP2021/086348
(87) 국제공개번호 WO 2022/129438
국제공개일자 2022년06월23일
(30) 우선권주장
20214696.5 2020년12월16일
유럽특허청(EPO)(EP)
21209719.0 2021년11월22일
유럽특허청(EPO)(EP)

(71) 출원인
우니베르시타트 폼페우 파브라
스페인 08002 바르셀로나 12 카레르 데 라 메르세
(72) 발명자
구엘 카르골 마르크
스페인, 08002 바르셀로나, 플라자 데 라 메르세
10-12, 씨/오우니베르시타트 폼페우 파브라
산체즈-메히아스 가르시아 아벤치아
스페인, 08002 바르셀로나, 플라자 데 라 메르세
10-12, 씨/오우니베르시타트 폼페우 파브라
(뒷면에 계속)
(74) 대리인
특허법인한얼

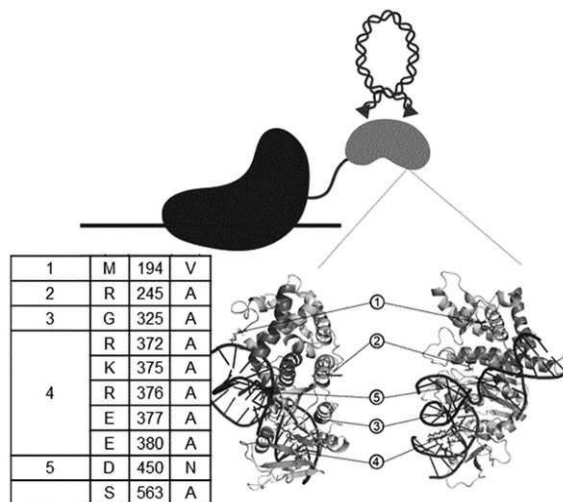
전체 청구항 수 : 총 22 항

(54) 발명의 명칭 프로그래밍 가능한 트랜스포사제 및 이의 용도

(57) 요약

(i) 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질; 또는 상기 제1 단백질을 코딩하는 핵산 작제물; 및 (ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물을 포함하며; 여기서 상기 트랜스포사제가 변형된(modified) 과활성(hyperactive) PiggyBac인, 조성물을 기반으로 하는 효율적이고 정밀한 프로그래밍 가능한 유전자 전달 기술을 제공한다.

대표도 - 도1



(52) CPC특허분류

C12N 15/113 (2013.01)

C12N 15/907 (2013.01)

C12N 9/1241 (2013.01)

C12N 9/22 (2013.01)

C07K 2319/80 (2013.01)

C12N 2310/20 (2017.05)

C12N 2800/90 (2013.01)

(72) 발명자

팔라레스 마스미타 마리아

스페인, 08002 바르셀로나, 플라자 데 라 메르세
10-12, 씨/오우니베르시타트 폼페우 파브라

이반치크 데르마노빅 디미트리

스페인, 08002 바르셀로나, 플라자 데 라 메르세
10-12, 씨/오우니베르시타트 폼페우 파브라

라흐메 아말

스페인, 08002 바르셀로나, 플라자 데 라 메르세
10-12, 씨/오우니베르시타트 폼페우 파브라

명세서

청구범위

청구항 1

(i) 표적 핵산 서열(target nucleic acid sequence)에 결합하고 이를 절단할 수 있는 부위-특이적(site-specific) DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질; 또는 상기 제1 단백질을 코딩하는(encoding) 핵산 작제물(construct); 및

(ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물

을 포함하며; 여기서 상기 트랜스포사제가 서열번호: 9의 과활성(hyperactive) PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는, 변형된 과활성 PiggyBac인, 조성물.

청구항 2

제1항에 있어서, 제1 및 제2 단백질이 임의로 링커(linker)를 통해 함께 융합(fusing)되어 융합 단백질을 형성하는, 조성물.

청구항 3

제2항에 있어서, 제1 단백질이, 임의로 링커를 통해, 제2 단백질의 C 말단 단부(terminal end)에 융합되는, 조성물.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 트랜스포사제가 변형되지 않은 과활성 PiggyBac과 비교하여 절제 활성(excision activity)을 증가시키기 위한 하나 이상의 아미노산 돌연변이, 및/또는 변형되지 않은 과활성 PiggyBac과 비교하여 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인 조성물.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 하나 이상의 아미노산 돌연변이가 R372A, K375A, 및 D450N으로 이루어지지 않는, 조성물.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 하나 이상의 아미노산 돌연변이가 M194, D450, T560, S564 S573, S592 또는 F594의 위치에서 절제 활성을 증가시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환 M194V 및/또는 D450N 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응하는, 조성물.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 상기 하나 이상의 아미노산 돌연변이가 M194 또는 D450의 위치에서 절제 활성을 증가시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호가 바람직하게는 아미노산 치환 M194V 및/또는 D450N 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응하는, 조성물.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 하나 이상의 아미노산 돌연변이가 위치 R275, R277, R347, R372, K375, R376, E377, 및/또는 E380에서 DNA 결합 활성을 감소시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환 R275A, R277, R347S, R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응하는, 조성물.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 하나 이상의 아미노산 돌연변이가 위치 R372, K375, R376, E377, 및/또는 E380에서 DNA 결합 활성을 감소시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환 R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응하는, 조성물.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 변형된 과활성 PiggyBac이 이중 돌연변이 N347S 및 D450N을 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응하는, 조성물.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 변형된 과활성 PiggyBac 돌연변이가 하기 아미노산 치환 또는 아미노산 치환의 조합 중 하나를 포함하는, 조성물: R372A/K375A/R376A/D450N, K375A/R376A/E377A/E380A/D450N, R372A/K375A/R376A/E377A/E380A/D450N, M194V, R376A, E377A, E380A, M194V/R372A/K375A, S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, T43I/R372A/K375A/A411T/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G; 여기서, 위치 번호는 서열번호: 9의 과활성 PiggyBac의 아미노산 번호에 상응하며, 전형적으로 상기 변형된 트랜스포사제는 서열번호: 2-8, 10-18 및 135-149 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 제2 트랜스포사제를 포함하거나 이로 이루어진 제3 단백질; 또는 상기 제3 단백질을 코딩하는 핵산 작제물을 포함하며; 여기서 상기 제2 트랜스포사제는 서열번호: 9의 과활성 PiggyBac, 또는 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인, 조성물.

청구항 13

제12항에 있어서, 제1, 제2 및 제3 단백질이 임의로 링커를 통해, 함께 융합되어 삼중 융합 단백질을 형성하는, 조성물.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 제1 단백질이 RNA-가이드된(guided) 뉴클레아제 또는 니카제(nickase), 또는 징크 핑거(zinc finger) 뉴클레아제를 포함하거나 이로 이루어진, 조성물.

청구항 15

제1항 내지 제14항 중 어느 한 항에 있어서, 상기 제1 단백질이 활성 DNA 절단(cleavage) 도메인 및 가이드 RNA 결합 도메인을 포함하며 서열번호: 31의 스트렙토코커스 피로게네스(*Streptococcus pyogenes*) Cas9 (SpCas9), 서열번호: 72의 스태필로코커스 아우레우스(*Staphylococcus aureus*) Cas9 (SaCas9), 서열번호: 74의 Cpf1, 서열번호: 29의 캄필로박터 제주니(*Campylobacter jejuni*) Cas9 (CjCas9), 서열번호: 70의 스트렙토코커스 피로게네스 Cas9 니카제 (nCas9), 서열번호: 75의 CasX, 또는 서열번호: 76의 스태필로코커스 아우레우스 Cas9 니카제에 대해 적어도 80%, 90%, 95%, 99% 또는 적어도 100% 동일성(identity)을 갖는 뉴클레아제 단백질이며; 바람직하게는 상기 제1 단백질은 서열번호: 72의 스태필로코커스 아우레우스 Cas9 (SaCas9) 및 서열번호: 31의 스트렙토코커스 피로게네스 Cas9 (SpCas9)로 이루어진 군으로부터 선택된 Cas9 단백질인, 조성물.

청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서, 가이드 RNA, 및 게놈에 삽입하기 위한 외인성 핵산을 추가로 포함하는, 조성물.

청구항 17

제16항에 있어서, 트랜스포사제가 가이드 RNA에 포함된 적어도 하나의 특이적 RNA 서열에 결합할 수 있는 RNA-결합 단백질에 융합되고;

임의로 여기서 상기 RNA-결합 단백질은 MS2 박테리오파지 코트 단백질(bacteriophage coat protein) (MCP)이고, 여기서 가이드 RNA는 바람직하게는 서열번호: 153과 적어도 75% 동일성을 공유하는 MS2 RNA 테트라루프(tetraloop) 결합 서열을 포함하는, 조성물.

청구항 18

제16항 또는 제17항에 있어서, 상기 외인성 핵산이, 전형적으로 크기가 5 kb 내지 25 kb, 그리고 보다 바람직하게는 8 kb 내지 20 kb인, 큰 DNA 단편인, 조성물.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, 조성물이 나노입자에 포함되는, 조성물.

청구항 20

전형적으로 메신저(messenger) RNA (mRNA)인, 제2항 내지 제17항 중 어느 한 항에 정의된 바와 같은 융합 단백질을 코딩하는 핵산.

청구항 21

제1항 내지 제19항 중 어느 한 항에 따른 조성물, 가이드 RNA, 및 외인성 핵산을 세포에 전달하는 단계를 포함하는, 세포 게놈으로의 외인성 핵산 서열의 부위 특이적 통합(site specific integration)을 위한 시험관내 방법.

청구항 22

세포 게놈으로의 외인성 핵산 서열의 부위-특이적 통합에 의한, 질환의 치료에 사용하기 위한, 제1항 내지 제19항 중 어느 한 항에 따른 조성물, 가이드 RNA, 및 외인성 핵산.

발명의 설명

기술 분야

[0001] 본 발명은 유전자 편집(gene editing) 및 유전자 요법(gene therapy)의 분야에 관한 것이다.

배경 기술

[0002] 암, 발달 장애(developmental disorder), 및 일부 감염과 같은 많은 질환은 공통적으로 유전적 및 후생적 이상(aberration)을 가지고 있다. 유전자 요법은 유전적으로 기능장애가 있는 세포(dysfunctional cell)를 교정하여 연관된 질환을 치료하기 위해 직접 게놈을 표적으로 하고 편집하기 위해 세포에 유전 물질(genetic material)을 도입하도록 설계되어 있다.

[0003] 유전자 편집 도구상자(toolbox)는 지난 몇 년에 걸쳐 상당히 확장되어 이를 필요한 대상체에서 장애를 치료하기 위해 결합이 있는 유전자를 복구(repair)하는 유전자 요법에 더하여 유망한 도구를 가지고 있다.

[0004] 전통적으로, 유전자 편집은 게놈에서 관심 서열에 이중-가닥 파단(double-strand break) (DSB)을 유도하는 인공 엔도뉴클레아제의 설계를 기반으로 한다¹. 세포는 비-상동 말단-연결(Non-Homologous End-Joining) (NHEJ) 또는 상동성 지향 복구(Homology Directed Repair) (HDR)²의 두 가지 주요 경로 중 하나를 통해 DSB를 복구한다. 최근에, DSB에 독립적인 편집이 개발되었다. 데아미나제로 DNA 염기를 직접 편집하는 것을 기반으로 하는 방법

론, 즉 염기 편집자(base editor) (BE)³; 및 역전사 효소(reverse transcriptase) (RT), 즉 프라임 편집자(prime editor) (PE)⁴의 도움으로 DNA 염기를 계내(in situ) 대체하는 것이 이용가능해졌다.

[0005] 그러나, 병리학적 유전적 결함(pathological genetic defect)은 소수의 염기에서 대규모 결실의 범위까지 다양할 수 있다. 염기 편집자 또는 프라임 편집자는 소수의 염기만을 표적으로 하며, HDR-기반 편집은 크기로 제대로 크기 조정(scale)되지 않는다⁵. 상동성 비의존성 표적화된 통합(Homology Independent Targeted Integration) (HITI)과 같은 NHEJ를 기반으로 하는 방법론이 개발되었다⁶. 이 방법론은 수 킬로베이스(several kilobases)의 삽입에 대해 입증되었으나 매우 대규모 편집에는 여전히 비효율적이다⁵. HITI는 엑손을 전달하기 위해 작동할 수 있지만, 디스트로핀(Dystrophin) (~14 kb) 또는 라미닌(Laminin)-α2 (~9 kb)와 같은 유전자의 cDNA를 강력하게 전달하기에는 충분히 효율적이지 않을 수 있다.

[0006] 고정밀 CRISPR 프로그래밍 가능한(programmable) 트랜스포존은 박테리아에서 설명되었으나^{7,8}, 포유동물 세포에서는 이용가능하지 않다. 징크 핑거(zinc finger) 또는 스트렙토코커스 피로게네스(*Streptococcus pyogenes*) Cas9 (SpCas9)를 포유동물 호환성 piggyBac (PB) 또는 슬리핑 뷰티(sleeping beauty) 트랜스포사제에 융합하려는 이전의 시도는 상대적으로 낮은 수준의 정밀도로 시스템을 전달하였다⁹⁻¹¹. PB 시스템은 효율이 크기로 잘 크기 조정되기 때문에¹², 유전자 요법을 위한 매력적인 도구이며, 돌연변이 독립적인 기술이며, DNA 복구 메커니즘에 대한 의존도가 낮기 때문에 임의의 조직에서 작동한다.

[0007] 따라서, 시험관내, 생체외 또는 생체외에서 포유동물 세포에서 표적화된 유전자 전달을 위한 새로운 시스템을 개발할 필요가 여전히 존재한다.

발명의 내용

[0008] 특정 프로그래밍 가능한 트랜스포사제 및 표적화된 유전자 편집(targeted gene editing)에서의 이의 용도는 WO2020250181에 개시되었으며, 그 내용은 본원에 참조로 포함된다.

[0009] 본 개시내용은 이제 (i) 표적 핵산 서열(target nucleic acid sequence)에 결합하고 이를 절단할 수 있는 부위-특이적(site-specific) DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질; 또는 상기 제1 단백질을 코딩하는(encoding) 핵산 작제물(construct); 및 (ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물을 포함하며; 여기서 상기 트랜스포사제가 변형된(modified) 과활성(hyperactive) PiggyBac인, 조성물을 기반으로 하는 추가의 효율적이고 정밀한 프로그래밍 가능한 유전자 전달 기술을 제공한다. 이러한 기술은 작은 핵산 단편뿐만 아니라 큰 핵산 단편을 전달하는 능력을 가지고 있다. 본 발명자들은 포유동물 세포 및 생체내 마우스 간에서 기술을 시험하였고 놀랍게도 이들 모두에서 부위 지향 통합(site directed integration)의 고효율 (5-10%)을 달성하였다.

[0010] 한 실시양태에서, 조성물은 (i) 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질; 또는 상기 제1 단백질을 코딩하는 핵산 작제물; 및 (ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물을 포함하며; 여기서 상기 트랜스포사제가 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는, 변형된 과활성 PiggyBac이다.

[0011] 한 실시양태에서, 제1 단백질 및 제2 단백질은, 임의로 링커(linker)를 통해, 함께 융합(fusing)되어 융합 단백질을 형성한다. 한 실시양태에서, 제1 단백질은, 임의로 링커를 통해, 제2 단백질의 C 말단 단부(terminal end)에 융합된다.

[0012] 한 실시양태에서, 상기 트랜스포사제는 변형되지 않은 과활성 PiggyBac과 비교하여 절제 활성(excision activity)을 증가시키기 위한 하나 이상의 아미노산 돌연변이, 및/또는 변형되지 않은 과활성 PiggyBac과 비교하여 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac이다.

[0013] 한 실시양태에서, 상기 하나 이상의 아미노산 돌연변이는 R372A, K375A, 및 D450N으로 이루어지지 않는다. 한 실시양태에서, 상기 하나 이상의 아미노산 돌연변이는 M194, D450, T560, S564 S573, S592 또는 F594의 위치에서 절제 활성을 증가시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환

M194V 및/또는 D450N 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다. 한 실시양태에서, 상기 하나 이상의 아미노산 돌연변이가 M194 또는 D450의 위치에서 절제 활성을 증가시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환 M194V 및/또는 D450N 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다. 한 실시양태에서, 상기 하나 이상의 아미노산 돌연변이는 위치 R275, R277, R347, R372, K375, R376, E377, 및/또는 E380에서 DNA 결합 활성을 감소시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환 R275A, R277, R347S, R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다. 한 실시양태에서, 상기 하나 이상의 아미노산 돌연변이는 위치 R372, K375, R376, E377, 및/또는 E380에서 DNA 결합 활성을 감소시키는 아미노산 치환 중에서 선택되며, 상기 위치 번호는 바람직하게는 아미노산 치환 R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택된, 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다.

[0014] 한 실시양태에서, 변형된 과활성 PiggyBac은 이중 돌연변이 N347S 및 D450N을 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다. 한 실시양태에서, 변형된 과활성 PiggyBac 돌연변이는 하기 아미노산 치환 또는 아미노산 치환의 조합 중 하나를 포함한다: R372A/K375A/R376A/D450N, K375A/R376A/E377A/E380A/D450N, R372A/K375A/R376A/E377A/E380A/D450N, M194V, R376A, E377A, E380A, M194V/R372A/K375A, S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, T43I/R372A/K375A/A411T/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G; 위치 번호는 서열번호: 9의 과활성 PiggyBac의 아미노산 번호에 상응하며, 전형적으로 상기 변형된 트랜스포사제는 서열번호: 2-8, 10-18 및 135-149 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.

[0015] 한 실시양태에서, 조성물은 제2 트랜스포사제를 포함하거나 이로 이루어진 제3 단백질; 또는 상기 제3 단백질을 코딩하는 핵산 작제물을 추가로 포함하며; 여기서 상기 제2 트랜스포사제는 서열번호: 9를 가진 과활성 PiggyBac, 또는 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac이다. 한 실시양태에서, 제1, 제2 및 제3 단백질은, 임의로 링커를 통해, 함께 융합되어 삼중 융합 단백질을 형성한다.

[0016] 한 실시양태에서, 제1 단백질은 RNA-가이드된(guided) 뉴클레아제 또는 니카제(nickase), 또는 징크 핑거 뉴클레아제를 포함하거나 이로 이루어진다. 한 실시양태에서, 상기 제1 단백질은 활성 DNA 절단(cleavage) 도메인(domain) 및 가이드 RNA 결합 도메인을 포함하며 서열번호: 31의 스트렙토코커스 피로게네스 Cas9 (SpCas9), 서열번호: 72의 스태필로코커스 아우레우스(*Staphylococcus aureus*) Cas9 (SaCas9), 서열번호: 74의 Cpf1, 서열번호: 29의 캄필로박터 제주니(*Campylobacter jejuni*) Cas9 (CjCas9), 서열번호: 70의 스트렙토코커스 피로게네스 Cas9 니카제 (nCas9), 서열번호: 75의 CasX, 또는 서열번호: 76의 스태필로코커스 아우레우스 Cas9 니카제에 대해 적어도 80%, 90%, 95%, 99% 또는 적어도 100% 동일성(identity)을 갖는 뉴클레아제 단백질이며; 바람직하게는 상기 제1 단백질은 서열번호: 72의 스태필로코커스 아우레우스 Cas9 (SaCas9) 및 서열번호: 31의 스트렙토코커스 피로게네스 Cas9 (SpCas9)로 이루어진 군으로부터 선택된 Cas9 단백질이다.

[0017] 한 실시양태에서, 조성물은 가이드 RNA, 및 게놈에 삽입하기 위한 외인성 핵산을 추가로 포함한다.

[0018] 한 실시양태에서, 트랜스포사제는 가이드 RNA에 포함된 적어도 하나의 특이적 RNA 서열에 결합할 수 있는 RNA 결합 단백질에 융합되고; 임의로 여기서 상기 RNA 결합 단백질은 MS2 박테리오파지 코트 단백질(bacteriophage coat protein) (MCP)이고 여기서 가이드 RNA는 바람직하게는 서열번호: 153과 적어도 75% 동일성을 공유하는 MS2 RNA 테트라루프(tetraloop) 결합 서열을 포함한다.

[0019] 한 실시양태에서, 외인성 핵산은, 전형적으로 크기가 5 kb 내지 25 kb, 그리고 보다 바람직하게는 8 kb 내지 20 kb인, 큰 DNA 단편이다.

[0020] 한 실시양태에서, 조성물은 나노입자에 포함된다.

- [0021] 본 발명은 또한 전형적으로 메신저(messenger) RNA (mRNA)의 형태로 본원에 개시된 융합 단백질 중 어느 하나를 코딩하는 핵산에 관한 것이다.
- [0022] 본 발명은 또한 본 발명의 조성물, 가이드 RNA 및 외인성 핵산을 세포에 전달하는 것을 포함하는, 세포 계놈으로의 외인성 핵산 서열의 부위-특이적 통합을 위한 시험관내 방법에 관한 것이다.
- [0023] 본 발명은 또한 세포 계놈으로의 외인성 핵산 서열의 부위-특이적 통합에 의한, 질환 치료에 사용하기 위한, 본 발명의 조성물, 가이드 RNA, 및 외인성 핵산에 관한 것이다.

도면의 간단한 설명

- [0024] **도 1.** 프로그래밍 가능한 트랜스포사제 기술: cas9 (적색)는 조작된 PB 트랜스포사제 도메인 (분홍색)과 조합된다. PB의 코어 모델(core model)에서 이의 상응하는 위치 (위치 563은 모델에 포함되지 않은 C-t에 있음)를 가진 실험에 사용된 돌연변이체의 표.
- 도 2.** cas9의 변이체에 대한 프로그래밍 가능한 트랜스포사제 의존성. 뉴클레아제 cas9 및 PB 융합은 데드(dead) cas9 (dcas9) 또는 니카제 cas9 (ncas9) 융합과 대조적으로 표적 삽입(targeted insertion) 및 전반적인 삽입에서 더 양호한 결과를 제시한다. 청색은 표적 삽입 및 황색 오프-타겟(off-target) 삽입을 나타냈다. B, PB의 변이체의 프로그래밍 가능한 트랜스포사제 의존성. DNA 결합이 감소된 절제 항상 돌연변이는 최상의 온-타겟(on-target):오프-타겟 비 (주황색)를 나타낸다. 온-타겟 삽입은 AAV 부위 (녹색), 및 TRAC 부위 (청색)에서 수행되었다. C. 상이한 링커의 시험. 링커 길이 및 토폴로지(topology)는 Spcas9 및 PB 융합의 온-타겟 활성에 유의하게 영향을 미치지 않는다.
- 도 3.** 허쉬(Hershey) 리포터 세포주: HEK293T 세포주는 하나의 스플라이싱 수용체 및 gRNA 표적 부위가 선행하는 GFP의 C-말단 단편을 함유하도록 조작되었다. PB 트랜스포존은 CAG 프로모터, GFP의 N-말단 단편에 이어서 스플라이싱 공여자를 결합하여 생성되었다. 회색 삼각형에서, PB ITR; SA: 스플라이싱 수용체; SD: 스플라이싱 공여자; 표적: 표적 삽입 부위; * 삽입 과정은 ITR을 분열시킴(disrupted).
- 도 4.** A PB 변이체의 프로그래밍 가능한 트랜스포사제 의존성. DNA 결합을 감소시키기 위해 상이한 돌연변이의 맥락에서 절단 항상 돌연변이체 (450)은 최상의 표적을 제시한다. R372와 R376의 A로의 동시 돌연변이는 잘 용인되지 않는다. E377은 DNA 결합에 관여하지 않긴 하지만, A로의 돌연변이는 K375 및 R376에서 A로의 돌연변이 이시 그 영역에서 음전하 축적(build-up)을 피하는데 유익할 수 있다. B R372A/K375A는 표적 DNA에 대한 결합 감소의 결과로서 PB의 통합 활성을 감소시킨다 (D450N에서도 관찰된 바와 같음). 진행 중인 오프-타겟 통합 시험.
- 도 5.** 표적 삽입에서 이중 가닥 파단(break) 및 프로그래밍 가능한 DNA 결합 도메인 효과. 온-타겟된 삽입을 위해 삽입 부위에서 이중 가닥 파단 및 PB의 공동-국부화(Co-localization)가 필요하다.
- 도 6.** 다중 삽입의 생어(Sanger) 시퀀싱(sequencing) 검증 (NGS에 의해 측정된 보다 포괄적인 분포는 도 2a 참조). ITR TTAA는 표적 삽입의 과정에서 손실된다. NGG Pam은 적색으로 강조 표시된다.
- 도 7.** cas9가 없는 삽입 활성 PB K375A_R376A_E377A_E380A_D450N. 표적 삽입 메커니즘을 추가로 조사하기 위해, hyPB K375A_R376A_E377A_E380A_D450N을 cas9 없이 클로닝하고 hek293T 세포에서 RFP 트랜스포존을 사용하여 hyPB WT와 비교하여 이의 삽입 효율을 시험하였다. 결과는 융합된 cas9 없이 이 돌연변이의 어떤 삽입 활성도 나타내지 않는다.
- 도 8.** 프로그래밍 가능한 트랜스포사제는 다수의 indel에 의해 ITR 부위를 불활성화하여 비가역적 삽입을 생성한다. B 가이드-seq를 사용한 전반적인 삽입 부위의 특성화. TCR 유전자좌 (상단 패널)에서 온-타겟만 검출되었다. 생어 시퀀싱은 4개의 클론에 대해 표시된다 (하단 패널).
- 도 9.** 가이드-seq에 의한 삽입 프로파일링의 프로그래밍 가능한 트랜스포사제 특성화는 Cas9와 조합된 hyPB 돌연변이가 정밀 트랜스포존 삽입을 수행했음을 나타낸다.
- 도 10.** Cas9 유도 HDR (300 bp 상동성 아암(homology arm)이 사용됨)과 같은 다른 표적 삽입 플랫폼에 대한 Cas9-hyPB R372A-K375A-D450N의 벤치마킹(Benchmarking).
- 도 11.** 마우스 간에서 Cas9-hyPB R372A-K375A-D450N의 생체내 전개(*in vivo* deployment). qPCR에 의해 측정된 상대 카피 수(Relative copy number)를 보고한다.

도 12. 프로그래밍 가능한 트랜스포사제를 상이한 Cas 변이체, 예컨대 CasX, CjCas9 Cpf1 또는 SaCas9로 조작할 수 있으며, 이들 중 일부는 SpCas9와 같이 표적 부위에서 프로그래밍 가능한 삽입의 면에서 유사한 결과를 달성하였다. 시험된 Cas 변이체 각각은 3개의 독립적인 gRNA를 가진 분할(split) GFP 리포터 세포주의 특이적 표적 영역을 표적으로 하였다.

도 13. Cas9 및 단일 gRNA (gRNA-TCR1 또는 AAVS1-3)에 의해 또는 Nickase Cas9 및 인접 위치 (gRNA-TCR1 및 AAVS1-3)에서 표적화하는 2개의 gRNA, 및 변형된 hyPB (돌연변이체 R372A-K375A-D405N)에 대한 융합으로 프로그래밍 가능한 DNA 결합 도메인 (ZnF)에 의한 이중 가닥 파단은 표적 삽입을 발생시킨다. 효율적인 온-타깃된 삽입을 위해 삽입 부위에서 이중 가닥 파단 및 PB의 공동-국부화가 필요하다. 이것은 뉴클레아제 Cas9 또는 Nickase Cas9에 의한 이중 컷(double cut)에 의해 달성될 수 있다.

도 14. 프로그래밍 가능한 트랜스포사제는 2개의 hyPB 도메인과 Cas9 뉴클레아제의 이량체 폴리펩티드로 조작되어, Cas9-hyPB와 비교하여 더 양호한 프로그래밍 가능한 삽입을 발생시킬 수 있다. 분할 GFP 리포터 세포주는 분할 GFP 트랜스포존의 표적 부위로의 프로그래밍 가능한 삽입에 사용하였다. hyPB R372A-K375A-D450N의 돌연변이체는 Cas9에 대한 단량체 또는 이량체 융합에 사용되었다. 조건: 1-삽입 기구(insertion machinery)로서 hyPB만을 사용하는 음성 대조군; 2: pcDNA 발현 벡터에서 Cas9-hyPB R372A-K375A-D450N의 양성 대조군; 3: 렌티바이러스 발현 벡터에서 Cas9-hyPB R372A-K375A-D450N의 양성 대조군; 4: C-말단에서 hyPB R372A-K375A-D450N의 2개의 단위에 융합된 Cas9 뉴클레아제; 5: 하나는 C-말단에 있고 다른 하나는 N-말단에 있는 hyPB R372A-K375A-D450N의 2개의 단위에 융합된 Cas9 뉴클레아제.

도 15. 라이브러리로부터 최상의 돌연변이체 조합을 선택할 수 있도록 프로그래밍 가능한 전위가 발생한 세포 선택의 수개의 주기. 우리는 Cas9에 융합될 때 Cas9-hyPB R372A-K375A-D450N보다 더 양호한 농축화(enrichment) 및 프로그래밍 가능한 삽입 용량을 가진 수개의 돌연변이체를 확인하였다.

도 16. 선택 주기에 따라 온-타깃 효율이 증가한다. 각각의 주기로부터 선택된 벌크(Bulk) 변이체는 AAVS1 및 1/2 GFP 트랜스포존을 표적으로 하는 gRNA로 리포터 세포주에 공동-형질감염되었다. 플라스미드의 수량은 클로닝 효율을 위해 정규화하기 위해 PB 카피 수에 의해 교정되었다.

도 17. (A) 상위로 선택된 후보의 온-타깃 효율. 마지막 주기로부터 선택된 96개의 무작위 클론 중 가장 높은 온-타깃 활성을 기준으로 하여 6개의 개별 후보가 선택되었다. 개별 온-타깃 활성을 Cas9-hyPB R372A-K375A-D450N과 비교하였다. (B) 상위 온-타깃 활성 변이체에서 우세한 PB 잔기를 나타내는 로고.

도 18. 상동성-비의존성 표적화된 삽입 (HITI)에 대한 Cas9-hyPB R372A-K375A-D450N (FiCAT)의 벤치마킹.

도 19. 네 가지 상이한 뉴클레아제 단백질을 사용한 FiCAT R372A-K375A-D450N의 프로그래밍 가능한 삽입 활성. SpCas9는 gRNA-TRAC-1 단독으로 프로그래밍 가능한 삽입을 위한 대조군으로서 사용된다 (좌측). 각각의 뉴클레아제는 1/2 GFP 리포터 세포주에 표적 삽입을 위해 3개의 독립적인 gRNA (1-3)와 함께 사용되었다.

도 20. 미니썬클 루시퍼라제 트랜스포존의 간 통합. 미니썬클 루시퍼라제 트랜스포존, Rosa26 유전자좌를 표적으로 하는 sgRNA 및 FiCAT (Cas9-hyPB R372A-K375A-D450N) mRNA를 유체역학적 주입에 의해 전달하고 루시퍼라제 신호를 모니터링하였다.

도 21. (a) CasX (좌측) 및 Cpf1 (중간)에 의한 편집 활동. (b) SaCas9 (좌측), CjCas9 (중간)에 의한 편집 활동. indels +/- SD를 가진 판독의 평균 %는 2개의 기술적 반복, N=3 생물학적 복제의 대표 이미지에 대해 표시된다. TRAC-1 부위를 표적으로 하는 SpCas9는 참조용으로 사용되었다 (우측).

도 22. 선택 주기에 따른 온-타깃 효율의 증가. (A) 각각의 주기로부터 선택된 벌크 변이체는 AAVS1 및 1/2 GFP 트랜스포존을 표적으로 하는 gRNA로 리포터 세포주에 공동-형질감염되었다. 플라스미드의 수량은 클로닝 효율을 위해 정규화하기 위해 PB 카피 수에 의해 교정되었다. (B) 각각의 주기의 벌크 변이체를 발현하는 렌티바이러스를 생산하여 리포터 세포주를 감염시키는 데 사용하였다.

도 23. gRNA tcr1 및 1/2 GFP MC 트랜스포존으로 공동 형질감염된 cas9_PB 라이브러리 농축화의 4주기 및 5주기 후 벌크 변이체로부터 단리된 단일 돌연변이체의 FiCAT (hyPB R372A-K375A-D450N)에 대한 특이적 표적 통합.

도 24. 1/2 GFP 리포터 세포주에서 표적 삽입을 위한 SpCas9 또는 SaCas9와 융합된 이량체 hyPB R372A-K375A-D450N의 프로그래밍 가능한 삽입 활성.

도 25. 1/2 GFP 리포터 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 상대적 비교. (A) SpCas9

단백질과 융합된 hyPB R372A-K375A-D450N (좌측) 및 SpCas9가 별도로 첨가된 MCP 단백질과 융합된 hyPB R372A-K375A-D450N (우측) 사이의 비교. (B) SpCas9가 별도로 첨가된 MCP 단백질에 융합된 3개의 hyPB 돌연변이체 (R372A-K375A-D450N; R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L)간의 비교.

도 26. ½ GFP 리포터 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 비교. (A) hyPB R372A-K375A-D450N 및 SpCas9 단백질 (좌측)과 HyPB R372A-K375A-D450N과 SpCas9 단백질을 포함하는 융합 단백질 (우측)의 공동-발현 사이의 비교. (B) SpCas9와 공동-발현된 3개의 hyPB 돌연변이체 (R372A-K375A-D450N; R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L) 간의 상대적 비교.

도 27. SpCas 및 hyPB R372A-K375A-D450N을 포함하는 제1 융합 단백질, 및 MCP 단백질 및 hyPB 돌연변이체 (R372A-K375A-D450N; R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L)를 포함하는 제2 융합 단백질의 공동-발현과 함께 ½ GFP 리포터 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 상대적 비교.

도 28. SpCas 및 hyPB R372A-K375A-D450N, 및 3개의 hyPB 돌연변이체 R372A-K375A-D450N; R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L을 포함하는 융합 단백질의 공동-발현과 함께 ½ GFP 리포터 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 상대적 비교; R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L.

도 29. hyPB R272A-K275A-D450N의 이량체에 융합된 SpCas9 (좌측)와 제1 hyPB R272A-K275A-D450N에 그리고 제2 돌연변이체에 융합된 SpCas9 (우측) 사이에 ½ GFP 리포터 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 비교.

발명을 실시하기 위한 구체적인 내용

정의

본원에 사용된 바와 같이, 단수형 "a", "an", 및 "the"는 문맥상 분명히 달리 나타내지 않는 한 단수형 및 복수형 언급대상을 포함한다. 따라서, 예를 들어 "제제(agent)"에 대한 언급은 단일 제제 및 복수의 이러한 제제를 포함한다.

용어 "**핵산 서열**" 및 "**뉴클레오타이드 서열**"은 단량체 뉴클레오타이드로 이루어지거나 그를 포함하는 임의의 분자를 지칭하기 위해 상호교환적으로 사용될 수 있다. 핵산은 올리고뉴클레오타이드 또는 폴리뉴클레오타이드일 수 있다. 뉴클레오타이드 서열은 DNA, RNA, 또는 이의 혼합일 수 있다. 뉴클레오타이드 서열은 화학적으로 변형되거나 인공적일 수 있다. 뉴클레오타이드 서열은 펩티드 핵산 (PNA), 모르폴리노 및 잠금 핵산 (LNA), 뿐만 아니라 글리콜 핵산 (GNA) 및 트레오스 핵산 (TNA)을 포함한다. 이들 서열 각각은 분자의 백본에 대한 변화에 의해 자연적으로 발생하는 DNA 또는 RNA와 구별된다. 또한, 포스포로티오에이트 뉴클레오타이드가 사용될 수 있다. 다른 데옥시뉴클레오타이드 유사체는, 제한 없이, 메틸포스포네이트, 포스포르아미데이트, 포스포로티오에이트, N3'P5'-포스포르아미데이트 및 올리고리보뉴클레오타이드 포스포로티오에이트 및 이의 2'-O-알릴 유사체 및 2'-O-메틸리보뉴클레오타이드 메틸포스포네이트를 포함하며, 이들은 본 개시내용의 뉴클레오타이드에 사용될 수 있다.

용어 "**트랜스진**"은 외인성 핵산 서열, 특히 유전자 산물을 코딩하는 외인성 DNA 또는 cDNA를 지칭한다. 유전자 산물은 RNA, 펩티드 또는 단백질일 수 있다. 유전자 산물 (CDS)에 대한 코딩 영역(coding regio)에 더하여, 트랜스진은 발현을 용이하게 하거나 향상시키기 위한 하나 이상의 작동 서열, 예컨대 프로모터, 인핸서(들), 반응 요소(response element)(들), 리포터 요소(들), 절연체 요소(들), 폴리아데닐화 신호(들) 및/또는 기타 기능 요소를 포함하거나 이와 연관될 수 있다. 본 개시내용의 실시양태는 달리 특정되지 않는 한, 임의의 공지된 적합한 프로모터, 인핸서(들), 반응 요소(들), 리포터 요소(들), 절연체 요소(들), 폴리아데닐화 신호(들) 및/또는 다른 기능적 요소를 이용할 수 있다. 적합한 요소 및 서열은 관련 기술분야의 통상의 기술자에게 널리 공지될 것이다.

용어 "**폴리펩티드**", "**펩티드**" 및 "**단백질**"은 아미노산 잔기의 중합체를 지칭하기 위해 상호교환적으로 사용된다. 상기 용어는 또한 하나 이상의 아미노산이 상응하는 자연-발생 아미노산의 화학적 유사체이거나 변형된 유도체인 아미노산 중합체에 또한 적용된다.

용어 "**결합 단백질**"은 또 다른 분자에 비공유결합으로 결합할 수 있는 단백질을 지칭한다. 결합 단백질은, 예를 들어, DNA 분자 (DNA-결합 단백질), RNA 분자 (RNA-결합 단백질) 및/또는 단백질 분자 (단백질-결합 단백질)

질)에 결합할 수 있다. 단백질-결합 단백질의 경우에, 동일한 단백질의 하나 이상의 분자에 결합하여 동종이량체, 동종삼량체 등을 형성할 수 있고/거나; 상이한 단백질 또는 단백질들의 하나 이상의 분자에 결합할 수 있다. 결합 단백질은 하나 초과 유형의 결합 활성을 가질 수 있다. 예를 들어, 징크 핑거 단백질은 DNA-결합, RNA-결합 및 단백질-결합 활성을 갖는다.

[0031] 용어 "Cas9" 또는 "Cas9 뉴클레아제"는 Cas9 단백질, 또는 이의 단편 (예를 들어, Cas9의 활성 또는 불활성 DNA 절단 도메인, 및/또는 Cas9의 gRNA 결합 도메인을 포함하는 단백질)을 포함하는 RNA-가이드된 뉴클레아제를 지칭한다. Cas9 뉴클레아제는 또한 때때로 casn1 뉴클레아제 또는 CRISPR (클러스터된 규칙적으로 간격을 둔 짧은 회문 반복부(clustered regularly interspaced short palindromic repeat))-연관 뉴클레아제라고도 지칭된다. CRISPR은 이동성 유전 요소 (바이러스, 전이 요소(transposable element) 및 접합 플라스미드(conjugative plasmid))에 대한 보호를 제공하는 적응 면역 시스템이다. CRISPR 클러스터는 스페이서, 선행 이동 요소에 상보적인 서열, 및 표적 침습 핵산을 함유한다. CRISPR 클러스터는 전사되어 CRISPR RNA (crRNA)로 프로세스된다. 유형 II CRISPR 시스템에서, pre-crRNA의 올바른 처리는 트랜스-코딩된 소형 RNA (tracrRNA), 내인성 리보뉴클레아제 3 (rnc) 및 Cas9 단백질을 필요로 한다. tracrRNA는 pre-crRNA의 리보뉴클레아제 3-보조 처리(aided processing)를 위한 가이드 역할을 한다. 후속적으로, Cas9/crRNA/tracrRNA는 스페이서에 상보적인 선행 또는 원형 dsDNA 표적을 내핵분해적으로(endonucleolytically) 절단한다. crRNA에 상보적이지 않은 표적 가닥은 먼저 내핵분해적으로 커팅된 다음에, 외핵분해적으로(exonucleolytically) 3'-5' 트리밍된다. 자연에서, DNA-결합 및 절단은 전형적으로 단백질 및 두 RNA 모두를 필요로 한다. 그러나, 단일 가이드 RNA ("sgRNA" 또는 간단히 "gRNA")는 crRNA 및 tracrRNA 둘 다의 측면을 단일 RNA 종으로 혼입되도록 조작될 수 있다.

[0032] Cas9는 CRISPR 반복 서열에서 짧은 모티프(motif) (PAM 또는 프로토스페이스(protospacer) 인접 모티프)를 인식하여 자기와 비-자기(non-self)를 구별하는 데 도움을 준다. Cas9 뉴클레아제 서열 및 구조는 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. Cas9 오르토로그(ortholog)는 에스. 피오케네스 및 에스. 썬모필러스(*S. thermophilus*)를 포함하나 이에 제한되지는 않는 다양한 종에서 기재되어 있다. 추가적인 적합한 Cas9 뉴클레아제 및 서열은 본 개시내용을 기반으로 하여 관련 기술분야의 통상의 기술자에게 명백할 것이며, 이러한 Cas9 뉴클레아제 및 서열은 문헌 [Chylinski *et al.*, 2013. (*RNA Biol.* **10**(5):726-37)]에 개시된 유기체 및 유전자로부터의 Cas9 서열을 포함하며, 상기 문헌의 전체 내용은 본원에 참조로 포함된다.

[0033] 일부 실시양태에서, Cas9 뉴클레아제는 불활성 (예를 들어, 불활성화된) DNA 절단 도메인을 갖는다. 뉴클레아제-불활성화된 Cas9 단백질은 "dCas9" 단백질 (뉴클레아제-"데드" Cas9의 경우)로서 상호교환적으로 지칭될 수 있다. 불활성 DNA 절단 도메인을 갖는 Cas9 단백질 (또는 이의 단편)을 생성하는 방법은 관련 기술분야에 공지되어 있다 (예를 들어, 문헌 [Jinek *et al.*, 2012. *Science*. **337**(6096):816-821; Qi *et al.*, 2013. *Cell*. **152**(5):1173-83]을 참조하며, 각각의 전체 내용은 본원에 참조로 포함됨).

[0034] 용어 "징크 핑거 단백질"은 아연 이온의 배위를 통해 구조가 안정화되는 징크 핑거 단백질의 결합 도메인 내의 아미노산 서열의 영역인 하나 이상의 징크 핑거를 통해 서열-특이적 방식으로 DNA에 결합하는, 단백질, 또는 더 큰 단백질 내의 도메인을 지칭한다. 용어 "징크 핑거 단백질"은 종종 "ZFP"로 약칭된다.

[0035] 용어 "징크 핑거 뉴클레아제"는 징크 핑거 DNA-결합 도메인을 DNA-절단 도메인에 융합시킴으로써 생성된 인공 제한 효소를 지칭한다. 징크 핑거 도메인은 원하는 특정 DNA 서열을 표적으로 하도록 조작될 수 있으며, 이를 통해 징크 핑거 뉴클레아제가 복잡한 게놈 내 고유한 서열을 표적으로 삼을 수 있다. "징크 핑거 뉴클레아제"는 종종 "ZFN" 또는 "ZNP"로 약칭된다.

[0036] 본원에 사용된 바와 같은 용어 "아미노산 서열" 또는 "폴리펩티드" 또는 "단백질"은 아미노산 잔기의 중합체를 지칭한다. 특정되지 않는 한, 아미노산 잔기의 중합체는 임의의 길이일 수 있다.

[0037] 본원에 사용된 바와 같이, 용어 "외인성"은 세포에 자연적으로 존재하지 않으나, 하나 이상의 유전적, 생화학적 또는 다른 방법에 의해 세포 내로 도입될 수 있는 분자를 지칭한다. 세포 내 자연적 존재는 또한 세포의 특정한 발달 단계 및 환경 조건과 관련하여 결정될 수 있다. 따라서, 예를 들어, 근육의 배아 발달 동안에만 존재하는 분자는 성체 근육 세포에 대해 외인성 분자이다. 유사하게, 열충격에 의해 유도된 분자는 열충격을 받지 않은 세포에 대해 외인성 분자이다. 외인성 분자는 예를 들어 오작동 내인성 분자의 기능 버전(functioning version) 또는 정상적으로 기능하는 내인성 분자의 오작동 버전(malfunctioning version)을 포함할 수 있다.

[0038] 대조적으로, "내인성" 분자는 특정한 환경 조건 하에 특정한 발달 단계에서 특정한 세포에 통상 존재하는 분자

이다. 예를 들어, 내인성 핵산은 염색체, 미토콘드리아의 게놈, 엽록체 또는 다른 소기관, 또는 자연 발생 에피솜(episomal) 핵산을 포함할 수 있다. 추가적인 내인성 분자는 단백질, 예를 들어 전사 인자 및 효소를 포함할 수 있다.

- [0039] 본원에 사용된 바와 같이, "**표적 서열**" 또는 "**표적 핵산 서열**" 또는 "**표적 부위**"는 예를 들어 게놈에서 결합 분자가 결합할 핵산의 일부를 정의하는 서열이되, 단, 결합에 대한 충분한 조건이 존재한다. 예를 들어, 서열 5'-GAATTC-3'은 EcoRI 제한 엔도뉴클레아제의 표적 부위이다.
- [0040] 용어 "**융합**"은 2개 이상의 서브유닛 분자가 연결된 분자를 지칭한다. 일부 실시양태에서, 둘 사이의 연결은 공유적이며; 대안적으로, 둘 사이의 연결은 비공유적일 수 있고, 예를 들어 분자간 상호작용에 의존할 수 있다. 서브유닛 분자는 동일한 화학적 유형의 분자이거나, 상이한 화학적 유형의 분자일 수 있다.
- [0041] 용어 "**융합 단백질**"은 적어도 2개의 상이한 단백질로부터의 단백질 도메인을 포함하는 하이브리드 폴리펩티드를 지칭한다. 예를 들어, 하나의 단백질 도메인은 융합 단백질의 아미노-말단 (N-말단) 부분 또는 카르복시 말단 (C-말단) 단백질에 위치하여, "아미노-말단 융합 단백질" 또는 "카르복시-말단 융합 단백질"을 각각 형성할 수 있다. 바람직한 실시양태에서, 융합 단백질은 핵산 서열에 의해 완전히 코딩될 수 있는 단일쇄 폴리펩티드이고, 펩티드에 의해 직접 공유결합되거나 펩티드 링커를 통해 임의로 공유결합된 적어도 2개의 단백질 도메인을 포함한다.
- [0042] 본원에 사용된 바와 같은 용어 "유전자" 또는 "게놈"은 유전자 산물을 코딩하는 DNA 영역뿐만 아니라 유전자 산물의 생성을 조절하는 모든 DNA 영역을 이러한 조절 서열이 코딩 및/또는 전사된 서열에 인접 여부와 관계 없이 포함한다. 따라서, 유전자는 프로모터 서열, 종결자(terminator), 번역 조절 서열 예컨대 리보솜 결합 부위 및 내부 리보솜 진입 부위, 인핸서(enhancer), 사일런서(silencer), 절연체, 경계 요소, 복제 기점, 매트릭스 부착 부위(matrix attachment site) 및 유전자와 제어 영역을 포함하지만 반드시 이에 제한되지는 않는다.
- [0043] 용어 "진핵 세포"는 진균 세포 (예컨대 효모), 식물 세포, 동물 세포, 포유동물 세포 및 인간 세포 (예를 들어, T-세포)를 포함하나 이에 제한되지는 않는다.
- [0044] 본원에 사용된 바와 같이, 용어 "연결된"은 두 개 이상의 구성요소 (예를 들어, 서열 요소)의 병렬 배치를 지칭하며, 여기서 구성요소는 두 구성요소가 모두 정상적으로 기능하고 구성요소 중 적어도 하나가 다른 구성요소 중 적어도 하나에 발휘되는 기능을 매재할 수 있는 가능성을 허용하도록 배열된다.
- [0045] 단백질, 폴리펩티드 또는 핵산의 "기능적 단편"은 각각 단백질, 폴리펩티드 또는 핵산이며, 이의 서열은 전장 단백질, 폴리펩티드 또는 핵산과 동일하지 않지만, 전장 단백질, 폴리펩티드 또는 핵산과 동일한 기능을 유지한다. 기능적 단편은 상응하는 천연 분자보다 많거나, 적거나, 동일한 수의 잔기를 가질 수 있고/있거나, 하나 이상의 아미노산 또는 뉴클레오타이드 치환을 함유할 수 있다.
- [0046] 본원에 사용된 바와 같이, 용어 "형질감염"은 핵산 (DNA 또는 RNA)을 진핵 또는 원핵 세포 또는 유기체에 도입하는 것을 지칭한다.
- [0047] 용어 "**절단(cleavage)**"은 DNA 분자의 공유 백본의 파손(breakage)을 지칭한다. 절단은 포스포디에스테르 결합의 효소적 또는 화학적 가수분해를 포함하나 이에 제한되지는 않는 다양한 방법에 의해 개시될 수 있다. 단일-가닥 절단 및 이중-가닥 절단 둘 다 가능하고, 이중-가닥 절단은 2개의 별개의 단일-가닥 절단 사건의 결과로서 발생할 수 있다. DNA 절단은 무딘 말단(blunt end) 또는 엇갈린 말단(staggered end)을 생성할 수 있다. 특정 실시양태에서, 융합 폴리펩티드는 표적화된 이중-가닥 DNA 절단에 사용된다.
- [0048] 용어 "**특이성**"은 선택된 서열에 대해 어느 정도의 서열 동일성을 공유하는 서열에 선택 적으로 결합하는 능력을 지칭한다.
- [0049] 용어 "**삽입**" 및 "**통합**"은 핵산 서열을 제2 핵산 서열에 또는 게놈 또는 이의 일부에 첨가하는 것을 지칭한다. 삽입 또는 통합과 관련하여 용어 "**특이적**", "**부위-특이적**", "**표적화된**" 및 "**온-타겟된**"은 본원에서 핵산의 삽입을 제2 핵산의 특정 부위에 또는 게놈 또는 이의 일부의 특정 부위에 삽입하는 것을 지칭하기 위해 상호교환적으로 사용된다. 반대로, 용어 "**무작위**", "**비표적화된(non-targeted)**" 및 "**오프-타겟된(off-targeted)**"은 원하지 않는 부위로의 핵산의 비특이적 및 의도하지 않은 삽입을 지칭한다. 용어 "**총**" 또는 "**전반적인**"은 총 삽입수를 지칭한다.
- [0050] 용어 "**돌연변이**"는 서열, 예를 들어 핵산 또는 아미노산 서열 내의 잔기의, 또 다른 잔기로의 치환; 및/또는 핵산 또는 아미노산 서열 내의 하나 이상의 잔기의 결실 또는 삽입을 지칭한다. 돌연변이는 전형적으로 원래의

잔기에 이어서 서열 내 잔기의 위치를 확인한 다음에, 새로 치환된 잔기의 정체를 확인함으로써 본원에 기재된 다. 본원에서 제공되는 아미노산 치환 (돌연변이)을 만드는 다양한 방법은 관련 기술분야에 널리 공지 있으며, 예를 들어 문헌 [Green & Sambrook, 2012 (*Molecular cloning: a laboratory manual* (4th Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.)]에 의해 제공된다. 바람직한 실시양태에서, 용어 단백질에서의 돌연변이는 아미노산 치환을 지칭한다.

[0051] 용어 "**트랜스포사제**"는 트랜스포존의 말단에 결합하고 잘라내기(cut)- 및 붙여넣기(paste) 메커니즘 또는 복제 전치 메커니즘(replicative transposition mechanism)에 의해 게놈의 또 다른 부분으로의 이의 이동을 촉매하는 효소를 지칭한다.

[0052] 용어 "**변형된**"은 상응하는 변형되지 않은 단백질 또는 핵산 서열과 상이한 단백질 또는 핵산 서열을 지칭한다.

[0053] 용어 "**링커**"는 2개의 인접한 분자 또는 모이어터를 연결하는 화학적 기 또는 분자를 지칭한다.

[0054] 본원에 사용된 바와 같은 용어 "벡터" 및 "플라스미드"는 예를 들어 제2 관심 폴리뉴클레오티드를 운반할 수 있고 예를 들어 유전자 서열을 표적 세포로 전달할 수 있는 임의의 폴리뉴클레오티드를 지칭한다. 따라서, 상기 용어는 클로닝, 및 발현 비히클, 뿐만 아니라 통합 벡터를 포함한다. 특히, 본원에 사용된 바와 같은 용어 "발현 벡터"는 핵산의 발현을 지향할 수 있는 임의의 폴리뉴클레오티드를 지칭한다. 일부 측면에서, 용어 "벡터" 및 "플라스미드"는 용어 "핵산 작제물"과 상호교환적으로 사용된다.

[0055] 본원에 사용된 바와 같이, 2개의 서열 사이의 퍼센트 동일성은 갭(gap)의 수, 및 및 두 서열의 최적 정렬을 위해 도입될 필요가 있는 각 갭의 길이를 고려하여, 서열에 의해 공유되는 동일한 위치의 수의 함수이다 (즉, % 동일성 = 동일한 위치의 수/위치의 총 수 x 100). 서열의 비교 및 2개의 서열 사이의 퍼센트 동일성의 결정은 하기에 기재된 바와 같이 수학적 알고리즘을 사용하여 달성될 수 있다.

[0056] 2개의 아미노산 서열 간의 퍼센트 동일성은 PAM120 중량 잔기 표, 12의 갭 길이 페널티 및 4의 갭 페널티를 사용하여, ALIGN 프로그램 (버전 2.0)에 통합된 이. 마이어즈 앤드 더블우. 밀러(E. Meyers and W. Miller) (Comput. Appl. Biosci., 4:11-17, 1988)]의 알고리즘을 사용하여 결정될 수 있다. 대안적으로, 2개의 아미노산 서열 사이의 퍼센트 동일성은 블로섬(Blossom) 62 매트릭스 또는 PAM250 매트릭스, 16, 14, 12, 10, 8, 6, 또는 4의 갭 가중치 및 1, 2, 3, 4, 5, 또는 6의 길이 가중치를 사용하여, GCG 소프트웨어 패키지의 GAP 프로그램 (<http://www.gcg.com>에서 입수가능)에 통합된 니들만 앤드 분쉬(Needleman and Wunsch) (J. Mol. Biol. 48:444-453, 1970) 알고리즘을 사용하여 결정될 수 있다.

[0057] 2개의 뉴클레오티드 아미노산 서열 사이의 퍼센트 동일성은 또한 예를 들어 디폴트(default)로서 11의 단어 길이 (W), 10의 기대치 (E), M=5, N=4, 및 두 가닥의 비교를 사용하여 핵산에 대한 BLASTN 프로그램과 같은 알고리즘을 사용하여 결정될 수 있다.

[0058] 본원에 사용된 바와 같은 용어 "재조합" 또는 "조작된"은 인위적으로 생성된 단백질 또는 핵산 서열을 지칭한다.

[0059] 본원에 사용된 바와 같이, 용어 "대상체"는 개별 유기체, 예를 들어 개별 포유동물을 지칭한다. 일부 실시양태에서, 대상체는 인간이다. 일부 실시양태에서, 대상체는 인간이 아닌 포유동물이다. 일부 실시양태에서, 대상체는 인간이 아닌 영장류이다. 일부 실시양태에서, 대상체는 설치류이다. 일부 실시양태에서, 대상체는 양, 염소, 소, 고양이, 또는 개이다. 일부 실시양태에서, 대상체는 척추동물, 양서류, 파충류, 어류, 곤충, 파리, 또는 선충류이다. 일부 실시양태에서, 대상체는 연구 동물이다.

[0060] 용어 "치료", "치료하다", 및 "치료하는"는 본원에 기재된 바와 같이, 질환 또는 장애, 또는 이의 하나 이상의 증상의 역전, 완화, 개시 지연, 또는 진행 억제를 목적으로 하는 임상 개입을 지칭한다. 본원에 사용된 바와 같이, 용어 "치료", "치료하다", 및 "치료하는"는 본원에 기재된 바와 같이, 질환 또는 장애, 또는 이의 하나 이상의 증상의 역전, 완화, 개시 지연, 또는 진행 억제를 목적으로 하는 임상 개입을 지칭한다. 일부 실시양태에서, 치료는 하나 이상의 증상이 발생한 후 및/또는 질환이 진단된 후에 투여될 수 있다. 다른 실시양태에서, 예를 들어, 증상의 부재에서 치료가 투여될 수 있고, 예를 들어, 예방하거나, 발생 가능성을 감소시키거나, 증상의 개시를 지연시키거나, 질환의 개시 또는 진행을 억제할 수 있다. 예를 들어, 치료는 증상 개시 전에 감수성이 있는 개체에게 투여될 수 있다 (예를 들어, 증상의 이력 및/또는 유전적 또는 기타 감수성 요인에 비추어). 예를 들어 재발을 예방하거나 지연시키기 위해 증상이 해결된 후에도 치료를 계속할 수 있다.

[0061] **상세한 설명**

- [0062] 본 발명은
- [0063] (i) 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질; 또는 상기 제1 단백질을 코딩하는 핵산 작제물; 및
- [0064] (ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물
- [0065] 을 포함하며; 여기서 상기 트랜스포사제가 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는, 변형된 과활성 PiggyBac인, 조성물에 관한 것이다.
- [0066] 조작된 징크 핑거 단백질(zinc finger protein) (ZFP), 전사 활성화제 유사 이펙터 뉴클레아제(transcription activator like effector 뉴클레아제) (TALEN), 그리고 보다 최근에는 Cas9와 같은 RNA-가이드된 DNA 뉴클레아제를 포함한 현재의 게놈 공학 도구(genome engineering tool)는 게놈에서 서열-특이적 DNA 절단에 영향을 미친다. 이 프로그래밍 가능한 절단은 비-상동 말단 연결 (NHEJ)을 통해 절단 부위에서 DNA의 돌연변이 또는 상동성-지향 복구(homology-directed repair) (HDR)를 통해 절단 부위 주변의 DNA의 대체를 발생시킬 수 있다.
- [0067] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 RNA-가이드된 DNA 뉴클레아제, 징크 핑거 단백질 및 전사 활성화제 유사 이펙터 뉴클레아제를 포함하거나 이로 이루어진 군으로부터 선택된다.
- [0068] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 RNA-가이드된 DNA 뉴클레아제 및 징크 핑거 단백질을 포함하거나 이로 이루어진 군으로부터 선택된다.
- [0069] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 RNA-가이드된 뉴클레아제이다.
- [0070] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 이의 변이체 및 기능적 단편을 포함하여, Cas9 단백질 (예를 들어, 제외 없이, 스트렙토코커스 피오게네스(*Streptococcus pyogenes*) Cas9 (SpCas9), 스태필로코커스 아우레우스(*Staphylococcus aureus*) Cas9 (SaCas9), 또는 캄필로박터 제주니(*Campylobacter jejuni*) Cas9 (CjCas9); 일부 다른 적합한 예가 하기에 기재될 것임), 또는 이의 변이체 (예를 들어, 닉카제 Cas9 (nCas9) 또는 데드(dead) Cas9 (dCas9)), Cas12a 단백질, Cas12b 단백질, Cpf1 단백질, 또는 CasX 단백질이다.
- [0071] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 이의 변이체 및 기능적 단편을 포함하여, Cas9 단백질이다.
- [0072] CRISPR-Cas9 시스템은 서열-특이적 이중-가닥 파단(double-strand break) (DSB)을 통해 유전자를 불활성화하거나 수정하는 데 매우 효과적인 도구이다. 이들 DSB는 세포 DNA 손상 반응 기계에 의해 인식되며 내인성 DSB 복구 경로에 의해 복구될 수 있다. 우세한 복구 경로는 비-상동 말단 연결 (NHEJ)이며, 이는 종종 작은 삽입 및/또는 결실을 발생시켜 프레임이동(frameshift) 돌연변이를 생성하고 유전자의 기능을 방해할 수 있다. 이 경로를 이용하여 유전적 녹아웃(genetic knockout) 돌연변이를 생성할 수 있다. 대안적으로, 복구 주형(repair template) (예컨대, 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스젠을 포함하거나 이로 이루어진 핵산 작제물)이 있는 경우, 손상은 상동성-지향 복구 (HDR)에 의해 원활하게(seamlessly) 복구될 수 있다. 그러나, 현저한 진전에도 불구하고, 정밀한 유전자 변형을 도입하기 위한 HDR-매개 게놈 편집은 NHEJ-매개 유전자 파괴(gene disruption)보다 훨씬 덜 효율적이다. 더욱이, HDR 경로에 의한 대규모 멀티-kb 대체는 도전과제이고 선택 및/또는 대규모 집단 세포 분류를 필요로 한다. 결과적으로, HDR 경로의 주요 적용은 현재 유전자 내 핵심 영역의 로컬 대체로 제한되나, 대규모 전장 유전자는 아니다. 상기에서 설명된 바와 같이, 본 발명은 이러한 결점을 해결한다.
- [0073] 한 실시양태에서, Cas9 단백질은 (i) 활성 DNA 절단 도메인 및 (ii) 가이드 RNA 결합 도메인을 포함한다.
- [0074] 공지된 Cas9 단백질 중에서, 에스. 피오게네스 Cas9 단백질은 게놈 공학을 위한 도구로 널리 사용되어 왔다. 이 Cas9 단백질은 2개의 별개의 뉴클레아제 도메인을 함유하는 대형, 멀티-도메인(multi-domain) 단백질이다.
- [0075] 한 실시양태에서, Cas9 단백질은 서열번호: 19를 가진 코리네박테리움 울세란스(*Corynebacterium ulcerans*) (NCBI 참조번호: NC_015683.1, NC_017317.1); 서열번호: 20을 가진 코리네박테리움 디프테리아(*Corynebacterium diphtheria*) (NCBI 참조번호: NC_016782.1, NC_016786.1); 서열번호: 21을 가진 스피로플라즈마 시르피디콜라(*Spiroplasma syrophidicola*) (NCBI 참조번호: NC_021284.1); 서열번호: 22를 가진 프레보텔라 인터메디아(*Prevotella intermedia*) (NCBI 참조번호: NC_017861.1); 서열번호: 23을 가진 스피로플라즈마 타이완렌스(*Spiroplasma taiwanense*) (NCBI 참조번호: NC_021846.1); 서열번호: 24를 가진 스트렙토코커스 이니아에(*Streptococcus iniae*) (NCBI 참조번호: NC_021314.1); 서열번호: 25를 가진 벨리엘라 발티카(*Beiliiella baltica*) (NCBI Ref: NC_018010.1); 서열번호: 26을 가진 사이크로플렉서스 토르퀴시

(*Psychroflexus torquisi*) (NCBI Ref: NC_018721.1); 서열번호: 27을 가진 스트렙토코커스 써모필러스 (*Streptococcus thermophilus*) (NCBI 참조번호: YP_820832.1); 서열번호: 28을 가진 리스테리아 이노쿠아 (*Listeria innocua*) (NCBI 참조번호: NP_472073.1); 서열번호: 29를 가진 캄필로박터 제주니 (CjCas9) (NCBI 참조번호: YP_002344900.1) (서열번호: 81에 의해 코딩됨); 서열번호: 30을 가진 네이세리아 메닝기티디스 (*Neisseria meningitidis*) (NCBI 참조번호: YP_002342100.1); 서열번호: 72을 가진 스타필로코커스 아우레우스 (SaCas9) (서열번호: 77에 의해 코딩됨); 및 서열번호: 31을 가진 스트렙토코커스 피오게네스 (SpCas9) (NCBI 참조번호: NC_017053.1)로부터의 Cas 9 단백질을 포함하거나 이로 이루어진 군으로부터 선택된다.

[0076] 한 실시양태에서, 본원에서 야생형 Cas9 단백질을 언급할 때, 상기 야생형 Cas9 단백질은 달리 특정되지 않는 한, 서열번호: 31을 가진 스트렙토코커스 피오게네스 (spCas9)로부터의 Cas9에 상응한다.

[0077] 한 실시양태에서, Cas9 단백질은 "Cas9 변이체"일 수 있다. 본원에 사용된 바와 같은 "Cas9 변이체"는 본원에 기재된 바와 같은 Cas9 단백질에 대한 상동성을 공유하는 단백질이며, 이의 단편을 포함한다.

[0078] 한 실시양태에서, Cas9 변이체는 서열번호: 31을 가진 야생형 Cas9 단백질 또는 서열번호: 19-30 또는 72을 가진 임의의 다른 Cas9 단백질과 적어도 약 70% 동일, 적어도 약 80% 동일, 적어도 약 90% 동일, 적어도 약 95% 동일, 적어도 약 96% 동일, 적어도 약 97% 동일, 적어도 약 98% 동일, 적어도 약 99% 동일, 적어도 약 99.5% 동일할 수 있거나, 적어도 약 99.9%일 수 있다.

[0079] 한 실시양태에서, Cas9 변이체는 하나 또는 수개의 아미노산 치환을 가진 Cas9 단백질의 아미노산 서열을 포함한다. 예를 들어, Cas9의 DNA 절단 도메인은 2개의 서브도메인, HNH 뉴클레아제 서브도메인 및 RuvC1 서브도메인을 포함하는 것으로 공지되어 있다. HNH 서브도메인은 gRNA에 상보적인 가닥을 절단하며, 한편 RuvC1 서브도메인은 비상보적인(non-complementary) 가닥을 절단한다.

[0080] 이들 서브도메인 내의 돌연변이는 Cas9의 뉴클레아제 활성을 침묵시킬 수 있다. 예를 들어, D10A 및 H841A 치환은 서열번호: 31을 가진 에스. 피오게네스 Cas9 단백질의 뉴클레아제 활성을 완전히 불활성화하여, sgRNA-프로그래밍된 방식에서 DNA에 결합하는 이의 능력을 여전히 유지하는 데드 Cas9 (dCas9)를 발생시키는 것으로 공지되어 있다. 원칙적으로, 또 다른 단백질이나 도메인에 융합될 때, dCas9는 적절한 sgRNA와의 공동-발현(co-expression)에 의해 그 단백질을 거의 모든 DNA 서열로 표적화할 수 있다. 한 실시양태에서, dCas9 단백질은 서열번호: 66에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 한 실시양태에서, dCas9 단백질은 서열번호: 71에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함하거나 이로 이루어진다.

[0081] Cas9 니카제 (nCas9)에 관해서, Cas9 니카제 (nCas9)는 RuvC 뉴클레아제 도메인의 점 돌연변이 (D10A)가 상이한 Cas9 뉴클레아제의 변종으로, DNA를 니킹할 수 있으나(nick), 절단할 수는 없다. 한 실시양태에서, nCas9 단백질은 서열번호: 65에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 한 실시양태에서, nCas9 단백질은 서열번호: 70에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함하거나 이로 이루어진다. 일부 실시양태에서, SaCas9 니카제 (SanCas9)는 서열번호: 80에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 일부 실시양태에서, SaCas9 니카제 (SanCas9)는 서열번호: 76에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0082] 한 실시양태에서, Cas9 변이체는 Cas9의 단편을 포함하여, 단편이 서열번호: 31을 가진 야생형 Cas9 단백질 또는 서열번호: 19-30 또는 72을 가진 임의의 다른 Cas9 단백질의 상응하는 단편과 적어도 약 70% 동일, 적어도 약 80% 동일, 적어도 약 90% 동일, 적어도 약 95% 동일, 적어도 약 96% 동일, 적어도 약 97% 동일, 적어도 약 98% 동일, 적어도 약 99% 동일, 적어도 약 99.5% 동일할 수 있거나, 적어도 약 99.9%일 수 있도록 한다.

[0083] 한 실시양태에서, Cas9 변이체는 DNA 절단 도메인 또는 가이드 RNA 결합 도메인 중 하나만을 포함한다.

- [0084] 한 실시양태에서, 예시적인 Cas9 변이체는 인간화 Cas9 (hCas9) 또는 이의 변이체 또는 기능적 단편이다. 본원에 사용되는 바와 같이, 용어 "**인간화 Cas9**" 또는 "**hCas9**"는 인간 세포에 대해 서열-최적화된 Cas9 단백질을 지칭한다.
- [0085] 한 실시양태에서, hCas9 단백질은 서열번호: 64에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 한 실시양태에서, hCas9 단백질은 서열번호: 69에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.
- [0086] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 cpf1 단백질이다. 한 실시양태에서, cpf1 단백질은 서열번호: 78에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 한 실시양태에서, cpf1 단백질은 서열번호: 74에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.
- [0087] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 CasX 단백질이다. 한 실시양태에서, 상기 CasX는 서열번호: 79에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 한 실시양태에서, 상기 CasX는 서열번호: 75에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.
- [0088] 하기에 더 상세히 설명되는 바와 같이, 본 개시내용의 특정 측면은 또한 부위-특이적 DNA 결합 단백질, 특히 RNA-가이드된 뉴클레아제, 특히 본원에 기재된 Cas9 단백질 중 임의의 것을 코딩하는 핵산 작제물을 포함하는 벡터 또는 플라스미드 (예를 들어, 발현 벡터, 패키징 벡터 등)에 관한 것이며; 상기 벡터 또는 플라스미드는 바람직하게는 숙주 세포, 예를 들어 포유동물 세포, 효모 세포, 곤충 세포, 식물 세포, 진균 세포, 또는 조류 세포에서의 발현에 적합하다.
- [0089] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 징크 핑거 단백질 (ZFP)이다.
- [0090] 징크 핑거 단백질은 서열-특이적 방식으로 DNA에 결합할 수 있는 단백질이다. ZFP는 진핵생물에 고르지 않게 분포되어 있다. ZFP는 DNA 인식, RNA 결합, 및 단백질 결합에 관여하는 것으로 확인되었다. 징크 핑거 단백질에 대한 특정 분류는 폴딩된(folded) 도메인에서 단백질 백본(backbone)의 전반적인 형상의 관점에서 "폴드 그룹"을 기반으로 한다. 징크 핑거의 가장 통상적인 "폴드 그룹(fold group)"은 C2H2 또는 Cys2His2-유사 ("클래식(classic) 징크 핑거"), 트레블 클랩(treble clef), 및 징크 리본(zinc ribbon)이다. 이들 단백질을 특징짓는 대표적인 모티프는 문헌 [Li & Liu, 2020 (*Int J Mol Sci.* 21(4):1361)]의 표 1에 개시되어 있으며, 이 표는 본원에 참조로 포함된다.
- [0091] ZFP는 게놈에서 특이적 게놈 DNA 서열에 결합할 수 있는, 임의의 ZFP, 이의 변이체 또는 기능적 단편일 수 있다. ZFP의 비제한적 예는 C2H2, gag 너클(knuckle), 트레블 클랩, 징크 리본, Zn2/Cys6-유사, 또는 TAZ2 도메인-유사, 또는 이의 임의의 조합으로부터 선택된 폴드 그룹 또는 징크 핑거 모티프를 포함하는 ZFP를 포함한다. 한 실시양태에서, ZFP는 C2H2 징크 핑거 단백질이다.
- [0092] 한 실시양태에서, ZFP는 조작된 ZFP이다. 조작된 징크 핑거 어레이는 DNA 절단 도메인 (대개 FokI의 절단 도메인)에 융합되어 징크 핑거 뉴클레아제를 생성할 수 있다. 이러한 징크 핑거-FokI 융합은 게놈 조작에 유용한 시약이 되었다.
- [0093] ZFP는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12개, 또는 그 초과 of 징크 핑거 도메인을 포함할 수 있다. ZFP는 2 내지 12개, 2 내지 10개, 2 내지 8개, 3 내지 8개, 4 내지 8개, 또는 5 내지 8개의 징크 핑거 도메인을 포함할 수 있다. 한 실시양태에서, ZFP는 6개의 징크 핑거 도메인을 포함한다.
- [0094] 통상적인 모듈 어셈블리(modular assembly) 공정은 9 염기쌍에서 18 염기쌍 길이 범위의 표적 부위를 인식하는 3-핑거, 4-, 5-, 또는 6-핑거 어레이를 생성하기 위해 각각 3-염기쌍 DNA 서열을 인식할 수 있는 별도의 징크

핑거를 조합하는 것을 포함한다. 또 다른 방법은 2-핑커 모듈을 사용하여 최대 6개의 개별 징크 핑거를 가진 징크 핑거 어레이를 생성한다.

- [0095] 한 실시양태에서, ZFP의 결합 도메인은 관심 서열에 결합하도록 조작될 수 있다. 조작된 징크 핑거 결합 도메인은 자연-발생 ZFP와 비교하여, 개선된 결합 특이성을 가질 수 있다.
- [0096] 한 실시양태에서, ZFP를 코딩하는 예시적인 핵산 서열은 서열번호: 32, 서열번호: 34, 서열번호: 36, 또는 서열번호: 38을 포함하거나 이로 이루어진다. 한 실시양태에서, 이들 서열에 의해 코딩되는 예시적인 아미노산 서열은 서열번호: 33, 서열번호: 35, 서열번호: 37, 또는 서열번호: 39를 포함하거나 이로 이루어진다.
- [0097] 한 실시양태에서, ZFP는 서열번호: 33, 35, 37 또는 39 중 어느 한 서열에 대해 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.
- [0098] 한 실시양태에서, ZFP는 Gal4 DNA 결합 도메인을 갖지 않는다. Gal4는 N이 임의의 염기일 수 있는 CGG-N11-CCG에 결합한다. 이 단백질은 갈락토스를 글루코스로 전환시키는 데 사용되는 효소를 코딩하는 갈락토스-유도 유전자 예컨대 *GAL1*, *GAL2*, *GAL7*, *GAL10*, 및 *MEL1*의 유전자 발현에 대한 양성 조절자이다. 이는 이들 유전자의 상류 활성화 서열 (UAS-G)에서 17-염기쌍 서열을 인식한다. 따라서, Gal4는 게놈에서 짧고 매우 빈번한 서열을 인식하므로, 부위-특이적이지 않는다. 한 실시양태에서, ZFP는 부위-특이적하도록 조작된 Gal4 DNA 결합 도메인을 갖는다.
- [0099] 하기에 더 상세히 설명되는 바와 같이, 본 개시내용의 특정 측면은 부위-특이적 DNA 결합 단백질, 특히 본원에 기재된 ZFP를 코딩하는 핵산 작제물을 포함하는 벡터 또는 플라스미드 (예를 들어, 발현 벡터, 패키징 벡터 등)에 관한 것이며; 상기 벡터 또는 플라스미드는 바람직하게는 숙주 세포, 예를 들어 포유동물 세포, 효모 세포, 곤충 세포, 식물 세포, 진균 세포, 또는 조류 세포에서의 발현에 적합하다.
- [0100] 본 발명에 따르면, 제2 단백질은 트랜스포사제를 포함하거나 이로 이루어진다.
- [0101] 트랜스포존은 전위를 겪을 수 있는 염색체 세그먼트, 예를 들어, 숙주 DNA에 상보적인 서열의 부재 하에 전체적으로 전위될 수 있는 DNA이다. 트랜스포존은 인간 세포에서 장거리 DNA 조작을 수행하는 데 사용할 수 있다. 포유동물 세포에 사용되는 통상적인 트랜스포존 시스템은 불활성 트랜스포존으로부터 재구성된, 슬리핑 뷰티 (SB), 및 나방(moth) 트리코플루시아(*Trichoplusia*)로부터 단리된, PiggyBac (PB)을 제한 없이 포함한다. PiggyBac은 SB보다 전위 활성도가 높으며 흉터 없이 절제할 수 있다.
- [0102] 천연 DN트랜스포존은 전형적으로 트랜스포사제 결합 부위를 운반하는 역 말단 반복부(inverted terminal repeat) (ITR)에 의해 플랭킹되는 트랜스포사제 단백질을 코딩하는 단일 유전자를 함유한다. 이의 전위 동안, 트랜스포사제 단백질은 이들 ITR을 인식하여 임의의 방식으로 다른 곳에서 요소의 절단 및 후속 재통합을 촉매한다. 게다가, 이들 트랜스포존 중 일부는 이들을 이중-구성요소 시스템(bi-component system)으로서 이용하여, 유전자 요법 프로토콜에 사용하기 위해 적합화될 수 있으며, 여기서 플라스미드는 트랜스포존 ITR 사이에 배치된 관심 DNA 서열 이 시험관 내에서 합성된 트랜스포사제 효소 또는 그 mRNA를 코딩하는 서열을 함유하는 공동-형질감염된 플라스미드에 의해 지향되는 숙주 게놈으로 도입될 수 있는 발현 카세트를 함유한다. 본 개시내용에 따르면, 트랜스포존-기반 시스템은 세포에서 치료 유전자로서, 트랜스진의 안전한 통합 및 지속적인 발현을 효율적으로 매개하는데 사용된다.
- [0103] 본 개시내용의 트랜스포사제 또는 변형된 트랜스포사제는 게놈의 특정 부위에 외인성 핵산을 삽입할 수 있는 임의의 트랜스포사제일 수 있다. 본 개시내용의 일부 측면은 본원에 기재된 방법 및 전략을 사용하여 설계된 트랜스포사제 융합 단백질을 제공한다. 본 개시내용의 일부 실시양태는 이러한 트랜스포사제 또는 변형된 트랜스포사제 및/또는 이를 포함하는 융합 단백질을 코딩하는 핵산을 제공한다. 본 개시내용의 일부 실시양태는 트랜스포사제 또는 변형된 트랜스포사제 및/또는 이를 포함하는 융합 단백질을 코딩하는 이러한 핵산 작제물을 포함하는 플라스미드 또는 발현 벡터를 제공한다.
- [0104] 트랜스포사제의 비제한적인 예는 프로그 프린스(Frog Prince), 슬리핑 뷰티, 과활성 과활성 슬리핑 뷰티, PiggyBac, 및 과활성 PiggyBac을 포함한다.
- [0105] 한 실시양태에서, 트랜스포사제는 과활성 PiggyBac 트랜스포사제이다. 일부 실시양태에서, 트랜스포사제는 서열번호: 9에 상응하거나 서열번호: 67에 의해 코딩된 과활성 PiggyBac 트랜스포사제 (본 개시내용에서 또한 hyPB 또는 간단히 PB로 지칭됨)이다.

- [0106] 한 실시양태에서, 트랜스포사제는 변형된 과활성 PiggyBac 트랜스포사제이다.
- [0107] 본원에 사용된 바와 같이, "**변형된 과활성 PiggyBac 트랜스포사제**"는 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제와 비교하여, 하나 이상의 아미노산 치환, 전형적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 아미노산 치환을 포함하는 트랜스포사제를 지칭한다. 보다 구체적으로, 변형된 과활성 PiggyBac은 (i) 야생형 과활성 PiggyBac 트랜스포사제와 비교하여 절제 활성을 증가시키기 위한 하나 이상의 아미노산 치환, 및/또는 (ii) 야생형 과활성 PiggyBac 트랜스포사제와 비교하여 DNA 결합 활성을 감소시키는 하나 이상의 아미노산 치환을 포함한다. 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9에 제시된 서열에 대해 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 아미노산 서열을 갖는다.
- [0108] 일부 실시양태에서, 과활성 PiggyBac 트랜스포사제에 대한 하나 이상의 돌연변이는 삼중 돌연변이 R372A/K375A/D450N으로 이루어지지 않으며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다.
- [0109] 일부 실시양태에서, 변형된 과활성 PiggyBac은 절제 활성을 증가시키기 위해 하나 이상의 아미노산 돌연변이를 포함한다.
- [0110] 일부 실시양태에서, 변형된 과활성 Piggybac은 아미노산 위치 번호 [194-200], [214-222], [434-442] 또는 [446-456]에 의해 정의된 영역 내에 아미노산 돌연변이, 예를 들어 위치 D198, D201, R202, M212 및/또는 S213에서의 아미노산 치환 중에서 선택된 절제 활성을 증가시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며; 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.
- [0111] 일부 실시양태에서, 변형된 과활성 PiggyBac은 위치 450, 560, 564, 573, 589, 592, 및/또는 594에서 아미노산 돌연변이 중에서 선택되는 절제 활성을 증가시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며; 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.
- [0112] 일부 실시양태에서, 변형된 과활성 PiggyBac은 M194 및/또는 D450의 위치에서 아미노산 돌연변이 중에서 선택된 절제 활성을 증가시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며, 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하며, 바람직하게는 아미노산 치환은 M194V 및/또는 D450N 중에서 선택된다.
- [0113] 일부 실시양태에서, 변형된 과활성 PiggyBac은 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함한다.
- [0114] 일부 실시양태에서, 변형된 과활성 PiggyBac은 위치 254, 275, 277, 347, 372, 375, 및/또는 465에서의 아미노산 돌연변이 중에서 선택된 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며; 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.
- [0115] 일부 실시양태에서, 변형된 과활성 PiggyBac은 R275, N347, R372, K375, R376, E377, 및 E380 중에서 선택된 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.
- [0116] 일부 실시양태에서, 변형된 과활성 PiggyBac은 R372, K375, R376, E377, 및 E380 중에서 선택된 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하며, 바람직하게는 아미노산 치환 R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택된다.
- [0117] 일부 실시양태에서, 변형된 과활성 PiggyBac은 N347, R372, 및 K375 중에서 선택된 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하며, 바람직하게는 아미노산 치환 N347S, N347A, R372A, K375A 중에서 선택되고, 보다 바람직하게는 아미노산 치환 N347S, N347A 중에서 선택된다.
- [0118] 일부 실시양태에서, 변형된 과활성 Piggybac은 상기에 정의된 바와 같이 절제 활성을 증가시키기 위한 하나 이상의 아미노산 돌연변이를 포함한다; 및 상기에 정의된 바와 같이 DNA 결합 활성을 감소시키는 하나 이상의 아미노산 돌연변이를 포함한다.
- [0119] 일부 실시양태에서, 변형된 과활성 PiggyBac은 위치 D450에서 절제 활성을 증가시키기 위한 적어도 하나의 아미노산 치환, 및 위치 N347, R372 및 K375에서 DNA 결합 활성을 감소시키기 위한 적어도 2개의 아미노산 치환을

포함하며, 바람직하게는 과활성 Piggybac의 상기 변형된 트랜스포사제는 이중 돌연변이 N347S 및 D450N 또는 삼중 돌연변이 D450N, R372A 및 K375A를 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다. 보다 바람직한 실시양태에서, 과활성 Piggybac의 변형된 트랜스포사제는 이중 돌연변이 N347S 및 D450N을 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.

[0120] 일부 실시양태에서, 이전 실시양태에서 개시된 바와 같은 변형된 과활성 Piggybac은 아미노산 위치 번호 [158-169]에 의해 정의된 영역에 적어도 하나의 돌연변이, 예를 들어 A166S; 및/또는 위치 Y527, R518, K525, N463에서 적어도 하나의 돌연변이를 추가로 포함한다.

[0121] 전형적으로, 상기 변형된 과활성 PiggyBac은 서열번호: 1의 변형된 과활성 Piggybac과 적어도 85%, 적어도 90%, 적어도 95% 동일성, 또는 100% 동일성을 갖는 아미노산 서열을 포함한다.

[0122] 일부 실시양태에서, 상기 변형된 과활성 PiggyBac은 서열번호: 9와 비교하여 하나 이상의 아미노산 치환, 전형적으로 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10 이하의 아미노산 치환을 가진 서열번호: 9의 과활성 Piggybac의 변이체이다.

[0123] 일부 실시양태에서, 상기 변형된 과활성 PiggyBac은 위치 34, 43, 117, 202, 230, 245, 268, 275, 277, 287, 290, 315, 325, 341, 346, 347, 350, 351, 356, 357, 388, 409, 411, 412, 432, 447, 460, 461, 465, 517, 560, 564, 571, 573, 576, 586, 587, 589, 592, 및/또는 594에서 하기 아미노산 돌연변이 중 하나 이상을 추가로 포함하며, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응한다.

[0124] 일부 실시양태에서, 상기 변형된 과활성 PiggyBac은 하기 돌연변이 또는 돌연변이의 조합을 포함한다: V34M, T43I, Y177H, R202K, S230N, R245A, D268N, K287A, K290A, K287A/K290A, R315A, G325A, R341A, D346N, N347A, N347S, T350A, S351E, S351P, S351A, K356E, N357A, R388A, K409A, A411T, K412A, K432A, D447A, D447N, D450N, R460A, K461A, W465A, S517A, T560A, S564P, S571N, S573A, K576A, H586A, I587A, M589V, S592G, 또는 F594L, D450N/R372A/K375A, R275A/R277A, K409A/K412A, R460A/K461A, R275A/R277A/N347S/K375A/T560A/S573A/M589V/S592G 및 R245A/R275A/R277A/R372A/W465A.

[0125] 일부 실시양태에서, 상기 변형된 과활성 PiggyBac은 하기 아미노산 치환 또는 아미노산 치환의 조합을 포함한다: R372A/K375A/D450N, R372A/K375A/R376A/D450N, K375A/R376A/E377A/E380A/D450N, R372A/K375A/R376A/E377A/E380A/D450N, M194V, M194V/R372A/K375A, S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, T43I/R372A/K375A/A411T/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응함.

[0126] 본 개시내용에 따라 사용하기 위한 매우 바람직한 변형된 과활성 PiggyBac 트랜스포사제는 하기 아미노산 치환의 조합을 포함하는 변형된 과활성 PiggyBac을 포함한다: R372A/K375A/D450N, S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, T43I/R372A/K375A/A411T/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G 및 R275A/325A/R372A/T560A, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응함.

- [0127] 일부 실시양태에서, 상기 변형된 과활성 PiggyBac은 하기 아미노산 치환 또는 아미노산 치환의 조합을 포함한다: R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G, , 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응함.
- [0128] 보다 바람직한 실시양태에서, 상기 변형된 과활성 PiggyBac은 하기 아미노산 치환의 조합을 포함한다: N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응함.
- [0129] 일부 실시양태에서, 상기 변형된 트랜스포사제는 서열번호: 1-8, 10-18 및 135-149 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0130] 일부 실시양태에서, 상기 변형된 트랜스포사제는 서열번호: 1-8 및 10-18 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0131] 일부 실시양태에서, 상기 변형된 트랜스포사제는 서열번호: 90-99 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0132] 일부 실시양태에서, 상기 변형된 트랜스포사제는 서열번호: 135-149 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다. 일부 실시양태에서, 상기 변형된 트랜스포사제는 서열번호: 135-140 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다. 일부 실시양태에서, 상기 변형된 트랜스포사제는 서열번호: 141-149 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0133] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 11의 아미노산 넘버링에 상응하는 예를 들어 아미노산 268 및/또는 346 (예를 들어, D268N 및/또는 D346N)에서 보존된 촉매 트리아드 (catalytic triad)에 관여하는 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0134] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 12의 아미노산 넘버링에 상응하는 예를 들어 아미노산 287, 287/290 및/또는 460/461 (예를 들어, K287A, K287A/K290A, 및/또는 R460A/K461A)에서 절제에 중대한 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0135] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 13의 아미노산 넘버링에 상응하는 아미노산 351, 356 및/또는 379 (예를 들어 S351E, S351P, S351A, 및/또는 K356E)에서 표적 연결 (target joining)에 관여하는 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0136] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 14의 아미노산 넘버링에 상응하는 아미노산 560, 564, 571, 573, 589, 592, 및/또는 594 (예를 들어, T560A, S564P, S571N, S573A, M589V, S592G, 및/또는 F594L)에서 통합에 중대한 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0137] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 15의 아미노산 넘버링에 상응하는 아미노산 325, 347, 350, 357 및/또는 465 (예를 들어, G325A, N347A, N347S, T350A 및/또는 W465A)에서 정렬에 관여하는 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0138] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 16의 아미노산 넘버링에 상응하는 아미노산 576 및/또는 587 (예를 들어, K576A 및/또는 I587A)에서 잘 보존된 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0139] 일부 실시양태에서, 변형된 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 17의 아미노산 넘버링에 상

응하는 586 (예를 들어, H586A)에서 Zn^{2+} 결합에 관여하는 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.

- [0140] 일부 실시양태에서, 프로그래밍 가능한 트랜스포사제는 예를 들어 서열번호: 9 또는 서열번호: 18의 아미노산 넘버링에 상응하는 315, 341, 372, 및/또는 375 (예를 들어, R315A, R341A, R372A, 및/또는 K375A)에서 통합에 관여하는 hyPB에 대한 하나 이상의 돌연변이를 포함할 수 있다.
- [0141] 일부 실시양태에서, 변형된 과활성 PiggyBac은 서열번호: 9에 제시된 서열과 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 아미노산 서열을 포함한다. 일부 실시양태에서, 변형된 과활성 PiggyBac은 과활성 PiggyBac과 비교하여 게놈으로의 DNA 통합의 이의 높은 특이성을 위해 선택된다. 일부 실시양태에서, 변형된 과활성 PiggyBac은 서열번호: 9, 10, 11, 12, 13, 14, 15, 16, 17, 또는 18과 관련하여 본원에 개시된 변형 중 하나 이상을 갖는 아미노산 서열을 포함하며, 서열번호: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 또는 18에 제시된 서열과 각각 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일함을 유지한다.
- [0142] 일부 실시양태에서, 과활성 PiggyBac 트랜스포사제는 서열번호: 67에 대해 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다. 일부 실시양태에서, SB100 트랜스포사제는 서열번호: 68에 대해 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열에 의해 코딩된다.
- [0143] 일부 실시양태에서, SB100 트랜스포사제는 서열번호: 73에 대해 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.
- [0144] 일부 실시양태에서, 변형된 트랜스포사제는 하나 이상의 돌연변이를 포함하는 변형된 슬리핑 뷰티 트랜스포사제이다. 일부 실시양태에서, 과활성 슬리핑 뷰티 트랜스포사제 또는 SB100에서의 하나 이상의 돌연변이는 서열번호: 9 또는 서열번호: 73의 L25F, R36A, I42K, G59D, I212K, N245S, K252A 및 Q271L에 상응한다.
- [0145] 특정 실시양태에서, 변형된 트랜스포사제는 Himar1C9 돌연변이체가 아니다.
- [0146] 본 개시내용의 특정 측면은 숙주 세포, 예를 들어 포유동물 세포, 효모 세포, 곤충 세포, 식물 세포, 진균 세포, 또는 조류 세포에서의 발현에 적합한 개시내용의 트랜스포사제 또는 변형된 트랜스포사제를 포함하는 핵산 작제물을 포함하는 벡터 또는 플라스미드 (예를 들어, 발현 벡터 또는 패키징 벡터)에 관한 것이다. 일부 실시양태에서, 변형된 트랜스포사제는 Cas9와의 융합 단백질로서 발현된다. 일부 실시양태에서, 변형된 트랜스포사제는 별개의 벡터로부터의 Cas9와 공동-발현되나, 동일한 세포로 전달된다. 일부 실시양태에서, 변형된 트랜스포사제 또는 이를 포함하는 융합 단백질은 세포로의 전달을 위해 렌티바이러스 입자에 패키징된다.
- [0147] 예에서 나타난 바와 같이, 새로 개발된 과활성 PiggyBac 트랜스포사제 돌연변이 라이브러리를 사용하여 특이적 표적화된 전위를 수행하는 변형된 과활성 PiggyBac을 확인하였다. 이러한 라이브러리를 사용하여 양성 표적화된 전위를 가진 변형된 과활성 PiggyBac을 확인하였다.
- [0148] 일부 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 번호에 상응하는 아미노산: 245, 275, 277, 325, 347, 351, 372, 375, 388, 450, 465, 560, 564, 573, 589, 592, 594로부터 선택된 아미노산 중 하나 이상의 돌연변이를 포함할 수 있다.
- [0149] 일부 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 번호에 상응하는 R245A, R275A, R277A, R275A/R277A, G325A, N347A, N347S, S351E, S351P, S351A, R372A, K375A, R388A, D450N, W465A, T560A, S564P, S573A, M589V, S592G, 또는 F594L로부터 선택된 아미노산 변형 중 하나 이상을 포함할 수 있다.
- [0150] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는 아미노산 변형 D450N을 포함한다.
- [0151] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 1에 상응하며 아미노산 변형 R372A, K375A 및 D450을 포함한다.
- [0152] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는, 아미노산 변형 R245A 및 D450을 포함한다.
- [0153] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는, 아미노

산 변형 R245A, G325A, 및 S573P를 포함한다.

- [0154] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는, 아미노산 변형 R245A, G325A, D450 및 S573P를 포함한다.
- [0155] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는, 아미노산 변형 N347S 또는 N347A를 포함한다.
- [0156] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는, 아미노산 변형 N347S 및 D450N을 포함한다.
- [0157] 또 다른, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 넘버링에 상응하는, 아미노산 변형 N347A 및 D450N을 포함한다. 한 실시양태에서, 이러한 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 137의 아미노산 서열을 포함한다.
- [0158] 이전에 언급한 바와 같이, 본원에 개시된 요소에 융합될 수 있지만 단독으로 또는 상이한 요소와 함께 사용될 수도 있는 변형된 과활성 PiggyBac 트랜스포사제가 제공된다. 상기 트랜스포사제는 본 발명자들에 의해 생성되었다. 따라서, 아미노산 서열 서열번호: 9를 포함하는 변형된 과활성 PiggyBac 트랜스포사제가 제공되며, 여기서:
 - [0159] - 34번 위치의 아미노산이 V 또는 M이고,
 - [0160] - 43번 위치의 아미노산이 T 또는 I이고,
 - [0161] - 177번 위치의 아미노산이 Y 또는 H이고,
 - [0162] - 202번 위치의 아미노산이 R 또는 K이고,
 - [0163] - 230번 위치의 아미노산이 S 또는 N이고,
 - [0164] - 245번 위치의 아미노산이 A이고,
 - [0165] - 268번 위치의 아미노산이 D 또는 N이고,
 - [0166] - 277번 위치의 아미노산이 R 또는 A이고,
 - [0167] - 275번 위치의 아미노산이 R 또는 A이고,
 - [0168] - 277번 위치의 아미노산이 R 또는 A이고
 - [0169] - 325번 위치의 아미노산이 A 또는 G이고,
 - [0170] - 347번 위치의 아미노산이 S, 또는 A이고,
 - [0171] - 351번 위치의 아미노산이 E, P 또는 A이고,
 - [0172] - 372번 위치의 아미노산이 R 또는 A이고,
 - [0173] - 375번 위치의 아미노산이 K 또는 A이고,
 - [0174] - 388번 위치의 아미노산이 R 또는 A이고,
 - [0175] - 409번 위치의 아미노산이 K 또는 A이고,
 - [0176] - 411번 위치의 아미노산이 A 또는 T이고,
 - [0177] - 412번 위치의 아미노산이 K 또는 A이고,
 - [0178] - 450번 위치의 아미노산이 D 또는 N이고,
 - [0179] - 460번 위치의 아미노산이 R 또는 A이고,
 - [0180] - 465번 위치의 아미노산이 W 또는 A이고,
 - [0181] - 517번 위치의 아미노산이 S 또는 A이고,
 - [0182] - 560번 위치의 아미노산이 T 또는 A이고,

- [0183] - 564번 위치의 아미노산이 P 또는 S이고,
- [0184] - 571번 위치의 아미노산이 S 또는 N이고,
- [0185] - 573번 위치의 아미노산이 S 또는 A이고,
- [0186] - 576번 위치의 아미노산이 K 또는 A이고,
- [0187] - 586번 위치의 아미노산이 H 또는 A이고,
- [0188] - 587번 위치의 아미노산이 I 또는 A이고,
- [0189] - 589번 위치의 아미노산이 M 또는 V이고,
- [0190] - 592번 위치의 아미노산이 G 또는 S이고/이거나,
- [0191] - 594번 위치의 아미노산이 L 또는 F이다.
- [0192] 본 개시내용은 또한 특히 유전자 요법, 생체외 또는 생체내에서 의약으로서 사용하기 위해 본원에 제공된 변형된 과활성 PiggyBac 트랜스포사제에 관한 것이다.
- [0193] 한 실시양태에서, 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질 (상기에 기재된 바와 같음)을 포함하거나 이로 이루어진 제1 단백질, 및 트랜스포사제 (상기에 기재된 바와 같음)를 포함하거나 이로 이루어진 제2 단백질이, 함께 융합되어 직접 또는 링커를 통해 간접적으로 융합 단백질을 형성한다.
- [0194] 한편으로는 부위-특이적 DNA 결합 단백질, 및 다른 한편으로는 트랜스포사제에 관련된 임의의 실시양태는 본원에 기재된 융합 단백질의 경우에 필요한 부분만 약간 수정하여 적용한다.
- [0195] 따라서, 한 실시양태에서, 융합 단백질은
- [0196] (i) 상기에 기재된 바와 같은, RNA-가이드된 DNA 뉴클레아제, 징크 핑거 단백질 또는 전사 활성화제 유사 이펙터 뉴클레아제를 포함하거나 이로 이루어진 제1 단백질, 및
- [0197] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질로서, 상기 트랜스포사제가 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인, 제2 단백질
- [0198] 을 포함하거나 이로 이루어진다.
- [0199] 한 실시양태에서, 융합 단백질은
- [0200] (i) 상기에 기재된 바와 같은, RNA-가이드된 DNA 뉴클레아제 또는 징크 핑거 단백질을 포함하거나 이로 이루어진 제1 단백질, 및
- [0201] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질로서, 상기 트랜스포사제가 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인, 제2 단백질
- [0202] 을 포함하거나 이로 이루어진다.
- [0203] 한 실시양태에서, 융합 단백질은
- [0204] (i) 상기에 기재된 바와 같은, RNA-가이드된 DNA 뉴클레아제를 포함하거나 이로 이루어진 제1 단백질, 및
- [0205] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질로서, 상기 트랜스포사제가 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인, 제2 단백질
- [0206] 을 포함하거나 이로 이루어진다.
- [0207] 한 실시양태에서, 융합 단백질은
- [0208] (i) 상기에 기재된 바와 같은, Cas9 단백질 또는 이의 변이체를 포함하거나 이로 이루어진 제1 단백질, 및
- [0209] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질로서, 상기 트랜스포사제

가 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인, 제2 단백질

- [0210] 을 포함하거나 이로 이루어진다.
- [0211] 한 실시양태에서, 제1 단백질 및 제2 단백질은 어느 순서로든 융합 단백질에서 배향될 수 있다.
- [0212] 한 실시양태에서, 융합 단백질은 직접적으로 또는 링커를 통해 간접적으로 제2 단백질의 C-말단 단부에 융합된 제1 단백질을 포함하거나 이로 이루어진다. 환언하면, 융합 단백질은 N-말단에서 C-말단까지: (i) 제2 단백질 (즉, 트랜스포사제); (ii) 임의로, 링커; 및 (iii) 제1 단백질 (즉, 부위-특이적 DNA 결합 단백질, 바람직하게는 RNA-가이드된 DNA 뉴클레아제; 보다 바람직하게는 Cas9 단백질 또는 이의 변이체)을 포함하거나 이로 이루어진다.
- [0213] 한 실시양태에서, 융합 단백질은 직접적으로 또는 링커를 통해 간접적으로 제2 단백질의 N-말단 단부에 융합된 제1 단백질을 포함하거나 이로 이루어진다. 환언하면, 융합 단백질은, N-말단에서 C-말단까지: (i) 제1 단백질 (즉, 부위-특이적 DNA 결합 단백질, 바람직하게는 RNA-가이드된 DNA 뉴클레아제; 보다 바람직하게는 Cas9 단백질 또는 이의 변이체); (ii) 임의로, 링커; 및 (iii) 제2 단백질 (즉, 트랜스포사제)을 포함하거나 이로 이루어진다.
- [0214] 한 실시양태에서, 융합 단백질은 링커를 포함한다.
- [0215] 링커의 적합한 예는 제1 단백질과 제2 단백질 사이 (임의의 순서로)에 펩티드 링커를 포함한다.
- [0216] 한 실시양태에서, 펩티드 링커는 $(\text{GGG})_n$, 서열번호: 133를 가진 $(\text{GGGG})_n$, $(\text{G})_n$, 서열번호: 134를 가진 $(\text{EAAAK})_n$, XTEN 링커, 및 $(\text{XP})_n$ 모티프, 및 이들 중 임의의 것의 임의의 것의 조합을 포함하거나 이로 이루어진 군으로부터 선택되며, 여기서 n 은 독립적으로 1 내지 50의 정수이다.
- [0217] 한 실시양태에서, 링커는 12- 내지 24-아미노산 길이어거나, 36- 내지 72- 뉴클레오티드 길이인 핵산 서열에 의해 코딩된다.
- [0218] 한 실시양태에서, 링커는 XTEN 링커 또는 $(\text{GGG})_n$ 링커이다.
- [0219] 한 실시양태에서, 링커는 표 1에 나타난 링커 중에서 선택된다.

[0220] [표 1]

링커	핵산 서열 (서열번호)	아미노산 서열 (서열번호)
GGs3	ggtagatctggcggtagatctggcgggt (서열번호: 48)	GGSGGSGGG (서열번호: 49)
GGs4x	ggaggagtagtggtggcggtagtgcgatcc (서열번호: 50)	GGSGGSGSGGS (서열번호: 51)
GGs5x	ggaggtctcgggtggtcggtagcggtagtg cgatcc (서열번호: 52)	GGSGGSGSGSGGS (서열번호: 53)
GGs6x	ggaggcagtggtgggagcgggttcgggtagtg gtggttcggggatcc (서열번호: 54)	GGSGGSGSGSGSGGS (서열번호: 55)
GGs7x	ggaggttctggaggtcgggttcggggaagt gggggtcaggcgatcaggagatcc (서열번호: 56)	GGSGGSGSGSGSGGS GGS (서열번호: 57)
GGs8x	ggaggtagcggaggttcgggagcggcggagt gggggaagcggggaagtggagatccggggagg atcc (서열번호: 58)	GGSGGSGSGSGSGGS GGS (서열번호: 59)
링커 XTEN	tccgtagcgaacaccgggactcagaatcgccac cccgagtct (서열번호: 60)	SGSETPGTSESATPES (서열번호: 61)
링커 B	ggaagcggcggtagtgcggtcgggtcggcagttc (서열번호: 62)	GSAGSAAGSGEF (서열번호: 63)

[0221]

[0222] 한 실시양태에서, 링커는 선택된 서열번호: 49, 서열번호: 51, 서열번호: 53, 서열번호: 55, 서열번호: 57, 서열번호: 59, 서열번호: 61, 서열번호: 63, 또는 이의 임의의 조합을 포함하거나 이로 이루어진 군으로부터 선택된 아미노산 서열을 포함하며; 각각 서열번호: 48, 서열번호: 50, 서열번호: 52, 서열번호: 54, 서열번호: 56, 서열번호: 58, 서열번호: 60, 서열번호: 62의 예시적인 핵산 서열에 의해 코딩된다.

[0223] 한 실시양태에서, 링커는 서열번호 49의 아미노산 서열을 포함하거나 이로 이루어지며; 서열번호: 48의 예시적인 핵산 서열에 의해 코딩된다..

[0224] 또한, 본 개시내용에 제공된 핵산 작제물 중 임의의 것의 발현으로부터 수득된 융합 단백질이 본원에 제공된다.

[0225] 한 실시양태에서, 융합 단백질은 삼중 융합 단백질이다.

[0226] 이러한 삼중 융합 단백질은 다음을 포함하거나 이루어질 수 있다:

[0227] - 1개의 제1 단백질 (즉, 1개의 부위-특이적 DNA 결합 단백질) 및 2개의 제2 단백질 (즉, 2개의 트랜스포사제); 또는

[0228] - 2개의 제1 단백질 (즉, 2개의 부위-특이적 DNA 결합 단백질) 및 1개의 제2 단백질 (즉, 1개의 트랜스포사제).

[0229] 한 실시양태에서, 삼중 융합은 1개의 제1 단백질 (즉, 1개의 부위-특이적 DNA 결합 단백질) 및 2개의 제2 단백질 (즉, 2개의 트랜스포사제)을 포함하거나 이로 이루어지고, 삼중 융합은 N-말단에서 C-말단까지 다음을 포함한다:

[0230] - (i) 부위-특이적 DNA 결합 단백질, (ii) 제1 트랜스포사제; (iii) 제2 트랜스포사제; 또는

- [0231] - (i) 제1 트랜스포사제; (ii) 부위-특이적 DNA 결합 단백질, (iii) 제2 트랜스포사제; 또는
- [0232] - (i) 제1 트랜스포사제; (ii) 제2 트랜스포사제, (iii) 부위-특이적 DNA 결합 단백질.
- [0233] 한 실시양태에서, 제1 및 제2 트랜스포사제는 동일하다. 한 실시양태에서, 제1 및 제2 트랜스포사제는 상이하다. 예를 들어, 제1 트랜스포사제는 과활성 PiggyBac 트랜스포사제일 수 있고 제2 트랜스포사제는 본원에 기재된 변형된 과활성 PiggyBac 트랜스포사제 중 임의의 것 중에서 선택된, 변형된 과활성 PiggyBac 트랜스포사제일 수 있다. 대안적으로, 제1 및 제2 트랜스포사제는 변형된 과활성 PiggyBac 트랜스포사제일 수 있으나, 각각은 본원에 기재된 바와 같이 상이한 치환 또는 치환의 상이한 조합을 갖는다.
- [0234] 한 실시양태에서, 제1 및 제2 트랜스포사제는 기능적 이량체를 형성할 수 있다.
- [0235] 한 실시양태에서, 삼중 융합은 2개의 제1 단백질 (즉, 2개의 부위-특이적 DNA 결합 단백질) 및 1개의 제2 단백질 (즉, 1개의 트랜스포사제)을 포함하거나 이로 이루어지고, 삼중 융합은 N-말단에서 C-말단까지를 포함한다:
- [0236] - (i) 제1 부위-특이적 DNA 결합 단백질, (ii) 제2 부위-특이적 DNA 결합 단백질; (iii) 트랜스포사제; 또는
- [0237] - (i) 제1 부위-특이적 DNA 결합 단백질; (ii) 트랜스포사제, (iii) 제2 부위-특이적 DNA 결합 단백질; 또는
- [0238] - (i) 트랜스포사제; (ii) 제1 부위-특이적 DNA 결합 단백질, (iii) 제2 부위-특이적 DNA 결합 단백질.
- [0239] 한 실시양태에서, 제1 및 제2 부위-특이적 DNA 결합 단백질은 동일하다. 한 실시양태에서, 제1 및 제2 부위-특이적 DNA 결합 단백질은 상이하다. 예를 들어, 제1 부위-특이적 DNA 결합 단백질은 Cas9 단백질일 수 있고 제2 부위-특이적 DNA 결합 단백질은 본원에 기재된 Cas9 단백질 변이체 중 임의의 것 중에서 선택된 Cas9 단백질의 변이체일 수 있다. 대안적으로, 제1 및 제2 부위-특이적 DNA 결합 단백질은 둘 다 Cas9 단백질 변이체일 수 있으나, 각각은 상이한 변이체이다..
- [0240] 한 실시양태에서, 삼중 융합 단백질은 이의 단백질 중 2개 또는 3개의 단백질 사이에 링커를 임의로 포함한다.
- [0241] 또한, 다음을 포함하는 융합 단백질이 본원에 개시된다:
- [0242] (i) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기에 기재된 바와 같은 상기 제2 단백질을 코딩하는 핵산 작제물, 및
- [0243] (ii) 적어도 하나의 특이적 RNA 서열에 결합할 수 있는 RNA-결합 단백질; 또는 상기 RNA-결합 단백질을 코딩하는 핵산 작제물.
- [0244] 한 실시양태에서, 융합 단백질은 상기에 기재된 바와 같은 링커를 포함한다.
- [0245] 한 실시양태에서, 제2 단백질은 트랜스포사제를 포함하거나 이로 이루어지며, 상기 트랜스포사제는 서열번호: 9를 가진 과활성 PiggyBac이다. 한 실시양태에서, 제2 단백질은 트랜스포사제를 포함하거나 이로 이루어지며, 상기 트랜스포사제는 서열번호: 9를 갖는 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac이다. 특히, 변형된 과활성 PiggyBac은 본원에 개시된 것들 중 임의의 것일 수 있다.
- [0246] 한 실시양태에서, 트랜스포사제/RNA 결합 단백질 융합은 상기에 기재된 바와 같이 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질에 추가로 융합될 수 있다.
- [0247] 일부 실시양태에서, RNA-결합 단백질은 MS2 박테리오파지 코트 단백질 (MCP) 또는 이의 단편이다.
- [0248] 일부 실시양태에서, MCP는 서열번호: 151 (예를 들어, 서열번호: 150을 가진 핵산 서열에 의해 코딩됨)과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성을 갖는다.
- [0249] 일부 실시양태에서, RNA 결합 단백질은 적어도 하나의 특이적 RNA 서열에 결합할 수 있으며, 상기 RNA 서열은 테트라루프를 포함한다. 용어 "**테트라루프**"는 용어 "스텝 루프(stem loop)" 및 "헤어핀 루프(hairpin loop)"와 상호교환적으로 사용된다.
- [0250] 일부 실시양태에서, 적어도 하나의 테트라루프는 MS2 RNA 테트라루프-결합 서열이다.
- [0251] 일부 실시양태에서, 테트라루프는 가이드 RNA (gRNA) 내에 포함된다. 특정 실시양태에서, gRNA는 상기에 기재된 바와 같이, Cas9 단백질과 복합체를 이룬다.
- [0252] 일부 실시양태에서, gRNA는 적어도 하나의 MS2 RNA 테트라루프-결합 서열을 포함한다. 일부 실시양태에서,

gRNA는 하나 초과 MS2 RNA 테트라루프-결합 서열을 포함한다.

- [0253] 일부 실시양태에서, 적어도 하나의 MS2 RNA 테트라루프-결합 서열을 포함하는 gRNA는 서열번호: 153과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성 (예를 들어, 서열번호: 152를 가진 DNA 서열에 의해 코딩됨)을 가진다.
- [0254] 일부 실시양태에서, 융합 단백질의 MCP는 Cas9 단백질에 비공유결합으로 결합된 gRNA 자체에 포함된 적어도 하나의 MS2 RNA 테트라루프-결합 서열에 비공유결합으로 결합하고; 특히, Cas9/gRNA 복합체에 대한 융합 단백질의 결합은 변형된 과활성 PiggyBac 트랜스포사제의 절단 활성을 Cas9/gRNA 복합체에 의해 특이적으로 인식되는 부위로 지향한다.
- [0255] 하기에 더 상세히 설명되는 바와 같이, 본 개시내용의 특정 측면은 또한 본원에 기재된 융합 단백질을 코딩하는 핵산 작제물을 포함하는 벡터 또는 플라스미드 (예를 들어, 발현 벡터, 패키징 벡터 등)에 관한 것이며; 상기 벡터 또는 플라스미드는 바람직하게는 숙주 세포, 예를 들어 포유동물 세포, 효모 세포, 곤충 세포, 식물 세포, 진균 세포, 또는 조류 세포에서의 발현에 적합하다.
- [0256] 본 발명에 따르면, 조성물은 상기에 기재된 바와 같이 단백질로서; 또는 이들 단백질을 코딩하는 핵산 작제물로서, 제1 단백질 및/또는 제2 단백질 (또는 이 둘 다를 포함하는 융합 단백질)을 포함한다.
- [0257] 예를 들어, 게놈 DNA에 특정 변형 (예를 들어, 외인성 핵산 삽입)을 도입하는 것과 같은 핵산 서열의 표적 편집은 인간 유전 질환(genetic disease)을 치료하기 위한 유망한 접근법이다. 이를 위해, 본 발명자들은 원하는 변형을 설치하는 데 매우 효율적인 게놈 편집에 사용하기 위한 개선된 핵산 작제물을 제공하는 것; 최소한의 오프-타겟 활성; 및 인간 게놈 내의 한 부위를 정밀하게 편집하도록 프로그래밍되는 능력을 목표로 하였다.
- [0258] 따라서 본 출원의 특정 측면은 외인성 핵산, 예를 들어 관심 유전자 (GOI)의 게놈 내로의 부위-특이적 삽입을 개선하는데 사용하기 위한 핵산 작제물에 관한 것이다. 일부 실시양태에서, GOI는 치료 유전자, 예를 들어 치료 단백질을 코딩하는 유전자이다. 관심 치료 유전자의 예는 낭포성 섬유증 질환(Cystic Fibrosis disease)을 치료하기 위한 CFTR 유전자(낭포성 섬유증 막투과 전도 조절자(Cystic fibrosis transmembrane conductance regulator)); 척수성 근위축증(Spinal muscular atrophy) (SMA)을 치료하기 위한 *SMN1* 유전자 (생존 운동 뉴런 (Survival motor neuron) 1); 골다공증(osteoporosis) 및 골절(bone fracture)을 예방하기 위한 *LRP5* 유전자 (LDL 수용체 관련 단백질 5) 변이체 G171V; 및 알츠하이머병 소인(Alzheimer's predisposition)을 감소시키는 *APP* 유전자 (아밀로이드 베타 전구체 단백질) 변이체를 포함한다.
- [0259] 일부 실시양태에서, 삽입을 위한 외인성 핵산 (예를 들어, GOI)은 최대 약 10 kb, 최대 약 15 kb, 최대 약 20 kb 길이, 최대 약 25 kb 길이, 최대 약 30 kb 길이, 최대 약 35 kb 길이, 또는 최대 약 40 kb 길이이다.
- [0260] 일부 실시양태에서, 삽입을 위한 외인성 핵산은 예를 들어, 최대 10 kb, 최대 15 kb, 최대 20 kb 길이, 최대 25 kb 길이, 최대 30 kb 길이, 최대 35 kb 길이, 또는 최대 40 kb 길이, 예를 들어, 약 1 kb 내지 약 40 kb, 약 1 kb 내지 약 39 kb, 약 1 kb 내지 약 38 kb, 약 1 kb 내지 약 37 kb, 약 1 kb 내지 약 36 kb, 또는 약 1 kb 내지 약 35 kb 그리고 보다 바람직하게는 5 내지 25 kb, 전형적으로 8 내지 20 kb이다.
- [0261] 따라서, 한 실시양태에서, 본 발명의 조성물은
- [0262] a. 상기에 기재된 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진, 상기에 기재된 제1 단백질을 코딩하는 핵산 작제물;
- [0263] b. 상기에 기재된 바와 같이, 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는, 변형된 과활성 PiggyBac인 트랜스포사제를 포함하거나 이로 이루어진, 제2 단백질을 코딩하는 핵산 작제물
- [0264] 을 포함하거나 이로 이루어진다.
- [0265] 또 다른 실시양태에서, 본 발명의 조성물은 (i) 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질, 및 (ii) 상기에 기재된 바와 같이, 서열번호: 9의 과활성 PiggyBac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는, 변형된 과활성 PiggyBac인 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질을 포함하거나 이로 이루어진, 상기에 기재된 융합 단백질을 코딩하는 핵산 작제물을 포함하거나 이로 이루어진다.
- [0266] 한 실시양태에서, 융합 단백질을 코딩하는 핵산 작제물은 상기에 기재된 바와 같이 제1 단백질과 제2 단백질 사이의, 또는 삼중 융합 단백질의 경우에, 이의 단백질 둘 사이 또는 세 단백질 사이의 링커를 코딩하는 핵산 서

열을 추가로 포함한다.

- [0267] 본 개시내용에 따르면, 제1 및 제2 단백질, 또는 상기 제1 및 제2 단백질을 포함하거나 이로 이루어진 융합 단백질은 외인성 핵산의 부위-특이적 삽입을 가능하게 하고/하거나 촉진한다.
- [0268] 일부 실시양태는 다음 중 하나를 포함하는 플라스미드 또는 벡터 (예컨대, 예를 들어, 발현 벡터)에 관한 것이다:
- [0269] - 제1 단백질을 코딩하는 핵산 작제물; 또는
- [0270] - 제2 단백질을 코딩하는 핵산 작제물; 또는
- [0271] - 제1 단백질을 코딩하는 핵산 작제물 및 제2 단백질을 코딩하는 핵산 작제물; 또는
- [0272] - 융합 단백질 또는 삼중 융합 단백질을 코딩하는 핵산 작제물.
- [0273] 일부 실시양태에서, 플라스미드는 패키징 플라스미드이다. 일부 실시양태에서, 플라스미드는 캡시드 단백질, 예를 들어 gag 및 pol을 코딩하는 폴리뉴클레오티드를 추가로 포함한다. 일부 실시양태에서, 플라스미드는 바이러스 외피(envelope) (외피 플라스미드)에 대한 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 제2 플라스미드; 및 외인성 핵산 트랜스진을 포함하는 핵산 작제물을 포함하는 제3 플라스미드와 조합되며, 여기서 조합이 생산 세포주 (예를 들어, 진핵 세포, 원핵 세포 및/또는 세포주)에 도입될 때, 외인성 핵산 트랜스진을 코딩하는 핵산 작제물 및 제1 단백질, 제2 단백질, 제1 및 제2 단백질 둘 다 또는 융합 단백질 중 어느 하나를 코딩하는 핵산 작제물을 포함하는 바이러스 입자가 생성된다.
- [0274] 일부 실시양태에서, 플라스미드는 캡시드 단백질, 예를 들어 gag 및 pol(패키징 플라스미드, 여기서 패키징 플라스미드는 기능성 인테그라제를 결여함)을 코딩하는 폴리뉴클레오티드를 포함하는 제2 플라스미드, 바이러스 외피 (외피 플라스미드)에 대한 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 제3 플라스미드 및 외인성 핵산 트랜스진을 포함하는 핵산 작제물을 포함하는 제4 플라스미드와 조합되며, 여기서 조합이 생산 세포주 (예를 들어, 진핵 세포, 원핵 세포 및/또는 세포주)에 도입될 때, 외인성 핵산 트랜스진을 포함하는 핵산 작제물 및 제1 단백질, 제2 단백질, 제1 및 제2 단백질 둘 다 또는 융합 단백질 중 어느 하나를 코딩하는 핵산 작제물을 포함하는 바이러스 입자가 생성된다.
- [0275] 한 실시양태에서, 제1 단백질, 제2 단백질, 제1 및 제2 단백질 둘 다 또는 융합 단백질, 및/또는 외인성 핵산 트랜스진은 렌티바이러스 입자를 사용하여 세포로 전달된다.
- [0276] 한 실시양태에서, 핵산 작제물은 표적 핵산 서열에 결합하도록 조작된 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질을 코딩하는 제1 폴리뉴클레오티드 서열, 외인성 핵산 트랜스진을 게놈에 삽입할 수 있게 하는 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질을 코딩하는 제2 폴리뉴클레오티드 서열, 및 임의로, 제1 폴리뉴클레오티드와 제2 폴리뉴클레오티드 사이의 링커를 코딩하는 핵산 서열을 포함하는 제3 폴리뉴클레오티드 서열을 포함한다. 일부 실시양태에서, 제1 단백질은 상기에 기재된 바와 같은, 징크 핑거 단백질 또는 Cas9 단백질 또는 이의 변이체이고/거나; 제2 단백질은 상기에 기재된 바와 같은, 변형된 과활성 PiggyBac 트랜스포사제이다.
- [0277] 융합 단백질을 생성하기 위한 적합한 링커의 예는 상기에 기재되어 있다.
- [0278] 일부 실시양태에서, 제1 단백질이 제2 단백질과 별도의 플라스미드로부터 발현되기 때문에 링커가 필요하지 않다.
- [0279] 한 실시양태에서, 링커를 사용하는 대신에, 제1 및/또는 제2 폴리뉴클레오티드 서열은 각각 제1 및 제2 단백질을 코딩하는 핵산을 포함하며, 이의 적어도 하나의 말단에 링커의 기능을 만드는 추가 뉴클레오티드를 추가로 포함한다.
- [0280] 한 실시양태에서, 핵산 작제물은 DNA 또는 RNA 형태이다.
- [0281] 또한, 본 개시내용에 제공된 핵산 작제물 중 임의의 것을 포함하는 벡터가 본원에 제공된다. 특히, 벡터는 포유동물 세포, 효모 세포, 곤충 세포, 식물 세포, 진균 세포, 또는 조류 세포에서의 발현에 적합하다. 또한, 본 개시내용에 제공된 핵산 작제물 또는 벡터 중 임의의 것을 포함하는 숙주 세포가 본원에 제공된다.
- [0282] 일부 실시양태에서, 본 개시내용의 핵산 작제물은 숙주 세포에서 발현된다. 적합한 숙주 세포는 진핵 및 원핵 세포 및/또는 세포주를 포함하나 이에 제한되지는 않는다. 이러한 숙주 세포 또는 이러한 세포로부터 생성된

세포주의 비제한적 예는 COS, CHO (예를 들어, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NSO, SP2/0-Ag14, HeLa, HEK293 (예를 들어, HEK293-F, HEK293-H, HEK293-T), 및 perC6 세포뿐만 아니라 곤충 세포 예컨대 스포도페르다 푸기페르다(*Spodoptera fugiperda*) (Sf), 또는 진균 세포 예컨대 사카로마이세스(*Saccharomyces*), 피치아(*Pichia*) 및 스킴조사카로마이세스(*Schizosaccharomyces*)를 포함한다.

[0283] 일부 실시양태에서, 숙주 세포는 미생물로부터 유래한다. 본원에 개시된 특정 방법에 유용한 미생물은 예를 들어 박테리아 (예를 들어, 이. 콜라이(*E. coli*)), 효모 (예를 들어, 사카로마이세스 세레비시아에(*Saccharomyces cerevisiae*)), 및 식물을 포함한다. 숙주 세포는 원핵 또는 진핵일 수 있다. 일부 실시양태에서, 숙주 세포는 진핵 세포이다. 적합한 진핵 숙주 세포는 효모 세포, 곤충 세포, 식물 세포, 진균 세포, 및 조류 세포를 포함하나 이에 제한되지는 않는다.

[0284] 일부 실시양태에서, 숙주 세포는 적격(competent) 숙주 세포이다. 일부 실시양태에서, 숙주 세포는 자연적으로 적격이다. 일부 실시양태에서, 숙주 세포는 예를 들어 염화칼슘 및 열 충격(heat shock)을 사용하는 공정에 의해 적격하게 제조된다. 사용되는 세포는 임의의 세포 적격, 특히 진핵 세포, 특히 포유동물, 예를 들어 인간 또는 동물일 수 있다. 이들은 체세포 또는 배아 줄기이거나 분화될 수 있다. 일부 측면에서, 세포는 293T 세포, 섬유모세포, 간세포, 근육 세포 (골격, 심장, 평활, 혈관 등), 상피 세포의 신경 세포 (뉴런, 신경아교세포, 성상세포), 신장, 안구 등을 포함한다. 이는 또한 곤충, 식물 세포, 효모, 또는 원핵 세포를 포함할 수 있다. 대안적으로, 1차 세포는 뉴클레아제 (예를 들어, ZFN 또는 TALEN) 또는 뉴클레아제 시스템 (예를 들어, CRISPR/Cas)으로 치료한 후 치료할 대상체에 재도입하기 위해 생체외에서 단리 및 사용할 수 있다. 적합한 1차 세포는 말초 혈액 단핵 세포 (PBMC), 및 CD4+ T 세포 또는 CD8+ T 세포와 같은 T-림프구와 같으나 이에 제한되지는 않는 다른 혈액 세포 서브세트를 포함한다. 적합한 세포는 또한 예로서 배아 줄기 세포, 유도 만능 줄기 세포, 조혈 줄기 세포 (CD34+), 신경 줄기 세포 및 중간엽 줄기 세포와 같은 줄기 세포를 포함한다.

[0285] 일부 실시양태에서, 숙주 세포는 본원에 개시된 핵산 작제물을 포함하는 플라스미드로 형질감염된다. 일부 실시양태에서, 핵산 작제물을 포함하는 플라스미드는 패키징 플라스미드이다. 일부 실시양태에서, 핵산 작제물을 포함하는 플라스미드는 캡시드 단백질, 예를 들어 gag 및 pol을 코딩하는 폴리뉴클레오티드를 추가로 포함한다. 일부 실시양태에서, 숙주 세포는 (i) 숙주 세포에서 조합되는 핵산 작제물을 포함하는 플라스미드 (ii) 바이러스 외피용 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 플라스미드 (외피 플라스미드); 및 (iii) 외인성 핵산 서열 (예를 들어, GOI)을 포함하는 플라스미드로 형질감염되며, 여기서 외인성 핵산, 예를 들어, GOI, 및 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 바이러스 입자가 생성된다.

[0286] 일부 실시양태에서, 숙주 세포는 (i) 핵산 작제물을 포함하는 플라스미드 (ii) 캡시드 단백질, 예를 들어 gag 및 pol을 코딩하는 폴리뉴클레오티드를 추가로 포함하는 핵산 작제물을 포함하는 플라스미드 (패키징 플라스미드, 여기서 패키징 플라스미드는 기능적 인테그라제(integrase)가 결합되어 있음); (iii) 바이러스 외피용 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 플라스미드 (외피 플라스미드) 및 (iv) 외인성 핵산 서열 (예를 들어, GOI)을 포함하는 플라스미드로 형질감염되며, 여기서 외인성 핵산, 예를 들어, GOI, 및 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 바이러스 입자가 생성된다.

[0287] 추가 실시양태에서, 벡터, 예를 들어, 본 개시내용에 따른 렌티바이러스 벡터는 본 개시내용의 핵산 작제물에 의해 코딩된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서) 및 외인성 핵산을 유기체, 예를 들어 포유동물에, 그리고 보다 특히 관심 포유동물 표적 세포에 전달하기 위해 사용될 수 있다. 제1 및 제2 단백질을 포함하는 렌티바이러스 벡터 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)는 예를 들어 간 세포 (예를 들어 간세포), 근육 세포, 뇌 세포, 신장 세포, 망막 세포, 및 조혈 세포와 같은 다양한 세포 유형을 형질도입할 수 있다. 일부 실시양태에서, 본 발명의 표적 세포는 "비분할(non-dividing)" 세포이다. 이들 세포는 정상적으로 분열하지 않는 신경 세포와 같은 세포를 포함한다. 그러나, 본 개시내용이 비분할 세포 (근육 세포, 백혈구, 비장 세포, 간 세포, 눈 세포(eye cell), 상피 세포 등을 포함하나 이에 제한되지는 않음)에 제한되는 것은 아닌 것으로 의도된다.

[0288] 특정 실시양태에서, 패키징된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 예를 들어 유기체 DNA의 유전자 편집을 위해 유기체에 투여된다. 일부 실시양태에서, 유기체는 인간이다. 일부 실시양태에서, 유기체는 인간이 아닌 포유동물이다. 일부 실시양태에서, 유기체는 인간이 아닌 영장류이다. 일부 실시양태에서 유기체는 설치류이다. 일부 실시양태에서, 유기체는 양, 염소, 소, 고양이, 또는 개이다. 일부 실시양태에서, 유기체는 척추동물, 양서류, 파충류, 어류, 곤충, 파리, 또는 선충류이다. 일부 실시양태에

서, 유기체는 연구 동물이다. 일부 실시양태에서, 유기체는 유전적으로 조작된, 예를 들어 유전적으로 조작된 비-인간 대상체이다. 유기체는 성별 어느 쪽이든 및 임의의 발달 단계에 있을 수 있다.

- [0289] 핵산, 예를 들어 외인성 핵산을 게놈에 삽입하는 방법이 기재되어 있다. 예를 들어, 문헌 [Yusa et al. PNAS 4(108):1531-1536 (2011); Feng et al. Nuc. Acid Res. 4(38):1204-1216 (2009); Kettlun et al. Amer. Soc. Gene and Cell Ther. 9(19):1636-1644 (2011); Skipper et al. 20(92):1-23 (2013); Li et al. PNAS 25:E2279-E2287 (2013); Máts et al. Nature Genetics 41(6):753-761 (2009); Mali et al. Nat. Methods 10(10):957-963; Vargas et al. J. Trans. Med. 14(288):1-15 (2016); Gersbach et al. Acc. Chem. Res. 47:2309-2318 (2014); Chandrasegaran et al. Cell Gene Ther. Ins. 3(1):33-41 (2017); Wilson et al. 649:353-363 (2010); Zhao Zhang, et al. Mol Ther Nucleic Acids. 9: 230-241 (2017); Naldini L. EMBO Mol Med. 11(3) (2019)]; 및 [Naldini L, et al. Hum Gene Ther. 27(10):727-728 (2016)]을 참조하며, 상기 문헌들 각각은 본원에 참조로 포함된다.
- [0290] 본 개시내용은 핵산 (전형적으로 외인성 핵산)을 게놈의 특정 부위에 삽입하기 위한, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 코딩하는 핵산 작제물을 제공한다. 본 발명은 또한 외인성 핵산을 게놈의 특정 부위에 삽입하기 위한, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 제공한다. 일부 실시양태에서 삽입을 위한 외인성 핵산은 최대 5 kb 길이, 최대 10 kb 길이, 최대 15 kb 길이, 최대 20 kb 길이, 최대 25 kb 길이, 최대 30 kb 길이, 최대 35 kb 길이, 또는 최대 40 kb 길이, 그리고 특히 긴 핵산의 경우, 예를 들어 5 kb 내지 25 kb, 전형적으로 8 kb 내지 20 kb일 수 있다.
- [0291] 또 다른 실시양태에서, 게놈으로 부위-특이적 핵산 삽입을 위한 방법이 제공된다.
- [0292] 따라서, 본 개시내용은 세포 게놈으로의 외인성 핵산 서열의 부위 특이적 통합(site specific integration) 방법으로서, 세포에
- [0293] (i) 본원에 개시된 바와 같은 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), 또는 본원에 개시된 바와 같은 핵산 작제물,
- [0294] (ii) 세포 게놈에 통합될 외인성 핵산, 및
- [0295] (iii) 세포 게놈으로의 상기 외인성 핵산의 부위-특이적 통합을 결정하기 위한 가이드 RNA
- [0296] 를 포함하는 조성물을 전달하는 것을 포함하며,
- [0297] 여기서 상기 제1 및 제2 단백질의 결합 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 세포 게놈 내의 특이적 게놈 DNA 서열에 결합하여 게놈의 절단 및 가이드 RNA에 의해 결정된 바와 같이 세포 게놈으로의 상기 외인성 핵산의 부위-특이적 통합을 발생시킨다.
- [0298] 상기 방법의 구체적 실시양태에서, 상기 외인성 핵산은 크기가 적어도 5 kb, 적어도 6 kb, 적어도 7 kb, 적어도 8 kb, 적어도 9 kb, 전형적으로 5 내지 25 kb, 바람직하게는 8 내지 20 kb로 포함되는 핵산 단편이다
- [0299] 상기 방법의 구체적 실시양태에서, 상기 외인성 핵산은 유전적 장애(genetic diorder)의 결핍을 교정하기 위해 이를 필요로 하는 대상체의 게놈에 삽입되는 치료적 트랜스젠이다.
- [0300] 상기 방법의 구체적 실시양태에서, 상기 조성물은 전형적으로 포유동물 세포, 바람직하게는 인간 세포에서, 그리고 보다 바람직하게는 유전 질환을 앓고 있는 인간 대상체로부터 수득된 인간 세포에서 시험관내 또는 생체의 전달된다.
- [0301] 상기 방법의 구체적 실시양태에서, 상기 조성물은 유전적 장애의 치료적 치료를 위해 이를 필요로 하는 포유동물, 예를 들어 인간 대상체 내로 생체내 전달된다.
- [0302] 일부 실시양태에서, 방법은 표적 DNA를 본원에 기재된 Cas9 및 트랜스포사제를 포함하는 융합 단백질 중 임의의 것과 접촉시키는 것을 포함한다. 예를 들어, 일부 실시양태에서, 방법은 2개의 연결된 폴리펩티드: (i) Cas9; 및 (ii) 트랜스포사제를 포함하는 융합 단백질과 DNA를 접촉시키는 것을 포함하며, 여기서 활성 Cas9는 DNA의 영역, 예를 들어 게놈 DNA에 하이브리드화하는 gRNA에 결합한다.
- [0303] 일부 실시양태에서, 방법은 표적 DNA를 본원에 기재된 ZFP 및 인테그라제를 포함하는 융합 단백질 중 임의의 것과 접촉시키는 것을 포함한다. 예를 들어, 일부 실시양태에서, 방법은 2개의 연결된 폴리펩티드: (i) ZFP; 및 (ii) 인테그라제를 포함하는 융합 단백질과 DNA를 접촉시키는 것을 포함하며, 여기서 활성 ZFP는 DNA의 영역,

예를 들어 게놈 DNA에 하이브리드화한다.

- [0304] 일부 실시양태에서, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 표적 DNA, 예를 들어 게놈 DNA를 포함하는 유기체 및/또는 세포에 바이러스 벡터, 예를 들어 렌티 바이러스 입자를 사용하여 전달된다.
- [0305] 렌티바이러스 패키징 방법이 기재되어 있다. 문헌 [Grandchamp et al. 9(6):1-13 (2014); Voelkel et al. 107(17):7805-7810 (2010); Tan et al. 80(4):1939-1948; Li et al. 9(8):1-9 (2014); Mátés et al. Nature Genetics 41(6):753-761 (2009)]; 및 [Robert H Kutner1, et al. NATURE PROTOCOLS 4(4):495 (2009)]을 참조하며, 상기 문헌의 각각은 본원에 참조로 포함된다.
- [0306] 전형적으로, 렌티바이러스 전달 시스템은 바이러스성 질환을 유발하는 데 필요한 유전적 구성요소를 함유하지 않는 완전한 바이러스를 생산하기 위해 사용되는 별도의 플라스미드에 상이한 렌티바이러스 유전자를 가진 분할 시스템을 사용한다. 예를 들어, 하나의 플라스미드 (외피 플라스미드)는 바이러스 외피 (env)에 대한 단백질을 코딩할 수 있으며; 또 다른 플라스미드 (패키징 플라스미드)는 캡시드 단백질 (예를 들어, gag 및 pol)과 역전사 효소 및/또는 인테그라제와 같은 효소를 코딩할 수 있으며; 추가 플라스미드 (전이 플라스미드)는 긴-말단 반복부 (게놈 통합을 위한) 및 psi-서열 (유전자를 바이러스에 패키징하는 신호를 표시함)에 의해 플랭킹되는 관심 유전자 (GOI)를 포함한다. 이들 플라스미드를 동시에 세포에 도입할 경우, 질환을 유발하는데 필요한 바이러스 유전자 없이 관심 유전자를 함유하는 바이러스가 생성될 것이다.
- [0307] 본 개시내용의 특정 측면에서, 본 발명의 렌티바이러스 벡터 (또는 입자)는 분할 시스템, 예를 들어 트랜스상보성 시스템(transcomplementation system) (벡터/패키징 시스템)에 의해, 렌티바이러스 벡터 게놈의 특정 구성요소를 포함하는 플라스미드로, 그리고 폴리펩티드 GAG, POL 및 외피 단백질(들)을 코딩하는 gag, pol 및 env 서열을 도중에(in trans) 제공하거나, 또는 레트로바이러스 입자의 형성을 가능하게 하기에 충분한 이들 폴리펩티드의 일부를 제공하는 적어도 하나의 다른 플라스미드로, 허용 세포(permissive cell) (예를 들어, 293T 세포)를 시험관내에서 형질감염시킴으로써 수득가능하다.
- [0308] 예로서, 숙주 세포는 a) 렌티바이러스 gag 및 pol 서열을 포함하는 패키징 플라스미드, b) 외피 단백질(들) (예컨대 VSV-G)을 코딩하는 유전자를 포함하는 제2 플라스미드 (외피 발현 플라스미드 또는 슈도타이핑(pseudotyping) env 플라스미드), c) 5' 및 3' LTR 서열, psi 캡시드화(encapsidation) 서열, 및 트랜스진을 포함하는 플라스미드 벡터, 및 d) 본원에 개시된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부)을 코딩하는 핵산 작제물을 포함하는 플라스미드 벡터로 형질감염된다. 일부 실시양태에서, 본원에 개시된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 코딩하는 핵산 작제물은 별도의 플라스미드 대신에 패키징 플라스미드 상에 있다. gag, pol 및 env cDNA를 코딩하는 핵산은 선행 기술 및 데이터베이스에서 이용 가능한 바이러스 유전자 서열로부터 통상적인 기술에 따라 유리하게 제조될 수 있다.
- [0309] 일부 실시양태에서, 렌티바이러스 벡터는 본원에 기재된 바와 같은 핵산 작제물을 포함한다. 일부 실시양태에서, 렌티바이러스 벡터는 본원에 기재된 바와 같은 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함한다.
- [0310] 플라스미드에 사용된 프로모터는 동일하거나 상이할 수 있다. 일부 실시양태에서, 플라스미드 트랜스상보성 시스템에서, 외피 플라스미드 및 플라스미드 벡터는 각각 코트 단백질의 gag 및 pol의 발현을 촉진하기 위해, 벡터 게놈의 mRNA 및 트랜스진은 동일하거나 상이할 수 있는 프로모터이다. 이러한 프로모터는 유비쿼터스 프로모터 또는 특정, 예를 들어, 바이러스 프로모터 CMV, TK, RSV LTR 프로모터 및 RNA 폴리머라제 III 프로모터 예컨대 U6 또는 H1 또는 env, gag 및 pol을 코딩하는 헬퍼 바이러스 (즉, 아데노바이러스, 바칼로바이러스, 헤르페스 바이러스)의 프로모터로부터 유리하게 선택될 수 있다.
- [0311] 본 개시내용의 렌티바이러스 벡터의 생산을 위해, 본원에 기재된 플라스미드를 숙주 세포에 도입할 수 있고 바이러스를 생산하고 수확할 수 있다. 적합한 세포는 진핵 및 원핵 세포 및/또는 세포주를 포함하나 이에 제한되지는 않는다. 이러한 세포 또는 이러한 세포로부터 생성된 세포주의 비제한적 예는 예를 들어, COS, CHO (예를 들어, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (예를 들어, HEK293-F, HEK293-H, HEK293-T), 및 perC6 세포뿐만 아니라 곤충 세포 예컨대 스포둠테라 푸기페르다 (Sf), 또는 진균 세포 예컨대 사카로마이세스, 피치아 및 스킴조사카로마이세스를 포함한다.
- [0312] 일단 숙주 세포가 플라스미드로 형질감염되고 본 발명의 렌티바이러스 벡터(또는 입자)가 생성되면, 본 발명의

렌티바이러스 벡터 (또는 입자)는 세포의 상청액으로부터 정제될 수 있다. 농도를 향상시키기 위한 렌티바이러스 벡터의 정제는 임의의 적합한 방법에 의해, 예컨대 밀도 구배 정제 (예를 들어, 염화세슘(CsCl)), 크로마토그래피 기술 (예를 들어, 컬럼 또는 배치 크로마토그래피)에 의해, 또는 초원심분리에 의해 완수될 수 있다. 예를 들어, 본 발명의 벡터는 2개 또는 3개의 CsCl 밀도 구배 정제 단계를 거칠 수 있다. 벡터는 바람직하게는 세포를 용해시키고, 용해물을 크로마토그래피 수지에 적용하고, 크로마토그래피 수지로부터 바이러스를 용리하고, 본 발명의 렌티바이러스 벡터를 함유하는 분획을 수집하는 것을 포함하는 방법을 사용하여 감염된 세포로부터 정제된다.

[0313] 렌티바이러스 벡터의 전달 방법이 기재되어 있다. 예를 들어, 문헌 [Vargas et al. J. Trans. Med. 14(288):1-15 (2016); Mali et al. Nat. Methods 10(10):957-963; Mátés et al. Nature Genetics 41(6):753-761 (2009); Skipper et al. 20(92):1-23 (2013)]을 참조한다.

[0314] 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서) 또는 이를 코딩하는 핵산 작제물을 포함하는 렌티바이러스 벡터를 임의의 경로에 의해 대상체에게 투여할 수 있다. 일부 실시양태에서, 본 개시내용의 렌티바이러스 벡터를 생체내 또는 생체외에서 대상체의 세포로 전달할 수 있다.

[0315] 일부 실시양태에서, 본 개시내용의 렌티바이러스 벡터는 생체내에서 전달될 수 있다. 일부 실시양태에서, 본 개시내용의 핵산 작제물에 의해 코딩된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 렌티바이러스 벡터는 관심 유전자를 전달하고/하거나 대상체의 DNA에 있는 유전적 결함을 표적으로 하는데 사용될 수 있다. 일부 실시양태에서, 렌티바이러스 벡터는 대상체에게 비경구적으로, 바람직하게는 혈관내 (정맥내 포함)로 투여된다. 비경구적으로 투여하는 경우, 벡터는 멸균 수용액 또는 분산액과 같은 주사에 적합한 제약 비히클에 제공되는 것이 바람직하다.

[0316] 일부 실시양태에서, 본 개시내용의 렌티바이러스 벡터는 생체외에서 사용될 수 있다.

[0317] 일부 실시양태에서, 본 개시내용의 핵산 작제물에 의해 코딩된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 렌티바이러스 벡터는 관심 유전자를 전달하고/하거나 대상체의 DNA에 있는 유전적 결함을 표적으로 하는데 사용될 수 있다. 일부 실시양태에서, 세포는 대상체로부터 제거되고 본 개시내용의 핵산 작제물에 의해 코딩된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 렌티바이러스 벡터를 생체외에서 세포에 투여되어 세포의 DNA를 변형시킨다. 변형된 DNA를 운반하는 세포는 확장되어 대상체에게 다시 주입된다. 특정 실시양태에서, 본 개시내용의 핵산 작제물에 의해 코딩된 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 렌티바이러스 벡터는 키메라 항원 수용체(Chimeric Antigen Receptor) (CAR) T-세포 요법에 사용되어 종양 항원에 특이적인 CAR을 발현하도록 환자의 자가 T-세포를 유전적으로 변형시킨다. 추가 실시양태에서, 변형된 CAR-T 세포는 생체외에서 확장되고 다시 환자에게 재주입된다. 일부 실시양태에서, 변형된 T 세포는 암 세포를 보다 특이적으로 표적화한다. 항체 요법과 달리 CAR-T 세포는 생체내에서 복제할 수 있어 장기간 지속된다.

[0318] 본 개시내용의 렌티바이러스 벡터 또는 본 개시내용의 렌티바이러스 벡터를 사용하여 생체외에서 변형된 세포의 투여 후, 대상체를 모니터링하여 트랜스진의 발현을 검출할 수 있다. 치료 용량 및 지속기간은 치료할 병태 또는 질환에 따라 개별적으로 결정된다. 본 발명의 벡터 내 관심 유전자의 투여에 의해 생성된 유전자 발현을 기준으로 하여 다양한 상태 또는 질환을 치료할 수 있다. 본 발명의 방법을 사용하여 전달되는 벡터의 투여량은 숙주 및 사용되는 벡터에 의한 원하는 반응에 따라 달라질 것이다.

[0319] 일부 유전자 요법 적용에서, 유전자 요법 벡터는 특정한 조직 유형에 대해 고도의 특이성을 가지고 전달되는 것이 바람직하다. 따라서, 바이러스의 외부 표면에 바이러스 코트 단백질과 융합 단백질로서 리간드를 발현함으로써 바이러스 벡터가 주어진 세포 유형에 대한 특이성을 갖도록 변형될 수 있다. 리간드는 관심 세포 유형에 존재하는 것으로 공지된 수용체에 대한 친화도를 갖도록 선택된다.

[0320] 본 개시내용의 특정 측면은 유기체의 게놈 DNA에 외인성 핵산 서열을 삽입하는 방법으로서, 유기체의 게놈에서 특정 게놈 DNA 서열을 확인하는 단계; 본 개시내용의 핵산 작제물을 포함하는 렌티바이러스 입자를 유기체에 투여하여 특정 게놈 DNA 서열에 결합하고 외인성 핵산을 게놈 DNA에 삽입하는 단계를 포함하며; 여기서 외인성 핵산은 특정 게놈 DNA 서열에 통합되는, 방법에 관한 것이다.

[0321] 본 개시내용의 특정 측면은 외인성 핵산 서열의 단일 카피 또는 다수 카피의 세포 내로의 제어된 부위-특이적 통합을 위한 방법으로서, a) 핵산 작제물, 벡터, 또는 본 개시내용의 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 세포에 전달하는 단계, 및 b) 외인성 핵산을 세포에 전달하는 단계를 포함하

며; 여기서 세포 계놈에서 특정 계놈 DNA 서열에 대한 제1 및 제2 단백질(별도로 또는 상기에 기재된 융합 단백질의 일부로서)의 결합은 세포 계놈으로의 계놈의 절단 및 외인성 핵산의 하나 이상의 카피의 통합을 발생시키는, 방법에 관한 것이다. 일부 측면에서, 세포로의 전달은 렌티바이러스 입자에 의해 이루어진다.

[0322] 통합 부위에 대해 시험하고 직접 통합을 위한 최상의 기구(best machinery)를 스크리닝하기 위해 몇몇 전략을 사용할 수 있다.

[0323] 본원에 개시된 변형된 트랜스포존의 분석을 위해, 프로모터, GFP 코딩 서열의 절반 및 계놈에서 표적화된 삽입 부위의 스플라이스 부위 공여자 하류를 가진 리포터 세포주가 사용될 수 있다. 예를 들어, 렌티바이러스 페이로드는 융합 인테그라제 변이체에 이어서 역 스플라이스 부위 수용체(inverted splice site acceptor) 및 GFP의 나머지 절반을 가질 수 있다. GFP의 발현은 직접 삽입이 일어나고 삽입 부위 및 통합된 페이로드로부터 생성된 mRNA를 함유하는 GFP의 스플라이싱이 전체 GFP CDS를 생기게 할 때 발생할 것이다.

[0324] VPR 트랜스상보성 시스템은 또한 통합 돌연변이를 스크리닝하고 비교하는 데 사용할 수 있다. 트랜스상보성 시스템은 융합 인테그라제 변이체를 함유하는 렌티바이러스 페이로드의 표적 삽입에 사용될 수 있으며, 이는 발현되고 입자에 로딩될(loaded) 때 VPR 융합을 사용하여 바이러스 입자에 로딩될 그 자체 통합을 촉진한다. 이는 입자 생산에 사용되는 패키징 벡터에 코딩된 통합 결합 IN을 도중에 보완한다. IC, 또는 FISH 프로브를 포함하여 통합 매핑에 사용할 수 있는 다른 방법. 표적 삽입은 또한 TCRA 또는 RFP 표적 파괴(targeted disruption), 또는 표적 스플라이스 부위 통합에 의한 GFP 활성화에 의해 스크리닝될 수 있다.

[0325] 염색질의 삽입 및 표적 영역의 공동-염색에 대한 FISH 접근법의 경우, Hek293T 계놈에서 관심 트랜스포존 유전자를 국부화하기 위한 형광 계내(in situ) 하이브리드화를 수행할 수 있다. Hek293T는 1) GOI-트랜스포존 2) 프로그래밍 가능한 트랜스포사제 및 3) gRNA를 PPP1R12로 형질감염시킬 수 있다. 프로브는 PPP1R12 유전자, CD46 유전자 (음성 대조군으로서) 및 GOI를 표적으로 하여 설계되었으며, PCR 증폭 DNA로부터 Nick Translation Mix(Nick Translation Mix) (시그마(Sigma))로 합성할 수 있다.

[0326] 일부 실시양태에서, 본원에 개시된 바와 같은 변형된 트랜스포사제를 포함하는 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 야생형 트랜스포사제 (또는 상응하는 야생형 트랜스포사제를 함유하는 융합 단백질)와 비교하여, 예를 들어 진트랩 검정(Genetrap assay)에 의해 결정된 바와 같이, 계놈으로의 외인성 핵산의 삽입의 특이성을 개선한다. 일부 실시양태에서, HEK293T 세포, 또는 임의의 다른 허용가능한 세포는 하기 플라스미드 또는 페이로드를 가진 렌티바이러스 입자로 형질감염 또는 형질도입된다: (i) DNA의 특정 영역을 표적화하는 gRNA를 포함하는 플라스미드, (ii) 다음을 포함하는 플라스미드 제2 단백질이 변형된 트랜스포사제인 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 코딩하는 본 개시내용의 핵산 작제물, 및 (iii) 리포터 단백질을 코딩하는 핵산 서열을 포함하는 진트랩(genetrap) 플라스미드, 예를 들어 프로모터가 없는 GFP. 일부 실시양태에서, 진트랩 플라스미드는 역 반복부를 가진 트랜스포존을 추가로 포함한다.

[0327] 일부 실시양태에서, GFP 삽입을 함유하는 세포의 퍼센트 유동 세포측정법에 의해 결정될 수 있다. 일부 실시양태에서, 제2 단백질이 변형된 트랜스포사제인 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 상응하는 야생형 단백질과 비교하여 GFP의 삽입을 함유하는 세포의 퍼센트를 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 또는 적어도 30% 증가시킨다. 일부 실시양태에서, 제2 단백질이 변형된 트랜스포사제인 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 GFP 삽입을 함유하는 세포의 퍼센트를 약 15-30% 증가시킨다.

[0328] 일부 실시양태에서, 표적 부위에서의 삽입 퍼센트 및 표적 부위에서의 커버리지(coverage) 퍼센트 (삽입 부위당 관독 수)는 계놈 DNA 추출 및 바이러스 LTR에 특이적인 올리고뉴클레오타이드를 사용한 표적화된 시퀀싱(targeted sequencing)에 의해 결정될 수 있다. 일부 실시양태에서, 제2 단백질이 변형된 트랜스포사제인 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 표적 부위에서의 삽입 퍼센트를 상응하는 야생형 단백질과 비교하여 적어도 10배, 적어도 20배, 적어도 30배, 적어도 40배, 적어도 50배, 적어도 60배, 적어도 70배, 적어도 80배, 적어도 90배, 또는 적어도 100배 증가시킨다. 일부 실시양태에서, 표적 부위에서의 삽입 퍼센트는 약 10-100배 증가된다. 일부 실시양태에서, 제2 단백질이 변형된 트랜스포사제인 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 상응하는 야생형 단백질과 비교하여 표적 부위에서의 커버리지 퍼센트 (삽입 부위당 관독 수)를 적어도 10배, 적어도 20배, 적어도 30배, 적어도 40배, 적어도 50배, 적어도 60배, 적어도 70배, 적어도 80배, 적어도 90배, 적어도 100배, 적어도 110배, 적어도 120배, 적어도 130배, 적어도 140배, 적어도 150배, 적어도 160배, 적어도 170배, 적어도 180배, 적어도 190배, 또는 적어도 200배 증가시킨다.

배 증가시킨다. 일부 실시양태에서, 표적 부위에서의 커버리지의 퍼센트 (삽입 부위당 관독 수)는 적어도 100 배이다.

[0329] 일부 실시양태에서, 표적 부위에서의 삽입 퍼센트 및 표적 부위에서의 커버리지 퍼센트 (삽입 부위당 관독 수)는 게놈 DNA 추출 및 바이러스 삽입된 LTR에 특이적인 올리고뉴클레오타이드를 사용한 표적화된 시퀀싱에 의해 결정될 수 있다.

[0330] 본 개시내용의 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 포함하는 렌티바이러스 벡터의 가능한 적용은 유전자 요법, 즉 임의의 포유동물 세포, 특히 인간 세포에서의 유전자 전달을 포함한다. 이는 분할 세포 또는 정지 세포(quiescent cell)일 수 있으며, 세포는 간, 췌장, 근육, 심장 등과 같은 중추 기관 또는 말초 기관에 속한다. 유전자 요법은 단백질, 예를 들어 신경영양 인자, 효소, 전사 인자, 수용체 등의 발현을 허용할 수 있다. 본 발명에 따른 렌티바이러스 벡터는 또한 연구 목적에 특히 적합할 수 있다.

[0331] 일부 실시양태에서, 핵산 작제물, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), 및/또는 본 개시내용의 렌티바이러스 벡터는 질환을 치료하기 위해 대상체에게 투여된다. 일부 실시양태에서, 질환은 유전자 요법으로부터 이익을 수득할 수 있는 유전적 장애이다.

[0332] 일부 실시양태에서, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), 또는 이를 코딩하는 핵산 작제물, 또는 이후 개시되는 바와 같은 키트 또는 조성물 또는 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서) 또는 본 개시내용에 따른 핵산 작제물을 포함하는 렌티바이러스 벡터가 의약으로서 사용될 수 있다.

[0333] 본 개시내용에 따른 렌티바이러스 벡터는 대상체에서 유전 질환을 치료하는데 특히 적합할 수 있다.

[0334] 본 발명은 또한

[0335] (i) RNA 가이드된 뉴클레아제 또는 징크 핑거 뉴클레아제,

[0336] (ii) 트랜스포사제,

[0337] (iii) 가이드 RNA, 및

[0338] (iv) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산

[0339] 을 포함하는 조성물로서,

[0340] 여기서 상기 트랜스포사제는 서열번호: 9의 과활성 Piggybac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 Piggybac인, 조성물에 관한 것이다.

[0341] 한 실시양태에서, 변형된 과활성 PiggyBac 돌연변이는 아미노산 치환 R372A/K375A/D450N을 포함한다.

[0342] 바람직한 실시양태에서, 변형된 과활성 PiggyBac 돌연변이는 아미노산 치환 R372A/K375A/D450N을 포함하지 않는다.

[0343] 일부 실시양태에서, 변형된 과활성 PiggyBac 돌연변이는 S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L, R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G, R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L, V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P, G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N, T43I/R372A/K375A/A411T/D450N, G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G의 하기 아미노산 치환 또는 아미노산 치환의 조합을 포함하며, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응한다.

[0344] 바람직한 실시양태에서, 변형된 과활성 PiggyBac 돌연변이는 R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G,

R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L,
 R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G,
 R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L,
 G325A/N347S/K375A/D450N/S573A/M589V/S592G, S230N/R277A/N347S/K375A/D450N,
 G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G의 하
 기 아미노산 치환 또는 아미노산 치환의 조합을 포함하며, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의
 아미노산 번호에 상응한다.

[0345] 일부 실시양태에서 RNA 가이드된 뉴클레아제는 Cas9 단백질이다. 일부 실시양태에서 RNA 가이드된 뉴클레아제는 SpCas9 단백질이다. 일부 실시양태에서 RNA 가이드된 뉴클레아제는 SaCas9 단백질이다.

[0346] 본 발명은 또한 다음을 코딩하는 핵산을 포함하는 조성물에 관한 것이다:

[0347] (i) 상기에 기재된 바와 같은, RNA 가이드된 뉴클레아제 또는 징크 핑거 뉴클레아제,

[0348] (ii) 상기에 기재된 바와 같은, 트랜스포사제,

[0349] (iii) 가이드 RNA, 및

[0350] (iv) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산.

[0351] 일부 실시양태에서, 조성물의 핵산은 적합한 발현 벡터를 통해 세포에서 발현된다. 본원에 사용된 바와 같이, 용어 "발현 벡터"는 발현될 뉴클레오타이드 서열에 작동가능하게 연결된 발현 제어 서열을 포함하는 폴리뉴클레오타이드를 포함하는 벡터를 지칭한다. 발현 벡터는 발현을 위한 충분한 시스-작용 요소를 포함하며; 발현을 위한 다른 요소는 숙주 세포에 의해 또는 시험관내 발현 시스템에서 공급될 수 있다. 발현 벡터는 재조합 폴리뉴클레오타이드를 포함하는 코스미드, 플라스미드 (예를 들어, 네이키드(naked) 또는 리포솜에 함유됨) 및 바이러스 (예를 들어, 렌티바이러스, 레트로바이러스, 아데노바이러스 및 아데노-연관 바이러스)를 포함하는 관련 기술분야에 공지된 모든 것들을 포함한다.

[0352] 바람직한 실시양태에서, 2개의 핵산은 동일한 세포 또는 세포 집단에서 공동-발현된다. 일부 실시양태에서, 2개의 핵산은 동시에 공동-발현된다. 또 다른 실시양태에서, RNA 가이드된 뉴클레아제 또는 징크 핑거 뉴클레아제를 코딩하는 핵산이 먼저 발현된다. 또 다른 실시양태에서, 트랜스포사제를 코딩하는 핵산이 먼저 발현된다.

[0353] 본 발명은 추가로

[0354] (i) 본원에 개시된 바와 같은 RNA-가이드된 뉴클레아제 및 트랜스포사제를 포함하는 융합 단백질, 또는 이를 코딩하는 핵산,

[0355] (ii) 본원에 개시된 바와 같은 트랜스포사제, 또는 이를 코딩하는 핵산,

[0356] (iii) 가이드 RNA, 및

[0357] (iv) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산

[0358] 을 포함하는 조성물에 관한 것이다.

[0359] 본 발명은 추가로

[0360] (i) 본원에 개시된 바와 같은 RNA-가이드된 뉴클레아제 및 트랜스포사제를 포함하는 제1 융합 단백질, 또는 이를 코딩하는 핵산,

[0361] (ii) 본원에 개시된 바와 같은 적어도 하나의 특이적 RNA 서열에 결합하도록 조작된 RNA 결합 단백질, 및 트랜스포사제를 포함하는 제2 융합 단백질, 또는 이를 코딩하는 핵산,

[0362] (iii) 가이드 RNA, 및

[0363] (iv) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산을 포함하는 조성물에 관한 것이다.

[0364] 본 개시내용은 또한 본원에 기재된 개시된 방법을 실시하기 위한 조성물을 제공한다. 일부 실시양태에서, 조성물은 본 개시내용에 정의된 바와 같은 핵산 작제물 또는 벡터, 및 패키징 벡터에 함유되거나 결합된, 게놈에 삽입하기 위한 외인성 핵산을 코딩하는 폴리뉴클레오타이드 서열을 포함한다.

[0365] 본 개시내용은 추가로

- [0366] (i) 본원에 개시된 바와 같은 RNA-가이드된 뉴클레아제 및 트랜스포사제를 포함하는 융합 단백질 또는 상기 융합 단백질을 코딩하는 핵산,
- [0367] (ii) 가이드 RNA, 및
- [0368] (iii) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산
- [0369] 을 포함하는 조성물에 관한 것이다.
- [0370] 구체적 실시양태에서, 상기 관심 핵산 또는 유전자는 전형적으로 크기가 5 kb 내지 25 kb, 그리고 보다 바람직하게는 8 kb 내지 20 kb인 큰 DNA 단편이다.
- [0371] 또한, 본원에 기재된 바와 같이, 개시된 방법을 실시하기 위한 키트가 본 개시내용에 의해 제공된다. 키트는 본원에 기재된 바와 같은 핵산 작제물 또는 융합 단백질을 함유할 수 있다. 일부 측면에서, 키트는 본원에 기재된 바와 같은 핵산 작제물 또는 융합 단백질을 함유하는 렌티바이러스 입자를 함유할 수 있다.
- [0372] 대상 키트는 대상 방법을 실시하기 위해 키트의 구성요소를 사용하기 위한 설명서를 추가로 포함할 수 있다. 대상 방법을 실행하기 위한 설명서는 일반적으로 적합한 기록 매질에 기록된다. 예를 들어, 설명서는 종이 또는 플라스틱 등과 같은 기재 상에 인쇄될 수 있다. 이와 같이, 설명서는 패키지 삽입물로서 키트에, 키트 또는 그 구성요소의 용기의 라벨링 (즉, 패키징 또는 서브패키징과 연관됨) 등에 존재할 수 있다. 다른 실시양태에서, 설명서는 적합한 컴퓨터 판독가능 저장 매체에 존재하는 전자 저장 데이터 파일로서 존재한다. 또 다른 실시양태에서, 실제 설명서는 키트에 존재하지 않지만, 예를 들어 인터넷을 통해 원격 공급원으로부터 설명서를 얻기 위한 수단이 제공된다. 이 실시양태의 예는 설명서를 볼 수 있고/거나 설명서를 다운로드할 수 있는 웹 주소를 포함하는 키트이다. 설명서와 마찬가지로, 설명서를 얻기 위한 이 수단은 적합한 기관에 기록된다.
- [0373] 본 개시내용은 전형적으로 키트로서,
- [0374] (i) 본원에 정의된 바와 같은 제1 융합 단백질, 또는 상기 제1 융합 단백질을 코딩하는 핵산, 그리고 여기서 상기 제1 융합 단백질은 제1 가이드된 RNA 니카제 Cas9의 아미노산 서열, 전형적으로 변형된 과활성 Piggybac에 융합된 서열번호: 70의 SpCas9 니카제를 포함함, 및,
- [0375] (ii) 제1 가이드된 RNA 핵산
- [0376] 을 포함하는 제1 조성물,
- [0377] (iii) 본원에 정의된 바와 같은 제2 융합 단백질, 또는 상기 제2 융합 단백질을 코딩하는 핵산, 그리고 여기서 상기 제2 융합 단백질은 제2 가이드된 RNA 니카제 Cas9의 아미노산 서열, 전형적으로 변형된 과활성 Piggybac에 융합된 서열번호: 76의 SaCas9 니카제를 포함함,
- [0378] (iv) 제2 가이드된 RNA 핵산
- [0379] 을 포함하는 제2 조성물,
- [0380] 임의로, 게놈에 삽입하기 위한 핵산, 예를 들어 크기가 5 kb 내지 25 kb, 그리고 보다 바람직하게는 8 kb 내지 20 kb인 핵산
- [0381] 을 포함하며;
- [0382] 여기서 상기 제1 및 제2 융합 단백질은 이중이량체화를 형성할 수 있고, 게놈 DNA 영역의 인접 부위에서 상기 제1 및 제2 가이드된 RNA에 의해 결정된 이중 커트를 만들 수 있고, 임의로 인접 부위 사이에 상기 핵산을 삽입할 수 있는, 키트에 관한 것이다.
- [0383] 구체적 실시양태에서, 조성물 또는 키트는 미니씨클, 플라스미드 또는 바이러스 벡터, 특히 비통합 바이러스 벡터, 예를 들어 비통합 렌티바이러스 벡터에 외인성 핵산을 포함한다.
- [0384] 구체적 실시양태에서, 본원에 개시된 바와 같은 조성물 또는 키트는 나노입자에 포함된다.
- [0385] 구체적 실시양태에서, 상기 조성물은
- [0386] (i) 본원에 개시된 바와 같은 융합 단백질을 코딩하는 핵산 작제물,
- [0387] (ii) 가이드 RNA, 및

- [0388] (iii) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산
- [0389] 을 포함하는 핵산 조성물이다.
- [0390] 구체적 실시양태에서, 상기 키트는
- [0391] ● (i) 본원에 개시된 바와 같은 제1 융합 단백질을 코딩하는 핵산 작제물, 그리고 여기서 상기 제1 융합 단백질은 제1 가이드된 RNA 니카제 Cas9의 아미노산 서열, 전형적으로 변형된 과활성 Piggybac에 융합된 서열번호: 70의 SpCas9 니카제를 포함함, 및,
- [0392] (ii) 제1 가이드된 RNA 핵산
- [0393] 을 포함하는 제1 조성물,
- [0394] ● (iii) 본원에 개시된 바와 같은 상기 제2 융합 단백질을 코딩하는 핵산 작제물, 그리고 여기서 상기 제2 융합 단백질은 제2 가이드된 RNA 니카제 Cas9의 아미노산 서열, 전형적으로 변형된 과활성 Piggybac에 융합된 서열번호: 76의 SaCas9 니카제를 포함함, 및,
- [0395] (iv) 제2 가이드된 RNA 핵산
- [0396] 을 포함하는 제2 조성물을 포함한다,
- [0397] 구체적 실시양태에서, 상기 키트 또는 조성물은 특히 인간에서 장애를 치료하는데, 예를 들어 치료를 필요한 인간 대상체에서 유전적 결함을 치료하기 위한 약물로서 사용하기 위한 것이다.
- [0398] 일부 실시양태에서, 핵산 작제물은 RNA, DNA 또는 단백질의 형태이고, 외인성 핵산을 코딩하는 폴리뉴클레오티드 서열은 전달 방법에 따라 RNA 또는 DNA의 형태이다. 특히, 외인성 핵산을 코딩하는 폴리뉴클레오티드 서열은 RNA 형태이다.
- [0399] 일부 실시양태에서, 조성물 또는 키트는 바이러스가 없고(viral-free) 패키징 벡터는 나노입자, 예를 들어 중합체 또는 지질 나노입자이다. 패키징 벡터는 또한 조성물의 요소에 결합되는 담체일 수 있다. 일부 실시양태에서, 조성물은 바이러스 벡터, 특히 렌티바이러스 입자에 함유된다.
- [0400] 일부 실시양태에서, 조성물 또는 키트는 (a) RNA 형태로 본원에 기재된 (예를 들어, Cas9 및 트랜스포사제를 포함함) 핵산 작제물, (b) 필요한 경우 가이드 RNA (예를 들어, 별도의 직계(lineal) 단일 가닥 RNA 분자), 및 (c) 패키징 벡터 내에 함유되거나 그에 결합된, DNA 형태 (예를 들어, 벡터)로 삽입을 위해 외인성 유전자를 포함하는 폴리뉴클레오티드를 포함한다.
- [0401] 일부 실시양태에서, 조성물은 (a) 단백질 형태로 본원에 기재된 (예를 들어, Cas9 및 트랜스포사제를 포함함) 융합 단백질, (b) 필요한 경우 가이드 RNA (예를 들어, 별도의 직계 단일 가닥 RNA 분자로서), 여기서 융합 단백질 및 가이드 RNA는 리보핵 단백질 복합체 (RNP)를 형성함, 및 (c) 패키징 벡터 내에 함유되거나 그에 결합된, DNA 형태 (예를 들어, 벡터)로 삽입을 위해 외인성 유전자를 포함하는 폴리뉴클레오티드를 포함한다.
- [0402] 일부 실시양태에서, 조성물은 (a) DNA 형태로 본원에 기재된 (예를 들어, Cas9 및 트랜스포사제를 포함함) 핵산 작제물, (b) 필요한 경우 가이드 RNA (예를 들어, 별도의 직계 RNA 분자로서 또는 벡터 내 DNA로서), 및 (c) 패키징 벡터 내에 함유되거나 그에 결합된, DNA 형태 (예를 들어, 벡터)로 삽입을 위해 외인성 유전자를 포함하는 폴리뉴클레오티드를 포함한다.
- [0403] 일부 실시양태에서, 조성물은 (a) 단백질 형태로 본원에 기재된 (예를 들어, Cas9 및 트랜스포사제를 포함함) 융합 단백질, (b) 필요한 경우 가이드 RNA (예를 들어, 융합 단백질과 복합체를 형성하는 별도의 RNA 분자로서), 및 (c) 패키징 벡터에 함유되거나 결합된, 삽입을 위한 외인성 유전자를 포함하는 폴리뉴클레오티드를 포함한다. 특정한 실시양태에서, 패키징 벡터는 렌티바이러스 입자이다. 일부 실시양태에서, (a) 융합 단백질은 gag-pol 또는 VPR (바이러스 단백질 R)에 의해 렌티바이러스 캡시드에 결합된다. 일부 실시양태에서, (c) 폴리뉴클레오티드는 인테그라제의 페이로드로서 RNA 형태이다.
- [0404] 특정한 실시양태에서, ZFP가 사용될 때, (b) 가이드 RNA는 필요하지 않을 수 있다.
- [0405] 본 발명은 추가로
- [0406] (i) 적어도 하나의 특정 RNA 서열에 결합하도록 조작된 RNA 결합 단백질, 외인성 핵산의 게놈으로의 삽입을 가능하게 하는 DNA 결합 단백질, 및 제1 및 제2 단백질을 연결하는 링커를 포함하는 융합 단백질,

- [0407] (ii) Cas9 단백질, 및
- [0408] (iii) 융합 단백질 결합을 위한 상기 적어도 하나의 특정 RNA 서열을 포함하는 가이드 RNA,
- [0409] (iv) 관심 핵산 또는 유전자, 예를 들어 게놈에 삽입하기 위한 외인성 핵산
- [0410] 을 포함하며,
- [0411] 여기서 DNA 결합 단백질은 본 개시내용의 변형된 트랜스포사제, 전형적으로 변형되지 않은 과활성 PiggyBac과 비교하여 절제 활성을 증가시키기 위한 하나 이상의 아미노산 돌연변이, 및 서열번호: 9에 따른 변형되지 않은 과활성 Piggybac과 비교하여 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 Piggybac인 조성물에 관한 것이다.
- [0412] 일부 실시양태에서, RNA 결합 단백질은 MS2 박테리오파지 코트 단백질 (MCP)이다.
- [0413] 일부 실시양태에서, 융합 단백질의 MCP에 의해 인식되는 적어도 하나의 RNA 서열은 테트라루프이다. 본원에 사용된 바와 같이, 용어 "테트라루프"는 용어 "스텝 루프(stem loop)" 및 "헤어핀 루프(hairpin loop)"와 상호교환적으로 사용된다. 일부 실시양태에서, 적어도 하나의 RNA 테트라루프는 MS2 RNA 테트라루프 결합 서열이다.
- [0414] 일부 실시양태에서 가이드 RNA는 적어도 하나의 MS2 RNA 테트라루프 결합 서열을 포함한다. 일부 실시양태에서, gRNA는 하나 초과 MS2 RNA 테트라루프 결합 서열을 포함한다. 본원에 사용된 바와 같이, 용어 "하나 초과"는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 또는 그 초과를 의미한다.
- [0415] 다양한 실시양태가 청구범위에 기재되어 있다. 일부 추가 실시양태가 이하에 개시된다:
- [0416] E1. (i) RNA 가이드된 뉴클레아제 또는 징크 핑거 뉴클레아제로 이루어진 제1 단백질,
- [0417] (ii) 트랜스포사제로 구성된 제2 단백질, 및
- [0418] (iii) 임의로, 제1 및 제2 단백질을 연결하는 링커
- [0419] 를 포함하는 융합 단백질로서,
- [0420] 여기서 상기 트랜스포사제는 서열번호: 9의 과활성 Piggybac과 비교하여 하나 이상의 아미노산 돌연변이를 포함하는, 변형된 과활성 PiggyBac이고,
- [0421] 상기 제1 단백질은 직접적으로, 또는 링커를 통해 간접적으로 제2 단백질의 C-말단 단부에 융합되는, 융합 단백질.
- [0422] E2. 실시양태 1에 있어서, 상기 트랜스포사제가, 변형되지 않은 과활성 PiggyBac과 비교하여 절제 활성을 증가시키기 위한 하나 이상의 아미노산 돌연변이, 및 변형되지 않은 과활성 Piggybac과 비교하여 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac인, 융합 단백질.
- [0423] E3. 실시양태 2에 있어서, 절제 활성을 증가시키기 위한 상기 하나 이상의 아미노산 돌연변이가 아미노산 위치 번호 [194-200], [214-222], [434-442] 또는 [446-456]에 의해 정의된 영역 내에 아미노산 돌연변이, 예를 들어 위치 D198, D201, R202, M212, 또는 S213에서의 아미노산 치환 중에서 선택되며; 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하는, 융합 단백질.
- [0424] E4. 실시양태 1-3 중 어느 하나에 있어서, 상기 하나 이상의 아미노산 돌연변이가 M194 또는 D450의 위치에서 절제 활성을 증가시키는 아미노산 치환 중에서 선택되고, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하며, 바람직하게는 아미노산 치환 M194V 및/또는 D450N 중에서 선택되는, 융합 단백질.
- [0425] E5. 실시양태 1-4 중 어느 하나에 있어서, 상기 하나 이상의 아미노산 돌연변이가 위치 R372, K375, R376, E377, 및/또는 E380에서 DNA 결합 활성을 감소시키는 아미노산 치환 중에서 선택되고, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하며, 바람직하게는 아미노산 치환 R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택되는, 융합 단백질.
- [0426] E6. 실시양태 1-4 중 어느 하나에 있어서, 변형된 과활성 Piggybac이 위치 D450에서 절제 활성을 증가시키기 위한 적어도 하나의 아미노산 치환, 및 위치 R372 및 K375에서 DNA 결합 활성을 감소시키기 위한 적어도 2개의 아미노산 치환을 포함하고, 바람직하게는 과활성 Piggybac의 상기 변형된 트랜스포사제가 삼중 돌연변이 D450N, R372A 및 K375A를 포함하고, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호

에 상응하는, 융합 단백질.

- [0427] E7. 실시양태 1 내지 6 중 어느 하나에 있어서, 변형된 과활성 Piggybac이 아미노산 위치 번호 [158-169], 예를 들어 A166S에 의해 정의되는 영역에 적어도 하나의 돌연변이; 및/또는 위치 Y527, R518, K525, N463에서 적어도 하나의 돌연변이를 추가로 포함하는, 융합 단백질.
- [0428] E8. 실시양태 1 내지 7 중 어느 하나에 있어서, 상기 변형된 과활성 Piggybac이 서열번호: 1의 변형된 과활성 Piggybac과 적어도 85%, 적어도 90%, 적어도 95% 동일성, 또는 100% 동일성을 갖는 아미노산 서열을 포함하는, 융합 단백질.
- [0429] E9. 실시양태 1 내지 8 중 어느 하나에 있어서, 상기 변형된 과활성 Piggybac이 하나 이상의 아미노산 치환, 전형적으로 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 이하의 아미노산 치환을 가진 서열번호: 1의 과활성 Piggybac의 변이체인, 융합 단백질.
- [0430] E10. 실시양태 1 내지 9 중 어느 하나에 있어서, 상기 변형된 과활성 Piggybac이 위치 번호 245, 268, 275, 277, 287, 290, 315, 325, 341, 346, 347, 350, 351, 356, 357, 388, 409, 412, 432, 447, 460, 461, 465, 517, 560, 564, 571, 573, 576, 586, 587, 589, 592, 및/또는 594에서 하나 이상의 아미노산 치환을 추가로 포함하며, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응하는, 융합 단백질.
- [0431] E11. 실시양태 10에 있어서, R245A, D268N, R275A, R277A, K287A, K290A, K287A/K290A, R315A, G325A, R341A, D346N, N347A, N347S, T350A, S351E, S351P, S351A, K356E, N357A, R388A, K409A, K412A, K432A, D447A, D447N, D450N, R460A, K461A, W465A, S517A, T560A, S564P, S571N, S573A, K576A, H586A, I587A, M589V, S592G, 또는 F594L, D450N/R372A/K375A, R275A/R277A, K409A/K412A, R460A/K461A, R275A/R277A/N347S/K375A/T560A/S573A/M589V/S592G 및 R245A/R275A/R277A/R372A/W465A의 하기 아미노산 치환 또는 아미노산 치환의 조합을 포함하며, 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응하는, 융합 단백질.
- [0432] E12. 실시양태 10에 있어서, 변형된 과활성 PiggyBac 돌연변이가 R372A/K375A/D450N, R372A/K375A/R376A/D450N, K375A/R376A/E377A/E380A/D450N, R372A/K375A/R376A/E377A/E380A/D450N, M194V, R376A, E377A, E380A M194V/R372A/K375A, S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L의 하기 아미노산 치환 또는 아미노산 치환의 조합을 포함하며; 위치 번호는 과활성 PiggyBac 서열 (서열번호: 9)의 아미노산 번호에 상응하며 전형적으로 상기 변형된 트랜스포사제가 서열번호: 1-8 및 10-18 중 임의의 것 중에서 선택된 아미노산 서열을 갖는, 융합 단백질.
- [0433] E13. 실시양태 1-12 중 어느 하나에 있어서, 링커가 XTEN 서열 또는 GGS 서열, 바람직하게는 XTEN 서열을 포함하는 펩티드 링커인, 융합 단백질.
- [0434] E14. 실시양태 1 내지 13 중 어느 하나에 있어서, 링커가 전형적으로 서열번호: 49, 51, 53, 55, 57, 59, 61 중 임의의 것 중에서 선택되는 3 내지 50개의 아미노산 길이를 갖는 펩티드 링커인, 융합 단백질.
- [0435] E15. 실시양태 1 내지 14 중 어느 하나에 있어서, 상기 제1 단백질이 활성 DNA 절단 도메인 및 가이드 RNA 결합 도메인을 포함하는 Cas9 단백질인, 융합 단백질.
- [0436] E16. 실시양태 1 내지 15 중 어느 하나에 있어서, 상기 제1 단백질이 활성 DNA 절단 도메인 및 가이드 RNA 결합 도메인을 포함하며 서열번호: 31의 스트렙토코커스 피로게네스 Cas9, 서열번호: 72의 스탕필로코커스 아우레우스 SaCas9, 서열번호: 74의 Cpf1, 서열번호: 29의 CjCas9, 서열번호: 70의 SpCas9, 서열번호: 75의 CasX, 또는 서열번호: 76의 SaCas9 닉카제에 대해 적어도 80%, 90%, 95%, 99% 또는 적어도 100% 동일성을 갖는 뉴클레아제 단백질인, 융합 단백질.
- [0437] E17. 실시양태 1 내지 16 중 어느 하나에 있어서, 상기 제1 단백질이 서열번호: 72의 SaCas9 또는 서열번호: 31의 스트렙토코커스 피로게네스 Cas9로 이루어진 군으로부터 선택되는 Cas9 단백질인, 융합 단백질.
- [0438] E18. 실시예 1 내지 17 중 어느 하나에 있어서,

- [0439] (i) RNA 가이드된 뉴클레아제 또는 니카제로 이루어진 제1 단백질,
- [0440] (ii) 제1 트랜스포사제로 이루어진 제2 단백질,
- [0441] (iii) 제2 트랜스포사제로 이루어진 제3 단백질 및,
- [0442] (iv) 임의로, 제1 및 제2 단백질과 제2 및 제3 단백질 사이의 펩티드 링커,
- [0443] 를 포함하며,
- [0444] 여기서 제1 및 제2 트랜스포사제는 예를 들어 실시양태 3-12 중 어느 하나에 정의된 바와 같은 변형된 Piggybac 트랜스포사제의 동일하거나 상이한 서열을 갖는, 삼중 융합 단백질인 융합 단백질.
- [0445] E19. 실시양태 1 내지 14 중 어느 하나에 있어서, 상기 제1 단백질이 서열번호: 33의 징크 핑거 단백질인 융합 단백질.
- [0446] E20. (i) 실시양태 1-19 중 어느 하나의 융합 단백질, 또는 상기 융합 단백질을 코딩하는 핵산,
- [0447] (ii) 가이드 RNA, 및
- [0448] (iii) 게놈에 삽입하기 위한 외인성 핵산
- [0449] 을 포함하는 조성물.
- [0450] E21. 실시양태 20에 있어서, 상기 외인성 핵산이 전형적으로 크기가 5 kb 내지 25 kb, 그리고 보다 바람직하게는 8 kb 내지 20 kb인 큰 DNA 단편인, 조성물.
- [0451] E22.
- [0452] (i) - 실시양태 1 및 18 중 어느 하나에 정의된 바와 같은 제1 융합 단백질, 또는 상기 제1 융합 단백질을 코딩하는 핵산, 그리고 여기서 상기 제1 융합 단백질은 제1 가이드된 RNA 니카제 Cas9의 아미노산 서열, 전형적으로 변형된 과활성 Piggybac에 융합된 서열번호: 70의 SpCas9 니카제를 포함함,
- [0453] - 제1 가이드된 RNA 핵산
- [0454] 을 포함하는 제1 조성물,
- [0455] (ii) - 실시양태 1 및 18 중 어느 하나에 정의된 바와 같은 제2 융합 단백질, 또는 상기 제2 융합 단백질을 코딩하는 핵산, 그리고 여기서 상기 제2 융합 단백질은 제2 가이드된 RNA 니카제 Cas9의 아미노산 서열, 전형적으로 변형된 과활성 Piggybac에 융합된 서열번호: 76의 SaCas9 니카제를 포함함,
- [0456] - 제2 가이드된 RNA 핵산
- [0457] 을 포함하는 제2 조성물,
- [0458] (iii) 임의로, 게놈에 삽입하기 위한 외인성 핵산, 예를 들어 크기가 5 kb 내지 25 kb, 그리고 보다 바람직하게는 8 kb 내지 20 kb인 외인성 핵산
- [0459] 을 포함하며;
- [0460] 여기서 상기 제1 및 제2 융합 단백질은 이중이량체화를 형성할 수 있고, 게놈 DNA 영역의 인접 부위에서 상기 제1 및 제2 가이드된 RNA에 의해 결정된 이중 커트를 만들 수 있고, 임의로 인접 부위 사이에 상기 외인성 핵산을 삽입할 수 있는, 키트.
- [0461] E23. 실시양태 20 내지 22에서, 상기 외인성 핵산이 미니씨클, 플라스미드 또는 바이러스 벡터, 특히 비통합 바이러스 벡터, 예를 들어 비통합 렌티바이러스 벡터에 포함되는, 조성물 또는 키트.
- [0462] E24. 실시양태 20-23에서, 조성물이 나노입자에 포함되는, 조성물 또는 키트.
- [0463] E25. 변형되지 않은 과활성 Piggybac과 비교하여 절제 활성을 증가시키기 위한 적어도 하나의 아미노산 돌연변이, 및/또는 변형되지 않은 과활성 Piggybac과 비교하여 DNA 결합 활성을 감소시키기 위한 적어도 하나의 아미노산 돌연변이를 포함하는, 변형된 과활성 Piggybac 트랜스포사제로서,
- [0464] 여기서 절제 활성을 증가시키기 위한 상기 적어도 하나의 돌연변이는 위치 194에서 M의 아미노산 치환, 전형적으로 M194V이고/이거나 여기서 DNA 결합 활성을 감소시키기 위한 적어도 하나의 아미노산 돌연변이는 위치

R376, E377, 및 E380, 전형적으로 R376A, E377A, 및/또는 E380A에서의 아미노산 치환 중에서 선택되는, 변형된 과활성 Piggybac 트랜스포사제.

- [0465] E26. 서열번호: 9의 변형되지 않은 과활성 Piggybac과 비교하여 하기 돌연변이 R372A/K375A/D450N, R372A/K375A/R376A/D450N, K375A/R376A/E377A/E380A/D450N, R372A/K375A/R376A/E377A/E380A/D450N, M194V, R376A, E377A, E380A, M194V/R372A/K375A, S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L, R245A/R275A/R277A/R372A/W465A/M589V, R275A/325A/R372A/T560A, N347A/D450N, N347S/D450N/T560A/S573A/F594L, R202K/R275A/N347S/R372A/D450N/T560A/F594L, R275A/N347S/K375A/D450N/S592G, R275A/N347S/R372A/D450N/T560A/F594L, R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L의 조합을 포함하는 변형된 과활성 Piggybac 트랜스포사제.
- [0466] E27. 실시양태 25 또는 26에 있어서, 하기로 이루어진 군으로부터 선택되는 하나 이상의 아미노산 돌연변이를 추가로 포함하는, 변형된 과활성 Piggybac 트랜스포사제: 245번 위치의 아미노산이 A이고, 275번 위치의 아미노산이 R 또는 A이고, 277번 위치의 아미노산이 R 또는 A이고, 325번 위치의 아미노산이 A 또는 G이고, 347번 위치의 아미노산이 N 또는 A이고, 351번 위치의 아미노산이 E, P 또는 A이고, 372번 위치의 아미노산이 R이고, 375번 위치의 아미노산이 A이고, 450번 위치의 아미노산이 D 또는 N이고, 465번 위치의 아미노산이 W 또는 A이고, 560번 위치의 아미노산이 T 또는 A이고, 564번 위치의 아미노산이 P 또는 S이고, 573번 위치의 아미노산이 S 또는 A이고, 592번 위치의 아미노산이 G 또는 S이고, 594번 위치의 아미노산은 L 또는 F임.
- [0467] E28. 실시양태 25 또는 26에 있어서, 서열번호: 9의 변형되지 않은 과활성 PiggyBac와 비교하여 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 추가적인 돌연변이를 포함하고, 여기서 변형된 과활성 PiggyBac이 서열번호: 9의 과활성 PiggyBac와 비교하여 감소된 DNA 결합 활성 및/또는 증가된 절제 활성을 나타내는, 변형된 과활성 PiggyBac 트랜스포사제.
- [0468] E29. 실시양태 1 내지 18 중 어느 하나의 융합 단백질을 코딩하는 핵산, 전형적으로 메신저 RNA (mRNA).
- [0469] E30. 실시양태 28에 있어서, 서열번호: 110-112로 이루어진 군으로부터 선택된 서열 또는 이의 상응하는 mRNA 서열을 포함하는, 실시양태 28의 융합 단백질을 코딩하는 핵산.
- [0470] E31. 실시양태 25-28 중 어느 하나의 변형된 과활성 Piggybac을 코딩하는 핵산.
- [0471] E32. 실시양태 29-31 중 어느 하나의 핵산을 포함하는 발현 벡터.
- [0472] E33. 실시양태 29-31 중 어느 하나의 핵산 또는 실시양태 32의 발현 벡터를 포함하는 숙주 세포.
- [0473] E34. 세포 계놈으로의 외인성 핵산 서열의 부위 특이적 통합 방법으로서, 세포에
- [0474] (i) 실시양태 1 내지 18 중 어느 하나의 융합 단백질, 또는 실시양태 29 내지 31 중 어느 하나의 핵산,
- [0475] (ii) 세포 계놈에 통합될 외인성 핵산, 및
- [0476] (iii) 세포 계놈으로의 상기 외인성 핵산의 부위-특이적 통합을 결정하기 위한 가이드 RNA
- [0477] 를 포함하는 조성물을 전달하는 것을 포함하며,
- [0478] 여기서 상기 융합 단백질을 세포 계놈 내의 특이적 계놈 DNA 서열에 결합하여 계놈의 절단 및 가이드 RNA에 의해 결정된 바와 같이 세포 계놈으로의 상기 외인성 핵산의 부위-특이적 통합을 발생시키는, 방법.
- [0479] E35. 실시양태 34의 방법에 있어서, 상기 외인성 핵산은 크기가 적어도 5 kb, 적어도 6 kb, 적어도 7 kb, 적어도 8 kb, 적어도 9 kb, 전형적으로 5 내지 25 kb, 바람직하게는 8 내지 20 kb인 핵산 단편인 방법.
- [0480] E36. 실시양태 34 또는 35에 있어서, 상기 외인성 핵산이 유전적 장애의 결핍을 교정하기 위해 이를 필요로 하는 대상체의 계놈에 삽입되는 치료적 트랜스جين인 방법.
- [0481] E37. 실시양태 34 내지 36 중 어느 하나에 있어서, 상기 조성물이 전형적으로 포유동물 세포, 바람직하게는 인간 세포, 그리고 보다 바람직하게는 유전적 장애를 앓고 있는 인간 대상체로부터 얻은 인간 세포에서, 시험관내 또는 생체외에서 전달되는 방법.
- [0482] E38. 실시양태 34 내지 37 중 어느 하나에 있어서, 상기 조성물이 전형적으로 유전적 장애의 치료적 치료를 위해 이를 필요로 하는 포유동물, 예를 들어 인간 대상체 내로 생체내 전달되는 방법.

- [0483] E39. 외인성 핵산 서열을 유기체의 게놈 DNA에 삽입하는 방법으로서, 실시양태 20-24 중 어느 하나에 정의된 바와 같은 하나 이상의 조성물 또는 키트를 유기체에 투여하여 상기 하나 이상의 조성물 또는 키트에 포함된 융합 단백질이 특정 게놈 DNA 서열에 결합하고 상기 조성물에 포함된 외인성 핵산의 게놈 DNA로의 삽입을 가능하게 되도록 하는 단계를 포함하고,
- [0484] 여기서 외인성 핵산이 상기 유기체, 예를 들어 비인간 유기체 또는 이를 필요로 하는 인간 대상체의 세포 게놈의 특정 부위에 통합되는 방법.
- [0485] **실시예**
- [0486] 프로그래밍 가능한 트랜스포사제를 달성하기 위해, 우리의 설계 원칙은 CRISPR 시스템의 정밀-게놈 표적화를 향상된 삽입 및 절제 활성과 낮은 표적 DNA 결합 활성을 나타내는 PB 변이체와 조합하려고 시도하였다. 우리는 뉴클레아제 SpCas9를 PB 트랜스포사제와 융합함으로써 시작하였다 (도 1). 우리는 PB 및 SpCas9 변이체의 다양한 라이브러리를 구축하였으며 (도 2a, b, c), 여기서 PB 변이체, 3가지 SpCas9 변이체 (뉴클레아제 SpCas9 (cas9), 데드 Spcas9 (dcas9), Nickase SpCas9 (ncas9))의 라이브러리, 및 6개의 링커 (4xGGG, 5xGGG, 7xGGG, 8xGGG, XTEN, FOKI)를 시험하였다.
- [0487] PB 트랜스포사제의 상이한 부분은 다양화되었다. 트랜스포사제-표적 DNA 상호작용에 관여할 가능성이 있는 15개의 아르기닌 및 리신 잔기로 둘러싸인 3개의 아스파테이트 잔기 (D268, D346, D447)의 촉매 트리어드에 의해 형성되는 촉매 코어 도메인에 특별히 중점을 두었다¹⁴. 과거⁹에 탐구된 과활성 버전에 더하여 PB 절제 및 통합 활동¹⁵에 영향을 미칠 수 있는 추가 잔기들을 다양화하였다 (SEQ 1-N). 최상의 성능을 발휘하는 돌연변이체를 단리하기 위해, 온-타겟 통합시 에머랄드(Emerald) GFP (emGFP) 재구성을 기반으로 하여, 우리는 표적 유전자 삽입을 위한 민감한 리포터 시스템을 개발하였다. 스플라이싱 수용체가 선행하는 emGFP의 프로모터가 없는 C-말단 (C-t) 절반을 Hek293T 세포 게놈에 삽입하여 리포터 세포주 (허쉬라고 함)를 구축하였다. 상보적인 '삽입-트랩 리포터'는 emGFP의 N-말단 (N-t) 앞 절반부를 코딩하고 상류 프로모터에 이어서 스플라이스 공여자가 뒤따르고 모두 PB 역 반복부에 플랭킹되도록 구축하였다. 특정 gRNA-유도 cas9는 PB가 C-t 절반에 인접한 N-t 절반을 삽입하도록 지향하였으며, 생성된 전사물의 스플라이싱 시 녹색 형광이 생성되었다 (도 3). cas9-PB 키메라 단백질의 라이브러리를 어셈블리하고 허쉬 리포터를 함유하는 세포에 '삽입-트랩 리포터' 및 가이드-RNA (gRNA)로 형질감염시켰다. 더 높은 프로그래밍 가능한 삽입을 제시하는 변이체는 동일한 리포터 세포주를 사용하여 별도로 시험하였다 (도 2a, 2b; 4a, 4b). 이 검정은 N-t 절반 emGFP 및 전체 RFP 서열 상류를 함유하는 이중 트랜스포존을 사용하여 온-타겟 활성 (emGFP 양성 세포) 및 전체 전위 활성 (RFP 양성 세포)을 시험하였다. RFP 양성 세포의 총 백분율 (총 삽입 활성)에 대한 emGFP 양성 세포의 백분율 (온-타겟 활성)을 나누어 온-타겟 대 오프-타겟 비율을 계산하였다.
- [0488] 허쉬 세포주에서 온-타겟 및 오프-타겟 전위를 위해 PB 변이체에 융합된 cas9 변이체의 다양한 조합을 시험하였다. 우리는 i) 무손상 뉴클레아제 활성을 갖는 cas9 (도 2A) 및 ii) 증가된 절제 활성을 갖는 PB 변이체 및 iii) 감소된 t-DNA 결합 활성을 함유하는 Nter-cas9-PB-Cter 융합에서 최고 수준의 프로그래밍 가능한 삽입을 관찰하였다 (도 2B; 4A, 4B). 세 가지 파라미터 모두 정밀하고 효율적인 표적 삽입을 달성하는 데 중요하다는 것을 관찰하였다. 첫째, PB에 융합된 ncas9 (D10A) 및 dcas9 (D10A 및 H840A)에 의한 표적 삽입의 상당히 낮은 효율에 의해 cas9 뉴클레아제 활성에 대한 요구 사항이 입증되었다 (도 2A). 표적 통합을 촉진하는 데 있어 cas9의 이중 가닥 파단 (DSB) 활동의 역할을 추가로 탐구하기 위해, 우리는 트랜스포존의 지향 국부화를 위해 징크 핑거-PB 융합을 사용하여 현장 표적화 및 DSB 활동을 분리하고 이를 현장 DSC로 -독립적인 Cas9 뉴클레아제에 의한 보완하였다. Znf-PB 융합은 Znf 결합 부위 근처에 gRNA가이드된-cas9 (도 5)를 사용하여 DSB를 도입하는 것과 조합할 때 구제된(rescued) 표적 삽입 활성이 없거나 매우 낮았다. 이들 결과는 PB 부근에서 cas9에 의한 DSB 생성이 PB의 삽입 활성을 촉진하고 삽입 부위에서 표준 TTAA에 대한 그 요구 사항을 우회하는 메커니즘과 일치한다. 실제로, cas9-PB 삽입 부위에 대한 우리의 분석은 역전 말단 반복부 (ITR) 서열이 표적 부위 근처에 작은 indel 존재로 인해 파괴된다는 것을 나타냈다 (도 6). 이러한 파괴의 중요한 결과는 PB-매개 통합 메커니즘의 비가역성이다. 파괴된 ITR 또는 TTAA가 있는 트랜스포존의 동원은 제거되거나 감소한다¹⁶. 이 메커니즘은 cas9-PB에 의한 프로그래밍 가능한 삽입의 효율에 기여할 가능성이 높으며, 변형된 PB의 '붙여넣기 (pasting)' 활동과 cas9의 '찾기(find)' 및 '잘라내기(cut)' 활동의 결합은 높은 수준의 정밀도에 기여한다. 둘째, 높은 절제 PB 돌연변이체는 더 높은 프로그래밍 가능한 삽입에 기여하는 것으로 보인다 (D450N, M194V; 도 2B, 4A, 4B). 이 향상된 절단은 통합을 위해 프라이밍된 공여자 DNA 기질이 증가한 결과일 수 있다. 더욱

이, 바람직한 절제 기질의 파괴는 더 높은 절제 활성을 갖는 PB 돌연변이체로부터도 충분한 절제를 방지할 수 있다 (도 6). 셋째, t-DNA 결합이 감소된 PB 돌연변이체는 가장 낮은 온-타겟 수준을 발생시켰다 (도 2B, 4A, 4B). 이 결과는 PB 단독 삽입 활성을 불활성화하는 cas9-PB 융합에서 cas9의 서열-특이적 DNA 결합에 의해 보완되는 PB에 의한 내재적 비특이적 DNA 결합의 감소와 일치한다 (도 7).

[0489] 게놈 가이드 접근법을 사용하여, 우리는 가이드-seq¹⁷ 기반 프로토콜의 변형된 버전을 사용하여 프로그래밍 가능한 트랜스포사제 삽입 부위 및 온-타겟 수준을 특성화할 수 있었다 (도 8A). TTAA 부위에서 감지된 온-타겟 삽입은 전혀 발생하지 않았으며, 이는 cas9에 의해 생성된 DSB 부위에 대한 통합을 추가로 입증하고 선호하는 절제 기질의 손실을 발생시켰다 (도 8B). 또한, TRAC 유전자좌를 표적으로 하는 프로그래밍 가능한 트랜스포사제 기술로 변형된 세포에 대해 가이드-seq 분석을 시행하였다. 우리는 1-10%까지의 감도로 온-타겟 모든 삽입을 검출하였다 (도 9).

[0490] 우리는 cas9 기반 HDR (도 10)과 같은 정밀 유전자 전달을 위한 현재 방법으로 프로그래밍 가능한 트랜스포사제 기술을 벤치마킹하였으나 프로그래밍 가능한 트랜스포사제는 더 높은 효율, 큰 페이로드로 넓혀지는 캡을 나타낸다. 최상의 돌연변이는 HDR보다 2배 더 높은 효율과 높은 정확도로 삽입 (최대 8 kb)을 달성했다. 우리는 또한 프로그래밍 가능한 트랜스포사제를 Cas9를 PB의 촉매 데드 버전에 융합한 HITI 변이체와 비교하였다. 이는 최근 SB100 트랜스포사제의 DNA 결합 도메인을 사용하는 유사한 접근 방식에서 제안된 것처럼 삽입 부위에 DNA를 모집하는 데 도움이 될 수 있다. 프로그래밍 가능한 트랜스포사제는 대안적인 보조 HITI 방법에 비해 두 배 더 높은 효율을 나타낸다 (도 18). 생체내 마우스 모델에서 프로그래밍 가능한 트랜스포사제 활성을 입증하기 위해, 우리는 프로그래밍 가능한 트랜스포사제의 mRNA 버전을 구축하였다. 우리는 생체내 JetPEI 시약을 사용하여 Rosa26 게놈 세이프 하버(safe harbour)를 표적으로 하는 마우스 간으로 프로그래밍 가능한 트랜스포사제를 전달했고 내인성 유전자와 비교하여 높은 카피 수의 트랜스진을 관찰했으며 (도 11) 시간이 지남에 따라 트랜스진 발현을 유지하였다 (도 20).

[0491] 요약하자면, 우리는 변형된 PB의 DNA 잘라내기 및 붙여넣기 활동과 CRISPR 분자 인식 및 절단을 결합하여 정밀하고 효율적인 유전자 전달을 수행하는 효율적인 도구를 생성하였다. 이 기술은 페이로드 크기에 따라 매우 잘 확장된다. 우리는 Hek293T와 마우스 간에서 그 효능을 입증하였다. 우리는 고급 요법 및 기타 적용을 위한 치료 유전자 전달을 위한 일반화된 플랫폼으로 프로그래밍 가능한 트랜스포사제 기술을 구상한다.

[0492] **hyPB에 융합된 Cas 변이체의 결과**

[0493] 프로그래밍 가능한 전위를 수행하기 위해 조작된 hyPB의 능력을 추가로 특성화하기 위해, 우리는 다른 Cas 단백질이 뉴클레아제 활성을 가진 상이한 유기체를 형성하는, 시험된 프로그래밍 가능한 트랜스포사제의 SpCas9 머뮬(mudile) (즉, 서열번호: 72의 SaCas9, 서열번호: 74의 cpf1, 서열번호: 75의 CasX, 및 서열번호: 29의 CjCas9)을 치환하였다. 분할 GFP 리포터의 상류 영역을 표적으로 하는 특이적 gRNA를 설계하고 허쉬 세포주 형질감염을 위해 클로닝하였다 (하기 표 2 참조). 표적 전위는 GFP 발현에 의해 측정되었다 (도 12).

[0494] 이들 결과는 또 다른 실험 라인에서 확인되었다. 우리는 CjCas9 및 LbCpf1에 대해 우수한 프로그램 가능한 삽입 활성을 획득하였으며, 한편 CasX는 우리의 검정에서 프로그래밍 가능한 통합을 달성하지 못하였다. 특히, SaCas9는 변형된 hyPB에 융합된 SpCas9와 유사한 수준으로, 시험된 Cas 단백질 중에서 가장 높은 수준의 프로그래밍 가능한 삽입을 가졌다 (도 19). 사용된 상이한 Cas 단백질 및 일루미나(Illumina) NGS (도 21)에 의해 각각의 단백질에 대해 설계된 세 가지 상이한 gRNA에 대해 Indels이 결정되었으며, 정규화 목적으로 표시되었다.

[0495] 이들 긍정적인 결과는 임의의 서열 특이적 뉴클레아제 모듈에 유용하도록 프로그래밍 가능한 전위를 위해 조작된 hyPB를 검증한다.

[0496] **PB 돌연변이체의 이량체에 융합된 Cas9의 추가 결과**

[0497] 전위를 수행할 때 이량체로서 작용하는 PB의 특성을 감안하여, Cas9와 hyPB R372A-K375A-D450N 돌연변이체의 융합 단백질 생성을 시도하였다. 우리는 이들 융합의 온-타겟 활성을 Cas9-PB 돌연변이체 단독과 비교하였다. 우리는 구성 Cas9-PB-PB의 더 양호한 성능을 관찰하였으며; 한편 PB-Cas9-PB 구성은 Cas9-PB 단량체 융합을 능가하지 못하였다 (도 14). Cas9에 대한 이량체 융합을 위해 우리는 클로닝 및 발현을 용이하게 하기 위해 기록된 버전의 hyPB R372A-K375A-D450N 돌연변이체를 사용하였다.

[0498] 흥미롭게도, 이량체 hyPB R372A-K375A-D450N에 융합된 Cas9가 SpCas9 대신에 SaCas9인 경우, 활성이 추가로 증가된다 (도 24). 이량체 hyPB를 사용하여 SpCas9에 비해 SaCas9의 증가된 성능은 단량체 hyPB로 획득된 결과와

일치한다 (도 12 및 19).

[0499] **ZNF-PB 돌연변이체 구제의 결과**

[0500] 우리는 오프 타겟 부위에서 비유도 DSB에 의한 오프 타겟 활성을 저하시키면서, 표적 통합을 용이하게 하는데 SpCas9 니카제 변이체 (D10A)에 의한 단일 가닥 커트를 촉진하는 2개의 gRNA의 가까운 (4개의 뉴클레오타이드) 표적 부위에 의해 유도된 이중-가닥 파단 (DSB) 활성의 역할을 추가로 탐구하고 싶었다. 우리는 D450N 돌연변이체 및 R372A-K375A-D450N 돌연변이체와의 융합에 의해, 트랜스포존의 지향된 국부화를 위해 징크 핑거-PB 융합을 사용하고 독립적인 니카제 Cas9에 의한 2개의 현장(on-site) 단일 가닥 브레이크(brake), 또는 Cas9 뉴클레아제에 의한 단일 DSB로 이를 보완하였다.

[0501] Znf-PB 융합은 단일 또는 이중 gRNA 가이드-cas9 뉴클레아제 또는 니카제와 함께 Znf 결합 부위 근처에 DSB를 도입하는 것과 조합될 때 구제된 표적 삽입 활성을 전혀 나타내지 않거나 매우 낮게 나타냈다 (도 13).

[0502] **Cas9-hyPB 돌연변이체 변이체의 결과**

[0503] 더 양호한 효율로 프로그래밍 가능한 전위를 수행할 수 있는 돌연변이체 조합을 추가로 탐구하기 위해, 분할 GFP 리포터 시스템의 프로그래밍 가능한 삽입에 의해 GFP가 재구성되는 세포 선택의 수개의 주기를 수행하였다. 흥미롭게도 우리는 Cas9-hyPB R372A-K375A-D450N을 능가하는 수개의 조합을 관찰하였다 (도 15). 특히 언급할 가치가 있는 것은 hyPB에서 AA: A351-A372-A375-A388-N450-A465-A573-V589-G592-L594 (서열번호: 2로도 식별됨)에서 돌연변이된 Cas9에 융합된 hyPB의 변이체이다. R372A-K375A-D450N (서열번호: 1)과 비교하여 양성 세포 집단에서 수배(several fold)가 농축화되며; 또한 A245-A275-A277-A372-A465-V589 (서열번호: 3) 및 A275-A325-A372-A560 (서열번호: 4)은 더 적은 정도로 농축화된다.

[0504] 또 다른 실험 라인에서, PiggyBac DNA 라이브러리는 트위스트 바이오사이언스(Twist Bioscience)에 의해 생산되었으며, cas9와 융합하여 렌티바이러스 벡터로 클로닝되고 stb4 유능한 세포로 형질전환되어, x100 변이체 복잡성(variant complexity)을 보장하였다. 플라스미드를 maxiprep에 의해 정제하고 렌티바이러스 패키징 플라스미드로 Hek293T 세포에 공동형질감염시켰다. 렌티바이러스는 1/2 GFP 리포터 세포주를 감염시키는 데 사용되었다. 감염된 세포를 AAVS1 서열을 표적화하는 1/2 GFP 트랜스포존 및 gRNA로 형질감염시켰다. GFP 양성 세포는 유세포 분석 분류(flow cytometry sorting)에 의해 선택되었고 게놈 DNA가 추출되었다. 추출된 gDNA로부터 PB를 증폭하고, 렌티바이러스 벡터에 재클로닝하여 새로운 주기를 다시 시작하였다. 최상 성능의 프로그래밍 가능한 트랜스포사제 변이체를 선택하여 AAVS1 gRNA 및 MC 1/2 GFP로 개별적으로 형질감염시켰다.

[0505] 먼저, 96개의 변이체의 무작위 선택을 수행하고 최상의 성능의 변이체를 별도로 스크리닝하였다 (도 16). 높은 온-타겟 삽입을 위한 최상의 PB 아미노산 변이체의 요약은 돌연변이 D450N, R372A 및 K375A의 중요성을 확인해 주나; 증가된 표적 효율에 기여하는 다른 중요한 잔기를 강조한다 (도 17B). 최상의 온-타겟 효율을 가진 6개의 PB 변이체를 선택하였다 (도 17A). 개별 온-타겟 활성은 FiCAT (Cas9-hyPB R372A-K375A-D450N)와 비교하여 하기 변이체로 유의하게 개선되었다: N347A-D450N; N347S-D450N-T560A-S573A-F594L; R202K-R275A-N347S-R372A-D450N-T560A-F594L; R275A-N347S-K375A-D450N-S592G; R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-R277A-N347S-R372A-D450N-T560A-S564P-F594L (양측 t-검정(two-sided t-test)).

[0506] 이 실험을 반복하여 확인하였다 (도 22A). 우리는 또한 각각의 주기의 별크 변이체를 발현하는 렌티바이러스 및 PB 변이체 CN에 의해 이의 역가를 교정하는 감염된 리포터 세포주를 생산하여, 주기에 걸쳐 유사한 온-타겟 효율 증가를 입증하였다 (도 22B). 단일 돌연변이체는 cas9_PB 라이브러리 농축화의 4주기 및 5주기 후에 별크 변이체로부터 단리되었다. 온-타겟 리포터 세포주를 FiCAT 돌연변이체, gRNA tcr1 및 1/2 GFP MC 트랜스포존으로 형질감염하여 돌연변이를 개별적으로 시험하였다. 최상의 FiCAT 돌연변이는 FiCAT R372A_K375A_D450N과 비교하여 표시된다 (도 23). 개별 온-타겟 활성은 FiCAT (Cas9-hyPB R372A-K375A-D450N)과 비교하여 하기 변이체로 유의하게 개선되었다: R202K-R275A-N347S-R372A-D450N-T560A-F594L; R245A-N347S-R372A-D450N-T560A-S564P-S573A-S592G; R275A-N347S-R372A-D450N-T560A-F594L; N347A-D450N; R277A-G325A-N347A-K375A-D450N-T560A-S564P-S573A-S592G-F594L; N347S-D450N-T560A-S573A-F594L; V34M-R275A-G325A-N347S-S351A-R372A-K375A-D450N-T560A-S564P; G325A-N347S-K375A-D450N-S573A-M589V-S592G; S230N-R277A-N347S-K375A-D450N; T43I-R372A-K375A-A411T-D450N; G325A-N347S-S351A-K375A-D450N-S573A-M589V-S592G; Y177H-R275A-G325A-K375A-D450N-T560A-S564P-S592G.

[0507] 돌연변이체 R372A-K375A-D450N과 비교하여 돌연변이체 R202K-R275A-N347S-R372A-D450N-T560A-F594L, R275A-R277A-N347S-R372A-D450N-T560A-S564P-F594L 및 R275A-N347S-R372A-D450N-T560A-F594L의 우월성은 SpCas9 및

2개의 hyPB를 포함하는 삼중 융합 단백질에서 추가로 입증되었다 (도 29).

[0508] **Cas9 및 hyPB 비공유 연결의 결과**

[0509] 링커를 통해 Cas9와 hyPB R372A-K375A-D450N의 공유 결합에 더하여, 우리는 MS2-MCP 시스템을 사용하여 MCP 단백질과 hyPB R372A-K375A-D450N으로 이루어진 융합 단백질과 Cas9를 MCP 단백질에 결합하는 MS2 서열의 테트라루프를 함유하는 변형된 gRNA를 통해 연결하였다.

[0510] MCP-hyPB R372A-K375A-D450N 융합 단백질과 Cas9의 조합은 Cas9-hyPB R372A-K375A-D450N 융합 단백질과 비교하여 프로그램 가능한 삽입 활성이 증가하였다 (도 25A). 또한, MCP 단백질을 hyPB의 다른 돌연변이체에 융합하여 SpCas9와 함께 프로그래밍 가능한 전위를 수행하였다. 사용된 두 변이체 (R202K-R275A-N347S-R372A-D450N-T560A-F594L, 및 R275A-N347S-R372A-D450N-T560A-F594L) 모두 R372A-K375A-D450N을 능가하였다 (도 25B).

[0511] **프로그래밍 가능한 전위를 위한 Cas9 및 디커플링(decoupling)의 결과**

[0512] 우리는 또한, 링커 없이, 또는 MS2-MCP 시스템도 없이 hyPB R372A-K375A-D450N으로 SpCas9의 성능을 시험하였다. 우리는 동일한 세포 SpCas9 및 hyPB R372A-K375A-D450N에서 공동-발현시켰고, Cas9-hyPB R372A-K375A-D450N 융합 단백질과 비교하여 증가된 프로그래밍 가능한 삽입 활성이 등록되었다 (도 26A). 우리는 Cas9에 융합되지 않았으나, 동시에 발현되고, 프로그래밍 가능한 전위의 활성을 달성하기 위해 함께 작용하는 시험된 hyPB 돌연변이체 변이체의 수를 확장하였다 (도 26B).

[0513] **Cas9-hyPB 및 MCP-hyPB 융합 단백질의 공동-발현의 결과**

[0514] 우리는 비공유결합으로 연결된 단량체 중 하나와의 융합의 이량체 버전을 수득하기 위해, MCP 단백질에 융합된 hyPB R372A-K375A-D450N 돌연변이체 및 SpCas9에 융합된 hyPB 돌연변이체를 공동-형질감염시켰다. SpCas9에 융합된 수개의 hyPB 돌연변이체를 특이적 표적 통합에 대해 비교하였다 (도 27).

[0515] **Cas9-hyPB 및 hyPB 변이체의 공동-발현 결과**

[0516] 유사한 방식으로, 우리는 SpCas9 hyPB R372A-K375A-D450N 융합 단백질 및 hyPB 돌연변이체를 독립적으로 공동-형질감염시켜, 연결되지 않은 단량체 중 하나와의 융합의 이량체 버전을 수득하였다 (도 28).

[0517] **방법**

[0518] **클로닝 및 플라스미드.** 무작위 삽입 모니터링을 위한 RFP 트랜스포존 PB512-B는 시스템 바이오사이언시즈 인크.(System Biosciences Inc.)에서 구입하였다. hyPB 벡터는 웰컴 트러스트 생어 인스티튜트(Wellcome Trust Sanger Institute) (pCMV_hyPBase)9에서 구입하였다. 플라스미드 벡터 pCRTM-Blunt II-TOPO®는 인비트로젠(Invitrogen)에서 구입하였고 cas9, ncas9 및 SP-dcas9-VPR은 애드진(Addgene)에서 구입하였다 (애드진 플라스미드 #41815, #41816, #63798). 마지막으로 SB100X와 pT4-HB는 닥터. 주즈사나 지트스바크(Dr. Zsuzsana Zizsvak)의 친절한 선물이었다. 더 제로 블런트(The Zero Blunt) TOPO PCR 클로닝 키트 (인비트로젠)를 사용하여 gRNA를 생산하였다. U6 프로모터, 20 nt 표적 부위, gRNA 스캐폴드(scaffold) 및 종결자를 함유하는 gblock 유전자 단편 (인테그레이티드 DNA 테크놀로지(Integrated DNA Technologies))를 사용한다. gRNA TRAC는 실험실에서 설계 및 검증되었으며 gRNA aavs1 3 서열은 이전에 기재되었다¹⁸.

[0519] HyPB 및 PB RFP ½ emGFP SMN1 트랜스포존에 대한 뉴클레아제, 니카제 및 테드 cas9 융합은 BspQI 효소 및 표준 방법을 사용하여 골든 게이트 어셈블리(Golden Gate assembly)에 의해 수행되었다.

[0520] 퀵체인지 라이트닝(Quickchange Lightning) 돌연변이유발 키트 (애질런트(Agilent)의 지침에 따라 부위 지향 돌연변이유발에 의해 cas9 (cas9_PB 플라스미드)에 융합된 hyPB 서열에 상이한 돌연변이가 도입되었다. 프라이머는 HyPB 서열에 대한 하기 돌연변이를 달성하기 위해 퀵체인지 프라이머 디자인(QuickChange Primer Design)으로 설계되었다: M194V, R245A, G325A, R372A, K375A, R376A, E377A, E380A, D450N, S564P (서열번호: 90-99 참조). 모든 플라스미드는 요청시 이용가능하다. PB ½ emGFP SMN1은 emGFP 서열의 앞 절반부와 SMN1 인트론 6 서열을 piggyBac 수용체 벡터에 도입하여 얻었다. pT4 SMN1 2/2 emGFP는 SB100X 트랜스포존 벡터에 제2 절반부 SMN1 인트론 6과 부분 emGFP를 추가하여 얻었다. SMN1을 함유하는 emGFP 서열은 스리 코수리(Sri Kosuri)의 친절한 선물인 DYP004reporter¹⁹에서 얻었다.

[0521] 분할 emGFP 리포터 시스템의 부분 cDNA (NC_000006.¹²) 단편 상류를 복제하여 상이한 크기의 트랜스포존 및 HDR 템플릿을 생성하였다.

- [0522] **세포 배양, 형질감염 및 전기천공법.** Hek293T 세포주 (써모 피셔 사이언티픽) 및 C2C12 세포주 (ATCC)를 높은 글루코스 (깁코(Gibco), 써모 피셔(Therm Fisher), 10% 태아 소 혈청 (FBS), 2 mM 글루타민 및 100U 페니실린 /0.1 mg/mL 스트렙토마이신이 보충된 돌베코 변형 이글 배지(Dulbecco's modified eagle medium) (DMEM)와 함께 5% CO2 인큐베이터에서 37°C에서 배양하였다. 주르카트(Jurkat) 세포주를 글루맥스(Glutamax) 및 HEPES (깁코, 써모 피셔) 및 10% FBS가 보충된 로스웰 파크 메모리얼 인스티튜트(Roswell Park Memorial Institute) 1640 배지 (RPMI)와 함께 5% CO2 인큐베이터에서 37°C에서 배양하였다. 이 세포주는 마넬 후안(Manel Juan) (바르셀로나, 병원 클리닉)로부터의 선물이었다. Hek293T 세포의 형질감염 실험은 제조업체의 설명서에 따라 리포펙타민 3000 시약을 사용하거나 OptiMem에서 1:3 DNA-PEI 비율의 폴리에틸렌이민 (PEI, 써모 피셔 사이언티픽)을 사용하여 수행하였다. 형질감염 당일에 70% 전면생장을 달성하기 위해 전날에 세포를 시딩하였다 (대개 부착 p12 웰 플레이트에서 290,000개 세포). 플라스미드 DNA 비율은 1 트랜스포사제: 2,5 gRNA: 2,5 트랜스포존 또는 1 Cas9: 2,5 gRNA: 2,5 HDR 템플릿 (0,076 pmol 프로그래밍 가능한 트랜스포사제 또는 p12 웰 플레이트용 Cas9 사용)이었다.
- [0523] **emGFP 스피라이싱 기반 재구성 검증.** pT4 SMN1 2/2 emGFP를 함유하는 Hek293T 세포주는 SB100X 및 pT4 SMN1 2/2 emGFP DNA 작제물의 PEI 매개 형질감염에 의해 생성된 후 단일 클론 확장 및 PCR 유전자형 분석 (보충 표 3)에 이어서 생성되었다. 양성 클론을 선택하고 확장하여 후속 검증에 사용하였다. emGFP 재구성 분석을 위해, 프로그래밍 가능한 트랜스포사제, gRNA 및 트랜스포존 플라스미드를 12웰 플레이트에 대해 0,076 pmol 프로그래밍 가능한 트랜스포사제 또는 hyPB 및 0,19 pmols 트랜스포존 및 gRNA를 사용하여 1 프로그래밍 가능한 트랜스포사제: 2,5 gRNA: 2,5 트랜스포존 비율로 형질감염시켰다. 온-타깃 삽입은 형질감염 5일 후 emGFP 형광에 의해 측정되었다. RFP 형광에 의해 형질감염 15일 후 오프-타깃 전이를 측정하였다. 온 타깃 삽입을 PB 변이체, gRNA 및 PB 1/2SMN1 RFP emGFP (최종 작제물 명칭 삽입) 발현 플라스미드를 공동 형질감염시키고 14일 후 (에피솜 붕괴에 대한 평균 n일 설정) emGFP 및 RFP 형광을 (BD LSR 포트테사(Fortessa), BD 바이오사이언시스(Biosciences). 530/30 필터가 있는 청색 488nm 레이저 및 610/20 필터가 있는 황색 녹색 561 nm 레이저)에서 측정된 emGFP 발현에 의해 결정하였다.
- [0524] **삽입 부위 시퀀싱을 위한 접합 PCR** 접합 PCR을 BD FACSAria (바이오사이언시스)를 사용하여 emGFP 분류된 세포에서 수행하였다. 선택된 세포를 리포터 세포주에서 TRAC 표적 부위를 표적으로 하는 PB 1/2 emGFP SMN1 트랜스포존의 표적 삽입을 가졌다. 게놈 DNA는 DNeasy 혈액 및 조직 키트 (퀴아젠(Qiagen))를 사용하여 추출하였다. 프라이머는 트랜스포존 (정방향)의 3' ITR에 의해 설계되었으며 리포터 세포주의 2/2 emGFP 또는 내인성 T 세포 수용체 (TRAC) (역방향)의 인트론을 표적으로 하였다 (보충 표 4).
- [0525] **가이드-SEQ 실험의 생물정보학 분석.** 일루미나 판독을 usearch²⁰으로 클러스터링하고 bwa-mem²¹을 사용하여 참조에 매핑하였다. 온-타깃 삽입 특성화를 위해, 피톤(Python) 스크립팅(scripting) 및 Samtools²²를 사용하여 표적 삽입 부위에서 5' 및 3' 접합부를 포함하는 판독을 선택하였다. indel 수는 CRISPR-GA²³으로 얻었다. 온-타깃 및 오프-타깃 실험을 위해, 백터에 매핑된 클러스터된 판독을 선택하고 bwa-mem을 사용하여 참조 게놈에 매핑하였다. 삽입 피크의 유의성을 macs2²⁴ 알고리즘으로 평가하였다.
- [0526] **표적 삽입에 적합화된 가이드-seq 라이브러리 준비.** 적합화된 가이드-seq¹⁵ 프로토콜 구현은 DNeasy 혈액 및 조직 키트 (퀴아젠)를 사용하여 게놈 DNA를 추출하고 Q800R3 소니케이터(Sonicator)를 사용하여 500 bp 조각으로 단편화하여 수행되었다. KAPA Hyper Prep Kit (KR0961 - v5.16) 및 3 ug의 단편화된 게놈 DNA를 사용하여 말단 복구, A-테일링, 및 Y-어댑터 라이게이션(ligation)을 수행한 후 1X 비율로 AMPure XP SPRI 비드 정제를 수행하였다. 어댑터 라이게이션 후, 각각의 샘플을 2개로 분할하고 GSP5' 또는 GSP3'로 증폭하여 각각 5' 및 3' 접합부를 캡처하였다. 5' 및 3' 트랜스포존-게놈 접합부를 캡처하기 위해, KAPA HiFi DNA를 사용하여 두 개의 중첩된 PCR을 수행하였다. 제조업체 프로토콜에 따른 중합효소: 25ul 최종 부피에서 P5_1 및 PB_5_GSP1 또는 PB_3_GSP1을 사용한 PCR1; 및 25ul 최종 부피에서 P5_2 PB_5_GSP2 또는 PB_3_GSP2를 사용한 PCR2. 5' 및 3' PCR 산물을 1X 비율의 AMPure XP SPRI 비드 정제로 정제하고 등몰비로 혼합한 후 일루미나 Miseq Reagent Kit V2 - 500 사이클 (2 x 250 bp 쌍형성 말단)으로 시퀀싱하였다. 3ul의 100 μM 커스텀 프라이머 인덱스 1 및 Read 2를 시퀀싱 반응에 첨가하였다.
- [0527] **마우스 간으로 생체내 표적 삽입.** 동물 실험 절차는 바르셀로나 바이오메디컬 연구 단지(Barcelona Biomedical Research Park)의 동물 실험 윤리위원회(Animal Experimentation Ethic Committee)의 승인을 받았다. C57BL/6J, 8-10주령이 이 연구에 사용되었다. 동물은 잭슨 레보리토리즈(Jackson Laboratories)로부터 구입하

였으며, 수컷과 암컷을 구분하지 않고 사용하였다. 프로그래밍 가능한 트랜스포사제 mRNA를 제조업체의 설명서에 따라 대규모 RNA 생산 시스템즈-T7 (프로메가)로 생산하였다. Rosa26 gRNA²⁵를 IDT로부터 구입하였다. Rosa26 및 PB512-B 트랜스포존을 표적으로 하는 프로그래밍 가능한 트랜스포사제 mRNA, gRNA를 1:1:2,5 비율로 후안와를 통해 주입하였다. 총 60 ug의 핵산을 복합체화하여; NP 비율 7에서 생체내-JetPEI (폴리플러스 형질 감염)시켰다. 주사 후 10일에 동물을 안락사시키고 간을 단리하고 균질화하였다. 게놈 DNA를 DNeasy 혈액 및 조직 키트 (키아젠)를 사용하여, 간 샘플로부터 추출하였다. Tfrc 내인성 유전자에 대한 트랜스포존 상대 카피 수는 qPCR (보충 표 1에 열거된 프라이머)에 의해 측정하였다. IVIS 스펙트럼 이미징 시스템 (칼리퍼 라이프 사이언시스(Caliper Life Sciences))을 사용하여 FiCAT-gRNA-트랜스포존 또는 트랜스포존 대조군 투여 후 상이한 시점에서 루시페라제 발현의 이미징을 수행하였다. 제조업체의 설명서에 따라 D-루시페린 포타슘 염 (골드 바이오테크놀로지(Gold Biotechnology))을 복강내 주사한 후 5분 후에 이미지를 촬영하였다.

[0528]

PB 구조 모델링. 트코플루시아 니(*Trichoplusia ni*) piggyBac 트랜스포사제 단백질의 3D 구조를 로베타 웹(Robetta Web) 단백질 구조 예측 서버 (<http://rosetta.bakerlab.org>)에서 얻었다. 코어 도메인 (131-550aa)은 데카르트 공간 최소화 및 전체-원자 최적화²⁶가 매립된 몬테 카를로(Monte Carlo) 알고리즘을 기반으로 하는 로세타 비교 모델링(Rosetta Comparative Modeling) 방법으로 예측되었다. 3차 구조 폴드는 SPSServer 및 ProSa-Web 지식-기반 방법으로 분석 및 검증되었다 (보충 도면 2). 이차 구조는 PSIPRED 및 HHPred 기계-학습 기반 방법으로 분석되었다. 이어서 PB의 코어를 비교 단백질 모델링 방법을 통해 PyMOL을 사용한 개선을 위해 모델링하였다. 정제 과정(refinement process)은 바이러스 DNA 및 표적 숙주 DNA에 결합된 HIV 인테그라제 테트라머로 이루어진 Cryo-EM HIV-1 가닥 전달 복합체 인타솜(Strand Transfer Complex Intasome) (PDB ID: 5U1C)과 piggyBac 모델의 중첩과 X선 회절 Tn5 트랜스포사제 복합체 구조 (PDB ID: 1MUS27)에 의해 가이드되었다. 가닥-전달 DNA와 공여자 DNA는 각각 HIV-1 인타솜 및 Tn5의 중첩으로부터 외삽되었다. 단백질과 접촉하는 인터페이스의 뉴클레오티드는 이중 가닥 DNA인 X3DNA로 분석되었다. 우리는 통계적 잠재력을 사용하여 단백질과 DNA 간의 상호 작용을 평가하고 이론적인 PWM²⁸을 생성했다. 이론적인 PWM은 인터페이스에서 모든 잠재적인 이중 가닥 DNA 서열을 시험하고 통계적 잠재력으로 순위를 매기고 상단을 선택하여 다중 서열 정렬을 만들어 측정하였다. 이 원고를 제출하는 동안 cryo-EM 구조를 사용할 수 있게 되었으며, 이는 수행된 모델링과 중요한 일치를 나타낸다²⁹. piggyBac 트랜스포사제 가닥 전달 복합체 (PDB ID: 6X67)의 Cryo-EM 구조는 모델의 일반적인 접힘과 우리가 가정한 도메인이 공여자 및 표적 DNA와의 접촉을 담당한다는 것을 확인해 주었다.

[0529]

[표 2]

[0530] Cas 변이체 gRNA

Cas 변이체	서열	PAM
SaCas9	TATGTACACTTCTGACCCAC (서열번호: 113)	TGGAAT
	GCCTTTAAGCTTGATATCCA (서열번호: 114)	TGGAAT
	GTATCACAATTCCAGTGGGT (서열번호: 115)	CAGAA
	GGACAGGATCGGCATAACCG (서열번호: 116)	GTGAAT
	GTGCTCGGGGCCACTAGGGA (서열번호: 117)	CAGGAT
CpfI	ACTTATAATTCAGTGTATCA (서열번호: 118)	TTTC
	AGCTTGATATCCATGGAATT (서열번호: 119)	TTTA
	TGCTCGGGGCCACTAGGGAC (서열번호: 120)	TTTG
	CTTTTGTAAGTCTTATGGT (서열번호: 121)	TTTA
	CAAAAGTAAATAGCCCGGCT (서열번호: 122)	TTTA
CjCas9	GCCGATCCTGTCCCTAGTGGCC (서열번호: 123)	CCGAGCAC
	ACAATTCCAGTGGGTCAGAAGT (서열번호: 124)	GTACATAC
	ACACTTCTGACCCACTGGAAT (서열번호: 125)	TGTGATAC
	GAATTCATGGATATCAAGCTT (서열번호: 126)	AAAGGCAC
	AATTCAGTGGGTCAGAAGTGT (서열번호: 127)	ACATACAC
CasX	TCAAGCGCGTGTATGTACAC (서열번호: 128)	TTCT
	GGATCGGCATAACCGGTGAA (서열번호: 129)	TTCC
	TAGACATGAGGTCTATGGAC (서열번호: 130)	TTCA
	TAAGCTTGATATCCATGGAA (서열번호: 131)	TTCA
	TATAATTCAGTGTATCACA (서열번호: 132)	TTCC

[0531]

[0532] 참고문헌

[0533] 1. Porteus, M.H. & Carroll, D. *Nat. Biotechnol.* **23**, 967-973 (2005).

[0534] 2. Sander, J.D. & Joung, J.K. *Nat. Biotechnol.* **32**, 347-355 (2014).

[0535] 3. Rees, H.A. & Liu, D.R. *Nat. Rev. Genet.* doi:10.1038/s41576-018-0059-1

[0536] 4. Anzalone, A.V. et al. *Nature* **576**, 149-157 (2019).

[0537] 5. He, X. et al. *Nucleic Acids Res.* **44**, e85 (2016).

[0538] 6. Suzuki, K. et al. *Nature* **540**, 144-149 (2016).

[0539] 7. Klompe, S.E., Vo, P.L.H., Halpin-Healy, T.S. & Sternberg, S.H. *Nature* **1** (2019).

[0540] 8. Strecker, J. et al. *Science* (2019).doi:10.1126/science.aax9181

[0541] 9. Yusa, K., Zhou, L., Li, M.A., Bradley, A. & Craig, N.L. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1531-1536 (2011).

[0542] 10. Hew, B.E., Sato, R., Mauro, D., Stoytchev, I. & Owens, J.B. *Synth. Biol.* **4**, ysz018 (2019).

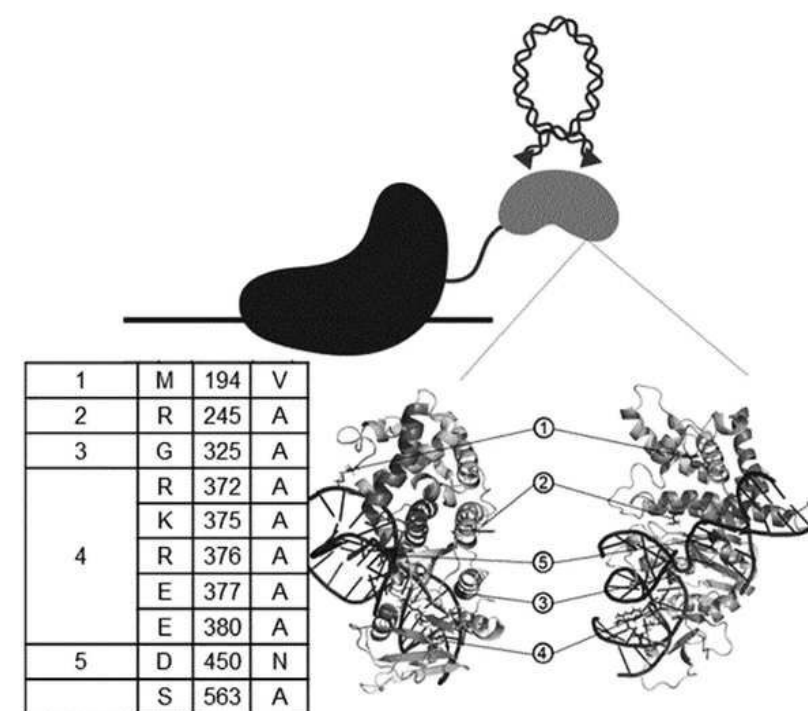
[0543] 11. Kováč, A. et al. *Elife* **9**, (2020).

[0544] 12. Loperfido, M. et al. *Nucleic Acids Research* **44**, 744-760 (2016).

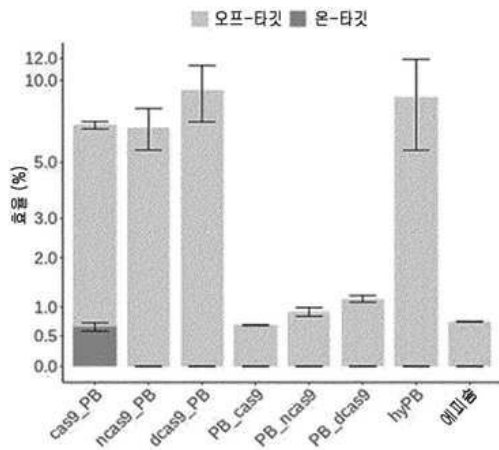
- [0545] 13. Passos, D.O. et al. *Science* **355**, 89-92 (2017).
- [0546] 14. Li, X. et al. doi:10.1073/pnas.1305987110
- [0547] 15. Morellet, N. et al. *Nucleic Acids Res.* **46**, 2660-2677 (2018).
- [0548] 16. Li, M.A. et al. *Mol. Cell. Biol.* **33**, 1317-1330 (2013).
- [0549] 17. Tsai, S.Q. et al. *Nat. Biotechnol.* **33**, 187-197 (2015).
- [0550] 18. Mali, P. et al. *Science* **339**, 823-826 (2013).
- [0551] 19. Cheung, R. et al. *Molecular Cell* **73**, 183-194.e8 (2019).
- [0552] 20. Edgar, R.C. *Bioinformatics* **26**, 2460-2461 (2010).
- [0553] 21. Li, H. *arXiv [q-bio.GN]* (2013).at <<http://arxiv.org/abs/1303.3997>>
- [0554] 22. Li, H. et al. *Bioinformatics* **25**, 2078-2079 (2009).
- [0555] 23. Guell, M., Yang, L. & Church, G.M. *Bioinformatics* **30**, 2968-2970 (2014).
- [0556] 24. Gaspar, J.M.496521 (2018).doi:10.1101/496521
- [0557] 25. Chu, V.T. et al. *BMC Biotechnol.* **16**, 4 (2016).
- [0558] 26. Fu, D.Y. (2018) at <https://etd.library.vanderbilt.edu/available/etd-08012018-164524/unrestricted/DarwinYFu_Thesis_Submit.pdf>
- [0559] 27. Steiniger-White, M., Rayment, I. & Reznikoff, W.S. *Curr. Opin. Struct. Biol.* **14**, 50-57 (2004).
- [0560] 28. Meseguer, A. et al. *NAR Genom Bioinform* **2**, (2020).
- [0561] 29. Chen, Q. et al. *Nat. Commun.* **11**, 3446 (2020).

도면

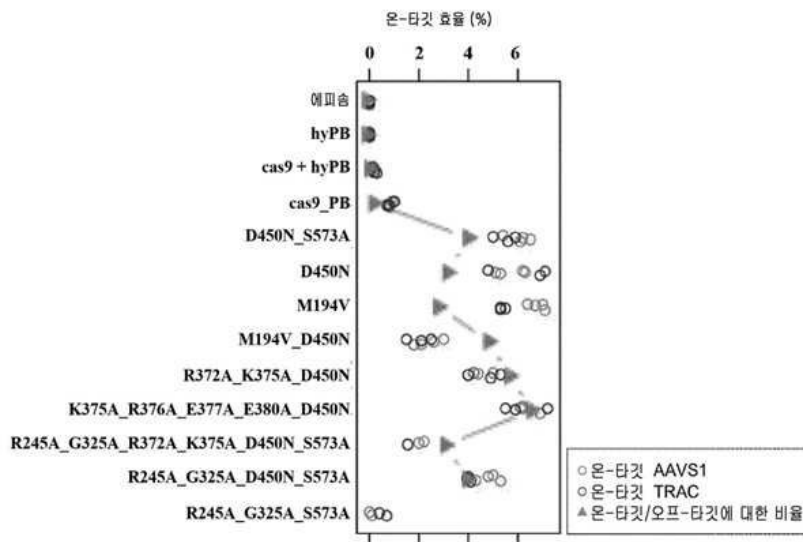
도면1



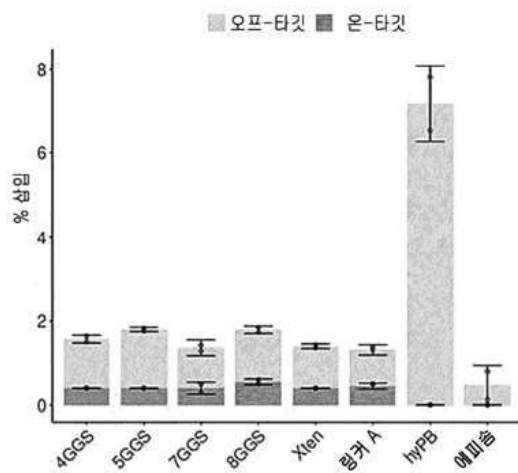
도면2a



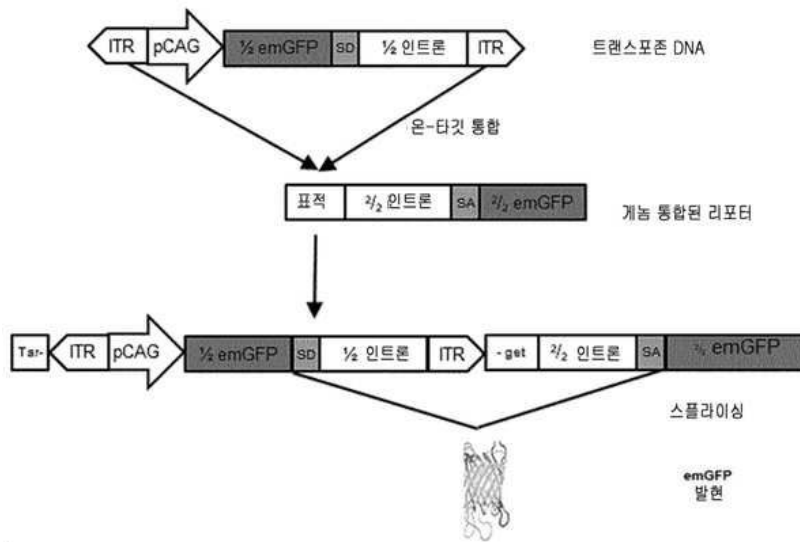
도면2b



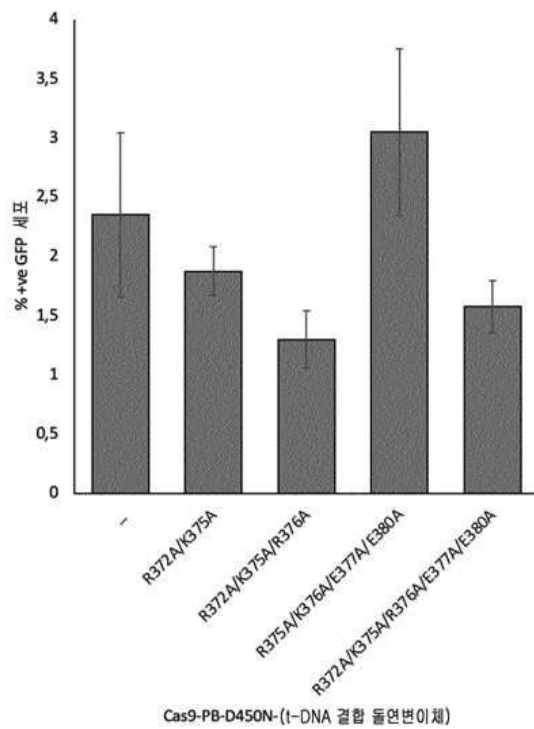
도면2c



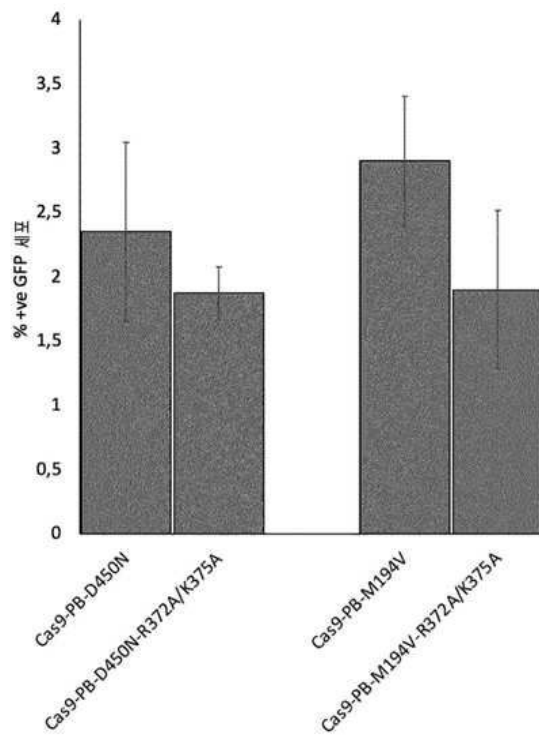
도면3



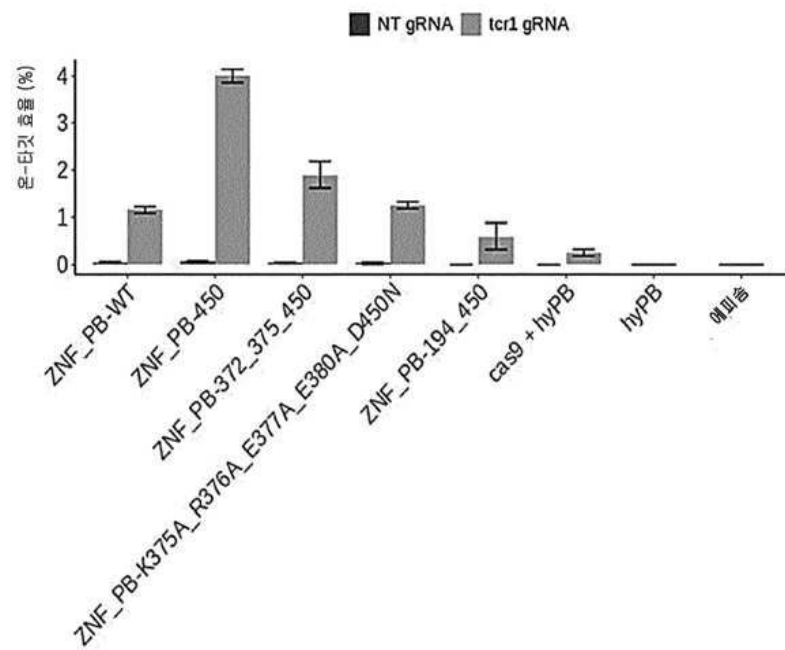
도면4a



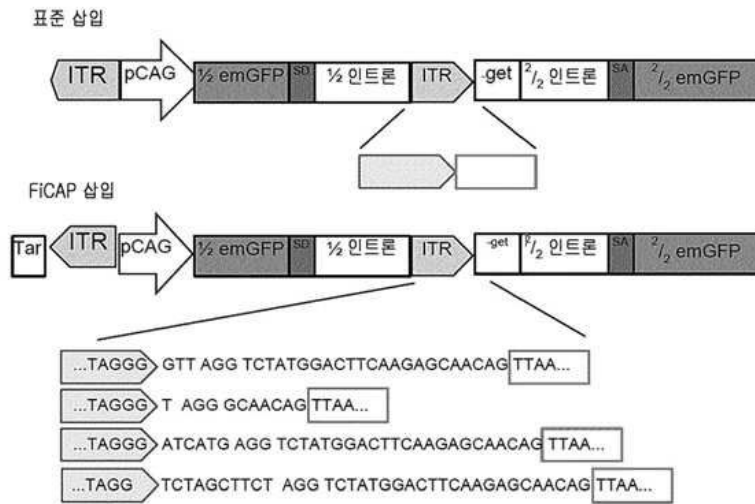
도면4b



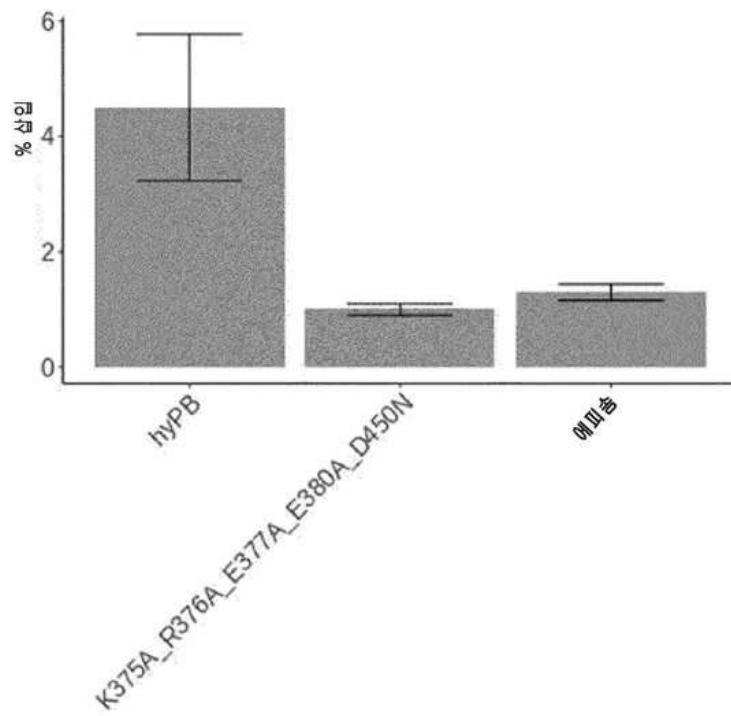
도면5



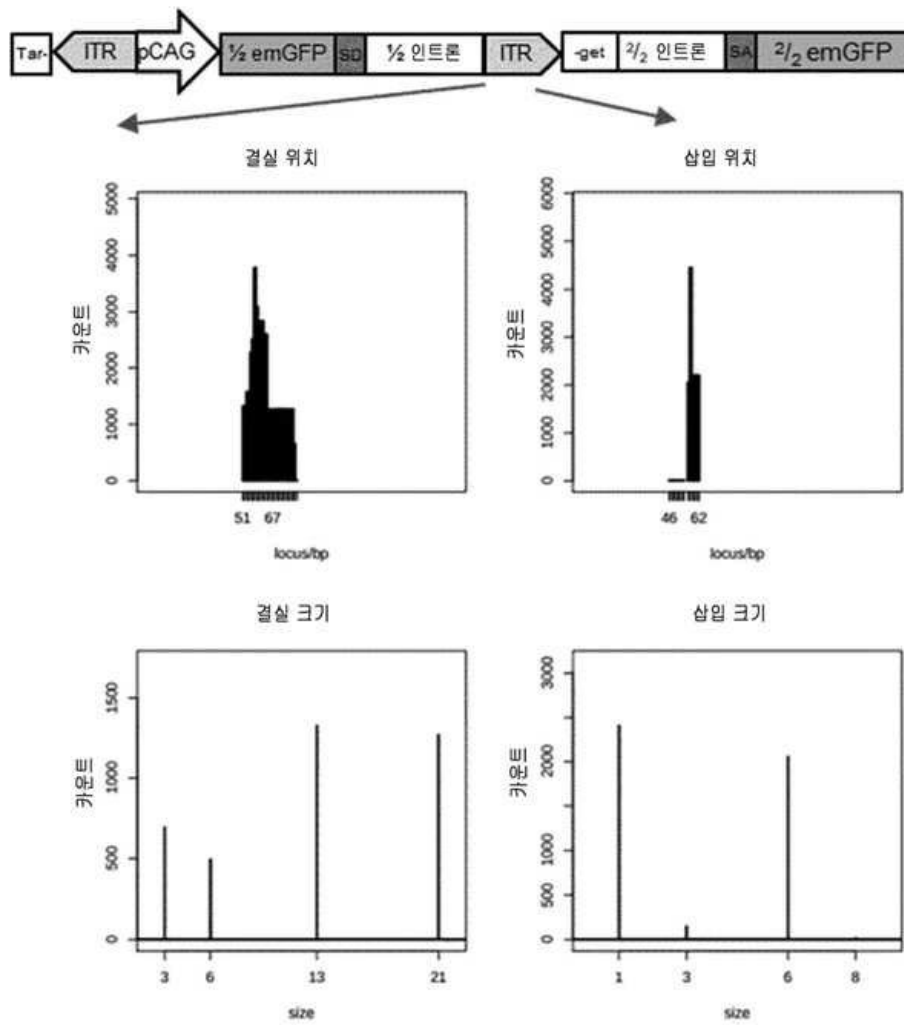
도면6



도면7



도면8a

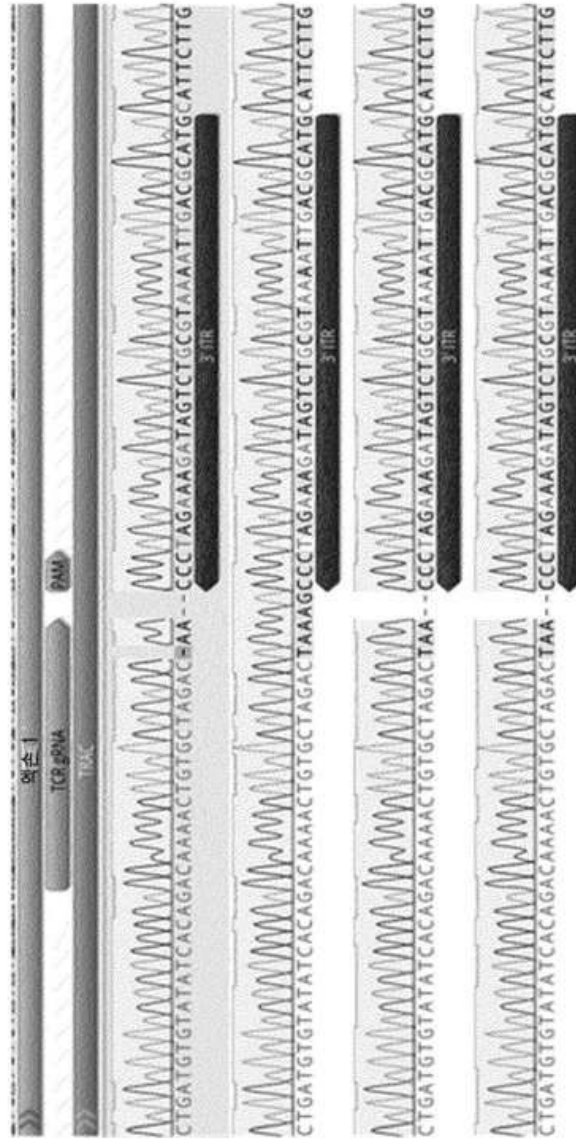


도면8b

TCR-지향된 FiCAP* 삽입의 가이드-seq 프로파일 (엑손 1 TRAC 유전자좌):

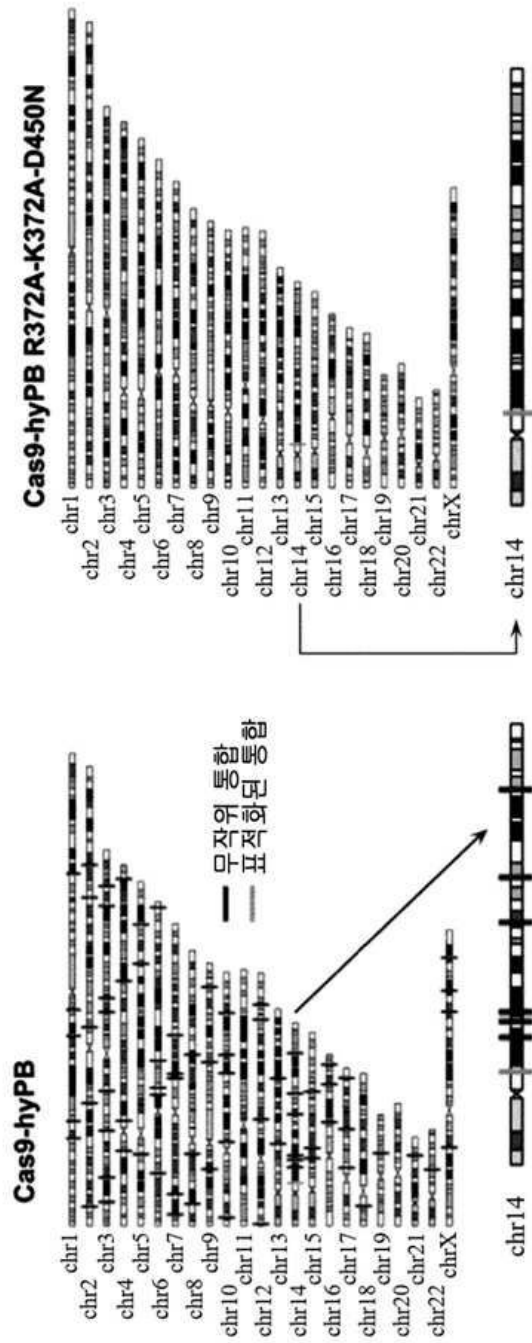


4개의 클론의 생어 시퀀싱:

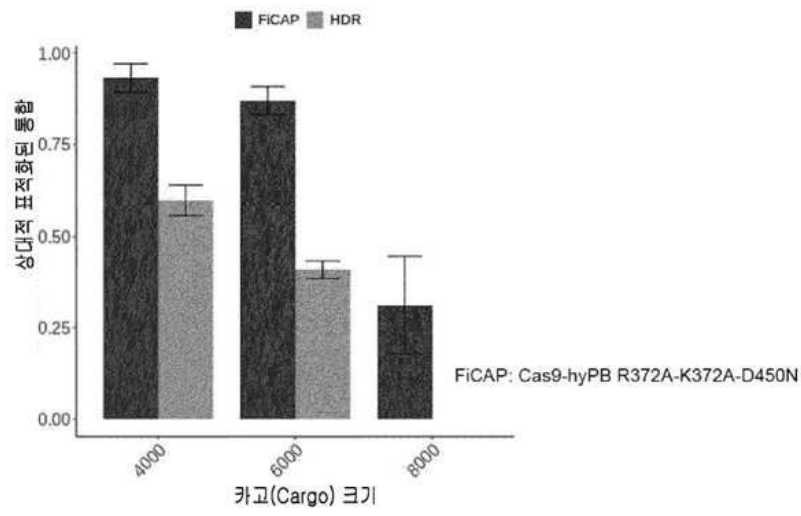


* FiCAP: Cas9-hyPB R372A-K372A-D450N

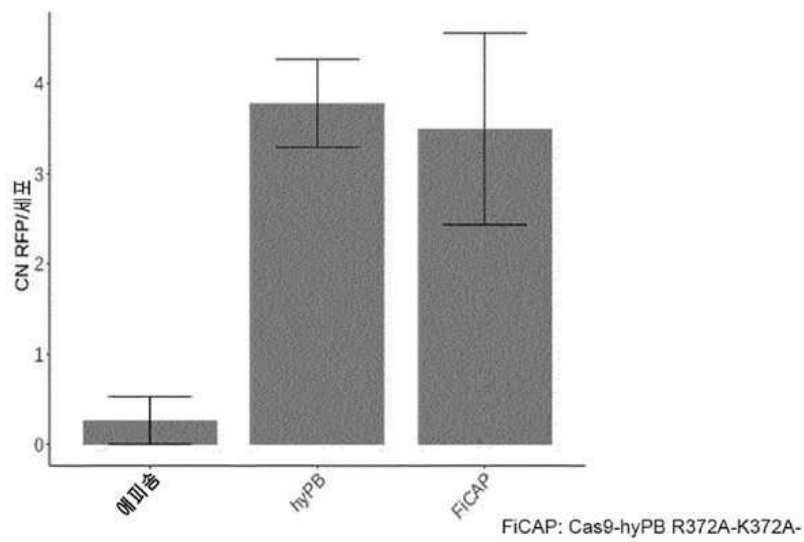
도면9



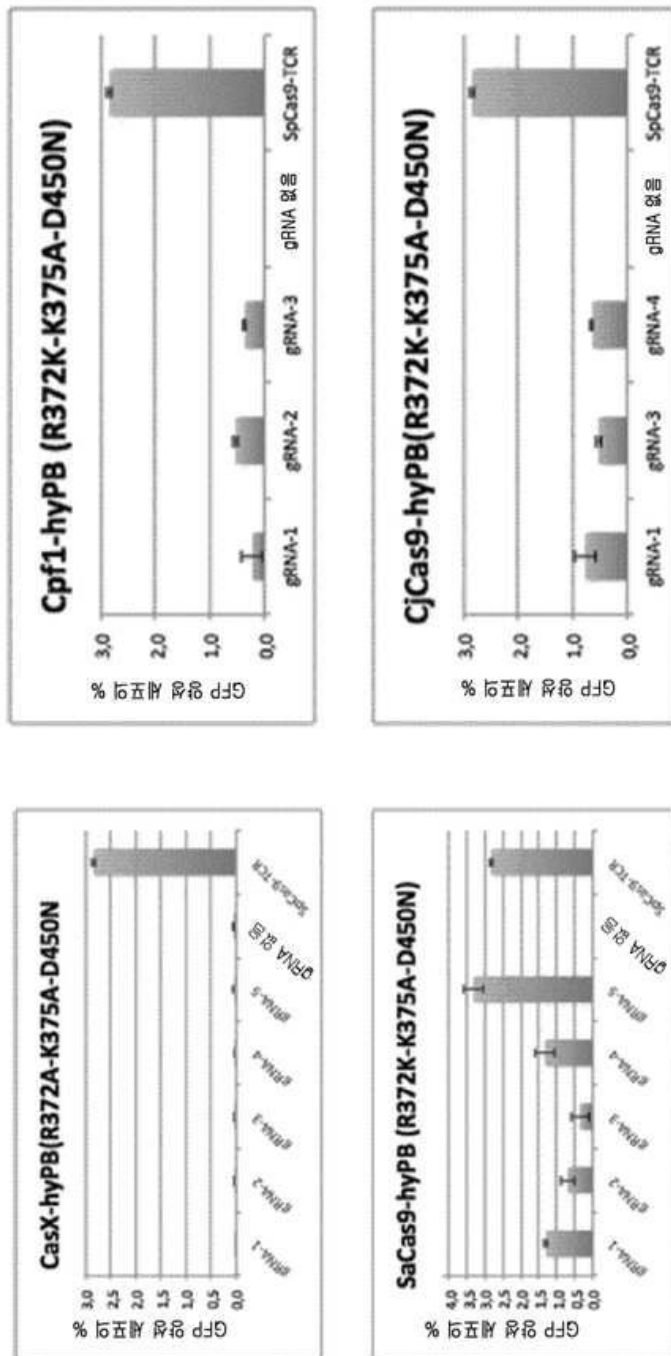
도면10



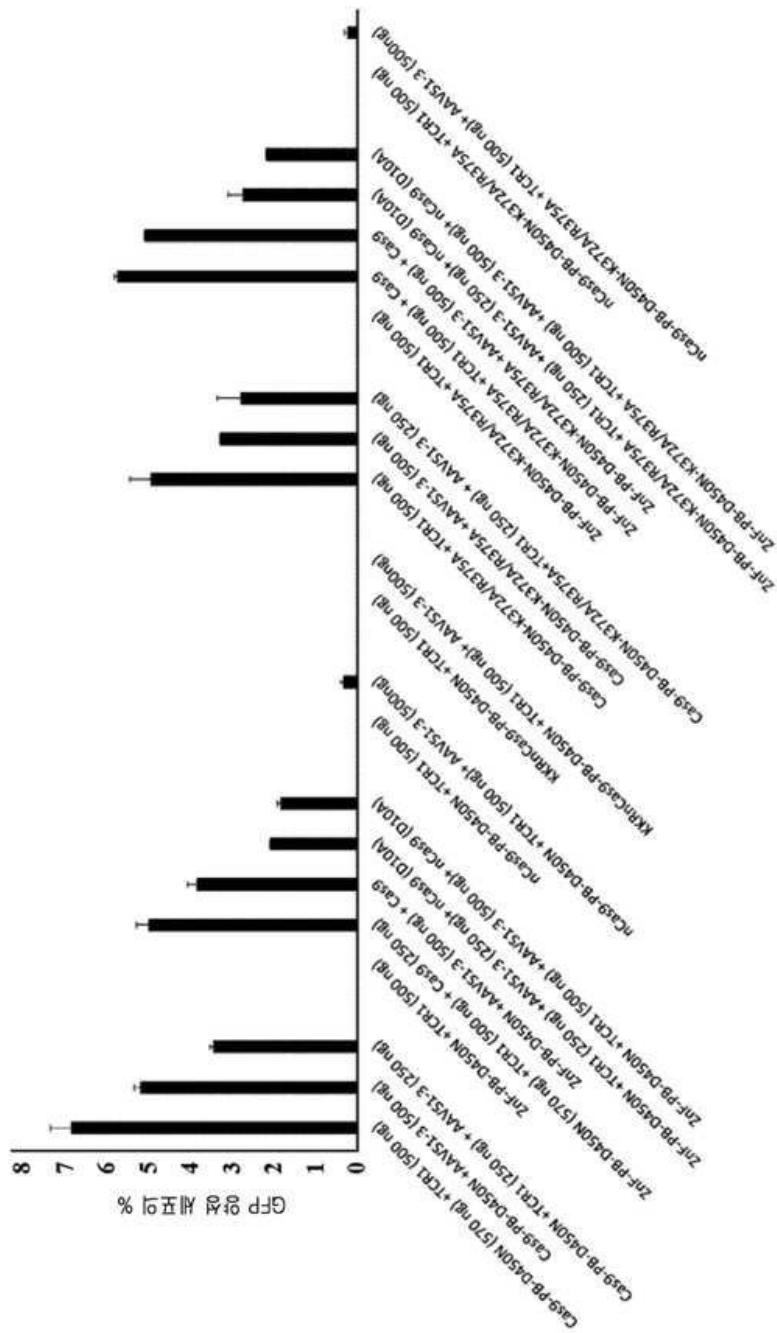
도면11



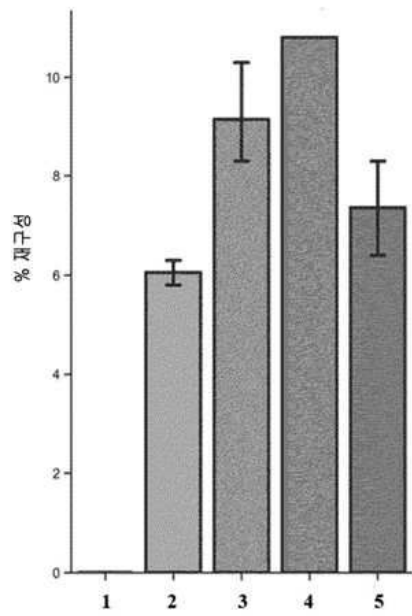
도면12



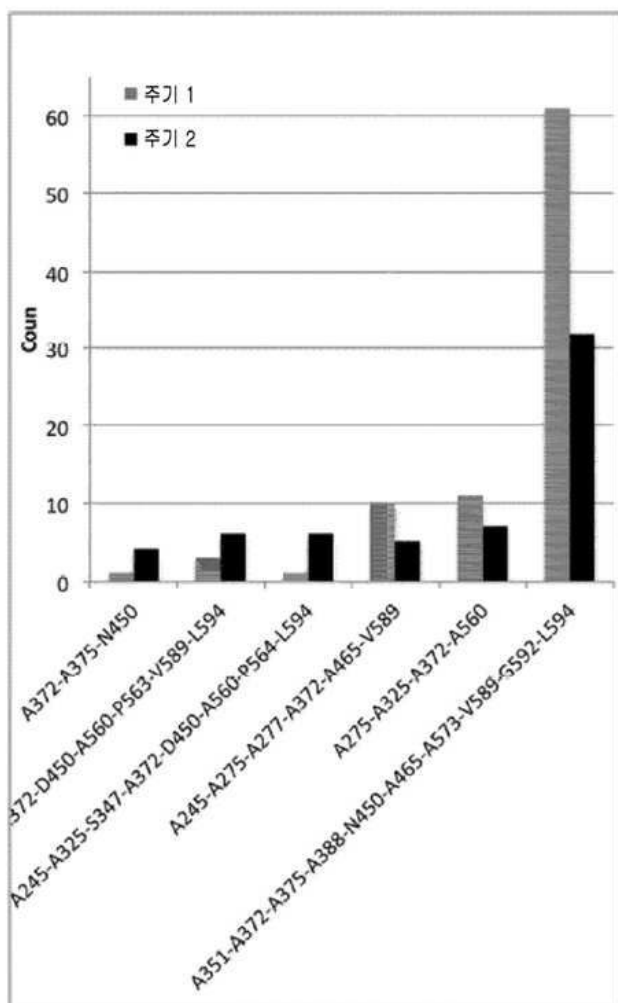
도면13



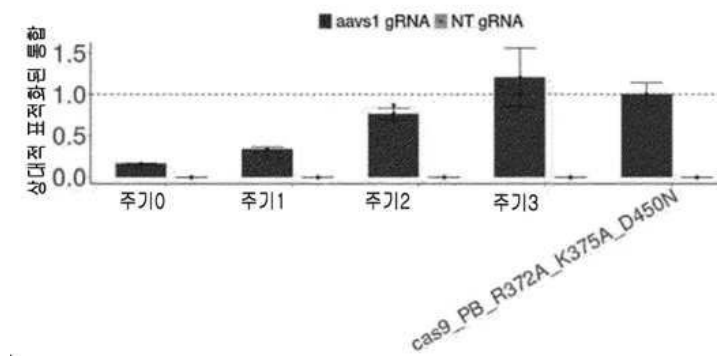
도면14



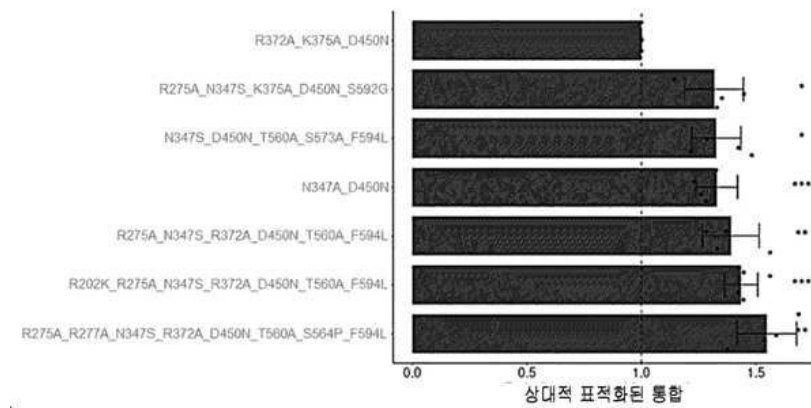
도면15



도면16



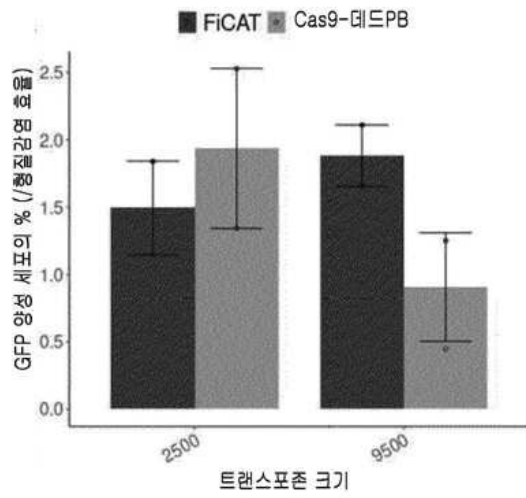
도면17a



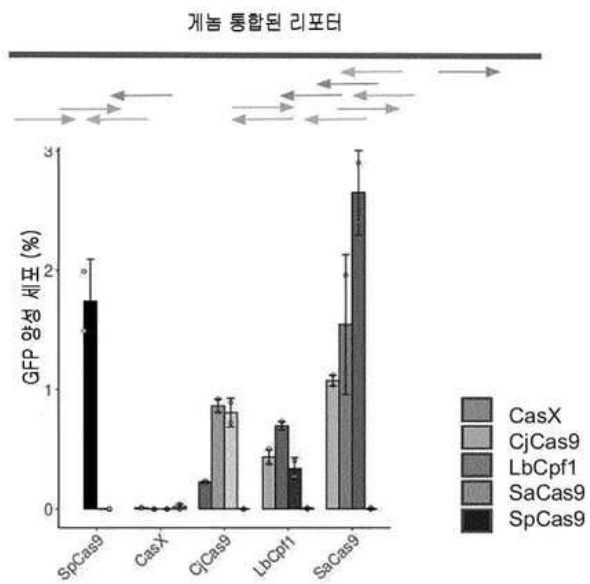
도면17b



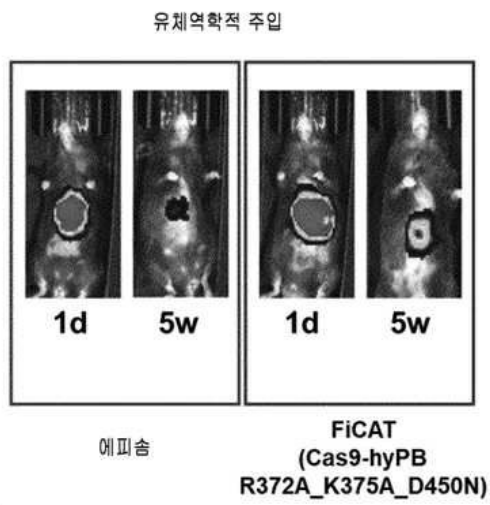
도면18



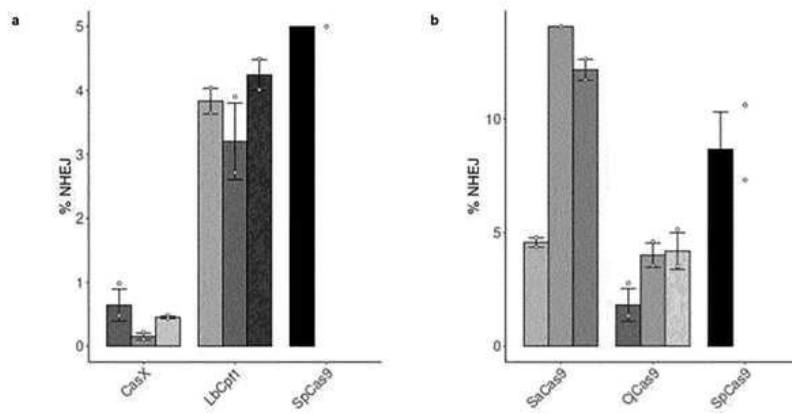
도면19



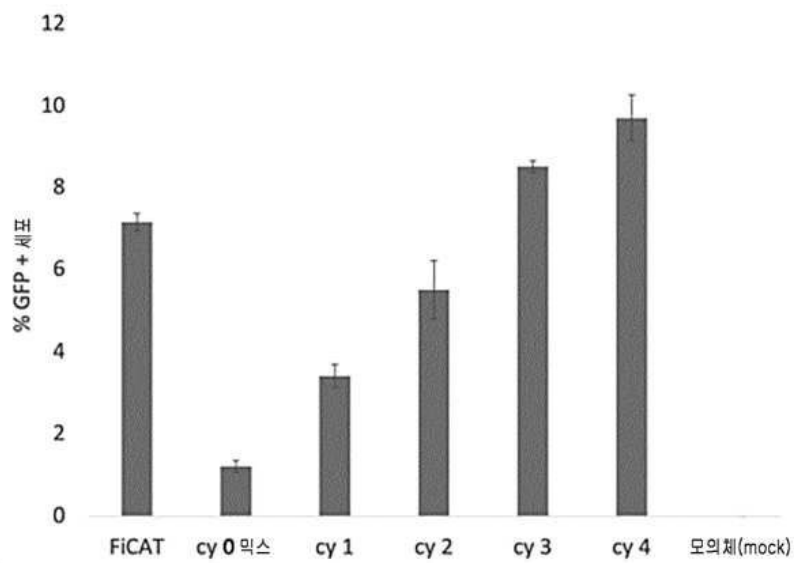
도면20



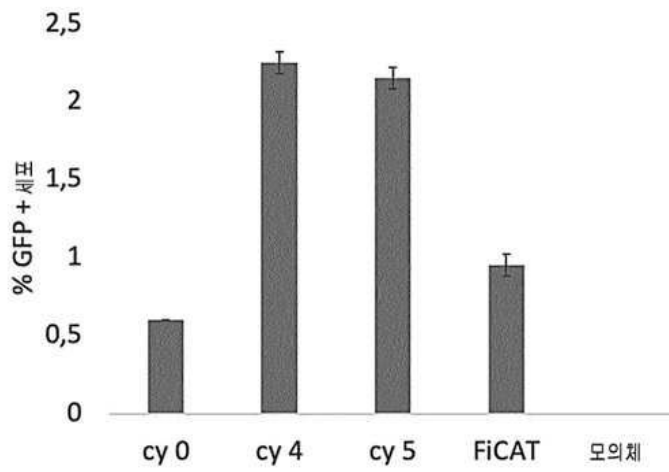
도면21



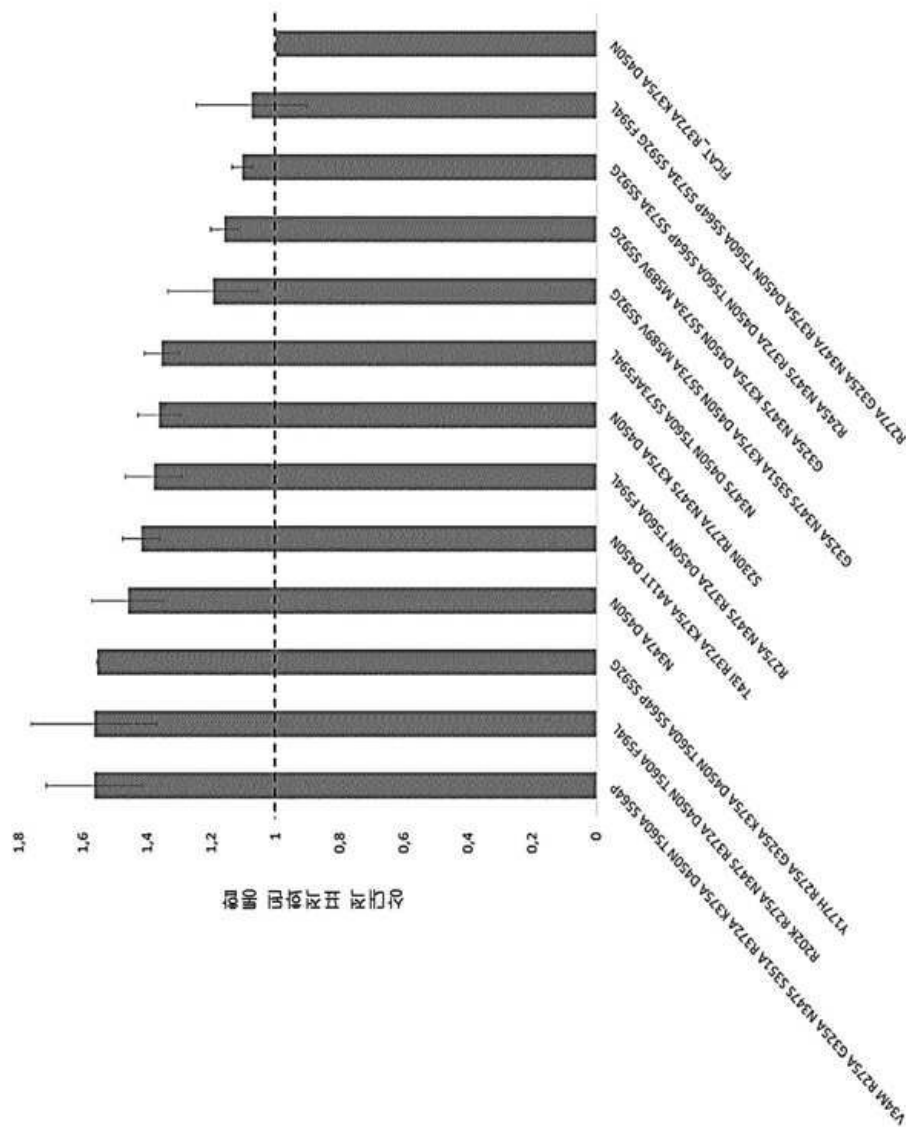
도면22a



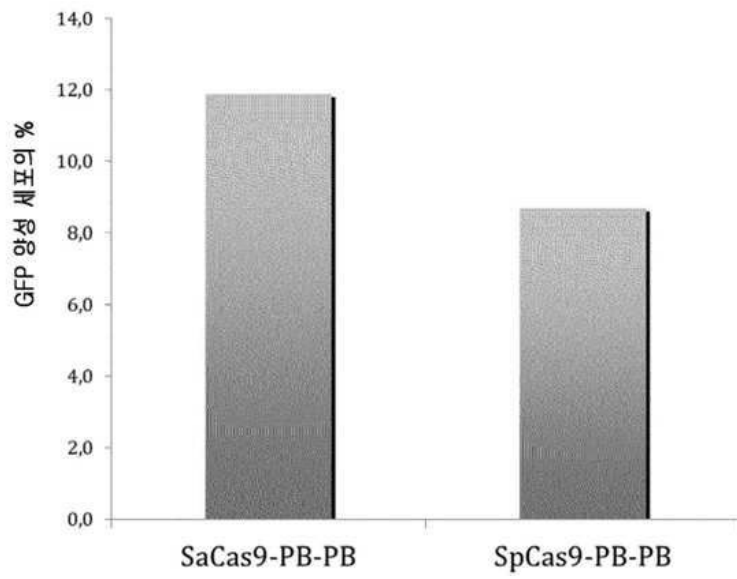
도면22b



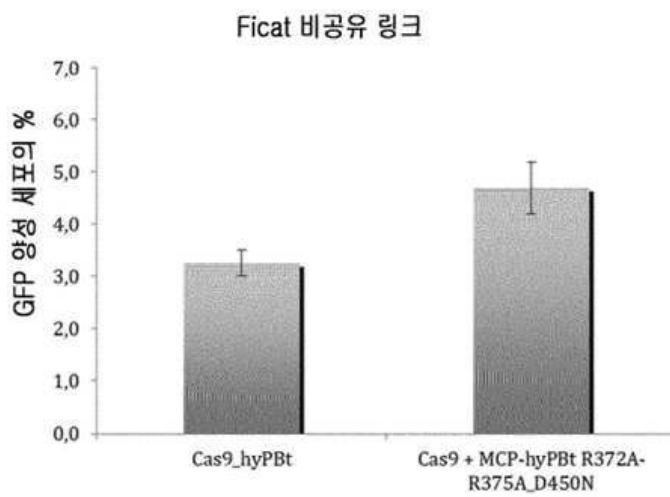
도면23



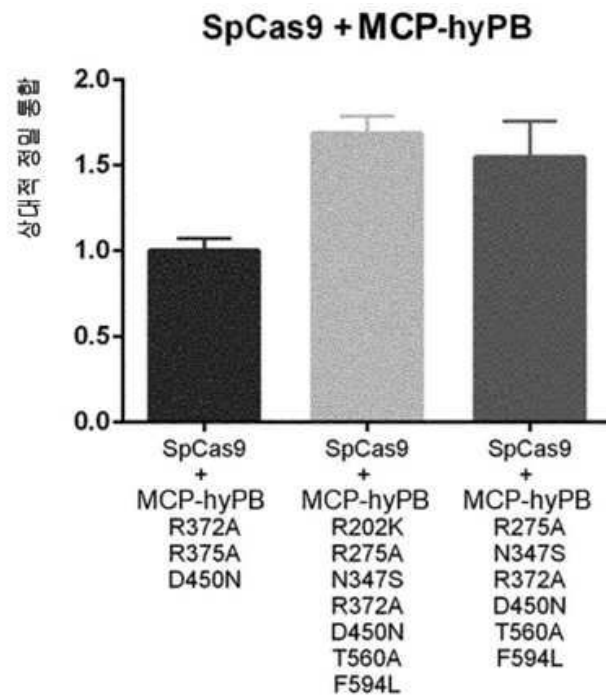
도면24



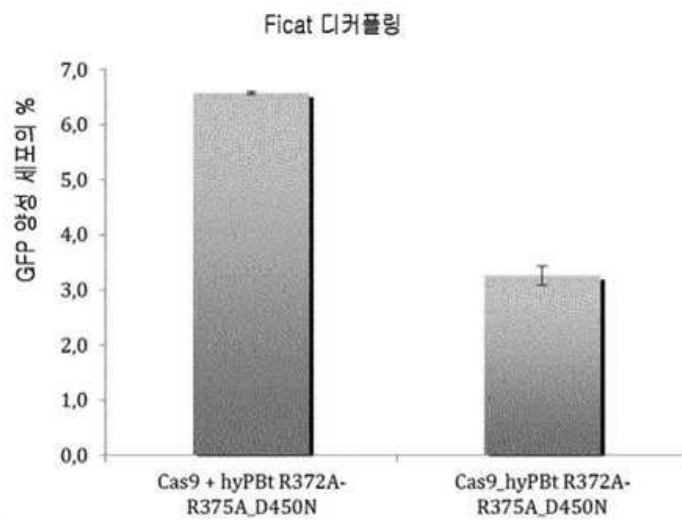
도면25a



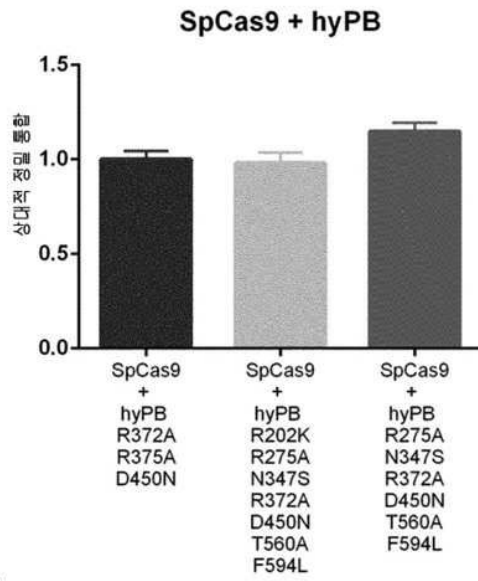
도면25b



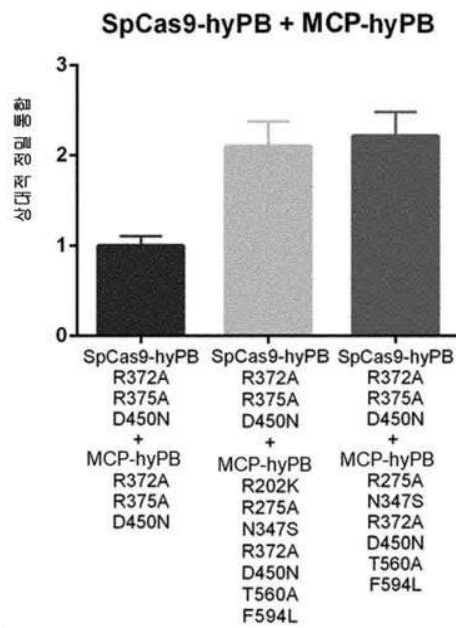
도면26a



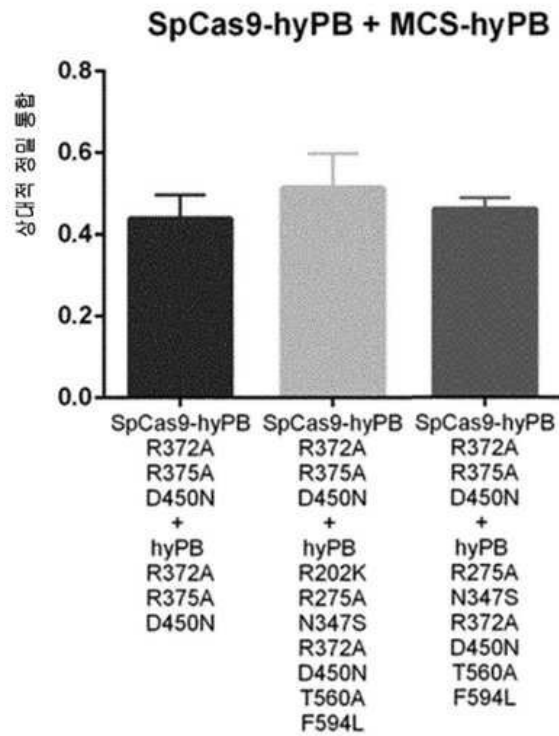
도면26b



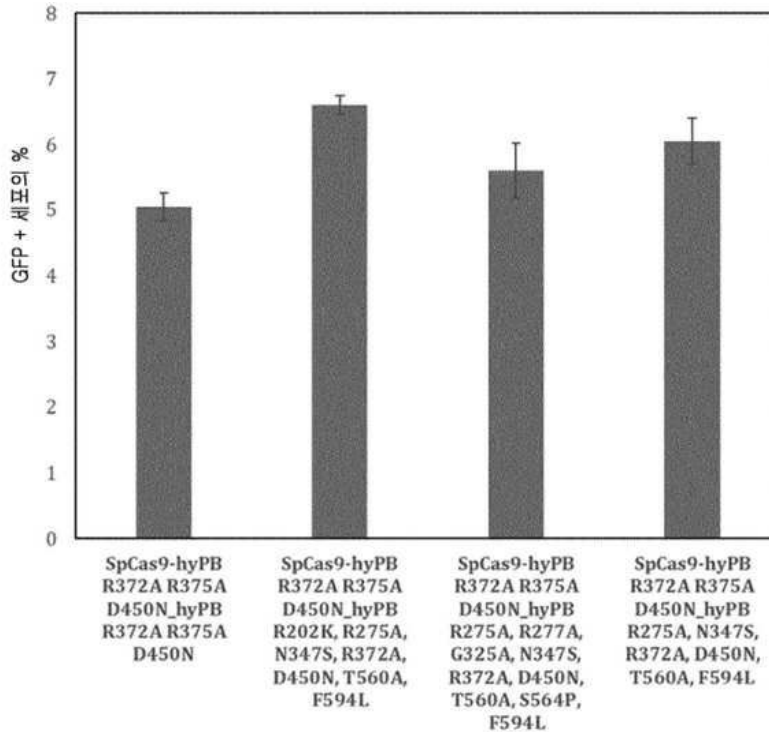
도면27



도면28



도면29



서열 목록

SEQUENCE LISTING

<110> Universitat Pompeu Fabra

<120> PROGRAMMABLE TRANSPOSASES AND USES THEREOF

<130> IBIO-2199/PCT

<150> EP21209719.0

<151> 2021-11-22

<150> EP20214696.5

<151> 2020-12-16

<160> 153

<170> BiSSAP 1.3.6

<210> 1

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified Hyperactive Piggybac

<400> 1

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

210 215 220

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

225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile

340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val

355 360 365

Gly Thr Val Ala Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn

370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro

385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu

405 410 415

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

420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525

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

530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser

580 585 590

Cys Phe

 $\langle 210 \rangle$ 2

<211> 594

 $\langle 212 \rangle$

PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<400> 2

```

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1           5           10           15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
           20           25           30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
           35           40           45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

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

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

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

```

225 230 235 240
 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

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

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

 435 440 445
 Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460
 Ala Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525

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

530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr

545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys

565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Gly

580 585 590

Cys Leu

<210> 3

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified Hyperactive Piggy

<400> 3

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

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

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335
His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
340 345 350
Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365
Gly Thr Val Ala Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400
Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415
Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430
Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445
Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460
Ala Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480
Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495
Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560
Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys

565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Ser
580 585 590
Cys Phe

<210> 4

<211> 594

<212>

PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<400> 4

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

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

115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240
 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Ala Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Ala Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335
 His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
 340 345 350
 Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
 355 360 365
 Gly Thr Val Ala Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
 370 375 380
 Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu

405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

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

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

 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 5

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<400> 5

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240
Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

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

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

 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460
 Ala Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560
Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Arg
580 585 590
Cys Leu

<210> 6

<211> 594

<212>

PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<400> 6

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

85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110
 Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

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

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

 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Ala Gly Ala Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

 305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Ser Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Ala Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Ala Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Ala
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Ser

580 585 590

Cys Leu

<210> 7

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<400> 7

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240
 Phe Thr Pro Val Ala Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Ala Gly Ala Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335
 His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
 340 345 350
 Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
 355 360 365
 Gly Thr Val Ala Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
 370 375 380
 Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys

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

 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495
 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525
 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser

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

<210> 8

<211> 594

<212>

PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<400> 8

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45
 Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60
 Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
 65 70 75 80
 Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
 85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110
 Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
 115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
 130 135 140
 Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240
 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

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

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

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

 500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Arg
580 585 590

Cys Leu

<210> 9

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Hyperactive PiggyBac aa sequence

<400> 9

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

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

 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

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

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

 340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Phe

<210> 10

<211> 594

<212

> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

<220>

<221> SITE

<222> 245..245

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 268..268

<223> Wherein amino acid can be Asp, Asn

<220>

<221> SITE

<222> 275..275

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 277..277

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 287..287

<223> Wherein amino acid can be Ala, Lys

<220>

<221> SITE

<222> 290..290

<223> Wherein amino acid can be Ala, Lys

<220>

<221> SITE

<222> 315..315

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 325..325

<223> Wherein amino acid can be Gly, Ala

<220>

<221> SITE

<222> 341..341

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 346..346

<223> Wherein amino acid can be Asp, Asn

<220>

<221> SITE

<222> 347..347

<223> Wherein amino acid can be Asn, Ala, Ser

<220>

<221> SITE

<222> 350..350

<223> Wherein amino acid can be Thr, Ala

<220>

<221> SITE

<222> 351..351

<223> Xaa can be Ser, Glu, Pro, Ala

<220>

<221> SITE

<222> 356..356

<223> Wherein amino acid can be Lys, Glu

<220>

<221> SITE

<222> 357..357

<223> Wherein amino acid can be Asn, Ala

<220>

<221> SITE
 <222> 372..372
 <223> Wherein amino acid can be Arg, Ala
 <220>
 <221> SITE
 <222> 375..375
 <223> Wherein amino acid can be Lys, Ala
 <220>
 <221> SITE
 <222> 388..388
 <223> Wherein amino acid can be Arg, Ala
 <220>
 <221> SITE
 <222> 409..409
 <223> Wherein amino acid can be Lys, Ala
 <220>
 <221> SITE
 <222> 412..412
 <223> Wherein amino acid can be Lys, Ala
 <220>
 <221> SITE
 <222> 432..432
 <223> Wherein amino acid can be Lys, Ala
 <220>
 <221> SITE
 <222> 460..460
 <223> Wherein amino acid can be Arg, Ala
 <220>
 <221> SITE
 <222> 461..461
 <223> Wherein amino acid can be Ala, Lys
 <220>
 <221> SITE
 <222> 465..465

<223> Wherein amino acid can be Trp, Ala

<220>

<221> SITE

<222> 560..560

<223> Wherein amino acid can be Thr, Ala

<220>

<221> SITE

<222> 564..564

<223> Wherein amino acid can be Ser, Pro

<220>

<221> SITE

<222> 571..571

<223> Wherein amino acid can be Asn, Ser

<220>

<221> SITE

<222> 573..573

<223> Wherein amino acid can be Ser, Ala

<220>

<221> SITE

<222> 576..576

<223> Wherein amino acid can be Lys, Ala

<220>

<221> SITE

<222> 586..586

<223> Wherein amino acid can be His, any naturally occurring amino acid

<220>

<221>

> SITE

<222> 587..587

<223> Wherein amino acid can be Ile, Ala

<220>

<221> SITE

<222> 589..589

<223> Wherein amino acid can be Met, Val

<220>

<221> SITE

<222> 592..592

<223> Wherein amino acid can be Ser, Gly

<220>

<221> SITE

<222> 594..594

<223> Wherein amino acid can be Phe, Leu

<400> 10

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

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

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

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

 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Xaa
 420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445
Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Xaa Xaa Thr Asn Arg
450 455 460
Xaa Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480
Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495
Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Xaa
545 550 555 560
Tyr Cys Pro Xaa Lys Ile Arg Arg Lys Ala Xaa Ala Xaa Cys Lys Xaa
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Xaa Xaa Asp Xaa Cys Gln Xaa
580 585 590
Cys Xaa

<210> 11

<211> 594

<212>

> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations in the catalytic
triad

<220>

<221> SITE

<222> 268..268

<223> Xaa can be Asp, Asn

<220>

<221> SITE

<222> 346..346

<223> Xaa can be Asp, Asn

<400> 11

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Xaa Glu Gln Leu Leu
260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Xaa Asn Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450 455 460
 Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 470 475 480

 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495
 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525
 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 535 540

 Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 550 555 560
 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 12

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations in amino acids that

are critical for excision

<220>

<221> SITE

<222> 287..287

<223> Xaa can be Lys, Ala

<220>

<221> SITE

<222> 290..290

<223> Xaa can be Lys, Ala

<220>

<221> SITE

<222> 460..460

<223> Xaa can be Arg, Ala

<220>

<221> SITE

<222> 461..461

<223> Xaa can be Lys, Ala

<400> 12

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

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

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

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

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

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

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495
Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560
Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Phe

<210> 13

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutation that are involved in

target joining

<220>

<221> SITE

<222> 351..351

<223> Xaa can be Ser, Glu, Pro, Ala

<220>

<221> SITE

<222> 356..356

<223> Xaa can be Lys, Glu

<400> 13

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240

 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300

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

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

 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

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

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

 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 14

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations that are critical
for integration

<220>

<221> SITE

<222> 560..560

<223> Wherein amino acid can be Thr, Ala,

<220>

<221> SITE

<222> 564..564

<223> Wherein amino acid can be Ser, Pro

<220>

<221> SITE

<222> 571..571

<223> Wherein amino acid can be Asn, Ser

<220>

<221> SITE

<222> 573..573

<223> Wherein amino acid can be Ser, Ala

<220>

<221> SITE

<222> 589..589

<223> Wherein amino acid can be Met, Val

<220>

<221> SITE

<222> 592..592

<223> Wherein amino acid can be Ser, Gly

<220>

<221> SITE

<222> 594..594

<223> Wherein amino acid can be Phe, Leu

<400> 14

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

1	5	10	15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp			
	20	25	30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile			
	35	40	45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu			
50	55	60	

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

210 215 220

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

225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

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

 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

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

 515 520 525
 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 535 540
 Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Xaa
 545 550 555 560

Tyr Cys Pro Xaa Lys Ile Arg Arg Lys Ala Xaa Ala Xaa Cys Lys Lys

565

570

575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Xaa Cys Gln Xaa

580

585

590

Cys Xaa

<210> 15

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations that are involved in alignment

<220>

<221> SITE

<222> 325..325

<223> Xaa can be Gly, Ala

<220>

<221> SITE

<222> 347..347

<223> Xaa can be Asn, Ala, Ser

<220>

<221> SITE

<222> 350..350

<223> Xaa can be Thr, Ala

<220>

<221> SITE

<222> 465..465

<223> Xaa can be Trp, Ala

<400> 15

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

1

5

10

15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45
 Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60

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

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

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

 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Xaa Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Xaa Trp Phe Xaa Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Xaa Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

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

 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 16

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations at well conserved
amino acids

<220>

<221> SITE

<222> 576..576

<223> Xaa can be lys, Ala

<220>

<221> SITE

<222> 587..587

<223> Xaa can be Ile, Ala

<400> 16

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

1 5 10 15
 Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60
 Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
 65 70 75 80
 Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
 85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110
 Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
 115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
 130 135 140
 Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240
 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

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

 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

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

 515 520 525
 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 550 555 560
 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Xaa
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Xaa Asp Met Cys Gln Ser

580 585 590
 Cys Phe

<210> 17

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations involved in Zn2+
 binding

<220>

<221> SITE

<222> 586..586

<223> Wherein amino acid can be His, any naturally occurring amino acid

<400> 17

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

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

 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

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

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

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

 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460
 Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495
 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525
 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser

 530 535 540
 Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 550 555 560
 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Xaa Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 18

<211> 594

<212>

> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac with mutations that are involved in
integration

<220>

<221> SITE

<222> 315..315

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 341..341

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 372..372

<223> Wherein amino acid can be Arg, Ala

<220>

<221> SITE

<222> 375..375

<223> Wherein amino acid can be Lys, Ala

<400> 18

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

1	5	10	15
Ser	Asp	Asp	Glu
Leu	Val	Gly	Glu
Asp	Ser	Asp	Ser
Glu	Val	Ser	Asp
20	25	30	
His	Val	Ser	Glu
Asp	Asp	Val	Gln
Ser	Asp	Thr	Glu
Glu	Ala	Phe	Ile
35	40	45	
Asp	Glu	Val	His
Glu	Val	Gln	Pro
Thr	Ser	Ser	Gly
Ser	Glu	Ile	Leu
50	55	60	

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

210 215 220

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

225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Xaa Gly Thr Gln Thr Asn

305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

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

 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

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

 515 520 525
 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 535 540
 Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser

580 585 590

Cys Phe

<210> 19

<211> 897

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from *Corynebacterium ulcerans*

<400> 19

Met Thr Asn Ala Val Ala Asn His His Val Leu Trp Ala Lys Phe Asp

1 5 10 15

Asn Val Ser Glu Pro Tyr Pro Leu Leu Ala His Leu Leu Asp Thr Ala

20 25 30

Thr Ala Ala Thr Cys Leu Phe Asn His Trp Leu Arg Lys Gly Leu Arg

35 40 45

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

50 55 60

Phe Val Ala Gly Ile His Asp Leu Gly Lys Ala Asn Pro Tyr Phe Gln

65 70 75 80

Ala Gln Arg Arg Asn Lys Lys Glu Glu Trp Ile Thr Leu Arg Asp Ala

85 90 95

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

100 105 110

Glu Glu Thr Lys Glu Lys Arg Arg His Glu Asn Ile Thr Leu Ser Ile

115 120 125

Leu Gly Trp Glu Ile Thr Lys Phe Leu Gln Val Lys Asp Val Trp Pro

130 135 140

Gln Leu Ala Ile Ile Gly His His Gly Asn Phe Ser Ala Pro Gly Phe

145 150 155 160
 Leu Ser Asp Glu Asp Asp Leu Glu Asp Ile Glu Asp Ile Phe Asp Asp

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

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

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

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

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

645 650 655
 Glu Val Arg Val Pro Gly Met Val Gly Lys Lys Leu Thr Ile Ala Val
 660 665 670
 Val Asp Ser Pro Ser Thr Gly Gln Thr Leu Pro Tyr Leu Arg Ser Gln

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

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

805 810 815
 Thr Ile Pro Gly Ala Trp Thr Asp Ser Val Glu Lys Leu Ile Ala Ile
 820 825 830
 Ser Ala Lys Asp Ser Glu Ser Leu Arg Arg Ala Leu Arg Ala Ser Ile
 835 840 845
 Ser Ile Pro His Ser Lys Lys Phe Leu Pro Ile Thr Ser Arg Glu Ile
 850 855 860
 Pro Leu Ser Glu Ala Lys Thr Leu Leu Ser Gly Tyr Ser Ala Val His

865 870 875 880
 Ile Gln Pro Asp Glu Tyr Asp Leu Gln Ser Gly Leu Lys Gly Pro Gln
 885 890 895

Lys

<210> 20

<211> 876

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from Corynebacterium diphtheria

<400> 20

Met Asn Pro His Glu Glu Leu Trp Ala Lys Gln Lys Gly Leu Ala Lys

1 5 10 15

Pro Tyr Pro Leu Leu Ala His Leu Leu Asp Ser Ala Ala Val Ala Gly

20 25 30

Ala Leu Trp Asp His Trp Leu Arg Gln Asp Leu Arg Gln Met Phe Ile

35 40 45

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

50 55 60

Ser His Asp Ile Gly Lys Ala Thr Pro Leu Phe Gln Tyr Gln Lys Ala

65 70 75 80

Gln Lys Gly Glu Val Trp Asp Ser Ile Arg Tyr Ala Ile Asp Arg Thr

85 90 95

Gly Arg Tyr Gln Lys Pro Leu Pro Ser Ser Tyr Leu Val Lys Lys Thr

100 105 110

Ser Gly Gly Pro Asn Arg His Glu Gln Trp Ser Ser Phe Ala Ser Lys

115 120 125

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

130 135 140

Leu Ala Ile Gly Gly His His Gly Arg Phe Glu Pro Val Gly Tyr Gly

145 150 155 160

Arg His Gln Arg Lys Ala Ala Glu Asp Leu Ala Lys Ser Gly Trp Ser

165 170 175

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

180 185 190
 Thr Arg Ala Ser Leu Pro Ser Glu Leu Ser Pro Glu Leu Thr Leu Val
 195 200 205
 Leu Ser Gly Leu Thr Ile Leu Ala Asp Arg Ile Ser Ser Thr Glu Ser

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

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

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

405 410 415
 Ala Leu Ala Thr Ala Leu Pro Gly Lys Trp Ile His Leu Arg Ile Leu
 420 425 430

Ala Leu Ala Asn Ala His Ile Val Ile Asp Glu Val His Thr Leu Asp
435 440 445

His Tyr Gln Thr Ala Leu Leu Glu Asn Ile Leu Pro Ile Leu Ala Lys
450 455 460

Leu Lys Thr Lys Ile Thr Phe Leu Thr Ala Thr Met Pro Ser Trp Gln

465 470 475 480

Arg Thr Lys Leu Leu Thr Ala Tyr Gly Gly Glu Asp Leu Gln Ile Pro
485 490 495

Pro Thr Val Phe Pro Ala Ala Glu Thr Val Leu Pro Gly Gln Phe Asn
500 505 510

Arg Thr Leu Ile Asp Ser Asp Ser Thr Thr Ile Asp Phe Thr Met Glu
515 520 525

Glu Thr Ser Tyr Asp His Leu Val Glu Ser His Val Lys Trp His Gln

530 535 540

Thr Thr Arg Leu Asn Ala Pro His Ala Arg Ile Gly Leu Ile Cys Asn

545 550 555 560

Thr Val Lys Arg Ala Gln Glu Ile Ala Ala Ala Leu Glu Lys Thr Asn
565 570 575

Asp Arg Ile Val Leu Leu His Ser Arg Met Thr Thr Glu His Arg Arg
580 585 590

Arg Ser Ala Glu Leu Leu Glu Ser Leu Leu Gly Pro Asn Gly Asn Arg

595 600 605

Lys Thr Ile Thr Val Val Gly Thr Gln Ala Ile Glu Ala Ser Leu Asp
610 615 620

Ile Asp Leu Asp Ile Leu Arg Thr Glu Leu Cys Pro Ala Pro Ser Leu
625 630 635 640

Val Gln Arg Ala Gly Arg Val Trp Arg Arg Asn Asp Pro Tyr Arg Ser
645 650 655

Ser Arg Ile Thr Ala Asp His Lys Pro Ile Ser Val Val Phe Ile Ala

660 665 670

Glu Ala Lys Asp Trp Gln Val Leu Pro Tyr Leu Arg Ala Glu Thr Ser

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

725 730 735
 Met Lys Ala Asp Gly Ala Lys Ala Arg Ile Gln Asp Val Leu Asn Ser
 740 745 750
 Asp Ser Lys Val Ser Asp Phe Ala Leu Leu Thr Ser Lys Asn Glu Ile
 755 760 765
 Asp Glu Ala Gln Thr Arg Leu Ile Glu Glu Gly Thr His Leu Arg Ile
 770 775 780
 Ile Leu Gly Asp Glu Asn Glu Ser Ile Pro Gly Gly Trp Lys His Gly

785 790 795 800
 Leu Ser Ser Leu Leu Lys Leu Lys Ala Ser Asp Arg Glu Ser Leu Arg
 805 810 815
 Thr Ala Leu Leu Ala Ser Ile Pro Leu Leu Val Ser Glu Lys Gln Lys
 820 825 830
 Gln Leu Leu Tyr Gln His Asn Leu Val Pro Leu Ser Ser Ser Lys Thr
 835 840 845
 Val Leu Ala Gly Phe Tyr Phe Leu Pro Lys Ala Gln Asn Phe Tyr Ser

850 855 860
 Lys Asn Leu Gly Phe Ile Trp Pro Glu Glu Lys Asp

865 870 875

<210> 21

<211> 773

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from *Spiroplasma syrophidicola*

<400> 21

Met Asn Tyr Lys Lys Leu Ile Leu Gly Leu Asp Leu Gly Ile Ala Ser
1 5 10 15
Cys Gly Trp Ala Val Thr Gly Gln Met Glu Asp Gly Asn Trp Val Leu
20 25 30
Asp Asp Phe Gly Val Arg Leu Phe Gln Thr Pro Glu Asn Ser Lys Asp
35 40 45
Gly Thr Thr Asn Ala Ala Ala Arg Arg Leu Lys Arg Gly Ala Arg Arg
50 55 60
Leu Ile Lys Arg Arg Lys Asn Arg Ile Lys Asp Leu Lys Asn Leu Phe
65 70 75 80
Glu Lys Ile Asn Phe Ile Asn Lys Ala Ser Leu Asp Lys Tyr Ile Asn
85 90 95
Glu His Ser Ala Thr Asn Leu Val Glu Asp Phe Asn Arg His Glu Leu
100 105 110
Tyr Asn Pro Tyr Phe Leu Arg Ser Ile Gly Ile Thr Glu Lys Leu Thr
115 120 125
Arg Glu Glu Leu Val Trp Ser Leu Ile His Ile Ala Asn Arg Arg Gly
130 135 140
Tyr Lys Asn Lys Phe Ala Phe Asp Ile Glu Gly Asp Gly Lys Lys Arg
145 150 155 160
Glu Thr Lys Leu Asp Glu Ala Ile Ser Asn Ala Leu Ile Ser Ser Asn
165 170 175
Leu Thr Ile Ser Gln Glu Ile Val Arg Asn Lys Lys Phe Arg Asp Ala
180 185 190
Lys Asn Lys Lys Ala Leu Leu Val Arg Asn Lys Gly Gly Lys Glu Gly
195 200 205
Glu Asn Asn Phe Gln Phe Leu Phe Ala Arg Asp Asp Tyr Lys Lys Glu
210 215 220
Val Asp Leu Leu Leu Ala Lys Gln Ala Lys Phe Tyr Pro Glu Leu Thr
225 230 235 240
Glu Glu Ile Arg Ala Lys Ala Ala Asp Ile Ile Phe Arg Gln Arg Asp

245 250 255
 Phe Glu Asp Gly Pro Gly Pro Lys Lys Gln Glu Leu Arg Glu Ile Tyr
 260 265 270
 Lys Lys Glu Asn Lys Gln Phe Ser Lys Asn Phe Thr Gln Leu Glu Gly
 275 280 285

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

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

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

 Ile Lys Glu Asp Asn Phe Lys Leu Pro Glu Ser Tyr Ala Lys Gln Tyr
 485 490 495

Gln Lys Tyr Leu Thr Gly Glu Lys Thr Phe Glu Met Phe Ala Pro Ile
500 505 510

Ile Asp Pro Asp Leu Trp Arg Asn Pro Ile Val Phe Arg Ala Ile Asn
515 520 525

Gln Ala Arg Lys Val Ile Lys Lys Leu Phe Glu Lys Tyr Thr Phe Ile
530 535 540

Asp Gln Ile Asn Ile Glu Leu Thr Arg Glu Met Gly Leu Ser Phe Ser
545 550 555 560

Asp Arg Lys Lys Val Lys Glu Arg Gln Asp Asp Ser Leu Lys Glu Asn
565 570 575

Ala Lys Ala Lys Glu Phe Leu Met Ala Asn Gly Ile Ile Val Asn Asp
580 585 590

Thr Asn Val Leu Lys Tyr Lys Leu Trp Ile Gln Gln Asn Lys Lys Ser
595 600 605

Leu Tyr Ser Gly Lys Glu Ile Thr Ile Ala Asp Leu Gly Ala Ser Asn
610 615 620

Val Leu Gln Ile Asp His Ile Ile Pro Tyr Ser Lys Leu Ala Asp Asp
625 630 635 640

Ser Phe Asn Asn Lys Val Leu Val Phe Ser Lys Glu Asn Gln Glu Lys
645 650 655

Gly Asn Gln Phe Ala Asp Gln Tyr Val Lys Ser Leu Gly Thr Glu Asn
660 665 670

Tyr Asn Asn Tyr Lys Lys Arg Val Asn Tyr Leu Leu Phe Gln Asn Gln
675 680 685

Ile Asn Gln Lys Lys Ala Glu Tyr Leu Leu Cys Ser Asn Gln Asn Glu
690 695 700

Glu Ile Leu Asn Asp Phe Val Ser Arg Asn Leu Asn Asp Thr Arg Tyr
705 710 715 720

Ile Thr Arg Tyr Val Thr Asn Trp Leu Lys Ala Glu Phe Glu Leu Gln
725 730 735

Ser Arg Phe Gly Leu Ala Lys Pro Lys Ile Met Thr Leu Asn Gly Ala

740 745 750
 Ile Thr Ser Arg Phe Arg Arg Thr Trp Leu Arg Asn Ser Pro Trp Gly
 755 760 765
 Leu Glu Lys Lys Ser
 770
 <210> 22
 <211> 1380
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Cas9 from Prevotella intermedia
 <400> 22
 Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Thr Ser Ile Gly Trp
 1 5 10 15

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

 Leu His Asp Cys Leu Gln Lys Gln Gly Trp Leu Gly Thr Glu Ala Met
 85 90 95
 Tyr Glu Glu Gly Lys Ala Ser Thr Phe Glu Thr Tyr Lys Leu Arg Ala
 100 105 110
 Lys Ala Ala Glu Glu Glu Ile Ser Leu His Glu Phe Ala Arg Val Leu
 115 120 125
 Phe Met Leu Asn Lys Lys Arg Gly Tyr Lys Ser Asn Arg Lys Ala Asn
 130 135 140

 Asn Lys Glu Asp Gly Gln Leu Phe Asp Gly Met Thr Ile Ala Lys Lys
 145 150 155 160

Leu Tyr Glu Glu His Leu Thr Pro Ala Glu Tyr Ser Leu Gln Leu Leu
 165 170 175
 Asn Lys Gly Lys Lys Phe Thr Gln Gly Tyr Tyr Arg Ser Asp Leu Asn
 180 185 190
 Ala Glu Leu Glu Arg Ile Trp Asp Glu Gln Lys Lys Tyr Tyr Pro Glu
 195 200 205

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

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

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

 Pro Lys Ser Ser Pro Val Phe Gln Glu Phe Arg Ile Trp Gln Asn Leu

405 410 415
 Asn Asn Val Leu Leu Ile Asp Asn Asp Thr Asn Glu Lys Arg Pro Leu
 420 425 430
 Asp Glu Val Glu Arg Asn Leu Leu Tyr Lys Glu Leu Ser Ile Lys Ala
 435 440 445
 Lys Leu Ser Lys Thr Glu Ala Leu Lys Ile Leu Asn Lys Lys Gly Lys
 450 455 460

 Gln Trp Asp Leu Asn Tyr Arg Glu Leu Glu Gly Asn Arg Thr Gln Ala
 465 470 475 480
 Ile Leu Phe Asp Cys Tyr Asn Arg Ile Ile Thr Leu Thr Gly His Glu
 485 490 495
 Glu Cys Asp Phe Lys Lys Ile Lys Ala Ser Glu Ile Arg His Tyr Val
 500 505 510
 Ser Thr Ile Phe Lys Asn Leu Gly Phe Ser Thr Glu Ile Leu Asp Phe
 515 520 525

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

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

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

 Leu Ser Tyr Ile Ser Arg Asn Asp Ile Ile Lys Tyr Lys Leu Tyr Glu
 725 730 735
 Glu Leu Glu Pro Asn Tyr Tyr Lys Thr Leu Tyr Ser Asp Thr Tyr Ile
 740 745 750
 Thr Lys Asp Lys Leu Phe Ser Lys Asp Phe Asp Ile Glu His Ile Ile
 755 760 765
 Pro Lys Ala Arg Leu Phe Asp Asp Ser Phe Ser Asn Lys Thr Leu Glu
 770 775 780

 Ala Arg Asn Ile Asn Leu Glu Lys Ser Asn Lys Thr Ala Phe Asp Phe
 785 790 795 800
 Ile Lys Glu Lys Tyr Gly Glu Asp Gly Ala Glu Ala Tyr Lys Lys Lys
 805 810 815
 Leu Asp Met Leu Leu Glu Asn Asp Ala Ile Ser Arg Pro Lys Tyr Asn
 820 825 830
 Asn Leu Leu Arg Ala Glu Ala Asp Ile Pro Ser Asp Phe Ile Asn Arg
 835 840 845

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

900	905	910	
Asp Gly Arg Lys Ile Lys Arg Ile Lys Asp Trp Thr Lys Arg Asn Asp			
915	920	925	
His Arg His His Ala Met Asp Ala Leu Ala Ile Ala Phe Thr Lys Pro			
930	935	940	
Ser Phe Ile Gln Tyr Leu Asn Asn Leu Asn Ala Arg Ser Asn Lys Gly			
945	950	955	960
Asp Ser Ile Tyr Ala Ile Glu Asn Lys Glu Leu His Tyr Glu Glu Gly			
965	970	975	
Lys Leu Arg Phe Asn Ala Pro Ile Pro Val Asn Glu Phe Arg Ala Glu			
980	985	990	
Ala Lys Arg His Leu Ser Ala Ile Leu Val Ser Ile Lys Ala Lys Asn			
995	1000	1005	
Lys Val Met Thr Gln Asn Val Asn Lys Ile Lys Thr Lys His Gly Ile			
1010	1015	1020	
Ile Lys Lys Ile Gln Leu Thr Pro Arg Gly Pro Leu His Asn Glu Thr			
1025	1030	1035	1040
Ile Tyr Gly Thr Lys Met Arg Pro Ile Ile Lys Met Val Lys Val Gly			
1045	1050	1055	
Ala Ala Leu Asp Glu Ala Thr Ile Asn Lys Val Ser Ser Pro Ala Ile			
1060	1065	1070	
Arg Glu Ala Leu Leu Lys Arg Leu Asn Glu Tyr Ser Gly Asn Ala Lys			
1075	1080	1085	
Lys Ala Phe Thr Gly Lys Asn Thr Leu Glu Lys Asn Pro Ile Tyr Leu			
1090	1095	1100	
Asn Ala Gly Arg Thr Lys Thr Val Pro Ser Leu Val Lys Thr Val Glu			
1105	1110	1115	1120
Trp Glu Ser Phe His Pro Thr Arg Lys Leu Ile Asp Lys Asp Leu Asn			
1125	1130	1135	
Val Asp Lys Val Val Asp Lys Gly Ile Arg Glu Ile Leu Lys Ala Arg			
1140	1145	1150	

Leu Glu Glu Phe Asn Gly Asp Ala Lys Lys Ala Phe Ser Asn Leu Glu
1155 1160 1165

Glu Asn Pro Ile Tyr Leu Asp Glu Ala Lys Lys Ile Ala Leu Lys Arg
1170 1175 1180

Val Ser Ile Glu Gly Val Leu Ser Ala Ile Pro Leu His Thr Leu Lys
1185 1190 1195 1200

Asn Gln Ala Gly Lys Pro Ile Thr Gly Lys Asp Gly Lys Pro Val Leu
1205 1210 1215

Gly Asn Tyr Val Gln Thr Ser Asn Asn His His Ile Ala Phe Tyr Tyr
1220 1225 1230

Asp Glu Asp Gly Asn Leu Gln Asp Asn Ala Val Ser Phe Phe Glu Ala
1235 1240 1245

Ala Glu Arg Lys Ser Gln Gly Ile Pro Val Ile Asp Lys Asp Tyr Asn
1250 1255 1260

Arg Asp Lys Gly Trp Arg Phe Leu Phe Thr Met Lys Gln Asn Glu Tyr
1265 1270 1275 1280

Phe Val Phe Pro Asn Glu Ala Thr Gly Phe Ile Pro Ser Glu Val Asp
1285 1290 1295

Leu Thr Asp Glu Ala Asn Tyr Gly Ile Ile Ser Pro Asn Leu Tyr Arg
1300 1305 1310

Val Gln Lys Val Ser Arg Ile Asp Lys Gly Thr Ser Ala Ser Arg Asp
1315 1320 1325

Tyr Trp Phe Arg His His Leu Glu Thr Ile Leu Asn Asp Asp Ala Lys
1330 1335 1340

Leu Lys Asn Leu Ala Phe Lys Arg Ile Arg Gly Leu Leu Glu Leu Lys
1345 1350 1355 1360

Asp Ile Ile Lys Val Arg Ile Asn Ser Thr Gly Lys Ile Val Ala Val
1365 1370 1375

Gly Glu Tyr Asp
1380

<210> 23

<211> 535

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from *Spiroplasma taiwanense*

<400> 23

Met Trp Ser Arg Lys Ile Leu Lys Ala Gly Ser Arg Leu Phe Asp Glu

1 5 10 15

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

20 25 30

Arg Arg Arg Asn Leu Arg Arg Lys Ile Thr Trp Lys Gln Asp Leu Ile

35 40 45

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

50 55 60

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

65 70 75 80

Asn Ser Lys Ile Glu Leu Glu Gln Leu Leu Ile Ile Leu Phe Asn Tyr

85 90 95

Ile Lys His Arg Gly Ser Phe Asn Tyr Arg Glu Asp Leu Ser Glu Leu

100 105 110

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

115 120 125

Pro Val Asp Ile Gln Phe Glu Leu Lys Glu Glu Asn Asn Lys Phe Arg

130 135 140

Glu Ile Asn Asn Glu Lys Ser Leu Ile Asn His Glu Trp Tyr Val Lys

145 150 155 160

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

165 170 175

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

180 185 190

Tyr Tyr Asp Gly Pro Gly Pro Lys Asp Lys Asn Leu Leu Asn Pro Ser

195 200 205

Lys Tyr Gly Trp Lys Asn Gln Glu Glu Phe Phe Asp Arg Phe Ala Gly
 210 215 220

Lys Asp Thr Tyr Asp Ser Lys Glu Gln Arg Ala Pro Lys His Ser Leu
 225 230 235 240

Thr Ser Tyr Leu Phe Asn Ile Leu Asn Asp Leu Asn Asn Leu Ser Ile
 245 250 255

Asn Gly Asp Arg Asn Gln Leu Thr Tyr Glu Asn Lys Lys Asp Leu Ile
 260 265 270

Asn Leu Thr Leu Ile Asn Gln Lys Glu Lys Ala Glu Asn Ile Thr Leu
 275 280 285

Lys Lys Ile Ala Lys Tyr Leu Lys Ile Asn Glu Lys Asn Ile Thr Gly
 290 295 300

Tyr Arg Leu Lys Pro Asn Ser Asn Glu Ser Ile Phe Thr Val Phe Glu
 305 310 315 320

Ser Ala Asn Lys Met Arg Ser Ile Leu Val Lys Asn Asn Lys Ser Ile
 325 330 335

Asp Phe Ile Cys Leu Glu Asn Ile Asp Lys Ile Asp Lys Ile Val Asp
 340 345 350

Ile Leu Thr Lys Tyr Gln Ser Ile Glu Asp Lys Ser Leu Lys Leu Glu
 355 360 365

Glu Leu Asn Phe Asp Phe Phe Asp Lys Glu Thr Cys Glu Lys Leu Ala
 370 375 380

Val Ile Ser Leu Thr Gly Thr His Ala Leu Ser Lys Lys Thr Met Ser
 385 390 395 400

Lys Leu Ile Glu Glu Met Phe His Asp Asn Leu Asn His Met Glu Ala
 405 410 415

Leu Ala Lys Leu Lys Ile Lys Pro Asp Tyr Lys Leu Lys Val Asp Leu
 420 425 430

Thr Asn Phe Lys Thr Ile Pro Ile Leu Arg Glu Lys Ile Asn Glu Met
 435 440 445

Tyr Ile Ser Pro Val Val Lys Arg Ala Leu Ile Glu Ser Leu Lys Ile

450 455 460
 Ile Lys Glu Leu Glu Arg His Phe Lys Asp Phe Glu Ile Lys Asp Ile
 465 470 475 480

 Val Ile Glu Met Ala Lys Lys Asn Ser Ala Glu Lys Lys Gln Phe Ile
 485 490 495
 Ser Lys Ile Gln Arg Gln Asn Val Asp Leu Val Lys Lys Leu Ser Asn
 500 505 510
 Asp Tyr Ser Leu Asp Glu Asn Lys Leu Asn Phe Lys Met Lys Glu Lys
 515 520 525
 Phe Leu Leu Leu Ser Glu Gln
 530 535

 <210> 24
 <211> 1281
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> Cas9 from Streptococcus iniae
 <400> 24

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

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

Phe Glu Asn His Pro Ile Phe Gly Asn Leu Lys Asp Glu Ile Thr Tyr

115 120 125

His Asn Asp Tyr Pro Thr Ile Tyr His Leu Arg Gln Thr Leu Ala Asp

130 135 140

Ser Asp Gln Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

Glu Lys Ile Leu Thr Ser Lys Thr Ser Lys Ser Arg Arg Leu Glu Asn

210 215 220

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

225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

Leu Phe Ile Ala Ala Lys Lys Leu Ser Asp Ala Ile Leu Leu Ser Asp

290 295 300

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

305 310 315 320

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

325 330 335

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

340 345 350

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

355 360 365
Gln Glu Glu Phe Tyr Lys Tyr Ile Lys Pro Ile Leu Leu Lys Leu Asp

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

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

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

565 570 575
Thr Val Thr Ile Leu Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590
Thr Tyr His Asp Leu Leu Lys Ile Phe Lys Asp Lys Ala Phe Leu Asp
595 600 605

Asp Glu Ala Asn Gln Asp Ile Leu Glu Glu Ile Val Trp Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Gln Ala Met Ile Glu Arg Arg Leu Val Lys Tyr Ala

 625 630 635 640
 Asp Val Phe Glu Lys Ser Val Leu Lys Lys Leu Lys Lys Arg His Tyr
 645 650 655
 Thr Gly Trp Gly Arg Leu Ser Gln Lys Leu Ile Asn Gly Ile Lys Asp
 660 665 670
 Lys Gln Thr Gly Lys Thr Ile Leu Gly Phe Leu Lys Asp Asp Gly Val
 675 680 685
 Ala Asn Arg Asn Phe Met Gln Leu Ile Asn Asp Ser Ser Leu Asp Phe

 690 695 700
 Ala Lys Ile Ile Lys Asn Glu Gln Glu Lys Thr Ile Lys Asn Glu Ser
 705 710 715 720
 Leu Glu Glu Thr Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys
 725 730 735
 Gly Ile Leu Gln Ser Ile Lys Ile Val Asp Glu Ile Val Lys Ile Met
 740 745 750
 Gly Gln Asn Pro Asp Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

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

 820 825 830
 Leu Ser Gln Tyr Asp Ile Asp His Ile Ile Pro Gln Ser Phe Ile Lys
 835 840 845
 Asp Asn Ser Ile Asp Asn Thr Val Leu Thr Thr Gln Ala Ser Asn Arg

850 855 860
 Gly Lys Ser Asp Asn Val Pro Asn Ile Glu Thr Val Asn Lys Met Lys
 865 870 875 880
 Ser Phe Trp Tyr Lys Gln Leu Lys Ser Gly Ala Ile Ser Gln Arg Lys

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

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

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

1075 1080 1085
 Tyr Pro Gln Val Asn Ile Val Met Lys Thr Glu Val Gln Thr Gly Gly
 1090 1095 1100

Phe Ser Lys Glu Ser Ile Trp Pro Lys Gly Asp Ser Asp Lys Leu Ile
 1105 1110 1115 1120
 Ala Arg Lys Lys Ser Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser
 1125 1130 1135
 Pro Ile Ile Ala Tyr Ser Val Leu Val Val Ala Lys Ile Ala Lys Gly
 1140 1145 1150
 Lys Thr Gln Lys Leu Lys Thr Ile Lys Glu Leu Val Gly Ile Lys Ile
 1155 1160 1165
 Met Glu Gln Asp Glu Phe Glu Lys Asp Pro Ile Ala Phe Leu Glu Lys
 1170 1175 1180
 Lys Gly Tyr Gln Asp Ile Gln Thr Ser Ser Ile Ile Lys Leu Pro Lys
 1185 1190 1195 1200
 Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Leu Leu Ala Ser
 1205 1210 1215
 Ala Lys Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Asn Lys Tyr
 1220 1225 1230
 Val Lys Phe Leu Tyr Leu Ala Ser His Tyr Thr Lys Phe Thr Gly Lys
 1235 1240 1245
 Glu Glu Asp Arg Glu Lys Lys Arg Ser Tyr Val Glu Ser His Leu Tyr
 1250 1255 1260
 Tyr Phe Asp Val Arg Leu Ser Gln Val Phe Arg Val Thr Asn Val Glu
 1265 1270 1275 1280
 Phe

<210> 25

<211> 1352

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from *Belliella baltica*

<400> 25

Met Lys Lys Ile Leu Gly Leu Asp Leu Gly Thr Thr Ser Ile Gly Trp

1 5 10 15
 Ala Phe Ile Lys Glu Pro Glu Lys Asp Val Val Gly Ser Glu Ile Val
 20 25 30
 Asp Met Gly Val Arg Ile Val Pro Leu Ser Ser Asp Glu Glu Asn Asp

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

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

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

 225 230 235 240
 Glu Asn Lys Gly Lys Arg Glu Glu Arg Arg Leu Gln Glu Tyr Lys Trp
 245 250 255

Arg Ala Glu Ala Val Asn Phe Lys Ile Asp Leu Ser Glu Ile Ala Leu
260 265 270

Ile Leu Ser Gln Ile Asn Ser Gln Ile Ser Asn Ser Ser Gly Tyr Leu
275 280 285

Gly Ala Ile Ser Asp Arg Ser Lys Glu Leu Tyr Phe Lys Asn Leu Thr
290 295 300

Val Gly Gln Tyr Leu Tyr Gln Gln Ile Lys Lys Asn Pro His Thr Arg
305 310 315 320

Leu Lys Gly Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe Glu
325 330 335

Arg Ile Trp Ser Val Gln Ser Ser Phe Tyr Pro Gln Leu Asn Asp Ala
340 345 350

Leu Lys Arg Glu Val Arg Asp Ile Thr Ile Phe Phe Gln Arg Arg Leu
355 360 365

Lys Ser Gln Lys His Leu Ile Ser Asn Cys Glu Phe Glu Asp His His
370 375 380

Lys Val Val Pro Lys Ser His Pro Val Phe Gln Glu Phe Arg Ile Trp
385 390 395 400

Gln Asn Leu Asn Asn Leu Leu Leu Ile Lys Lys Asp Asn Leu Asn Glu
405 410 415

Lys Phe Asp Leu Glu Leu Glu Ser Lys Ile Ala Leu Ala Asn Glu Leu
420 425 430

Ala Phe Lys Arg Glu Leu Asn Val Lys Asp Ala Leu Lys Ile Leu Gly
435 440 445

Leu Lys Pro Asn Glu Trp Glu Phe Asn Phe Thr Lys Ile Glu Gly Asn
450 455 460

Arg Thr Asn Gln Ala Phe Phe Asp Ala Phe Ala Lys Ile Ile Glu Leu
465 470 475 480

Glu Asp Gly Glu Pro Ile Asp Leu Gly Asp Leu Lys Ala Asp Asp Ile
485 490 495

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

500 505 510
 Leu Leu Gln Val Asn Ser Asp Ile Glu Gly Ala Glu Tyr Glu Lys Gln
 515 520 525
 Ser Tyr Ile Gln Phe Trp His Leu Leu Tyr Ser Ser Glu Asp Asp Gln
 530 535 540
 Lys Leu Lys Leu Asn Leu Ile Arg Lys Phe Gly Phe Lys Pro Glu His

 545 550 555 560
 Ala Lys Ile Leu Ala Ser Ile Ser Leu Gln Asp Asp His Ala Ser Leu
 565 570 575
 Ser Ser Arg Ala Ile Lys Lys Ile Leu Pro His Leu Gln Ser Gly Leu
 580 585 590
 Ile Tyr Asp Lys Ala Cys Thr Tyr Ala Gly Tyr Asn His Ser Ser Ser
 595 600 605
 Phe Thr Lys Asp Glu Asn Glu Lys Arg Glu Leu Arg Ala Glu Leu Glu

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

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

 740 745 750

Leu Phe Asp Asp Ser Phe Ser Asn Lys Thr Ile Cys Glu Arg Gln Leu
 755 760 765
 Asn Ile Asp Lys Ala Asn Val Thr Ala Phe Ser Phe Leu Gln Asn Lys
 770 775 780
 Leu Ser Ala Asp Glu Phe Glu Gln Tyr Gln Ser Arg Val Lys Ser Leu
 785 790 795 800
 Tyr Gly Lys Leu Ser Lys Ala Lys Ile Gln Lys Leu Leu Met Ala Asn
 805 810 815
 Asp Lys Ile Pro Glu Asp Phe Ile Ala Arg Gln Leu Gln Glu Thr Arg
 820 825 830
 Tyr Ile Ser Lys Lys Ala Lys Glu Ile Leu Phe Glu Ile Ser Arg Arg
 835 840 845
 Val Ser Val Thr Thr Gly Thr Ile Thr Asp Lys Leu Arg Glu Asp Trp
 850 855 860
 Gly Leu Val Glu Ile Met Lys Glu Leu Asn Trp Glu Lys Tyr Asp Lys
 865 870 875 880
 Leu Gly Leu Thr Tyr Thr Ile Glu Gly Lys His Gly Glu Arg Leu Asn
 885 890 895
 Lys Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala Met
 900 905 910
 Asp Ala Leu Thr Val Ala Leu Thr Lys Pro Ala Tyr Ile Gln Tyr Leu
 915 920 925
 Asn Asn Leu Asn Ala Lys Gly Leu Asn Asn Lys Lys Gly Thr Glu Val
 930 935 940
 Phe Ala Ile Glu Gln Lys Tyr Leu Lys Arg Glu Asn Gly Lys Leu Cys
 945 950 955 960
 Phe Ile Pro Pro Ile Glu Asn Ile Arg Ser Glu Ala Lys Lys His Leu
 965 970 975
 Ser Arg Ile Leu Val Ser Tyr Lys Ala Lys Asn Lys Val Val Thr Ile
 980 985 990
 Asn Lys Asn Lys Thr Lys Ser Lys Ala Gly Leu Asn Glu Gln Ile Ala

995	1000	1005	
Leu Thr Pro Arg Gly Gln	Leu His Lys Glu Thr Val Tyr Gly Lys Ser		
1010	1015	1020	
Phe His Tyr Ser Thr Lys Phe Glu Lys Ile Gly Ala Ser Phe Asn Val			
1025	1030	1035	1040
Gln Lys Ile Asn Thr Val Ala Lys Lys Glu Glu Arg Glu Ala Leu Leu			
1045	1050	1055	
Lys Arg Leu Ala Glu Asn Gly Asn Asp Pro Lys Lys Ala Phe Thr Gly			
1060	1065	1070	
Lys Asn Thr Leu Asn Lys Met Pro Ile Tyr Leu Asp Leu Gly Lys Asn			
1075	1080	1085	
Ile Lys Leu Ser Glu Lys Val Lys Thr Val Val Leu Glu Gln Asn Tyr			
1090	1095	1100	
Thr Ile Arg Lys Asn Ile Asp Pro Asp Leu Lys Val Asp Lys Val Ile			
1105	1110	1115	1120
Asp Val Gly Ile Lys Arg Ile Leu Glu Ser Arg Leu Glu Glu Phe Gly			
1125	1130	1135	
Gly Asn Ala Lys Leu Ala Phe Ser Asn Leu Glu Glu Asn Pro Ile Trp			
1140	1145	1150	
Leu Asn Lys Glu Lys Gly Ile Ser Ile Lys Arg Val Lys Ile Ser Gly			
1155	1160	1165	
Val Ser Asn Val Glu Ser Leu His Val Lys Lys Asp His Phe Gly Glu			
1170	1175	1180	
Pro Ile Leu Asp Gln Glu Gly Asn Glu Ile Pro Val Asp Phe Val Ser			
1185	1190	1195	1200
Thr Gly Asn Asn His His Val Ala Ile Tyr Glu Asp Glu Asn Gly Asn			
1205	1210	1215	
Leu Gln Glu Glu Val Val Ser Phe Phe Glu Ala Val Val Arg Gln Asn			
1220	1225	1230	
Gln Gly Leu Pro Ile Ile Lys Lys Asn His Thr Leu Gly Trp Lys Phe			
1235	1240	1245	

Leu Phe Thr Leu Lys Gln Asn Glu Tyr Phe Val Phe Pro Ser Asp Asp

1250

1255

1260

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

1265

1270

1275

1280

Ile Ser Pro Asn Leu Phe Arg Val Gln Lys Ile Ala Arg Lys Asn Tyr

1285

1290

1295

Val Phe Asn Asn His Leu Glu Thr Lys Ala Val Asp Asn Asp Leu Leu

1300

1305

1310

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

1315

1320

1325

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

1330

1335

1340

Lys Ile Ile Gln Ile Gly Glu Tyr

1345

1350

<210> 26

<211> 1509

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from Psychroflexus torquisI

<400> 26

Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly Trp

1

5

10

15

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

20

25

30

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

35

40

45

Asp Ala Gly Gln Ser Ile Ser Gln Thr Ala Asp Arg Thr Lys Tyr Arg

50

55

60

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

65

70

75

80

His Arg Val Leu Lys Ile Leu Asp Phe Leu Pro Lys His Tyr Ser Glu
 85 90 95
 Ser Ile Asp Phe Gln Asp Lys Val Gly Gln Phe Lys Pro Lys Gln Glu
 100 105 110
 Val Lys Leu Asn Tyr Arg Lys Asn Glu Lys Asn Lys His Glu Phe Val
 115 120 125
 Phe Met Asn Ser Phe Ile Glu Met Val Ser Glu Phe Lys Asn Ala Gln
 130 135 140

 Pro Glu Leu Phe Tyr Asn Lys Gly Asn Gly Glu Glu Thr Lys Ile Pro
 145 150 155 160
 Tyr Asp Trp Thr Leu Tyr Tyr Leu Arg Lys Lys Ala Leu Thr Gln Gln
 165 170 175
 Ile Thr Lys Glu Glu Leu Ala Trp Leu Ile Leu Asn Phe Asn Gln Lys
 180 185 190
 Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Asp Ile Asp Glu Asp Lys Asn
 195 200 205

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

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

				325				330				335			
Lys	Leu	Leu	Ser	Lys	Gln	Ile	Glu	Leu	Gln	Pro	Glu	Leu	Phe	Asn	Glu
340				345				350							
Ser	Leu	Tyr	Lys	Ala	Cys	Ile	Lys	Glu	Leu	Tyr	Pro	Arg	Asn	Glu	Asn
355				360				365							
His	Gln	Ser	Asn	Asn	Lys	Lys	Gln	Gly	Phe	Glu	Tyr	Leu	Phe	Thr	Glu
370				375				380							
Asp	Ile	Ile	Phe	Tyr	Gln	Arg	Pro	Leu	Lys	Ser	Gln	Lys	Ser	Asn	Ile
385				390				395				400			
Ser	Gly	Cys	Gln	Phe	Glu	His	Lys	Ile	Tyr	Lys	Gln	Lys	Asn	Lys	Lys
405				410				415							
Thr	Gly	Lys	Leu	Glu	Leu	Ile	Lys	Glu	Pro	Ile	Lys	Thr	Ile	Ser	Arg
420				425				430							
Ser	His	Pro	Leu	Phe	Gln	Glu	Phe	Arg	Ile	Trp	Gln	Trp	Leu	Gln	Asn
435				440				445							
Leu	Lys	Ile	Tyr	Asn	Lys	Glu	Lys	Ile	Glu	Asn	Gly	Lys	Leu	Glu	Asp
450				455				460							
Val	Thr	Thr	Gln	Leu	Leu	Pro	Asn	Asn	Glu	Ala	Tyr	Val	Thr	Leu	Phe
465				470				475				480			
Asp	Phe	Leu	Asn	Thr	Lys	Lys	Glu	Leu	Glu	Gln	Lys	Gln	Phe	Ile	Glu
485				490				495							
Tyr	Phe	Val	Lys	Lys	Lys	Leu	Ile	Asp	Lys	Lys	Glu	Lys	Glu	His	Phe
500				505				510							
Arg	Trp	Asn	Phe	Val	Glu	Asp	Lys	Lys	Tyr	Pro	Phe	Ser	Glu	Thr	Arg
515				520				525							
Ala	Gln	Phe	Leu	Ser	Arg	Leu	Ala	Lys	Val	Lys	Gly	Ile	Lys	Asn	Thr
530				535				540							
Glu	Asp	Phe	Leu	Asn	Lys	Asn	Thr	Gln	Val	Gly	Ser	Lys	Glu	Asn	Ser
545				550				555				560			
Pro	Phe	Ile	Lys	Arg	Ile	Glu	Gln	Leu	Trp	His	Ile	Ile	Tyr	Ser	Val
565				570				575							

Ser Asp Leu Lys Glu Tyr Glu Lys Ala Leu Glu Lys Phe Ala Glu Lys
580 585 590

His Asn Leu Glu Lys Asp Ser Phe Leu Lys Asn Phe Lys Lys Phe Pro
595 600 605

Pro Phe Val Ser Asp Tyr Ala Ser Tyr Ser Lys Lys Ala Ile Ser Lys
610 615 620

Leu Leu Pro Ile Met Arg Met Gly Lys Tyr Trp Ser Glu Ser Ala Val
625 630 635 640

Pro Thr Gln Val Lys Glu Arg Ser Leu Ser Ile Met Glu Arg Val Lys
645 650 655

Val Leu Pro Leu Lys Glu Gly Tyr Ser Asp Lys Asp Leu Ala Asp Leu
660 665 670

Leu Ser Arg Val Ser Asp Asp Asp Ile Pro Lys Gln Leu Ile Lys Ser
675 680 685

Phe Ile Ser Phe Lys Asp Lys Asn Pro Leu Lys Gly Leu Asn Thr Tyr
690 695 700

Gln Ala Asn Tyr Leu Val Tyr Gly Arg His Ser Glu Thr Gly Asp Ile
705 710 715 720

Gln His Trp Lys Thr Pro Glu Asp Ile Asp Arg Tyr Leu Asn Asn Phe
725 730 735

Lys Gln His Ser Leu Arg Asn Pro Ile Val Glu Gln Val Val Met Glu
740 745 750

Thr Leu Arg Val Val Arg Asp Ile Trp Glu His Tyr Gly Asn Asn Glu
755 760 765

Lys Asp Phe Phe Lys Glu Ile His Val Glu Leu Gly Arg Glu Met Lys
770 775 780

Ser Pro Ala Gly Lys Arg Glu Lys Leu Ser Gln Arg Asn Thr Glu Asn
785 790 795 800

Glu Asn Thr Asn His Arg Ile Arg Glu Val Leu Lys Glu Leu Met Asn
805 810 815

Asp Ala Ser Val Glu Gly Gly Val Arg Asp Tyr Ser Pro Ser Gln Gln

820 825 830
 Glu Ile Leu Lys Leu Tyr Glu Glu Gly Ile Tyr Gln Asn Pro Asn Thr
 835 840 845

Asn Tyr Leu Lys Val Asp Glu Asp Glu Ile Leu Lys Ile Arg Lys Lys
 850 855 860
 Asn Asn Pro Thr Gln Lys Glu Ile Gln Arg Tyr Lys Leu Trp Leu Glu
 865 870 875 880
 Gln Gly Tyr Ile Ser Pro Tyr Thr Gly Lys Ile Ile Pro Leu Thr Lys
 885 890 895
 Leu Phe Thr His Glu Tyr Gln Ile Glu His Ile Ile Pro Gln Ser Arg
 900 905 910

Tyr Tyr Asp Asn Ser Leu Gly Asn Lys Ile Ile Cys Glu Ser Glu Val
 915 920 925
 Asn Glu Asp Lys Asp Asn Lys Thr Ala Tyr Glu Tyr Leu Lys Val Glu
 930 935 940
 Lys Gly Ser Ile Val Phe Gly His Lys Leu Leu Asn Leu Asp Glu Tyr
 945 950 955 960
 Glu Ala His Val Asn Lys Tyr Phe Lys Lys Asn Lys Thr Lys Leu Lys
 965 970 975

Asn Leu Leu Ser Glu Asp Ile Pro Glu Gly Phe Ile Asn Arg Gln Leu
 980 985 990
 Asn Asp Ser Arg Tyr Ile Ser Lys Leu Val Lys Gly Leu Leu Ser Asn
 995 1000 1005
 Ile Val Arg Glu Asn Gly Glu Gln Glu Ala Thr Ser Lys Asn Leu Ile
 1010 1015 1020
 Pro Val Thr Gly Val Val Thr Ser Lys Leu Lys Gln Asp Trp Gly Leu
 1025 1030 1035 1040

Asn Asp Lys Trp Asn Glu Ile Ile Ala Pro Arg Phe Lys Arg Leu Asn
 1045 1050 1055
 Lys Leu Thr Asn Ser Asn Asp Phe Gly Phe Trp Asp Asn Asp Ile Asn
 1060 1065 1070

Ala Phe Arg Ile Gln Val Pro Asp Ser Leu Ile Lys Gly Phe Ser Lys

1075 1080 1085

Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Val Ala

1090 1095 1100

Cys Thr Ser Arg Asn His Thr His Tyr Leu Ser Ala Leu Asn Ala Glu

1105 1110 1115 1120

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

1125 1130 1135

Gly Asp Tyr Thr Lys Thr Phe Gln Ile Pro Trp Gln Gly Phe Thr Ile

1140 1145 1150

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

1155 1160 1165

Leu Arg Val Ile Asn Lys Thr Asn Asn Lys Phe Trp Ser Tyr Lys Asp

1170 1175 1180

Glu Asn Gly Asn Leu Asn Leu Gly Lys Asp Gly Lys Pro Lys Lys Lys

1185 1190 1195 1200

Leu Arg Lys Gln Thr Lys Gly Tyr Asn Trp Ala Ile Arg Lys Pro Leu

1205 1210 1215

His Lys Glu Thr Val Ser Gly Ile Tyr Asn Ile Asn Ala Pro Lys Asn

1220 1225 1230

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

1235 1240 1245

Lys His Leu Ala Lys Ile Thr Asp Leu Arg Ile Arg Glu Thr Ile Leu

1250 1255 1260

Pro Asn His Leu Lys His Tyr Leu Asn Asn Lys Gly Glu Ala Asn Phe

1265 1270 1275 1280

Ser Glu Ala Phe Ser Gln Gly Gly Ile Glu Asp Leu Asn Lys Lys Ile

1285 1290 1295

Thr Thr Leu Asn Glu Gly Lys Lys His Gln Pro Ile Tyr Arg Val Lys

1300 1305 1310

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

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

 Ile Pro Leu Asn Glu Val Ile Thr His Gln Lys Gln Val Ala Gly Phe
 1365 1370 1375
 Pro Lys Ser Glu Arg Leu Ser Val Gln Pro Asp Ser Gln Lys Gly Thr
 1380 1385 1390
 Phe Leu Phe Thr Leu Ser Pro Asn Asp Leu Val Tyr Val Pro Asn Asn
 1395 1400 1405
 Glu Glu Leu Glu Asn Arg Asp Leu Phe Asn Leu Gly Asn Leu Asn Val
 1410 1415 1420

 Glu Gln Ile Ser Arg Ile Tyr Lys Phe Thr Asp Ser Ser Asp Lys Thr
 1425 1430 1435 1440
 Cys Asn Phe Ile Pro Phe Gln Val Ser Lys Leu Ile Phe Asn Leu Lys
 1445 1450 1455
 Lys Lys Glu Gln Lys Lys Leu Asp Val Asp Phe Ile Ile Gln Asn Glu
 1460 1465 1470
 Phe Gly Leu Gly Ser Pro Gln Ser Lys Asn Gln Lys Ser Ile Asp Asp
 1475 1480 1485

 Val Met Ile Lys Glu Lys Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly
 1490 1495 1500
 Asn Ile Ser Lys Ala
 1505
 <210> 27
 <211> 1388
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Cas9 from Streptococcus thermophilus
 <400> 27

Met Thr Lys Pro Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15
Gly Trp Ala Val Thr Thr Asp Asn Tyr Lys Val Pro Ser Lys Lys Met
20 25 30
Lys Val Leu Gly Asn Thr Ser Lys Lys Tyr Ile Lys Lys Asn Leu Leu
35 40 45
Gly Val Leu Leu Phe Asp Ser Gly Ile Thr Ala Glu Gly Arg Arg Leu
50 55 60
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Arg Asn Arg Ile Leu
65 70 75 80
Tyr Leu Gln Glu Ile Phe Ser Thr Glu Met Ala Thr Leu Asp Asp Ala
85 90 95
Phe Phe Gln Arg Leu Asp Asp Ser Phe Leu Val Pro Asp Asp Lys Arg
100 105 110
Asp Ser Lys Tyr Pro Ile Phe Gly Asn Leu Val Glu Glu Lys Ala Tyr
115 120 125
His Asp Glu Phe Pro Thr Ile Tyr His Leu Arg Lys Tyr Leu Ala Asp
130 135 140
Ser Thr Lys Lys Ala Asp Leu Arg Leu Val Tyr Leu Ala Leu Ala His
145 150 155 160
Met Ile Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Glu Phe Asn Ser
165 170 175
Lys Asn Asn Asp Ile Gln Lys Asn Phe Gln Asp Phe Leu Asp Thr Tyr
180 185 190
Asn Ala Ile Phe Glu Ser Asp Leu Ser Leu Glu Asn Ser Lys Gln Leu
195 200 205
Glu Glu Ile Val Lys Asp Lys Ile Ser Lys Leu Glu Lys Lys Asp Arg
210 215 220
Ile Leu Lys Leu Phe Pro Gly Glu Lys Asn Ser Gly Ile Phe Ser Glu
225 230 235 240
Phe Leu Lys Leu Ile Val Gly Asn Gln Ala Asp Phe Arg Lys Cys Phe

245 250 255
 Asn Leu Asp Glu Lys Ala Ser Leu His Phe Ser Lys Glu Ser Tyr Asp
 260 265 270
 Glu Asp Leu Glu Thr Leu Leu Gly Tyr Ile Gly Asp Asp Tyr Ser Asp
 275 280 285

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

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

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

 Val Ile Asp Lys Glu Ser Ser Ala Glu Ala Phe Ile Asn Arg Met Thr
 485 490 495

Ser Phe Asp Leu Tyr Leu Pro Glu Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Thr Phe Asn Val Tyr Asn Glu Leu Thr Lys Val Arg
515 520 525

Phe Ile Ala Glu Ser Met Arg Asp Tyr Gln Phe Leu Asp Ser Lys Gln
530 535 540

Lys Lys Asp Ile Val Arg Leu Tyr Phe Lys Asp Lys Arg Lys Val Thr
545 550 555 560

Asp Lys Asp Ile Ile Glu Tyr Leu His Ala Ile Tyr Gly Tyr Asp Gly
565 570 575

Ile Glu Leu Lys Gly Ile Glu Lys Gln Phe Asn Ser Ser Leu Ser Thr
580 585 590

Tyr His Asp Leu Leu Asn Ile Ile Asn Asp Lys Glu Phe Leu Asp Asp
595 600 605

Ser Ser Asn Glu Ala Ile Ile Glu Glu Ile Ile His Thr Leu Thr Ile
610 615 620

Phe Glu Asp Arg Glu Met Ile Lys Gln Arg Leu Ser Lys Phe Glu Asn
625 630 635 640

Ile Phe Asp Lys Ser Val Leu Lys Lys Leu Ser Arg Arg His Tyr Thr
645 650 655

Gly Trp Gly Lys Leu Ser Ala Lys Leu Ile Asn Gly Ile Arg Asp Glu
660 665 670

Lys Ser Gly Asn Thr Ile Leu Asp Tyr Leu Ile Asp Asp Gly Ile Ser
675 680 685

Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ala Leu Ser Phe Lys
690 695 700

Lys Lys Ile Gln Lys Ala Gln Ile Ile Gly Asp Glu Asp Lys Gly Asn
705 710 715 720

Ile Lys Glu Val Val Lys Ser Leu Pro Gly Ser Pro Ala Ile Lys Lys
725 730 735

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

740	745	750	
Gly Gly Arg Lys Pro Glu Ser Ile Val Val Glu Met Ala Arg Glu Asn			
755	760	765	
Gln Tyr Thr Asn Gln Gly Lys Ser Asn Ser Gln Gln Arg Leu Lys Arg			
770	775	780	
Leu Glu Lys Ser Leu Lys Glu Leu Gly Ser Lys Ile Leu Lys Glu Asn			
785	790	795	800
Ile Pro Ala Lys Leu Ser Lys Ile Asp Asn Asn Ala Leu Gln Asn Asp			
805	810	815	
Arg Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Lys Asp Met Tyr Thr Gly			
820	825	830	
Asp Asp Leu Asp Ile Asp Arg Leu Ser Asn Tyr Asp Ile Asp His Ile			
835	840	845	
Ile Pro Gln Ala Phe Leu Lys Asp Asn Ser Ile Asp Asn Lys Val Leu			
850	855	860	
Val Ser Ser Ala Ser Asn Arg Gly Lys Ser Asp Asp Val Pro Ser Leu			
865	870	875	880
Glu Val Val Lys Lys Arg Lys Thr Phe Trp Tyr Gln Leu Leu Lys Ser			
885	890	895	
Lys Leu Ile Ser Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg			
900	905	910	
Gly Gly Leu Ser Pro Glu Asp Lys Ala Gly Phe Ile Gln Arg Gln Leu			
915	920	925	
Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Arg Leu Leu Asp Glu			
930	935	940	
Lys Phe Asn Asn Lys Lys Asp Glu Asn Asn Arg Ala Val Arg Thr Val			
945	950	955	960
Lys Ile Ile Thr Leu Lys Ser Thr Leu Val Ser Gln Phe Arg Lys Asp			
965	970	975	
Phe Glu Leu Tyr Lys Val Arg Glu Ile Asn Asp Phe His His Ala His			
980	985	990	

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

995 1000 1005

Pro Lys Leu Glu Pro Glu Phe Val Tyr Gly Asp Tyr Pro Lys Tyr Asn

1010 1015 1020

Ser Phe Arg Glu Arg Lys Ser Ala Thr Glu Lys Val Tyr Phe Tyr Ser

1025 1030 1035 1040

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

1045 1050 1055

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

1060 1065 1070

Val Trp Asn Lys Glu Ser Asp Leu Ala Thr Val Arg Arg Val Leu Ser

1075 1080 1085

Tyr Pro Gln Val Asn Val Val Lys Lys Val Glu Glu Gln Asn His Gly

1090 1095 1100

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

1105 1110 1115 1120

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

1125 1130 1135

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

1140 1145 1150

Thr Val Leu Val Lys Gly Thr Ile Glu Lys Gly Ala Lys Lys Lys Ile

1155 1160 1165

Thr Asn Val Leu Glu Phe Gln Gly Ile Ser Ile Leu Asp Arg Ile Asn

1170 1175 1180

Tyr Arg Lys Asp Lys Leu Asn Phe Leu Leu Glu Lys Gly Tyr Lys Asp

1185 1190 1195 1200

Ile Glu Leu Ile Ile Glu Leu Pro Lys Tyr Ser Leu Phe Glu Leu Ser

1205 1210 1215

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

1220 1225 1230

Arg Gly Glu Ile His Lys Gly Asn Gln Ile Phe Leu Ser Gln Lys Phe

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

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

 Glu Thr Arg Ile Asp Leu Ala Lys Leu Gly Glu Gly
 1380 1385

 <210> 28
 <211> 1334
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Cas9 from *Listeria innocua*
 <400> 28

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

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

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

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

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

Gly Asp Glu Tyr Ala Glu Leu Phe Val Ala Ala Lys Asn Ala Tyr Ser

290

295

300

Ala Val Val Leu Ser Ser Ile Ile Thr Val Ala Glu Thr Glu Thr Asn

305

310

315

320

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

325

330

335

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

340

345

350

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

355

360

365

Ile Asp Gly Lys Thr Lys Gln Ala Asp Phe Tyr Lys Tyr Met Lys Met

370

375

380

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

385

390

395

400

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

405

410

415

Pro His Gln Leu His Leu Glu Glu Leu Glu Ala Ile Leu His Gln Gln

420

425

430

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

435

440

445

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

450

455

460

Gln Ser Glu Phe Ala Trp Leu Thr Arg Lys Ala Asp Gly Glu Ile Arg

465

470

475

480

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

485

490

495

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

500

505

510

Val Leu Pro Lys His Ser Leu Cys Tyr Gln Lys Tyr Leu Val Tyr Asn

515

520

525

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

530 535 540
 Phe Ser Gly Gln Glu Lys Glu Gln Ile Phe Asn Asp Leu Phe Lys Gln

 545 550 555 560
 Lys Arg Lys Val Lys Lys Lys Asp Leu Glu Leu Phe Leu Arg Asn Met

 565 570 575
 Ser His Val Glu Ser Pro Thr Ile Glu Gly Leu Glu Asp Ser Phe Asn

 580 585 590
 Ser Ser Tyr Ser Thr Tyr His Asp Leu Leu Lys Val Gly Ile Lys Gln

 595 600 605
 Glu Ile Leu Asp Asn Pro Val Asn Thr Glu Met Leu Glu Asn Ile Val

 610 615 620
 Lys Ile Leu Thr Val Phe Glu Asp Lys Arg Met Ile Lys Glu Gln Leu
 625 630 635 640
 Gln Gln Phe Ser Asp Val Leu Asp Gly Val Val Leu Lys Lys Leu Glu

 645 650 655
 Arg Arg His Tyr Thr Gly Trp Gly Arg Leu Ser Ala Lys Leu Leu Met

 660 665 670
 Gly Ile Arg Asp Lys Gln Ser His Leu Thr Ile Leu Asp Tyr Leu Met

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

 725 730 735
 Ile Lys Lys Gly Ile Leu Gln Ser Leu Lys Ile Val Asp Glu Leu Val

 740 745 750
 Ser Val Met Gly Tyr Pro Pro Gln Thr Ile Val Val Glu Met Ala Arg

 755 760 765
 Glu Asn Gln Thr Thr Gly Lys Gly Lys Asn Asn Ser Arg Pro Arg Tyr

 770 775 780

Lys Ser Leu Glu Lys Ala Ile Lys Glu Phe Gly Ser Gln Ile Leu Lys
 785 790 795 800
 Glu His Pro Thr Asp Asn Gln Glu Leu Arg Asn Asn Arg Leu Tyr Leu

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

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

 930 935 940
 Glu Lys Asp Asp His Gly Asn Thr Met Lys Gln Val Arg Ile Val Thr
 945 950 955 960
 Leu Lys Ser Ala Leu Val Ser Gln Phe Arg Lys Gln Phe Gln Leu Tyr
 965 970 975
 Lys Val Arg Asp Val Asn Asp Tyr His His Ala His Asp Ala Tyr Leu
 980 985 990
 Asn Gly Val Val Ala Asn Thr Leu Leu Lys Val Tyr Pro Gln Leu Glu

 995 1000 1005
 Pro Glu Phe Val Tyr Gly Asp Tyr His Gln Phe Asp Trp Phe Lys Ala
 1010 1015 1020
 Asn Lys Ala Thr Ala Lys Lys Gln Phe Tyr Thr Asn Ile Met Leu Phe

1025 1030 1035 1040
 Phe Ala Gln Lys Asp Arg Ile Ile Asp Glu Asn Gly Glu Ile Leu Trp
 1045 1050 1055
 Asp Lys Lys Tyr Leu Asp Thr Val Lys Lys Val Met Ser Tyr Arg Gln

 1060 1065 1070
 Met Asn Ile Val Lys Lys Thr Glu Ile Gln Lys Gly Glu Phe Ser Lys
 1075 1080 1085
 Ala Thr Ile Lys Pro Lys Gly Asn Ser Ser Lys Leu Ile Pro Arg Lys
 1090 1095 1100
 Thr Asn Trp Asp Pro Met Lys Tyr Gly Gly Leu Asp Ser Pro Asn Met
 1105 1110 1115 1120
 Ala Tyr Ala Val Val Ile Glu Tyr Ala Lys Gly Lys Asn Lys Leu Val

 1125 1130 1135
 Phe Glu Lys Lys Ile Ile Arg Val Thr Ile Met Glu Arg Lys Ala Phe
 1140 1145 1150
 Glu Lys Asp Glu Lys Ala Phe Leu Glu Glu Gln Gly Tyr Arg Gln Pro
 1155 1160 1165
 Lys Val Leu Ala Lys Leu Pro Lys Tyr Thr Leu Tyr Glu Cys Glu Glu
 1170 1175 1180
 Gly Arg Arg Arg Met Leu Ala Ser Ala Asn Glu Ala Gln Lys Gly Asn

 1185 1190 1195 1200
 Gln Gln Val Leu Pro Asn His Leu Val Thr Leu Leu His His Ala Ala
 1205 1210 1215
 Asn Cys Glu Val Ser Asp Gly Lys Ser Leu Asp Tyr Ile Glu Ser Asn
 1220 1225 1230
 Arg Glu Met Phe Ala Glu Leu Leu Ala His Val Ser Glu Phe Ala Lys
 1235 1240 1245
 Arg Tyr Thr Leu Ala Glu Ala Asn Leu Asn Lys Ile Asn Gln Leu Phe

 1250 1255 1260
 Glu Gln Asn Lys Glu Gly Asp Ile Lys Ala Ile Ala Gln Ser Phe Val
 1265 1270 1275 1280

Asp Leu Met Ala Phe Asn Ala Met Gly Ala Pro Ala Ser Phe Lys Phe
1285 1290 1295
Phe Glu Thr Thr Ile Glu Arg Lys Arg Tyr Asn Asn Leu Lys Glu Leu
1300 1305 1310
Leu Asn Ser Thr Ile Ile Tyr Gln Ser Ile Thr Gly Leu Tyr Glu Ser
1315 1320 1325
Arg Lys Arg Leu Asp Asp
1330
<210> 29
<211> 984
<212> PRT
<213> Artificial Sequence
<220>
<223> Cas9 from *Campylobacter jejuni*
<400> 29
Met Ala Arg Ile Leu Ala Phe Asp Ile Gly Ile Ser Ser Ile Gly Trp
1 5 10 15
Ala Phe Ser Glu Asn Asp Glu Leu Lys Asp Cys Gly Val Arg Ile Phe
20 25 30
Thr Lys Val Glu Asn Pro Lys Thr Gly Glu Ser Leu Ala Leu Pro Arg
35 40 45
Arg Leu Ala Arg Ser Ala Arg Lys Arg Leu Ala Arg Arg Lys Ala Arg
50 55 60
Leu Asn His Leu Lys His Leu Ile Ala Asn Glu Phe Lys Leu Asn Tyr
65 70 75 80
Glu Asp Tyr Gln Ser Phe Asp Glu Ser Leu Ala Lys Ala Tyr Lys Gly
85 90 95
Ser Leu Ile Ser Pro Tyr Glu Leu Arg Phe Arg Ala Leu Asn Glu Leu
100 105 110
Leu Ser Lys Gln Asp Phe Ala Arg Val Ile Leu His Ile Ala Lys Arg
115 120 125
Arg Gly Tyr Asp Asp Ile Lys Asn Ser Asp Asp Lys Glu Lys Gly Ala

130 135 140
 Ile Leu Lys Ala Ile Lys Gln Asn Glu Glu Lys Leu Ala Asn Tyr Gln
 145 150 155 160
 Ser Val Gly Glu Tyr Leu Tyr Lys Glu Tyr Phe Gln Lys Phe Lys Glu

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

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

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

 355 360 365
 Lys Lys Ala Leu Ala Lys Tyr Asp Leu Asn Gln Asn Gln Ile Asp Ser
 370 375 380

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

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

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

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

 805 810 815
 Ile Arg Lys Val Asn Gly Lys Ile Val Lys Asn Gly Asp Met Phe Arg
 820 825 830
 Val Asp Ile Phe Lys His Lys Lys Thr Asn Lys Phe Tyr Ala Val Pro
 835 840 845
 Ile Tyr Thr Met Asp Phe Ala Leu Lys Val Leu Pro Asn Lys Ala Val
 850 855 860
 Ala Arg Ser Lys Lys Gly Glu Ile Lys Asp Trp Ile Leu Met Asp Glu

 865 870 875 880

Asn Tyr Glu Phe Cys Phe Ser Leu Tyr Lys Asp Ser Leu Ile Leu Ile
885 890 895
Gln Thr Lys Asp Met Gln Glu Pro Glu Phe Val Tyr Tyr Asn Ala Phe
900 905 910
Thr Ser Ser Thr Val Ser Leu Ile Val Ser Lys His Asp Asn Lys Phe
915 920 925
Glu Thr Leu Ser Lys Asn Gln Lys Ile Leu Phe Lys Asn Ala Asn Glu
930 935 940
Lys Glu Val Ile Ala Lys Ser Ile Gly Ile Gln Asn Leu Lys Val Phe
945 950 955 960
Glu Lys Tyr Ile Val Ser Ala Leu Gly Glu Val Thr Lys Ala Glu Phe
965 970 975
Arg Gln Arg Glu Asp Phe Lys Lys
980

<210> 30

<211> 1082

<212> PRT

<213> Artificial Sequence

<220>

<223> Cas9 from Neisseria. meningitidis

<400> 30

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

Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu
65 70 75 80
Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp

85	90	95	
Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln			
100	105	110	
Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser			
115	120	125	
Ala Val Leu Leu His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg			
130	135	140	
Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys			
145	150	155	160
Gly Val Ala Asp Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr			
165	170	175	
Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile			
180	185	190	
Arg Asn Gln Arg Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu			
195	200	205	
Gln Ala Glu Leu Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn			
210	215	220	
Pro His Val Ser Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met			
225	230	235	240
Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly			
245	250	255	
His Cys Thr Phe Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr			
260	265	270	
Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile			
275	280	285	
Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr			
290	295	300	
Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala			
305	310	315	320
Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg			
325	330	335	

Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala
 340 345 350
 Tyr His Ala Ile Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys
 355 360 365
 Lys Ser Pro Leu Asn Leu Ser Pro Glu Leu Gln Asp Glu Ile Gly Thr
 370 375 380

 Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys
 385 390 395 400
 Asp Arg Ile Gln Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser
 405 410 415
 Phe Asp Lys Phe Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val
 420 425 430
 Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile
 435 440 445

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

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

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

580 585 590
 Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly
 595 600 605
 Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn
 610 615 620
 Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu
 625 630 635 640

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

 Leu Leu Arg Gly Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp
 705 710 715 720
 Arg His His Ala Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala
 725 730 735
 Met Gln Gln Lys Ile Thr Arg Phe Val Arg Tyr Lys Glu Met Asn Ala
 740 745 750
 Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr Gly Glu Val Leu His Gln
 755 760 765

 Lys Thr His Phe Pro Gln Pro Trp Glu Phe Phe Ala Gln Glu Val Met
 770 775 780
 Ile Arg Val Phe Gly Lys Pro Asp Gly Lys Pro Glu Phe Glu Glu Ala
 785 790 795 800
 Asp Thr Pro Glu Lys Leu Arg Thr Leu Leu Ala Glu Lys Leu Ser Ser
 805 810 815
 Arg Pro Glu Ala Val His Glu Tyr Val Thr Pro Leu Phe Val Ser Arg
 820 825 830

Ala Pro Asn Arg Lys Met Ser Gly Gln Gly His Met Glu Thr Val Lys
835 840 845

Ser Ala Lys Arg Leu Asp Glu Gly Val Ser Val Leu Arg Val Pro Leu
850 855 860

Thr Gln Leu Lys Leu Lys Asp Leu Glu Lys Met Val Asn Arg Glu Arg
865 870 875 880

Glu Pro Lys Leu Tyr Glu Ala Leu Lys Ala Arg Leu Glu Ala His Lys
885 890 895

Asp Asp Pro Ala Lys Ala Phe Ala Glu Pro Phe Tyr Lys Tyr Asp Lys
900 905 910

Ala Gly Asn Arg Thr Gln Gln Val Lys Ala Val Arg Val Glu Gln Val
915 920 925

Gln Lys Thr Gly Val Trp Val Arg Asn His Asn Gly Ile Ala Asp Asn
930 935 940

Ala Thr Met Val Arg Val Asp Val Phe Glu Lys Gly Asp Lys Tyr Tyr
945 950 955 960

Leu Val Pro Ile Tyr Ser Trp Gln Val Ala Lys Gly Ile Leu Pro Asp
965 970 975

Arg Ala Val Val Gln Gly Lys Asp Glu Glu Asp Trp Gln Leu Ile Asp
980 985 990

Asp Ser Phe Asn Phe Lys Phe Ser Leu His Pro Asn Asp Leu Val Glu
995 1000 1005

Val Ile Thr Lys Lys Ala Arg Met Phe Gly Tyr Phe Ala Ser Cys His
1010 1015 1020

Arg Gly Thr Gly Asn Ile Asn Ile Arg Ile His Asp Leu Asp His Lys
1025 1030 1035 1040

Ile Gly Lys Asn Gly Ile Leu Glu Gly Ile Gly Val Lys Thr Ala Leu
1045 1050 1055

Ser Phe Gln Lys Tyr Gln Ile Asp Glu Leu Gly Lys Glu Ile Arg Pro
1060 1065 1070

Cys Arg Leu Lys Lys Arg Pro Pro Val Arg
1075 1080

<210> 31

<211> 1367

<212>

PRT

<213> Artificial Sequence

<220>

<223> Cas9 from Streptococcus pyogenes

<400> 31

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

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

100 105 110

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

115 120 125

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

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175

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

180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Arg Val Asp Ala
 195 200 205
 Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220
 Leu Ile Ala Gln Leu Pro Gly Glu Lys Arg Asn Gly Leu Phe Gly Asn
 225 230 235 240
 Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255
 Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270
 Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300
 Ile Leu Arg Val Asn Ser Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile

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

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

 565 570 575
 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590
 Ala Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605
 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Arg Gly Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala

 625 630 635 640
 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655
 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685

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

690

695

700

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

705

710

715

720

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

725

730

735

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

740

745

750

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

755

760

765

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

770

775

780

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

785

790

795

800

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

805

810

815

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

820

825

830

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

835

840

845

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

850

855

860

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

865

870

875

880

Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe

885

890

895

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

900

905

910

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

915

920

925

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

930 935 940
 Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys

 945 950 955 960
 Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu

 965 970 975
 Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val

 980 985 990
 Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val

 995 1000 1005
 Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser

 1010 1015 1020
 Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn
 1025 1030 1035 1040
 Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile

 1045 1050 1055
 Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val

 1060 1065 1070
 Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met

 1075 1080 1085
 Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe
 1090 1095 1100
 Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala
 1105 1110 1115 1120
 Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro

 1125 1130 1135
 Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys

 1140 1145 1150
 Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met

 1155 1160 1165
 Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys

 1170 1175 1180

Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr
1185 1190 1195 1200
Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala
1205 1210 1215
Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
1220 1225 1230
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro
1235 1240 1245
Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr
1250 1255 1260
Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile
1265 1270 1275 1280
Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His
1285 1290 1295
Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe
1300 1305 1310
Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr
1315 1320 1325
Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala
1330 1335 1340
Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp
1345 1350 1355 1360
Leu Ser Gln Leu Gly Gly Asp
1365
<210> 32
<211> 528
<212> DNA
<213> Artificial Sequence
<220>
<223> zfp coding sequence
<400> 32
atggcccagg ctgctcttga gcccggagag aaacctaca agtggccgga gtgcgaaag 60

tccttctctg agcggagtca cctccgagag caccagcgga ctcatacggg cgaaaaacca 120

tacaagtgcc cagaatgtgg taaatctttt tctcgggctg acaacctgac tgaacatcag 180

cgcacgcaca ccggtgaaaa accttacaag tgtccagagt gtggcaagag cttttctagt 240

agaaggacct gtcgagcgca tcagcggact cacaccggcg aaaaacccta taagtgtccg 300

gaatgtggaa agagctttag cgcacacgac acccttactg aacaccagcg aacacacacg 360

ggagaaaaac catataaatg tccggaatgt ggcaaaagtt ttagtcggag tgataaactt 420

acggagcacc aacggacaca caccggagag aagccatata agtgcctga atgtggaaag 480

tccttctcac agcttgctca tctgcgagca catcagcgca cacacacc 528

<210> 33

<211> 176

<212> PRT

<213> Artificial Sequence

<220>

<223> ZFP amino acid sequence

<400> 33

Met Ala Gln Ala Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro

1 5 10 15

Glu Cys Gly Lys Ser Phe Ser Glu Arg Ser His Leu Arg Glu His Gln

20 25 30

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

35 40 45

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

50 55 60

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser

65 70 75 80

Arg Arg Thr Cys Arg Ala His Gln Arg Thr His Thr Gly Glu Lys Pro

85 90 95

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn Asp Thr Leu

100 105 110

Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

115 120 125

Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Lys Leu Thr Glu His Gln

130

135

140

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

145

150

155

160

Ser Phe Ser Gln Leu Ala His Leu Arg Ala His Gln Arg Thr His Thr

165

170

175

<210> 34

<211> 555

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNF-E2C Nucleotide

<400> 34

atggcgagg cggcgctgga accgggagaa aaaccgtata aatgcccgga atgcggcaaa	60
agcttttagcc gcaaagatag cctggtgagc catcagcgca ccataaccgg cgaaaaaccg	120
tataaatgcc cggaatgcgg caaaagcttt agccagagcg gcgatctgag ccgccatcag	180
cgcaccata ccggcgaaaa accgtataaa tgcccggaat gcggcaaaag ctttagcgat	240
tgccgcgagc tggcgcgcca tcagcgagc cataccggcg aaaaaccgta taaatgcccg	300
gaatgcggca aaagcttttag ccagagcagc catctggtgc gccatcagcg caccataacc	360

ggcgaaaaac cgtataaatg cccggaatgc ggcaaaagct ttagcgattg ccgcgatctg	420
gcgcgccatc agcgacacca taccggcgaa aaaccgtata aatgcccgga atgcggcaaa	480
agcttttagcc gcagcgataa actggtgagc catcagcgca ccataaccgg caaaaaaacc	540
agcgggcagg cgggc	555

<210> 35

<211> 185

<212> PRT

<213> Artificial Sequence

<220>

<223> ZNF-E2C Aminoacid

<400> 35

Met Ala Gln Ala Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro

1

5

10

15

Glu Cys Gly Lys Ser Phe Ser Arg Lys Asp Ser Leu Val Arg His Gln
 20 25 30
 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 35 40 45
 Ser Phe Ser Gln Ser Gly Asp Leu Arg Arg His Gln Arg Thr His Thr
 50 55 60
 Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp
 65 70 75 80

Cys Arg Asp Leu Ala Arg His Gln Arg Thr His Thr Gly Glu Lys Pro
 85 90 95
 Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser His Leu
 100 105 110
 Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
 115 120 125
 Glu Cys Gly Lys Ser Phe Ser Asp Cys Arg Asp Leu Ala Arg His Gln
 130 135 140

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 145 150 155 160
 Ser Phe Ser Arg Ser Asp Lys Leu Val Arg His Gln Arg Thr His Thr
 165 170 175
 Gly Lys Lys Thr Ser Gly Gln Ala Gly
 180 185

<210> 36

<211> 555

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNF-E3 Nucleotide

<400> 36

atggcgcagg cggcgcctgga accgggcgaa aaaccgtata aatgcccgga atgcggcaaa 60

agcttttagcg atccgggcgc gctggtgcmc catcagcgca cccataccgg cgaaaaaccg 120

tataaatgcc cggaatgcgg caaaagcttt agccagagca gccatctggt gcgccatcag 180

cgcacccata ccggcgaaaa accgtataaa tgcccgggaat gcggcaaaag ctttagcgat 240
 tgccgcgatac tggcgcgccca tcagcgcacc cataccggcg aaaaaccgta taaatgcccg 300
 gaatgcggca aaagcttttag ccagagcagc catctggtgc gccatcagcg cacccatacc 360
 ggcgaaaaac cgtataaatg cccggaatgc ggcaaaagct ttagcgattg ccgcgatctg 420
 gcgcgccatc agcgcaccca taccggcgaa aaaccgtata aatgcccgga atgcggcaaa 480

agcttttagcc agagcagcca tctggtgcgc catcagcgca cccataccgg caaaaaaacc 540
 agcggccagg cgggc 555

<210> 37

<211> 185

<212> PRT

<213> Artificial Sequence

<220>

<223> ZNF-E3 Amino acid

<400> 37

Met Ala Gln Ala Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro

1 5 10 15

Glu Cys Gly Lys Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln

20 25 30

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

35 40 45

Ser Phe Ser Gln Ser Ser His Leu Val Arg His Gln Arg Thr His Thr

50 55 60

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp

65 70 75 80

Cys Arg Asp Leu Ala Arg His Gln Arg Thr His Thr Gly Glu Lys Pro

85 90 95

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser His Leu

100 105 110

Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

115 120 125

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

130 135 140

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
145 150 155 160

Ser Phe Ser Gln Ser Ser His Leu Val Arg His Gln Arg Thr His Thr
165 170 175

Gly Lys Lys Thr Ser Gly Gln Ala Gly
180 185

<210> 38

<211> 528

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNF-TRCa Nucleotide

<400> 38

atggcgagg cggtcttga acccgaggag aaacctata aatgcctga gtgtggcaag 60
agtttttcaa ccacaggaaa ctgacagtc caccaacgga cccacaccgg cgagaaacca 120
tacaagtgtc cggagtgtgg taagtcttc tcaagtctg cgcaccttac cagacatcaa 180

cgcacacata caggtgaaaa acctacaag tgcccagagt gcggaaaaag tttttcacia 240
tctggcgacc tccgaggca ccagcgact cacaccgtg aaaaaccata caagtgtcct 300
gagtgcggga agagttttag tcaacgagct catctggagc gacaccaaag gactcatact 360
ggggagaaac cgtacaaatg tcccgaatgt gggaagagct tctctacaa gaattccctt 420
acagagcacc agcgacgca tacgggagag aagccgtata agtgtccgga atgtggcaag 480
agcttttcca gaagtgacca cttacaacc caccagagga cgcacacc 528

<210> 39

<211> 176

<212> PRT

<213> Artificial Sequence

<220>

<223> ZNF-TRCa Aminoacid

<400> 39

Met Ala Gln Ala Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro
1 5 10 15
Glu Cys Gly Lys Ser Phe Ser Thr Thr Gly Asn Leu Thr Val His Gln

20 25 30
 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 35 40 45
 Ser Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr
 50 55 60

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

Glu Cys Gly Lys Ser Phe Ser Thr Lys Asn Ser Leu Thr Glu His Gln
 130 135 140
 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 145 150 155 160
 Ser Phe Ser Arg Ser Asp His Leu Thr Thr His Gln Arg Thr His Thr
 165 170 175

<210> 40

<211> 18

<212> DNA

<213> Artificial sequence

<220>

<223> AAVS1 site

<400> 40

agacggccgc gtcagagc

18

<210> 41

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Zinc Finger 1

<400> 41

Glu Arg Ser His Leu Arg Glu

1 5

<210> 42

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Zinc Finger 2

<400> 42

Arg Ala Asp Asn Leu Thr Glu

1 5

<210> 43

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Zinc Finger 3

<400> 43

Ser Arg Arg Thr Cys Arg Ala

1 5

<210> 44

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Zinc Finger 4

<400> 44

Arg Asn Asp Thr Leu Thr Glu

1 5

<210> 45

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Zinc Finger 5

<400> 45

Arg Ser Asp Lys Leu Thr Glu

1 5

<210> 46

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Zinc Finger 6

<400> 46

Gln Leu Ala His Leu Arg Ala

1 5

<210> 47

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Nuclear Localization Signal

<400> 47

atggctccaa agaaaaagag gaaagtggga atccacggag tccccgccgc t 51

<210> 48

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS linker Nucleic Acid

<400> 48

ggtggatctg gcggtggatc tggcggcggt 30

<210> 49

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS linker Amino Acid

<400> 49

Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly

1 5 10

<210> 50

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS4x Linker Nucleic Acid

<400> 50

ggagggagtgtgtgggtccgg tggtagtggc ggatcc 36

<210> 51

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS4x Linker Amino Acid

<400> 51

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

1 5 10

<210> 52

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS5x Linker Nucleic Acid

<400>

> 52

ggaggtctcgtgtgggtctgg tgggagcgggt gtagtggcg gatcc 45

<210> 53

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS5x Linker Amino Acid

<400> 53

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

1 5 10 15

<210> 54

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS6x Linker Nucleic Acid

<400> 54

ggaggcagtg gtgggagcgg tgggtccggg ggtagtggtg gttccggggg atcc 54

<

210> 55

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS6x Linker Amino Acid

<400> 55

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

1 5 10 15

Gly Ser

<210> 56

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS7x Linker Nucleic Acid

<400> 56

ggaggttctg gaggtccgg tgggtccggg ggaagtgggg ggtcaggcgg atcaggagga 60

tcc

63

<210> 57

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS7x Linker Amino Acid

<400> 57

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

1 5 10 15

Gly Ser Gly Gly Ser

20

<210> 58

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS8x Linker Nucleic Acid

<400> 58

ggaggttagcg gaggttccgg agggagcggc gggagtgggg gaagcggggg aagtggagga 60

tccgggggag gatcc 75

<210> 59

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS8x Linker Amino Acid

<400> 59

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

1 5 10 15

Gly Ser Gly Gly Ser

20

<210> 60

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker XTEN Nucleic Acid

<400> 60

tccggtagcg aaacaccggg gacttcagaa tcggccaccc cggagtct 48

<210> 61

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Linker XTEN Amino Acid

<400> 61

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser

1 5 10 15

<210> 62

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker B Nucleic Acid

<400> 62

ggaagcgccg gtagtgccgc tgggtctggc gagttc 36

<210> 63

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Linker B Amino Acid

<400> 63

Gly Ser Ala Gly Ser Ala Ala Gly Ser Gly Glu Phe

1 5 10

<210> 64

<211> 4104

<212> DNA

<213> Artificial Sequence

<220>

<223> human Cas9 (hCas9) nucleic acid

<400> 64

atggacaaga agtactccat tgggctcgat atcggcacaa acagcgctcg ctgggccgtc	60
attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc	120
cacagcataa agaagaacct cattggcgcc ctctgttcg actccgggga gacggccgaa	180
gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc	240
tacctgcagg agatctttag taatgagatg gctaagggtg atgactcttt cttccatagg	300
ctggaggagt ctttttttgt ggaggaggat aaaaagcacg agcgccaccc aatctttggc	360
aatatcgtgg acgaggtggc gtaccatgaa aagtacccaa ccatatatca tctgaggaag	420
aagctttag acagtactga taaggctgac ttgcggttga tctatctcgc gctggcgcat	480
atgatcaaat ttccggggaca cttctcatc gagggggacc tgaaccaga caacagcgat	540
gtcgacaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg	600
atcaacgcat ccgaggttga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
cggtcga aa accatcatgc acagctccct ggggagaaga agaacggcct gtttggtaat	720
cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa	780
gatgccaaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc	840
cagatcggcg accagtacgc agaccttttt ttggcggaaga agaacctgtc agacgccatt	900
ctgctgagt atattctcgc agtgaacacg gagatcacca aagctccgct gagcgctagt	960
atgatcaagc gctatgatga gcaccacaa gacttgactt tgctgaaggc cttgtcaga	1020
cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggctacgcc	1080
ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gcccatcttg	1140
gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttgccg	1200
aaacagcgca ctttcgacaa tggaagcatc cccaccaga ttacactggg cgaactgcac	1260
gctatcctca ggcggcaaga ggatttctac ccctttttga aagataacag ggaaaagatt	1320
gagaaaaatc tcacatttcg gataccctac tatgtaggcc cctcgcgccg gggaaattcc	1380
agattcgcgt ggatgactcg caaatcagaa gagacatca ctccctggaa cttcgaggaa	1440
gtcgtggata agggggcctc tgcccagtc ttcacgaaa ggatgactaa ctttgataaa	1500

aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacgagta cttcacagtt	1560
tataacgagc tcaccaaggt caaatacgtc acagaaggga tgagaaagcc agcattcctg	1620
tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaacg gaaagttacc	1680
gtgaaacagc tcaaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc	1740
agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc	1800
attaaagaca aggacttcct ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
ctcaccctta cgttgtttga agataggag atgattgaag aacgcttgaa aacttacgct	1920
catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg	1980
cggtgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg	2040
gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac	2100
tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt	2160
cacgagcaca tcgctaactt tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220
gttaaggctg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt	2280
atcgagatgg cccgagagaa ccaactacc cagaaggagc agaagaacag tagggaaagg	2340
atgaagagga ttgaagaggg tataaaagaa ctgggggtccc aaatccttaa ggaacaccca	2400
gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaaaata gagggaaagag tgataacgtc cctcagaag aagtgtgtcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggttcat caaaaggcag	2760
cttgttgaga cacgccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac	2820
accaagtacg atgaaaatga caaactgatt cgagaggatga aagtatttac tctgaagtct	2880
aagctggtct cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat	2940
taccaccatg cgcatgatgc ctacctgaat gcagtggtag gcactgcact tatcaaaaaa	3000
tatccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa	3060
atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc	3120
aatattatga attttttcaa gaccgagatt aactggcca atggagagat tcggaagcga	3180
ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240

gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta 3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgac 3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgt 3420
tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc 3480
aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac 3540
tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcatcattaa gcttcccaag 3600
tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg 3660

cagaaaggta acgagctggc actgccctct aaatacgtta atttcttgta tctggccagc 3720
cactatgaaa agctcaaagg gtctcccgaa gataatgagc agaagcagct gttcgtggaa 3780
caacacaaac actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagtg 3840
atcctcgccg acgctaacct cgataagggt ctttctgctt acaataagca cagggataag 3900
cccatcaggg agcaggcaga aaacattatc cacttggtta ctctgaccaa ctgggcgcg 3960
cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtagac ctctacaaag 4020
gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc 4080

gacctctctc agctcgggtgg agac 4104

<210> 65

<211> 4104

<212> DNA

<213> Artificial Sequence

<220>

<223> nickase Cas9 (nCas9) nucleic acid

<400> 65

atggacaaga agtactccat tgggctcgct atcggcacaa acagcgtcgg ctgggcccgtc 60
attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc 120
cacagcataa agaagaacct cattggcgcc ctctgttctg actccgggga gacggccgaa 180
gccacgcggc taaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc 240
tacctgcagg agatctttag taatgagatg gctaagggtg atgactcttt cttccatagg 300

ctggaggagt cctttttggt ggaggaggat aaaaagcacg agcgccaccc aatctttggc 360
aatatcgtgg acgaggtggc gtacatgaa aagtaccaa ccatatatca tctgaggaag 420
aagctttagt acagtactga taagctgac ttgcggttga tctatctcgc gctggcgcat 480
atgatcaaat ttcggggaca cttctcatc gagggggacc tgaaccaga caacagcgat 540
gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaaccg 600

atcaacgcat ccggagtga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat	720
cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa	780
gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc	840
cagatcggcg accagtacgc agaccttttt ttggcggcaa agaacctgtc agacgccatt	900
ctgctgagtg atattctcgc agtgaacacg gagatcacca aagctccgct gagcgctagt	960
atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc cttgttcaga	1020
cagcaactgc ctgagaagta caaggaaatt ttcttcgacg agtctaaaaa tggttacgcc	1080
ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gcccattctg	1140
gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttgccg	1200
aaacacgcga ctttcgacaa tggaagcatc ccccaccaga ttacactggg cgaactgcac	1260
gctatcctca ggcggaaga ggatttctac ccctttttga aagataacag ggaaaagatt	1320
gagaaaatcc tcacatttcg gataccctac tatgtaggcc ccctcgcccg gggaaattcc	1380
agattcgctg ggatgactcg caaatcagaa gagacatca ctccctggaa cttcgaggaa	1440
gtcgtggata agggggcctc tgcccagtc ttcatcgaaa ggatgactaa ctttgataaa	1500
aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacgagta cttcacagtt	1560
tataacgagc tcaccaaggt caaatcgtc acagaaggga tgagaaagcc agcattcctg	1620
tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagttacc	1680
gtgaaacagc tcaaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc	1740
agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc	1800
attaaagaca aggacttctt ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
ctcacctta cgttgtttga agataggag atgattgaag aacgcttgaa aacttacgct	1920
catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg	1980
cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg	2040
gattttctta agtccgatgg atttccaac aggaacttca tgcagttgat ccatgatgac	2100
tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt	2160
cacgagcaca tcgctaact tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220
gttaaggctg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt	2280
atcgagatgg cccgagagaa ccaaactacc cagaaggagc agaagaacag tagggaaagg	2340

atgaagagga ttgaagaggg tataaaagaa ctgggtccc aaatccttaa ggaacaccca	2400
gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaaaata gagggaaagag tgataacgtc cctcagaag aagtgtgcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggcttcat caaaaggcag	2760
cttgttgaga cacgccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac	2820
accaagtacg atgaaaatga caaactgatt cgagaggtga aagtattac tctgaagtct	2880
aagctggtct cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat	2940
taccaccatg cgcatgatgc ctacctgaat gcagtggtag gcactgcact tatcaaaaaa	3000
tatccaagc ttgaatctga atttgtttac ggagactata aagtgtaga tgtaggaaa	3060
atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc	3120
aatattatga atttttcaa gaccgagatt aactggcca atggagagat tcggaagcga	3180
ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgatc	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct	3420
tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc	3480
aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac	3540
tttctcgagg cgaagagata taaagaggtc aaaaaagacc tcatcattaa gtttccaag	3600
tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg	3660
cagaaaggta acgagctggc actgcctctt aaatacgtta atttcttgta tctggccagc	3720
cactatgaaa agctcaaagg gtctcccgaa gataatgagc agaagcagct gttcgtggaa	3780
caacacaaac actaccttga tgagatcatc gagcaataa gcgaattctc caaaagagtg	3840
atcctcgccg acgctaacct cgataagggt ctttctgctt acaataagca cagggataag	3900
cccatcaggg agcaggcaga aaacattatc cacttgttta ctctgaccaa ctggggcgcg	3960
cctgcagcct tcaagtactt cgacaccacc atagacagaa agcgggtacac ctctacaaag	4020
gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc	4080

gacctctctc agctcggtgg agac 4104

<210> 66

<211> 4104

<212> DNA

<213> Artificial Sequence

<220>

<223> dead Cas9 (dCas9) nucleic acid

<400> 66

atggacaaga agtactccat tgggctcgt atcggcacaa acagcgtcgg ctgggccgtc 60

attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc 120

cacagcataa agaagaacct cattggcgcc ctctgttcg actccgggga gacggccgaa 180

gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc 240

tacctgcagg agatctttag taatgagatg gctaagggtg atgactcttt cttccatagg 300

ctggaggagt ctttttttgt ggaggaggat aaaaagcacg agcgccaccc aatctttggc 360

aatatcgtgg acgaggtggc gtacatgaa aagtaccaa ccatatatca tctgaggaag 420

aagctttag acagtactga taagctgac ttgcggtga tctatctgc gctggcgcat 480

atgatcaaat ttcggggaca ctctctcgc gggggggacc tgaaccaga caacagcgat 540

gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg 600

atcaacgcat ccggagtga cgcaaagca atcctgagcg ctaggctgtc caaatcccgg 660

cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat 720

cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttga cctggccgaa 780

gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc 840

cagatcggcg accagtacgc agaccttttt ttggcggaag agaacctgtc agacgccatt 900

ctgctgagtg atattctcgc agtgaacacg gagatcacca aagctccgct gagcgctagt 960

atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc ccttgtcaga 1020

cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggctacgcc 1080

ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gcccattctg 1140

gaaaaaatgg acggcaccca ggagctgctg gttaaagctta acagagaaga tctgttcgcg 1200

aaacagcgca ctttcgacaa tggaagcatc cccaccaga ttcacctggg cgaactgcac 1260

gctatcctca ggcggaaga ggatttctac ccctttttga aagataacag ggaaaagatt 1320

gagaaaaatc tcacatttcg gataccctac tatgtaggcc ccctcgcccg gggaaattcc 1380

agattcgcgt ggatgactcg caaatcagaa gagaccatca ctccctggaa cttcgaggaa 1440

gtcgtggata agggggcctc tgcccagtc ttcacgaaa ggatgactaa ctttgataaa	1500
aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacagta cttcacagtt	1560
tataacgagc tcaccaaggt caaatacgtc acagaaggga tgagaaagcc agcattcctg	1620
tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagttacc	1680
gtgaaacagc tcaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc	1740
agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc	1800
attaaagaca aggacttcct ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
ctcaccctta cgttggttga agataggag atgattgaag aacgcttgaa aacttacgct	1920
catctcttcg acgacaaagt catgaacag ctcaagggc gccgatatac aggatggggg	1980
cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg	2040
gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac	2100
tctctacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt	2160
cacgagcaca tcgctaactc tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220
gttaaggtcg tggatgaact cgtcaaagta atgggaaggc ataagccga gaatatcgtt	2280
atcgagatgg cccgagagaa ccaactacc cagaaggac agaagaacag tagggaaagg	2340
atgaagagga ttgaagagg tataaaagaa ctgggtccc aaatcctta ggaacacca	2400
gttgaaaaca ccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtct ccgactacga cgtggctgct	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaagcta gagggaagag tgataacgtc ccctcagaag aagtgtgcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggttcat caaaaggcag	2760
cttggtgaga cagccagat caccaagcac gtggcccaaa ttctcgattc acgatgaac	2820
accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct	2880
aagctggtct cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat	2940
taccaccatg cgcgatgac ctacctgaat gcagtgttag gcactgcact tatcaaaaaa	3000
tatcccaagc ttgaatcga atttgtttac ggagactata aagtgtacga tgttaggaaa	3060
atgatcgcaa agictgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc	3120
aatattatga attttttcaa gaccgagatt aactggcca atggagagat tcggaagcga	3180

ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgac	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct	3420
tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc	3480
aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac	3540
tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcatcattaa gcttcccaag	3600
tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg	3660
cagaaaggta acgagctggc actgccctct aaatacgtta atttcttgta tctggccagc	3720
cactatgaaa agctcaaagg gtctcccgaa gataatgagc agaagcagct gttcgtggaa	3780
caacacaaac actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagtg	3840
atcctcgccg acgctaacct cgataagggt ctttctgctt acaataagca cagggataag	3900
cccatcaggg agcaggcaga aaacattatc cacttggtta ctctgaccaa ctgggcgcg	3960
cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtacac ctctacaaag	4020
gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc	4080
gacctctctc agctcgggtg agac	4104
<210> 67	
<211> 1782	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> hyperactive PiggyBac (PB) transposase nucleic acid	
<400> 67	
atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgagggt agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac	240
aggatcctga ccctgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagcccc ccaggcggag cagggtgtcc gcctgaaca tcgtgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac ccctgtgtgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480

gagagcatga cctccgccac cttcagggac accaacgagg acgagatcta cgccttcttc 540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660

atcagatgcc ttaggatgga cgacaagagc atcaggccca cctgcggga gaacgacgtg 720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc 780
cctggcgccc acctgacat cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840
agggctctata tccccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac 900
agcggcacca agtatcatgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960
ggcgtgcccc tggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc 1020
agaaacatca cctgcgacaa ctgggtcacc agcatcccc tggccaagaa cctgctgcag 1080

gagccctaca agctgacat cgtgggcacc gtgagaagca acaagagaga gatccccgag 1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc 1200
ctgaccctgg tgcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac 1320
cagaccaagg gcggcggtga caccctggac cagatgtgca gcgtgatgac ctgcagcaga 1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
agcttcatca tctacagcca caacgtgagc agcaagggcg agaaggtgca gagccggaaa 1500

aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560
gccccacccc tgaagagata cctgagggac aacatcagca acatcctgcc caaagaggtg 1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 68

<211> 1020

<212> DNA

<213> Artificial Sequence

<220>

<223> hyperactive Sleeping Beauty (SB100) transposase nucleic acid

<400> 68

atgggaaaat caaaagaaat cagccaagac ctcagaaaaa gaattgtaga cctccacaag 60
tctgggtcat ccttgggagc aatttccaaa cgctggcggt taccacgttc atctgtacaa 120

acaatagtac gcaagtataa acaccatggg accacgcagc cgtcataccg ctcaggaagg 180
 agacgcgttc tgtctcctag agatgaacgt actttggtgc gaaaagtgc aatcaatccc 240
 agaacaacag caaaggacct tgtgaagatg ctggaggaaa caggtacaaa agtatctata 300
 tccacagtaa aacgagtcct atatcgacat aacctgaaag gccactcagc aaggaagaag 360
 ccactgctcc aaaaccgaca taagaaagcc agactacggt ttgcaactgc acatggggac 420

aaagatcgta ctttttggag aaatgtcctc tggctctgatg aaacaaaaat agaactgttt 480
 ggccataatg accatcgtaa tgtttggagg aagaaggggg aggcttgcaa gccgaagaac 540
 accatcccaa ccgtgaagca cgggggtggc agcatcatgt tgtgggggtg ctttgcigca 600
 ggagggactg gtgcacttca caaaatagat ggcatcatgg acgccgtgca gtatgtggat 660
 atattgaagc aacatctcaa gacatcagtc aggaagttaa agcttggtcg caaatgggtc 720
 ttccaacagc acaatgacce caagcatact tccaaagtgt tggcaaaatg gcttaaggac 780
 aacaaagtca aggtattgga gtggccatca caaagccctg acctcaatcc tatagaaaat 840

ttgtgggcag aactgaaaaa gcgtgtgcga gcaaggaggc ctacaaacct gactcagtta 900
 caccagctct gtcaggagga atgggccaaa attcacccaa attattgtgg gaagcttggtg 960
 gaaggctacc cgaacgttt gaccaagt tt aaacaattta aaggcaatgc taccaatac 1020

<210> 69

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<223> human Cas9 (hCas9) amino acid

<400> 69

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

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85

90

95

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

100

105

110

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

115

120

125

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

130

135

140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145

150

155

160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165

170

175

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

180

185

190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala

195

200

205

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

210

215

220

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

225

230

235

240

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

245

250

255

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

260

265

270

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

275

280

285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp

290

295

300

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

305

310

315

320

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

325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe

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

 405 410 415
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

 465 470 475 480
 Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495
 Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510
 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525
 Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln

 530 535 540
 Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560
 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

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

820 825 830
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys

835 840 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg

850 855 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865 870 875 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp

900 905 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr

915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp

930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser

945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe

995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys

1010 1015 1020
Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser

1025 1030 1035 1040
Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu

1045 1050 1055
Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile

1060 1065 1070

Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser

1075 1080 1085

Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly

1090 1095 1100

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

1105 1110 1115 1120

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser

1125 1130 1135

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

1140 1145 1150

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

1155 1160 1165

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

1170 1175 1180

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

1185 1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275 1280

Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys

1285 1290 1295

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu

1300 1305 1310

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

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

<210> 70

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<223> nickase Cas9 (nCas9) amino acid

<400> 70

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

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

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125
 His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160
 Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

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

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255
 Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270
 Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300

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

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg

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

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

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

 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575
 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590
 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605
 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620

 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640

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

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

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

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

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

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

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

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

Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser
1075 1080 1085
Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly
1090 1095 1100
Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
1105 1110 1115 1120
Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser
1125 1130 1135

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

1140 1145 1150

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

1155 1160 1165

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

1170 1175 1180

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

1185 1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275 1280

Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys

1285 1290 1295

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu

1300 1305 1310

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

1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360

Asp Leu Ser Gln Leu Gly Gly Asp

1365

<210> 71

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<223> dead Cas9 (dCas9) amino acid

<400> 71

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

1 5 10 15
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

35 40 45
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

50 55 60
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr

180 185 190
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala

195 200 205

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

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

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

 580 585 590
 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605
 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640
 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr

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

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

705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

725 730 735

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

740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

755 760 765

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

770 775 780

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

785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu

805 810 815

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

820 825 830

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

835 840 845

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

850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

 965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

 980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe

 995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys

 1010 1015 1020
Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser

1025 1030 1035 1040
Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu

 1045 1050 1055
Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile

 1060 1065 1070
Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser

 1075 1080 1085
Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly

 1090 1095 1100
Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile

1105 1110 1115 1120
Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser

 1125 1130 1135
Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly

 1140 1145 1150
Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile

 1155 1160 1165
Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala

 1170 1175 1180
Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys

1185 1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275 1280

Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys

1285 1290 1295

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu

1300 1305 1310

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

1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360

Asp Leu Ser Gln Leu Gly Gly Asp

1365

<210> 72

<211> 1069

<212> PRT

<213> Staphylococcus aureus

<400> 72

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile

65 70 75 80

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His

85 90 95

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu

100 105 110

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu

115 120 125

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

130 135 140

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

145 150 155 160

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

165 170 175

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr

180 185 190

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln

195 200 205

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

210 215 220

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys

225 230 235 240

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe

245 250 255

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr

260 265 270

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn

275 280 285
Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe

290 295 300
Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
305 310 315 320

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
325 330 335

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
340 345 350

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

355 360 365
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
370 375 380

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
385 390 395 400

Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
405 410 415

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala

420 425 430
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
435 440 445

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
450 455 460

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
465 470 475 480

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

485 490 495
Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
500 505 510

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
515 520 525

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
530 535 540

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu

545 550 555 560

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
565 570 575

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
580 585 590

Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
595 600 605

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile

610 615 620

Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
625 630 635 640

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
645 650 655

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
660 665 670

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys

675 680 685

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
690 695 700

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
705 710 715 720

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
725 730 735

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys

740 745 750

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
755 760 765

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp

770 775 780
 Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
 785 790 795 800
 Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu

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

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

 930 935 940
 Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
 945 950 955 960
 Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
 965 970 975
 Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
 980 985 990
 Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile

 995 1000 1005
 Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
 1010 1015 1020

Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr
 1025 1030 1035 1040
 Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr Glu
 1045 1050 1055
 Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly

1060 1065
 <210> 73
 <211> 340
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> hyperactive Sleeping Beauty (SB100) transposase amino acid
 <400> 73

Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Arg Ile Val
 1 5 10 15
 Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
 20 25 30
 Ala Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His

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

100 105 110
 Lys Gly His Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
 115 120 125
 Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
 130 135 140
 Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe

145 150 155 160
Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys

 165 170 175
Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile

 180 185 190
Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys

 195 200 205
Ile Asp Gly Ile Met Asp Ala Val Gln Tyr Val Asp Ile Leu Lys Gln

 210 215 220
His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val

225 230 235 240
Phe Gln His Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys

 245 250 255
Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser

 260 265 270
Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg

 275 280 285
Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys

 290 295 300
Gln Glu Glu Trp Ala Lys Ile His Pro Asn Tyr Cys Gly Lys Leu Val

305 310 315 320
Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn

 325 330 335
Ala Thr Lys Tyr

340

<210> 74

<211> 1236

<212> PRT

<213> Artificial Sequence

<220>

<223> cpf1

<400> 74

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

1 5 10 15

Asn Cys Tyr Ser Leu Ser Lys Thr Leu Arg Phe Lys Ala Ile Pro Val

20 25 30

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

35 40 45

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

50 55 60

Tyr Tyr Leu Ser Phe Ile Asn Asp Val Leu His Ser Ile Lys Leu Lys

65 70 75 80

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

85 90 95

Lys Glu Asn Lys Glu Leu Glu Asn Leu Glu Ile Asn Leu Arg Lys Glu

100 105 110

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

115 120 125

Lys Asp Ile Ile Glu Thr Ile Leu Pro Glu Phe Leu Asp Asp Lys Asp

130 135 140

Glu Ile Ala Leu Val Asn Ser Phe Asn Gly Phe Thr Thr Ala Phe Thr

145 150 155 160

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

165 170 175

Thr Ser Ile Ala Phe Arg Cys Ile Asn Glu Asn Leu Thr Arg Tyr Ile

180 185 190

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

195 200 205

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

210 215 220

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

225 230 235 240

Gly Ile Asp Val Tyr Asn Ala Ile Ile Gly Gly Phe Val Thr Glu Ser

245 250 255
Gly Glu Lys Ile Lys Gly Leu Asn Glu Tyr Ile Asn Leu Tyr Asn Gln

260 265 270
Lys Thr Lys Gln Lys Leu Pro Lys Phe Lys Pro Leu Tyr Lys Gln Val

275 280 285
Leu Ser Asp Arg Glu Ser Leu Ser Phe Tyr Gly Glu Gly Tyr Thr Ser

290 295 300
Asp Glu Glu Val Leu Glu Val Phe Arg Asn Thr Leu Asn Lys Asn Ser

305 310 315 320
Glu Ile Phe Ser Ser Ile Lys Lys Leu Glu Lys Leu Phe Lys Asn Phe

325 330 335
Asp Glu Tyr Ser Ser Ala Gly Ile Phe Val Lys Asn Gly Pro Ala Ile

340 345 350
Ser Thr Ile Ser Lys Asp Ile Phe Gly Glu Trp Asn Val Ile Arg Asp

355 360 365
Lys Trp Asn Ala Glu Tyr Asp Asp Ile His Leu Lys Lys Lys Ala Val

370 375 380
Val Thr Glu Lys Tyr Glu Asp Asp Arg Arg Lys Ser Phe Lys Lys Ile

385 390 395 400
Gly Ser Phe Ser Leu Glu Gln Leu Gln Glu Tyr Ala Asp Ala Asp Leu

405 410 415
Ser Val Val Glu Lys Leu Lys Glu Ile Ile Ile Gln Lys Val Asp Glu

420 425 430
Ile Tyr Lys Val Tyr Gly Ser Ser Glu Lys Leu Phe Asp Ala Asp Phe

435 440 445
Val Leu Glu Lys Ser Leu Lys Lys Asn Asp Ala Val Val Ala Ile Met

450 455 460
Lys Asp Leu Leu Asp Ser Val Lys Ser Phe Glu Asn Tyr Ile Lys Ala

465 470 475 480
Phe Phe Gly Glu Gly Lys Glu Thr Asn Arg Asp Glu Ser Phe Tyr Gly

485 490 495

Asp Phe Val Leu Ala Tyr Asp Ile Leu Leu Lys Val Asp His Ile Tyr
 500 505 510
 Asp Ala Ile Arg Asn Tyr Val Thr Gln Lys Pro Tyr Ser Lys Asp Lys
 515 520 525
 Phe Lys Leu Tyr Phe Gln Asn Pro Gln Phe Met Gly Gly Trp Asp Lys
 530 535 540
 Asp Lys Glu Thr Asp Tyr Arg Ala Thr Ile Leu Arg Tyr Gly Ser Lys
 545 550 555 560
 Tyr Tyr Leu Ala Ile Met Asp Lys Lys Tyr Ala Lys Cys Leu Gln Lys
 565 570 575
 Ile Asp Lys Asp Asp Val Asn Gly Asn Tyr Glu Lys Ile Asn Tyr Lys
 580 585 590
 Leu Leu Pro Gly Pro Asn Lys Met Leu Pro Lys Val Phe Phe Ser Lys
 595 600 605
 Lys Trp Met Ala Tyr Tyr Asn Pro Ser Glu Asp Ile Gln Lys Ile Tyr
 610 615 620
 Lys Asn Gly Thr Phe Lys Lys Gly Asp Met Phe Asn Leu Asn Asp Cys
 625 630 635 640
 His Lys Leu Ile Asp Phe Phe Lys Asp Ser Ile Ser Arg Tyr Pro Lys
 645 650 655
 Trp Ser Asn Ala Tyr Asp Phe Asn Phe Ser Glu Thr Glu Lys Tyr Lys
 660 665 670
 Asp Ile Ala Gly Phe Tyr Arg Glu Val Glu Glu Gln Gly Tyr Lys Val
 675 680 685
 Ser Phe Glu Ser Ala Ser Lys Lys Glu Val Asp Lys Leu Val Glu Glu
 690 695 700
 Gly Lys Leu Tyr Met Phe Gln Ile Tyr Asn Lys Asp Phe Ser Asp Lys
 705 710 715 720
 Ser His Gly Thr Pro Asn Leu His Thr Met Tyr Phe Lys Leu Leu Phe
 725 730 735
 Asp Glu Asn Asn His Gly Gln Ile Arg Leu Ser Gly Gly Ala Glu Leu

740 745 750
 Phe Met Arg Arg Ala Ser Leu Lys Lys Glu Glu Leu Val Val His Pro
 755 760 765
 Ala Asn Ser Pro Ile Ala Asn Lys Asn Pro Asp Asn Pro Lys Lys Thr

 770 775 780
 Thr Thr Leu Ser Tyr Asp Val Tyr Lys Asp Lys Arg Phe Ser Glu Asp
 785 790 795 800
 Gln Tyr Glu Leu His Ile Pro Ile Ala Ile Asn Lys Cys Pro Lys Asn
 805 810 815
 Ile Phe Lys Ile Asn Thr Glu Val Arg Val Leu Leu Lys His Asp Asp
 820 825 830
 Asn Pro Tyr Val Ile Gly Ile Asp Arg Gly Glu Arg Asn Leu Leu Tyr

 835 840 845
 Ile Val Val Val Asp Gly Lys Gly Asn Ile Val Glu Gln Tyr Ser Leu
 850 855 860
 Asn Glu Ile Ile Asn Asn Phe Asn Gly Ile Arg Ile Lys Thr Asp Tyr
 865 870 875 880
 His Ser Leu Leu Asp Lys Lys Glu Lys Glu Arg Phe Glu Ala Arg Gln
 885 890 895
 Asn Trp Thr Ser Ile Glu Asn Ile Lys Glu Leu Lys Ala Gly Tyr Ile

 900 905 910
 Ser Gln Val Val His Lys Ile Cys Glu Leu Val Glu Lys Tyr Asp Ala
 915 920 925
 Val Ile Ala Leu Glu Asp Leu Asn Ser Gly Phe Lys Asn Ser Arg Val
 930 935 940
 Lys Val Glu Lys Gln Val Tyr Gln Lys Phe Glu Lys Met Leu Ile Asp
 945 950 955 960
 Lys Leu Asn Tyr Met Val Asp Lys Lys Ser Asn Pro Cys Ala Thr Gly

 965 970 975
 Gly Ala Leu Lys Gly Tyr Gln Ile Thr Asn Lys Phe Glu Ser Phe Lys
 980 985 990

Ser Met Ser Thr Gln Asn Gly Phe Ile Phe Tyr Ile Pro Ala Trp Leu

995 1000 1005

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

1010 1015 1020

Lys Tyr Thr Ser Ile Ala Asp Ser Lys Lys Phe Ile Ser Ser Phe Asp

1025 1030 1035 1040

Arg Ile Met Tyr Val Pro Glu Glu Asp Leu Phe Glu Phe Ala Leu Asp

1045 1050 1055

Tyr Lys Asn Phe Ser Arg Thr Asp Ala Asp Tyr Ile Lys Lys Trp Lys

1060 1065 1070

Leu Tyr Ser Tyr Gly Asn Arg Ile Arg Ile Phe Arg Asn Pro Lys Lys

1075 1080 1085

Asn Asn Val Phe Asp Trp Glu Glu Val Cys Leu Thr Ser Ala Tyr Lys

1090 1095 1100

Glu Leu Phe Asn Lys Tyr Gly Ile Asn Tyr Gln Gln Gly Asp Ile Arg

1105 1110 1115 1120

Ala Leu Leu Cys Glu Gln Ser Asp Lys Ala Phe Tyr Ser Ser Phe Met

1125 1130 1135

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

1140 1145 1150

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

1155 1160 1165

Phe Tyr Asp Ser Arg Asn Tyr Glu Ala Gln Glu Asn Ala Ile Leu Pro

1170 1175 1180

Lys Asn Ala Asp Ala Asn Gly Ala Tyr Asn Ile Ala Arg Lys Val Leu

1185 1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Ser Val Lys His

1235

<210> 75

<211> 990

<212> PRT

<213> Artificial Sequence

<220>

<223> CasX

<400> 75

Met Ala Pro Lys Lys Lys Arg Lys Val Ser Met Gln Glu Ile Lys Arg

1 5 10 15

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

20 25 30

Ala Gly Lys Thr Gly Pro Met Lys Thr Leu Leu Val Arg Val Met Thr

35 40 45

Pro Asp Leu Arg Glu Arg Leu Glu Asn Leu Arg Lys Lys Pro Glu Asn

50 55 60

Ile Pro Gln Pro Ile Ser Asn Thr Ser Arg Ala Asn Leu Asn Lys Leu

65 70 75 80

Leu Thr Asp Tyr Thr Glu Met Lys Lys Ala Ile Leu His Val Tyr Trp

85 90 95

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

100 105 110

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

115 120 125

Gly Asn Glu Arg Leu Thr Ser Ser Gly Phe Ala Cys Ser Gln Cys Cys

130 135 140

Gln Pro Leu Tyr Val Tyr Lys Leu Glu Gln Val Asn Asp Lys Gly Lys

145 150 155 160

Pro His Thr Asn Tyr Phe Gly Arg Cys Asn Val Ser Glu His Glu Arg

165 170 175

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

180 185 190

Thr Tyr Ser Leu Gly Lys Phe Gly Gln Arg Ala Leu Asp Phe Tyr Ser
 195 200 205
 Ile His Val Thr Arg Glu Ser Asn His Pro Val Lys Pro Leu Glu Gln
 210 215 220
 Ile Gly Gly Asn Ser Cys Ala Ser Gly Pro Val Gly Lys Ala Leu Ser
 225 230 235 240
 Asp Ala Cys Met Gly Ala Val Ala Ser Phe Leu Thr Lys Tyr Gln Asp
 245 250 255
 Ile Ile Leu Glu His Gln Lys Val Ile Lys Lys Asn Glu Lys Arg Leu
 260 265 270
 Ala Asn Leu Lys Asp Ile Ala Ser Ala Asn Gly Leu Ala Phe Pro Lys
 275 280 285
 Ile Thr Leu Pro Pro Gln Pro His Thr Lys Glu Gly Ile Glu Ala Tyr
 290 295 300
 Asn Asn Val Val Ala Gln Ile Val Ile Trp Val Asn Leu Asn Leu Trp
 305 310 315 320
 Gln Lys Leu Lys Ile Gly Arg Asp Glu Ala Lys Pro Leu Gln Arg Leu
 325 330 335
 Lys Gly Phe Pro Ser Phe Pro Leu Val Glu Arg Gln Ala Asn Glu Val
 340 345 350
 Asp Trp Trp Asp Met Val Cys Asn Val Lys Lys Leu Ile Asn Glu Lys
 355 360 365
 Lys Glu Asp Gly Lys Val Phe Trp Gln Asn Leu Ala Gly Tyr Lys Arg
 370 375 380
 Gln Glu Ala Leu Leu Pro Tyr Leu Ser Ser Glu Glu Asp Arg Lys Lys
 385 390 395 400
 Gly Lys Lys Phe Ala Arg Tyr Gln Phe Gly Asp Leu Leu Leu His Leu
 405 410 415
 Glu Lys Lys His Gly Glu Asp Trp Gly Lys Val Tyr Asp Glu Ala Trp
 420 425 430
 Glu Arg Ile Asp Lys Lys Val Glu Gly Leu Ser Lys His Ile Lys Leu

435 440 445
 Glu Glu Glu Arg Arg Ser Glu Asp Ala Gln Ser Lys Ala Ala Leu Thr
 450 455 460
 Asp Trp Leu Arg Ala Lys Ala Ser Phe Val Ile Glu Gly Leu Lys Glu
 465 470 475 480
 Ala Asp Lys Asp Glu Phe Cys Arg Cys Glu Leu Lys Leu Gln Lys Trp

485 490 495
 Tyr Gly Asp Leu Arg Gly Lys Pro Phe Ala Ile Glu Ala Glu Asn Ser
 500 505 510
 Ile Leu Asp Ile Ser Gly Phe Ser Lys Gln Tyr Asn Cys Ala Phe Ile
 515 520 525
 Trp Gln Lys Asp Gly Val Lys Lys Leu Asn Leu Tyr Leu Ile Ile Asn
 530 535 540
 Tyr Phe Lys Gly Gly Lys Leu Arg Phe Lys Lys Ile Lys Pro Glu Ala

545 550 555 560
 Phe Glu Ala Asn Arg Phe Tyr Thr Val Ile Asn Lys Lys Ser Gly Glu
 565 570 575
 Ile Val Pro Met Glu Val Asn Phe Asn Phe Asp Asp Pro Asn Leu Ile
 580 585 590
 Ile Leu Pro Leu Ala Phe Gly Lys Arg Gln Gly Arg Glu Phe Ile Trp
 595 600 605
 Asn Asp Leu Leu Ser Leu Glu Thr Gly Ser Leu Lys Leu Ala Asn Gly

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

675 680 685

Ser Arg Phe Lys Asp Ser Leu Gly Asn Pro Thr His Ile Leu Arg Ile
 690 695 700
 Gly Glu Ser Tyr Lys Glu Lys Gln Arg Thr Ile Gln Ala Ala Lys Glu
 705 710 715 720
 Val Glu Gln Arg Arg Ala Gly Gly Tyr Ser Arg Lys Tyr Ala Ser Lys
 725 730 735
 Ala Lys Asn Leu Ala Asp Asp Met Val Arg Asn Thr Ala Arg Asp Leu
 740 745 750
 Leu Tyr Tyr Ala Val Thr Gln Asp Ala Met Leu Ile Phe Glu Asn Leu
 755 760 765
 Ser Arg Gly Phe Gly Arg Gln Gly Lys Arg Thr Phe Met Ala Glu Arg
 770 775 780
 Gln Tyr Thr Arg Met Glu Asp Trp Leu Thr Ala Lys Leu Ala Tyr Glu
 785 790 795 800
 Gly Leu Pro Ser Lys Thr Tyr Leu Ser Lys Thr Leu Ala Gln Tyr Thr
 805 810 815
 Ser Lys Thr Cys Ser Asn Cys Gly Phe Thr Ile Thr Ser Ala Asp Tyr
 820 825 830
 Asp Arg Val Leu Glu Lys Leu Lys Lys Thr Ala Thr Gly Trp Met Thr
 835 840 845
 Thr Ile Asn Gly Lys Glu Leu Lys Val Glu Gly Gln Ile Thr Tyr Tyr
 850 855 860
 Asn Arg Tyr Lys Arg Gln Asn Val Val Lys Asp Leu Ser Val Glu Leu
 865 870 875 880
 Asp Arg Leu Ser Glu Glu Ser Val Asn Asn Asp Ile Ser Ser Trp Thr
 885 890 895
 Lys Gly Arg Ser Gly Glu Ala Leu Ser Leu Leu Lys Lys Arg Phe Ser
 900 905 910
 His Arg Pro Val Gln Glu Lys Phe Val Cys Leu Asn Cys Gly Phe Glu
 915 920 925
 Thr His Ala Asp Glu Gln Ala Ala Leu Asn Ile Ala Arg Ser Trp Leu

930 935 940
 Phe Leu Arg Ser Gln Glu Tyr Lys Lys Tyr Gln Thr Asn Lys Thr Thr
 945 950 955 960
 Gly Asn Thr Asp Lys Arg Ala Phe Val Glu Thr Trp Gln Ser Phe Tyr
 965 970 975
 Arg Lys Lys Leu Lys Glu Val Trp Lys Pro Ala Val Thr Ser
 980 985 990
 <210> 76
 <211> 1069
 <212> PRT
 <213> Staphylococcus aureus

 <400> 76
 Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
 1 5 10 15
 Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
 20 25 30
 Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
 35 40 45
 Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
 50 55 60

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

 Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
 130 135 140
 Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
 145 150 155 160

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

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

 Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
 260 265 270
 Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
 275 280 285
 Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
 290 295 300
 Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
 305 310 315 320

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

 Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
 385 390 395 400
 Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile

405 410 415
 Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
 420 425 430
 Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
 435 440 445

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

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

 Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
 580 585 590
 Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
 595 600 605
 Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
 610 615 620
 Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
 625 630 635 640

 Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
 645 650 655

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

 Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
 705 710 715 720
 Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
 725 730 735
 Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
 740 745 750
 Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
 755 760 765

 Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
 770 775 780
 Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
 785 790 795 800
 Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
 805 810 815
 Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
 820 825 830

 Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
 835 840 845
 Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
 850 855 860
 Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
 865 870 875 880
 Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
 885 890 895

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

900 905 910
Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
915 920 925
Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
930 935 940
Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
945 950 955 960

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
965 970 975
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
980 985 990
Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
995 1000 1005
Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
1010 1015 1020

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

<210> 77

<211> 3207

<212> DNA

<213> Staphylococcus aureus

<400> 77

atggcccca aagaagcgc gaagtcggt atccacggag tcccagcagc caagcggaac 60

tacatcctgg gcctggacat cggcatcacc agcgtgggct acggcatcat cgactacgag 120

acacgggacg tgatcgatgc cggcgtgcgg ctgttcaaag aggccaacgt ggaaaacaac 180

gagggcagcg ggagcaagag aggcgccaga aggctgaagc ggcggaggcg gcatagaatc 240

cagagagtga agaagctgct gtctgactac aacctgctga ccgaccacag cgagctgagc 300

ggcatcaacc cctacgagcg cagagtgaag ggcttagacc agaagctgag cgaggaagag 360

ttctctgccg cctgtctgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg	420
gaagaggaca cgggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc	480
ctggaagaga aatacgtggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagtg	540
cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg	600
aaggtgcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg	660
ctggaaaccc ggcggaccta ctatgaggga cctggcgagg gcagcccctt cggctggaag	720
gacatcaaag aatggtacga gatgtgatg ggccactgca cctacttccc cgaggaaactg	780
cggagcgtga agtacgccta caacccgac ctgtacaacg ccctgaacga cctgaacaat	840
ctcgtgatca ccagggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc	900
gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatcctc	960
gtgaacgaag aggatattaa gggctacaga gtgaccagca cgggcaagcc cgagttcacc	1020
aacctgaagg tgiaccagca catcaaggac attaccgccc ggaaagagat tattgagaac	1080
gccgagctgc tggatcagat tgccaagatc ctgaccatct accagagcag cgaggacatc	1140
caggaagaac tgaccaatct gaactccgag ctgaccaggg aagagatcga gcagatctct	1200
aatctgaagg gctataccgg caccacaac ctgagcctga aggccatcaa cctgatcctg	1260
gacgagctgt ggcacaccaa cgacaaccag atcgtatct tcaaccggct gaagctggtg	1320
cccaagaagg tggacctgtc ccagcagaaa gagatcccca ccacctggt ggacgacttc	1380
atcctgagcc ccgtcgtgaa gagaagcttc atccagagca tcaaagtgat caacgccatc	1440
atcaagaagt acggcctgcc caacgacatc attatcgagc tggcccgcga gaagaactcc	1500
aaggacgccc agaaaatgat caacgagatg cagaagcgga accggcagac caacgagcgg	1560
atcgaggaaa tcatccggac caccggcaaa gagaacgcca agtacctgat cgagaagatc	1620
aagctgcacg acatgcagga aggcaagtgc ctgtacagcc tggaagccat ccctctggaa	1680
gatctgctga acaaccctt caactatgag gtggaccaca tcatccccag aagcgtgtcc	1740
ttcgacaaca gcttcaacaa caagtgctc gtgaagcagg aagaaaacag caagaagggc	1800
aaccggaccc cattccagta cctgagcagc agcgacagca agatcagcta cgaaaccttc	1860
aagaagcaca tcctgaatct ggccaagggc aagggcagaa tcagcaagac caagaaagag	1920
tatctgctgg aagaacggga catcaacagg ttctccgtgc agaaagactt catcaaccgg	1980
aacctggtgg ataccagata cgccaccaga ggccctgatga acctgctgcg gagctacttc	2040
agagtgaaca acctggacgt gaaagtgaag tccatcaatg gcggcttcac cagctttctg	2100

cggcggaagt ggaagtttaa gaaagagcgg aacaaggggt acaagcacca cgccgaggac	2160
gcctgatca ttgccaacgc cgatttcac ttcaaagagt ggaagaaact ggacaaggcc	2220
aaaaaagtga tggaaaacca gatgttcgag gaaaagcagg ccgagagcat gcccagatc	2280
gaaaccgagc aggagtacaa agagatcttc atcaccccc accagatcaa gcacattaag	2340
gacttcaagg actacaagta cagccaccgg gtggacaaga agcctaatag agagctgatt	2400
aacgacaccc tgtactccac ccggaaggac gacaagggca acaccctgat cgtgaacaat	2460
ctgaacggcc tgiacgacaa ggacaatgac aagctgaaaa agctgatcaa caagagcccc	2520
gaaaagctgc tgatgtacca ccacgacccc cagacctacc agaaactgaa gctgattatg	2580
gaacagtacg gcgacgagaa gaatccccig tacaagtact acgaggaaac cgggaactac	2640
ctgaccaagt actccaaaaa ggacaacggc cccgtgatca agaagattaa gtattacggc	2700
aacaaactga acgcccactt ggacatcacc gacgactacc ccaacagcag aaacaaggtc	2760
gtgaagctgt cctgaagcc ctacagattc gacgtgtacc tggacaatgg cgtgtacaag	2820
tctgtgaccg tgaagaatct ggatgtgatc aaaaaagaaa actactacga agtgaatagc	2880
aagtgtatg aggaagctaa gaagctgaag aagatcagca accaggccga gtttatcgcc	2940
tcctttaca acaacgatct gatcaagatc aacggcgagc tgtatagagt gatcggcgtg	3000
aacaacgacc tgctgaaccg gatcgaagtg aacatgatcg acatcaccta ccgcgagtac	3060
ctggaaaaca tgaacgacaa gaggcccccc aggatcatta agacaatcgc ctccaagacc	3120
cagagcatta agaagtacag cacagacatt ctgggcaacc tgtatgaagt gaaatctaag	3180
aagcacctc agatcatcaa aaagggc	3207
<210> 78	
<211> 3708	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> cpfl coding sequence	
<400> 78	
atggcccaaa agaagaagcg gaaggtcagc aagctggaga agtttacaaa ctgtactcc	60
ctgtctaaga ccctgaggtt caaggccatc cctgtgggca agaccagga gaacatcgac	120
aataagcggc tgctggtgga ggacgagaag agagccgagg attataaggc cgtgaagaag	180
ctgtcggatc gctactatct gtcttttacc aacgacgtgc tgcacagcat caagctgaag	240
aatctgaaca attacatcag cctgttccgg aagaaaacca gaaccgagaa ggagaataag	300

gagctggaga acctggagat caatctgcgg aaggagatcg ccaaggcctt caagggaac 360
 gagggctaca agtccctgtt taagaaggat atcatcgaga caatcctgcc agagtctctg 420
 gacgataagg acgagatcgc cctggtgaac agcttcaatg gctttaccac agccttcacc 480
 ggcttctttg ataacagaga gaatatgttt tccgaggagg ccaagagcac atccatcgcc 540

 ttcaggtgta tcaacgagaa tctgaccgc tacatctcta atatggacat cttcgagaag 600
 gtggacgcca tctttgataa gcacgagggtg caggagatca aggagaagat cctgaacagc 660
 gactatgatg tggaggattt ctttgagggc gagttcttta actttgtgct gacacaggag 720
 ggcatcgacg tgtataacgc catcatcggc ggcttcgtga ccgagagcgg cgagaagatc 780
 aagggcctga acgagtacat caacctgtat aatcagaaaa ccaagcagaa gctgcctaag 840
 ttttaagccac tgtataagca ggtgctgagc gatcgggagt ctctgagctt ctacggcgag 900
 ggctatacat ccgatgagga ggtgctggag gtgttttagaa acaccctgaa caagaacagc 960

 gagatcttca gctccatcaa gaagctggag aagctgttca agaattttga cgagtactct 1020
 agcgccggca tctttgtgaa gaacggcccc gccatcagca caatctccaa ggatatcttc 1080
 ggcgagtggg acgtgatccg ggacaagtgg aatgccgagt atgacgatat ccacctgaag 1140
 aagaaggccg tggtagccga gaagtagcag gacgatcgga gaaagtcctt caagaagatc 1200
 ggctcctttt ctctggagca gctgcaggag tacgccgacg ccgatctgtc tgtggtggag 1260
 aagctgaagg agatcatcat ccagaagggt gatgagatct acaagggtga tggctcctct 1320
 gagaagctgt tcgacgccga ttttgtgctg gagaagagcc tgaagaagaa cgacgccgtg 1380

 gtggccatca tgaaggacct gctggattct gtgaagagct tcgagaatta catcaaggcc 1440
 ttctttggcg agggcaagga gacaacagg gacgagtcct tctatggcga ttttgtgctg 1500
 gcctacgaca tctgtctgaa ggtggaccac atctacgatg ccatccgcaa ttatgtgacc 1560
 cagaagccct actctaagga taagttcaag ctgtattttc agaaccctca gttcatgggc 1620
 ggctgggaca aggataagga gacagactat cgggccacca tcctgagata cggctccaag 1680
 tactatctgg ccatcatgga taagaagtac gccaaagtcc tgcagaagat cgacaaggac 1740
 gatgtgaacg gcaattacga gaagatcaac tataagctgc tgcccggccc taataagatg 1800

 ctgccaaagg tgttcttttc taagaagtgg atggcctact ataaccacag cgaggacatc 1860
 cagaagatct acaagaatgg cacattcaag aaggcgata tgtttaacct gaatgactgt 1920
 cacaagctga tcgacttctt taaggatagc atctcccgtt atccaaagtg gtccaatgcc 1980
 tacgatttca acttttctga gacagagaag tataaggaca tcgccggctt ttacagagag 2040
 gtggaggagc agggctataa ggtgagcttc gactctgcca gcaagaagga ggtggataag 2100
 ctggtggagg agggcaagct gtatatgttc cagatctata acaaggactt ttccgataag 2160

tctcacggca cacccaatct gcacaccatg tacttcaagc tgctgtttga cgagaacaat 2220

cacggacaga tcaggctgag cggaggagca gagctgttca tgaggcgcgc ctccctgaag 2280

aaggaggagc tgggtggtgca cccagccaac tccccatatcg ccaacaagaa tccagataat 2340

ccaagaaaa ccacaacct gtcctacgac gtgtataagg ataagagggt tcttgaggac 2400

cagtacgagc tgcacatccc aatcgccatc aataagtgcc ccaagaacat cttcaagatc 2460

aatacagagg tgcgcgtgct gctgaagcac gacgataacc cctatgtgat cggcatcgat 2520

aggggcgagc gcaatctgct gtatatctgt gtggtggacg gcaagggcaa catcgtggag 2580

cagtattccc tgaacgagat catcaacaac ttcaacggca tcaggatcaa gacagattac 2640

cactctctgc tggacaagaa ggagaaggag aggttcgagg cccgccagaa ctggacctcc 2700

atcgagaata tcaaggagct gaaggccggc tatatctctc aggtggtgca caagatctgc 2760

gagctggtgg agaagtacga tgccgtgatc gccctggagg acctgaactc tggctttaag 2820

aatagccgcg tgaagggtga gaagcagggt tatcagaagt tcgagaagat gctgatcgat 2880

aagctgaact acatggtgga caagaagtct aatccttgtg caacaggcgg cggcctgaag 2940

ggctatcaga tcaccaataa gttcgagagc ttaagtcca tgtctacca gaacggcttc 3000

atcttttaca tccctgcctg gctgacatcc aagatcgatc catctaccgg ctttgtgaac 3060

ctgctgaaaa ccaagtatac cagcatcgcc gattccaaga agttcatcag ctcccttgac 3120

aggatcatgt acgtgcccgga ggaggatctg ttcgagtttg cctggacta taagaacttc 3180

tctcgcacag acgccgatta catcaagaag tggaagctgt actcctacgg caaccggatc 3240

agaatcttcc ggaatcctaa gaagaacaac gtgttcgact gggaggaggt gtgcctgacc 3300

agcgcctata aggagctgtt caacaagtac ggcatcaatt atcagcaggg cgatatcaga 3360

gccctgctgt gcgagcagtc cgacaaggcc ttctactcta gctttatggc cctgatgagc 3420

ctgatgctgc agatgcggaa cagcatcaca ggccgcaccg acgtggattt tctgatcagc 3480

cctgtgaaga actccgacgg catcttctac gatagccgga actatgaggc ccaggagaat 3540

gccatcctgc caaagaacgc gcagccaat ggcgccata acatcgccag aaaggtgctg 3600

tgggccatcg gccagttcaa gaaggccgag gacgagaagc tggataaggt gaagatcgcc 3660

atctctaaca aggagtggct ggagtacgcc cagaccagcg tgaagcac 3708

<210> 79

<211> 2970

<212> DNA

<213> Artificial Sequence

<220>

<223> CasX coding sequence

<400> 79

atggcccaa agaagaagcg gaaggtcagc atgcaagaga tcaagagaat caacaagatc	60
agaaggagac tggtaagga cagcaacaca aagaaggccg gcaagacagg ccccatgaaa	120
accctgctcg tcagagtgat gacccctgac ctgagagagc ggctggaaaa cctgagaaag	180
aagcccagaga acatccctca gcctatcagc aacaccagca gggccaacct gaacaagctg	240
ctgaccgact acaccgagat gaagaaagcc atcctgcacg tgtactggga agagtccag	300
aaagaccccc tgggctgat gagcagagtt gctcagcccg ctctaagaa catcgaccag	360
agaaagctga tccccgtgaa ggacggcaac gagagactga cctctagcgg ctttgcctgc	420
agccagtgtt gccagcctct gtacgtgtac aagctggaac aagtgaacga caagggaag	480
ccccacacca actacttcgg cagatgcaac gtgtccgagc acgagaggct gatcctgctg	540
tctcctcaca agcccagggc caacgatgag ctggtcacat acagcctggg caagttcgga	600
cagagagccc tggacttcta cagcatccac gtgaccaggg agagcaatca ccctgtgaag	660
cccctggaac agatcggcgg caatagctgt gcctctggac ctgtgggaaa agccctgagc	720
gacgcctgta tgggagccgt ggcatccttc ctgaccaagt accaggacat catcctggaa	780
caccagaaag tgatcaagaa gaacgagaaa agactggcca acctcaagga tatcgccagc	840
gctaacggcc tggcctttcc taagatcacc ctgcctccac agcctcacac caaagagggc	900
atcgaggcct acaacaacgt ggtggcccag atcgtgattt gggtaacct gaatctgtgg	960
cagaagctga agatcggcag ggacgaagcc aagccactgc agagactgaa gggcttcctt	1020
agcttccttc tgggtgaaag acaggccaat gaagtggatt ggtgggacat ggtctgcaac	1080
gtgaagaagc tgatcaagca gaagaaagag gatggcaagg ttttctggca gaacctggcc	1140
ggctacaaga gacaagaagc cctgctgcct tacctgagca gcgaagagga ccggaagaag	1200
ggcaagaagt tcgccagata ccagttcggc gacctgctgc tgcacctgga aaagaagcac	1260
ggcgaggact ggggcaaagt gtacgatgag gcctgggaga gaatcgaca gaaggtggaa	1320
ggcctgagca agcacattaa gctggaagag gaaagaagga gcgaggacgc ccaatctaaa	1380
gccgctctga ccgattggct gagagccaag gccagctttg tgatcgaggg cctgaaagag	1440
gccgacaagg acgagttctg cagatgcgag ctgaagctgc agaagtggta cggcgatctg	1500
agaggcaagc ccttcgcat tgaggccgag aacagcatcc tggacatcag cggcttcagc	1560
aagcagtaca actgcgcctt catttggcag aaagacggcg tcaagaaact gaacctgtac	1620
ctgatcatca attacttcaa aggcggcaag ctgcggttca agaagatcaa acccgaggcc	1680
ttcgaggcta acagattcta caccgtgatc acaaaaaagt ccggcgagat cgtgcccattg	1740

gaagtgaact tcaacttcga cgaccccaac ctgattatcc tgcctctggc cttcggaag 1800

agacaggga gagagttcat ctggaacgat ctgctgagcc tggaaaccgg ctctctgaag 1860

ctggccaatg gcagagtgat cgagaaaacc ctgtacaaca ggagaaccag acaggacgag 1920

cctgctctgt ttgtggccct gaccttcgag agaagagagg tgctggacag cagcaacatc 1980

aagcccatga acctgatcgg catcgaccgg ggcgagaata tccctgctgt gatcgccctg 2040

acagaccctg aaggatgccc actgagcaga ttcaaggact ccctgggcaa ccctacacac 2100

atcctgagaa tcggcgagag ctacaaagag aagcagagga caatccaggc cgccaaagag 2160

gtggaacaga gaagagccgg cggatactct aggaagtacg ccagcaaggc caagaatctg 2220

gccgacgaca tgggccgaaa caccgccaga gatctgctgt actacgccgt gacacaggac 2280

gccatgctga tcttcgagaa tctgagcaga ggcttcggcc ggcagggcaa gagaaccttt 2340

atggccgaga ggcagtacac cagaatggaa gattggctca cagctaaact ggcctacgag 2400

ggactgcccc gcaagacctt cctgtccaaa aactggccc agtatacctc caagacctgc 2460

agcaattgcg gcttcacat caccagcgcc gactacgaca gagtgtgga aaagctcaag 2520

aaaaccgcca cgggctggat gaccaccatc aacggcaaag agctgaaggt tgagggccag 2580

atcacctact acaacaggta caagaggcag aacgtcgtga aggatctgag cgtggaactg 2640

gacagactga gcaagagag cgtgaacaac gacatcagca gctggacaaa gggcagatca 2700

ggcgaggctc tgagcctgct gaagaagagg tttagccaca gacctgtgca agagaagttc 2760

gtgtgctga actgcggctt cgagacacac gccgatgaac aggtcgccct gaacattgcc 2820

agaagctggc tgttcctgag aagccaagag tacaagaagt accagaccaa caagaccacc 2880

ggcaacaccg acaagagggc ctttgtggaa acctggcaga gcttctacag aaaaaagctg 2940

aaagaagtct ggaagccgc cgtgactagt 2970

<210> 80

<211> 3213

<212> DNA

<213> Staphylococcus aureus

<400> 80

gccaccatgg ccccaaagaa gaagcggaag gtcggtatcc acggagtccc agcagccaag 60

cggaactaca tcttgggcct ggccatcggc atcaccagcg tgggctacgg catcatcgac 120

tacgagacac gggacgtgat cgatgccggc gtgcggctgt tcaaagagc caacgtggaa 180

aacaacgagg gcaggcggag caagagaggc gccagaaggc tgaagcggcg gaggcggcat 240

agaatccaga gagtgaagaa gctgctgttc gactacaacc tgctgaccga ccacagcgag 300

ctgagcggca tcaacccta cgaggccaga gtgaagggcc tgagccagaa gctgagcgag	360
gaagagtctt ctgccgcctt gctgcacctg gccaaagaaa gaggcgtgca caacgtgaac	420
gaggtggaag aggacaccgg caacgagctg tccaccaaag agcagatcag ccggaacagc	480
aaggcccttg aagagaaata cgtggccgaa ctgcagctgg aacggctgaa gaaagacggc	540
gaagtgcggg gcagcatcaa cagattcaag accagcgact acgtgaaaga agccaaacag	600
ctgctgaagg tgcagaaggc ctaccaccag ctggaccaga gcttcacga cacctacatc	660
gacctgctgg aaaccggcg gacctactat gagggacctg gcgagggcag ccccttcggc	720
tggaaggaca tcaagaatg gtacgagatg ctgatgggcc actgcaccta cttccccgag	780
gaactgcgga gcgtgaagta cgcctacaac gccgacctgt acaacgcctt gaacgacctg	840
aacaatctcg tgatcaccag ggacgagaac gagaagctgg aatattacga gaagttccag	900
atcatcgaga acgtgttcaa gcagaagaag aagccacccc tgaagcagat cgccaaagaa	960
atcctcgtga acgaagagga tattaagggc tacagagtga ccagcaccgg caagcccgag	1020
ttaccaacc tgaagggtga ccacgacatc aaggacatta ccgcccgaa agagattatt	1080
gagaacgccg agctgctgga tcagattgcc aagatcctga ccatctacca gagcagcgag	1140
gacatccagg aagaactgac caatctgaac tccgagctga cccaggaaga gatcgagcag	1200
atctctaate tgaagggtga taccggcacc cacaacctga gcctgaaggc catcaacctg	1260
atcctggacg agctgtggca caccaacgac aaccagatcg ctatcttcaa ccggctgaag	1320
ctggtgcccc agaagggtga cctgtcccag cagaaagaga tccccaccac cctggtggac	1380
gacttcaccc tgagccccgt cgtgaagaga agcttcaccc agagcatcaa agtgatcaac	1440
gccatcatca agaagtacgg cctgccccac gacatcatta tcgagctggc ccgcgagaag	1500
aactccaagg acgcccagaa aatgatcaac gagatgcaga agcggaaccg gcagaccaac	1560
gagcggatcg aggaaatcat ccggaccacc ggcaaagaga acgccaagta cctgatcgag	1620
aagatcaagc tgcacgacat gcaggaaggc aagtgcctgt acagcctgga agccatccct	1680
ctggaagatc tgctgaacaa ccccttcaac tatgaggtgg accacatcat cccagaagc	1740
gtgtccttcg acaacagctt caacaacaag gtgtctctga agcaggaaga aaacagcaag	1800
aagggaacc ggacccatt ccagtacctg agcagcagcg acagcaagat cagctacgaa	1860
accttcaaga agcacaacct gaatctggcc aagggaagg gcagaatcag caagaccaag	1920
aaagagtatc tgctggaaga acgggacatc aacaggttct ccgtgcagaa agacttcac	1980
aaccggaacc tgggtgatac cagatacgcc accagaggcc tgatgaacct gctgcggagc	2040

tacttcagag tgaacaacct ggacgtgaaa gtgaagtcca tcaatggcgg cttcaccagc 2100

tttctgcggc ggaagtggaa gtttaagaaa gagcgggaaca aggggtacaa gcaccacgcc 2160

gaggacgccc tgatcattgc caacccgat ttcattctca aagagtggaa gaaactggac 2220

aaggccaaaa aagtgatgga aaaccagatg ttcgaggaaa agcaggccga gagcatgccc 2280

gagatcgaaa ccgagcagga gtacaaagag atcttcatca cccccacca gatcaagcac 2340

attaaggact tcaaggacta caagtacagc caccgggtgg acaagaagcc taatagagag 2400

ctgattaacg acacctgta ctccaccggg aaggacgaca agggcaacac cctgatcgtg 2460

aacaatctga acggcctgta cgacaaggac aatgacaagc tgaaaaagct gatcaacaag 2520

agccccgaaa agctgctgat gtaccaccac gacccccaga cctaccagaa actgaagctg 2580

attatggaac agtacggcga cgagaagaat ccctgtaca agtactacga ggaaaccggg 2640

aactacatga ccaagtactc caaaaaggac aacggccccc tgatcaagaa gattaagtat 2700

tacggcaaca aactgaacgc ccatctggac atcaccgacg actaccccaa cagcagaaac 2760

aaggtcgtga agctgtccct gaagccctac agattcgacg tgtacctgga caatggcgtg 2820

tacaagttcg tgacctgaa gaattctggat gtgatcaaaa aagaaaacta ctacgaagtg 2880

aatagcaagt gctatgagga agctaagaag ctgaagaaga tcagcaacca ggccgagttt 2940

atcgctcct tctacaacaa cgatctgatc aagatcaacg gcgagctgta tagagtgatc 3000

ggcgtgaaca acgacctgct gaaccggatc gaagtgaaca tgatcgacat cacctaccgc 3060

gagtacctgg aaaacatgaa cgacaagagg cccccagga tcattaagac aatcgctcc 3120

aagaccaga gcattaagaa gtacagcaca gacattctgg gcaacctgta tgaagtgaaa 3180

tctaagaagc acctcagat catcaaaaag ggc 3213

<210> 81

<211> 2976

<212> DNA

<213> Campylobacter jejuni

<400> 81

atggcccaa agaagaagcg gaaggtcgcc agaatcctgg ccttcgacat cggcatcagc 60

agcatcggtt gggccttcag cgagaacgac gagctgaagg actgcggcgt gcggtcttc 120

accaaggtgg aaaaccccaa gaccggcgag agcctggccc tgcccagaag gctggccaga 180

agcggccgga agagactggc cagacggaag gcccggctga accacctgaa gcacctgatc 240

gccaacgagt tcaagctgaa ctacgaggac taccagagct tcgacgagtc cctggccaag 300

gcctacaagg gcagcctgat cagccctac gagctgcggt tccgggcctt gaacgagctg 360

ctgagcaagc aggacttcgc cagagtgatc ctgcacattg ccaagcggag aggctacgac	420
gacatcaaga acagcgacga caaagagaag ggcgccatcc tgaaggccat caagcagaac	480
gaggaaaagc tggccaacta ccagtcctgt ggcgagtacc tgtacaaaga gtacttcag	540
aagttcaaag agaacagcaa agaattcacc aacgtgcgga acaagaaaga aagctacgag	600
cgggtgatcg cccagagctt cctgaaggat gagctgaagc tgatcttcaa gaagcagaga	660
gagttcggct tcagcttcag caagaaattc gaggaagagg tgctgagcgt cgccttctac	720
aagagagccc tgaaggactt cagccacctc gtgggcaact gcagcttctt caccgacgag	780
aagagagccc ccaagaacag cccctggcc ttcattgtcg tggccctgac ccggatcatc	840
aacctgctga acaatctgaa gaacaccgag ggcatcctgt acaccaagga cgacctgaac	900
gccttctga atgaggtgct gaagaacggc accctgacct acaagcagac caagaagctg	960
ctgggcctga gcgacgacta cgagtttaag ggcgagaagg gcacctactt catcgagttc	1020
aagaagtaca aagagttcat caaggccctg ggcgagcaca acctgagcca ggacgatctg	1080
aatgagatcg ccaaggacat caccctgatc aaggacgaga ttaagctgaa gaaggccctg	1140
gccaatacgc acctgaatca gaaccagatc gacagcctga gcaagctgga attcaaggat	1200
cacctgaaca tcagcttcaa ggctctgaag ctggtcacc cctgatgct ggaaggcaag	1260
aagtacgacg aggcctgcaa cgagctgaac ctgaaggtgg ccatcaacga ggacaagaag	1320
gacttcctgc ccgcttcaa cgaaacctac tacaaggacg aagtaccaa cccctgggtg	1380
ctgcgggcca tcaaagaata ccggaagggt ctgaatgccc tgctcaagaa atacggcaag	1440
gtgcacaaga tcaacatcga gctggcccgg gaagtgggca agaaccacag ccagcgggcc	1500
aagatcgaga aagagcagaa cgaaaactac aaggccaaga aggacgctga gctggaatgc	1560
gagaagctgg gactgaagat caacagcaag aacatcctga agctgcggct gttcaaagaa	1620
cagaaagagt tctgcgcta cagcggcgag aagatcaaga tcagcgatct gcaggacgag	1680
aagatgctgg aaatcgacca catctacccc tacagccggt ccttcgacga cagctacatg	1740
aacaaggtgc tgggtgtcac caaacagaac caggaaaaac tgaaccagac ccccttcgag	1800
gccttcggca acgacagcgc caagtggcag aaaatcgagg tgctggccaa gaacctgccc	1860
accaagaaac agaagagaat cctggacaag aattacaagg acaaagagca gaagaacttc	1920
aaggaccgga acctgaacga caccgggtat atgcgccggt tgggtgctgaa ctacacaaag	1980
gactacctgg atttcctgcc cctgtccgac gacgagaaca ccaagctgaa cgataccag	2040
aaaggtcca aggtgcacgt ggaagccaag agcggcatgc tgaccagcgc cctgagacac	2100

acctggggct tcagcgccaa ggatcggaac aaccatctgc accacgcat cgacgccgtg 2160

atcattgcct acgccaacaa cagcatctg aaggccttct ccgacttcaa gaaagaacag 2220

gaaagcaaca gcgccgagct gtacccaag aagatctctg agctggacta caagaacaag 2280

cggaagtctt tcgagccctt cagcggttct cggcagaagg tgctggataa gatcgacgag 2340

atcttctgtt ccaagcccca gcggaagaag cctcttggcg cctgcacga ggaacatttc 2400

agaaaagagg aagagttcta ccagtcctac ggcggaag aaggcgtgct gaaggccctc 2460

gagctgggca agatcagaaa agtgaacggc aagatctga agaacgggga catgttccgg 2520

gtggacatct tcaagcaca aaagaccaac aagtctacg ccgtgcccac ctacaccatg 2580

gacttcgccc tgaagggtgt gccaacaag gccgtggccc ggtccaagaa gggcgagatc 2640

aaggactgga ttctgatgga cgagaactac gatttctgt ttagcctgta caaggactcc 2700

ctgatcctga tccagacca ggacatgcag gaacccgagt tcgtctacta caacgccttc 2760

accagcagca ccgtgtccct gatcgtgtct aagcacgaca acaagttcga gacactgagc 2820

aagaaccaga agatcctgtt caagaacgcc aacgagaaag aagtgatcgc caagagcatc 2880

ggcatccaga atctgaaggt gttcgagaag tacatcgtgt ccgccctggg agaagtgaca 2940

aaggccgagt tccggcagag agaggacttc aaaaag 2976

<210> 82

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac D450N/R372A/K375A

<400> 82

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60

ctggtcggcg aggacagcga cagcgaggtg agcgaccacg tgagcgagga cgacgtgcag 120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac 240

aggatcctga ccttgcacca gaggaccatc aggggcaaga acaagcactg ctggtccacc 300

tccaagccca ccaggcgag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac ccctgtgtgt gttcaagct gttcttcacc 420

gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

gagagcatga cctccgccac cttcagggac accaacgagg acgagatcta cgccttcttc 540

ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
atcagatgcc ttaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg 720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattcca gaactacacc 780

cctggcgccc acctgacctat cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840
agggtctata tccccaacaa gcccagcaag tacggcatca agatcctgat gatgtgcgac 900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc 1020
agaaacatca cctgcgacaa ctgggttcacc agcatcccc tggccaagaa cctgctgcag 1080
gagccctaca agctgacctat cgtgggcacc gtggccagca acgccagaga gatccccgag 1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc 1200

ctgacctggg tgtcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320
cagaccaagg gcggcgtgga caccctgaac cagatgtgca gcgtgatgac ctgcagcaga 1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
agttcatca tctacagcca caacgtgagc agcaaggcg agaaggtgca gagccggaaa 1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560
gcccccccc tgaagagata cctgaggggac aacatcagca acatcctgcc caaagaggtg 1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 83

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R275A/R277A

<400> 83

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60
ctggtcggcg aggacagcga cagcgaggtg agcgaccacg tgagcgagga cgacgtgcag 120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgac gagcagcccg gcagctccct ggccagcaac 240
aggatcctga cctgccccca gaggaccatc aggggcaaga acaagcactg ctggtccacc 300
tccaagcccc ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360
cccaccagga tgtgcaggaa catctacgac ccctgctgt gcttcaagct gtcttcacc 420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480
gagagcatga cctccgccac cttcaggga accaacgagg acgagatcta cgccttcttc 540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600

gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
atcagatgcc tgaggatgga cgacaagagc atcaggccca cctgcgggga gaacgacgtg 720
ttccccccg tgagaaagat ctgggacctg ttcatccacc agtgcattca gaactacacc 780
cctggcgccc acctgaccat cgacgagcag ctgctgggct tcgccggcgc ctgccccttc 840
agggtctata tccccaaaca gcccagcaag tacggcatca agatcctgat gatgtgcgac 900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgcca cggcagctgc 1020

agaaacatca cctgcgacaa ctgggtcacc agcatcccc tggccaagaa cctgctgcag 1080
gagccctaca agctgaccat cgtgggcacc gtgagaagca acaagagaga gatccccgag 1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc 1200
ctgacctgg tgcctacaa gcccagccc gccaaagtgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320
cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga 1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440

agcttcatca tctacagcca caacgtgagc agcaaggcg agaaggtgca gagccgaaa 1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560
gccccaccc tgaagagata cctgaggga aacatcagca acatcctgcc caaagaggtg 1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcggcagct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 84

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac K409A/K412A

<400> 84

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac	240
aggatcctga ccttgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480
gagagcatga cctccgccac cttcagggac accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc tgaggatgga cgacaagagc atcaggccca cctgcggga gaacgacgtg	720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc	780
cctggcgccc acctgacct cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggtctata tccccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac	900
agcggcacca agtacatgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc	1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacct cgtgggcacc gtgagaagca acaagagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggccccgtg ggcaccagca tgttctgctt cgacggcccc	1200
ctgacctgg tgtcctacaa gcccgcccc gccgcatgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac	1320
cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatgc ctgcatcaac	1440
agcttcatca tctacagcca caacgtgagc agcaaggcg agaagggtgca gagccgaaa	1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggctggag	1560
gccccacc tgaagagata cctgagggac aacatcagca acatcctgcc caaagagggtg	1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaaggagac ctactgcacc	1680

tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc	1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc	1782
<210> 85	
<211> 1782	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Engineered Piggybac R460A/K461A	
<400> 85	
atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgaggtg agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggcagcaac	240
aggatcctga cctgcacca gaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480
gagagcatga cctccgccac cttcaggac accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc tgaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg	720
ttcacccccg tgagaaagat ctgggacctg ttcaccacc agtgcatcca gaactacacc	780
cctggcgccc acctgacct cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggtctata tcccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac	900
agcggcacca agtacatgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgctca cggcagctgc	1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacct cgtgggcacc gtgagaagca acaagagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggccccgtg ggcaccagca tgttctgctt cgacggcccc	1200
ctgacctgg tgtcctacaa gcccaagccc gccaaagtgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac	1320
cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcgcc	1380

gccaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440

agcttcatca tctacagcca caacgtgagc agcaagggcg agaaggtgca gagccggaaa 1500

aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggttgag 1560

gccccaccc tgaagagata cctgaggac aacatcagca acatcctgcc caaagaggtg 1620

cccgccacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680

tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740

atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 86

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac S351A/R372A/K375A/R388A/
D450N/W465A/S573A/M589V/S594L

<400> 86

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60

ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag 120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac 240

aggatcctga ccctgcccc gagggaccatc aggggcaaga acaagcactg ctggtccacc 300

tccaagcccc ccaggcggag cagggtgtcc gcctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac cccctgctgt gcttcaagct gttcttcacc 420

gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

gagagcatga cctccgccac cttcaggac accaacgagg acgagatcta cgccttcttc 540

ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600

gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660

atcagatgcc ttaggatgga cgacaagagc atcaggcccc cctgcggga gaacgacgtg 720

ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc 780

cctggcgccc acctgacat cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840

agggtctata tccccaaaa gccagcaag tacggcatca agatcctgat gatgtcgac 900

agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960

ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgcca cggcagctgc 1020
 agaaacatca cctgcgacaa ctggttcacc gccatcccc tggccaagaa cctgtctcag 1080
 gagccctaca agctgacat cgtgggcacc gtggccagca acgccagaga gatccccgag 1140
 gtcctgaaga acagcaggtc cgccccctg ggcaccagca tgttctgctt cgacggcccc 1200

ctgaccctgg tgtcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc 1260
 gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac 1320
 cagaccaagg gcggcgtgga caccctgaac cagatgtgca gcgtgatgac ctgcagcaga 1380
 aagaccaaca gggcccccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
 agcttcatca tctacagcca caacgtgagc agcaaggcg agaagggtgca gagccggaaa 1500
 aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560
 gcccccccc tgaagagata cctgaggggac aacatcagca acatcctgcc caaagagggtg 1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaaggagc ctactgcacc 1680
 tactgtccca gcaagatcag aagaaaggcc agcgccgcct gcaagaagt taagaaggtc 1740
 atctgccggg agcacaacat cgacgtgtgc cagagggtgc tc 1782

<210> 87

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac

R275A/R277A/N347S/K375A/T560A/S573A/M589V/S594L

<400> 87

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60
 ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag 120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180
 tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac 240
 aggatcctga ccttgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc 300
 tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360
 cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gttcttcacc 420
 gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480
 gagagcatga cctccgccac cttcaggggac accaacgagg acgagatcta cgccttcttc 540

ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
atcagatgcc ttaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg 720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc 780
cctggcgccc acctgacct cgacgagcag ctgctgggct tcgccggcgc ctgccccttc 840
agggtctata tccccaacaa gcccagcaag tacggcatca agatcctgat gatgtgcgac 900
agcggcacca agtacctgat caacggcatg ccctacctgg gcagggggcac ccagaccaac 960

ggcgtgcccc tggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc 1020
agaaacatca cctgcgacaa ctgggtcacc agcatcccc tggccaagaa cctgctgcag 1080
gagccctaca agctgacct cgtgggcacc gtgagaagca acaagagaga gatccccgag 1140
gtcctgaaga acagcaggtc gcggccctg ggcaccagca tgttctgctt cgacggcccc 1200
ctgacctgg tgcctacaa gcccagccc gccaagatgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320
cagaccaagg gcggcgtgga cacctgaac cagatgtgca gcgtgatgac ctgcagcaga 1380

aagaccaaca gggcccccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
agttcatca tctacagcca caacgtgagc agcaaggcg agaaggtgca gagccggaaa 1500
aagttcatgc ggaacctgta catgggctg acctccagct tcatgaggaa gaggtggag 1560
gcccccccc tgaagagata cctgagggac aacatcagca acatcctgcc caaagaggtg 1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcgccgcct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacgtgtgc cagagctgtc tc 1782

<210> 88

<211> 1133

<212> DNA

<213> Artificial Sequence

<220>

<223> engineered PiggyBac R245A/R275A/R277A/R372A/W465A

<400> 88

tcgacttcct gatcagatgc ctgaggatgg acgacaagag catcaggccc accctgcggg 60
agaacgacgt gttcaccccc gtggccaaga tctgggacct gttcatccac cagtgcattc 120
agaactacac ccctggcgcc cacctgacca tcgacgagca gctgctgggc ttgccggcg 180
cctgcccctt cagggtctat atccccaaca agcccagcaa gtacggcatc aagatcctga 240

tgatgtgcga cagcggcacc aagtacatga tcaacggcat gccctacctg ggcaggggca 300

cccagaccaa cggcgtgccc ctgggcgagt actacgtgaa ggagctgtcc aagcccgtcc 360

acggcagctg cagaaacatc acctgcgaca actggttcac cagcatcccc ctggccaaga 420

acctgctgca ggagccctac aagctgacca tcgtgggcac cgtggccagc aacaagagag 480

agatccccga ggtcctgaag aacagcaggt ccaggcccgt gggcaccagc atgttctgct 540

tcgacggccc cctgacctg gtgtcctaca agcccaagcc cgccaagatg gtgtacctgc 600

tgtccagctg cgacgaggac gccagcatca acgagagcac cggcaagccc cagatggtga 660

tgtactacaa ccagaccaag ggcggcgtgg acaccctgga ccagatgtgc agcgtgatga 720

cctgcagcag aaagaccaac agggccccc tggccctgct gtacggcatg atcaacatcg 780

cctgcatcaa cagcttcac atctacagcc acaacgtgag cagcaagggc gagaagggtc 840

agagccggaa aaagtctatg cggaacctgt acatgggcct gacctccagc ttcattgagga 900

agaggctgga ggccccacc ctgaagagat acctgaggga caacatcagc aacatcctgc 960

ccaaagaggt gcccggcacc agcgacgaca gcaccgagga gcccgtagtg aagaagagga 1020

cctactgcac ctactgtccc agcaagatca gaagaaaggc cagcgccagc tgcaagaagt 1080

glaagaaggt catctgccgg gagcacaaca tcgacatgtg ccagagctgt ttc 1133

<210> 89

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> engineered Piggybac

R372A/K375A/R388A/D450N/W465A/S573A/M589V/S594L

<400> 89

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60

ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag 120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac 240

aggatcctga ccctgcccc aaggaccatc aggggcaaga acaagcactg ctggtccacc 300

tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac ccctgtgtgt gcttcaagct gttcttcacc 420

gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

gagagcatga cctccgccac cttcagggac accaacgagg acgagatcta cgccttcttc 540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
atcagatgcc tgaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg 720

ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc 780
cctggcgccc acctgacat cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840
agggctctata tcccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac 900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960
ggcgtgcccc tggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc 1020
agaaacatca cctgcgacaa ctgggtcacc agcatcccc tggccaagaa cctgctgcag 1080
gagccctaca agctgacat cgtgggcacc gtggccagca acgccagaga gatccccgag 1140

gtcctgaaga acagcaggtc cgccccgtg ggcaccagca tgttctgctt cgacggcccc 1200
ctgaccctgg tgcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac 1320
cagaccaagg gcggcgtgga caccctgaac cagatgtgca gcgtgatgac ctgcagcaga 1380
aagaccaaca gggcccccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
agcttcatca tctacagcca caacgtgagc agcaagggcg agaaggtgca gagccggaaa 1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560

gcccccccc tgaagagata cctgagggac aacatcagca acatcctgcc caaagaggtg 1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcggcgct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacgtgtgc cagaggtgtc tc 1782

<210> 90

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac M194A

<400> 90

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

1

5

10

15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190

His Ala Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335

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

 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460

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

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Phe

<210> 91

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R245A

<400> 91

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100	105	110	
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile			
115	120	125	
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile			
130	135	140	
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg			
145	150	155	160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile			
165	170	175	
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn			
180	185	190	
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr			
195	200	205	
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu			
210	215	220	
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val			
225	230	235	240
Phe Thr Pro Val Ala Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile			
245	250	255	
Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu			
260	265	270	
Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro			
275	280	285	
Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys			
290	295	300	
Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn			
305	310	315	320
Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val			
325	330	335	
His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile			
340	345	350	

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Phe

<210> 92

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R325A

<400> 92

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Ala Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495
 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525

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

Cys Phe

<210> 93

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R372A

<400> 93

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15
 Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300

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

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

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

 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 535 540
 Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 94

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R375A

<400> 94

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45
 Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60
 Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
 65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
 85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110
 Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
 130 135 140

 Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205

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

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

 His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
 340 345 350
 Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
 355 360 365

Gly Thr Val Arg Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn

370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro

385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu

405 410 415

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

420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525

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

530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr

545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys

565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser

580 585 590

Cys Phe

<210> 95

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R376A

<400> 95

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Ala Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450 455 460
 Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495

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

 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 96

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac E377A

<400> 96

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60
 Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
 65 70 75 80

 Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
 85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110
 Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
 115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
 130 135 140

 Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205

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

 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335

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

 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460

 Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495
 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525

 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 550 555 560
 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590

Cys Phe

<210> 97

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac E380A

<400> 97

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15
 Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45

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

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
 115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130	135	140	
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg			
145	150	155	160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile			
	165	170	175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn			
	180	185	190
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr			
	195	200	205
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu			
	210	215	220
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val			
225	230	235	240
Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile			
	245	250	255
Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu			
	260	265	270
Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro			
	275	280	285
Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys			
290	295	300	
Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn			
305	310	315	320
Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val			
	325	330	335
His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile			
	340	345	350
Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val			
	355	360	365
Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Ala Val Leu Lys Asn			
370	375	380	

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

<210> 98

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Engineered Piggybac D450N

<400> 98

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

210 215 220

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

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

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

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

 485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

 500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

 515 520 525

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

Cys Phe

<210> 99

<211> 594

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Engineered Piggybac S564P

<400> 99

Met	Gly	Ser	Ser	Leu	Asp	Asp	Glu	His	Ile	Leu	Ser	Ala	Leu	Leu	Gln
1				5					10					15	
Ser	Asp	Asp	Glu	Leu	Val	Gly	Glu	Asp	Ser	Asp	Ser	Glu	Val	Ser	Asp
			20					25					30		
His	Val	Ser	Glu	Asp	Asp	Val	Gln	Ser	Asp	Thr	Glu	Glu	Ala	Phe	Ile
			35					40					45		

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
 65 70 75 80
 Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
 85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110

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

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

 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300

 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

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

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

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

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

Tyr Cys Pro Pro Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590
Cys Phe

<210> 100

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> engineered Piggybac

<400> 100

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac	240
aggatcctga ccctgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagcccc ccaggcggag cagggtgtcc gcctgaaca tcgtgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480
gagagcatga cctccgccac cttcaggac accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccacg ccagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc ttaggatgga cgacaagagc atcaggcccc ccctgcggga gaacgacgtg	720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc	780
cctggcgccc acctgaccat cgacgagcag ctgctgggct tcaggggcag gtgcccttc	840
agggtctata tccccaacaa gccagcaag tacggcatca agatcctgat gatgtcgac	900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc	1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgaccat cgtgggcacc gtgagaagca acaagagaga gatccccgag	1140

gtcctgaaga acagcaggtc caggccccgtg ggcaccagca tgttctgctt cgacggcccc 1200
ctgacctgg tgtcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320
cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga 1380

aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
agcttcatca tctacagcca caactgagc agcaaggcg agaaggtgca gagccggaaa 1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggctggag 1560
gccccacc tgaagagata cctgagggac aacatcagca acatcctgcc caaagaggtg 1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 101

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> engineered Piggybac

<400> 101

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60
ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag 120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180
tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac 240
aggatcctga ccctgcccc gagggaccatc aggggcaaga acaagcactg ctggtccacc 300
tccaagcccc ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac cccctgctgt gttcaagct gttcttcacc 420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480
gagagcatga cctccgccac cttcagggac accaacgagg acgagatcta cgccttcttc 540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
atcagatgcc ttaggatgga cgacaagagc atcaggcccc cctgcggga gaacgacgtg 720
ttcacccccg tggccaagat ctgggacctg ttcattccacc agtgcattca gaactacacc 780

cctggcgccc acctgacctat cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840
 aggtgtctata tccccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac 900
 agcggcacca agtacctgat caacggcatg cctacctgg gcaggggcac ccagaccaac 960
 ggctgtcccc tgggcgagta ctacgtgaag gagctgtcca agcccggtcca cggcagctgc 1020
 agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag 1080
 gagccctaca agctgacctat cgtgggcacc gtgagaagca acaagagaga gatccccgag 1140
 gtctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc 1200

ctgacctgg tgtctacaa gcccaagccc gccaatgg tgtacctgct gtccagctgc 1260
 gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac 1320
 cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga 1380
 aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatgc ctgcatcaac 1440
 agcttcatca tctacagcca caacgtgagc agcaaggcg agaagggtgca gagccggaaa 1500
 aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560
 gccccacc tgaagagata cctgaggac aacatcagca acatcctgcc caaagaggtg 1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaaggagac ctactgcacc 1680
 tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagt taagaaggtc 1740
 atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 102

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> engineered Piggybac

<400> 102

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60
 ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag 120
 tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac 240
 aggatcctga cctgccccca gaggacctat aggggcaaga acaagcactg ctggtccacc 300
 tccaagccca ccaggcggag cagggtgtcc gccctgaaca tctgagaag ccagaggggc 360
 cccaccagga tgtgcaggaa catctacgac cccctgctgt gcttcaagct gttcttcacc 420
 gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

gagagcatga cctccgccac cttcaggga accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc tgaggatgga cgacaagagc atcaggccca cctgcggga gaacgacgtg	720
ttcaccccc tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc	780
cctggcgccc acctgacat cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggctctata tcccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac	900
agcggcacca agtatcatg caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tggccgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc	1020
agaaacatca cctgcgacaa ctgggtcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacat cgtgggcacc gtgagaagca acaagagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc	1200
ctgaccctgg tgcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatgggtgat gtactacaac	1320
cagaccaagg gcggcggtga caccctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac	1440
agcttcatca tctacagcca caacgtgagc agcaaggcg agaagggtgca gagccggaaa	1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag	1560
gcccccccc tgaagagata cctgaggga aacatcagca acatcctgcc caaagaggtg	1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc	1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc	1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc	1782
<210> 103	
<211> 1782	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> engineered Piggybac	
<400> 103	
atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgaggtg agcgaccacg tgagcgagga cgactgcag	120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac	240
aggatcctga ccttgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480
gagagcatga cctccgccac cttcaggac accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc tggagatgga cgacaagagc atcaggccca cctgcggga gaacgacgtg	720
ttcaccccc tgagaaagat ctgggacctg ttcattccacc agtgcacca gaactacacc	780
cctggcgccc acctgacct cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggtctata tcccaacaa gccagcaag tacggcatca agatcctgat gatgtcgac	900
agcggcacca agtatcatgat caacggcatg cctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc	1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacct cgtgggcacc gtggccagca acaagagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggcccggt ggccaccagca tgttctgctt cgacggcccc	1200
ctgaccctgg tgtcctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac	1320
cagaccaagg gcggcgtgga cacctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac	1440
agcttcatca tctacagcca caacgtgagc agcaagggcg agaagggtgca gagccgaaa	1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag	1560
gccccaccc tgaagagata cctgaggac aacatcagca acatcctgcc caaagaggtg	1620
cccggcacca gcgacgacag caccgaggag ccctgatga agaaggagac ctactgcacc	1680
tactgtccca gcaagatcag aagaaaggcc agcgcagct gcaagaagtg taagaaggtc	1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc	1782
<210> 104	
<211> 1782	
<212> DNA	

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R375A

<400> 104

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgcagag	60
ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac	240
aggatcctga ccctgcccc aaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagcccc ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480
gagagcatga cctccgccac cttcaggac accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgactgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc ttaggatgga cgacaagagc atcaggcccc ccctgcggga gaacgacgtg	720
ttcacccccg tgagaaagat ctgggacctg ttcattcacc agtgcattca gaactacacc	780
cctggcgccc acctgacct cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggtctata tccccaaaa gcccagcaag tacggcatca agatcctgat gatgtgcgac	900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc	1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacct cgtgggcacc gtgagaagca acgccagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc	1200
ctgaccctgg tgcctacaa gcccagccc gccaatggtg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac	1320
cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac	1440
agcttcatca tctacagcca caacgtgagc agcaaggcg agaaggtgca gagccggaac	1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag	1560
gccccacccc tgaagagata cctgaggggac aacatcagca acatcctgcc caaagaggtg	1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
 tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740
 atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782
 <210> 105
 <211> 1782
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Engineered Piggybac R376A
 <400> 105
 atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60

 ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag 120
 tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180
 tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac 240
 aggatcctga ccctgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc 300
 tccaagcccc ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360
 cccaccagga tgtgcaggaa catctacgac ccctgtctgt gcttcaagct gtctttcacc 420
 gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

 gagagcatga cctccgccac cttcagggac accaacgagg acgagatcta cgcctttctc 540
 ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
 gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
 atcagatgcc ttaggatgga cgacaagagc atcaggcccc ccctgcggga gaacgacgtg 720
 ttcacccccg tgagaaagat ctgggacctg ttcattcacc agtgcattca gaactacacc 780
 cctggcgccc acctgacatc cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840
 aggtctata tccccaacaa gcccagcaag tacggcatca agatcctgat gatgtcgac 900

 agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960
 ggctgtcccc tgggcgagta ctacgtgaag gagctgtcca agcccgctca cggcagctgc 1020
 agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag 1080
 gagccctaca agctgacatc cgtgggcacc gtgagaagca acaaggccga gatccccgag 1140
 gtctgaaga acagcaggtc caggccctg ggccaccagca tgttctgctt cgacggcccc 1200
 ctgaccctgg tgtcctacaa gcccaagccc gccaaagtgg tgtacctgct gtccagctgc 1260

gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320

cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga 1380

aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440

agcttcatca tctacagcca caacgtgagc agcaagggcg agaaggtgca gagccggaaa 1500

aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtctggag 1560

gccccacccc tgaagagata cctgagggac aacatcagca acatcctgcc caaagaggtg 1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680

tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740

atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 106

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac E377A

<400> 106

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60

ctggtcggcg aggacagcga cagcaggtg agcgaccacg tgagcgagga cgacgtgcag 120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac 240

aggatcctga cctgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc 300

tccaagcccc ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac ccctgctgt gcttcaagct gttcttcacc 420

gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

gagagcatga cctccgccac cttcaggga accaacgagg acgagatcta cgccttcttc 540

ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600

gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660

atcagatgcc tgaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg 720

ttcaccccc tgagaaagat ctgggacctg ttcattcacc agtgcattca gaactacacc 780

cctggcgccc acctgacct cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840

agggtctata tcccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac 900

agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960

ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc 1020

agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag 1080

gagccctaca agctgacat cgtgggcacc gtgagaagca acaagagagc catccccgag 1140

gtcctgaaga acagcaggtc caggcccggtg ggaccagca tgttctgctt cgacggcccc 1200

ctgaccctgg tgtcctacaa gcccagccc gccaagatgg tgtacctgct gtccagctgc 1260

gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320

cagaccaagg gcggcggtga caccctggac cagatgtgca gcgtgatgac ctgcagcaga 1380

aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440

agcttcatca tctacagcca caacgtgagc agcaaggcg agaagggtgca gagccgaaa 1500

aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560

gccccaccc tgaagagata cctgaggac aacatcagca acatcctgcc caaagaggtg 1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680

tactgtccca gcaagatcag aagaaaggcc agcggcagct gcaagaagtg taagaaggtc 1740

atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 107

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac E380A

<400> 107

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag 60

ctggtcggcg aggacagcga cagcgaggtg agcgaccacg tgagcgagga cgacgtgcag 120

tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc 180

tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac 240

aggatcctga ccttgcccca gaggaccatc aggggcaaga acaagcactg ctggtccacc 300

tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac ccctgctgt gcttcaagct gttcttcacc 420

gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480

gagagcatga cctccgccac cttcaggac accaacgagg acgagatcta cgccttcttc 540

ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc ttaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg	720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc	780
cctggcgccc acctgacctg cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggtctata tccccaacaa gcccagcaag tacggcatca agatcctgat gatgtcgac	900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc	1020
agaaacatca cctgcgacaa ctgggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacctg cgtgggcacc gtgagaagca acaagagaga gatccccgcc	1140
gtcctgaaga acagcaggtc caggccctg ggcaccagca tgttctgctt cgacggcccc	1200
ctgacctggt tgcctacaa gcccagccc gccaagatgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac	1320
cagaccaagg gggcgctgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac	1440
agttcatca tctacagcca caacgtgagc agcaaggcg agaaggtgca gagccggaaa	1500
aagttcatgc ggaacctgta catgggctg acctccagct tcatgaggaa gaggtggag	1560
gcccccccc tgaagagata cctgagggac aacatcagca acatcctgcc caaagaggtg	1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc	1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc	1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc	1782

<210> 108

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac D450N

<400> 108

atgggcagca gcctggacga cgagcacatc ctgagcgccc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgaggtg agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccc gcagctccct ggccagcaac	240

aggatcctga ccctgcccc gagggaccatc aggggcaaga acaagcactg ctggtccacc 300
tccaagcccc ccaggcggag cagggtgtcc gccctgaaca tcgtgagaag ccagaggggc 360

cccaccagga tgtgcaggaa catctacgac cccctgctgt gcttcaagct gttcttcacc 420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg 480
gagagcatga cctccgccac cttcaggggac accaacgagg acgagatcta cgccttcttc 540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc 600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg 660
atcagatgcc tgaggatgga cgacaagagc atcaggcccc ccctgcggga gaacgacgtg 720
ttcacccccg tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc 780

cctggcgccc acctgacat cgacgagcag ctgctgggct tcaggggcag gtgccccttc 840
agggtctata tccccaacaa gcccagcaag tacggcatca agatcctgat gatgtcgac 900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac 960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgtcca cggcagctgc 1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag 1080
gagccctaca agctgacat cgtgggcacc gtgagaagca acaagagaga gatccccgag 1140
gtcctgaaga acagcaggtc caggcccgtg ggcaccagca tgttctgctt cgacggcccc 1200

ctgaccttgg tgcctacaa gcccagccc gccaatgg tgtacctgct gtccagctgc 1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac 1320
cagaccaagg gcggcgtgga cacctgaac cagatgtgca gcgtgatgac ctgcagcaga 1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac 1440
agcttcatca tctacagcca caacgtgagc agcaagggcg agaaggtgca gagccggaaa 1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag 1560
gccccaccg tgaagagata cctgaggggac aacatcagca acatcctgcc caaagaggtg 1620

cccggcacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc 1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc 1740
atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 109

<211> 1782

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac S564P

<400> 109

atgggcagca gcctggacga cgagcacatc ctgagcgcgc tgctgcagag cgacgacgag	60
ctggtcggcg aggacagcga cagcgagggtg agcgaccacg tgagcgagga cgacgtgcag	120
tccgacaccg aggaggcctt catcgacgag gtgcacgagg tgcagcctac cagcagcggc	180
tccgagatcc tggacgagca gaacgtgatc gagcagcccg gcagctccct ggccagcaac	240
aggatcctga ccctgcccc aaggaccatc aggggcaaga acaagcactg ctggtccacc	300
tccaagccca ccaggcggag cagggtgtcc gccctgaaca tcttgagaag ccagaggggc	360
cccaccagga tgtgcaggaa catctacgac cccctgctgt gcttcaagct gttcttcacc	420
gacgagatca tcagcgagat cgtgaagtgg accaacgccg agatcagcct gaagaggcgg	480
gagagcatga cctccgccac cttcaggggac accaacgagg acgagatcta cgccttcttc	540
ggcatcctgg tgatgaccgc cgtgaggaag gacaaccaca tgagcaccga cgacctgttc	600
gacagatccc tgagcatggt gtacgtgagc gtgatgagca gggacagatt cgacttcctg	660
atcagatgcc tgaggatgga cgacaagagc atcaggccca ccctgcggga gaacgacgtg	720
ttcaccccc tgagaaagat ctgggacctg ttcattccacc agtgcattca gaactacacc	780
cctggcgccc acctgacat cgacgagcag ctgctgggct tcaggggcag gtgcccttc	840
agggtctata tcccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac	900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgcca cggcagctgc	1020
agaaacatca cctgcgacaa ctgggtcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacat cgtgggcacc gtgagaagca acaagagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc	1200
ctgacctgg tctctacaa gccaagccc gccaagatgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac	1320
cagaccaagg gcggcgtgga cacctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatcgc ctgcatcaac	1440
agcttcatca tctacagcca caacgtgagc agcaaggcg agaaggtgca gagccggaaa	1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggctggag	1560
gccccaccc tgaagagata cctgaggggac aacatcagca acatcctgcc caaagaggtg	1620
cccgccacca gcgacgacag caccgaggag cccgtgatga agaagaggac ctactgcacc	1680
tactgtcccc ccaagatcag aagaaaggcc agcgccagct gcaagaagtg taagaaggtc	1740

atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 110

<211> 5931

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac

S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/592G/S594L

<400> 110

atggacaaga agtactccat tgggctcgat atcggcacia acagcgtcgg ctgggccgtc 60

attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc 120

cacagcataa agaagaacct cattggcgcc ctctgttcg actccgggga aacggccgaa 180

gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc 240

tacctgcagg agatctttag taatgagatg gctaaggtgg atgactcttt ctccatagg 300

ctggaggagt cttttttggt ggaggaggat aaaaagcacg agcgccaccc aatctttggc 360

aatatcgtgg acgaggtggc gtacatgaa aagtaccaa ccatatatca tctgaggaag 420

aagcttgtag acagtactga taagctgac ttgcggttga tctatctcgc gctggcgcat 480

atgatcaaat ttcggggaca ctctctcgc gagggggacc tgaaccaga caacagcgat 540

gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg 600

atcaacgcat ccggagttag cgccaaagca atcctgagcg ctaggctgtc caaatcccgg 660

cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat 720

cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa 780

gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgacia tctgctggcc 840

cagatcggcg accagtacgc agaccttttt ttggcggcaa agaacctgtc agacgccatt 900

ctgctgagtg atattctcgc agtgaacacg gagatcacca aagctccgct gagcgctagt 960

atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc cttgtcaga 1020

cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggttacgcc 1080

ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gcccattctg 1140

gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttcgcg 1200

aaacagcgca ctttcgacia tggaagcatc ccccaccaga ttcacctggg cgaactgcac 1260

gctatcctca ggcggcaaga ggatttctac ccctttttga aagataacag ggaaaagatt 1320

gagaaaaatcc tcacatttcg gataccctac tatgtaggcc ccctcgcccg gggaaattcc	1380
agattcgcgt ggatgactcg caaatcagaa gagaccatca ctccctggaa cttcgaggaa	1440
gtcgtggata agggggcctc tgcccagtc ttcacgaaa ggatgactaa ctttgataaa	1500
aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacgagta cttcacagtt	1560
tataacgagc tcaccaaggt caaatacgtc acagaaggga tgagaaagcc agcattcctg	1620
tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagttacc	1680
gtgaaacagc tcaaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc	1740
agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc	1800
attaaagaca aggacttcct ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
ctcaccctta cgttgtttga agataggag atgattgaag aacgcttgaa aacttacgct	1920
catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg	1980
cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg	2040
gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac	2100
tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt	2160
cacgagcaca tcgctaactt tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220
gttaaggtcg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt	2280
atcgagatgg cccgagagaa ccaactacc cagaagggaac agaagaacag tagggaaagg	2340
atgaagagga ttgaagaggg tataaaagaa ctggggtccc aaatccttaa ggaacaccca	2400
gttgaaaaca ccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaaaata gaggggaagag tgataacgtc ccctcagaag aagtgtgcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggttcat caaaaggcag	2760
cttgttgaga cacgccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac	2820
accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct	2880
aagctggtct cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat	2940
taccaccatg cgcatgatgc ctacctgaat gcagtggtag gcaactgact tatcaaaaaa	3000
tatccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa	3060
atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc	3120
aataattatga attttttcaa gaccgagatt acactggcca atggagagat tcggaagcga	3180

ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgac	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct	3420
tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc	3480
aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac	3540
tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcatcattaa gcttcccaag	3600
tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg	3660
cagaaaggta acgagctggc actgccctct aaatacgtta atttcttgta tctggccagc	3720
cactatgaaa agctcaaagg gtctcccgaa gataatgagc agaagcagct gttcgtggaa	3780
caacacaaa actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagtg	3840
atcctcgccg acgctaacct cgataagggt ctttctgctt acaataagca cagggataag	3900
cccatcaggg agcaggcaga aaacattatc cacttgttta ctctgaccaa ctgggcgcg	3960
cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtagac ctctacaaag	4020
gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc	4080
gacctctctc agctcgggtg agattccggt agcgaaacac cggggacttc agaatcggcc	4140
accccgaggt ctggcagcag cctggacgac gagcacatcc tgagcgcct gctgcagagc	4200
gacgacgagc tggtcggcga ggacagcgac agcgaggtga gcgaccacgt gagcgaggac	4260
gacgtgcagt ccgacaccga ggaggccttc atcgacgagg tgacagaggt gcagcctacc	4320
agcagcggct ccgagatcct ggacgagcag aacgtgatcg agcagcccgg cagctccctg	4380
gccagcaaca ggatcctgac cctgccccag aggaccatca ggggcaagaa caagcactgc	4440
tgggtccact ccaagccac caggcggagc aggtgtccg ccctgaacat cgtgagaagc	4500
cagaggggcc ccaccaggat gtgcaggaac atctacgacc ccctgctgtg cttcaagctg	4560
ttcttcaccg acgagatcat cagcgagatc gtgaagtgga ccaacgccga gatcagcctg	4620
aagaggcggg agagcatgac ctccgccacc ttcagggaca ccaacgagga cgagatctac	4680
gccttcttcg gcatcctggt gatgaccgcc gtgaggaagg acaaccacat gagcaccgac	4740
gacctgttcg acagatccct gagcatggtg tacgtgagcg tgatgagcag ggacagattc	4800
gacttcctga tcagatgcct gaggatggac gacaagagca tcaggccac cctgcgggag	4860
aacgacgtgt tcacccccgt gaaaaagatc tgggacctgt tcatccacca gtgcatccag	4920

aactacaccc ctggcgccca cctgaccatc gacgagcagc tgctgggctt caggggcagg 4980

tgcccttca gggctctatat cccaacaag ccagcaagt acggcatcaa gatcctgatg 5040

atgtgcgaca gcggcaccaa gtacatgatc aacggcatgc cctacctggg caggggcacc 5100

cagaccaacg gcgtgccctt gggcgagtac tacgtgaagg agctgtccaa gcccgccac 5160

ggcagctgca gaaacatcac ctgcgacaac tggttcaccg ccatccccct ggccaagaac 5220

ctgctgcagg agccctacaa gctgaccatc gtgggcaccg tggccagcaa cgccagagag 5280

atccccgagg tctgaagaa cagcaggtcc gccccgtgg gcaccagcat gttctgttc 5340

gacggcccc tgacctggt gtctacaag cccaagcccg ccaagatggt gtacctgtg 5400

tccagctgcg acgaggacgc cagcatcaac gagagcaccg gcaagccca gatggtgatg 5460

tactacaacc agaccaaggg cggcgtggac accctgaacc agatgtgcag cgtgatgacc 5520

tgccagcaga agaccaacag ggccccatg gcctgtgt acggcatgat caacatcgcc 5580

tgcatcaaca gcttcatcat ctacagccac aacgtgagca gcaagggcga gaaggtgcag 5640

agccgaaaa agttcatgcg gaacctgtac atgggcctga cctccagctt catgaggaag 5700

aggctggagg cccccacct gaagagatac ctgagggaca acatcagcaa catcctgccc 5760

aaagaggtgc ccggcaccag cgacgacagc accgaggagc ccgtgatgaa gaagaggacc 5820

tactgcacct actgtccag caagatcaga agaaaggcca gcgccgctg caagaagtgt 5880

aagaaggtca tctgccggga gcacaacatc gacgtgtgcc agggctgttt g 5931

<210> 111

<211> 5931

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R245A/R275A/R277A/R372A/W465A/M589V

<400> 111

atggacaaga agtactccat tgggctcgat atcggcacaa acagcgtcgg ctgggccgtc 60

attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc 120

cacagcataa agaagaacct cattggcgcc ctctgttcg actccgggga aacggccgaa 180

gccacgcggc taaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc 240

tacctgcagg agatctttag taatgagatg gctaaggtgg atgactcttt cttccatagg 300

ctggaggagt ctttttggg ggaggaggat aaaaagcacg agcgccaccc aatctttggc 360

aatatcgtgg acgaggtggc gtacatgaa aagtaccaa ccatatatca tctgaggaag 420

aagctttag acagtactga taaggctgac ttgcggttga tctatctcgc gctggcgcat	480
atgatcaaat ttctgggaca cttctcatc gagggggacc tgaaccaga caacagcgat	540
gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg	600
atcaacgcat ccggagtga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacgcct gtttgtaat	720
cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa	780
gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgaca tctgctggcc	840
cagatcggcg accagtacgc agacctttt ttggcggcaa agaacctgtc agacgccatt	900
ctgctgagt atattctcgc agtgaacacg gagatcacca aagctccgt gagcgtagt	960
atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc cttgtcaga	1020
cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggctacgcc	1080
ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gcccatttg	1140
gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttcgc	1200
aaacagcgca ctttcgaca tggaagcatc ccccaccaga ttcacctggg cgaactgcac	1260
gctatcctca ggcggaaga ggatttctac ccctttttga aagataacag ggaaaagatt	1320
gagaaaatcc tcacatttcg gataccctac tatgtaggcc ccctcgccg gggaaattcc	1380
agattcgct ggatgactcg caaatcagaa gagaccatca ctccctggaa cttcgaggaa	1440
gtcgtggata agggggcctc tgcccagtc ttcacgaaa gcatgactaa ctttgataaa	1500
aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacgagta cttcacagtt	1560
tataacgagc tcaccaaggt caaatacgtc acagaaggga tgagaaagcc agcattcctg	1620
tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagtacc	1680
gtgaaacagc tcaaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc	1740
agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc	1800
attaaagaca aggacttctt ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
ctcacctta cgttgtttga agataggag atgattgaag aacgcttgaa aacttacgct	1920
catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg	1980
cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg	2040
gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac	2100
tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt	2160
cacgagcaca tcgctaactt tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220

gttaaggctg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt	2280
atcgagatgg cccgagagaa ccaaactacc cagaaggac agaagaacag tagggaaagg	2340
atgaagagga ttgaagaggg tataaaagaa ctgggggtccc aaatccttaa ggaacaccca	2400
gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaaaata gaggggaagag tgataacgtc ccctcagaag aagtgtgcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacgccaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggcttcat caaaaggcag	2760
cttgttgaga cagccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac	2820
accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct	2880
aagctggtct cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat	2940
taccaccatg cgcatgatgc ctacctgaat gcagtggtag gcactgcact tatcaaaaaa	3000
tatcccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa	3060
atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc	3120
aatattatga attttttcaa gaccgagatt aacttgcca atggagagat tcggaagcga	3180
ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgac	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct	3420
tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc	3480
aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac	3540
tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcatcattaa gtttccaag	3600
tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg	3660
cagaaaggta acgagctggc actgccctct aaatacgtta atttcttgta tctggccagc	3720
cactatgaaa agctcaaagg gtctcccga gataatgagc agaagcagct gttcgtggaa	3780
caacacaaac actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagt	3840
atcctcgccg acgctaacct cgataaggig ctttctgctt acaataagca cagggataag	3900
cccatcaggg agcaggcaga aaacattatc cacttgttta ctctgaccaa cttgggcgcg	3960

cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtagac ctctacaaag 4020
 gaggctctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc 4080
 gacctctctc agctcgggtg agattccggt agcgaaacac cggggacttc agaatcggcc 4140
 acccgggagt ctggcagcag cctggacgac gaggacatcc tgagcgcctt gctgcagagc 4200
 gacgacgagc tggtcggcga ggacagcgac agcgagggtga gcgaccacgt gagcgaggac 4260
 gacgtgcagt ccgacaccga ggaggccttc atcgacgagg tgcacgaggt gcagcctacc 4320
 agcagcggct ccgagatcct ggacgagcag aacgtgatcg agcagcccgg cagctccctg 4380

gccagcaaca ggatcctgac cctgccccag aggaccatca ggggcaagaa caagcactgc 4440
 tgggtccacct ccaagcccac caggcgggagc aggggtgtccg ccctgaacat cgtgagaagc 4500
 cagagggggc ccaccaggat gtgcaggaac atctacgacc ccctgctgtg cttcaagctg 4560
 ttcttcaccg acgagatcat cagcgagatc gtgaagtgga ccaacgccga gatcagcctg 4620
 aagaggcggg agagcatgac ctccgccacc ttcagggaaca ccaacgagga cgagatctac 4680
 gccttctctg gcatcctggt gatgaccgcc gtgaggaagg acaaccacat gagcaccgac 4740
 gacctgttcg acagatccct gagcatggtg tacgtgagcg tgatgagcag ggacagattc 4800

gacttctga tcagatgcct gaggatggac gacaagagca tcaggccac cctgcgggag 4860
 aacgacgtgt tcacccccgt gaaaaagatc tgggacctgt tcatccacca gtgcatccag 4920
 aactacacc ctggcgccca cctgaccatc gacgagcagc tgctgggctt caggggcagg 4980
 tgcccttca ggggtctatat cccaacaag ccagcaagt acggcatcaa gatcctgatg 5040
 atgtgcgaca gcggcaccaa gtacatgatc aacggcatgc cctacctggg caggggcacc 5100
 cagaccaacg gcgtgccctt gggcgagtac tacgtgaagg agctgtcaa gcccgtccac 5160
 ggcagctgca gaaacatcac ctgcgacaac tggttcaccg ccatccccct ggccaagaac 5220

ctgctgcagg agccctacaa gctgaccatc gtgggcaccg tggccagcaa cgccagagag 5280
 atccccgagg tcctgaagaa cagcaggtcc gccccgtgg gcaccagcat gttctgcttc 5340
 gacggcccc tgacctggt gtccctacaag cccaagcccg ccaagatggt gtacctgctg 5400
 tccagctgcg acgaggacgc cagcatcaac gagagcaccc gcaagccca gatggtgatg 5460
 tactacaacc agaccaaggg cggcgtggac accctgaacc agatgtgcag cgtgatgacc 5520
 tgcagcagaa agaccaacag ggccccatg gccctgctgt acggcatgat caacatgcc 5580
 tgcataaca gcitcatcat ctacagccac aacgtgagca gcaagggcga gaaggtgcag 5640

agccggaaaa agttcatgcg gaacctgtac atgggcctga cctccagctt catgaggaag 5700
 aggctggagg cccccacct gaagagatac ctgagggaaca acatcagcaa catcctgccc 5760
 aaagaggtgc ccggcaccag cgacgacagc accgaggagc ccgtgatgaa gaagaggacc 5820

tactgcacct actgtccccc caagatcaga agaaaggcca gcgccgctg caagaagtgt 5880

aagaagtca tctgccggga gcacaacatc gacatgtgcc agaggtgtct c 5931

<210> 112

<211> 5931

<212> DNA

<213> Artificial Sequence

<220>

<223> Engineered Piggybac R275A/325A/R372A/T560A

<400>

> 112

atggacaaga agtactccat tgggctcgat atcggcacaa acagcgtcgg ctgggccgtc 60

attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc 120

cacagcataa agaagaacct cattggcgcc ctctgttcg actccgggga aacggccgaa 180

gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc 240

tacctgcagg agatctttag taatgagatg gctaaggtgg atgactcttt ctccatagg 300

ctggaggagt ccttttttgt ggaggaggat aaaaagcacg agcgccaccc aatctttggc 360

aatatcgtgg acgaggtggc gtacatgaa aagtaccaa ccatatatca tctgaggaag 420

aagctttag acagtactga taaggctgac ttgcggttga tctatctgc gctggcgcat 480

atgatcaaat ttcggggaca ctctctcacc gagggggacc tgaaccaga caacagcgat 540

gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg 600

atcaacgcat ccggagttag cgccaaagca atcctgagcg ctaggctgtc caaatcccgg 660

cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat 720

cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttga cctggccgaa 780

gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc 840

cagatcggcg accagtacgc agaccttttt ttggcggcaa agaacctgtc agacgccatt 900

ctgctgagtg atattctcgc agtgaacacg gagatcacca aagctccgt gagcgctagt 960

atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc cttgttcaga 1020

cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggttacgcc 1080

ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gcccatttg 1140

gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttcgc 1200

aaacagcgca ctttcgacaa tggaagcacc ccccaccaga ttcacctggg cgaactgcac 1260

gctatcctca ggcggcaaga ggatttctac ccctttttga aagataacag ggaaaagatt 1320

gagaaaaatcc tcacatttcg gataacctac tatgtaggcc ccctcgcccg gggaaattcc 1380
 agattcgctt ggatgactcg caaatcagaa gagaccatca ctccctggaa cttcgaggaa 1440
 gtctgtgata agggggcctc tgcccagtcc ttcacgaaa ggatgactaa ctttgataaa 1500
 aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacgagta cttcacagtt 1560
 tataacgagc tcaccaaggt caaatcgtc acagaaggga tgagaaagcc agcattcctg 1620
 tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagttacc 1680

 glgaaacagc tcaaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc 1740
 agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc 1800
 attaaagaca aggacttctt ggacaatgag gagaacgagg acattcttga ggacattgtc 1860
 ctccacctta cgttgtttga agataggagg atgattgaag aacgcttgaa aacttacgct 1920
 catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg 1980
 cggctgtcaa gaaaactgat caatgggacg cgagacaagc agagtggaaa gacaatcctg 2040
 gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac 2100

 tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt 2160
 cacgagcaca tcgctaactt tgcaggtagc ccagctatca aaaagggaat actgcagacc 2220
 gtttaaggtcg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt 2280
 atcgagatgg cccgagagaa ccaactacc cagaaggagc agaagaacag tagggaaagg 2340
 atgaagagga ttgaagaggg tataaaagaa ctggggtccc aaatccttaa ggaacaccca 2400
 gttgaaaaca ccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg 2460
 gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat 2520

 atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc 2580
 gataaaaata gaggggaagag tgataacgtc ccctcagaag aagttgtcaa gaaaatgaaa 2640
 aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg 2700
 actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggttcat caaaaggcag 2760
 ctgtttgaga cagccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac 2820
 accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct 2880
 aagctggtct cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat 2940

 taccaccatg cgcatgatgc ctacctgaat gcagtggtag gcactgcact tatcaaaaaa 3000
 tatccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa 3060
 atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc 3120
 aatattatga attttttcaa gaccgagatt acactggcca atggagagat tcggaagcga 3180

ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgac	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct	3420
tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc	3480
aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac	3540
tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcattcattaa gcttcccaag	3600
tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg	3660
cagaaaggta acgagctggc actgccctct aaatacgtta atttcttgta tctggccagc	3720
cactatgaaa agctcaaagg gtctcccga gataatgagc agaagcagct gttcgtggaa	3780
caacacaaac actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagtg	3840
atcctcgccg acgctaacct cgataagggt ctttctgctt acaataagca cagggataag	3900
cccatcaggg agcaggcaga aaacattatc cacttgctta ctctgaccaa ctgggcgcg	3960
cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtagac ctctacaaag	4020
gaggtcctgg acgccacct gattcatcag tcaattacgg ggctctatga aacaagaatc	4080
gacctctctc agctcgggtg agattccggt agcgaacac cggggacttc agaatcggcc	4140
accccgaggt ctggcagcag cctggacgac gagcacatcc tgagcgcct gctgcagagc	4200
gacgacgagc tggtcggcga ggacagcgac agcgaggtga gcgaccacgt gagcgaggac	4260
gacgtgcagt ccgacaccga ggaggccttc atcgacgagg tgacgaggt gcagcctacc	4320
agcagcggct ccgagatcct ggacgagcag aacgtgatcg agcagcccgg cagctccctg	4380
gccagcaaca ggatcctgac cctgccccag aggaccatca ggggcaagaa caagcactgc	4440
tgggtccact ccaagccac caggcggagc aggggtgtccg ccctgaacat cgtgagaagc	4500
cagaggggcc ccaccaggat gtgcaggaac atctacgacc ccctgctgtg cttcaagctg	4560
ttcttcaccg acgagatcat cagcgagatc gtgaagtgga ccaacgccga gatcagcctg	4620
aagaggcggg agagcatgac ctcccacc ttcagggaca ccaacgagga cgagatctac	4680
gccttcttcg gcatcctggt gatgaccgcc gtgaggaagg acaaccacat gagcaccgac	4740
gacctgttcg acagatccct gagcatggtg tacgtgagcg tgatgagcag ggacagattc	4800
gacttcctga tcagatgcct gaggatggac gacaagagca tcaggccac cctgcgggag	4860
aacgacgtgt tcacccccgt gagaagatc tgggacctgt tcattccacca gtgcatccag	4920
aactacacc ctggcgccca cctgaccatc gacgagcagc tgctgggctt cgccggcagg	4980

tgcccttca gggcttatat cccaacaag ccagcaagt acggcatcaa gatcctgatg	5040
atgtgcgaca gggcaccaa gtacatgatc aacggcatgc cctacctggg caggggcacc	5100
cagaccaacg gcgtgccct ggccgagtag tacgtgaagg agctgtccaa gcccgtccac	5160
ggcagctgca gaaacatcac ctgcgacaac tggttacca gcatccccct ggccaagaac	5220
ctgctgcagg agccctacaa gctgaccatc gtgggcaccg tggccagcaa caagagagag	5280
atccccgagg tectgaagaa cagcaggtcc agggccgtgg gcaccagcat gttctgcttc	5340
gacggcccc tgacctggt gtcctacaag cccaagcccg ccaagatggt gtacctgctg	5400
tccagctgcg acgaggacgc cagcatcaac gagagcaccg gcaagcccca gatggtgatg	5460
tactacaacc agaccaaggg cggcgtggac accctggacc agatgtgcag cgtgatgacc	5520
tgcagcagaa agaccaacag gtggcccatg gcctgtgtgt acggcatgat caacatcgcc	5580
tgcataca gcttcatcat ctacagccac aacgtgagca gcaagggcga gaaggtgcag	5640
agccggaaaa agttcatcg gaacctgtac atgggcctga cctccagctt catgaggaag	5700
aggctggagg cccccacct gaagagatac ctgagggaca acatcagcaa catcctgccc	5760
aaagaggtgc ccggcaccag cgacgacagc accgaggagc ccgtgatgaa gaagaggacc	5820
tacgcctact actgtcccag caagatcaga agaaaggcca gcgccagctg caagaagtgt	5880
aagaaggtca tctgccggga gcacaacatc gacatgtgcc agagctgttt c	5931
<210> 113	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> gRNA	
<400> 113	
tatgtacact tctgaccac	20
<210> 114	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> gRNA SaCas9	
<400> 114	
glatcacaat tccagtgggt	20

<210> 115

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> gRNA SaCas9

<400> 115

gtatcacaat tccagtgggt 20

<210> 116

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> gRNA SaCas9

<400> 116

ggacaggatc ggcataaccg 20

<210> 117

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> gRNA SaCas9

<400> 117

gtgctcgggg ccactaggga 20

<210> 118

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Cpf1 gRNA

<400> 118

acttataatt cactgtatca 20

<210> 119

<211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Cpf1 gRNA
 <400> 119
 agcttgatat ccatggaatt 20
 <210> 120
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Cpf1 gRNA
 <400> 120
 tgctcggggc cactagggac 20
 <210> 121
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Cpf1 gRNA
 <400> 121
 cttttgtaaa actttatggt 20
 <210> 122
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Cpf1 gRNA
 <400> 122
 caaaagtaaa tagcccggt 20
 <210> 123
 <211> 22
 <212> DNA

<213> Artificial Sequence

<220>

<223> CjCas9

<400> 123

gccgatacctg tccctagtgg cc 22

<210> 124

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> CjCas9 gRNA

<400> 124

acaattccag tgggtcagaa gt 22

<210> 125

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> CjCas9 gRNA

<400> 125

acacttctga cccactggaa t 21

<210> 126

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> CjCas9 gRNA

<400> 126

gaattccatg gatataaagg tt 22

<210> 127

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> CjCas9 gRNA

<400> 127

aattccagtg ggtcagaagt gt

22

<210> 128

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> CasX gRNA

<400> 128

tcaagcgcgt gtatgtacac

20

<210> 129

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> CasX gRNA

<400> 129

ggatcggcat aaccggtgaa

20

<210> 130

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> CasX gRNA

<400> 130

tagacatgag gtctatggac

20

<210> 131

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> CasX gRNA

<400> 131
taagcttgat atccatggaa 20
<210> 132
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> CasX gRNA
<400> 132
tataattcac tgtatcaca 20
<210> 133
<211> 5
<212> PRT
<213> Artificial sequence
<220>
<221> SITE
<222> 1..5
<223> wherein the sequence can be repeated 1 to 50 times
<220>
<223> Linker
<400> 133
Gly Gly Gly Gly Ser
1 5
<210> 134
<211> 5
<212> PRT
<213> Artificial sequence
<220>
<221> SITE
<222> 1..5
<223> wherein the sequence can be repeated 1 to 50 times
<220>
<223> linker
<400> 134

Glu Ala Ala Ala Lys

1 5

<210> 135

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac R275A/N347S/K375A/D450N/S592G

<400> 135

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

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

 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Ala Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320

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

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

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Gly
580 585 590

Cys Phe

<210> 136

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac N347S/D450N/T560A/S573A/F594L

<400> 136

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

				20						25					30				
His	Val	Ser	Glu	Asp	Asp	Val	Gln	Ser	Asp	Thr	Glu	Glu	Ala	Phe	Ile				
		35					40					45							
Asp	Glu	Val	His	Glu	Val	Gln	Pro	Thr	Ser	Ser	Gly	Ser	Glu	Ile	Leu				
	50					55					60								
Asp	Glu	Gln	Asn	Val	Ile	Glu	Gln	Pro	Gly	Ser	Ser	Leu	Ala	Ser	Asn				
65					70					75					80				
Arg	Ile	Leu	Thr	Leu	Pro	Gln	Arg	Thr	Ile	Arg	Gly	Lys	Asn	Lys	His				
				85					90					95					
Cys	Trp	Ser	Thr	Ser	Lys	Pro	Thr	Arg	Arg	Ser	Arg	Val	Ser	Ala	Leu				
		100					105					110							
Asn	Ile	Val	Arg	Ser	Gln	Arg	Gly	Pro	Thr	Arg	Met	Cys	Arg	Asn	Ile				
	115					120					125								
Tyr	Asp	Pro	Leu	Leu	Cys	Phe	Lys	Leu	Phe	Phe	Thr	Asp	Glu	Ile	Ile				
	130					135				140									
Ser	Glu	Ile	Val	Lys	Trp	Thr	Asn	Ala	Glu	Ile	Ser	Leu	Lys	Arg	Arg				
145				150					155					160					
Glu	Ser	Met	Thr	Ser	Ala	Thr	Phe	Arg	Asp	Thr	Asn	Glu	Asp	Glu	Ile				
			165					170					175						
Tyr	Ala	Phe	Phe	Gly	Ile	Leu	Val	Met	Thr	Ala	Val	Arg	Lys	Asp	Asn				
		180				185					190								
His	Met	Ser	Thr	Asp	Asp	Leu	Phe	Asp	Arg	Ser	Leu	Ser	Met	Val	Tyr				
	195					200					205								
Val	Ser	Val	Met	Ser	Arg	Asp	Arg	Phe	Asp	Phe	Leu	Ile	Arg	Cys	Leu				
	210					215				220									
Arg	Met	Asp	Asp	Lys	Ser	Ile	Arg	Pro	Thr	Leu	Arg	Glu	Asn	Asp	Val				
225				230					235					240					
Phe	Thr	Pro	Val	Arg	Lys	Ile	Trp	Asp	Leu	Phe	Ile	His	Gln	Cys	Ile				
			245					250					255						
Gln	Asn	Tyr	Thr	Pro	Gly	Ala	His	Leu	Thr	Ile	Asp	Glu	Gln	Leu	Leu				
	260					265						270							

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Ser Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Ala
545 550 555 560
Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590
Cys Leu

<210> 137

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac N347A/D450N

<400> 137

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Ser Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val

355 360 365
 Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
 370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445

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

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

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 138

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac R275A/N347S/R372A/D450N/T560A/F594L

<400> 138

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220

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

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

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

 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Ala
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Leu

<210> 139

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R202K/R275A/N347S/R372A/D450N/T560A/F594L

<400> 139

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190
His Met Ser Thr Asp Asp Leu Phe Asp Lys Ser Leu Ser Met Val Tyr
195 200 205
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240
Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255
Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270
Gly Phe Ala Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

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

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

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

 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Ala
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys

565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Leu

<210> 140

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L

<400> 140

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
 115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
 130 135 140

 Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205

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

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

 His Gly Ser Cys Arg Asn Ile Thr Cys Asp Ser Trp Phe Thr Ser Ile
 340 345 350
 Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val

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

 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
 450 455 460

 Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 470 475 480
 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 490 495
 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
 500 505 510
 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 520 525

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

 Cys Leu

<210> 141

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified Hyperactive Piggybac triple mutant

D450N/R372A/K375A/R376A

<400> 141

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val

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

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

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

 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

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

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

545 550 555 560
 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590
 Cys Phe

<210> 142

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G

<400> 142

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15
 Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45
 Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60

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

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

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

 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270

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

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

 Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445

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

 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

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

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Gly
 580 585 590
 Cys Phe

<210> 143

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L

<400> 143

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
 1 5 10 15
 Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

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

85 90 95
 Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

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

 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

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

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

 340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Ala
545 550 555 560

Tyr Cys Pro Pro Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Gly
580 585 590

Cys Leu

<210> 144

<211> 594

<212

> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P

<400> 144

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Met Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

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

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Ala Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300

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

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

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

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

Tyr Cys Pro Pro Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590
Cys Phe

<210> 145

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

G325A/N347S/K375A/D450N/S573A/M589V/S592G

<400> 145

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

1

5

10

15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
 20 25 30
 His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
 35 40 45
 Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
 50 55 60
 Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

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

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

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

260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Ala Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

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

 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430
 Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
 435 440 445
 Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

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

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser

530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr

545 550 555 560
Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys

565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Gly

580 585 590
Cys Phe

<210> 146

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac S230N/R277A/N347S/K375A/D450N

<400> 146

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

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

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

145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

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

180 185 190

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

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

210 215 220

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

225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270

Gly Phe Arg Gly Ala Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro

275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Ser Trp Phe Thr Ser Ile

340 345 350
 Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val

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

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

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

 545 550 555 560
 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys
 565 570 575
 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
 580 585 590

Cys Phe

<210> 147

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac T43I/R372A/K375A/A411T/D450N

<400> 147

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

1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp

20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Ile Glu Glu Ala Phe Ile

35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95
Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg

145 150 155 160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn

180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240
 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335
 His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
 340 345 350
 Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
 355 360 365
 Gly Thr Val Ala Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn
 370 375 380
 Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
 385 390 395 400
 Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Thr Lys Met Val Tyr Leu
 405 410 415
 Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
 420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

435

440

445

Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450

455

460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

465

470

475

480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

485

490

495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

500

505

510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515

520

525

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

530

535

540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr

545

550

555

560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys

565

570

575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser

580

585

590

Cys Phe

<210> 148

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G

<400> 148

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

1

5

10

15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260	265	270
Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro		
275	280	285
Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys		
290	295	300
Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn		
305	310	315
Gly Val Pro Leu Ala Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val		
325	330	335
His Gly Ser Cys Arg Asn Ile Thr Cys Asp Ser Trp Phe Thr Ala Ile		
340	345	350
Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val		
355	360	365
Gly Thr Val Arg Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn		
370	375	380
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro		
385	390	395
Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu		
405	410	415
Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys		
420	425	430
Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr		
435	440	445
Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg		
450	455	460
Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn		
465	470	475
Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val		
485	490	495
Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser		
500	505	510

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

<210> 149

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G

<400> 149

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Pro Thr Arg Arg Ser Arg Val Ser Ala Leu
 100 105 110
 Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
 115 120 125
 Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
 130 135 140
 Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
 145 150 155 160
 Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
 165 170 175
 His Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
 180 185 190
 His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
 195 200 205
 Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
 210 215 220
 Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
 225 230 235 240
 Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
 245 250 255
 Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
 260 265 270
 Gly Phe Ala Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
 275 280 285
 Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
 290 295 300
 Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
 305 310 315 320
 Gly Val Pro Leu Ala Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
 325 330 335
 His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile

340 345 350
Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val

355 360 365
Gly Thr Val Arg Ser Asn Ala Arg Glu Ile Pro Glu Val Leu Lys Asn

370 375 380
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro

385 390 395 400
Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu

405 410 415
Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys

420 425 430
Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

435 440 445
Leu Asn Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450 455 460
Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

465 470 475 480
Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

485 490 495
Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser

530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Ala

545 550 555 560
Tyr Cys Pro Gly Lys Ile Arg Arg Lys Ala Ser Ala Ser Cys Lys Lys

565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Gly

580 585 590

Cys Phe

<210> 150

<211> 390

<212> DNA

<213> Artificial Sequence

<220>

<223> MCP N55K

<400> 150

atggcttcaa accttactca gttcgtgctc gtggacaatg gtgggacagg ggatgtgaca 60
gtggctcctt ctaatttcgc taatggggtg gcagagtgga tcagctccaa ctcacggagc 120
caggcctaca aggtgacatg cagcgtcagg cagtctagtg cccagaagag aaagtatacc 180

atcaaggtgg aggtcccca agtggtctacc cagacagtgg gcggagtcca actgcctgtc 240
gccgcttgga ggtcctacct gaacatggag ctactatcc caattttcgc taccaattct 300
gactgtgaac tcacgtgaa ggcaatgcag gggctcctca aagacggtaa tcctatccct 360
tccgccatcg ccgctaactc aggtatctac 390

<210> 151

<211> 130

<212> PRT

<213> Artificial Sequence

<220>

<223> MCP N55K

<400> 151

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

1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu

20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45

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

50 55 60

Val Pro Lys Val Ala Thr Gln Thr Val Gly Gly Val Glu Leu Pro Val

[illegible]

130

<210> 152

<211> 162

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> gRNA-MS2(tetraloop)-AAVS1-3 spacer

<400> 152

ggggccacta gggacaggat gttttagagc taggccaaca tgaggatcac ccatgtctgc 60

agggcctagc aagttaaaat aaggctagtc cgttatcaac ttggccaaca tgaggatcac 120

ccatgtctgc agggccaagt ggcaccgagt cgggtcgtttt tt 162

<210> 153

<211> 162

<212> RNA

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> gRNA-MS2(tetraloop)-AAVS1-3 spacer

<400> 153

ggggccacua gggacaggau guuuuagagc uaggccaaca ugaggaucac ccaugucugc 60

agggccuagc aaguuaaaau aaggcuaguc cguuaucaac uggccaaca ugaggaucac 120

ccaugucugc agggccaagu ggcaccgagu cggugcuuuu uu 162