



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0125806
(43) 공개일자 2023년08월29일

- (51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) A61K 38/00 (2006.01)
A61P 21/04 (2006.01) C07K 14/705 (2006.01)
C12N 15/90 (2006.01) C12N 9/12 (2006.01)
- (52) CPC특허분류
C12N 9/22 (2013.01)
A61K 38/00 (2013.01)
- (21) 출원번호 10-2023-7023945
(22) 출원일자(국제) 2021년12월16일
심사청구일자 없음
(85) 번역문제출일자 2023년07월13일
(86) 국제출원번호 PCT/EP2021/086333
(87) 국제공개번호 WO 2022/129430
국제공개일자 2022년06월23일
- (30) 우선권주장
20214691.6 2020년12월16일
유럽특허청(EPO)(EP)
21209721.6 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
팔라레스 마스미타 마리아
스페인, 08002 바르셀로나, 플라자 데 라 메르세
10-12, 씨/오우니베르시타트 폼페우 파브라
- (74) 대리인
특허법인한얼

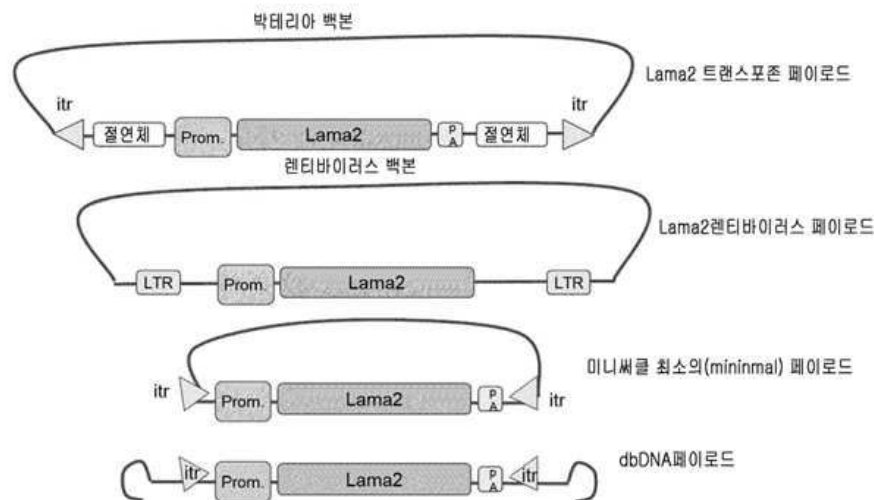
전체 청구항 수 : 총 21 항

(54) 발명의 명칭 선천성 근이영양증의 치료를 위한 치료용 LAMA2 페이로드

(57) 요약

본 발명은, a) 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질산 서열; 또는 상기 제1 단백질을 코딩하는 핵산 작제물; b) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물; 및 c) 라미닌-α2 단백질, 또는 이의 기능적 변이체 또는 단편을 코딩하는 트랜스젠을 포함하는 핵산 작제물을 포함하는 조성물에 관한 것이다. 본 발명은 또한 특히 선천성 근이영양증의 치료를 위해, LAMA2 트랜스젠을 세포의 게놈 내의 특이적 부위로 통합하기 위한 이러한 조성물의 치료적 용도에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

A61P 21/04 (2018.01)

C07K 14/705 (2013.01)

C12N 15/907 (2013.01)

C12N 9/1241 (2013.01)

C07K 2319/81 (2013.01)

명세서

청구범위

청구항 1

- a) 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(nucleic acid construct);
- b) 트랜스포사제(transposase)를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물; 및
- c) 바람직하게는 서열번호: 74의, 라미닌- α 2 단백질, 또는 이의 기능적 변이체(functional variant) 또는 단편을 코딩하는 트랜스젠(transgene)을 포함하는 핵산 작제물
- 을 포함하는 조성물.

청구항 2

제1항에 있어서, a) 및 b)의 제1 및 제2 단백질이 임의로 링커(linker)를 통해 함께 융합(fusing)되는 조성물.

청구항 3

제1항 또는 제2항에 있어서, c)의 상기 핵산 작제물이 서열번호: 76의 CMV 프로모터, 서열번호: 77의 CAG 프로모터, 서열번호: 78의 EF1-알파 프로모터, 서열번호: 79의 SV-40 프로모터 및 서열번호: 80의 EalBAAAT 프로모터로 이루어진 군으로부터 선택된 프로모터를 포함하는 조성물.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, c)의 상기 핵산 작제물이 서열번호: 81의 스플라이스 수용체(splice acceptor)를 포함하는 조성물.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 c)의 핵산 작제물이 바람직하게는 서열번호: 83-85로 이루어진 군으로부터 선택되는 폴리(A) 신호 서열(poly(A) signal sequence) 및/또는 바람직하게는 서열번호: 86-87로 이루어진 군으로부터 선택되는 절연체 요소(insulator element)를 포함하는 조성물.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 c)의 핵산 작제물이 역 말단 반복부(inverted terminal repeat; ITR), 바람직하게는 서열번호: 88 및 89의 5'-ITR 및 3'-ITR에 의해 플랭킹(flanking)되는 조성물.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 상기 c)의 핵산 작제물이 플라스미드 벡터, 미니써클(minicircle) 벡터, 도기 본(doggy bone) DNA 공여체 벡터(donor vector), 렌티바이러스 벡터 및 렌트로바이러스 벡터로 이루어진 군으로부터 선택되는 벡터에 포함되는 조성물.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 부위-특이적 DNA 결합 단백질이 Cas 단백질을 포함하는 RNA-가이드된(guided) 뉴클레아제이고, 여기서 상기 조성물이 세포 계놈의 특이적 부위에 상기 LAMA2 트랜스젠을 통합하기 위한 표적 핵산 서열에 대한 상보적 서열을 포함하는 가이드 RNA를 추가로 포함하며, 바람직하게는 상기 Cas 단백질이 에스. 피오게네스(*S. pyogenes*) Cas 9 단백질인 조성물.

청구항 9

제8항에 있어서, 가이드 RNA가 서열번호: 90 내지 97 중 어느 하나를 포함하는 조성물.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 상기 트랜스포사제가 변형된 과활성(hyperactive) Piggybac 트랜스포사제 또는 슬리피 뷰티(Sleepy Beauty) 트랜스포사제, 바람직하게는 변형되지 않은 과활성 Piggybac과 비교하여 절제 활성(excision activity)을 증가시키기 위한 하나 이상의 아미노산 돌연변이, 및 변형되지 않은 과활성 Piggybac과 비교하여 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac 트랜스포사제인 조성물.

청구항 11

제10항에 있어서, 상기 과활성 PiggyBac 트랜스포사제가 V34, T43, Y177, M194, R202, S230, R245, R275, R277, G325, S351, N347, R372, K375, R376, E377, E380, A411, D450, T560, S564, S573, M589, S592, 및 F594로 이루어진 군으로부터 선택된 아미노산의 적어도 하나의 돌연변이를 포함하며 바람직하게는 아미노산 Y177, R202, S230, R245, R275, R277, G325, N347, S351, E377, D450, R372 및 K375, E377, T560, S564, S573, M589, S592, F594의 돌연변이를 포함하는 변형된 과활성 PiggyBac 트랜스포사제이며, 상기 위치 번호가 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하는, 조성물.

청구항 12

제10항 또는 제11항에 있어서, 상기 과활성 PiggyBac 트랜스포사제가 M194, R245, R275, R277, G325, R372, K375, R376, E377, E380, D450 및 S573으로 이루어진 군으로부터 선택된 아미노산의 적어도 하나의 돌연변이를 포함하며, 바람직하게는 아미노산 D450, R372 및 R375의 돌연변이를 포함하는 변형된 과활성 PiggyBac 트랜스포사제이며, 상기 위치 번호가 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응하는, 조성물.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, 상기 트랜스포사제가 링커, 바람직하게는 GGS, XTEN 또는 FOKI를 포함하는 펩티드 링커, 보다 바람직하게는 서열번호: 53의 XTEN에 의해 상기 부위-특이적 DNA 결합 단백질에 N-말단으로 융합되는, 조성물.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 상기 조성물이 나노입자 내에 패키징(packaging)되는 조성물.

청구항 15

제1항 내지 제14항 중 어느 한 항에 따른 조성물을 세포 내로 도입하는 것을 포함하는, LAMA2 트랜스جين을 세포의 게놈 내의 표적 핵산 서열로 통합하기 위한 시험관내 방법.

청구항 16

제15항에 따른 방법에 의해 수득할 수 있는 조작된 세포(engineered cell)로서, 여기서 조작된 세포가 이의 게놈 내에 통합된 핵산을 포함하며, 상기 핵산은 인테그라제- 및/또는 트랜스포사제-매개 유전자 삽입을 위한 작동 서열에 의해 플랭킹되는 라미닌- α 2 단백질을 코딩하는 트랜스جين을 포함하는, 조작된 세포.

청구항 17

제16항에 있어서, 트랜스جين을 플랭킹하는 작동 서열(operational sequence)이 서열번호: 88 및 89를 포함하거나 이로 이루어진 조작된 세포.

청구항 18

제16항 또는 제17항에 있어서, ITR 간 크기(inter-ITR size)가 적어도 300 bp인 조작된 세포.

청구항 19

제16항 내지 제18항 중 어느 한 항에 있어서, 트랜스진이 전장(full-length) 라미닌- α 2 단백질을 코딩하는 조작된 세포.

청구항 20

제1항 내지 제14항 중 어느 한 항에 정의된 바와 같은 조성물, 또는 제16항 내지 제19항 중 어느 한 항에 따른 조작된 세포를, 임의로 하나 이상의 약제학적으로 허용되는 부형제와 조합하여 포함하는 약제학적 조성물.

청구항 21

이를 필요로 하는 대상체(subject)에서 요법(therapy)에 사용하기 위한, 특히 메로신-결핍 선천성 근이영양증 1A형(merosin-deficient congenital muscular dystrophy type 1A) (MDC1A)의 치료에 사용하기 위한, 제1항 내지 제14항 중 어느 한 항에 따른 조성물, 제14항 내지 제19항 중 어느 한 항에 따른 조작된 세포 또는 제20항의 약제학적 조성물.

발명의 설명

기술 분야

[0001] 본 발명은 *LAMA2* 트랜스진(transgene)을 세포의 게놈 내의 특이적 부위로 통합하기 위해 트랜스포사제(transposase) 및 *LAMA2* 트랜스진에 융합된 부위-특이적 DNA 결합 단백질을 포함하는 융합 단백질을 포함하는 조성물, 및 특히 선천성 근이영양증(congenital muscular dystrophy)의 치료를 위한 이의 치료적 용도에 관한 것이다.

배경 기술

[0002] 선천성 근이영양증 (CMD)은 골격/심장 근육뿐만 아니라 중추 신경계에 영향을 미치는 중증 조기-발병 근이영양증(severe early-onset muscular dystrophy)의 부류이다.

[0003] CMD 중에서, 메로신-결핍 선천성 이영양증 1A형(merosin-deficient congenital dystrophy type 1A) (MDC1A)으로도 공지된 라미닌- α 2 쇠-결핍 선천성 근이영양증(laminin- α 2 chain-deficient co-genital muscular dystrophy) (*LAMA2* MD)은 중증 저긴장증(severe hypotonia), 근육 쇠약(muscle weakness), 골격 기형(skeletal deformity), 보행 불가(non-ambulation), 호흡 부전(respiratory insufficiency)을 특징으로 한다. 유사한 증상을 나타내지만 더 야생형의 표현형 가변성을 가진, 나중에 발병하는 더 경증 형태의 질환이 있다.

[0004] 현재 MDC1A에 이용가능한 어떤 치유법(cure)도 없다. 클리닉에서의 현재 전략은 관리 (급식 보충(feeding supplementation), 호흡 부전에 대한 비침습적 인공호흡 지원(non-invasive ventilation support), 및 관절 구축(joint contracture), 척추 결함(spinal defect), 및 기타 문제에 대한 물리 요법(therapy))에 초점을 맞추고 있다 (Nguyen *et al.*, 2019. *Appl Clin Genet.* **12**:113-130).

[0005] MDC1A는 라미닌-211의 서브유닛(subunit) 중 하나인 라미닌- α 2 쇠를 코딩하는 *LAMA2* 유전자의 카피(copy) 둘다에서 기능 상실 열성 돌연변이(loss-of-function recessive mutation)에 의해 유발된다. 라미닌-211은 수축 동안 기저막과 근육 섬유를 안정화시키는 기능을 하는 세포의 기질 단백질이다. 라미닌- α 2 쇠 단백질의 길이는 3 122개의 아미노산 잔기이므로, 이의 코딩 DNA 서열 (CDS)의 길이는 ~9.3 kb이다. 따라서 다른 많은 근이영양증과 같이, MDC1A는 AAV 벡터 또는 렌티바이러스 벡터의 카고(cargo) 용량을 증가하는, 매우 큰 유전자의 돌연변이에 의해 유발된다. 더욱이, 전반적인 *LAMA2* 유전자에서 지금까지 단일 점 미스센스(single point missense), 스플라이스 부위(splice site) 또는 인프레임(in-frame) 돌연변이로부터, 다중-킬로베이스(multi-kilobase) 영역 결실(deletion)에 이르기까지 수십 가지의 상이한 돌연변이가 너무 많이 기재되어 있어 (Oliveira *et al.*, 2018. *Hum Mutat.* **39**(10):1314-1337), 치유적 유전자 요법 접근법이 도전과제이고 유전자 대체 전략은 지금까지 이용가능하지 않다.

[0006] 다른 교정 유전자-기반 전략이 MDC1A에 대해 시험되었다. 통상적으로 스플라이스 인핸서(enhancer) 부위를 억제하고 특정한 엑손이 스플라이싱에 참여하지 못하도록 하기 위해 합성 안티센스 올리고뉴클레오타이드의 사용을 지칭하는 엑손-스킵핑 접근법(Exon-skipping approach)이 사용되었다 (Aoki *et al.*, 2013. *Biomed Res Int.* **2013**:402369). 그러나, 이 엑손-스킵핑 접근법은 보편적이지 않으며 결실을 갖는 일부 환자에게만 사용할 수 있으며, 여기서 결실에 인접한 추가 엑손을 스킵함으로써 관독 프레임을 복원할 수 있다 (Shieh, 2018. *Neurotherapeutics.* **15**(4):840-848). 엑손-스킵핑은 말단절단된(truncated) 그러나 부분적으로 기능적인 단백

질의 번역을 복원하나 전체 길이 라미닌- $\alpha 2$ 단백질의 이점을 요약하지는 않는다.

- [0007] CRISPR/Cas9는 또한 *LAMA2* 돌연변이를 교정하는 데 사용되었으며, 특히 스플라이스-부위 돌연변이를 함유하는 *LAMA2* 인트론 2 영역을 절제하고 비-상동 말단-연결(non-homologous end-joining) (NHEJ)을 통해 기능적 공여체(donor) 스플라이스 부위를 생성하여, *LAMA2* 전사체의 엑손 2 포함 및 전장 라미닌- $\alpha 2$ 사슬 단백질 복원을 유도하였다 (Kemaladewi *et al.*, 2017. *Nat Med.* 23(8):984-989). 엑손-스킵핑으로서, 이 전략은 일부 환자에서만 사용할 수 있으며 보편적이지 않는다.
- [0008] 그러나, NHEJ는 큰 트랜스진을 게놈에 통합하는 데 적합하지 않으며, 한편 상동성 지향 복구(homology directed repair) (HDR) 경로에 의존하는 유전자 변형이 이론적으로 대부분의 병원성 점 돌연변이를 교정할 수 있지만, 이는 문헌 [Kemaladewi & Cohn (2019. *Emerg Top Life Sci.* 3(1):11-18)]에 의해 설명된 바와 같이 "극히 비효율적"인 것으로 밝혀졌다. 결론적으로, [Kemaladewi & Cohn]은 근이영양증을 유발하는 돌연변이를 교정하기 위한 대안적 전략의 개발을 촉구하였다.
- [0009] 마지막으로, CRISPR/Cas9는 *LAMA1* 유전자의 프로모터로 지향하는, VP64와 같은 전사 활성화 도메인에 융합된 촉매적으로 불활성인 Cas9 (*SadCas9*)를 사용함으로써, 유전자 발현을 유도하기 위해 또한 사용되어, 구조적으로 유사한 라미닌- $\alpha 1$ 쇠 단백질의 발현 증가를 허용한다 (Kemaladewi *et al.*, 2019. *Nature.* 572(7767):125-130). 그러나, 이 접근법은 트랜스진의 지속적인 발현이 필요하다.
- [0010] 따라서 보편적이어서 특정 돌연변이에 영향을 받는 특정 환자에게만 국한되지 않는, MDC1A를 치료하기 위한 영구적이고 효율적인 전략을 개발할 필요성이 남아 있다.

발명의 내용

- [0011] *LAMA2* 유전자의 유전자 교체는 유망한 치료 가치를 가지고 있으나; 카고(cargo) 크기 제약으로 인해 지금까지는 도전과제였다. 본 발명자들은 *LAMA2* 유전자의 건강한 전장 카피를 세포의 게놈 내의 특정 부위로 통합하기 위해 트랜스포사제 (예를 들어, 과활성(hyperactive) PiggyBac 트랜스포사제)에 융합된 부위-특이적 DNA-결합 단백질 (예를 들어, Cas9)을 포함하는 융합 단백질을 사용하였다. 이에 따라 본 발명자들은 상이한 유전자 전달 기술과 호환되는, 생체외 및 생체내 전달을 위한 대형 *LAMA2* 발현 카세트 (9.3 kb의 CDS)를 가진 치료 페이로드 (therapeutic payload)를 제공할 수 있었다.
- [0012] 따라서, 본 발명은 a) 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); b) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물; 및 c) 바람직하게는 서열번호: 74의, 라미닌- $\alpha 2$ 단백질, 또는 이의 기능적 변이체(functional variant) 또는 단편을 코딩하는 트랜스진을 포함하는 핵산 작제물을 포함하는 조성물을 포함한다.
- [0013] 한 실시양태에서, a) 및 b)의 제1 및 제2 단백질은 임의로 링커(linker)를 통해 함께 융합(fusing)된다.
- [0014] 한 실시양태에서, c)의 핵산 작제물은 서열번호: 76의 CMV 프로모터, 서열번호: 77의 CAG 프로모터, 서열번호: 78의 EF1-알파 프로모터, 서열번호: 79의 SV-40 프로모터 및 서열번호: 80의 EalbAAT 프로모터로 이루어진 군으로부터 선택된 프로모터를 포함한다. 한 실시양태에서, c)의 핵산 작제물은 서열번호: 81의 스플라이스 수용체(83-85)를 포함한다. 한 실시양태에서, c)의 핵산 작제물은 바람직하게는 서열번호: 83-85로 이루어진 군으로부터 선택되는 폴리(A) 신호 서열(poly(A) signal sequence) 및/또는 바람직하게는 서열번호: 86-87로 이루어진 군으로부터 선택되는 절연체 요소(insulator element)를 포함한다. 한 실시양태에서, c)의 핵산 작제물은 역말단 반복부(inverted terminal repeat) (ITR), 바람직하게는 서열번호: 88 및 89의 5'-ITR 및 3'-ITR에 의해 플랭킹(flanking)된다.
- [0015] 한 실시양태에서, c)의 핵산 작제물은 플라스미드 벡터, 미니써클 벡터(minicircle vector), 도기 본(doggy bone) DNA 공여체 벡터, 렌티바이러스 벡터 및 렌트로바이러스 벡터로 이루어진 군으로부터 선택된 벡터에 포함된다.
- [0016] 한 실시양태에서, 상기 부위-특이적 DNA 결합 단백질은 Cas 단백질을 포함하는 RNA-가이드된(guided) 뉴클레아제이고, 여기서 상기 조성물은 세포 게놈의 특이적 부위에 상기 *LAMA2* 트랜스진을 통합하기 위한 표적 핵산 서열에 대한 상보적 서열을 포함하는 가이드 RNA를 추가로 포함하며, 바람직하게는 상기 Cas 단백질은 에스. 피오게네스(*S. pyogenes*) Cas 9 단백질이다.

- [0017] 한 실시양태에서, 가이드 RNA는 서열번호: 90 내지 97 중 어느 하나를 포함한다.
- [0018] 한 실시양태에서, 상기 트랜스포사제는 변형된 과활성 Piggybac 트랜스포사제 또는 슬리피 뷰티(Sleepy Beauty) 트랜스포사제, 바람직하게는 변형되지 않은 과활성 Piggybac과 비교하여 절제 활성(excision activity)을 증가시키기 위한 하나 이상의 아미노산 돌연변이, 및 변형되지 않은 과활성 Piggybac과 비교하여 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 돌연변이를 포함하는 변형된 과활성 PiggyBac 트랜스포사제이다.
- [0019] 한 실시양태에서, 상기 과활성 PiggyBac 트랜스포사제는 V34, T43, Y177, M194, R202, S230, R245, R275, R277, G325, S351, N347, R372, K375, R376, E377, E380, A411, D450, T560, S564, S573, M589, S592, 및 F594로 이루어진 군으로부터 선택된 아미노산의 적어도 하나의 돌연변이를 포함하는, 바람직하게는 아미노산 Y177, R202, S230, R245, R275, R277, G325, N347, S351, E377, D450, R372 및 K375, E377, T560, S564, S573, M589, S592, F594의 돌연변이를 포함하는 변형된 과활성 PiggyBac 트랜스포사제이며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다. 한 실시양태에서, 상기 과활성 PiggyBac 트랜스포사제는 M194, R245, R275, R277, G325, R372, K375, R376, E377, E380, D450 및 S573으로 이루어진 군으로부터 선택된 아미노산의 적어도 하나의 돌연변이를 포함하는, 바람직하게는 아미노산 D450, R372 및 R375의 돌연변이를 포함하는, 변형된 과활성 PiggyBac 트랜스포사제이며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.
- [0020] 한 실시양태에서, 상기 트랜스포사제는 링커, 바람직하게는 GGS, XTEN 또는 FOKI를 포함하는 펩티드 링커, 보다 바람직하게는 서열번호: 53의 XTEN에 의해 상기 부위-특이적 DNA 결합 단백질에 N-말단으로 융합된다.
- [0021] 한 실시양태에서, 상기 조성물은 나노입자 내에 패키징(packaging)된다.
- [0022] 본 발명은 또한 세포 내로 본 발명의 조성물을 도입하는 것을 포함하는 세포의 계놈 내의 표적 핵산 서열 내로 LAMA2 트랜스진을 통합하기 위한 시험관내 방법에 관한 것이다.
- [0023] 본 발명은 또한 본 발명의 시험관내 방법에 의해 획득할 수 있는 조작된 세포(engineered cell)에 관한 것이며, 여기서 조작된 세포는 이의 계놈 내에 통합된 핵산을 포함하며, 상기 핵산은 인테그라제 및/또는 트랜스포사제 매개 유전자 삽입에 대한 작동 서열(operational sequence)에 의해 플랭킹된 라미닌- α 2 단백질을 코딩하는 트랜스진을 포함한다.
- [0024] 한 실시양태에서, 트랜스진을 플랭킹하는 작동 서열은 서열번호: 88 및 89를 포함하거나 이로 이루어진다.
- [0025] 한 실시양태에서, ITR-간 크기(inter-ITR size)는 적어도 300 bp이다.
- [0026] 한 실시양태에서, 트랜스진은 전장(full-length) 라미닌- α 2 단백질을 코딩한다.
- [0027] 본 발명은 또한 본 발명의 조성물, 또는 본 발명의 조작된 세포를, 임의로 하나 이상의 약제학적으로 허용되는 부형제(excipient)와 조합하여 포함하는 약제학적 조성물에 관한 것이다.
- [0028] 본 발명은 또한 요법에 사용하기 위한, 특히 이를 필요로 하는 대상체(subject)에서 메로신-결핍 선천성 근이영양증 1A형 (MDC1A)의 치료에 사용하기 위한 본 발명의 조성물, 본 발명의 조작된 세포 또는 본 발명의 약제학적 조성물에 관한 것이다.
- [0029] **정의**
- [0030] 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.
- [0031] 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, 포스포로디티오에이트, N3'P5'-포스포르아미데이트 및 올리고리보뉴클레오타이드 포스포로티오에이트 및 이의 2'-O-알릴 유사체 및 2'-O-메틸리

보뉴클레오티드 메틸포스포네이트를 포함하며, 이들은 본 개시내용의 뉴클레오티드에 사용될 수 있다.

[0032] 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 영역에 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.

[0033] 용어 "**핵산 작제물**"은 재조합 DNA 기술을 사용하여 생성된 사람이 제조한 핵산 분자를 지칭한다. 핵산 작제물은 핵산 서열의 세그먼트를 함유하도록 변형된, 단일- 또는 이중-가닥의 핵산 분자이며, 이는 자연계에 달리 존재하지 않을 방식으로 조합되고 병치된다. 핵산 작제물은 대개 "**벡터**", 즉 외인성으로 생성된 DNA를 숙주 세포로 전달하는 데 사용되는 핵산 분자이다.

[0034] 용어 "**벡터**" 또는 "**발현 벡터**"는 자율적으로 복제하는 벡터, 즉, 이의 복제가 염색체 복제와 무관한 여분의 염색체 실체로서 존재하는 벡터, 예를 들어, 플라스미드, 여분의 염색체 요소, 미니-염색체, 또는 자기-복제를 보장하기 위한 임의의 수단을 함유할 수 있는 인공 염색체; 또는 통합 벡터, 즉 숙주 세포에 도입될 때 이의 게놈에 통합되고 통합된 염색체(들)와 함께 복제되는 벡터를 지칭할 수 있다.

[0035] 용어 "**서열 동일성**" 또는 "**동일성**"은 2개의 폴리펩티드 또는 폴리뉴클레오티드 서열의 정렬로부터의 위치에서 매치 (동일한 아미노산 잔기 또는 뉴클레오티드)의 수(%)를 지칭한다. 서열 동일성은 서열 갭을 최소화하면서 중첩을 최대화하도록 정렬될 때 서열을 비교함으로써 결정된다. 특히, 서열 동일성은 두 서열의 길이에 따라, 임의의 다수의 수학적 글로벌 또는 로컬 정렬 알고리즘을 사용하여 결정될 수 있다. 유사한 길이의 서열은 바람직하게는 전체 길이에 걸쳐 서열을 최적으로 정렬하는 글로벌 정렬 알고리즘 (예를 들어, 니들만 앤드 분쉬(Needleman and Wunsch) 알고리즘 [Needleman & Wunsch, 1970. *J Mol Biol.* **48(3)**:443-53])을 사용하여 정렬되며, 한편 실질적으로 상이한 길이의 서열은 바람직하게는 로컬 정렬 알고리즘 (예를 들어, 스미스 앤드 워터만(Smith and Waterman) 알고리즘 [Smith & Waterman, 1981. *J Mol Biol.* **147(1)**:195-197] 또는 엘트설(Altschul) 알고리즘 [Altschul et al., 1997. *Nucleic Acids Res.* **25(17)**:3389-3402; Altschul et al., 2005. *FEBS J.* **272(20)**:5101-9])을 사용하여 정렬된다. 예컨대 <http://blast.ncbi.nlm.nih.gov/> 또는 <http://www.ebi.ac.uk/Tools/emboss/>에서 이용가능한 공개적으로 이용가능한 컴퓨터 소프트웨어를 사용하여, 관련 기술분야의 기술 범위 내에서 다양한 방식으로 달성될 수 있다. 관련 기술분야의 통상의 기술자는 비교되는 서열의 전장에 걸쳐 최대 정렬을 달성하는 데 필요한 임의의 알고리즘을 포함하여, 정렬을 측정하기 위한 적절한 파라미터를 결정할 수 있다. 본원의 목적상, % 뉴클레오티드 또는 아미노산 서열 동일성 값은 니들만-분쉬 알고리즘을 사용하여 두 서열의 최적 글로벌 정렬을 생성하는 쌍별(pair wise) 서열 정렬 프로그램 EMBOS 니들(Needle)을 사용하여 생성된 값을 지칭하며, 여기서 모든 검색 파라미터는 디폴트 값, 즉, 스코어링 매트릭스(Scoring matrix) = BLOSUM62, 갭 오픈(Gap open) = 10, 갭 확장(Gap extend) = 0.5, 엔드 갭 페널티(End gap penalty) = 거짓, 엔드 갭 오픈(End gap open) = 10 및 엔드 갭 확장(End gap extend) = 0.5로 설정된다.

[0036] 용어 "**융합**"은 2개 이상의 서브유닛 분자가 연결된 분자를 지칭한다. 일부 실시예에서, 둘 사이의 연결은 공유 결합이며; 대안적으로, 둘 사이의 연결은 비공유결합일 수 있고, 예를 들어 분자간 상호작용에 의존할 수 있다. 서브유닛 분자는 동일한 화학적 유형의 분자이거나, 상이한 화학적 유형의 분자일 수 있다.

[0037] 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 단일 쇠 폴리펩티드이고, 펩티드에 의해 직접 공유결합되거나 펩티드 링커를 통해 임의로 공유결합된 적어도 2개의 단백질 도메인을 포함한다.

[0038] 용어 "**링커**"는 2개의 인접한 분자 또는 모이어티를 연결하는 화학적 기 또는 또는 분자를 지칭한다.

[0039] 용어 "**결합 단백질**"은 또 다른 분자에 비공유 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)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.

[0040] 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.

[0041] Cas9는 CRISPR 반복 서열에서 짧은 모티프 (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. 써모필러스(*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.

[0042] 일부 실시양태에서, Cas9 뉴클레아제는 불활성 (예를 들어, 불활성화된) DNA 절단 도메인을 갖는다. 뉴클레아제-불활성화된 Cas9 단백질은 "dCas9" 단백질 (뉴클레아제-"죽은" 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).

[0043] 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.

[0044] 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.

[0045] 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 결과로 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.

[0046] 용어 "특이성"은 선택된 서열에 대해 어느 정도의 서열 동일성을 공유하는 서열에 선택적으로 결합하는 능력을 지칭한다.

[0047] 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(on-targeted)"은 본원에

서 핵산의 삽입을 제2 핵산의 특이적 부위에 또는 게놈 또는 이의 일부의 특이적 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 삽입 수를 지칭한다.

[0048] 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 기재된다. 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.

[0049] 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.

[0050] 용어 "변형된"은 상응하는 변형되지 않은 단백질 또는 핵산 서열과 상이한 단백질 또는 핵산 서열을 지칭한다.

[0051] 용어 "링크"는 2개의 인접한 분자 또는 모이어티를 연결하는 화학적 기 또는 분자를 지칭한다.

[0052] 상세한 설명

[0053] 본 발명자들은 *LAMA2* 트랜스진을 세포의 게놈 내의 특이적 부위에 통합시키기 위해 트랜스포사제에 융합된 부위-특이적 DNA-결합 단백질을 포함하는 융합 단백질을 사용하였다.

[0054] 따라서 본 발명은

[0055] a) 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, 바람직하게는 cDNA 또는 mRNA;

[0056] b) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기 제2 단백질을 코딩하는 핵산 작제물, 바람직하게는 cDNA 또는 mRNA; 및

[0057] c) 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스진을 포함하거나 이로 이루어진 핵산 작제물

[0058] 을 포함하거나 이로 이루어진 조성물에 관한 것이다.

[0059] 한 실시양태에서, 조성물은

[0060] a) (i) 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질, 및 (ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질을 포함하거나 이로 이루어진 융합 단백질; 또는 상기 융합 단백질을 코딩하는 핵산 작제물, 바람직하게는 cDNA 또는 mRNA; 및

[0061] b) 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스진을 포함하거나 이로 이루어진 핵산 작제물

[0062] 을 포함하거나 이로 이루어진다.

[0063] 본 발명에 따르면, 제1 단백질은 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진다.

[0064] 조작된 징크 핑거 단백질(zinc finger protein) (ZFP), 전사 활성화제 유사 이펙터 뉴클레아제(transcription activator like effector nuclease) (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.

[0065] 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.

- [0066] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 RNA-가이드된 DNA 뉴클레아제 및 징크 핑거 단백질을 포함하거나 이로 이루어진 군으로부터 선택된다.
- [0067] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 RNA-가이드된 뉴클레아제이다.
- [0068] 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.
- [0069] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 이의 변이체 및 기능적 단편을 포함하여, Cas9 단백질이다.
- [0070] 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 코절 대체로 제한되나, 대규모 전장 유전자는 아니다. 상기에서 설명된 바와 같이, 본 발명은 이러한 결점을 해결한다.
- [0071] 한 실시양태에서, Cas9 단백질은 (i) 활성 DNA 절단 도메인 및 (ii) 가이드 RNA 결합 도메인을 포함한다.
- [0072] 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.
- [0073] 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(*Belliella 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: 63에 의해 코딩됨); 서열번호: 30을 가진 네이세리아 메닝기티디스(*Neisseria meningitidis*) (NCBI 참조번호: YP_002342100.1); 서열번호: 68을 가진 스태필로코커스 아우레우스 (SaCas9) (서열번호: 60에 의해 코딩됨); 및 서열번호: 31을 가진 스트렙토코커스 피오게네스 (SpCas9) (NCBI 참조번호: NC_017053.1)로부터의 Cas 9 단백질을 포함하거나 이로 이루어진 군으로부터 선택된다.
- [0074] 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.
- [0075] 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.
- [0076] 한 실시양태에서, Cas9 변이체는 서열번호: 31을 가진 야생형 Cas9 단백질 또는 서열번호: 19-30 또는 68을 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% 동

일할 수 있거나, 적어도 약 99.9%일 수 있다.

- [0077] 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.
- [0078] 이들 서브도메인 내의 돌연변이는 Cas9의 뉴클레아제 활성을 침묵시킬 수 있다. 예를 들어, D10A 및 H841A 치환은 서열번호: 31을 가진 에스. 피오케네스 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: 59에 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: 67에 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.
- [0079] 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: 57에 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: 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 아미노산 서열을 포함하거나 이로 이루어진다. 일부 실시양태에서, 일부 실시양태에서, SaCas9 니카제 (SanCas9)는 서열번호: 58에 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: 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 아미노산 서열을 포함한다.
- [0080] 한 실시양태에서, Cas9 변이체는 Cas9의 단편을 포함하여, 단편이 서열번호: 31을 가진 야생형 Cas9 단백질 또는 서열번호: 19-30 또는 68을 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.
- [0081] 한 실시양태에서, Cas9 변이체는 DNA 절단 도메인 또는 가이드 RNA 결합 도메인 중 하나만을 포함한다.
- [0082] 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.
- [0083] 한 실시양태에서, hCas9 단백질은 서열번호: 56에 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: 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 아미노산 서열을 포함한다.
- [0084] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 cpf1 단백질이다. 한 실시양태에서, cpf1 단백질은 서열번호: 61에 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: 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.

98%, 적어도 약 99%, 또는 약 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.

- [0085] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 CasX 단백질이다. 한 실시양태에서, 상기 CasX는 서열번호: 62에 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: 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 포함한다.
- [0086] 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.
- [0087] 한 실시양태에서, 부위-특이적 DNA 결합 단백질은 징크 핑거 단백질 (ZFP)이다.
- [0088] 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.
- [0089] 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.
- [0090] 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.
- [0091] ZFP는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12개, 또는 그 초과와 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.
- [0092] 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-핑거 모듈을 사용하여 최대 6개의 개별 징크 핑거를 가진 징크 핑거 어레이를 생성한다.
- [0093] 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.
- [0094] 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.
- [0095] 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.
- [0096] 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-특이적이도록 조작된 Gal4 DNA 결합 도메인을 갖는다.

- [0097] 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.
- [0098] 본 발명에 따르면, 제2 단백질은 트랜스포사제를 포함하거나 이로 이루어진다.
- [0099] 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 (Sleeping Beauty) (SB), 및 나방(moth) 트리코플루시아(*Trichoplusia*)로부터 단리된, PiggyBac (PB)을 제한 없이 포함한다. PiggyBac은 SB보다 전위 활성도가 높으며 오프터 없이 절제할 수 있다.
- [0100] 천연 DN트랜스포존은 전형적으로 트랜스포사제 결합 부위를 운반하는 역 말단 반복부 (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 (예를 들어, *LAMA2* 트랜스진)이 시험관 내에서 합성된 트랜스포사제 효소 또는 그 mRNA를 코딩하는 서열을 함유하는 공동-형질감염된 플라스미드에 의해 지향되는 숙주 계놈으로 도입될 수 있는 발현 카세트를 함유한다. 본 개시내용에 따르면, 트랜스포존-기반 시스템은 세포에서 *LAMA2* 트랜스진의 안정한 통합 및 지속적인 발현을 효율적으로 매개하는데 사용된다.
- [0101] 한 실시양태에서, 본 발명에 사용되는 트랜스포사제 또는 변형된 트랜스포사제는 계놈의 특이적 부위에 *LAMA2* 트랜스진을 삽입할 수 있는 임의의 트랜스포사제일 수 있다.
- [0102] 트랜스포사제의 비제한적인 예는 프로그 프린스(Frog Prince), 슬리핑 뷰티, 과활성 과활성 슬리핑 뷰티, PiggyBac, 및 과활성 PiggyBac을 포함한다.
- [0103] 한 실시양태에서, 트랜스포사제는 과활성 PiggyBac 트랜스포사제 (hyPB)이다.
- [0104] 야생형 과활성 PiggyBac 트랜스포사제는 서열번호: 9를 포함하거나 이로 이루어진 아미노산 서열을 갖는다. 이 단백질을 코딩하는 예시적인 핵산 서열은 서열번호: 71에 제시된 바와 같다.
- [0105] 한 실시양태에서, 트랜스포사제는 변형된 과활성 PiggyBac 트랜스포사제이다.
- [0106] 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.
- [0107] 일부 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 절제 활성을 증가시키기 위해, 야생형 과활성 PiggyBac 트랜스포사제와 비교하여 하나 이상의 아미노산 치환을 포함한다.
- [0108] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 위치 D198, D201, R202, M212, 및/또는 S213에서 아미노산 위치 번호 [194-200], [214-222], [434-442] 및/또는 [446-456]에 의해 정의된 영역 내에 절제 활성을 증가시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9를 갖는 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0109] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 위치 450, 560, 564, 573, 589, 592, 및/또는 594로부터 선택된 아미노산 위치에서 절제 활성을 증가시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은

서열번호: 9를 갖는 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.

- [0110] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 위치 M194 및 D450으로부터 선택된 아미노산 위치에서 절제 활성을 증가시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9의 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 한 실시양태에서, 하나 이상의 아미노산 치환은 M194V 및 D450N 중에서 선택되며, 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0111] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 DNA 결합 활성을 감소시키기 위해, 야생형 과활성 PiggyBac 트랜스포사제와 비교하여, 하나 이상의 아미노산 치환을 포함한다.
- [0112] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 야생형 과활성 PiggyBac 트랜스포사제와 비교하여 위치 254, 275, 277, 347, 372, 375, 및/또는 465로부터 선택된 아미노산 위치에서 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0113] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 위치 R275, N347, R372, K375, R376, E377, 및/또는 E380으로부터 선택된 아미노산 위치에서 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0114] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 위치 R372, K375, R376, E377 및/또는 E380으로부터 선택된 아미노산 위치에서 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9의 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 한 실시양태에서, 하나 이상의 아미노산 치환은 R372A, K375A, R376A, E377A, 및/또는 E380A 중에서 선택되고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0115] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 위치 N347, R372, 및 K375로부터 선택된 아미노산 위치에서 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 한 실시양태에서, 하나 이상의 아미노산 치환은 N347S, N347A, R372A 및/또는 K375A 중에서 선택되고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 한 실시양태에서, 하나 이상의 아미노산 치환은 N347S 또는 N347A 중에서 선택되고; 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0116] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는, 상기에 정의된 바와 같이, 절제 활성을 증가시키기 위한 하나 이상의 아미노산 치환; 및 상기에 정의된 바와 같은 DNA 결합 활성을 감소시키기 위한 하나 이상의 아미노산 치환을 포함한다.
- [0117] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 위치 D450에서 절제 활성을 증가시키기 위한 적어도 하나의 치환 및 위치 N347, R372 및 K375에서 DNA 결합 활성을 감소시키는 위한 적어도 2개의 치환을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 하며; 바람직하게는 상기 변형된 과활성 PiggyBac 트랜스포사제는 이중 돌연변이 N347S 및 D450N 또는 삼중 치환 D450N, R372A 및 K375A를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 한 실시양태에서, 과활성 PiggyBac의 변형된 트랜스포사제는 이중 돌연변이 N347S 및 D450N을 포함하며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다.
- [0118] 일부 실시양태에서, 과활성 PiggyBac의 변형된 트랜스포사제는 삼중 돌연변이 D450N, R372A 및 K375A를 포함하지 않으며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 PiggyBac의 아미노산 번호에 상응한다.
- [0119] 한 실시양태에서, 이전 실시양태에서 개시된 바와 같은 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 위치 번호 [158-169]에 의해 정의된 영역에 적어도 하나의 치환, 예를 들어 A166S; 및/또는 위치 Y527, R518, K525 및/또는 N463에서 적어도 하나의 치환을 추가로 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0120] 전형적으로, 상기 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 1의 변형되지 않은 과활성 PiggyBac 트랜스포사제와 적어도 85%, 적어도 90%, 적어도 95% 동일성을 갖는 아미노산 서열을 포함한다.
- [0121] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 34, 43, 117, 202, 230, 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 594로부터 선택된 위치에서 하기 아미노산 치환 중 하나 이

상을 추가로 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.

- [0122] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 하기 아미노산 치환 또는 치환의 조합 중 하나를 포함한다: V34M, T43I, Y177H, R202K, S230N, R245A, D268N, R275A, R277A, 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, F594L, D450N/R372A/K375A, R275A/R277A, K409A/K412A, R460A/K461A, R275A/R277A/N347S/K375A/T560A/S573A/M589V/S592G 및 R245A/R275A/R277A/R372A/W465A; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0123] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 하기 아미노산 치환 또는 아미노산 치환의 조합 중 하나를 포함한다:
- [0124] - R372A/K375A/D450N,
- [0125] - R372A/K375A/R376A/D450N,
- [0126] - K375A/R376A/E377A/E380A/D450N,
- [0127] - R372A/K375A/R376A/E377A/E380A/D450N,
- [0128] - M194V,
- [0129] - M194V/R372A/K375A,
- [0130] - S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L,
- [0131] - R245A/R275A/R277A/R372A/W465A/M589V,
- [0132] - R275A/325A/R372A/T560A,
- [0133] - N347A/D450N,
- [0134] - N347S/D450N/T560A/S573A/F594L,
- [0135] - R202K/R275A/N347S/R372A/D450N/T560A/F594L,
- [0136] - R275A/N347S/K375A/D450N/S592G,
- [0137] - R275A/N347S/R372A/D450N/T560A/F594L,
- [0138] - R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L
- [0139] - R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G,
- [0140] - R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L,
- [0141] - V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P,
- [0142] - G325A/N347S/K375A/D450N/S573A/M589V/S592G,
- [0143] - S230N/R277A/N347S/K375A/D450N,
- [0144] - T43I/R372A/K375A/A411T/D450N,
- [0145] - G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, 또는
- [0146] - Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G;
- [0147] 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0148] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 하기 아미노산 치환 또는 아미노산 치환의 조합 중 하나를 포함한다:
- [0149] - K375A/R376A/E377A/E380A/D450N,
- [0150] - R372A/K375A/R376A/E377A/E380A/D450N,

- [0151] - M194V,
- [0152] - M194V/R372A/K375A,
- [0153] - R245A/R275A/R277A/R372A/W465A/M589V,
- [0154] - R275A/325A/R372A/T560A,
- [0155] - N347A/D450N,
- [0156] - N347S/D450N/T560A/S573A/F594L,
- [0157] - R202K/R275A/N347S/R372A/D450N/T560A/F594L,
- [0158] - R275A/N347S/K375A/D450N/S592G,
- [0159] - R275A/N347S/R372A/D450N/T560A/F594L,
- [0160] - R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L
- [0161] - R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G,
- [0162] - R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L,
- [0163] - G325A/N347S/K375A/D450N/S573A/M589V/S592G,
- [0164] - S230N/R277A/N347S/K375A/D450N,
- [0165] - G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, 또는
- [0166] - Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G;
- [0167] 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0168] 바람직한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 하기 아미노산 치환 또는 아미노산 치환의 조합 중 하나를 포함한다:
- [0169] - R372A/K375A/D450N,
- [0170] - S351A/R372A/K375A/R388A/D450N/W465A/S573A/M589V/S592G/F594L,
- [0171] - R245A/R275A/R277A/R372A/W465A/M589V,
- [0172] - N347A/D450N,
- [0173] - N347S/D450N/T560A/S573A/F594L,
- [0174] - R202K/R275A/N347S/R372A/D450N/T560A/F594L,
- [0175] - R275A/N347S/K375A/D450N/S592G,
- [0176] - R275A/N347S/R372A/D450N/T560A/F594L,
- [0177] - R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L,
- [0178] - R275A/325A/R372A/T560A,
- [0179] - R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G,
- [0180] - R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L,
- [0181] - G325A/N347S/K375A/D450N/S573A/M589V/S592G,
- [0182] - S230N/R277A/N347S/K375A/D450N,
- [0183] - G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G, 또는
- [0184] - Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G;
- [0185] 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.

- [0186] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 서열번호: 1-8, 10-18, 108-113, 및 122-130 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0187] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 1-8 및 10-18 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0188] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 108-113 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0189] 한 실시양태에서, 상기 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 122-130 중 임의의 것 중에서 선택된 아미노산 서열을 갖는다.
- [0190] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 268 및/또는 346 (예를 들어, D268N 및/또는 D346N)에서 보존된 촉매 트리아드(catalytic triad)에 관여하는 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 11을 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0191] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 287, 287/290 및/또는 460/461 (예를 들어, K287A, K287A/K290A, 및/또는 R460A/K461A)에서, 절제에 중대한 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 12를 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0192] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 351, 356, 및/또는 379 (예를 들어, S351E, S351P, S351A, 및/또는 K356E)에서 표적 연결(target joining)에 관여하는 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 13을 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0193] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 560, 564, 571, 573, 589, 592, 및/또는 594 (예를 들어, T560A, S564P, S571N, S573A, M589V, S592G, 및/또는 F594L)에서 통합에 중대한 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 14를 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0194] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 325, 347, 350, 357 및/또는 465 (예를 들어, G325A, N347A, N347S, T350A 및/또는 W465A)에서 정렬에 관여하는 hyPB에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 15를 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0195] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어 아미노산 576 및/또는 587 (예를 들어, K576A 및/또는 I587A)에서, 잘 보존된 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 16를 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0196] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 예를 들어, 586 (예를 들어, H586A)에서, Zn^{2+} 결합에 관여하는 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 17를 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0197] 한 실시양태에서, 프로그램가능한 트랜스포사제는 예를 들어 315, 341, 372, 및/또는 375 (예를 들어, R315A, R341A, R372A, 및/또는 K375A)에서 통합에 관여하는 야생형 과활성 PiggyBac 트랜스포사제에 비해 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제 또는 서열번호: 18을 가진 변형된 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0198] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 야생형 과활성 PiggyBac 트랜스포사제와 비교하여 게놈으로의 DNA 통합의 높은 특이성을 위해 선택된다. 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9-18 및 108-113 중 어느 하나에 비해 본원에 개시된 변형 중 하나 이상을 갖는 아미노산 서열을 포함하며, 서열번호: 1-18 및 108-113 각각에 제시된 서열과 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%,

적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일하다.

- [0199] 실시예에 볼 수 있듯이, 새로 개발된 과활성 PiggyBac 트랜스포사제 치환 라이브러리를 사용하여 특이적 표적화된 전위를 수행하는 변형된 과활성 PiggyBac을 확인하였다. 이러한 라이브러리를 사용하여 양성 표적화된 전위를 가진 변형된 과활성 PiggyBac을 확인하였다.
- [0200] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산: 245, 275, 275, 277, 325, 347, 351, 372, 375, 388, 450, 465, 560, 564, 573, 589, 592, 594로부터 선택된 아미노산의 하나 이상의 치환을 포함할 수 있고; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0201] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 R245A, R275A, R275A, R277A, R275A/R277A, G325A, N347A, N347S, S351E, S351P, S351A, R372A, K375A, R388A, D450N, W465A, T560A, S564P, S573A, M589V, S592G, 또는 F594L로부터 선택된 아미노산 치환 중 하나 이상을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0202] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 D450N을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0203] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 변형 N347A를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 또 다른 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 N347S를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0204] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 이중 아미노산 치환 D450N 및 N347A를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 또 다른 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 D450N 및 N347S를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0205] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 R372A, K375A 및 D450을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0206] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 R245A 및 D450을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0207] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 R245A, G325A, 및 S573P를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0208] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 R245A, G325A, D450 및 S573P를 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0209] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 N347S 및 D450N을 포함하며; 넘버링은 서열번호: 9의 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0210] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 아미노산 치환 N347A 및 D450N을 포함하며; 넘버링은 서열번호: 9의 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다. 한 실시양태에서, 이 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 110의 아미노산 서열을 포함한다.
- [0211] 한 실시양태에서, 변형된 과활성 PiggyBac 트랜스포사제는 L25F, R36A, I42K, G59D, I212K, N245S, K252A 및/또는 Q271L 중에서 선택된 하나 또는 수개의 아미노산 치환을 포함하며; 넘버링은 서열번호: 9를 가진 야생형 과활성 PiggyBac 트랜스포사제를 기반으로 한다.
- [0212] 본원에 제공된 변형된 과활성 PiggyBac 트랜스포사제는 본원에 개시된 다른 요소, 예컨대 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질에 융합할 수 있거나; 단독으로, 또는 이러한 다른 요소와 조합하여 사용할 수 있다. 한 실시양태에서, 본원에 개시된 변형된 과활성 PiggyBac 트랜스포사제는 서열번호: 9의 아미노산 서열을 포함하며, 여기서:
- [0213] - 34번 위치의 아미노산이 V 또는 M이고,
- [0214] - 43번 위치의 아미노산이 T 또는 I이고,

- [0215] - 177번 위치의 아미노산이 Y 또는 H이고,
- [0216] - 202번 위치의 아미노산이 R 또는 K이고,
- [0217] - 230번 위치의 아미노산이 S 또는 N이고,
- [0218] - 245번 위치의 아미노산이 A이고,
- [0219] - 268번 위치의 아미노산이 D 또는 N이고,
- [0220] - 275번 위치의 아미노산이 R 또는 A이고,
- [0221] - 277번 위치의 아미노산이 R 또는 A이고,
- [0222] - 325번 위치의 아미노산이 A 또는 G이고,
- [0223] - 347번 위치의 아미노산이 N, S 또는 A이고,
- [0224] - 351번 위치의 아미노산이 E, P 또는 A이고,
- [0225] - 372번 위치의 아미노산이 R 또는 A이고,
- [0226] - 375번 위치의 아미노산이 A 또는 K이고,
- [0227] - 388번 위치의 아미노산이 R 또는 A이고,
- [0228] - 409번 위치의 아미노산이 K 또는 A이고,
- [0229] - 411번 위치의 아미노산이 A 또는 T이고,
- [0230] - 412번 위치의 아미노산이 K 또는 A이고,
- [0231] - 450번 위치의 아미노산이 D 또는 N이고,
- [0232] - 460번 위치의 아미노산이 R 또는 A이고,
- [0233] - 465번 위치의 아미노산이 W 또는 A이고,
- [0234] - 517번 위치의 아미노산이 S 또는 A이고,
- [0235] - 560번 위치의 아미노산이 T 또는 A이고,
- [0236] - 564번 위치의 아미노산이 P 또는 S이고,
- [0237] - 571번 위치의 아미노산이 S 또는 N이고,
- [0238] - 573번 위치의 아미노산이 S 또는 A이고,
- [0239] - 576번 위치의 아미노산이 K 또는 A이고,
- [0240] - 586번 위치의 아미노산이 H 또는 A이고,
- [0241] - 587번 위치의 아미노산이 I 또는 A이고,
- [0242] - 589번 위치의 아미노산이 M 또는 V이고,
- [0243] - 592번 위치의 아미노산이 G 또는 S이고/이거나,
- [0244] - 594번 위치의 아미노산이 L 또는 F이다.
- [0245] 한 실시양태에서, 트랜스포사제는 슬리핑 뷰티 트랜스포사제이다.
- [0246] 한 실시양태에서, 슬리핑 뷰티 트랜스포사제는 서열번호: 73에 대해 적어도 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하거나 이로 이루어진다. 이 단백질을 코딩하는 예시적인 핵산 서열은 서열번호: 72에 제시된 바와 같다.
- [0247] 한 실시양태에서, 트랜스포사제는 변형된 슬리핑 뷰티 트랜스포사제이다. 한 실시양태에서, 변형된 슬리핑 뷰티 트랜스포사제는 야생형 슬리핑 뷰티 트랜스포사제와 비교하여 하나 이상의 치환을 포함한다.
- [0248] 한 실시양태에서, 슬리핑 뷰티 트랜스포사제에서 하나 이상의 치환은 L25F, R36A, I42K, G59D, I212K, N245S,

K252A 및/또는 Q271L 중에서 선택되고; 넘버링은 서열번호: 73을 가진 야생형 슬리핑 뷰티 트랜스포사제를 기반으로 한다.

- [0249] 한 실시양태에서, 트랜스포사제는 Himar1C9 돌연변이가 아니다.
- [0250] 하기에 더 상세히 설명되는 바와 같이, 본 개시내용의 특정 측면은 또한 특히 본원에 기재된 과활성 PiggyBac 트랜스포사제 또는 변형된 과활성 PiggyBac 트랜스포사제 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.
- [0251] 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.
- [0252] 한편으로는 부위-특이적 DNA 결합 단백질, 및 다른 한편으로는 트랜스포사제에 관련된 임의의 실시양태는 본원에 기재된 융합 단백질의 경우에 필요한 부분만 약간 수정하여 적용한다.
- [0253] 따라서, 한 실시양태에서, 융합 단백질은
- [0254] (i) 상기에 기재된 바와 같은, RNA-가이드된 DNA 뉴클레아제, 징크 핑거 단백질 또는 전사 활성화제 유사 이펙터 뉴클레아제를 포함하거나 이로 이루어진 제1 단백질, 및
- [0255] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질
- [0256] 을 포함하거나 이로 이루어진다.
- [0257] 한 실시양태에서, 융합 단백질은
- [0258] (i) 상기에 기재된 바와 같은, RNA-가이드된 DNA 뉴클레아제 또는 징크 핑거 단백질을 포함하거나 이로 이루어진 제1 단백질, 및
- [0259] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질
- [0260] 을 포함하거나 이로 이루어진다.
- [0261] 한 실시양태에서, 융합 단백질은
- [0262] (i) 상기에 기재된 바와 같은, RNA-가이드된 DNA 뉴클레아제를 포함하거나 이로 이루어진 제1 단백질, 및
- [0263] (ii) 상기에 기재된 바와 같은, 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질
- [0264] 을 포함하거나 이로 이루어진다.
- [0265] 한 실시양태에서, 융합 단백질은
- [0266] (i) 상기에 기재된 바와 같은, Cas9 단백질 또는 이의 변이체를 포함하거나 이로 이루어진 제1 단백질, 및
- [0267] (ii) 상기에 기재된 바와 같은, 과활성 PiggyBac 트랜스포사제 또는 변형된 과활성 PiggyBac 트랜스포사제; 특히 상기에 기재된 바와 같은, 변형된 과활성 PiggyBac 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질
- [0268] 을 포함하거나 이로 이루어진다.
- [0269] 한 실시양태에서, 제1 단백질 및 제2 단백질은 어느 순서로든 융합 단백질에서 배향될 수 있다.
- [0270] 한 실시양태에서, 융합 단백질은 직접적으로 또는 링커를 통해 간접적으로 제2 단백질의 C-말단 말단부 (terminal end)에 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.
- [0271] 한 실시양태에서, 융합 단백질은 직접적으로 또는 링커를 통해 간접적으로 제2 단백질의 N-말단 말단부에 융합된 제1 단백질을 포함하거나 이로 이루어진다. 환언하면, 융합 단백질은, N-말단에서 C-말단까지: (i) 제1 단

백질 (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

[0272] 한 실시양태에서, 융합 단백질은 링커를 포함한다.

[0273] 링커의 적합한 예는 제1 단백질과 제2 단백질 사이 (임의의 순서로)에 펩티드 링커를 포함한다.

[0274] 한 실시양태에서, 펩티드 링커는 (GGG)_n, 서열번호: 114를 가진 (GGGG)_n, (G)_n, 서열번호: 115를 가진 (EAAAK)_n, XTEN 링커, 및 (XP)_n 모티프, 및 이들 중 임의의 것의 임의의 것의 조합을 포함하거나 이로 이루어진 군으로부터 선택되며, 여기서 n은 독립적으로 1 내지 50의 정수이다.

[0275] 한 실시양태에서, 링커는 12- 내지 24-아미노산 길이어거나, 36- 내지 72- 뉴클레오타이드 길이인 핵산 서열에 의해 코딩된다.

[0276] 한 실시양태에서, 링커는 XTEN 링커 또는 (GGG)_n 링커이다.

[0277] 한 실시양태에서, 링커는 표 1에 나타난 링커 중에서 선택된다.

[0278] [표 1]

링커	핵산 서열 (서열번호)	아미노산 서열 (서열번호)
GG5x3	ggtgcatctggcgggtgcatctggtggcgt (서열번호: 40)	GGSGGSGGG (서열번호: 41)
GG54x	ggaggagtggtgggtccggtagtggtggatcc (서열번호: 42)	GGSGGSGGSGGS (서열번호: 43)
GG55x	ggaggctccggtgggtctggtggagcggtgtagtggtcc (서열번호: 44)	GGSGGSGGSGGSGGS (서열번호: 45)
GG56x	ggaggcagtggtgggagcggtggtccggggtagtggtggtccggggatcc (서열번호: 46)	GGSGGSGGSGGSGGSGGS (서열번호: 47)
GG57x	ggagggttctggaggctccggtgggtccggggaagtgggggtcaggcgatcaggaggatcc (서열번호: 48)	GGSGGSGGSGGSGGSGGSGGS (서열번호: 49)
GG58x	ggaggtagcggagggtccggaggagcggtgggagtggtgggaagcggggaagtggaggatccggggaggatcc (서열번호: 50)	GGSGGSGGSGGSGGSGGSGGS (서열번호: 51)
링커 XTEN	tccggtagcgaacaccggggacttcagaatcgccacc ccggagtct (서열번호: 52)	SGSETPGTSESATPES (서열번호: 53)
링커 B	ggaagcgccggtagtcgggtctgggtctggcgagtcc (서열번호: 54)	GSAGSAAGSGEF (서열번호: 55)

[0279]

[0280] 한 실시양태에서, 링커는 서열번호: 41, 서열번호: 43, 서열번호: 45, 서열번호: 47, 서열번호: 49, 서열번호: 51, 서열번호: 53, 서열번호: 55, 또는 이의 임의의 조합을 포함하거나 이로 이루어진 아미노산 서열을 포함하며; 각각 서열번호: 40, 서열번호: 42, 서열번호: 44, 서열번호: 46, 서열번호: 48, 서열번호: 50, 서열번호: 52, 서열번호: 54의 예시적인 핵산 서열에 의해 코딩된다.

- [0281] 한 실시양태에서, 링커는 서열번호: 41의 아미노산 서열을 포함하거나 이로 이루어지며; 서열번호: 40의 예시적인 핵산 서열에 의해 코딩된다.
- [0282] 또한, 본 개시내용에 제공된 핵산 작제물 중 임의의 것의 발현으로부터 수득된 융합 단백질이 본원에 제공된다.
- [0283] 한 실시양태에서, 융합 단백질은 삼중 융합 단백질이다.
- [0284] 이러한 삼중 융합 단백질은 다음을 포함하거나 이루어질 수 있다:
- [0285] - 1개의 제1 단백질 (즉, 1개의 부위-특이적 DNA 결합 단백질) 및 2개의 제2 단백질 (즉, 2개의 트랜스포사제); 또는
- [0286] - 2개의 제1 단백질 (즉, 2개의 부위-특이적 DNA 결합 단백질) 및 1개의 제2 단백질 (즉, 1개의 트랜스포사제).
- [0287] 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:
- [0288] - (i) 부위-특이적 DNA 결합 단백질, (ii) 제1 트랜스포사제; (iii) 제2 트랜스포사제; 또는
- [0289] - (i) 제1 트랜스포사제; (ii) 부위-특이적 DNA 결합 단백질, (iii) 제2 트랜스포사제; 또는
- [0290] - (i) 제1 트랜스포사제; (ii) 제2 트랜스포사제, (iii) 부위-특이적 DNA 결합 단백질.
- [0291] 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.
- [0292] 한 실시양태에서, 제1 및 제2 트랜스포사제는 기능적 이량체를 형성할 수 있다.
- [0293] 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:
- [0294] - (i) 제1 부위-특이적 DNA 결합 단백질, (ii) 제2 부위-특이적 DNA 결합 단백질; (iii) 트랜스포사제; 또는
- [0295] - (i) 제1 부위-특이적 DNA 결합 단백질; (ii) 트랜스포사제, (iii) 제2 부위-특이적 DNA 결합 단백질; 또는
- [0296] - (i) 트랜스포사제; (ii) 제1 부위-특이적 DNA 결합 단백질, (iii) 제2 부위-특이적 DNA 결합 단백질.
- [0297] 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..
- [0298] 한 실시양태에서, 삼중 융합 단백질은 이의 단백질 중 2개 또는 3개의 단백질 사이에 링커를 임의로 포함한다.
- [0299] 또한, 다음을 포함하는 융합 단백질이 본원에 개시된다:
- [0300] (i) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질; 또는 상기에 기재된 바와 같은 상기 제2 단백질을 코딩하는 핵산 작제물, 및
- [0301] (ii) 적어도 하나의 특이적 RNA 서열에 결합할 수 있는 RNA-결합 단백질; 또는 상기 RNA-결합 단백질을 코딩하는 핵산 작제물.
- [0302] 한 실시양태에서, 융합 단백질은 상기에 기재된 바와 같은 링커를 포함한다.
- [0303] 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.

- [0304] 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.
- [0305] 일부 실시양태에서, RNA-결합 단백질은 MS2 박테리오파지 코트 단백질 (MCP) 또는 이의 단편이다.
- [0306] 일부 실시양태에서, MCP는 서열번호: 132 (예를 들어, 서열번호: 131을 가진 핵산 서열에 의해 코딩됨)와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성을 갖는다.
- [0307] 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.
- [0308] 일부 실시양태에서, 적어도 하나의 테트라루프는 MS2 RNA 테트라루프-결합 서열이다.
- [0309] 일부 실시양태에서, 테트라루프는 가이드 RNA (gRNA) 내에 포함된다. 특정 실시양태에서, gRNA는 상기에 기재된 바와 같이, Cas9 단백질과 복합체를 이룬다.
- [0310] 일부 실시양태에서, gRNA는 적어도 하나의 MS2 RNA 테트라루프-결합 서열을 포함한다. 일부 실시양태에서, gRNA는 하나 초과와 MS2 RNA 테트라루프-결합 서열을 포함한다.
- [0311] 일부 실시양태에서, 적어도 하나의 MS2 RNA 테트라루프-결합 서열을 포함하는 gRNA는 서열번호: 134와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성 (예를 들어, 서열번호: 133을 가진 DNA 서열에 의해 코딩됨)을 가진다.
- [0312] 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.
- [0313] 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.
- [0314] 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.
- [0315] 따라서, 한 실시양태에서, 본 발명의 조성물은
- [0316] a) 상기에 기재된 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진, 상기에 기재된 제1 단백질을 코딩하는 핵산 작제물;
- [0317] b) 트랜스포사제를 포함하거나 이로 이루어진, 상기에 기재된 제2 단백질을 코딩하는 핵산 작제물; 및
- [0318] c) 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스진을 포함하거나 이로 이루어진 핵산 작제물
- [0319] 을 포함하거나 이로 이루어진다.
- [0320] 또 다른 실시양태에서, 본 발명의 조성물은
- [0321] a) (i) 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질, 및 (ii) 트랜스포사제를 포함하거나 이로 이루어진 제2 단백질을 포함하거나 이로 이루어진, 상기에 기재된 융합 단백질을 코딩하는 핵산 작제물; 및
- [0322] b) 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스진을 포함하거나 이로 이루어진 핵산 작제물
- [0323] 을 포함하거나 이로 이루어진다.
- [0324] 한 실시양태에서, 융합 단백질을 코딩하는 핵산 작제물은 상기에 기재된 바와 같이 제1 단백질과 제2 단백질 사

이의, 또는 삼중 융합 단백질의 경우에, 이의 단백질 둘 사이 또는 세 단백질 사이의 링커를 코딩하는 핵산 서열을 추가로 포함한다.

- [0325] 본 개시내용에 따르면, 제1 및 제2 단백질, 또는 상기 제1 및 제2 단백질을 포함하거나 이로 이루어진 융합 단백질은 트랜스진의 계놈, 특히 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 이의 단편을 코딩하는 트랜스진의 부위-특이적 삽입을 가능하게 하고/하거나 촉진한다.
- [0326] 일부 실시양태는 다음 중 하나를 포함하는 플라스미드 또는 벡터 (예컨대, 발현 벡터)에 관한 것이다:
- [0327] - 제1 단백질을 코딩하는 핵산 작제물; 또는
- [0328] - 제2 단백질을 코딩하는 핵산 작제물; 또는
- [0329] - 제1 단백질을 코딩하는 핵산 작제물 및 제2 단백질을 코딩하는 핵산 작제물; 또는
- [0330] - 융합 단백질 또는 삼중 융합 단백질을 코딩하는 핵산 작제물.
- [0331] 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 *LAMA2* 트랜스진을 포함하는 핵산 작제물을 포함하는 제3 플라스미드와 조합되며, 여기서 조합이 생산 세포주 (예를 들어, 진핵 세포, 원핵 세포 및/또는 세포주)에 도입될 때, *LAMA2* 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.
- [0332] 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 *LAMA2* 트랜스진을 포함하는 핵산 작제물을 포함하는 제4 플라스미드와 조합되며, 여기서 조합이 생산 세포주 (예를 들어, 진핵 세포, 원핵 세포 및/또는 세포주)에 도입될 때, *LAMA2* 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.
- [0333] 한 실시양태에서, 제1 단백질, 제2 단백질, 제1 및 제2 단백질 둘 다 또는 융합 단백질, 및/또는 *LAMA2* 트랜스진은 렌티바이러스 입자를 사용하여 세포로 전달된다.
- [0334] 한 실시양태에서, 핵산 작제물은 표적 핵산 서열에 결합하도록 조작된 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진 제1 단백질을 코딩하는 제1 폴리뉴클레오티드 서열, *LAMA2* 트랜스진을 계놈에 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, 과활성 PiggyBac 트랜스포사제 또는 변형된 과활성 PiggyBac 트랜스포사제이다.
- [0335] 융합 단백질을 생성하기 위한 적합한 링커의 예는 상기에 기재되어 있다.
- [0336] 일부 실시양태에서, 제1 단백질이 제2 단백질과 별도의 플라스미드로부터 발현되기 때문에 링커가 필요하지 않다.
- [0337] 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.
- [0338] 한 실시양태에서, 핵산 작제물은 DNA 또는 RNA 형태이다.
- [0339] 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.
- [0340] 본 발명에 따르면, 조성물은 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스진 (본원에서

"*LAMA2* 트랜스진"이라고도 함)를 포함하거나 이로 이루어진 핵산 작제물을 추가로 포함한다.

- [0341] 한 실시양태에서, *LAMA2* 트랜스진은 라미닌- $\alpha 2$ 단백질, 특히 야생형 포유동물, 바람직하게는 인간, 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 임의의 핵산 서열일 수 있다.
- [0342] 한 실시양태에서, *LAMA2* 트랜스진은 *LAMA2* 야생형 유전자 (또는 이의 기능적 변이체 또는 단편), *LAMA2* cDNA (또는 이의 기능적 변이체 또는 단편) 또는 *LAMA2* 미니유전자(minigene) (또는 이의 기능적 변이체 또는 단편)를 포함할 수 있다.
- [0343] 한 실시양태에서, *LAMA2* 트랜스진은 전장 *LAMA2* 유전자를 포함한다.
- [0344] 라미닌- $\alpha 2$ 쇠 단백질은 *LAMA2* 유전자 (유전자 ID: 3908, 2020년 11월 24일에 업데이트됨)에 의해 코딩되며, 이는 390-kDa 단백질로 전사되고 번역된다. 라미닌- $\alpha 2$ 는 이황화 결합에 의해 서로 결합된 3개의 서브유닛 (알파, 베타 및 감마)으로 구성된 라미닌-211 (또는 메로신)으로 공지된 헤테로삼량체, 십자형 분자의 구성요소이다.
- [0345] 인간, 돼지, 침팬지, 개, 소, 마우스, 토끼 또는 래트의 라미닌- $\alpha 2$ 단백질을 제한 없이 포함하는 다수의 상이한 포유동물 라미닌- $\alpha 2$ 단백질의 코딩 서열(coding sequence)은 관련 기술분야에 공지되어 있으며, 서열 데이터베이스에서 쉽게 찾을 수 있다. 대안적으로, 코딩 서열은 폴리펩티드 서열을 기반으로 하여 관련 기술분야의 통상의 기술자에 의해 쉽게 결정될 수 있다.
- [0346] 특히, *LAMA2* 트랜스진은 인간 라미닌 서브유닛 $\alpha 2$ 이소형(isoform) 전구체 (NCBI 참조: NP_000417, 2020년 10월 22일에 제출) 또는 라미닌 서브유닛 $\alpha 2$ 이소형 b 전구체 (NCBI 참조: NP_001073291.2, 2020년 10월 24일에 제출)을 코딩할 수 있다.
- [0347] 한 실시양태에서, *LAMA2* 트랜스진은 바람직하게는 서열번호: 74를 가진, 인간 라미닌 서브유닛 $\alpha 2$ 이소형 전구체를 코딩한다.
- [0348] 번역 후, 라미닌- $\alpha 2$ 쇠는 전형적으로 서로 비공유결합으로 연관된 N-말단 단편 및 C-말단 단편으로 절단된다.
- [0349] 한 실시양태에서, *LAMA2* 트랜스진은 바람직하게는 라미닌 서브유닛 $\alpha 2$ 이소형 X1 (NCBI 참조번호: XP_005267038.1, 서열번호: 116을 가짐), 라미닌 서브유닛 $\alpha 2$ 이소형 X2 (NCBI 참조번호: XP_011534122.1, 서열번호: 117을 가짐), 라미닌 서브유닛 $\alpha 2$ 이소형 X3 (NCBI 참조번호: XP_005267039.1, 서열번호: 118을 가짐), 라미닌 서브유닛 $\alpha 2$ 이소형 X4 (NCBI 참조번호: XP_016866340.1, 서열번호: 119를 가짐), 라미닌 서브유닛 $\alpha 2$ 이소형 X5 (NCBI 참조번호: XP_01686634.1, 서열번호: 120을 가짐) 및 라미닌 서브유닛 $\alpha 2$ 이소형 X6 (NCBI 참조번호: XP_005267038.1, 서열번호: 121을 가짐)을 포함하거나 이로 이루어진 군으로부터 선택된 라미닌- $\alpha 2$ 단백질의 상이한 인간 성숙 이소형을 코딩할 수 있다.
- [0350] 본원에 사용된 바와 같이, 용어 "단편" 또는 "기능적 단편"은 라미닌- $\alpha 2$ 단백질 또는 이를 코딩하는 핵산을 언급할 때 상기에 기재된 바와 같은 라미닌- $\alpha 2$ 단백질로부터 유래된 단백질, 폴리펩티드 또는 핵산을 지칭하며 이는 여전히 전장 라미닌- $\alpha 2$ 단백질의 활성을 유지하나 그 서열이 전장의 단백질, 폴리펩티드 또는 핵산과 100% 동일하지는 않다. 기능적 단편은 상응하는 천연 분자보다 많거나, 적거나, 동일한 수의 잔기를 가질 수 있고/있거나, 하나 이상의 아미노산 또는 뉴클레오티드 치환을 함유할 수 있다. 바람직하게는, 기능적 단편은 단백질의 N-말단 말단부에 신호 펩티드를 포함하지 않는 성숙한 형태의 라미닌- $\alpha 2$ 단백질을 지칭한다. 통상의 기술자는 예를 들어 유니프롯(Uniprot) 데이터베이스에서 공개적으로 이용가능한 정보를 기반으로 하여, 신호 펩티드가 라미닌- $\alpha 2$ 단백질의 어느 부분에 상응하는지 쉽게 결정할 수 있다.
- [0351] 한 실시양태에서, *LAMA2* 트랜스진은 서열번호: 75를 가진 서열을 포함한다. 한 실시양태에서, *LAMA2* 트랜스진은 서열번호: 82를 가진 합성 인트론을 포함하는 인간 *LAMA2* 미니유전자를 포함한다. 한 실시양태에서, 상기 합성 인트론은 서열번호: 75를 가진 코딩 서열의 위치 3925 및 3926에 있는 핵염기(nucleobase) 사이에 포함된다.
- [0352] 한 실시양태에서, *LAMA2* 트랜스진은 전장 라미닌- $\alpha 2$ 단백질의 활성을 유지하는 기능적 라미닌- $\alpha 2$ 단백질 변이체를 코딩하는 핵산 서열일 수 있다. 특히, 라미닌- $\alpha 2$ 단백질 변이체는 이황화 결합에 의해 라미닌-211 β 및 γ 서브유닛에 결합하여 기능적인 십자형 라미닌-211 또는 라미닌-221을 형성하는 능력을 유지한다. 라미닌은 독립적인 네트워크를 형성하고 엔탁틴, 피브로넥틴, 및 페르레칸을 통해 IV형 콜라겐 네트워크와 연관된다. 그들은 또한 디스트로글리칸 당단백질 복합체 및 루테란 혈액형(blood group) 당단백질과 같은 인테그린 수용체 및 기타 원형질막 분자를 통해 세포막에 결합한다. 이들 상호작용을 통해, 라미닌은 세포 부착 및 분화, 세포

형상 및 이동, 조직 표현형의 유지, 및 조직 생존의 촉진에 결정적으로 기여한다.

- [0353] 바람직하게는, 본원에 사용된 바와 같이, 용어 "**변이체**" 또는 "**기능적 변이체**"는 라미닌- $\alpha 2$ 단백질과 관련하여, 야생형 라미닌- $\alpha 2$ 단백질 서열에 대해 적어도 70, 75, 80, 85, 90, 95 또는 99% 서열 동일성을 갖는 아미노산 또는 뉴클레오티드 서열을 갖는 폴리펩티드를 지칭한다.
- [0354] 보다 바람직하게는, 용어 "**변이체**" 또는 "**기능적 변이체**"는 야생형 라미닌- $\alpha 2$ 단백질 서열과 30, 25, 20, 15, 10 또는 5개 미만의 치환, 삽입 및/또는 결실에 의해 상이한 아미노산 서열을 갖는 폴리펩티드를 지칭한다. 한 실시양태에서 변이체는 하나 이상의 보존적 치환, 바람직하게는 15, 10 또는 5개 미만의 보존적 치환에 의해 야생형 라미닌- $\alpha 2$ 단백질 서열과 상이하다. 보존적 치환의 예는 염기성 아미노산 (아르기닌, 리신 및 히스티딘), 산성 아미노산 (글루탐산 및 아스파르트산), 극성 아미노산 (글루타민 및 아스파라긴), 소수성 아미노산 (메티오닌, 류신, 이소류신 및 발린), 방향족 아미노산 (페닐알라닌, 트립토판 및 티로신), 및 소형 아미노산 (글리신, 알라닌, 세린 및 트레오닌)의 군 내에 있다.
- [0355] 따라서 한 실시양태에서, *LAMA2* 트랜스진은 서열번호: 75에 대해 적어도 70, 75, 80, 85, 90, 95 또는 99% 서열 동일성을 갖는 서열을 포함한다.
- [0356] 한 실시양태에서, *LAMA2* 트랜스진은 라미닌- $\alpha 2$ 단백질, 이의 변이체 또는 단편을 코딩하는 최적화된 서열을 포함할 수 있다.
- [0357] 핵산 서열과 관련하여 용어 "최적화된"은 코돈 최적화를 지칭하며 전형적으로 인간에 대한 편향을 나타내는 코돈이 인간에 대한 편향을 나타내지 않는 동의 코돈(synonymous codon) (즉, 동일한 아미노산 잔기를 코딩하는 코돈)으로 변경됨을 의미한다. 따라서, 코돈의 변화는 코딩된 단백질의 어떤 아미노산 변화도 발생시키지 않는다.
- [0358] 본 발명에 따르면, *LAMA2* 트랜스진은 핵산 작제물에 포함된다.
- [0359] 한 실시양태에서, 핵산 작제물은 세포에서 상기 트랜스진의 발현을 지향하는 하나 이상의 제어 서열에 작동가능하게 연결된 *LAMA2* 트랜스진을 포함한다.
- [0360] 프로모터는 숙주 세포에 도입시 *LAMA2* 트랜스진의 발현을 매개하는 전사 조절 서열을 포함한다. 프로모터는 돌연변이, 말단절단, 및 하이브리드 프로모터를 포함한 세포에서 전사 활성을 나타내는 임의의 폴리뉴클레오티드일 수 있다. 프로모터는 구성적 또는 유도성 프로모터, 바람직하게는 구성적 프로모터, 보다 바람직하게는 강한 구성적 프로모터일 수 있다.
- [0361] 적합한 프로모터의 예는 바람직하게는 서열번호: 76을 가진 CMV 프로모터, 바람직하게는 서열번호: 77을 가진 CAG 프로모터, 바람직하게는 서열번호: 78을 가진 EF1- α 프로모터, 바람직하게는 서열번호: 79를 가진 SV-40, 및 바람직하게는 서열번호: 80을 가진, 하이브리드 알부민 인헨서- $\alpha 1$ -항트립신 (Ea1bAAT) 간 특이적 프로모터를 포함하나, 이에 제한되지는 않는다.
- [0362] 대안적 실시양태에서, *LAMA2* 트랜스진은 또한 이의 발현이 통합 부위에서 또는 그 근처에서 내인성 프로모터에 의해 구동되는 이러한 방식으로 세포의 게놈으로 삽입될 수 있다. 특히, *LAMA2* 트랜스진을 포함하는 핵산 작제물은 트랜스진이 내인성 발현 제어를 갖기 위해, 내인성 *LAMA2* 또는 알부민 유전자와 같은 능동적으로 발현된 유전자의 인트론으로 통합될 때 유전자 발현을 보장하기 위해 스플라이스 수용체를 함유할 수 있다.
- [0363] 스플라이스 수용체 부위는 문헌 (Padgett *et al.*, 1988. *Annu Rev Biochem.* 55:1119-1150)과 같이 발현될 스플라이스 수용체 부위를 따르는 서열을 표적으로 하는 신호를 제공한다. 스플라이스 수용체 부위는 일반적으로 인트론 RNA 서열을 제거하기 위한 RNA 스플라이싱에 관여하는 뉴클레오티드 서열이다. 스플라이스 수용체 부위는 통상 인트론의 절제에 관여하며, 그 동안 스플라이소좀(spliceosome)이라고 하는 RNA-단백질 복합체에 의해 결합되고, 절단된 다음에, 이미 절단된 스플라이스 공여체 부위에 연결된다. 스플라이스 수용체 서열은 관련 기술분야에 널리 공지 있고 이들이 스플라이싱을 매개하는 엑손과 인트론 사이의 위치에 있는 유전자로부터 쉽게 획득될 수 있다. 대안적으로, 스플라이스 수용체 부위는 화학적으로 또는 효소적으로 합성될 수 있다. 스플라이스 수용체 (SA) 부위는 전형적으로 고도로 보존된 AG 디뉴클레오티드로 종료된다. 서열의 나머지 뉴클레오티드는 주로 시티딘 및/또는 티미딘이다. 한 실시양태에서, 스플라이스 수용체는 서열번호: 81의 서열을 갖는다.
- [0364] 제어 서열은 또한 적절한 전사 개시, 종결, 및 인헨서 서열; 스플라이싱 및 폴리아데닐화 신호와 같은 효율적인 RNA 처리 신호; 세포질 mRNA를 안정화시키는 서열; 번역 효율을 향상시키는 서열 (즉, 코작 공통 서열(Kozak

consensus sequence)); 및/또는 단백질 안정성을 향상시키는 서열을 포함할 수 있다. 매우 다수의 발현 제어 서열, 예를 들어 천연, 구성적, 유도성 및/또는 조직-특이적이 관련 기술분야에 공지되어 있고 *LAMA2* 트랜스진의 발현을 유도하기 위해 활용될 수 있다. 전형적으로, *LAMA2* 트랜스진은 전사 프로모터 및 전사 종결자에 작동가능하게 연결된다.

- [0365] 폴리(A) 신호는 전형적으로 a) 프리메신저(premessenger) RNA (pre-mRNA)의 3'-말단 세척(leavage) 및 폴리아데닐화 둘 다에 필요하고 하류 전사 종결을 촉진하는 것으로 밝혀진 공통 서열 AAUAAA, 및 b) AAAAAA를 폴리(A) 신호로 활용하는 효율을 제어하는 AAUAAA 서열의 상류 및 하류 추가 요소로 이루어진다. 포유동물 유전자에서 이들 모티프에는 상당한 가변성이 있다.
- [0366] 한 실시양태에서, 핵산 작제물 내의 폴리아데닐화 신호 서열은 포유동물 유전자 또는 바이러스 유전자의 폴리아데닐화 신호 서열이다. 적합한 폴리아데닐화 신호는 SV40 초기 폴리아데닐화 신호, SV40 후기 폴리아데닐화 신호, HSV 티미딘 키나제 폴리아데닐화 신호, 프로타민 유전자 폴리아데닐화 신호, 아데노바이러스 5 E1b 폴리아데닐화 신호, 성장 호르몬 폴리아데닐화 신호, PBGD 폴리아데닐화 신호, 뿐만 아니라 인실리코 설계된 폴리아데닐화 신호 (합성) 등을 포함하나, 이에 제한되지는 않는다. 한 실시양태에서, 폴리아데닐화 신호 서열은 서열 번호: 83, 84 및 85를 포함하거나 이로 이루어진 군으로부터 선택된 서열을 갖는 폴리(A)이다.
- [0367] 한 실시양태에서, *LAMA2* 트랜스진을 포함하는 핵산 작제물은 또한 트랜스진 발현 수준을 증가시키고, 침묵을 피하고, 발현 가변성을 감소시키기 위해 절연체 요소를 포함한다.
- [0368] 절연체는 헤테로크로마틴의 전개 및 유전자의 침묵 (장벽 활성화)을 방지하고 인핸서-차단 활성을 갖는 시스-작용 조절 서열의 복잡한 부류이다. 절연체는 전형적으로 길이가 300 bp 내지 2000 bp이다. CTCF 절연체, 집시(gypsy) 절연체 및 β -글로빈 유전자좌를 포함하여, 절연체의 많은 예가 있다.
- [0369] 한 실시예에서, 절연체 요소 서열은 핵산 작제물의 양 말단에 플랭킹한다.
- [0370] 한 실시양태에서, *LAMA2* 트랜스진을 포함하는 핵산 작제물은 서열번호: 86 및 87을 가진 절연체 요소를 포함한다.
- [0371] 한 실시양태에서, 본원에 개시된 트랜스포사제는 ITR 사이의 요소의 절제 및 후속 재통합을 촉매하기 위해 트랜스포사제-결합 부위를 운반하는 역 말단 반복부 (ITR) 요소를 인식한다.
- [0372] 따라서, 특정한 실시양태에서, *LAMA2* 트랜스진을 포함하는 상기 핵산 작제물은 트랜스포사제 결합 부위, 바람직하게는 서열번호: 88 및 89의 5'-ITR 및 3'-ITR을 운반하는 역 말단 반복부 (ITR)에 의해 플랭킹된다.
- [0373] 한 실시양태에서, ITR 사이의 크기 증가는 *LAMA2* 트랜스진의 게놈 DNA로의 전달을 긍정적으로 조절한다.
- [0374] 한 실시양태에서, ITR 사이의 크기는 적어도 105 bp이다. 한 실시양태에서, ITR 사이의 크기는 적어도 200 bp이다. 한 실시양태에서, ITR 사이의 크기는 적어도 300 bp이다.
- [0375] 본원에 사용된 바와 같이, 용어 "적어도 300 bp"는 300 bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 1 kb, 2 kb 또는 그 초과를 의미한다.
- [0376] 한 실시양태에서, *LAMA2* 트랜스진을 포함하는 핵산 작제물은 발현 벡터에 함유된다.
- [0377] 적절한 벡터의 예는 재조합 통합 또는 비통합 바이러스 벡터 및 재조합 박테리오파지 DNA, 플라스미드 DNA 또는 코스미드 DNA로부터 유래된 벡터를 포함하나, 이에 제한되지는 않는다. 한 실시양태에서, 상기 벡터는 플라스미드 벡터, 미니씨클 벡터 또는 도기 본 DNA 공여체 벡터이다. 바람직하게는, 벡터는 재조합 통합 또는 비통합 바이러스 벡터이다. 재조합 바이러스 벡터의 예는 헤르페스 바이러스, 레트로바이러스, 렌티바이러스, 백시니아 바이러스, 아데노바이러스, 아데노-연관 바이러스 또는 소 유두종 바이러스로부터 유래된 벡터를 포함하나 이에 제한되지는 않는다.
- [0378] 한 실시양태에서 상기 바이러스 벡터는 렌티바이러스 또는 레트로바이러스 벡터이다. 따라서 본 개시내용은 또한 상기에 기재된 바와 같이 *LAMA2* 트랜스진을 포함하는 핵산 작제물, 또는 상기 핵산 작제물을 포함하는 발현 벡터를 포함하는 바이러스 입자에 관한 것이다. 특히, 핵산 작제물 또는 발현 벡터는 "바이러스 벡터 입자"라고도 명명되는 "바이러스 입자"를 생성하기 위해 바이러스 캡시드에 패키징될 수 있다.
- [0379] 한 실시양태에서, 핵산 작제물 또는 발현 벡터는 "렌티바이러스 입자" 또는 "AAV 입자"를 생성하기 위해 렌티바이러스 또는 AAV-유래 캡시드에 패키징된다.

- [0380] 본 발명에 따르면, 제1 단백질은 표적 핵산 서열에 결합하고 이를 절단할 수 있는 부위-특이적 DNA 결합 단백질을 포함하거나 이로 이루어진다.
- [0381] 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.
- [0382] 상기 제1 단백질은 "**표적 핵산 서열**"로 불리는, 세포의 게놈 내의 임의의 선택 서열에 결합하도록 조작될 수 있다. 예를 들어, 서열 5'-GAATTC-3'은 EcoRI 제한 엔도뉴클레아제의 표적 부위이다.
- [0383] 한 실시양태에서, 표적 핵산 서열은 세포 게놈의 세이프 하버 유전자좌(safe harbor locus) 내에 있다. 한 실시양태에서, 표적 핵산 서열은 세포 게놈의 내인성 *LAMA2* 유전자 내에 있다. 특히, 제1 및/또는 제2 단백질은 세이프 하버 유전자좌 내에서 라미닌- $\alpha 2$ 단백질을 코딩하는 트랜스진에 결합 및/또는 통합되도록 조작될 수 있다.
- [0384] "**세이프 하버 유전자좌**"는 통합된 물질이 내인성 유전자 구조 또는 기능을 방해하지 않고 적절하게 발현될 수 있는 세포 게놈의 영역을 지칭한다. 세이프 하버 유전자좌는 AAVS1 (PPP1R12C의 인트론 1), HPRT, H11, hRosa26, 알부민 및 F-A 영역을 포함하나, 이에 제한되지는 않는다. 세이프 하버 유전자좌는 편재하여 발현되는 유전자 및/또는 조직 특이적 발현 (예를 들어, 근육)을 가진 유전자의 엑손 또는 인트론일 수 있다. 세이프 하버 유전자좌는 PPP1R12C의 엑손 1, 인트론 1 또는 엑손 2, HPRT의 엑손 1, 인트론 1 또는 엑손 2, 및 hRosa26의 엑손 1, 인트론 1 또는 엑손 2, **알부민** 유전자의 인트론 1으로 이루어진 군으로부터 선택될 수 있다. 세이프 하버 유전자좌는 또한 내인성 유전자가 없고 게놈 구조 또는 기능을 교란시키지 않고 삽입된 트랜스진의 발현을 가능하게 하는 개방 염색질을 가진 게놈의 영역을 포함할 수 있다.
- [0385] 한 실시양태에서, 제1 및/또는 제2 단백질은 *LAMA2* 트랜스진이 세포 게놈 내의 내인성 *LAMA2* 유전자, 특히 내인성 *LAMA2* 유전자의 인트론 1에 결합하고 통합되도록 조작된다.
- [0386] 한 실시양태에서, 임의로 가이드 RNA와 조합된 제1 단백질은 서열번호: 90-97 중 어느 하나로 이루어진 군으로부터 선택된 표적 핵산 서열에 결합한다.
- [0387] 한 실시양태에서, 본 발명의 조성물은 가이드 RNA를 추가로 포함한다.
- [0388] 본원에 사용된 바와 같이, "**가이드 RNA**", "**gRNA**" 또는 "**단일 가이드 RNA**"는 표적 핵산에 대한 gRNA/Cas 복합체의 특이적 표적화 또는 귀소를 촉진하는 핵산을 지칭한다.
- [0389] 특히, gRNA는 전이활성화 crRNA (tracrRNA) 및 crRNA를 포함하는 RNA 분자를 지칭한다. 바람직하게는, 상기 가이드 RNA는 별도로 사용되거나 함께 융합될 수 있는 crRNA 및 tracrRNA에 상응한다. 표적 서열과 쌍을 형성하는 상보적 서열은 Cas 단백질을 모집하여 표적 서열에서 DNA에 결합하고 이를 절단한다.
- [0390] 한 실시양태에서, 가이드 RNA는 바람직하게는 서열번호: 90-97 중 어느 하나로 이루어진 군으로부터 선택되는 표적 서열의 일부에 상보적인 서열을 포함하도록 조작된다.
- [0391] 본원에 사용된 바와 같이, 용어 "**상보적 서열**"은 표준 낮은 엄격한 조건 하에 폴리뉴클레오타이드의 또 다른 부분에 하이브리드화할 수 있는 폴리뉴클레오타이드의 서열 (예를 들어, crRNA 또는 tracrRNA의 일부)을 지칭한다. 바람직하게는, 서열은 가닥 사이의 왓슨-크릭(Watson-Crick) 염기 쌍형성, 즉 아데닌과 티민 (A-T) 뉴클레오타이드 및 구아닌과 시토신 (G-C) 뉴클레오타이드 사이의 고유 염기쌍형성에 의존하는 두 핵산 가닥 사이의 상보성에 따라 서로 상보적이다.
- [0392] 본 개시내용은 또한 *LAMA2* 트랜스진을 세포의 게놈 내의 표적 핵산 서열로 통합하는 방법에 관한 것이다. 상기 방법은 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)이 표적 핵산 서열을 절단하고 상기 표적 핵산 내에 *LAMA2* 트랜스진을 통합하도록 상기에 기재된 바와 같은 조성물을 세포 내로 도입하는 단계를 포함한다.
- [0393] 상기 방법은 세포에 다음을 도입하는 것을 포함한다: 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서) 또는 이를 코딩하는 하나 또는 여러 핵산 작제물; 가이드 RNA; 및 라미닌- $\alpha 2$ 단백질, 이의 기능적 변이체 또는 단편을 코딩하는 트랜스진을 포함하거나 이로 이루어진 핵산 작제물.
- [0394] 이들 요소는 상기에 기재된 바와 같이 상기 요소를 코딩하는 핵산 작제물(들) 또는 발현 벡터(들)를 세포 내로 도입한 결과로서 세포내 계내(*in situ*)에서 합성될 수 있다. 대안적으로, 상기 요소는 세포 외부에서 생산된

다음에 내부에 도입될 수 있다.

- [0395] 한 실시양태에서, 상기에 기재된 바와 같은 *LAMA2* 트랜스진을 포함하는 핵산 작제물, 발현 벡터 또는 바이러스 입자가 세포 내로 도입된다.
- [0396] 한 실시양태에서, 제1 및 제2 단백질을 코딩하는 하나 또는 수개의 폴리뉴클레오티드 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)는 예를 들어 전기천공법 또는 지질 나노입자에 의해 세포에 직접 도입되는 mRNA 형태로 형질감염될 수 있다.
- [0397] 한 실시양태에서, 가이드 RNA는 또한 세포 외부에서 생산된 다음에, 예를 들어 전기천공법 또는 지질 나노입자에 의해 내부에 도입될 수 있다.
- [0398] 한 실시양태에서, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 세포 내로 단독으로 도입되거나 가이드 RNA와 사전-복합체화될 수 있다.
- [0399] 한 실시양태에서, 가이드 RNA 및/또는 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 핵산 작제물 또는 발현 벡터에 의해 코딩된다.
- [0400] 본 발명에 따른 생성물 (단백질 및 핵산)은 비제한적 예로서 리포솜 전달 수단, 중합체 담체, 화학적 담체, 리포플렉스, 폴리플렉스, 덴드리머, 나노입자, 에멀전, 천연 엔도사이토시스 또는 식균작용 경로에 의해, 뿐만 아니라 물리적 방법 예컨대 전기천공법에 의해 세포 또는 세포하 구획 내부로 전달될 수 있다.
- [0401] 상기 핵산 작제물 또는 발현 벡터는 관련 기술분야에 공지된 임의의 방법에 의해 세포 내로 도입될 수 있으며, 비제한적 예로서, 핵산 작제물 또는 발현 벡터가 세포 계놈으로 통합되는 안전한 형질전환 방법, 핵산 작제물 또는 발현 벡터가 세포 및 바이러스-매개 방법의 계놈으로 통합되지 않는 일시적 형질전환 방법을 포함한다. 예를 들어, 일시적 형질전환 방법은 예를 들어 미세주입, 전기천공법 또는 유전자 총법(particle bombardment)을 포함한다.
- [0402] 본 발명에 따른 핵산 분자 또는 핵산 작제물 또는 발현 벡터는 인산칼슘 형질감염, DEAE-텍스트란 형질감염, 전기천공법, 미세주입, 바이오리스틱(biolistic), 바이러스 감염 또는 리포솜-매개 형질감염을 포함하나 이에 제한되지는 않는 임의의 공지 기술을 사용하여 세포 내로 전달될 수 있다.
- [0403] 한 실시양태에서, RNA, 바람직하게는 제1 및 제2 단백질을 코딩하는 가이드 RNA 또는 mRNA (별도로 또는 상기에 기재된 융합 단백질의 일부로서)는 예를 들어 시험관내 전사에 의해 시험관내에서 생성될 수 있다. 이어서 RNA는 전기천공에 의해 세포 내로 도입될 수 있다 (예를 들어 문헌 [Almås et al., 2011. *Cytotherapy*. 13(5):629-640; Rabinovich et al., 2009. *Hum Gene Ther*. 20(1):51-61]; 및 [Beatty et al., 2013. *Cancer Immunol Res*. 2(2):112-20]에 기재된 바와 같이).
- [0404] 대안적으로, RNA는 리포솜 또는 양이온 분자 등과 같은 다른 수단에 의해 도입될 수 있다.
- [0405] 한 실시양태에서, 세포 내로 도입된 핵산 작제물(들) 또는 벡터(들)은 에피솜으로(episomally) 발현될 수 있거나 세포의 계놈 내로 통합될 수 있다.
- [0406] 한 실시양태에서, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 코딩하는 가이드 RNA 및 mRNA와 같은 폴리뉴클레오티드 및/또는 *LAMA2* 트랜스진을 포함하거나 이로 이루어진 핵산 작제물은 나노입자, 바람직하게는 지질 나노입자 (LNP)에 의해 세포 또는 환자에게 전달될 수 있다. 관련 기술분야에 공지된 임의의 지질 또는 지질의 조합을 사용하여 LNP를 생성할 수 있다. LNP를 생산하는 데 사용되는 지질의 예는, DOTMA, DOSPA, DOTAP, DMRIE, DC-콜레르테롤, DOTAP-콜레르테롤, GAP-DMORIE-DPyPE, 및 GL67A-DOPE-DMPE-폴리에틸렌 글리콜 (PEG)이다. 양이온성 지질의 예는, 98N12-5, C12-200, Dlin-KC2- DMA (KC2), Dlin-MC3-DMA (MC3), XTC, MD1, 및 7C1이다. 중성 지질의 예는, DPSC, DPPC, POPC, DOPE, 및 SM이다. PEG-변형 지질의 예는, PEG-DMG, PEG- CerC14, 및 PEG-CerC20이다.
- [0407] 한 실시양태에서, 방법은 시험관내 또는 생체의 방법이다. 시험관내 방법은 세포 배양에서 수행된다.
- [0408] 또 다른 측면에서, 본 개시내용은 상기에 기재된 방법에 의해 수득가능하거나 수득된 분리된 조작된 세포에 관한 것이다. 상기 조작된 세포는 상기에 기재된 바와 같이 적어도 라미닌- α 2 단백질을 코딩하는 트랜스진을 포함한다.
- [0409] 한 실시양태에서, 단리된 조작된 세포는 상기에 기재된 바와 같이 세포에서 상기 트랜스진의 발현을 지향하는

제어 서열 중 임의의 하나 또는 여러 개를 포함한, 라미닌- $\alpha 2$ 를 코딩하는 적어도 하나의 트랜스진을 포함한다.

[0410] 일부 실시양태에서, 단리된 조작된 세포는 상기에 기재된 바와 같은 플랭킹 ITR을 포함한, 라미닌- $\alpha 2$ 를 코딩하는 적어도 하나의 트랜스진을 포함한다.

[0411] 한 실시양태에서, ITR 간 크기는 적어도 105 bp, 적어도 200 bp, 적어도 300 bp 또는 초과이다.

[0412] 본 개시내용의 조작된 세포는 생체의 유전자 요법 목적을 위해 사용될 수 있다. 이러한 실시양태에서, 상기에 기재된 바와 같은, 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), 가이드 RNA 및 LAMA2 트랜스진, 또는 핵산 작제물(들), 발현 벡터(들) 또는 바이러스 입자(들)이, 세포에 도입된다. 상기 단리된 세포는 후속적으로 환자 또는 대상체에게 이식될 수 있다. 이 이식 단계는 관련 기술분야에 공지된 임의의 이식 방법을 사용하여 완수될 수 있다. 예를 들어, 유전적으로 변형된 세포는 환자의 혈액에 직접 주입되거나 원하는 근육에 주입될 수 있으며, 그렇지 않으면 환자에게 투여될 수 있다. 이식된 세포는 자가, 동종 또는 이종 기원을 가질 수 있다. 특히, 상기 세포는 공여체 또는 환자의 골격근, 평활근 또는 심근으로부터 단리된 근세포이거나 골수 세포 또는 말초 혈액으로부터의 간엽 줄기 세포일 수 있다. 임상 사용을 위해, 세포 단리는 일반적으로 우수 의약품 제조 및 품질관리 기준(Good Manufacturing Practices) (GMP) 조건 하에 수행될 것이다.

[0413] 적합한 세포는 진핵 및 원핵 세포 및/또는 세포주를 포함하나 이에 제한되지는 않는다. 바람직하게는, 상기 세포는 포유동물 세포와 같은 진핵 세포이고, 이들은 인간, 비인간 영장류 예컨대 유인원; 침팬지; 원숭이, 및 오랑우탄, 개와 고양이를 포함한 길들인 동물(domesticated animal), 뿐만 아니라 말, 소, 돼지, 양, 및 염소와 같은 가축, 또는 마우스, 래트, 기니피그, 토끼, 햄스터 등을 포함하나 이에 제한되지는 않는 기타 포유동물 종들을 포함하나, 이에 제한되지는 않는다. 관련 기술분야의 통상의 기술자는 이식될 환자 또는 대상체에 따라 보다 적절한 세포를 선택할 것이다.

[0414] 상기 조작된 세포는 또한 근육 세포일 수 있다. 본원에 사용된 바와 같이, 용어 "근육"은 심근 (즉, 심장) 및 골격근을 지칭한다. 본원에 사용된 바와 같이, 용어 "근육 세포"는 근세포, 근관, 근모세포, 및/또는 위성 세포를 지칭한다.

[0415] 상기 조작된 세포는 줄기 세포 또는 유도 만능 줄기 세포와 같은 자기-재생 및 다능성 특성을 갖는 세포일 수 있다. 적합한 줄기 세포는 또한 예를 들어 배아 줄기 세포, 유도 만능 줄기 세포 (iPSC), 조혈 줄기 세포, 신경 줄기 세포 및 중간엽 줄기 세포를 포함한다. 줄기 세포는 바람직하게는 중간엽 줄기 세포이다. 중간엽 줄기 세포 (MSC)는 조골세포, 연골세포, 지방세포, 또는 근육세포 중 적어도 하나로 분화할 수 있으며 임의의 유형의 조직으로부터 단리될 수 있다. 일반적으로 MSC는 골수, 지방 조직, 제대, 또는 말초 혈액으로부터 단리될 것이다. 이를 수득하는 방법은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 유도 만능 줄기 세포 (iPS 세포 또는 iPSC라고도 함)는 성체 세포로부터 직접 생성될 수 있는 일종의 만능 줄기 세포이다. 야마나카 (Yamanaka) 등은 Oct3/4, Sox2, Klf4 및 c-Myc 유전자를 마우스 및 인간 섬유모세포로 전달하고 세포가 유전자를 발현하도록 강제하여 iPS 세포를 유도하였다 (WO 2007/069666). 톰슨(Thomson) 등은 후속적으로 Klf4 및 c-Myc 대신에 Nanog 및 Lin28을 사용하여 인간 iPS 세포를 생산하였다 (WO 2008/118820). 특정 실시양태에서, 세포는 근모세포이다. 근모세포는 줄기 세포, 예를 들어, 근이영양증과 같은 근육 장애를 가진 환자로부터 유래된 iPSC를 포함하는 iPSC로부터 유래될 수 있다.

[0416] 본 개시내용에 따른 조성물은 바람직하게는 본 발명에 따른 생성물 예컨대 상기 LAMA2 트랜스진을 코딩하는 핵산 작제물, 제1 및 제2 단백질을 코딩하는 핵산 작제물 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), 상기 생성물을 코딩하는 가이드 RNA, 발현 벡터 또는 바이러스 입자를 포함하는 폴리뉴클레오티드 또는 핵산 작제물의 치료학적 유효량을 포함하는 약제학적 조성물의 형태로 사용된다.

[0417] 본 개시내용과 관련하여, 치료학적 유효량은 그러한 용어가 적용되는 장애 또는 병태의 진행을 역전, 완화 또는 억제하거나 또는 그러한 용어가 적용되는 장애 또는 병태의 하나 이상의 증상의 진행을 역전, 완화 또는 억제하기에 충분한 용량을 지칭한다.

[0418] 유효 용량은 의료 분야의 통상의 기술자가 인정할, 사용되는 조성물, 투여 경로, 성별, 연령 및 체중과 같은 고려되는 개체의 신체적 특성, 병용 약물, 및 기타 요인과 같은 요인에 따라 결정되고 조정된다.

[0419] 본 개시내용의 다양한 실시양태에서, 약제학적 조성물은 약제학적으로 허용되는 담체 및/또는 비히클을 포함한다.

- [0420] "약제학적으로 허용되는 담체"는 포유동물, 특히 인간에게 적절하게 투여될 때 부작용, 알레르기 또는 기타 바람직하지 않은 반응을 초래하지 않는 비히클을 지칭한다. 약제학적으로 허용되는 담체 또는 부형제는 임의의 유형의 비독성 고체, 반-고체 또는 액체 충전제, 희석제, 캡슐화 물질 또는 보조 제형(formulation auxiliary)을 지칭한다.
- [0421] 바람직하게는, 약제학적 조성물은 주사될 수 있는 제형에 대해 약제학적으로 허용되는 비히클을 함유한다. 이들은 특히 등장성, 멸균, 식염수 (인산일나트륨 또는 이나트륨, 염화나트륨, 염화칼륨, 염화칼슘 또는 염화마그네슘 등 또는 이러한 염의 혼합물), 또는 건조, 특히 동결-건조된 조성물일 수 있으며, 이는 첨가 시 멸균수나 생리식염수의 경우에, 주사액의 구성이 가능하다.
- [0422] 주사 사용에 적합한 약제학적 형태는 멸균 수용액 또는 현탁액을 포함한다. 용액 또는 현탁액은 세포와 호환되는 첨가제를 포함할 수 있다. 용액 또는 현탁액은 비-바이러스 벡터, 바이러스-벡터 및 나노입자와 호환되고 성분이 표적 세포로 들어가는 것을 방해하지 않는 첨가제를 포함할 수 있다. 모든 경우에, 형태는 무균 상태여야 하며 쉽게 주사할 수 있는 정도의 유동성이 있어야 한다. 이는 제조 및 보관 조건 하에 안정적이어야 하며 박테리아 및 진균과 같은 미생물의 오염 작용에 대해 보존되어야 한다. 적절한 용액의 예는 완충제, 예컨대 인산염 완충 식염수 (PBS) 또는 링거 락테이트이다.
- [0423] 본 개시내용에 따른 조성물은 요법, 특히 *LAMA2*-관련 근이영양증의 치료를 위해 사용된다.
- [0424] *LAMA2*-관련 근이영양증은 운동에 사용되는 근육 (골격근)의 쇠약(weakness) 및 소모(wasting) (위축)를 유발하는 장애이다. 이 병태는 중증의, 조기-발병 유형에서 보다 경미한 후기-발병 형태에 이르기까지 중증도가 다양하다.
- [0425] 조기-발병 *LAMA2*-관련 근이영양증은 출생시 또는 생후 처음 몇 개월 이내에 명백하다. 이는 선천성 근이영양증이라고 하는 근육 장애 부류의 일부로 간주되며 때때로 선천성 근이영양증 1A형이라고 한다. 영향을 받은 영아는 중증 근육 쇠약, 근육 긴장도의 부족 (저긴장증), 자발적인 움직임이 거의 없고, 관절 기형 (구축(contracture))을 가질 수 있다. 얼굴과 목의 근육이 쇠약해지면 수유가 어려워지고 예상 속도로 성장 및 체중 증가가 불가능할 수 있다. 가슴의 근육이 쇠약해질 때 발생하는 호흡 부전은 약한 울음과 호흡 문제를 유발하며 이는 빈번하고, 잠재적으로 생명을 위협하는 폐 감염을 야기할 수 있다.
- [0426] 영향을 받은 소아가 성장함에 따라, 이들은 종종 척추의 좌우 만곡(side-to-side curvature) (척추측만증(scoliosis))과 등의 안쪽 만곡 (전만증(lordosis))이 비정상적으로 점차 악화된다. 조기-발병 *LAMA2*-관련 근이영양증을 가진 소아는 종종 걷는 능력이 발달하지 않는다. 언어 장애는 안면 근육의 쇠약과 혀가 커진 결과일 수 있다. 조기-발병 *LAMA2*-관련 근이영양증을 가진 개체의 약 1/3에서 발작(Seizure)이 발생하며; 드물게는, 이러한 형태의 장애에서 심장 합병증이 발생한다.
- [0427] 후기-발병 *LAMA2*-관련 근이영양증의 증상은 어린 시절이나 성인기에 분명해지며 사지-거들(limb-girdle) 근이영양증으로 분류되는 근육 장애의 그룹의 증상과 유사하다. 후기-발병 *LAMA2*-관련 근이영양증에서, 가장 영향을 받는 근육은 신체에 가장 가까운 근육 (근위 근육), 구체적으로 어깨, 상완, 골반 부위, 및 허벅지의 근육이다. 후기-발병 *LAMA2*-관련 근이영양증을 가진 소아는 때때로 걷기와 같은 운동 기술의 발달이 지연되나, 일반적으로 도움 없이 걸을 수 있는 능력을 달성한다. 시간이 지남에 따라, 이들은 등의 경직, 관절 구축, 척추측만증 및 호흡 문제가 발생할 수 있다. 그러나 영향을 받은 대부분의 개체는 걷고 계단을 오를 수 있는 능력을 유지한다.
- [0428] 본 개시내용은 또한 환자에게 치료학적 유효량의 상기에 기재된 바와 같은 조성물, 조작된 세포 또는 약제학적 조성물을 투여하는 것을 포함하는, 본 개시내용에 따른 *LAMA2*-관련 근이영양증을 치료하는 방법을 제공한다.
- [0429] "치료학적 유효량"은 원하는 치료 결과를 달성하고 쇠약 및 골격근의 위축과 같은 *LAMA2*-관련 근이영양증의 적어도 하나 이상의 징후 또는 증상을 예방, 지연 또는 역전시키는 데 필요한 투여량 및 기간 동안 효과적인 양을 지칭한다. 치료학적 유효량의 본 발명의 생성물 또는 이를 포함하는 약제학적 조성물은 개체의 질환 상태, 연령, 성별 및 체중, 및 생성물 또는 약제학적 조성물이 개체에서 원하는 반응을 도출하는 능력과 같은 요인에 따라 달라질 수 있다. 투여 요법은 최적의 치료 반응을 제공하도록 조정될 수 있다. 치료학적 유효량은 또한 전형적으로 생성물 또는 약제학적 조성물의 임의의 독성 또는 해로운 효과가 치료적으로 유의한 효과보다 더 중요한 양이다.
- [0430] 본원에 사용된 바와 같이, 용어 "**환자**" 또는 "**개체**"는 포유동물을 나타낸다. 바람직하게는, 본 개시내용에 따

른 환자 또는 개체는 인간이다.

- [0431] 본 개시내용과 관련하여, 본원에 사용된 바와 같은 용어 "**치료하는**" 또는 "**치료**"는 이러한 용어가 적용되는 *LAMA2*-관련 근이영양증 또는 병태의 진행을 역전, 완화 또는 억제하는 것을 의미하거나, 그러한 용어가 적용되는 장애 또는 병태의 하나 이상의 증상의 진행을 역전, 완화 또는 억제하는 것, 특히 골격근의 쇠약 및 위축의 감소 및/또는 골격근 기능의 개선을 의미한다.
- [0432] 본 발명의 약제학적 조성물은 일반적으로 환자에서 치료 효과를 유도하기에 효과적인 투여량 및 기간 동안 공지된 절차에 따라 투여된다.
- [0433] 투여는 전신적 또는 국소적일 수 있다. 전신 투여는 바람직하게는 비경구 예컨대 피하 (SC), 근육내 (IM), 혈관내 예컨대 정맥내 (IV) 또는 동맥내; 복강내 (IP); 피내 (ID), 간질 등이다. 투여는 예를 들어 주사 또는 관류에 의한 것일 수 있다. 일부 바람직한 실시양태에서, 투여는 비경구, 바람직하게는 혈관내 예컨대 정맥내 (IV) 또는 동맥내이다. 본 발명의 실시는 달리 나타나지 않는 한 관련 기술분야의 기술 범위 내에 있는 종래의 기술을 이용할 것이다. 이러한 기술은 문헌에 충분히 설명되어 있다.
- [0434] 또한, 본원에 기재된 바와 같이 개시된 방법을 실시하기 위한 키트가 본 개시내용에 의해 제공된다. 키트는 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서) 또는 이를 코딩하는 핵산 작제물, 및 본원에 기재된 바와 같은 *LAMA2* 트랜스진을 포함하는 핵산 작제물을 함유할 수 있다. 일부 측면에서, 키트는 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)을 코딩하는 핵산 작제물 및 본원에 기재된 바와 같은 *LAMA2* 트랜스진을 포함하거나 이로 이루어진 핵산 작제물을 함유하는 렌티바이러스 입자를 함유할 수 있다.
- [0435] 대상 키트는 대상 방법을 실시하기 위해 키트의 구성요소를 사용하기 위한 설명서를 추가로 포함할 수 있다. 대상 방법을 실행하기 위한 설명서는 일반적으로 적합한 기록 매체에 기록된다. 예를 들어, 설명서는 종이 또는 플라스틱 등과 같은 기재 상에 인쇄될 수 있다. 이와 같이, 설명서는 패키지 삽입물로서 키트에, 키트 또는 이의 구성요소의 용기의 라벨링 등 (즉, 패키징 또는 서브패키징과 연관됨) 등에 존재할 수 있다. 다른 실시양태에서, 설명서는 적합한 컴퓨터 판독가능 저장 매체에 존재하는 전자 저장 데이터 파일, 예를 들어, CD-ROM, 디스크 등으로서 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 기판에 기록된다.
- [0436] 본 개시내용은 전형적으로 다음을 포함하는 조성물 또는 키트에 관한 것이다:
- [0437] (i) 본원에 정의된 바와 같은 제1 단백질, 또는 상기 제1 단백질을 코딩하는 핵산,
- [0438] (ii) 본원에 정의된 바와 같은 제2 단백질, 또는 상기 제2 단백질을 코딩하는 핵산, 및
- [0439] (iii) *LAMA2* 트랜스진을 포함하거나 이로 이루어진 핵산 작제물
- [0440] 본 개시내용은 또한 다음을 포함하는 조성물 또는 키트에 관한 것이다:
- [0441] (i) 본원에 정의된 바와 같은 융합 단백질, 또는 상기 융합 단백질을 코딩하는 핵산,
- [0442] (ii) *LAMA2* 트랜스진을 포함하거나 이로 이루어진 핵산 작제물
- [0443] 을 포함하는 제1 조성물.
- [0444] 한 실시양태에서, 조성물 또는 키트는 미니씨클, 플라스미드 또는 바이러스 벡터, 특히 비통합 바이러스 벡터, 예를 들어 비통합 렌티바이러스 벡터에 외인성 핵산을 포함한다. 한 실시양태에서, 본원에 개시된 바와 같은 조성물 또는 키트는 나노입자에 포함된다.
- [0445] 한 실시양태에서, 제1 및/또는 제2 단백질을 코딩하는 핵산 작제물 (별도로 또는 상기에 기재된 융합 단백질의 일부로서)은 RNA, DNA 또는 단백질의 형태이고, *LAMA2* 트랜스진을 코딩하는 폴리뉴클레오타이드 서열은 전달 방법에 따라 RNA 또는 DNA의 형태이다. 특히, 외인성 핵산을 코딩하는 폴리뉴클레오타이드 서열은 RNA 형태이다.
- [0446] 한 실시양태에서, 조성물 또는 키트는 바이러스가 없고 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.

- [0447] 한 실시양태에서, 조성물 또는 키트는 (a) RNA 형태로 별도로 또는 상기에 기재된 융합 단백질의 일부로서 (예를 들어, Cas9 및 트랜스포사제를 포함함) 제1 및 제2 단백질을 코딩하는 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 *LAMA2* 트랜스진을 포함하거나 이로 이루어진 폴리뉴클레오티드를 포함한다.
- [0448] 한 실시양태에서, 조성물은 (a) 단백질 형태로 본원에 기재된 (예를 들어, Cas9 및 트랜스포사제를 포함함) 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), (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 *LAMA2* 트랜스진을 포함하거나 이로 이루어진 폴리뉴클레오티드를 포함한다.
- [0449] 한 실시양태에서, 조성물은 (a) DNA 형태로 별도로 또는 상기에 기재된 융합 단백질의 일부로서 (예를 들어, Cas9 및 트랜스포사제를 포함함) 제1 및 제2 단백질을 코딩하는 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 *LAMA2* 트랜스진을 포함하거나 이로 이루어진 폴리뉴클레오티드를 포함한다.
- [0450] 한 실시양태에서, 조성물은 (a) 단백질 형태로 본원에 기재된 (예를 들어, Cas9 및 트랜스포사제를 포함함) 제1 및 제2 단백질 (별도로 또는 상기에 기재된 융합 단백질의 일부로서), (b) 필요한 경우 가이드 RNA (예를 들어, 융합 단백질과 복합체를 형성하는 별도의 RNA 분자로서), 및 (c) 패키징 벡터에 함유되거나 결합된 삽입용 *LAMA2* 트랜스진을 포함하거나 이로 이루어진 폴리뉴클레오티드를 포함한다. 특정한 실시양태에서, 패키징 벡터는 렌티바이러스 입자이다. 일부 실시양태에서, 제1 및 제2 단백질은 별도로 또는 상기에 기재된 융합 단백질의 일부로서 gag-pol 또는 VPR (바이러스 단백질 R)에 의해 렌티바이러스 캡시드에 결합된다. 일부 실시양태에서, *LAMA2* 트랜스진을 포함하거나 이로 이루어진 폴리뉴클레오티드는 인테그라제의 페이로드로서 RNA 형태이다.
- [0451] 한 실시양태에서, 제1 단백질이 ZFP인 경우 가이드 RNA가 필요하지 않을 수 있다.
- [0452] 본 개시내용은 전형적으로 키트로서,
- [0453] (i) 별도로 또는 상기에 기재된 융합 단백질의 일부로서의 제1 및 제2 단백질, 또는 이를 코딩하는 핵산, 그리고 여기서 상기 제1 및 제2 단백질은 제1 가이드된 RNA Nickase Cas9의 아미노산 서열, 전형적으로 서열번호: 65를 가진 SpCas9 Nickase 및 변형된 과활성 PiggyBac를 포함함, 및,
- [0454] (ii) 제1 가이드된 RNA 핵산
- [0455] 을 포함하는 제1 조성물,
- [0456] (iii) 별도로 또는 상기에 기재된 융합 단백질의 일부로서의 제1 및 제2 단백질, 또는 이를 코딩하는 핵산, 그리고 여기서 상기 제1 및 제2 단백질은 제2 가이드된 RNA Nickase Cas9의 아미노산 서열, 전형적으로 서열번호: 66을 가진 SpCas9 Nickase 및 변형된 과활성 Piggybac를 포함함,
- [0457] (iv) 제2 가이드된 RNA 핵산
- [0458] 을 포함하는 제2 조성물,
- [0459] 및 *LAMA2* 트랜스진을 포함하는 핵산 작제물
- [0460] 을 포함하며;
- [0461] 여기서 제1 및 제2 단백질 각각 (제1 및 제2 조성물 각각 내)은 이중이량체화할 수 있고, 게놈 DNA 영역의 인접 부위에서 제1 및 제2 가이드된 RNA에 의해 결정된 이중 컷트를 만들 수 있고, 임의로 인접 부위 사이에 상기 핵산을 삽입할 수 있는, 키트에 관한 것이다.
- [0462] 일부 추가 실시양태가 이후에 개시된다.
- [0463] **E1:** a) 바람직하게는 알부민의 인트론 1; Lama2 또는 Rosa 26 유전자좌의 인트론 1 내에 포함된, 표적 핵산 서열에 결합할 수 있는 부위-특이적 DNA 결합 단백질로 이루어진 제1 단백질 및 트랜스포사제로 이루어진 제2 단백질을 포함하는 융합 단백질; 또는 핵산 작제물, 바람직하게는 상기 융합 단백질을 코딩하는 mRNA, b) 라미닌- $\alpha 2$ 단백질, 기능적 변이체 또는 단편을 코딩하는 트랜스진을 포함하는 핵산 작제물을 포함하는 조성물.

- [0464] E2: 트랜스진은 서열번호: 74의 라미닌- α 2 단백질을 코딩한다.
- [0465] E3: b)의 핵산 작제물은 서열번호: 76의 CMV 프로모터, 서열번호: 77의 CAG 프로모터, 서열번호: 78의 EF1-알파 프로모터, 서열번호: 79의 SV-40 프로모터 및 서열번호: 80의 EalbAAT 프로모터 또는 서열번호: 81의 스플라이스 수용체로 이루어진 군으로부터 선택된 프로모터를 포함한다.
- [0466] E4: b)의 핵산 작제물은 바람직하게는 서열번호: 83-85로 이루어진 군으로부터 선택된, 폴리(A) 신호 서열을 포함한다.
- [0467] E5: b)의 핵산 작제물은 바람직하게는 서열번호: 86-87로 이루어진 군으로부터 선택된, 절연체 요소를 또한 포함할 수 있다.
- [0468] E6: b)의 핵산 작제물은 역 말단 반복부 (ITR)에 의해, 바람직하게는 서열번호: 88 및 89의 5'-ITR 및 3'-ITR에 의해 플랭킹된다.
- [0469] E7: b)의 핵산 작제물은 플라스미드 벡터, 미니썬 벡터, 도기 본 DNA 공여체 벡터, 렌티바이러스 벡터 및 레트로바이러스 벡터로 이루어진 군으로부터 선택된 벡터에 포함될 수 있다.
- [0470] E8: 부위-특이적 DNA 결합 단백질은 Cas 단백질을 포함하는 RNA-가이드된 뉴클레아제이고 조성물은 세포 게놈의 특이적 부위에 상기 LAMA2 트랜스진을 통합하기 위한 표적 핵산 서열에 대한 상보적 서열을 포함하는 가이드 RNA를 추가로 포함하며, 바람직하게는 상기 Cas 단백질은 에스. 피오게네스 Cas 9 단백질이다.
- [0471] E9: 가이드 RNA는 서열번호: 90-97로 이루어진 군으로부터 선택된 서열의 상보적인 서열을 포함한다.
- [0472] E10: 트랜스포사제는 변형된 과활성 Piggybac 트랜스포사제 또는 슬리피 뷰티 트랜스포사제, 바람직하게는 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 PiggyBac 트랜스포사제.
- [0473] E11: 과활성 PiggyBac 트랜스포사제는 M194, R245, G325, R372, K375, R376, E377, E380, D450 및 S573으로 이루어진 군으로부터 선택된 아미노산의 적어도 하나의 돌연변이를 포함하며 바람직하게는 아미노산 D450, R372 및 K375의 돌연변이를 포함하는 변형된 과활성 PiggyBac 트랜스포사제이며, 상기 위치 번호는 서열번호: 9의 변형되지 않은 과활성 Piggybac의 아미노산 번호에 상응한다.
- [0474] E12: 트랜스포사제는 바람직하게는 링커, 바람직하게는 GGS, XTEN 또는 FOKI를 포함하는 펩티드 링커, 보다 바람직하게는 서열번호: 53의 XTEN에 의해 상기 부위-특이적 DNA 결합 단백질에 N-말단으로 융합된다.
- [0475] E13: 조성물이 나노입자 내에 패키징된다.
- [0476] E14: 상기에 기재된 바와 같은 조성물을 세포 내로 도입하는 것을 포함하는, LAMA2 트랜스진을 세포의 게놈 내의 표적 핵산 서열로 통합하기 위한 시험관내 방법 및 이의 게놈 내에 통합된 LAMA2 트랜스진을 포함하는 상기 방법에 의해 수득가능한 조작된 세포.
- [0477] E15: 상기에 기재된 바와 같은 조성물 또는 조작된 세포를, 임의로 하나 이상의 약제학적으로 허용되는 부형제와 조합하여 포함하는 약제학적 조성물.
- [0478] E16: 이를 필요로 하는 대상체에서 요법에 사용하기 위한, 특히 메로신-결핍 선천성 근이영양증 1A형 (MDC1A)의 치료에 사용하기 위한 조성물, 조작된 세포 또는 약제학적 조성물.
- [0479] 본 발명은 이제 첨부된 도면을 참조하여, 제한적이지 않은 하기 실시예로 예시될 것이다.

도면의 간단한 설명

- [0480] 도 1: MCD1A의 Lama2 동물 모델에서 치료 유전자 교체를 위해 생산된 상이한 Lama2 페이로드의 표시.
- 도 2: GFP 또는 Lama2 로드된 렌티바이러스로 감염된 Hek293T 세포 상의 바이러스 카피 수. 카피 수의 qPCR 추정치를 기반으로 하여 역가를 계산하였다. Lama2로 인한 증가된 카피 크기는 바이러스 생산 효율을 백배 감소시킨다.
- 도 3: 순환 백혈구(A) 및 침윤된 근육 세포 (B)의 FACS 분석; 이식 4주 후 마우스로부터 단리 후. 빈 렌티바이러스 형질도입과 GFP 로드된 바이러스의 비교가 표시된다.

도 4: MCD1A 모델을 치료하기 위해 생체외 렌티바이러스 매개 Lama2 유전자를 골수 유래 세포로 전달하기 위한 실험 설계 및 워크플로우.

도 5: Lama2 발현 골수 유래 세포로 이식한 후 6주 및 12주 후에 처리 및 비처리 MCD1A 모델 대 야생형 동물의 상대적 강도 (A). 동일한 동물에서 처리 후 근력 증가 (B). 두 경우 모두에서, 결과는 그립 강도(grip strength) 검정의 분석을 나타냈다.

도 6: 시간 경과에 따른 처리 및 비처리 MCD1A 모델 대 야생형 동물에 대한 카플란 마이어(Kaplan Meier) 생존 곡선.

도 7: MCD1A 모델을 치료하기 위해 간 세포로의 생체내, 무바이러스(viral free), 트랜스포사제 매개 Lama2 유전자 전달을 위한 실험 설계 및 워크플로우.

도 8: 트랜스진 카피 수 (A; C) 및 트랜스진 발현 (B; D); 간 세포에서 처리 4주 후 각각 qPCR 및 RT-qPCR에 의해 측정됨. 과활성 PiggyBac에 의해 안정적으로 통합된 RFP 유전자 리포터의 데이터 (A; B); 및 조작된 PiggyBac에 융합된 프로그래밍 가능한 트랜스포사제 Cas9에 의해 안정적으로 통합된 Lama2 유전자 (C; D)가 표시된다.

도 9: 안정적인 발현을 위한 삽입 기계의 유무에 관계없이 공동-형질감염된, MCD1A 모델을 Lama2로 처리한 후 1주 및 4주 후 순환 혈청에서 라미닌의 존재에 대한 Elisa 검정.

도 10: (A) RFP 트랜스포존 단독 (에피솜(episomal)) 또는 생체내 JetPEI 시약과 함께 전달된 hyPB 또는 FICAT R372A_K375A_D450N mRNA와 함께 마우스 게놈에서 Rosa26 세이프 하버를 표적화하는 데 사용되었다. 간에서 RFP 트랜스진의 상대 카피 수를 반정량적 qPCR에 의해 측정하고 Tfrc 유전자 (이배체 게놈)의 상대적 이중 카피로 정규화하였다. (B) 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. (C) 간 게놈 DNA에서 트랜스포존 3' ITR과 Rosa26 유전자좌 사이의 접합 PCR(Junction PCR). 마우스에 FiCAT R372A_K375A_D450N 플라스미드 DNA 또는 mRNA, Rosa26 유전자좌를 표적으로 하는 gRNA 및 미니썬클 트랜스포존 GFP 페이로드를 유체역학적으로 주입하고 주입 후 5주 후에 희생시켰다. 게놈+ 가닥 통합을 증폭시키는 PCR을 수행하였다. n = 2-3 동물/조건, 숫자는 상이한 개체에 상응한다. FiCAT mRNA 또는 pDNA로 처리된 마우스의 66%가 표적화 삽입을 나타낸다. FICAT에서 검출된 밴드(band)의 크기는 증폭된 삽입의 예상 크기에 상응한다. 배경으로 간주되는 에피솜 샘플에서 더 높은 크기의 밴드가 검출된다.

도 11: (a) C2C12 세포를 RFP 트랜스포존 단독 (에피솜)으로 형질도입하거나, FiCAT R372A_K375A_D450N 및 Lama2 유전자를 표적으로 하는 gRNA (스페이서 271.1)와 조합하여, 형질도입 후 2주 동안 RFP 양성 세포를 모니터링하였다. n = 2 기술 복제의 평균 $\pm$ SD가 플롯되었다. n = 3의 대표 이미지. (b) c2c12 게놈에서 검출된 삽입을 나타내는 카리오플롯(Karyoplot). (c) 3' ITR과 Lama2 유전자좌 사이의 접합 PCR을 FiCAT (3, 4)와 에피솜(1, 2) 처리된 농축화(enriched) 집단을 비교하여 + 가닥 (1, 3) 및 - 가닥 (2, 4) 페이로드 삽입에서 표시된다 (아래쪽 패널). n = 3의 대표 이미지. (d) 온-타겟(on-target) 접합부 (Lama2 부위)에서의 커버리지(Coverage).

도 12: 300 bp ITR 간 크기의 삽입과 관련하여, 리포터 세포주의 AAVS1 부위에 1/2 GFP 트랜스포존의 표적 삽입. 105 bp, 200 bp 및 300 bp의 ITR 간 크기는 과활성 PiggyBac (hyPB) R372A_K375A_D450N (Cas9-PBx3) hyPB R372A_K375A_D450N (Cas9-PBx3-PBx3)의 이량체에 융합된 Cas9와 함께 1/2 GFP의 삽입에 대해 시험되었다.

도 13. 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 14. 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 15. (a) CasX (좌측) 및 Cpf1 (중간)에 의한 편집 활동. (b) SaCas9 (좌측), CjCas9 (중간)에 의한 편집 활동. indels $\pm$ SD를 가진 판독의 평균 %는 2개의 기술적 반복, N=3 생물학적 복제의 대표 이미지에 대해 표

시된다. TRAC-1 부위를 표적으로 하는 SpCas9는 참조용으로 사용되었다 (우측).

도 16. 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 17. 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 18. 선택 주기에 따라 온-타겟 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 19. (A) 상위로 선택된 후보의 온-타겟 효율. 마지막 주기로부터 선택된 96개의 무작위 클론 중 가장 높은 온-타겟 활성을 기준으로 하여 6개의 개별 후보가 선택되었다. 개별 온-타겟 활성을 Cas9-hyPB R372A-K375A-D450N과 비교하였다. (B) 상위 온-타겟 활성 변이체에서 우세한 PB 잔기를 나타내는 로고.

도 20. Cas9 및 단일 gRNA (gRNA-TCR1 또는 AAVS1-3)에 의해 또는 니카제 Cas9 및 인접 위치 (gRNA-TCR1 및 AAVS1-3)에서 표적화하는 2개의 gRNA, 및 변형된 hyPB (돌연변이체 R372A-K375A-D450N)에 대한 융합으로 프로그래밍 가능한 DNA 결합 도메인 (ZnF)에 의한 이중 가닥 파단(break)은 표적 삽입을 발생시킨다. 효율적인 온-타겟된 삽입을 위해 삽입 부위에서 이중 가닥 파단 및 PB의 공동-국부화(Co-localization)가 필요하다. 이것은 뉴클레아제 Cas9 또는 니카제 Cas9에 의한 이중 커트에 의해 달성될 수 있다.

도 21. 1/2 GFP 리포터 세포주에서 표적 삽입을 위한 SpCas9 또는 SaCas9와 융합된 이량체 hyPB R372A-K375A-D450N의 프로그래밍 가능한 삽입 활성.

도 22. 선택 주기에 따른 온-타겟 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 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 25. 1/2 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 26. 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-발현을 가진 1/2 GFP 리포터 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 상대적 비교 (R372A-K375A-D450N; R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L).

도 27. SpCas 및 hyPB R372A-K375A-D450N을 포함하는 융합 단백질, 및 3 hyPB 돌연변이체 R372A-K375A-D450N;

R202K-R275A-N347S-R372A-D450N-T560A-F594L; 및 R275A-N347S-R372A-D450N-T560A-F594L의 공동-발현을 가진 ½ GFP 리porter 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 상대적 비교.

도 28. hyPB R272A-K275A-D450N의 이량체에 융합된 SpCas9 (좌측)와 제1 hyPB R272A-K275A-D450N에 및 제2 hyPB 돌연변이체 (우측)에 융합된 SpCas9 간의 ½ GFP 리porter 세포주에서 표적 삽입을 위한 프로그래밍 가능한 삽입 활성의 비교.

발명을 실시하기 위한 구체적인 내용

실시예

1. *LAMA2* 페이로드 아키텍처(architecture)

머스 머스쿨러스(*Mus musculus*) (서열번호: 94) 또는 호모 사피엔스(*Homo sapiens*) (서열번호: 75)에 대한, 외수송 (exporation)을 위한 리더 서열을 포함한 *LAMA2*-코딩 서열이 사용될 수 있다. *LAMA2*-코딩 서열은 우리가 유전자 편집을 수행하는 유기체에 존재하는 내인성 유전자와 구별하기 위해, 유전자 전달 후 페이로드 특이적 검출 및 추적(tracking)을 위한 서열의 일부에 대한 서열 리코딩(recording)을 함유할 수 있다.

이 *LAMA2* 유전자는 구성적 프로모터 (서열번호: 76-79), 조직 특이적 프로모터 (서열번호: 80), 또는 스플라이스 수용체 (서열번호: 81)의 하류에 클로닝된다. 스플라이스 수용체를 가진 버전은 능동적으로 발현되는 유전자의 제1 인트론에 통합되어 내인성 발현 제어를 가질 수 있다.

이 코딩 서열은 또한 페이로드의 미니-유전자(mini-gene) 버전을 위한 합성 인트론 (서열번호: 82)을 가질 수 있다. 발현 카세트는 폴리-A 신호 (서열번호: 83-85), 및 절연체 요소 (서열번호: 86-87)를 함유할 수 있다.

페이로드는 다음과 같은 형식일 수 있다:

1. 플라스미드 벡터; 박테리아 복제 기원 및 항생제 내성 함유

2. 플라스미드 벡터 성분에 더하여 32개의 *SceI* 제한 효소 표적 부위 및 *Lama2* 발현 카세트를 플랭킹하는 2개의 *AttB* 및 *AttP* 재조합 부위를 또한 함유하는 미니썬클 벡터.

3. 선형화 및 공여체 DNA의 카고 말단 보호를 위해, 플라스미드 벡터 성분에 더하여 *Lama2* 발현 카세트를 플랭킹하는 *TeIN* 표적 부위를 또한 함유하는 도기 본 DNA 공여체 벡터,

4. 렌티바이러스 또는 감마 레트로바이러스 페이로드 벡터, 치료 카고를 위한 추가 공간을 허용하도록 크기를 줄이도록 조작될 수 있는 렌티바이러스/레트로바이러스 작동 요소에 더하여 *Lama2* 발현 카세트를 플랭킹하는 긴 말단 반복부를 가짐.

이들 모든 경우에 *Lama2* 발현 카세트는 트랜스포사제 인식을 위한 역 말단 반복부 (ITR)에 의해 플랭킹될 수 있다 (Piggybac 및 슬리핑 뷰티).

도 1은 *Lama2* 트랜스진을 함유하는 상이한 카고의 표시를 나타낸다.

본 발명자들은 페이로드의 ITR 사이의 크기가 융합 단백질의 결합 능력에 영향을 미치고, 따라서 계놈으로의 이의 결과적인 전달에 영향을 미치는지 추가로 시험하였다. 세 가지 ITR 간 크기를 페이로드 효율의 표적 삽입을 위해 시도하였다: FiCAT (간기 R372A_K375A_D450N에서 돌연변이된 Cas9 + 과활성 PiggyBac 트랜스포사제) 및 FiCAT 이량체 (간기 R372A_K375A_D450N에서 돌연변이된 Cas9 + 과활성 PiggyBac 트랜스포사제의 이량체) 및 리porter 세포주에서 AAVS1 부위를 표적화하는 gRNA (도 12)를 가진 ½ GFP 트랜스포존의 미니썬클 (MC) 페이로드에서 대략 300 bp, 200 bp 및 105 bp. 최상의 ITR 간 크기는 300 bp 및 200 bp이었으며, 한편 105 bp의 표적 통합 효율은 더 낮았다.

렌티바이러스 및 레트로바이러스-기반 전달에 더하여, 상이한 *Lama2* 페이로드는 폴리플렉스 및 지질 나노입자를 사용하여 전달되거나 전기천공법 또는 미세주입에 의해 네이키드(naked) 핵산으로서 표적 세포에 전달될 수 있다.

2. *Lama2* 페이로드의 생체의 형질도입

머스 머스쿨러스의 *Lama2* 코딩 서열을 함유하는 렌티바이러스 입자는 CMV 프로모터의 제어 하에 2세대 렌티바이러스 생산을 위한 표준 프로토콜을 사용하여 생산되었다. 이 페이로드의 크기가 큰 것을 감안하여, 입자가 농축되었고, 바이러스 역가는 Hek293T 세포를 감염시킨 후 바이러스 카피 수 qPCR 추정에 의해 계산되었다 (도

2).

[0497] 생산 효율은 리포터 유전자 GFP와 비교하여 두 배 낮았긴 하지만, 이 역가는 Lama2 마우스 모델에서 세포 요법에 있어서 사용할 수 있는 세포의 생체외 형질도입에 충분하였다. 두 가지 상이한 접근법을 취하였다. 단리된 근육 줄기 세포를 수령자(수령자) 모델의 근육에 형질도입하고 이식하였다. 대안적으로, 골수 세포로부터 유래된 조혈 전구체는 Lama2 발현 카세트를 함유하는 렌티바이러스로 형질도입되었고, 이어서 이들 전구체는 정맥 주사를 통해 컨디셔닝된 모델로 이식될 수 있다.

[0498] 생체외 치료에 대한 개념의 증거로서, 본 발명자들은 골수 유래 조혈 전구체를 단리하고 이들을 컨디셔닝된 마우스에 이식하기 전에 먼저 GFP 발현 렌티바이러스로 형질도입하였다. 이어서 본 발명자들은 순환 혈액에서 백혈구의 GFP 발현을 조사하여, 치료 4주 후에 생착(engraftment)을 점검하였고, 양성 결과를 발견하였다 (도 3A). 또한, 염증이 생긴 근육으로의 순환 세포의 침윤이 또한 점검되었고 양성 결과가 발견되었다 (도 3B).

[0499] 생체외 형질도입된 조혈 전구체의 양성 침윤 및 페이로드 발현을 감안하여, 본 발명자들은 이어서 렌티바이러스 입자를 발현하는 Lama2로 이 실험을 반복하였다 (도 4). 마우스는 그림 검정에 의해 측정된 치료 후 12주 후에 근력을 회복하였다 (도 5A 및 B). 또한, 처리된 Lama2 동물 모델의 생존이 또한 관찰되었으며 (도 6), 이는 MDC1A 환자를 치료하기 위한 생체외 세포 형질도입 및 이식의 치료적 가치를 나타낸다.

[0500] 또 다른 라인의 실험에서, 본 발명자들은 ~20%의 효율로 C2C12 뮤린 근모세포 세포주에 RFP 리포터의 LAMA2-표적화 삽입을 수행하였다 (도 11). 접합 PCR 및 (STAT)-PCR을 사용하여 온-타겟 및 오프-타겟(off-target) 효율을 측정하였다. 이 개선된 효율은 FiCAT로 지칭되는 R372A_K375A_D450N 잔기에서 돌연변이된 Cas9 및 과활성 PiggyBac 트랜스포사제를 포함하는 융합 단백질을 사용한 결과였다.

[0501] 3. Lama2 페이로드의 생체내 형질도입

[0502] 간(liver)이 생체내에서 단백질 생산을 위한 생물반응기로서 기능하는 능력을 감안하여; 본 발명자들은 트랜스포사제를 기반으로 하는 무바이러스 유전자 전달 접근법을 사용하여 Lama2 발현 벡터로 동물 모델의 생체내 간 세포를 형질도입할 수 있는 가능성을 탐구하였다.

[0503] 생체내 JetPei (폴리플러스(PolyPlus)) (도 7)를 사용하여 제형화된 플라스미드 벡터 또는 mRNA 분자 및 트랜스포존 Lama2 플라스미드 벡터로서 프로그래밍 가능한 뉴클레아제 단백질 (SpCas9)을 가진 과활성 PiggyBac 트랜스포사제 또는 조작된 과활성 Piggybac 융합 트랜스포사제를 동물 모델에 정맥내 주사하고, 조직은 간에서 페이로드 카피 수 (개념 증명으로서의 RFP 데이터) (도 8A,C), 및 간에서의 발현 분석 (RFP, 도 8C; Lama2, 도 8D) 및 Lama2의 경우 순환 혈액 (도 9)을 위해 3 내지 4주 후에 수확하였다. 주사 1주 후 100 μ L의 혈액을 수집하여 종점(end-point)에서 전혈을 수집하였다.

[0504] 처리 4주 후 간 세포의 트랜스진 카피 수는 긍정적인 결과를 나타냈으며 RFP 리포터 유전자와 추가로 비교하였으며, 이는 상이한 크기의 페이로드로 유사한 형질도입 효율을 확인해 주었다 (도 8A, B). 에피솜 전달에 대한 참조는 어떤 트랜스포사제도 공동-전달되지 않았을 때 수행되었다. RFP 리포터 및 Lama2 유전자에 대한 트랜스진의 발현이 간세포에서 또한 확인되었으며 (도 8B, D), Lama2 발현은 전달 후 1주 및 4주 후에 혈액에서 검출되었다 (도 9).

[0505] 4. 생체내 유전자 전달을 위한 개선된 도구

[0506] 생체내 유전자 전달에 사용되는 프로그램 가능한 트랜스포사제를 개선하기 위해, 본 발명자들은 FiCAT (Cas9-hyPB R372A_K375A_D450N)에 플라스미드 DNA (pDNA)와 mRNA를 둘 다 사용하여, 플라스미드 또는 MC 형태로 Rosa26 게놈 세이프 하버 (개념 증명을 위한 트랜스포존을 코딩하는 RFP 또는 루시페라제)를 표적으로 하는 마우스 간으로 이를 전달하였다. 내인성 유전자 TFRC (도 10A)와 비교하여 트랜스진의 높은 카피수가 관찰되었고 시간이 지남에 따라 트랜스진 발현이 유지되었다 (도 10B). 3' ITR과 게놈 유전자와 사이의 접합부의 PCR을 사용하여 새로 형성된 온-타겟 삽입을 측정하였다 (도 10C).

[0507] 4.1. hyPB에 융합된 Cas 변이체의 결과

[0508] 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 머딜(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(Hershey)

세포주 형질감염을 위해 클로닝하였다. 표적 전위는 GFP 발현에 의해 측정되었다 (도 13).

[0509] 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 14). 사용된 상이한 Cas 단백질 및 일루미나(Illumina) NGS (도 15)에 의해 각각의 단백질에 대해 설계된 세 가지 상이한 gRNA에 대해 Indels이 결정되었으며, 정규화 목적으로 표시되었다.

[0510] 이들 긍정적인 결과는 임의의 서열 특이적 뉴클레아제 모듈에 유용하도록 프로그래밍 가능한 전위를 위해 조작된 hyPB를 검증한다.

[0511] 4.2. PB 돌연변이체의 이량체에 융합된 Cas9의 추가 결과

[0512] 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 16). Cas9에 대한 이량체 융합을 위해 우리는 클로닝 및 발현을 용이하게 하기 위해 기록된 버전의 hyPB R372A-K375A-D450N 돌연변이체를 사용하였다.

[0513] 흥미롭게도, 이량체 hyPB R372A-K375A-D450N에 융합된 Cas9가 SpCas9 대신에 SaCas9인 경우, 활성이 추가로 증가된다 (도 21). 이량체 hyPB를 사용하여 SpCas9에 비해 SaCas9의 증가된 성능은 단량체 hyPB로 획득된 결과와 일치한다 (도 13 및 14).

[0514] 4.3. ZNF-PB 돌연변이체 구제(rescue)의 결과

[0515] 우리는 오프 타겟 부위에서 비유도 DSB에 의한 오프 타겟 활성을 저하시키면서, 표적 통합을 용이하게 하는데 SpCas9 Nickase 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 Nickase Cas9에 의한 2개의 현장(on-site) 단일 가닥 브레이크(brake), 또는 Cas9 뉴클레아제에 의한 단일 DSB로 이를 보완하였다.

[0516] Znf-PB 융합은 단일 또는 이중 gRNA 가이드-cas9 뉴클레아제 또는 Nickase와 함께 Znf 결합 부위 근처에 DSB를 도입하는 것과 조합될 때 구제된 표적 삽입 활성을 전혀 나타내지 않거나 매우 낮게 나타냈다 (도 20).

[0517] 4.4. Cas9-hyPB 돌연변이 변이체의 결과

[0518] 더 양호한 효율로 프로그래밍 가능한 전위를 수행할 수 있는 돌연변이체 조합을 추가로 탐구하기 위해, 분할 GFP 리porter 시스템의 프로그래밍 가능한 삽입에 의해 GFP가 재구성되는 세포 선택의 수개의 주기를 수행하였다. 흥미롭게도 우리는 Cas9-hyPB R372A-K375A-D450N을 능가하는 수개의 조합을 관찰하였다 (도 17). 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.

[0519] 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 리porter 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.

[0520] 먼저, 96개의 변이체의 무작위 선택을 수행하고 최상의 성능의 변이체를 별도로 스크리닝하였다 (도 18). 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 19B). 최상의 온-타겟 효율을 가진 6개의 PB 변이체를 선택하였다 (도 19A). 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)).

[0521] 이 실험을 반복하여 확인하였다 (도 22A). 우리는 또한 각각의 주기의 벌크 변이체를 발현하는 렌티바이러스 및 PB 변이체 CN에 의해 이의 역가를 교정하는 감염된 리포터 세포주를 생산하여, 주기에 걸쳐 유사한 온-타깃 효율 증가를 입증하였다 (도 22B). 단일 돌연변이체는 cas9_PB 라이브러리 농축화의 4주기 및 5주기 후에 벌크 변이체로부터 단리되었다. 온-타깃 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-타깃 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-R375A-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.

[0522] 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 28).

[0523] 4.5. Cas9 및 hyPB 비공유 연결의 결과

[0524] 링커를 통해 Cas9와 hyPB R372A-R375A-D450N의 공유 결합에 더하여, 우리는 MS2-MCP 시스템을 사용하여 MCP 단백질과 hyPB R372A-R375A-D450N으로 이루어진 융합 단백질과 Cas9를 MCP 단백질에 결합하는 MS2 서열의 테트라루프를 함유하는 변형된 gRNA를 통해 연결하였다.

[0525] MCP-hyPB R372A-R375A-D450N 융합 단백질과 Cas9의 조합은 Cas9-hyPB R372A-R375A-D450N 융합 단백질과 비교하여 프로그램 가능한 삽입 활성이 증가하였다 (도 24A). 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-R375A-D450N을 능가하였다 (도 24B).

[0526] 4.6. 프로그래밍 가능한 전위를 위한 Cas9 및 디커플링(decoupling)의 결과

[0527] 우리는 또한, 링커 없이, 또는 MS2-MCP 시스템도 없이 hyPB R372A-R375A-D450N으로 SpCas9의 성능을 시험하였다. 우리는 동일한 세포 SpCas9 및 hyPB R372A-R375A-D450N에서 공동-발현시켰고, Cas9-hyPB R372A-R375A-D450N 융합 단백질과 비교하여 증가된 프로그래밍 가능한 삽입 활성이 등록되었다 (도 25A). 우리는 Cas9에 융합되지 않았으나, 동시에 발현되고, 프로그래밍 가능한 전위의 활성을 달성하기 위해 함께 작용하는 시험된 hyPB 돌연변이체 변이체의 수를 확장하였다 (도 25B).

[0528] 4.7. Cas9-hyPB 및 MCP-hyPB 융합 단백질의 공동-발현의 결과

[0529] 우리는 비공유결합으로 연결된 단량체 중 하나와의 융합의 이량체 버전을 수득하기 위해, MCP 단백질에 융합된 hyPB R372A-R375A-D450N 돌연변이체 및 SpCas9에 융합된 hyPB 돌연변이체를 공동-형질감염시켰다. SpCas9에 융합된 수개의 hyPB 돌연변이체를 특이적 표적 통합에 대해 비교하였다 (도 26).

[0530] 4.8. Cas9-hyPB 및 hyPB 변이체의 공동-발현 결과

[0531] 유사한 방식으로, 우리는 SpCas9 hyPB R372A-R375A-D450N 융합 단백질 및 hyPB 돌연변이체를 독립적으로 공동-형질감염시켜, 연결되지 않은 단량체 중 하나와의 융합의 이량체 버전을 수득하였다 (도 27).

[0532] **재료 및 방법**

[0533] 렌티바이러스 생산, 농도 및 역가

[0534] 바이러스를 생산하기 위해, 세포는 pSICO(GFP) 또는 pLV-Lama2 (Lama2) 및 pmd2.g (VSVG = 외피) 및 pax2 (IN을 포함한, 패키징 단백질 함유)로 그리고 때때로 단지 wt-인테그라제를 함유하는 플라스미드로 공동-형질감염시켜 비감염성 인테그라제를 구제하였다. 먼저 6×10^5 개 HEK392T 세포 (계대 8)를 6-웰 조직 판(tissue plat)의 웰에 시딩하고 밤새 인큐베이션하였다. 바이러스 생산을 시작하기 5시간 전에, 세포의 배지를 각각의 웰에서

1:1000 CD (클로로퀸 디포스페이트; 스톡 = 25 mM)를 함유하는 1,7 mL 배지로 교체하였다. 플라스미드를 몰비 1,6:1,32:0,72:3,32 (pSICO:pax2:VSVG:wt IN-구제)로 감염시켰다. PEI (폴리에틸렌이민; 스톡=1 mg/mL)을 형질 감염 시약으로 사용하였으며, 한편 형질감염에 사용된 총 DNA 1 µg에는 3 µL PEI가 사용되었다. DNA를 83 µL Opti-MEM (써모 사이언티픽(Thermo Scientific); #31985062)에 희석하고 PEI를 또 다른 83 µL에 희석하였다. 두 용액을 혼합한 후, 이들을 실온에서 15-20분 동안 인큐베이션하였다. 각각의 형질감염 혼합물을 CD-배지와 함께 세포에 적가하였다. 세포를 밤새 인큐베이션하였다. 다음날까지 배지를 교체하고 2,5 mL의 신선한 배지를 첨가하였다. 다음날 세포의 상청액을 1000 rpm에서 5분 동안 원심분리하고 45 µm 필터를 통과시켰다. 4°C에서 19.500 rpm으로 90분 동안 초원심분리하고 원래 배지 부피의 1:100에서 밤새 재현탁하였다. 상청액은 -80°C에 보관하였다.

[0535] 바이러스 역가를 결정하기 위해 HEK293T 세포를 생산된 바이러스로 감염시키고 GFP가 바이러스 패키징 서열에 코딩되어 있기 때문에 GFP 양성 세포의 양을 계수하였다. 따라서, 6-웰 플레이트의 웰당 75000개의 HEK393T 세포를 시딩하였다. 세포는 1:1000 폴리브렌과 이전에 생성된 바이러스 상청액 500 µL (1:3)를 포함하는 1 mL 배지의 혼합물로 감염되었다. 다음날 배지를 교체하였다. 다음날 배지를 흡인하고, 200 µL 트립신을 사용하여 세포를 탈착하였다. 800 µL 일반 배지를 첨가하여 반응을 중단하고 세포 측정법에 의해 분석하였다. 비형광 Lama2 페이로드의 경우에, gDNA를 세포로부터 추출하고 qPCR을 수행하여 페이로드 대 계놈의 카피 수를 결정하였다.

[0536] 생체내 Lama2 전달

[0537] 동물 실험 절차는 바르셀로나 바이오메디컬 연구 단지(Barcelona Biomedical Research Park)의 동물 실험 윤리 위원회(Animal Experimentation Ethic Committee)의 승인을 받았다. C57BL/6J 또는 Lama2 모델 dj2y/dj2y, 8-10주령이 이 연구에 사용되었다. 동물은 잭슨 레보라토리즈(Jackson Laboratories)로부터 구입하였으며, 수컷과 암컷을 구분하지 않고 사용하였다.

[0538] 공여체 마우스를 8주령에 희생시키고 뒷다리 뼈로부터 골수 세포를 수집하였으며; 계통 음성 고갈을 마우스 계통 세포 고갈 키트(Mouse Lineage Cell Depletion Kit) (밀테니 바이오텍(Miltenyi Biotec))을 사용하여 제조업체 설명서 따라 수행하고 적절한 자극 인자 (스텝셀 테크놀로지스(StemCell Technologies))를 가진 줄기 세포 배양 배지에서 밤새 렌티바이러스로 세포를 형질도입하였다. 이어서 이들 세포를 수확하고 후안와 주사(retro-orbital injection)에 의해 (방사선에 의한 컨디셔닝 후) 수령자 마우스에 이식하였다.

[0539] 이식 4주 후, 처리된 마우스로부터 혈액 및 경골근(tibialis muscle)을 수집하였다. 용해 완충제 (써모 피셔 사이언티픽(Thermo Fisher Scientific))을 사용하여 FACS 분석 전에 제조업체 설명서에 따라 혈액 샘플을 처리하였다. FACS 분석 전에 침윤된 세포를 단리하기 위해 경골근을 다지고(minced) 리베라제(liberase)/디스파제(dispace) 소화로 추가 처리하였다.

[0540] 처리 후 6주 및 12주 후에 제조업체 설명서 따라 바이오셉(Bioseb) 장치를 사용하여 그립 강도 검정을 수행하였다.

[0541] hyPB mRNA를 제조업체의 설명서에 따라 리보맥스(RiboMAX) 대규모 RNA 생산 시스템즈(Large Scale RNA Production Systems)-T7 (프로메가(Promega))로 생산하였다. Rosa26 gRNA (25)를 IDT로부터 구입하였다. Rosa26 및 PB512-B 또는 Lama2 트랜스포존을 표적으로 하는 hyPB mRNA, gRNA를 1:2,5:2,5 비율로 후안와를 통해 주입하였다. 총 55 ug의 핵산을 복합체화하여; NP 비율 7에서 생체내-JetPEI (폴리플러스 형질감염)시켰다. 주사 후 10일에 동물을 안락사시키고 간을 단리하고 균질화하였다. 계놈 DNA와 RNA를 간 샘플로부터 추출하였다. Tfrc 내인성 유전자에 대한 트랜스포존 상대 카피 수 및 트랜스진 상대 발현은 각각 qPCR, RT-qPCR 또는 Elisa 시험에 의해 획득하였다.

[0542] 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.

[0543] 간으로부터의 gDNA 및 RNA 수확

[0544] 계놈 DNA 추출은 DNeasy® 혈액 및 조직 키트 프로토콜 (키아젠(QIAGEN)®)에 따라 수행하였다. 간 조직을 PBS (인산염-완충 식염수)에서 균질화하였다. 20 µL의 프로테이나제 K (키트에 의해 제공됨)를 200 µL의 완충제 AL과 함께 첨가하였다. 볼텍싱 후, 샘플을 56°C에서 10분 동안 인큐베이션하였다. 200 µL의 에탄올 (96-

100%)을 첨가하고 짧게 볼텍싱(vortexing)한 후, 혼합물을 3 mL 수집 튜브에 배치된 Dneasy Mini 스핀 컬럼으로 옮기고 8000 rpm에서 1분 동안 원심분리하였다. 스핀 컬럼을 새로운 2 mL 수집 튜브로 옮기고 500 μ L의 완충제 AW1을 첨가하였다. 튜브를 8000 rpm에서 1분 동안 원심분리하였다. 완충제 AW2에 대해 이 세척 단계를 반복하였다 (3분의 원심분리). 이어서, 스핀을 새로운 1.5 mL 마이크로원심분리기 튜브로 옮기고 스핀 컬럼 멤브레인 중앙에 200 μ L의 완충제 AE를 첨가하여 튜브를 1분 동안 방치하여 DNA를 용리시킨 후 8000 rpm에서 1분 동안 원심분리하였다. 농도는 나노드롭 원(NanoDrop One)®(써모피셔 사이언티픽(ThermoFisher Scientific))으로 측정하였다.

[0545] qPCR 및 RT-qPCR

[0546] 게놈 DNA 샘플의 25 ng/ μ L 및 10 ng/ μ L 희석액을 qPCR에 의해 분석하였다. 5 μ L의 각각의 희석액을 4.4 μ L 파워(Power)UP SYBR 그린 마스터믹스(Green MasterMix) (피셔 사이언티픽) 및 0.3 μ L 올리고(oligo)와 앞뒤로 혼합하였다. Lama2, RFP 및 TfrC 내인성 유전자를 표적으로 하는 올리고를 사용하였다 (서열번호: 100-105)

[0547] 고용량 RNA-cDNA 키트 (써모피셔 사이언티픽)를 사용하여 mRNA에서 cDNA로의 역전사를 수행하였다. 1 μ g mRNA 샘플을 20 μ L 반응에서 10 μ L 완충제 및 1 μ L 효소 혼합물과 혼합하였다. 반응물을 37°C에서 2시간 동안 인큐베이션하고 80°C에서 10분 동안 불활성화하였다. RT 제품을 10배 희석하고 5 μ L의 샘플을 4.4 μ L 파워UP SYBR 그린 마스터믹스 (피셔 사이언티픽) 및 0.3 μ L 올리고 fw 및 rv와 혼합하여 qPCR을 수행하였다. Lama2, RFP 및 GAPDH 내인성 유전자를 표적으로 하는 올리고를 사용하였다 (서열번호: 102-107)

[0548] 효소 연결 면역흡착 검정(Enzyme linked immunosorbent assay) (ELISA)

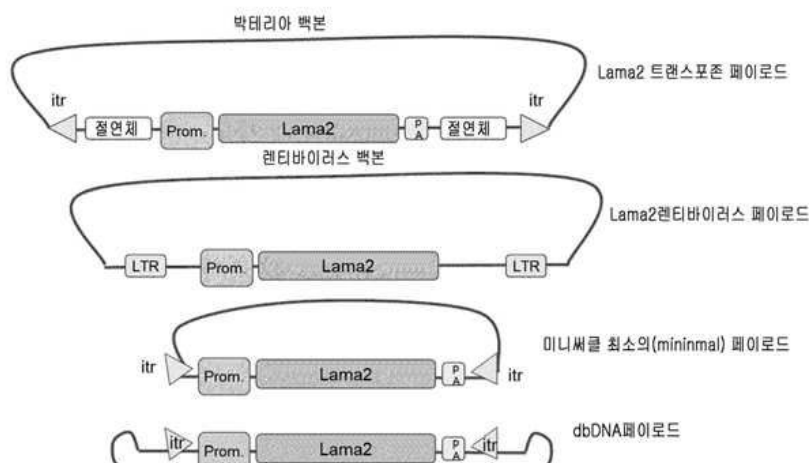
[0549] 전혈을 탁상용 원심분리기에서 4°C 10분 1200xg로 원심분리하였다. 혈장을 분리하여 ELISA 검정을 수행할 때까지 80°C에서 보관하였다. 인간 라미닌 서브유닛 알파-2 (LAMA-2) ELISA 키트 (쿠사비오(Cusabio))를 사용하였다. 20 μ L의 혈장을 80 μ L의 샘플 완충제로 희석하였다. 희석된 샘플을 플레이트에 배치하고 2시간 동안 인큐베이션하였다. 웰을 흡인하였으나 세척하지는 않았고 100 μ L의 비오틴(Biotin)-항체를 첨가하였다. 플레이트를 1시간 동안 인큐베이션하고 세척 완충제로 3회 세척하였다. 100 μ L의 HRP-아비딘을 각각의 웰에 첨가하고 37°C에서 1시간 동안 인큐베이션하였다. 5회 세척의 세척 단계를 조심스럽게 수행하였다. 90 μ L의 TMB 기질을 첨가하고 플레이트를 37°C에서 20분 동안 인큐베이션하였다. 50 μ L의 용액을 첨가하고, 플레이트를 완전히 두드린 다음에 플레이트 판독기에서 450 nm에서 5분 이내에 광학 밀도를 결정하였다.

[0550] 형광 활성화 세포계측법(Fluorescent activated cytometry analysis) 분석 (FACS)

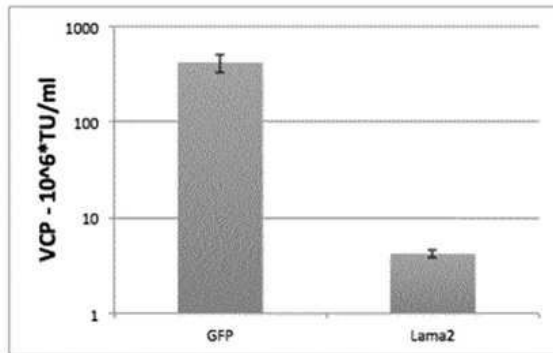
[0551] (BD LSR 포르테사(Fortessa); BD 바이오사이언시스(Biosciences). 530/30 필터를 가진 청색 488 nm 레이저 및 610/20 필터를 가진 황록색 561 nm 레이저)에서 emGFP 발현을 측정하기 전에 순환하는 백혈구 및 근육 내 침윤 세포를 분리하였다.

도면

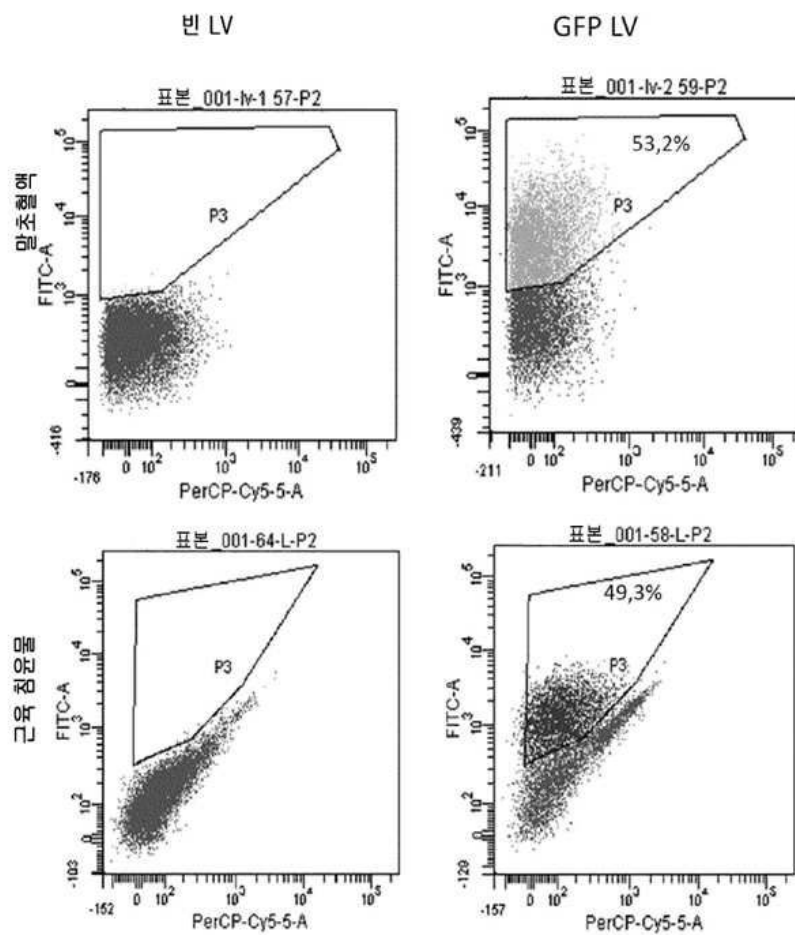
도면1



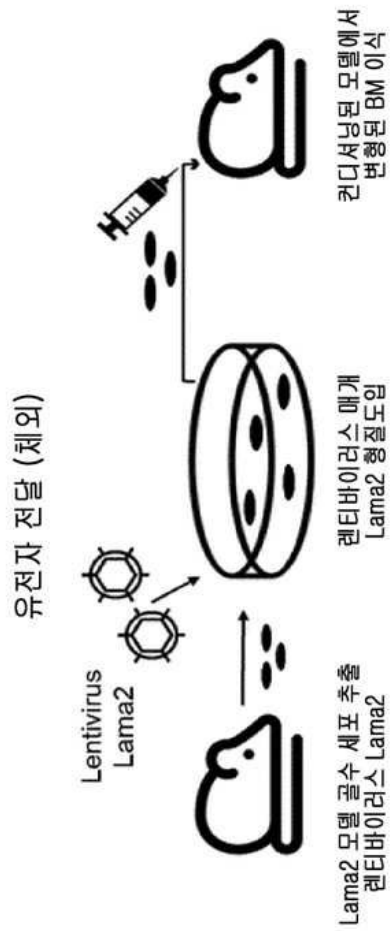
도면2



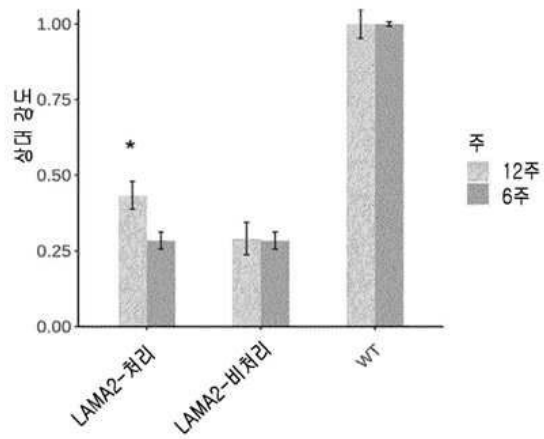
도면3



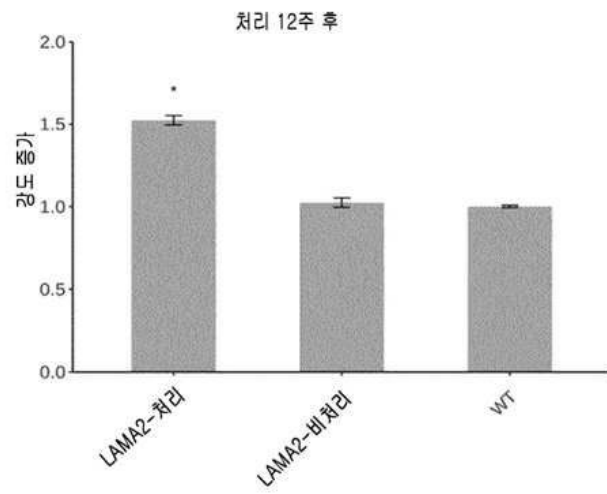
도면4



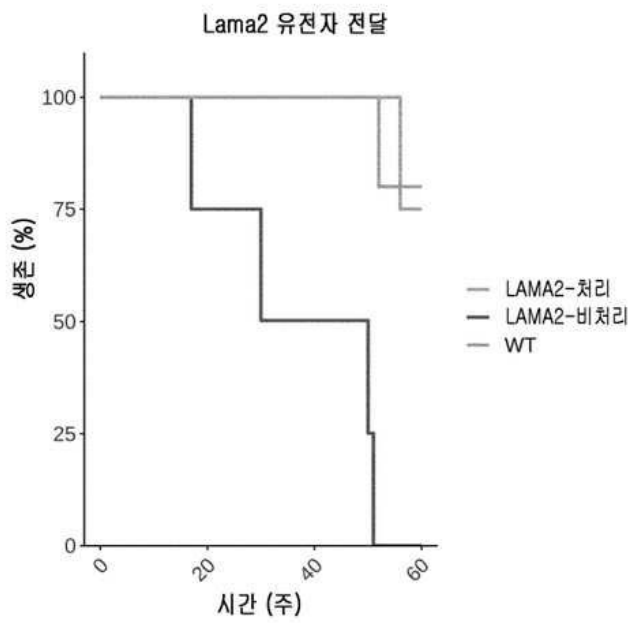
도면5a



도면5b



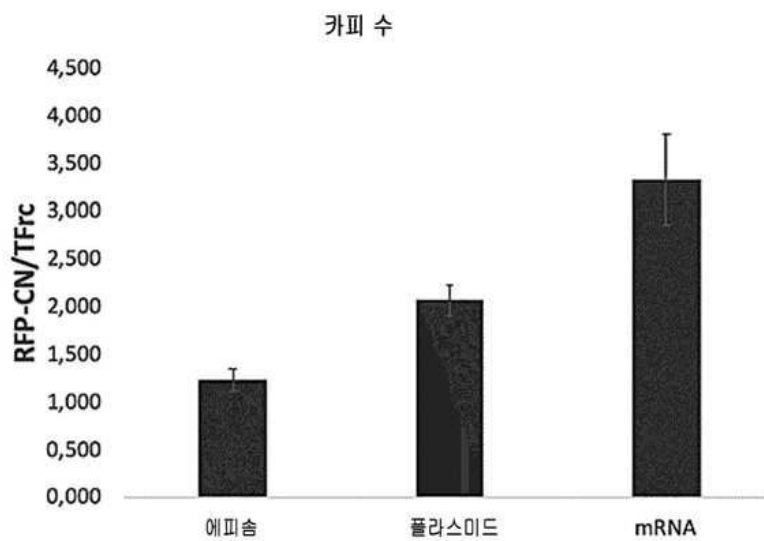
도면6



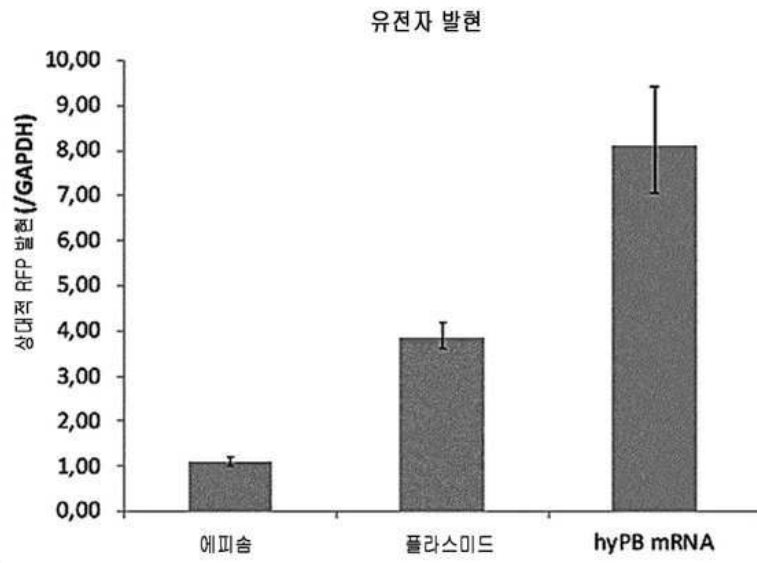
도면7



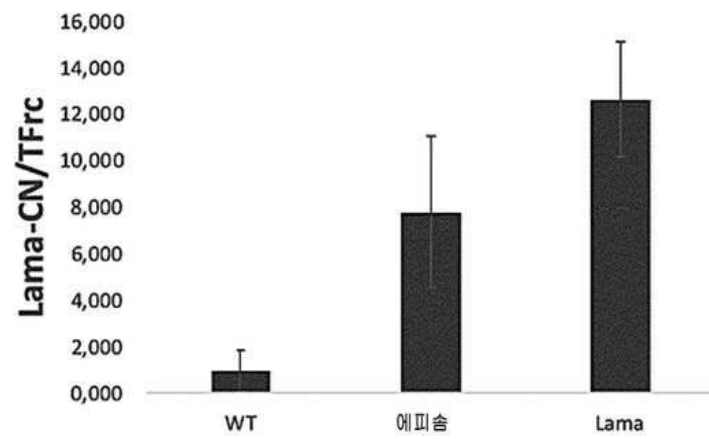
도면8a



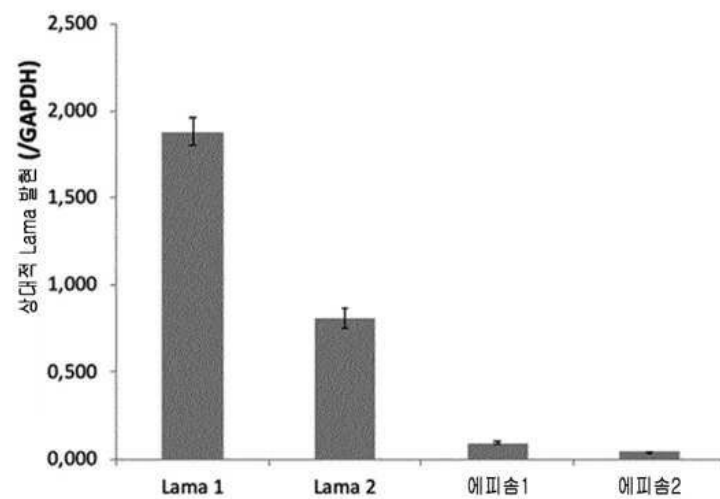
도면8b



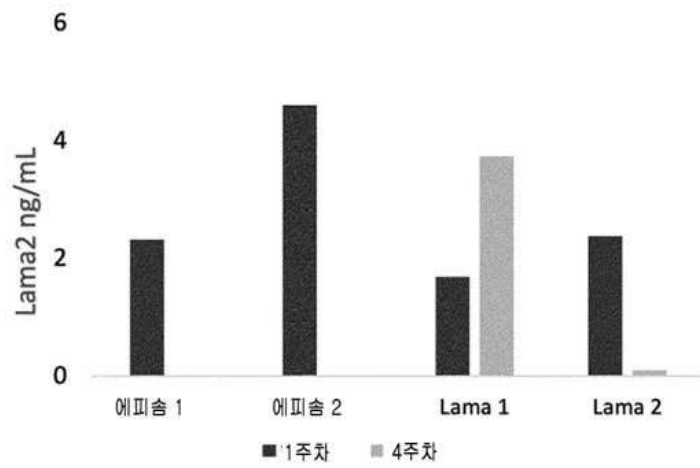
도면8c



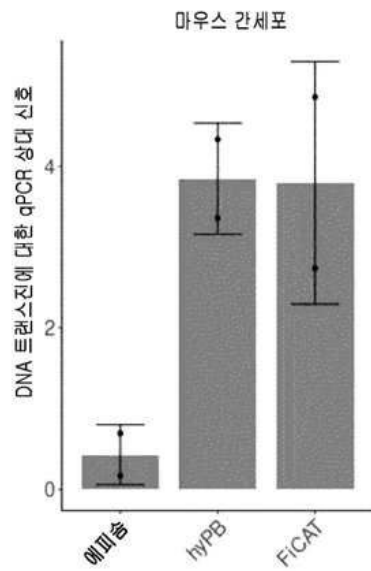
도면8d



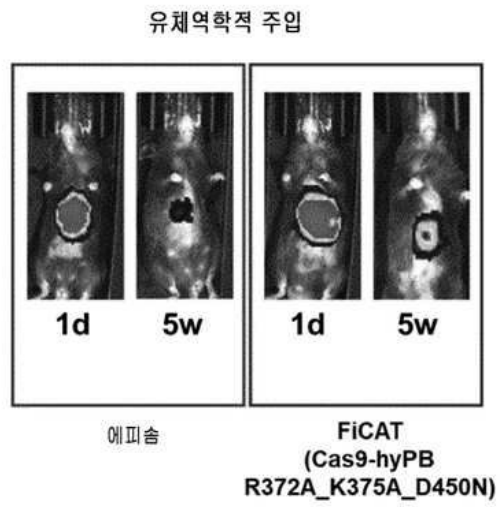
도면9



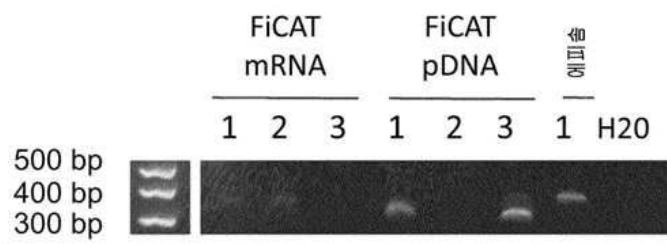
도면10a



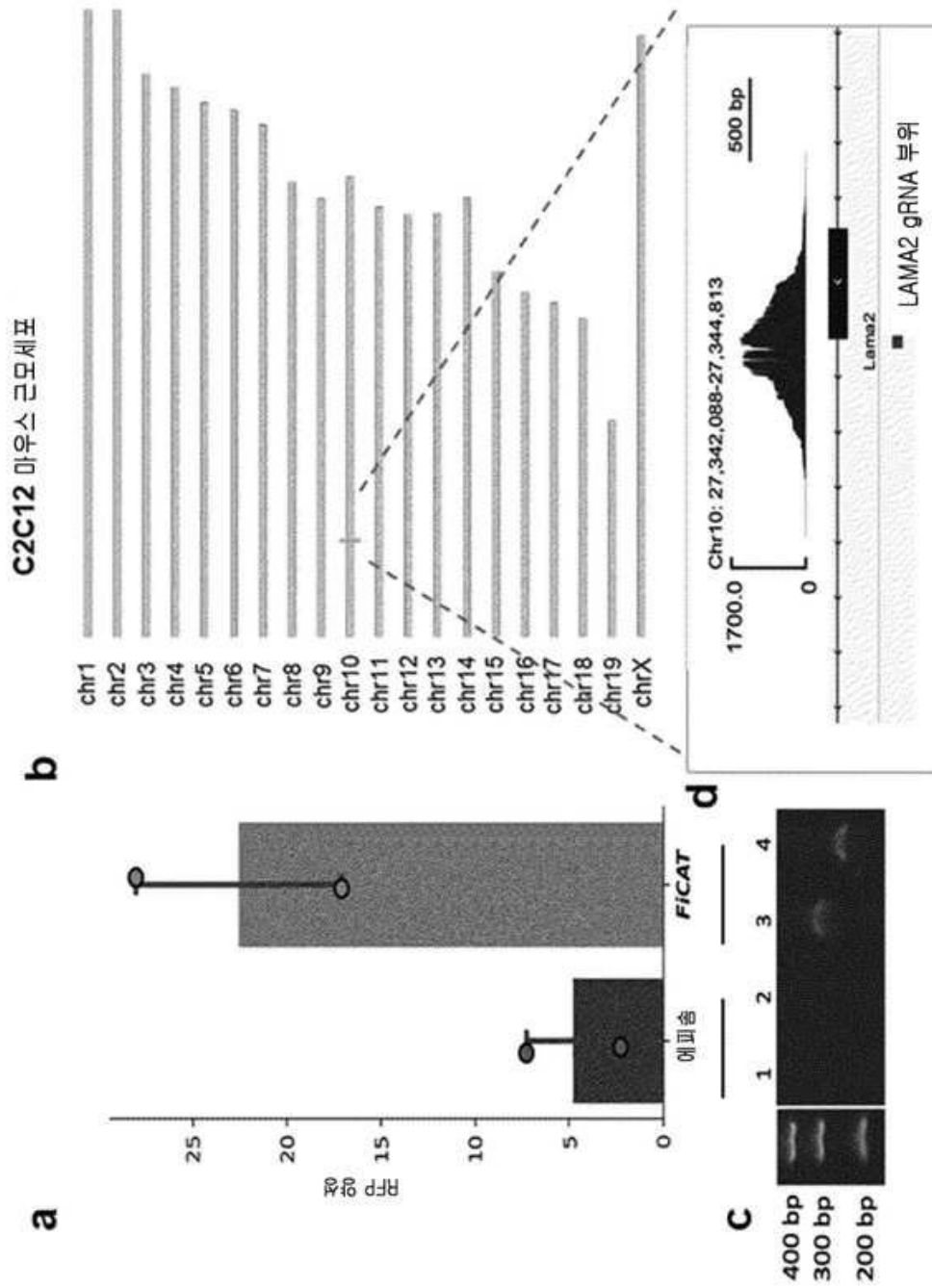
도면10b



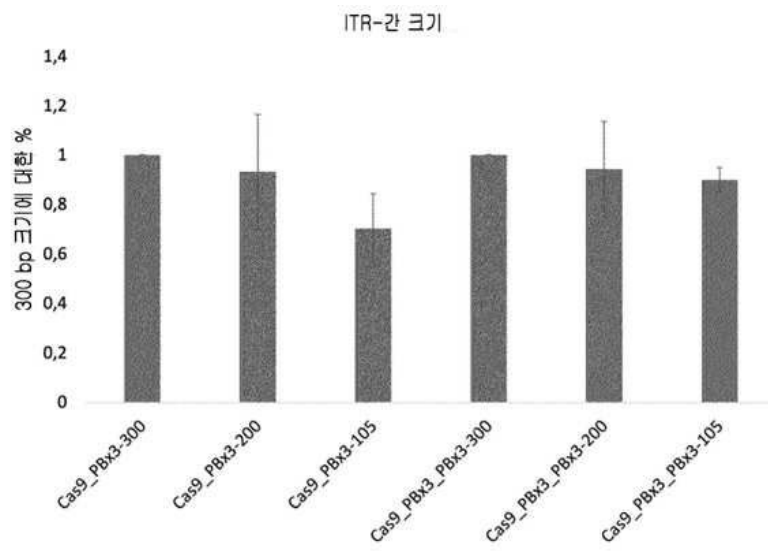
도면10c



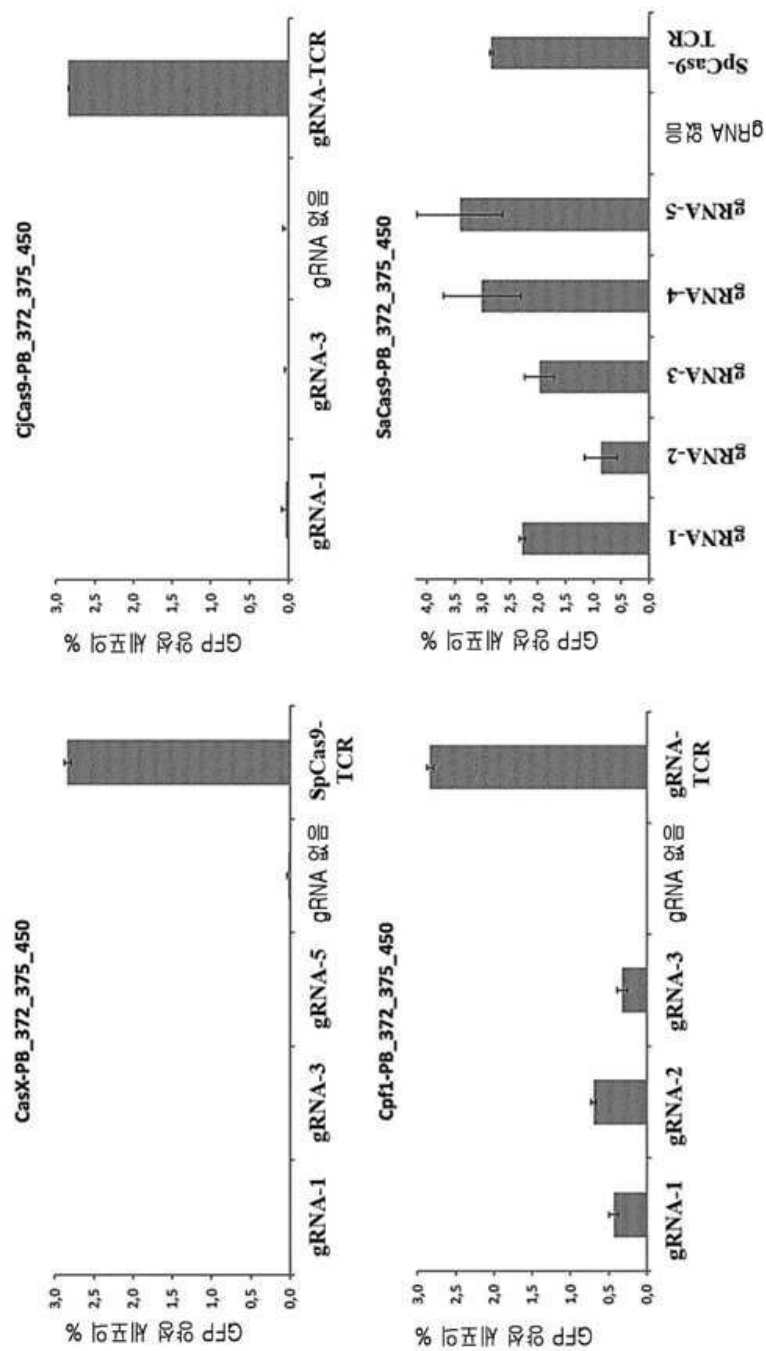
도면11



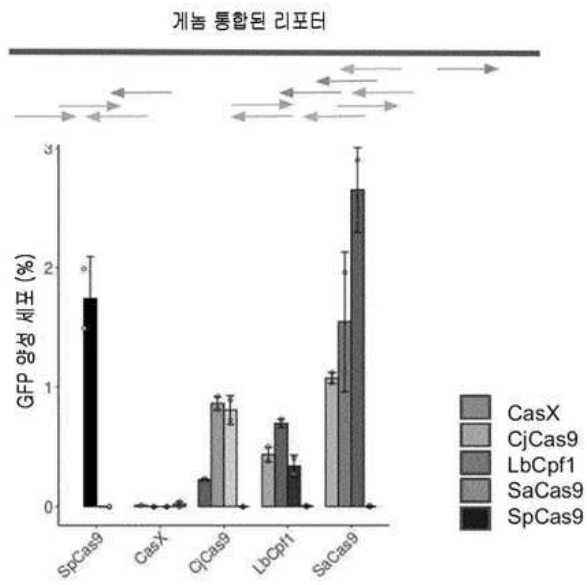
도면12



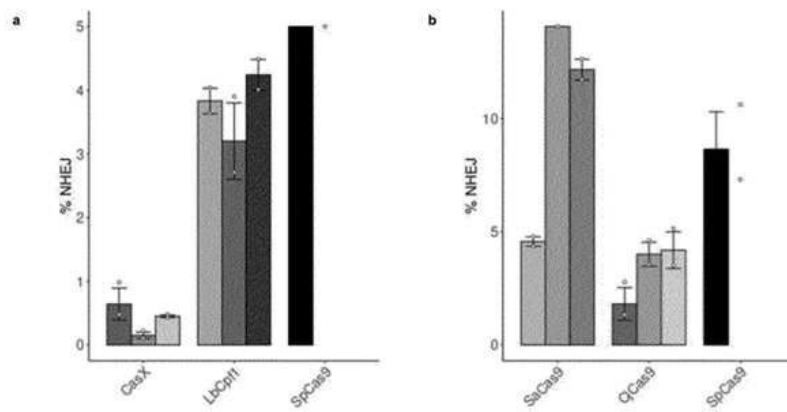
도면13



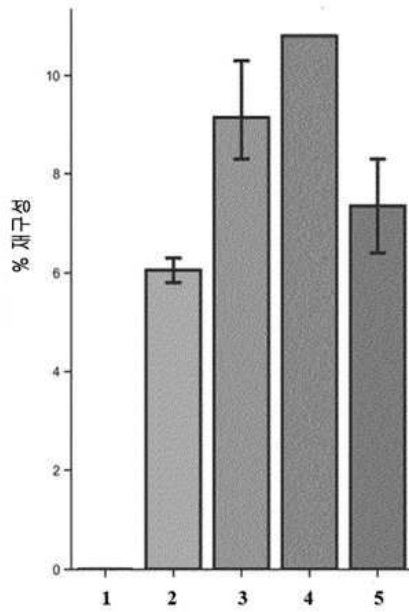
도면14



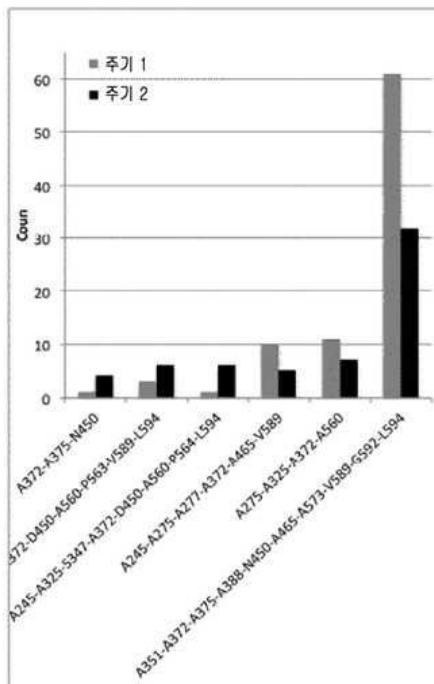
도면15



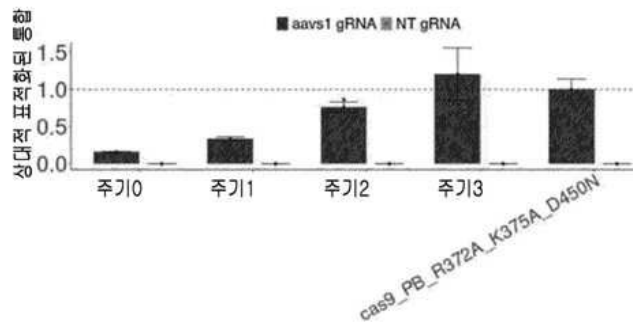
도면16



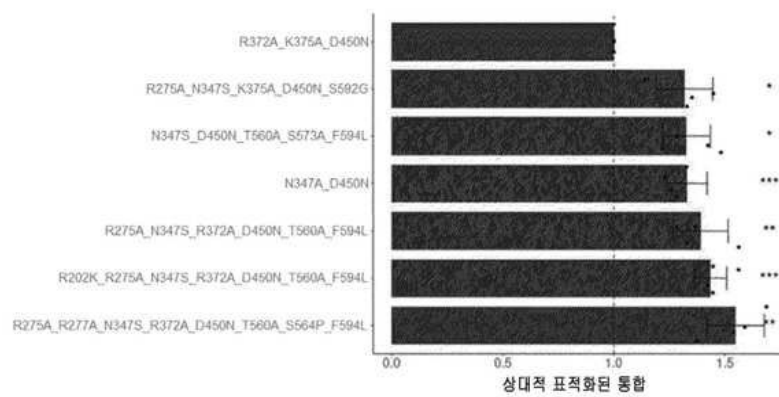
도면17



도면18



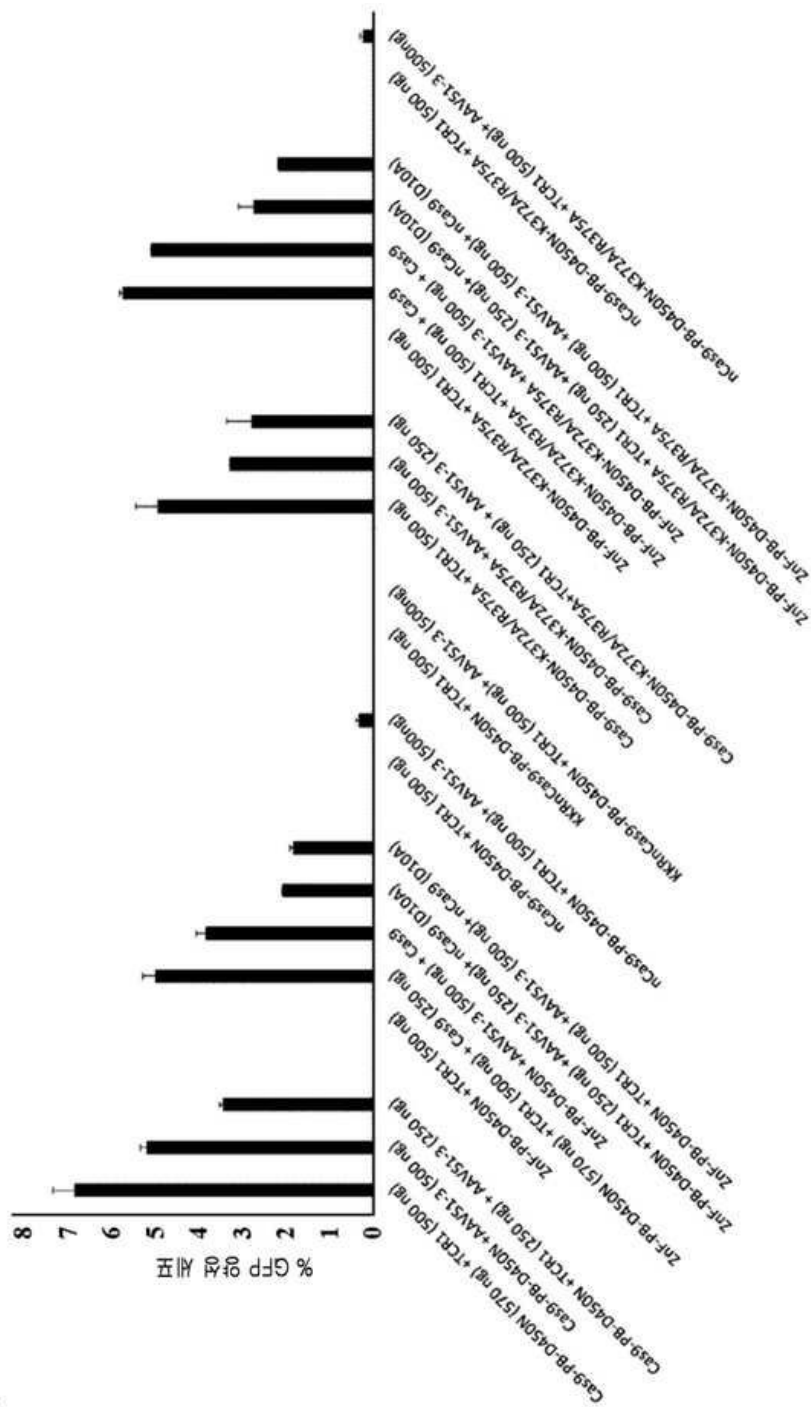
도면19a



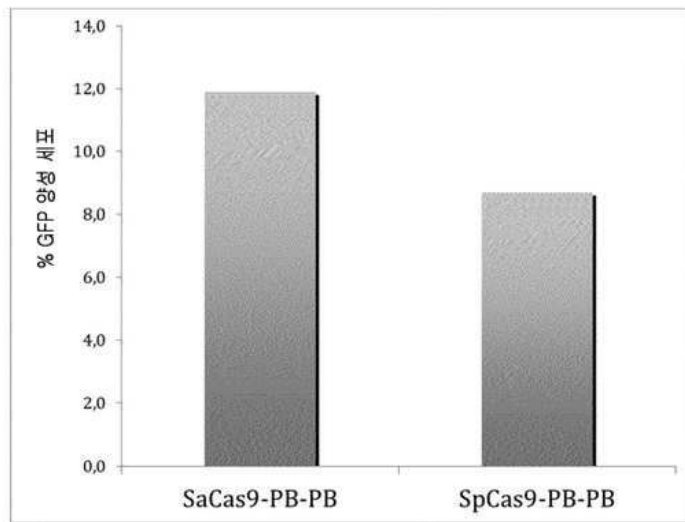
도면19b



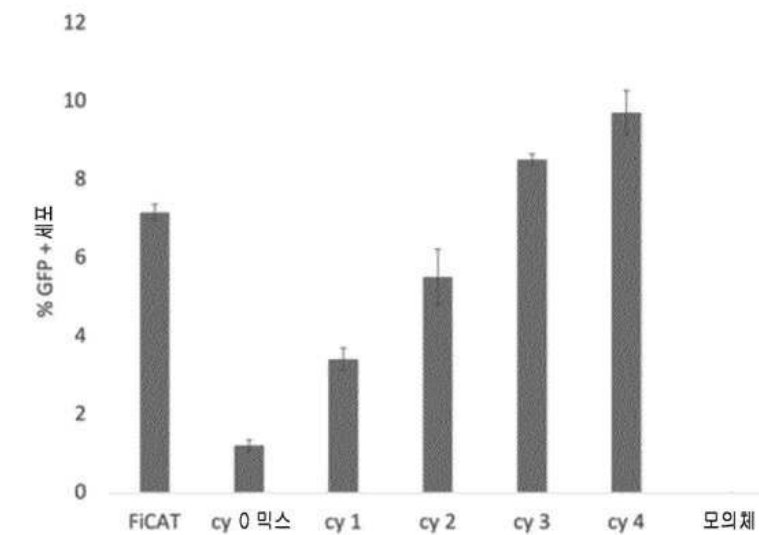
도면20



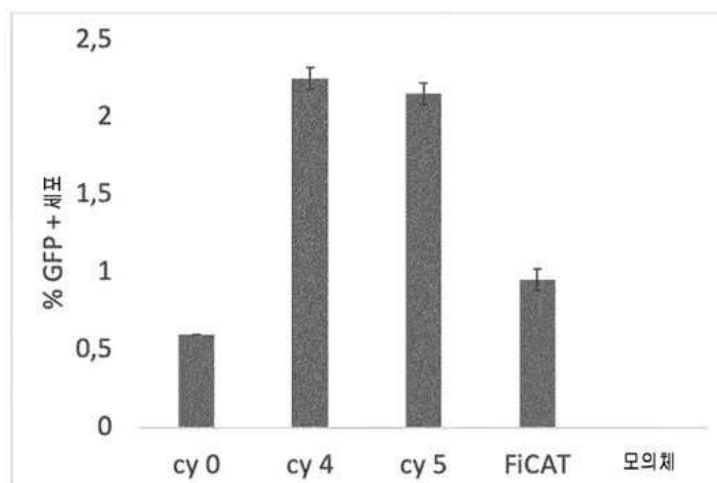
도면21



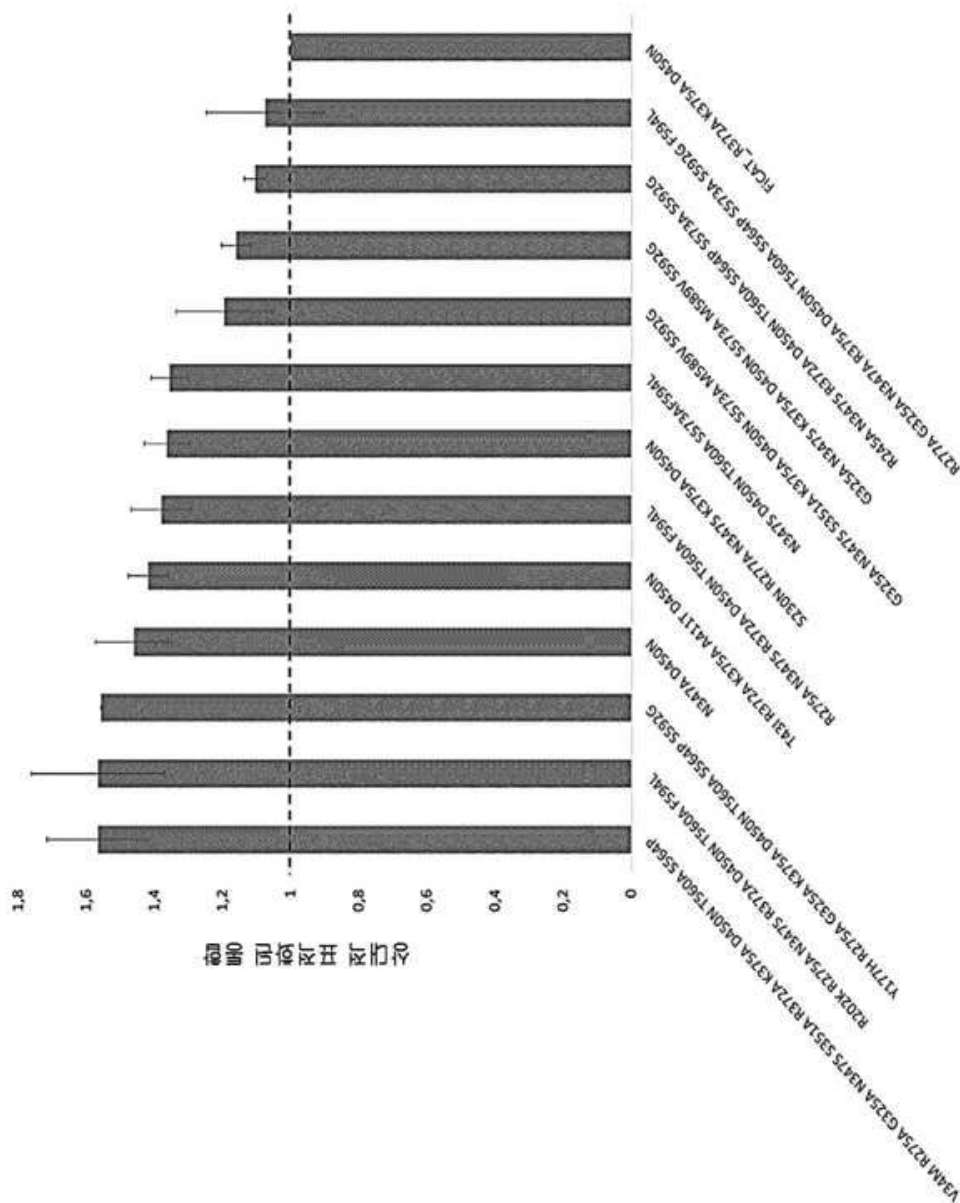
도면22a



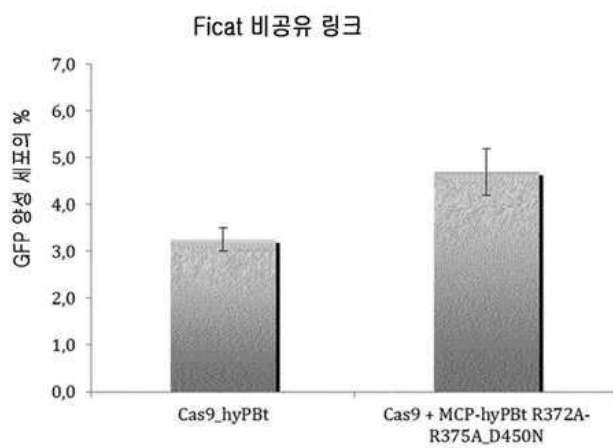
도면22b



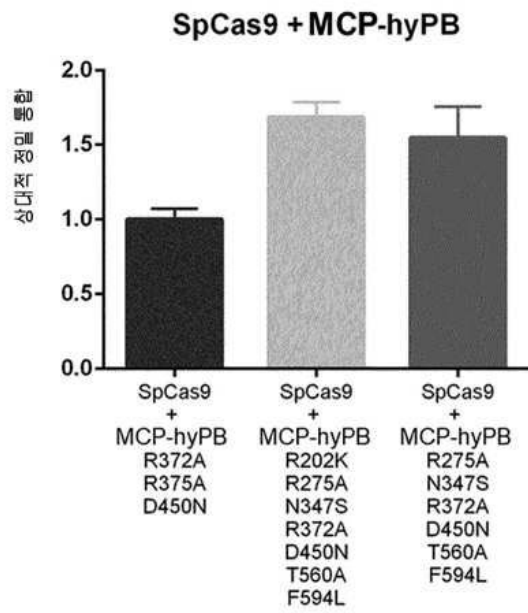
도면23



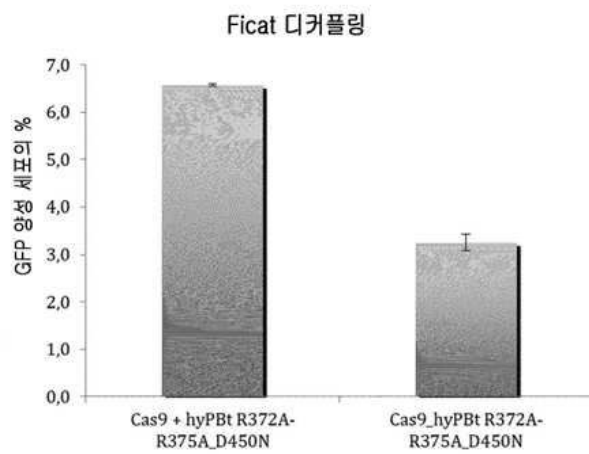
도면24a



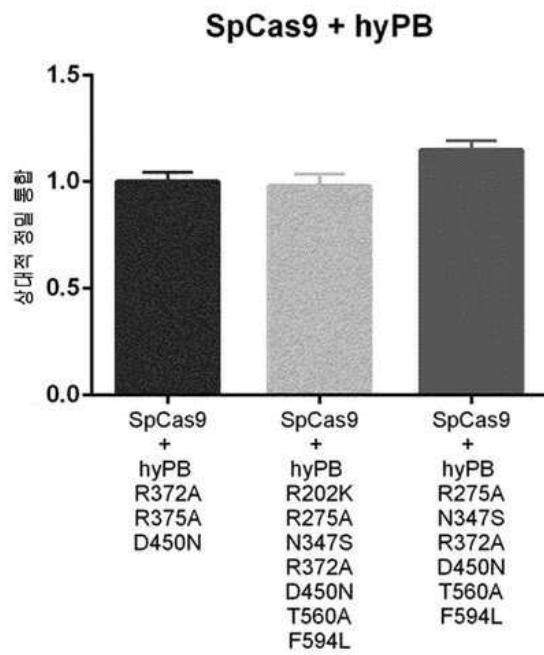
도면24b



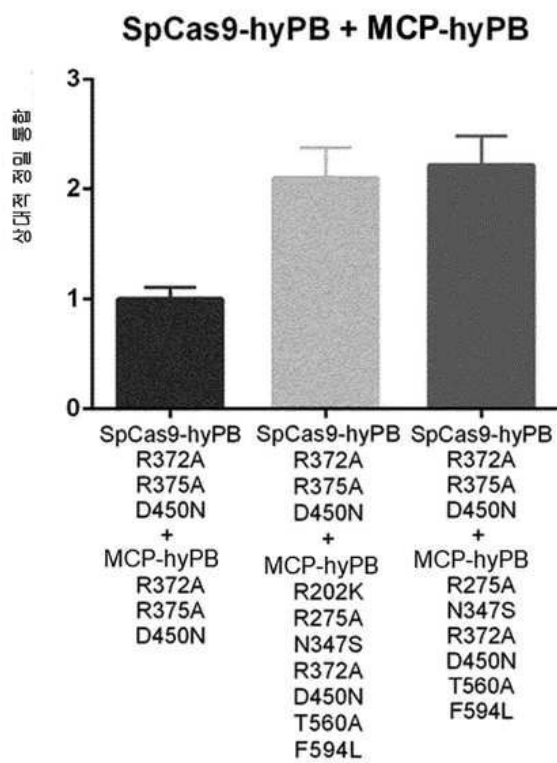
도면25a



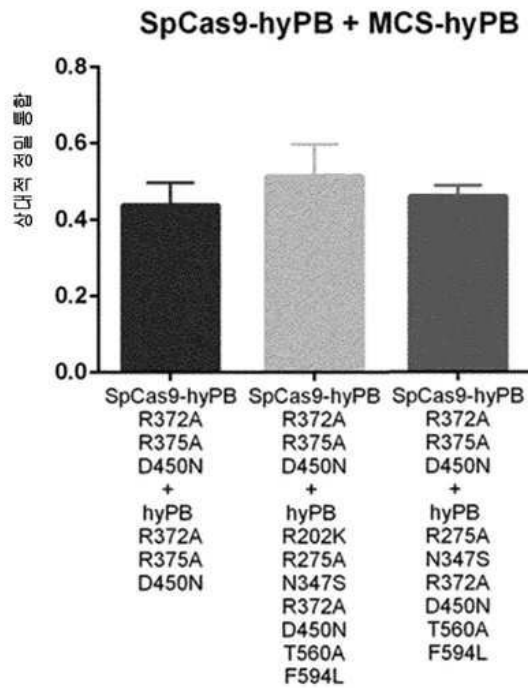
도면25b



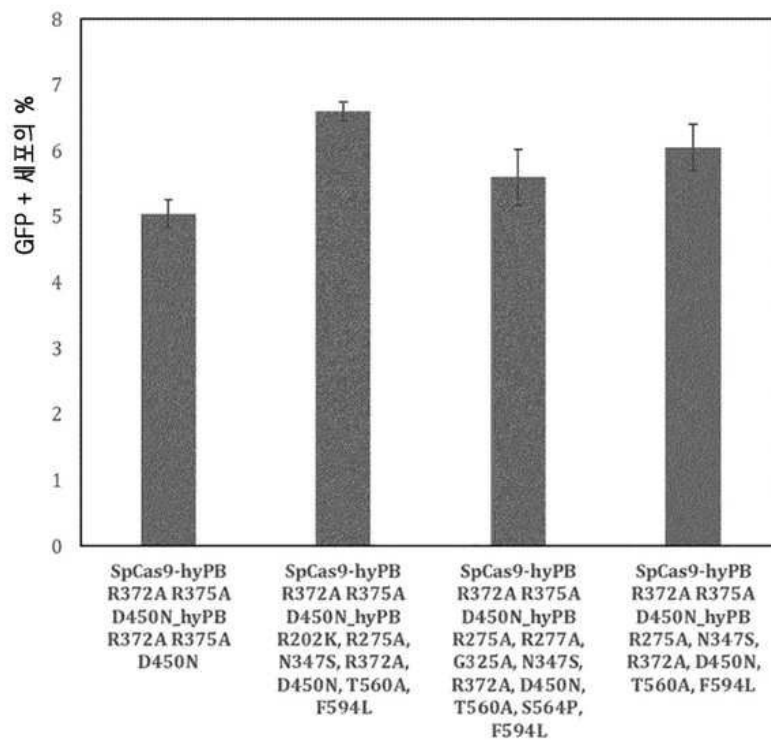
도면26



도면27



도면28



서열 목록

SEQUENCE LISTING

<110> Universitat Pompeu Fabra

<120> THERAPEUTIC LAMA2 PAYLOAD FOR TREATMENT OF CONGENITAL MUSCULAR DYSTROPHY

<130> IBIO-2200/PCT

<150> EP21209721.6

<151> 2021-11-22

<150> EP20214691.6

<151> 2020-12-16

<160> 134

<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

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

 $\langle 210 \rangle$ 2

<211> 594

<212> 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

[illegible]

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 PiggyBac

<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				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 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
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
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
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
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 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 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 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
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
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 gtgcggaag 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 gcatctgagc ccgccatcag	180
cgcaccata ccggcgaaaa accgtataaa tgcccggaat gcggcaaaag ctttagcgat	240
tgccgcgagc tggcgagcca tcagcgagc cataccggcg aaaaaccgta taaatgcccg	300
gaatgcggca aaagcttttag ccagagcagc catctggtgc gccatcagcg caccataacc	360

ggcgaaaaac cgtataaatg cccggaatgc ggcaaaagct ttagcgattg ccgcgagctg	420
gcgcgccatc agcgacacca taccggcgaa aaaccgtata aatgcccgga atgcggcaaa	480
agcttttagcc gcagcgataa actggtgagc catcagcgca ccataaccgg caaaaaaacc	540
agcggccagg 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 cggcgcgtgga accgggcgaa aaaccgtata aatgcccgga atgcggcaaa 60

agcttttagcg atccgggcgc gctggtgcmc catcagcgca cccataccgg cgaaaaaccg 120

tataaatgcc cggaatgcgg caaaagcttt agccagagca gccatctggt gcgccatcag 180

cgcacccata ccggcgaaaa accgtataaa tgcccggaat gcggcaaaag ctttagcgat 240
 tgccgcgatac tggcgcgcca 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 cgcacctac 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> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS linker Nucleic Acid

<400> 40

ggtggatctg gcggtggatc tgggtggcgg

30

<210> 41

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS linker Amino Acid

<400> 41

Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly

1 5 10

<210> 42

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS4x Linker Nucleic Acid

<400> 42

ggagggagtg gtgggtccgg tggtagtggc ggatcc

36

<210> 43

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS4x Linker Amino Acid

<400> 43

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

1 5 10

<210> 44

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS5x Linker Nucleic Acid

<400> 44

ggaggtccg gtgggtctgg tgggagcggg gtagtggcg gatcc

45

<210> 45

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS5x Linker Amino Acid

<400> 45

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

1 5 10 15

<210> 46

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS6x Linker Nucleic Acid

<400> 46

ggaggcagtg gtgggagcgg tgggtccggg gtagtggtg gttccggggg atcc 54

<210> 47

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS6x Linker Amino Acid

<400> 47

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

1 5 10 15

Gly Ser

<210> 48

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS7x Linker Nucleic Acid

<400> 48

ggaggttctg gaggtccgg tgggtccggg ggaagtgggg ggtcaggcgg atcaggagga 60

tcc 63

<210> 49

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS7x Linker Amino Acid

<400> 49

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> 50

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> GGS8x Linker Nucleic Acid

<400> 50

ggaggttagcg gaggttccgg agggagcggc gggagtgggg gaagcggggg aagtggagga 60

tccgggggag gatcc 75

<210> 51

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> GGS8x Linker Amino Acid

<400> 51

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> 52

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker XTEN Nucleic Acid

<400> 52

tccggtagcg aaacaccggg gacttcagaa tcggccaccc cggagtct 48

<210> 53

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Linker XTEN Amino Acid

<400> 53

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser

1 5 10 15

<210> 54

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker B Nucleic Acid

<400> 54

ggaagcgccg gtagtgcggc tgggtctggc gatttc 36

<210> 55

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Linker B Amino Acid

<400> 55

Gly Ser Ala Gly Ser Ala Ala Gly Ser Gly Glu Phe

1 5 10

<210> 56

<211> 4104

<212> DNA

<213> Artificial Sequence

<220>

<223> human Cas9 (hCas9) nucleic acid

<400> 56

atggacaaga agtactccat tgggctcgat atcggcacia 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 ctccatagg	300
ctggaggagt cctttttggt ggaggaggat aaaaagcacg agcgccaccc aatctttggc	360
aatatcgtgg acgaggtggc gtaccatgaa aagtacccaa ccatatatca tctgaggaag	420
aagcttgtag acagtactga taaggctgac ttgcggttga tctatctcgc gctggcgcat	480
atgatcaaat ttcggggaca ctctctcatc gagggggacc tgaaccaga caacagcgat	540
gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg	600
atcaacgcat ccggagtga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat	720
cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa	780
gatgccaaagc 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 ccttgtcaga	1020
cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggctacgcc	1080
ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gccatcttg	1140
gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttgccg	1200
aaacagcgca ctttcgacaa tggaagcatc cccaccaga ttcacctggg cgaactgcac	1260
gctatcctca ggcggaaga ggatttctac ccctttttga aagataacag ggaaaagatt	1320
gagaaaaatc tcacatttcg gataccctac tatgtaggcc cctcgcgcc gggaatttc	1380
agattcgcgt ggatgactcg caaatcagaa gagacatca ctccctggaa cttcgaggaa	1440
gtcgtggata agggggcctc tgcccagtc ttcatcgaaa ggatgactaa ctttgataaa	1500
aatctgccta acgaaaaggt gcttctctaaa cactctctgc tgtacagta 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
attaagaca aggacttcct ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
ctcaccctta cgttggttga agataggag atgattgaag aacgcttgaa aacttacgt	1920
catctcttcg acgacaaagt catgaacag ctcaagaggc gccgatatac aggatggggg	1980
cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg	2040
gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac	2100
tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtcct	2160
cacgagcaca tcgctaactc tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220
gttaaggtcg tggatgaact cgtcaaagta atgggaaggc ataagcccg gaatatcgtt	2280
atcgagatgg cccgagagaa ccaactacc cagaaggagc agaagaacag tagggaaagg	2340
atgaagagga ttgaagaggg tataaaagaa ctggggtccc aaatccttaa ggaacaccca	2400
gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtctc ccgactacga cgtggatcat	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaaaata gaggggaagag tgataacgtc ccctcagaag aagttgtcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacccaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggttcat caaaaggcag	2760
cttggtgaga cagccagat caccaagcac gtggccaaa 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 aagtgtaga 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 gcttctcaa ggaaagtatc ctcccgaaaa ggaacagcga caagctgatc	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct	3420

tacagtgtac tggttgtggc caaagtggag aaaggggaagt 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 cgataagtg 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 agctcggtag agac	4104
<210> 57	
<211> 4104	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> nickase Cas9 (nCas9) nucleic acid	
<400> 57	
atggacaaga agtactccat tgggctcgct atcggcacia acagcgctcg ctgggccgtc	60
attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc	120
cacagcataa agaagaacct cattggcgcc ctctgttctg actccgggga gacggccgaa	180
gccacgcggc taaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc	240
tacctgcagg agatcttttag taatgagatg gctaaggtag atgactcttt ctccatagg	300
ctggaggagt ctttttttgt ggaggaggat aaaaagcacg agcgccaccc aatctttggc	360
aatatcgtgg acgaggtggc gtacatgaa aagtaccaa ccatatatca tctgaggaag	420
aagctttag acagtactga taagctgac ttgcggttga tctatctcgc gctggcgcat	480
atgatcaaat ttcggggaca ctctctatc gagggggacc tgaaccaga caacagcgat	540
gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg	600
atcaacgcat ccggagttag cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
cggctcgaac acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat	720

cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa 780
 gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgtctggcc 840
 cagatcggcg accagtacgc agaccttttt ttggcggcaa agaacctgtc agacgccatt 900
 ctgctgagtg atattctgcg agtgaacacg gagatcacca aagctccgct gagcgctagt 960

 atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc cttgtcaga 1020
 cagcaactgc ctgagaagta caaggaaatt ttcttcgacg 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 ggcggaaga ggatttctac ccttttttga aagataacag ggaaaagatt 1320
 gagaaaatcc tcacatttcg gataccctac tatgtaggcc cctcgcccg gggaaattcc 1380

 agattcgcgt ggatgactcg caaatcagaa gagaccatca ctccctggaa cttcgaggaa 1440
 gtcgtggata agggggcctc tgcccagtcc ttcacgaaa ggatgactaa ctttgataaa 1500
 aatctgccta acgaaaaggt gcttctctaaa 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
 ctacacctta cgttgtttga agataggag atgattgaag aacgcttgaa aacttacgct 1920
 catctcttcg acgacaaagt catgaaacag ctcaagagcg gccgatatac aggatggggg 1980
 cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg 2040
 gattttctta agtccgatgg atttgccaac aggaacttca tgcagttgat ccatgatgac 2100
 tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt 2160
 cacgagcaca tcgctaattc tgcaggtagc ccagctatca aaaagggaat actgcagacc 2220

 gttaaggtcg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt 2280
 atcgagatgg cccgagagaa ccaaaactacc cagaaggagc agaagaacag tagggaaagg 2340
 atgaagagga ttgaagaggg tataaaagaa ctggggtccc aaatccttaa ggaacaccca 2400
 gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg 2460
 gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat 2520
 atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc 2580

gataaaaaata gagggaagag tgataacgtc ccctcagaag aagttgtcaa gaaaatgaaa 2640

aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg 2700

actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggttcat caaaaggcag 2760

cttggtgaga cagccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac 2820

accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct 2880

aagctggctc cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat 2940

taccaccatg cgcatgatgc ctacctgaat gcagtggtag gcactgcact tatcaaaaaa 3000

tatccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa 3060

atgatcgcaa agictgagca 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 caagctgatc 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 gcttccaag 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 gagcaataa gcgaattctc caaaagagtg 3840

atcctcgccg acgctaacct cgataaggctg ctttctgctt acaataagca cagggataag 3900

cccatcaggg agcaggcaga aaacattatc cacttgttta ctctgaccaa cttgggcgcg 3960

cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtacac ctctacaaag 4020

gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc 4080

gacctctctc agctcgggtg agac 4104

<210> 58

<211> 3213

<212> DNA

<213> Staphylococcus aureus

<400> 58

gccacatgg ccccaaagaa gaagcgggaag gtcggtatcc acggagtccc agcagccaag	60
cggaactaca tcctgggcct ggccatcggc atcaccagcg tgggctacgg catcatcgac	120
tacgagacac gggacgtgat cgatgccggc gtgcggctgt tcaaagaggc caacgtggaa	180
aacaacgagg gcaggcggag caagagaggc gccagaaggc tgaagcggcg gaggcggcat	240
agaatccaga gagtgaagaa gctgctgttc gactacaacc tgctgaccga ccacagcgag	300
ctgagcggca tcaacccta cgaggccaga gtgaagggcc tgagccagaa gctgagcgag	360
gaagagtctt ctgccgcct gctgcacctg gccaaagaa gaggcgtgca caacgtgaac	420
gaggtggaag aggacaccgg caacgagctg tccaccaaag agcagatcag ccggaacagc	480
aaggccctgg 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 tcaaagaatg gtacgagatg ctgatgggcc actgcaccta cttccccgag	780
gaactgcgga gcgtgaagta cgcctacaac gccgacctgt acaacgccct gaacgacctg	840
aacaatctcg tgatcaccag ggacgagaac gagaagctgg aatattacga gaagttccag	900
atcatcgaga acgtgttcaa gcagaagaag aagcccaccc tgaagcagat cgccaaagaa	960
atcctcgtga acgaagagga tattaagggc tacagagtga ccagcaccgg caagcccag	1020
ttaccaacc tgaaggtgta ccacgacatc aaggacatta ccgcccgaa agagattatt	1080
gagaacgccg agctgctgga tcagattgcc aagatcctga ccatctacca gagcagcgag	1140
gacatccagg aagaactgac caatctgaac tccgagctga cccaggaaga gatcgagcag	1200
atctctaate tgaagggcta taccggcacc cacaacctga gcctgaaggc catcaacctg	1260
atcctggacg agctgtggca caccaacgac aaccagatcg ctatcttcaa ccggctgaag	1320
ctggtgccca agaaggtgga cctgtcccag cagaaagaga tccccaccac cctggtggac	1380
gacttcaccc tgagccccgt cgtgaagaga agcttcaccc agagcatcaa agtgatcaac	1440
gccatcatca agaagtacgg cctgcccaac gacatcatta tcgagctggc ccgcgagaag	1500
aactccaagg acgccagaa aatgatcaac gagatgcaga agcggaaccg gcagaccaac	1560
gagcggatcg aggaaatcat ccggaccacc ggcaaagaga acgccaagta cctgatcgag	1620
aagatcaagc tgcacgacat gcaggaaggc aagtgcctgt acagcctgga agccatccct	1680
ctggaagatc tgctgaacaa ccccttcaac tatgaggtgg accacatcat ccccagaagc	1740

gtgtccttcg acaacagctt caacaacaag gtgctcgtga agcaggaaga aaacagcaag 1800

aagggaacc ggacccatt ccagtacctg agcagcagcg acagcaagat cagctacgaa 1860

accttcaaga agcacatcct gaatctggcc aagggaagg gcagaatcag caagaccaag 1920

aaagagtatc tgctggaaga acgggacatc aacaggttct ccgtgcagaa agacttcac 1980

aaccggaacc tggaggatc cagatagcc accagaggcc tgatgaacct gctgcggagc 2040

tacttcagag tgaacaacct ggacgtgaaa gtgaagtcca tcaatggcgg cttaccagc 2100

tttctgcggc ggaagtggaa gttaagaaa gagcgggaca aggggtacaa gcaccacgcc 2160

gaggacgccc tgatcattgc caacgccgat ttcatttca aagagtggaa gaaactggac 2220

aaggcaaaa aagtgatgga aaaccagatg ttcaggaaa agcaggccga gagcatgccc 2280

gagatcgaac ccgagcagga gtacaagag atcttcatca cccccacca gatcaagcac 2340

attaaggact tcaaggacta caagtacagc caccgggtgg acaagaagcc taatagagag 2400

ctgattaacg acacctgta ctccaccggc 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

aactacctga ccaagtactc caaaaaggac aacggccccg tgatcaagaa gattaagtat 2700

tacggcaaca aactgaacgc ccatctggac atcaccgacg actacccaa cagcagaaac 2760

aaggtcgtga agctgtccct gaagccctac agattcgacg tgtacctgga caatggcgtg 2820

tacaagttcg tgaccgtgaa gaatctggat gtgatcaaaa aagaaaacta ctacgaagtg 2880

aatagcaagt gctatgagga agctaagaag ctgaagaaga tcagcaacca ggccgagttt 2940

atcgctcct tctacaaca cgatctgatc aagatcaacg gcgagctgta tagagtgatc 3000

ggcgtgaaca acgacctgct gaaccggatc gaagtgaaca tgatcgacat cacctaccgc 3060

gagtacctgg aaaacatgaa cgacaagagg cccccagga tcattaagac aatcgctcc 3120

aagaccaga gcattaagaa gtacagcaca gacattctgg gcaacctgta tgaagtgaag 3180

tctaagaagc accctcagat catcaaaaag ggc 3213

<210> 59

<211> 4104

<212> DNA

<213> Artificial Sequence

<220>

<223> dead Cas9 (dCas9) nucleic acid

<400> 59

atggacaaga agtactccat tgggctcgct atcggcacaa acagcgctcg ctgggccgtc	60
attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc	120
cacagcataa agaagaacct cattggcgcc ctccgttctg actccgggga gacggccgaa	180
gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc	240
tacctgcagg agatctttag taatgagatg gctaagggtg atgactcttt ctccatagg	300
ctggaggagt ctttttgggt ggaggaggat aaaaagcacg agcgccaccc aatctttggc	360
aatatcgtgg acgaggtggc gtaccatgaa aagtacccaa ccatatatca tctgaggaag	420
aagcttgtag acagtactga taaggctgac ttgcggttga tctatctcgc gctggcgcat	480
atgatcaaat ttcggggaca ctccctcatc gagggggacc tgaaccaga caacagcgat	540
gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg	600
atcaacgcat ccggagtga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttggtaat	720
cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa	780
gatgccaaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc	840
cagatcggcg accagtacgc agaccttttt ttggcgcaaa agaacctgtc agacgccatt	900
ctgctgagtg atattctcgc agtgaacacg gagatcacca aagctccgct gagcgctagt	960
atgatcaagc gctatgatga gcaccaccaa gacttgactt tgcgaaggc ccttgtcaga	1020
cagcaactgc ctgagaagta caaggaaatt ttcttcgac agtctaaaaa tggctacgcc	1080
ggatacattg acggcggagc aagccaggag gaattttaca aatttattaa gccatcttg	1140
gaaaaaatgg acggcaccga ggagctgctg gtaaaagctta acagagaaga tctgttcgc	1200
aaacagcgca ctttcgacaa tggaagcatc cccaccaga ttacactggg cgaactgcac	1260
gctatcctca ggcggaaga ggatttctac cctttttga aagataacag ggaaaagatt	1320
gagaaaaatc tcacatttcg gataccctac tatgtaggcc cctcgcccg gggaaattcc	1380
agattcgcgt ggatgactcg caaatcagaa gagacatca ctccctggaa cttcgaggaa	1440
gtcgtggata agggggcctc tgcccagtcc ttcacgaaa ggatgactaa ctttgataaa	1500
aatctgccta acgaaaaggt gtttctctaa cactctctgc tgtacgagta cttcacagtt	1560
tataacgagc tcaccaaggt caaatcgtc acagaaggga tgagaaagcc agcattcctg	1620
tctggagagc agaagaaagc tatcgtggac ctccctttca 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 cgttgtttga agataggag atgattgaag aacgcttgaa aacttacgct	1920
catctcttcg acgacaaagt catgaacag 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 cagaagggac agaagaacag tagggaaagg	2340
atgaagagga ttgaagaggg tataaaagaa ctggggtccc aaatccttaa ggaacaccca	2400
gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggctgct	2520
atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
gataaagcta gaggggaagag tgataacgtc cctcagaag aagtgtgcaa gaaaatgaaa	2640
aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg	2700
actaaggctg aacgagggtg cctgtctgag ttggataaag ccggttcat caaaaggcag	2760
cttgttgaga cagccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac	2820
accaagtacg atgaaaatga caaactgatt cgagaggatga 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 aactggcca atggagagat tcggaagcga	3180
ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
gcgacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
cagaccggag gcttctccaa ggaaagtatc ctcccgaaaa ggaacagcga caagctgatc	3360
gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtgcgt	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 gtctgtggaa	3780
caacacaaac actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagt	3840
atcctcgccg acgctaacct cgataagggtg ctttctgctt acaataagca cagggataag	3900
cccatcaggg agcaggcaga aaacattatc cacttggtta ctcigaccaa cttgggcgcg	3960
cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtagac ctctacaaag	4020
gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc	4080
gacctctctc agctcgggtg agac	4104
<210> 60	
<211> 3207	
<212> DNA	
<213> Staphylococcus aureus	
<400> 60	
atggcccaa agaagaagcg gaaggtcggg atccacggag tcccagcagc caagcgaac	60
tacatcctgg gcttgacatc cggcatcacc agcgtgggct acggcatcat cgactacgag	120
acacgggacg tgatcgatgc cggcgtgcgg ctgttcaaag aggccaacgt ggaaaacaac	180
gagggcaggc ggagcaagag aggcgccaga aggtgaagc ggcgaggcg gcatagaatc	240
cagagagtga agaagctgct gtctgactac aacctgctga ccgaccacag cgagctgagc	300
ggcatcaacc cctacgagc cagagtgaag ggctgagcc agaagctgag cgaggaagag	360
ttctctgccg cctgctgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg	420
gaagaggaca ccggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc	480
ctggaagaga aatacgtggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagt	540
cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg	600
aaggtgcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg	660
ctggaaaccc ggcgaccta ctatgaggga cctggcgagg gcagccctt cggtggaag	720
gacatcaaag aatggtacga gatgctgatg ggccactgca cctacttccc cgaggaactg	780
cggagcgtga agtacgccta caacccgac ctgtacaacg ccctgaacga cctgaacaat	840
ctcgtgatca ccaggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc	900
gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatctc	960

gtgaacgaag aggatattaa gggctacaga gtgaccagca cggcaagcc cgagttcacc 1020

aacctgaagg tgtaccacga catcaaggac attaccgccc ggaaagagat tattgagaac 1080

gccgagctgc tggatcagat tgccaagatc ctgaccatct accagagcag cgaggacatc 1140

caggaagaac tgaccaatct gaactccgag ctgaccagg aagagatcga gcagatctct 1200

aatctgaagg gctataccgg caccacaaac ctgagcctga aggccatcaa cctgatcctg 1260

gacgagctgt ggcacaccaa cgacaaccag atcgctatct tcaaccggct gaagctggtg 1320

ccaagaagg 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 gcttcaaca caaggtgctc gtgaagcagg aagaaaacag caagaagggc 1800

aaccggaccc cattccagta cctgagcagc agcgacagca agatcagcta cgaaaccttc 1860

aagaagcaca tctgaatct ggccaagggc aagggcagaa tcagcaagac caagaaagag 1920

tatctgctgg aagaacggga catcaacagg ttctccgtgc agaaagactt catcaaccgg 1980

aacctggtgg ataccagata cgccaccaga ggctgatga acctgctgcg gagctacttc 2040

agagtgaaca acctggacgt gaaagtgaag tccatcaatg gcggttcac cagctttctg 2100

cggcggaagt ggaagtttaa gaaagagcgg aacaagggtt acaagcacca cgccgaggac 2160

ggcctgatca ttgccaacgc cgatttcac ttcaaagagt ggaagaaact ggacaaggcc 2220

aaaaaagtga tggaaaacca gatgttcgag gaaaagcagg ccgagagcat gcccagatc 2280

gaaaccgagc aggagtacaa agagatcttc atcccccc accagatcaa gcacattaag 2340

gatttcaagg actacaagta cagccaccgg gtggacaaga agcctaatag agagctgatt 2400

aacgacaccc tgiactccac ccggaaggac gacaaggga acacctgat cgtgaacaat 2460

ctgaacggcc tgtacgacaa ggacaatgac aagctgaaaa agctgatcaa caagagcccc 2520

gaaaagctgc tgatgtacca ccacgacccc cagacctacc agaaactgaa gctgattatg 2580

gaacagtacg gcgacgagaa gaatccccig tacaagtact acgaggaaac cgggaactac 2640

ctgaccaagt actcaaaaa ggacaacggc cccgtgatca agaagattaa gtattacggc 2700

aacaaactga acgcccattt ggacatcacc gacgactacc ccaacagcag aaacaaggtc	2760
gtgaagctgt cctgaagcc ctacagattc gacgtgtacc tggacaatgg cgtgtacaag	2820
ttcgtgaccg tgaagaatct ggatgtgac aaaaaagaaa actactacga agtgaatagc	2880
aagtgtctatg aggaagctaa gaagctgaag aagatcagca accaggccga gtttatcgcc	2940
tccttttaca acaacgatct gatcaagatc aacggcgagc tgtatagagt gatcggcgtg	3000
aacaacgacc tgctgaaccg gatcgaagtg aacatgatcg acatcaccta ccgcgagtac	3060
ctggaaaaca tgaacgacaa gaggccccc aggatcatta agacaatcgc ctccaagacc	3120
cagagcatta agaagtacag cacagacatt ctgggcaacc tgtatgaagt gaaatctaag	3180
aagcacctc agatcatcaa aaagggc	3207
<210> 61	
<211> 3708	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> cpf1	
<400> 61	
atggcccaa agaagaagcg gaaggtcagc aagctggaga agtttacaaa ctgctactcc	60
ctgtctaaga cctgaggtt caaggccatc cctgtgggca agaccagga gaacatcgac	120
aataagcggc tgctggtgga ggacgagaag agagccgagg attataaggc cgtgaagaag	180
ctgctggatc gctactatct gtctttttatc aacgacgtgc tgcacagcat caagctgaag	240
aatctgaaca attacatcag cctgttccgg aagaaaacca gaaccgagaa ggagaataag	300
gagctggaga acctggagat caatctgcgg aaggagatcg ccaaggcctt caaggcgaac	360
gagggtctaca agtccctgtt taagaaggat atcatcgaga caatcctgcc agagttcctg	420
gacgataagg acgagatcgc cctggtgaac agcttcaatg gctttaccac agccttcacc	480
ggcttctttg ataacagaga gaatatgttt tccgaggagg ccaagagcac atccatcgcc	540
ttcaggtgta tcaacgagaa tctgacccgc tacatctcta atatggacat cttcagaag	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
tttaagccac 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
ggcgagtgga acgtgatccg ggacaagtgg aatgccgagt atgacgatat ccacctgaag	1140
aagaaggccg tggtagccga gaagtacgag gacgatcgga gaaagtcctt caagaagatc	1200
ggctcctttt ctctggagca gctgcaggag tacgccgacg ccgatctgtc tgtggtggag	1260
aagctgaagg agatcatcat ccagaagggtg gatgagatct acaagggtga tggctcctct	1320
gagaagctgt tcgacgccga ttttgtgctg gagaagagcc tgaagaagaa cgacgccgtg	1380
gtggccatca tgaaggacct gctggattct gtgaagagct tcgagaatta catcaaggcc	1440
ttctttggcg agggcaagga gacaaacagg gacgagtcct tctatggcga ttttgtgctg	1500
gcctacgaca tctgtctgaa ggtggaccac atctacgatg ccatccgcaa ttatgtgacc	1560
cagaagccct actctaagga taagttcaag ctgtattttc agaaccctca gttcatgggc	1620
ggctgggaca aggataagga gacagactat cgggccacca tctgagata cggtccaag	1680
tactatctgg ccatcatgga taagaagtac gccaaagtgc 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 gagtctgcca gcaagaagga ggtggataag	2100
ctggtggagg agggcaagct gtatatgttc cagatctata acaaggactt ttccgataag	2160
tctcacggca cacccaatct gcacaccatg tacttcaagc tgctgtttga cgagaacaat	2220
cacggacaga tcaggctgag cggaggagca gagctgttca tgaggcgcg ctcctgaag	2280
aaggaggagc tggtagtgca ccagccaac tcccctatcg ccaacaagaa tccagataat	2340
cccaagaaaa ccacaacct gtectacgac gtgtataagg ataagaggtt ttctgaggac	2400
cagtacgagc tgcacatccc aatcgccatc aataagtgc ccaagaacat cttaagatc	2460
aatacagagg tgcgcgtgct gctgaagcac gacgataacc cctatgtgat cggcatcgat	2520
aggggagcgc 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 tgccgtgac gccctggagg acctgaactc tggctttaag 2820

aatagccgcg tgaagggtga gaagcaggtg tatcagaagt tcgagaagat gctgatcgat 2880

aagctgaact acatggtgga caagaagtct aatccttgtg caacaggcgg cgccctgaag 2940

ggctatcaga tcaccaataa gttcgagagc ttttaagtcca tgtctacca gaacggcttc 3000

atcttttaca tcctgcctg gctgacatcc aagatcgatc catctaccgg ctttgtgaac 3060

ctgctgaaaa ccaagtatac cagcatcgcc gattccaaga agttcatcag ctctttgac 3120

aggatcatgt acgtgcccga ggaggatctg ttcgagtgtg ccctggacta taagaacttc 3180

tctcgcacag acgccgatta catcaagaag tggaagctgt actcctacgg caaccggatc 3240

agaatcttcc ggaatcctaa gaagaacaac gtgttcgact gggaggaggt gtgcctgacc 3300

agcgctata 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 cgacgccaat ggcgccctata acatcgccag aaaggctgctg 3600

tgggccatcg gccagttcaa gaaggccgag gacgagaagc tggataaggt gaagatcgcc 3660

atctctaaca aggagtggct ggagtacgcc cagaccagcg tgaagcac 3708

<210> 62

<211> 2970

<212> DNA

<213> Artificial Sequence

<220>

<223> CasX

<400> 62

atggcccaa agaagaagcg gaagtcagc atgcaagaga tcaagagaat caacaagatc 60

agaaggagac tggtaagga cagcaacaca aagaaggccg gcaagacagg ccccatgaaa 120

accctgctcg tcagagtgat gaccctgac ctgagagagc ggctggaaaa cctgagaaag 180

aagcccgaga acatccctca gcctatcagc aacaccagca gggccaacct gaacaagctg 240

ctgaccgact acaccgagat gaagaaagcc atcctgcacg tgtactggga agagttccag 300

aaagacccc tgggcctgat gagcagagtt gctcagcccg ctcttaagaa catcgaccag 360

agaaagctga tccccgtgaa ggacggcaac gagagactga cctctagcgg ctttgctgctg 420

agccagtgtt gccagcctct gtacgtgtac aagctggaac aagtgaacga caagggaag	480
ccccacacca actacttcgg cagatgcaac gtgtccgagc acgagaggct gatcctgctg	540
tctcctcaca agcccagagc caacgatgag ctgggtcacat acagcctggg caagttcgga	600
cagagagccc tggacttcta cagcatccac gtgaccaggg agagcaatca ccctgtgaag	660
ccccctggaac agatcggcgg caatagctgt gcctctggac ctgtgggaaa agccctgagc	720
gacgcctgta tgggagccgt ggcatccttc ctgaccaagt accaggacat catcctggaa	780
caccagaaag tgatcaagaa gaacgagaaa agactggcca acctcaagga tatcgccagc	840
gctaaccggc tggcctttcc taagatcacc ctgcctccac agcctcacac caaagagggc	900
atcgaggcct acaacaacgt ggtggcccag atcgtgattt ggttcaacct gaatctgtgg	960
cagaagctga agatcggcag ggacgaagcc aagccactgc agagactgaa gggttcctc	1020
agcttccttc tgggtgaaag acaggccaat gaagtggatt ggtgggacat ggtctgcaac	1080
gtgaagaagc tgatcaacga gaagaaagag gatggcaagg tttcttgga gaacctggcc	1140
ggctacaaga gacaagaagc cctgtgcct tacctgagca gcgaagagga ccggaagaag	1200
ggcaagaagt tcgccagata ccagttcggc gacctgtgc tgcacctgga aaagaagcac	1260
ggcgaggact ggggcaaagt gtacgatgag gcctgggaga gaatcgaca gaagtgga	1320
ggcctgagca agcacattaa gctggaagag gaaagaagga gcgaggacgc ccaatctaaa	1380
gccgctctga ccgattggct gagagccaag gccagctttg tgatcgaggg cctgaaagag	1440
gccgacaagg acgagtctg cagatgcgag ctgaagctgc agaagtggta cggcgatctg	1500
agaggcaagc cttcgccat tgaggccgag aacagcatcc tggacatcag cggcttcagc	1560
aagcagtaca actgcgcctt catttggcag aaagacggcg tcaagaaact gaacctgtac	1620
ctgatcatca attacttcaa aggcggcaag ctgcggttca agaagatcaa acccgaggcc	1680
ttcgaggcta acagattcta caccgtgac acaaaaaagt ccggcgagat cgtgcccattg	1740
gaagtgaact tcaacttga cgacccaac ctgattatcc tgcctctggc cttcggaag	1800
agacaggga gagattcat 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 cctgggcaa ccctacacac	2100
atcctgagaa tcggcgagag ctacaaagag aagcagagga caatccagc cgccaaagag	2160
gtggaacaga gaagagccgg cggatactct aggaagtacg ccagcaagc caagaatctg	2220
gccgacgaca tgggtccgaaa caccgccaga gatctgctgt actacgccgt gacacaggac	2280

gccatgctga tcttcgagaa tctgagcaga ggcttcggcc ggcagggcaa gagaaccttt 2340

atggccgaga ggcagtagac cagaatggaa gattggctca cagctaaact ggcctacgag 2400

ggactgcccc gcaagacctc cctgtccaaa aacttgcccc agtatactc caagacctgc 2460

agcaattgcg gcttcacat caccagcgcc gactacgaca gagtgtgga aaagctcaag 2520

aaaaccgcca cgggtggat gaccaccatc aacggcaaag agctgaaggt tgagggccag 2580

atcacctact acaacaggta caagaggcag aacgtcgtga aggatctgag cgtggaactg 2640

gacagactga gcgaagagag cgtgaacaac gacatcagca gctggacaaa ggcagatca 2700

ggcgaggctc tgagcctgct gaagaagagg ttagccaca gacctgtgca agagaagttc 2760

gtgtgcctga actgcggctt cgagacacac gccgatgaac aggtgcctt gaacattgcc 2820

agaagctggc ttttctgag aagccaagag tacaagaagt accagaccaa caagaccacc 2880

ggcaacaccg acaagagggc ctttgtggaa acctggcaga gcttctacag aaaaaagctg 2940

aaagaagtct ggaagccgc cgtgactagt 2970

<210> 63

<211> 2976

<212> DNA

<213> Campylobacter jejuni

<400> 63

atggcccaa agaagaagcg gaaggtcgcc agaactctgg ccttcgacat cgcatcagc 60

agcatcggtt gggccttcag cgagaacgac gagctgaagg actgcggcgt gcggtcttc 120

accaaggttg aaaaccccaa gaccggcgag agcctggccc tgcccagaag gctggccaga 180

agcgcccgga agagactggc cagacggaag gcccggtga accacctgaa gcacctgac 240

gccaacagat tcaagctgaa ctacaggac taccagagct tcgacgagtc cctggccaag 300

gcctacaagg gcagcctgat cagccctac gagctgcggt tccgggacct gaacgagctg 360

ctgagcaagc aggacttcgc cagagtgatc ctgcacattg ccaagcgag aggtacgac 420

gacatcaaga acagcgacga caaagagaag ggcgcatcc tgaaggccat caagcagaac 480

gaggaaaagc tggccaacta ccagtccttg ggcgagtacc tgtacaaaga gtacttcag 540

aagttcaaag agaacagcaa agaattcacc aacgtgcgga acaagaaaga aagctacgag 600

cgggtgatcg cccagagctt cctgaaggat gagctgaagc tgatcttcaa gaagcagaga 660

gagttcggct tcagcttcag caagaaattc gaggaagagg tgctgagcgt gccttctac 720

aagagagccc tgaaggactt cagccacctc gtgggcaact gcagcttctt caccgacgag 780

aagagagccc ccaagaacag ccccttgccc ttcattgtcg tggccctgac ccggtatc 840

aacctgctga acaatctgaa gaacaccgag ggcatcctgt acaccaagga cgacctgaac	900
gccctgctga 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
gccaaatacg acctgaatca gaaccagatc gacagcctga gcaagctgga attcaaggat	1200
cacctgaaca tcagcttcaa ggctctgaag ctggtcacc cctgatgct ggaaggcaag	1260
aagtacgacg aggcctgcaa cgagctgaac ctgaaggtgg ccatcaacga ggacaagaag	1320
gacttcctgc ccgccttcaa cgaaacctac tacaaggacg aagtgaccaa ccccgtagtg	1380
ctcgccggcca tcaaagaata ccggaagggtg ctgaatgcc tgctcaagaa atacggcaag	1440
gtgcacaaga tcaacatcga gctggcccgg gaagtgggca agaaccacag ccagccggcc	1500
aagatcgaga aagagcagaa cgaaaactac aaggccaaga aggacgctga gctggaatgc	1560
gagaagctgg gactgaagat caacagcaag aacatcctga agctgcggct gttcaaagaa	1620
cagaaagagt tctgcgcta cagccggcag aagatcaaga tcagcgatct gcaggacgag	1680
aagatgctgg aaatcgacca catctacccc tacagccgg 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 atcgcccggc 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 cagcatcgtg aaggccttct ccgacttcaa gaaagaacag	2220
gaaagcaaca gcgccgagct gtacccaag aagatctctg agctggacta caagaacaag	2280
cggaagtctt tcgagccctt cagcggcttc cggcagaagg tgctggataa gatcgacgag	2340
atcttcgtgt ccaagcccga gcggaagaag ccctctggcg ccctgcacga ggaaaccttc	2400
agaaaagagg aagagttcta ccagtcctac ggcggaag aaggcgtgct gaaggccctc	2460
gagctgggca agatcagaaa agtgaacggc aagatcgtga agaacgggga catgttccgg	2520
giggacatct tcaagcaca aaagaccaac aagttctacg ccgtgcccat ctacaccatg	2580

gacttcgccc tgaagggtgct gcccaacaag gccgtggccc ggtccaagaa gggcgagatc 2640

aaggactgga ttctgatgga cgagaactac gagttctgct ttagcctgta caaggactcc 2700

ctgatcctga tccagaccaa ggacatgcag gaacccgagt tcgtctacta caacgccttc 2760

accagcagca ccgtgtccct gatcgtgtct aagcacgaca acaagttcga gacactgagc 2820

aagaaccaga agatcctgtt caagaacgcc aacgagaaag aagtatcgc caagagcatc 2880

ggcatccaga atctgaaggt gttcgagaag tacatcgtgt ccgccctggg agaagtgaca 2940

aaggccgagt tccggcagag agaggacttc aaaaag 2976

<210> 64

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<223> human Cas9 (hCas9) amino acid

<400> 64

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> 65

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<223> nickase Cas9 (nCas9) amino acid

<400> 65

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> 66
 <211> 1069
 <212> PRT
 <213> Staphylococcus aureus
 <400> 66
 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>
67
<211> 1368
<212> PRT
<213> Artificial Sequence
<220>
<223> dead Cas9 (dCas9) amino acid
<400> 67
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> 68

<211> 1069

<212> PRT

<213> Staphylococcus aureus

<400> 68

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> 69

<211> 1236

<212> PRT

<213> Artificial Sequence

<220>

<223> Cpf1

<400> 69

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> 70

<211> 990

<212> PRT

<213> Artificial Sequence

<220>

<223> Casx

<400> 70

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> 71
 <211> 1782
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> hyperactive PiggyBac (PB) transposase nucleic acid
 <400> 71

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 cctcgcccca 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 ttcattcacc agtgcattca gaactacacc	780
cctggcgccc acctgacctat cgacgagcag ctgctgggct tcaggggcag gtgccccttc	840
agggtctata tccccaacaa gccagcaag tacggcatca agatcctgat gatgtgcgac	900
agcggcacca agtacctgat caacggcatg ccctacctgg gcaggggcac ccagaccaac	960
ggcgtgcccc tgggcgagta ctacgtgaag gagctgtcca agcccgcca cggcagctgc	1020
agaaacatca cctgcgacaa ctggttcacc agcatcccc tggccaagaa cctgctgcag	1080
gagccctaca agctgacctat cgtgggcacc gtgagaagca acaagagaga gatccccgag	1140
gtcctgaaga acagcaggtc caggcccggtg ggcaccagca tgttctgctt cgacggcccc	1200
ctgacctgg tgtcctacaa gcccaagccc gccaatgg tgtacctgct gtccagctgc	1260
gacgaggacg ccagcatcaa cgagagcacc ggcaagcccc agatggtgat gtactacaac	1320
cagaccaagg gcggcgtgga caccctggac cagatgtgca gcgtgatgac ctgcagcaga	1380
aagaccaaca ggtggcccat ggccctgctg tacggcatga tcaacatgc ctgcatcaac	1440
agcttcatca tctacagcca caactgagc agcaaggcg agaagggtgca gagccgaaa	1500
aagttcatgc ggaacctgta catgggcctg acctccagct tcatgaggaa gaggtggag	1560
gccccaccc tgaagagata cctgaggac aacatcagca acatcctgcc caaagagggtg	1620
cccggcacca gcgacgacag caccgaggag cccgtgatga agaaggagc ctactgcacc	1680
tactgtccca gcaagatcag aagaaaggcc agcgccagct gcaagaagt taagaaggtc	1740

atctgccggg agcacaacat cgacatgtgc cagagctgtt tc 1782

<210> 72

<211> 1020

<212> DNA

<213> Artificial Sequence

<220>

<223> hyperactive Sleeping Beauty (SB100) transposase nucleic acid

<400> 72

atgggaaaat caaaagaaat cagccaagac ctcaaaaaa gaattgtaga cctccacaag 60

tctggttcat ccttgggagc aatttccaaa cgcttgccgg taccacgttc atctgtacaa 120

acaatagtac gcaagtataa acaccatggg accacgcagc cgtcataccg ctcaaggaagg 180

agacgcgttc tgtctcctag agatgaacgt actttggtgc gaaaagtga 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 aaatgtcttc tggctctgat aaacaaaaat agaactgttt 480

ggccataatg accatcgtaa tgtttggagg aagaaggggg aggcttgcaa gccgaagaac 540

accatcccaa ccgtgaagca cgggggtggc agcatcatgt tgtgggggtg ctttgctgca 600

ggagggactg gtgcacttca caaaatagat ggcatcatgg acgccgtgca gtatgtggat 660

atattgaagc aacatctcaa gacatcagtc aggaagttaa agcttggtcg caaatgggtc 720

ttccaacacg acaatgaccc caagcatact tccaaagtgt tggcaaaatg gcttaaggac 780

aacaaagtca aggtattgga gtggccatca caaagccctg acctcaatcc tatagaaaat 840

ttgtgggcag aactgaaaaa gcgtgtgcga gcaaggaggc ctacaaacct gactcagtta 900

caccagctct gtcaggagga atgggccaaa attcacccaa attattgtgg gaagcttgtg 960

gaaggtacc cgaaacgttt gaccaagtt aaacaattta aaggcaatgc taccaaatac 1020

<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> 3123

<212> PRT

<213> homo sapiens

<400> 74

Met Pro Gly Ala Ala Gly Val Leu Leu Leu Leu Leu Ser Gly Gly
 1 5 10 15
 Leu Gly Gly Val Gln Ala Gln Arg Pro Gln Gln Gln Arg Gln Ser Gln
 20 25 30

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

His Pro Ile Thr Asn Ala Ile Asp Gly Lys Asn Thr Trp Trp Gln Ser
 100 105 110

Pro Ser Ile Lys Asn Gly Ile Glu Tyr His Tyr Val Thr Ile Thr Leu
115 120 125

Asp Leu Gln Gln Val Phe Gln Ile Ala Tyr Val Ile Val Lys Ala Ala
130 135 140

Asn Ser Pro Arg Pro Gly Asn Trp Ile Leu Glu Arg Ser Leu Asp Asp
145 150 155 160

Val Glu Tyr Lys Pro Trp Gln Tyr His Ala Val Thr Asp Thr Glu Cys
165 170 175

Leu Thr Leu Tyr Asn Ile Tyr Pro Arg Thr Gly Pro Pro Ser Tyr Ala
180 185 190

Lys Asp Asp Glu Val Ile Cys Thr Ser Phe Tyr Ser Lys Ile His Pro
195 200 205

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

Ala Asp Asp Pro Ser Pro Glu Leu Leu Glu Phe Thr Ser Ala Arg Tyr
225 230 235 240

Ile Arg Leu Arg Phe Gln Arg Ile Arg Thr Leu Asn Ala Asp Leu Met
245 250 255

Met Phe Ala His Lys Asp Pro Arg Glu Ile Asp Pro Ile Val Thr Arg
260 265 270

Arg Tyr Tyr Tyr Ser Val Lys Asp Ile Ser Val Gly Gly Met Cys Ile
275 280 285

Cys Tyr Gly His Ala Arg Ala Cys Pro Leu Asp Pro Ala Thr Asn Lys
290 295 300

Ser Arg Cys Glu Cys Glu His Asn Thr Cys Gly Asp Ser Cys Asp Gln
305 310 315 320

Cys Cys Pro Gly Phe His Gln Lys Pro Trp Arg Ala Gly Thr Phe Leu
325 330 335

Thr Lys Thr Glu Cys Glu Ala Cys Asn Cys His Gly Lys Ala Glu Glu
340 345 350

Cys Tyr Tyr Asp Glu Asn Val Ala Arg Arg Asn Leu Ser Leu Asn Ile

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

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

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

 Leu Thr Asp Leu Pro Gly Arg Ile Arg Val Ala Pro Gln Gln Asp Asp
 545 550 555 560
 Leu Asp Ser Pro Gln Gln Ile Ser Ile Ser Asn Ala Glu Ala Arg Gln
 565 570 575
 Ala Leu Pro His Ser Tyr Tyr Trp Ser Ala Pro Ala Pro Tyr Leu Gly
 580 585 590
 Asn Lys Leu Pro Ala Val Gly Gly Gln Leu Thr Phe Thr Ile Ser Tyr
 595 600 605

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

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

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

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

850 855 860

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

 Val Cys His Ser Gln Thr Gly Gln Cys Glu Cys Arg Ala Asn Val Gln
 930 935 940
 Gly Gln Arg Cys Asp Lys Cys Lys Ala Gly Thr Phe Gly Leu Gln Ser
 945 950 955 960
 Ala Arg Gly Cys Val Pro Cys Asn Cys Asn Ser Phe Gly Ser Lys Ser
 965 970 975
 Phe Asp Cys Glu Glu Ser Gly Gln Cys Trp Cys Gln Pro Gly Val Thr
 980 985 990

 Gly Lys Lys Cys Asp Arg Cys Ala His Gly Tyr Phe Asn Phe Gln Glu
 995 1000 1005
 Gly Gly Cys Thr Ala Cys Glu Cys Ser His Leu Gly Asn Asn Cys Asp
 1010 1015 1020
 Pro Lys Thr Gly Arg Cys Ile Cys Pro Pro Asn Thr Ile Gly Glu Lys
 1025 1030 1035 1040
 Cys Ser Lys Cys Ala Pro Asn Thr Trp Gly His Ser Ile Thr Thr Gly
 1045 1050 1055

 Cys Lys Ala Cys Asn Cys Ser Thr Val Gly Ser Leu Asp Phe Gln Cys
 1060 1065 1070
 Asn Val Asn Thr Gly Gln Cys Asn Cys His Pro Lys Phe Ser Gly Ala
 1075 1080 1085
 Lys Cys Thr Glu Cys Ser Arg Gly His Trp Asn Tyr Pro Arg Cys Asn
 1090 1095 1100

Leu Cys Asp Cys Phe Leu Pro Gly Thr Asp Ala Thr Thr Cys Asp Ser
1105 1110 1115 1120

Glu Thr Lys Lys Cys Ser Cys Ser Asp Gln Thr Gly Gln Cys Thr Cys
1125 1130 1135

Lys Val Asn Val Glu Gly Ile His Cys Asp Arg Cys Arg Pro Gly Lys
1140 1145 1150

Phe Gly Leu Asp Ala Lys Asn Pro Leu Gly Cys Ser Ser Cys Tyr Cys
1155 1160 1165

Phe Gly Thr Thr Thr Gln Cys Ser Glu Ala Lys Gly Leu Ile Arg Thr
1170 1175 1180

Trp Val Thr Leu Lys Ala Glu Gln Thr Ile Leu Pro Leu Val Asp Glu
1185 1190 1195 1200

Ala Leu Gln His Thr Thr Thr Lys Gly Ile Val Phe Gln His Pro Glu
1205 1210 1215

Ile Val Ala His Met Asp Leu Met Arg Glu Asp Leu His Leu Glu Pro
1220 1225 1230

Phe Tyr Trp Lys Leu Pro Glu Gln Phe Glu Gly Lys Lys Leu Met Ala
1235 1240 1245

Tyr Gly Gly Lys Leu Lys Tyr Ala Ile Tyr Phe Glu Ala Arg Glu Glu
1250 1255 1260

Thr Gly Phe Ser Thr Tyr Asn Pro Gln Val Ile Ile Arg Gly Gly Thr
1265 1270 1275 1280

Pro Thr His Ala Arg Ile Ile Val Arg His Met Ala Ala Pro Leu Ile
1285 1290 1295

Gly Gln Leu Thr Arg His Glu Ile Glu Met Thr Glu Lys Glu Trp Lys
1300 1305 1310

Tyr Tyr Gly Asp Asp Pro Arg Val His Arg Thr Val Thr Arg Glu Asp
1315 1320 1325

Phe Leu Asp Ile Leu Tyr Asp Ile His Tyr Ile Leu Ile Lys Ala Thr
1330 1335 1340

Tyr Gly Asn Phe Met Arg Gln Ser Arg Ile Ser Glu Ile Ser Met Glu

1345 1350 1355 1360
 Val Ala Glu Gln Gly Arg Gly Thr Thr Met Thr Pro Pro Ala Asp Leu
 1365 1370 1375

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

 His His Thr Ala Gly Asp Phe Cys Glu Arg Cys Ala Leu Gly Tyr Tyr
 1445 1450 1455
 Gly Ile Val Lys Gly Leu Pro Asn Asp Cys Gln Gln Cys Ala Cys Pro
 1460 1465 1470
 Leu Ile Ser Ser Ser Asn Asn Phe Ser Pro Ser Cys Val Ala Glu Gly
 1475 1480 1485
 Leu Asp Asp Tyr Arg Cys Thr Ala Cys Pro Arg Gly Tyr Glu Gly Gln
 1490 1495 1500

 Tyr Cys Glu Arg Cys Ala Pro Gly Tyr Thr Gly Ser Pro Gly Asn Pro
 1505 1510 1515 1520
 Gly Gly Ser Cys Gln Glu Cys Glu Cys Asp Pro Tyr Gly Ser Leu Pro
 1525 1530 1535
 Val Pro Cys Asp Pro Val Thr Gly Phe Cys Thr Cys Arg Pro Gly Ala
 1540 1545 1550
 Thr Gly Arg Lys Cys Asp Gly Cys Lys His Trp His Ala Arg Glu Gly
 1555 1560 1565

 Trp Glu Cys Val Phe Cys Gly Asp Glu Cys Thr Gly Leu Leu Leu Gly
 1570 1575 1580
 Asp Leu Ala Arg Leu Glu Gln Met Val Met Ser Ile Asn Leu Thr Gly
 1585 1590 1595 1600

Pro Leu Pro Ala Pro Tyr Lys Met Leu Tyr Gly Leu Glu Asn Met Thr

1605 1610 1615

Gln Glu Leu Lys His Leu Leu Ser Pro Gln Arg Ala Pro Glu Arg Leu

1620 1625 1630

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

1635 1640 1645

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

1650 1655 1660

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

1665 1670 1675 1680

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

1685 1690 1695

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

1700 1705 1710

Asn Leu Glu Gly Leu Gln Lys Glu Ile Asp Gln Met Ile Lys Glu Leu

1715 1720 1725

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

1730 1735 1740

Val Ala Ala Glu Ala Leu Leu Lys Lys Val Lys Lys Leu Phe Gly Glu

1745 1750 1755 1760

Ser Arg Gly Glu Asn Glu Glu Met Glu Lys Asp Leu Arg Glu Lys Leu

1765 1770 1775

Ala Asp Tyr Lys Asn Lys Val Asp Asp Ala Trp Asp Leu Leu Arg Glu

1780 1785 1790

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

1795 1800 1805

Lys Asn Met Thr Ala Leu Glu Lys Lys Lys Glu Ala Val Glu Ser Gly

1810 1815 1820

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

1825 1830 1835 1840

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

1845	1850	1855	
Glu Asp Ile Gln Thr Lys Leu Pro Pro Met Ser Glu Glu Leu Asn Asp			
1860	1865	1870	
Lys Ile Asp Asp Leu Ser Gln Glu Ile Lys Asp Arg Lys Leu Ala Glu			
1875	1880	1885	
Lys Val Ser Gln Ala Glu Ser His Ala Ala Gln Leu Asn Asp Ser Ser			
1890	1895	1900	
Ala Val Leu Asp Gly Ile Leu Asp Glu Ala Lys Asn Ile Ser Phe Asn			
1905	1910	1915	1920
Ala Thr Ala Ala Phe Lys Ala Tyr Ser Asn Ile Lys Asp Tyr Ile Asp			
1925	1930	1935	
Glu Ala Glu Lys Val Ala Lys Glu Ala Lys Asp Leu Ala His Glu Ala			
1940	1945	1950	
Thr Lys Leu Ala Thr Gly Pro Arg Gly Leu Leu Lys Glu Asp Ala Lys			
1955	1960	1965	
Gly Cys Leu Gln Lys Ser Phe Arg Ile Leu Asn Glu Ala Lys Lys Leu			
1970	1975	1980	
Ala Asn Asp Val Lys Glu Asn Glu Asp His Leu Asn Gly Leu Lys Thr			
1985	1990	1995	2000
Arg Ile Glu Asn Ala Asp Ala Arg Asn Gly Asp Leu Leu Arg Thr Leu			
2005	2010	2015	
Asn Asp Thr Leu Gly Lys Leu Ser Ala Ile Pro Asn Asp Thr Ala Ala			
2020	2025	2030	
Lys Leu Gln Ala Val Lys Asp Lys Ala Arg Gln Ala Asn Asp Thr Ala			
2035	2040	2045	
Lys Asp Val Leu Ala Gln Ile Thr Glu Leu His Gln Asn Leu Asp Gly			
2050	2055	2060	
Leu Lys Lys Asn Tyr Asn Lys Leu Ala Asp Ser Val Ala Lys Thr Asn			
2065	2070	2075	2080
Ala Val Val Lys Asp Pro Ser Lys Asn Lys Ile Ile Ala Asp Ala Asp			
2085	2090	2095	

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

2100 2105 2110

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

2115 2120 2125

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

2130 2135 2140

Lys Val Ser Val Ser Ser Gly Gly Asp Cys Ile Arg Thr Tyr Lys Pro

2145 2150 2155 2160

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

2165 2170 2175

Ala Val Ala Asp Asn Leu Leu Phe Tyr Leu Gly Ser Ala Lys Phe Ile

2180 2185 2190

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

2195 2200 2205

Asp Val Gly Ser Gly Val Gly Arg Val Glu Tyr Pro Asp Leu Thr Ile

2210 2215 2220

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

2225 2230 2235 2240

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

2245 2250 2255

Pro Ser Thr His His Ser Thr Ser Pro Pro Gly Tyr Thr Ile Leu Asp

2260 2265 2270

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

2275 2280 2285

Lys Lys Ala Asp Ala Val Arg Val Ile Thr Phe Thr Gly Cys Met Gly

2290 2295 2300

Glu Thr Tyr Phe Asp Asn Lys Pro Ile Gly Leu Trp Asn Phe Arg Glu

2305 2310 2315 2320

Lys Glu Gly Asp Cys Lys Gly Cys Thr Val Ser Pro Gln Val Glu Asp

2325 2330 2335

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

2340 2345 2350
 Arg Pro Ile Arg Trp Tyr Pro Asn Ile Ser Thr Val Met Phe Lys Phe
 2355 2360 2365
 Arg Thr Phe Ser Ser Ser Ala Leu Leu Met Tyr Leu Ala Thr Arg Asp
 2370 2375 2380
 Leu Arg Asp Phe Met Ser Val Glu Leu Thr Asp Gly His Ile Lys Val
 2385 2390 2395 2400

 Ser Tyr Asp Leu Gly Ser Gly Met Ala Ser Val Val Ser Asn Gln Asn
 2405 2410 2415
 His Asn Asp Gly Lys Trp Lys Ser Phe Thr Leu Ser Arg Ile Gln Lys
 2420 2425 2430
 Gln Ala Asn Ile Ser Ile Val Asp Ile Asp Thr Asn Gln Glu Glu Asn
 2435 2440 2445
 Ile Ala Thr Ser Ser Ser Gly Asn Asn Phe Gly Leu Asp Leu Lys Ala
 2450 2455 2460

 Asp Asp Lys Ile Tyr Phe Gly Gly Leu Pro Thr Leu Arg Asn Leu Ser
 2465 2470 2475 2480
 Met Lys Ala Arg Pro Glu Val Asn Leu Lys Lys Tyr Ser Gly Cys Leu
 2485 2490 2495
 Lys Asp Ile Glu Ile Ser Arg Thr Pro Tyr Asn Ile Leu Ser Ser Pro
 2500 2505 2510
 Asp Tyr Val Gly Val Thr Lys Gly Cys Ser Leu Glu Asn Val Tyr Thr
 2515 2520 2525

 Val Ser Phe Pro Lys Pro Gly Phe Val Glu Leu Ser Pro Val Pro Ile
 2530 2535 2540
 Asp Val Gly Thr Glu Ile Asn Leu Ser Phe Ser Thr Lys Asn Glu Ser
 2545 2550 2555 2560
 Gly Ile Ile Leu Leu Gly Ser Gly Gly Thr Pro Ala Pro Pro Arg Arg
 2565 2570 2575
 Lys Arg Arg Gln Thr Gly Gln Ala Tyr Tyr Val Ile Leu Leu Asn Arg
 2580 2585 2590

Gly Arg Leu Glu Val His Leu Ser Thr Gly Ala Arg Thr Met Arg Lys

2595 2600 2605

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

2610 2615 2620

Ser Val His Val Glu Arg Thr Arg Gly Ile Phe Thr Val Gln Val Asp

2625 2630 2635 2640

Glu Asn Arg Arg Tyr Met Gln Asn Leu Thr Val Glu Gln Pro Ile Glu

2645 2650 2655

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

2660 2665 2670

Pro Leu Arg Asn Ile Pro Pro Phe Glu Gly Cys Ile Trp Asn Leu Val

2675 2680 2685

Ile Asn Ser Val Pro Met Asp Phe Ala Arg Pro Val Ser Phe Lys Asn

2690 2695 2700

Ala Asp Ile Gly Arg Cys Ala His Gln Lys Leu Arg Glu Asp Glu Asp

2705 2710 2715 2720

Gly Ala Ala Pro Ala Glu Ile Val Ile Gln Pro Glu Pro Val Pro Thr

2725 2730 2735

Pro Ala Phe Pro Thr Pro Thr Pro Val Leu Thr His Gly Pro Cys Ala

2740 2745 2750

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

2755 2760 2765

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

2770 2775 2780

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

2785 2790 2795 2800

Leu Leu Phe Tyr Met Ala Arg Ile Asn His Ala Asp Phe Ala Thr Val

2805 2810 2815

Gln Leu Arg Asn Gly Leu Pro Tyr Phe Ser Tyr Asp Leu Gly Ser Gly

2820 2825 2830

Asp Thr His Thr Met Ile Pro Thr Lys Ile Asn Asp Gly Gln Trp His

2835	2840	2845	
Lys Ile Lys Ile Met Arg Ser Lys Gln Glu Gly Ile Leu Tyr Val Asp			
2850	2855	2860	
Gly Ala Ser Asn Arg Thr Ile Ser Pro Lys Lys Ala Asp Ile Leu Asp			
2865	2870	2875	2880
Val Val Gly Met Leu Tyr Val Gly Gly Leu Pro Ile Asn Tyr Thr Thr			
	2885	2890	2895
Arg Arg Ile Gly Pro Val Thr Tyr Ser Ile Asp Gly Cys Val Arg Asn			
2900	2905	2910	
Leu His Met Ala Glu Ala Pro Ala Asp Leu Glu Gln Pro Thr Ser Ser			
2915	2920	2925	
Phe His Val Gly Thr Cys Phe Ala Asn Ala Gln Arg Gly Thr Tyr Phe			
2930	2935	2940	
Asp Gly Thr Gly Phe Ala Lys Ala Val Gly Gly Phe Lys Val Gly Leu			
2945	2950	2955	2960
Asp Leu Leu Val Glu Phe Glu Phe Arg Thr Thr Thr Thr Thr Gly Val			
2965	2970	2975	
Leu Leu Gly Ile Ser Ser Gln Lys Met Asp Gly Met Gly Ile Glu Met			
2980	2985	2990	
Ile Asp Glu Lys Leu Met Phe His Val Asp Asn Gly Ala Gly Arg Phe			
2995	3000	3005	
Thr Ala Val Tyr Asp Ala Gly Val Pro Gly His Leu Cys Asp Gly Gln			
3010	3015	3020	
Trp His Lys Val Thr Ala Asn Lys Ile Lys His Arg Ile Glu Leu Thr			
3025	3030	3035	3040
Val Asp Gly Asn Gln Val Glu Ala Gln Ser Pro Asn Pro Ala Ser Thr			
3045	3050	3055	
Ser Ala Asp Thr Asn Asp Pro Val Phe Val Gly Gly Phe Pro Asp Asp			
3060	3065	3070	
Leu Lys Gln Phe Gly Leu Thr Thr Ser Ile Pro Phe Arg Gly Cys Ile			
3075	3080	3085	

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

3090

3095

3100

Phe Ala Lys Ala Leu Glu Leu Arg Gly Val Gln Pro Val Ser Cys Pro

3105

3110

3115

3120

Ala Asn Ser

<210> 75

<211> 9369

<212> DNA

<213> homo sapiens

<400> 75

atgccgggag ccgccgggggt cctcctcctt ctgctgctct ccggaggcct cgggggctga 60

caggcgacaga ggccgcagca gcagcggcag tcacaggcac atcagcaaag aggtttattc 120

cctgctgtcc tgaatcttgc ttctaatgct ctatcacga ccaatgcaac atgtggagaa 180

aaaggacctg aaatgtactg caaatggta gaacatgtcc ctgggcagcc tgtgaggaac 240

ccgcagtgtc gaatctgcaa tcaaaacagc agcaatccaa accagagaca cccgattaca 300

aatgctattg atggaaagaa cacttggtgg cagagtccca gtattaagaa tggaatcgaa 360

taccattatg tgacaattac cctggattta cagcaggtgt tccagatcgc gtatgtgatt 420

gtgaaggcag ctaactcccc ccggcctgga aactggattt tggaacgctc tcttgatgat 480

gttgaataca agccctggca gtatcatgct gtgacagaca cggagtgcct aacgctttac 540

aatatttacc cccgcactgg gccaccgtca tatgccaaag atgatgaggt catctgcact 600

tcattttact ccaagataca ccccttagaa aatggagaga ttcacatctc tttaataaat 660

gggagaccaa gtgccgatga tccttctcca gaactgctag aatttacctc cgctcgctat 720

attcgctga gatttcagag gatccgcaca ctgaatgctg acttgatgat gtttgctcac 780

aaagacccaa gagaaattga ccccatgtc accagaagat attactactc ggtcaaggat 840

atttcagtgt gagggatgtg catctgctat ggtcatgcca gggcttgctc acttgatcca 900

gcgacaaata aatctcgctg tgagtgtgag cataacacat gtggcgatag ctgtgatcag 960

tgctgtccag gattccatca gaaaccctgg agagctggaa cttttctaac taaaactgaa 1020

tgtgaagcat gcaattgtca tggaaaagct gaagaatgct attatgatga aaatgttgcc 1080

agaagaaatc tgagtttgaa tatacgtgga aagtacattg gagggggtgt ctgcattaat 1140

tgtacccaaa aactgtctgg tataaactgc gagacatgta ctgatggctt cttcagaccc 1200

aaaggggtat ctccaaatta tccaaggcca tgccagccat gtcattgcga tccaattggt	1260
tccttaaatg aagtctgtgt caaggatgag aaacatgctc gacgaggttt ggcacctgga	1320
tcctgtcatt gcaaaactgg ttttggaggt gtgagctgtg atcggtgtgc caggggctac	1380
actggctacc cggactgcaa agccigtAAC tgCagtgggt tagggagcaa aaatgaggat	1440
ccttgttttg gccctgtat ctgcaaggaa aatgttgaag gaggagactg tagtcgttgc	1500
aaatccggct tcttcaattt gcaagaggat aattggaaag gctgcgatga gtgtttctgt	1560
tcaggggttt caaacagatg tcagagtcc tactggacct atggcaaat acaagatatg	1620
agtggctggt atctgactga ctttcttggc cgcattcgag tggctccca gcaggacgac	1680
ttggactcac ctCagcagat cagcatcagt aacgcggagg cccggcaagc cctgccgcac	1740
agctactact ggagcgcgcc ggctccctat ctgggaaca aactcccagc agtaggagga	1800
cagttgacat ttaccatata atatgacctt gaagaaggagg aagaagatac agaacgtgtt	1860
ctccagctta tgattatctt agagggtaat gacttgagca tcagcacagc ccaagatgag	1920
gtgtacctgc acccatctga agaacatact aatgtattgt tacttaaaga agaatcattt	1980
accatacatg gcacacattt tccagtccgt agaaaggaat ttatgacagt gcttgCGaat	2040
ttgaagagag tcctcttaca aatcacatac agctttggga tggatgccat cttcaggttg	2100
agctctgtta accttgaatc cgctgtctcc tatectactg atggaagcat tgcagcagct	2160
gtagaagtgt gtcagtgcc accagggtat actggctcct cttgtgaatc ttgttggcct	2220
aggcacaggc gagttaacgg cactatTTTT ggtggcatct gtgagccatg tcagtgtttt	2280
ggtcatgcgg agtcctgtga tgacgtcact ggagaatgcc tgaactgtaa ggatcacaca	2340
ggtggcccat attgtataa atgtcttctt ggtttctatg gcgagcctac taaaggaacc	2400
tctgaagact gtcaacctg tgctgtcca ctcaatatcc catccaataa ctttagccca	2460
acgtgccatt tagaccggag tcttggattg atctgtgatg gatgccctgt cgggtacaca	2520
ggaccacgct gtgagagggtg tgcagaaggc tattttggac aacctctgt acctggagga	2580
tcatgtcagc catgccaatg caatgacaac ctgacttct ccatccctgg cagctgtgac	2640
agcttgtctg gtcctgtct gatatgtaaa ccaggtaaa caggccggtA ctgtgagctc	2700
tgtgctgatg gatattttgg agatgcagtt gatgcgaaga actgtcagcc ctgtcgtgt	2760
aatgccggtg gctctttctc tgaggtttgc cacagtcaaa ctggacagtg tgagtgcaga	2820
gccaacgttc agggtcagag atgtgacaaa tgcaaggctg ggacctttgg cctacaatca	2880
gcaaggggct gtgttcctg caactgcaat tcttttgggt ctaagtcatt cgactgtgaa	2940
gagagtggac aatgttggtg ccaacctgga gtcacaggga agaaatgtga ccgctgtgcc	3000
cacggctatt tcaacttcca agaaggaggc tgcacagctt gtgaatgttc tcacttgggt	3060

aataattgtg acccaaagac tgggcgatgc atttgccttc ccaataccat tggagagaaa	3120
tgttctaata gtgcacccaa tacctggggc cacagcatta ccactggttg taaggcttgt	3180
aactgcagca cagtgggac cttggatttc caatgcaatg taaatacagg ccaatgcaac	3240
tgtcatccaa aattctctgg tgcaaaatgt acagagtga gtcgaggtca ctggaactac	3300
cctcgctgca atctctgtga ctgcttcttc cctgggacag atgccacaac ctgtgattca	3360
gagactaaaa aatgctcctg tagtgatcaa actgggcagt gcacttgtaa ggtgaatgtg	3420
gaaggcatcc actgtgacag atgccggcct ggcaaattcg gactcgatgc caagaatcca	3480
cttggctgca gcagctgcta ttgcttcggc actactaccc agtgctctga agcaaaagga	3540
ctgatccgga cgtgggtgac tctgaaggct gagcagacca ttctaccctt ggtagatgag	3600
gctctgcagc acacgaccac caagggcatt gtttttcaac atccagagat tgttgccac	3660
atggacctga tgagagaaga tctccatttg gaaccttttt attggaaact tccagaacaa	3720
tttgaaggaa agaagttgat ggcctatggg ggcaaaactca agtatgcaat ctatttcgag	3780
gctcgggaag aaacaggttt ctctacatat aatcctcaag tgatcattcg aggtgggaca	3840
cctactcatg ctagaattat cgtcaggcat atggctgctc ctctgattgg ccaattgaca	3900
aggcatgaaa ttgaaatgac agagaaagaa tggaaatatt atggggatga tcctcgagtc	3960
catagaactg tgacccgaga agacttcttg gatatactat atgatattca ttacattctt	4020
atcaaagcta cttatggaaa tttcatgcga caaagcagga tttctgaaat ctcaatggag	4080
gtagctgaac aaggacgtgg aacaacaatg actcctccag ctgacttgat tgaaaaatgt	4140
gattgtcccc tgggctattc tggcctgtcc tgtgaggcat gcttgccggg attttatcga	4200
ctgcgttctc aaccagggtgg ccgcaccctt ggaccaaccc tgggcacctg tgttccatgt	4260
caatgtaatg gacacagcag cctgtgtgac cctgaaacat cgatatgcca gaattgtcaa	4320
catcacactg ctggtgactt ctgtgaacga tgtgctcttg gatactatgg aattgtcaag	4380
ggattgcaa atgactgtca gcaatgtgcc tgcctctga tttcttcag taacaatttc	4440
agccctcttt gtgtgcaga aggacttgac gactaccgct gcacggcttg tccacgggga	4500
tatgaaggcc agtactgtga aagggtgtcc cctggctata ctggcagtc aggcaaccct	4560
ggaggtcctt gccagaatg tgagtgtgat ccctatggct cactgcctgt gccctgtgac	4620
cctgtcacag gattctgcac gtgccgacct ggagccacgg gaaggaagtg tgacggctgc	4680
aagcactggc atgcacgcga gggctgggag tgtgtttttt gtggagatga gtgcactggc	4740
cttcttctcg gtgacttggc tcgcctggag cagatggtca tgagcatcaa cctcactggt	4800

ccgtgcctg cgccatataa aatgctgtat ggtcttgaaa atatgactca ggagctaaag 4860

cacttgctgt cacctcagcg ggccccagag aggcttattc agctggcaga gggcaatctg 4920

aatacactcg tgaccgaaat gaacgagctg ctgaccaggg ctaccaaagt gacagcagat 4980

ggcgagcaga ccggacagga tgctgagagg accaacacaa gagcaaagtc cctgggagaa 5040

ttcattaagg agcttgcccg ggatgcagaa gctgtaaatg aaaaagctat aaaactaaat 5100

gaaactctag gaactcgaga tgaggccttt gagagaaatt tggaagggtc tcagaaagag 5160

attgaccaga tgattaaaga actgaggagg aaaaatctag agacacaaaa ggaaattgct 5220

gaagatgagt tggtagctgc agaagccctt ctgaaaaaag tgaagaagct gtttgagag 5280

tcccggggggg aaaatgaaga aatggagaag gatctccggg aaaaactggc tgactacaaa 5340

aacaaagttg atgatgcttg ggaccttttg agagaagcca cagataaaat cagagaagct 5400

aatcgccat ttgcagtaaa tcagaaaaac atgactgcat tggagaaaaa gaaggaggct 5460

gttgaagcg gcaaacgaca aattgagaac actttaaaag agggcaatga catactgat 5520

gaagccaacc gtcttcgaga tgaaatcaac tccatcatag actatgttga agacatccaa 5580

actaaattgc cacctatgtc tgaggagctt aatgataaaa tagatgacct ctcccaagaa 5640

ataaaggaca ggaaacttgc tgagaagggtg tcccaggctg agagccacgc agctcagttg 5700

aatgactcat ctgctgtcct tgatggaatc cttgatgagg ctaaaaacat ctcttcaat 5760

gccactgcag ccttcaaagc ctacagcaat attaggact atattgatga agctgagaaa 5820

gttgccaaag aagccaaaga tcttgacat gaagctacaa aactggcaac aggtcctcgg 5880

ggtttattaa aggaagatgc caaaggctgt cttcagaaaa gtttcaggat tcttaacgaa 5940

gccaaagaagt tagcaaatga tgtaaaagaa aatgaagacc atctaaatgg cttaaaaacc 6000

aggatagaaa atgctgatgc tagaaatggg gatctcttga gaactttgaa tgacactttg 6060

ggaaagttaa cagctattcc aaatgataca gctgctaaac tgcaagctgt taaggacaaa 6120

gccagacaag ccaacgacac agctaaagat gtactggcac agattacaga gctccaccag 6180

aacctcgatg gcctgaagaa gaattacaat aaactagcag acagcgtcgc caaaacgaat 6240

gctgtggtta aagatccttc caagaacaaa atcattgccg atgcagatgc cactgtcaaa 6300

aatttagaac aggaagctga ccggctaata gataaactca aacctatcaa ggaacttgag 6360

gataacctaa agaaaaacat ctctgagata aaggaattga taaaccaagc tcggaaacaa 6420

gccaatctta tcaaagtatc tgtgtcttca ggaggtgact gcattcgaac atacaaacca 6480

gaaatcaaga aaggaagtta caataatatt gttgtcaacg taaagacagc tgttgctgat 6540

aacctcctct tttatcttgg aagtgccaaa ttatttgact ttctggctat agaaatgcgt	6600
aaaggcaaag tcagcttctt ctgggatgtt ggatctggag ttggacgtgt agagtacca	6660
gatttgacta ttgatgactc atattggtac cgtatcgtag catcaagaac tgggagaaat	6720
ggaactatct ctgtgagagc cctggatgga cccaaagcca gcattgtgcc cagcacacac	6780
cattcgacgt cccctccagg gtacacgatt ctagatgtgg atgcaaatgc aatgctgttt	6840
gttggtggcc tgactgggaa attaaagaag gctgatgctg tacgtgtgat tacattcact	6900
ggctgcatgg gagaacata ctttgacaac aaacctatag gtttgtggaa tttccgagaa	6960
aaagaaggtg actgcaaagg atgcactgtc agtcctcagg tggaagatag tgaggggact	7020
attcaatttg atggagaagg ttatgcattg gtcagccgtc ccattcgctg gtacccaac	7080
atctccactg tcatgttcaa gttcagaaca ttttcttga gtgcccttct gatgtatctt	7140
gccacacgag acctgagaga tttcatgagt gtggagctca ctgatgggca cataaaagtc	7200
agttacgac tggtctcagg aatggcttcc gttgtcagca atcaaaacca taatgatggg	7260
aaatggaaat cattcactct gtcaagaatt caaaaacaag ccaatatac aattgtagat	7320
atagatacta atcaggagga gaatatagca acttcgtctt ctggaaacaa ctttggtctt	7380
gacttgaaag cagatgacaa aatatatttt ggtggcctgc caacgctgag aaacttgagt	7440
atgaaagcaa ggccagaagt aaatctgaag aaatatccg gctgcctcaa agatatgaa	7500
atttcaagaa ctccgtacaa tatactcagt agtcccgatt atgttggtgt taccaaagga	7560
tgttccctgg agaattgtta cacagttagc tttcctaagc ctggttttgt ggagctctcc	7620
cctgtgccaa ttgatgtagg aacagaaatc aacctgtcat tcagcaccaa gaatgagtc	7680
ggcatcattc ttttgggaag tggagggaca ccagcaccac ctaggagaaa acgaaggcag	7740
actggacagg cctattatgt aatactctc aacaggggcc gtctggaagt gcattctctc	7800
acaggggcac gaacaatgag gaaaattgtg atcagaccag agccgaatct gtttcatgat	7860
ggaagagaac attccgttca tgtagagcga actagaggca tctttacagt tcaagtggat	7920
gaaaacagaa gatacatgca aaacctgaca gttgaacagc ctatcgaagt taaaaaactt	7980
ttcgttgggg gtgctccacc tgaatttcaa ctttcccccac tcagaaatat tcctcctttt	8040
gaaggctgca tatggaatct tggttattaac tctgtcccca tggactttgc aaggcctgtg	8100
tccttcaaaa atgctgacat tggctcgtgt gcccatcaga aactccgtga agatgaagat	8160
ggagcagctc cagctgaaat agttatccag cctgagccag ttcccacccc agcctttcct	8220
acgccacccc cagttctgac acatggtcct tgtgtgcag aatcagaacc agctcttttg	8280
atagggagca agcagttcgg gctttcaaga aacagtcaca ttgcaattgc atttgatgac	8340
accaaagtta aaaaccgcct cacaattgag ttggaagtaa gaaccgaagc tgaatccggc	8400

ttgctttttt acatggctcg catcaatcat gctgattttg caacagttca gctgagaaat	8460
ggattgccct acttcagcta tgacttgggg agtggggaca cccacacat gatccccacc	8520
aaaatcaatg atggccagtgc gcacaagatt aagataatga gaagtaagca agaaggaatt	8580
ctttatgtag atggggcttc caacagaacc atcagtccca aaaaagccga catcctggat	8640
gtcgtgggaa tgctgtatgt tgggtgggta cccatcaact acactaccg aagaattggt	8700
ccagtgcct atagcattga tggctgcgtc aggaatctcc acatggcaga ggccctgcc	8760
gatctggaac aaccacctc cagcttccat gttgggacat gttttgcaa tgctcagagg	8820
ggaacatatt ttgacggaac cggttttgcc aaagcagttg gtggattcaa agtgggattg	8880
gaccttcttg tagaatttga attccgcaca actacaacga ctggagtctc tctggggatc	8940
agtagtcaa aaatggatgg aatgggtatt gaaatgattg atgaaaagt gatgtttcat	9000
gtggacaatg gtgcgggcag attcactgct gtctatgatg ctgggggttc agggcatttg	9060
tgtgatggac aatggcataa agtcactgcc aacaagatca aacaccgcat tgagctcaca	9120
gtcgatggga accaggtgga agcccaaagc ccaaaccag catctacatc agctgacaca	9180
aatgacctg tgtttgttgg aggtctccca gatgacctca agcagtttgg cctaacaacc	9240
agtattccgt tccgaggttg catcagatcc ctgaagctca ccaaaggcac aggcaagcca	9300
ctggaggtta attttgcaa ggccctggaa ctgaggggcg ttcaacctgt atcatgcca	9360
gccaactca	9369
<210> 76	
<211> 508	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> CMV promoter	
<400> 76	
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt	60
gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc attgacgtca	120
atgggtggag tatttacgtt aaactgcca ctgggcagta catcaagtgt atcatatgcc	180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta	240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	360
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420

ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt 480

acggtgggag gtctatataa gcagagct 508

<210> 77

<211> 1651

<212> DNA

<213> Artificial sequence

<220>

<223> CAG promoter

<400> 77

acataactta cggtaaatgg cccgcctggc tgaccgcca acgacccccg ccattgacg 60

tcaataatga cgtatgttcc catagtaacg ccaataggga ctttcattg acgtcaatgg 120

gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt 180

acgcccccta ttgacgtcaa tgacggtaaa tggcccgccg ggcatatgc ccagtacatg 240

accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg 300

gtcgaggatga gcccacgtt ctgcttact ctcccatct cccccctc cccacccca 360

atthtattt tttttttt ttaattatt ttgtcagcga tggggcgagg gggggggggg 420

gggcgcgcgc caggcggggc gggcggggc gagggcgagg gcggggcgag gcggagaggt 480

gcggcgagcag ccaatcagag cggcgcgctc cgaagtttc cttttatggc gaggcggcgg 540

cggcgcgagg cctataaaaa gcgaagcgcg cggcgggcgg ggagtcgtg cgacgtgcc 600

ttcgccccgt gcccgcctc gccgcgcct cgcgcgcgc gcccgcgtc tgactgaccg 660

cgttactccc acaggtgagc gggcgggagc gcccttctcc tccgggctgt aattagcgt 720

tggtttaatg acggcttgtt tcttttctgt ggctgcgtga aagccttgag gggtccggg 780

agggcccttt gtgcgggggg agcgctcgg ggggtgcgtg cgtgtgtgtg tgcgtgggga 840

gcgccgcgtg cgctccgcg ctccccgcg gctgtgagcg ctgcgggcgc gccgcggggc 900

tttgtgcgt cgcagtggtg cgcgagggga gcgcggccgg gggcggtgcc ccgcggtgcg 960

gggggggctg cgaggggaac aaaggctgcg tgcggggtgt gtgcgtgggg gggtgagcag 1020

ggggtgtggg cgcgtcggtc gggctgcaac ccccttca ccccttccc cgagttgctg 1080

agcacggccc ggttcgggt gcggggctcc gtacggggcg tggcggggg ctcgccgtgc 1140

cggcgggggg gtggcgagc gtgggggtgc cggcggggc ggggccgcct cgggcgggg 1200

agggctcggg ggagggcggc ggcggcccc ggagcgccgg cggctgtcga gccgcggcga 1260

gccgcagcca ttgcctttta tggtaatcgt gcgagagggc gcagggactt cttttgtccc 1320

aaatctgtgc ggagccgaaa tctgggaggc gccgcgcac cccctctagc gggcgcgggg 1380

cgaagcgggtg cggcgccggc aggaaggaaa tgggcgggga gggccttcgt gcgtcgccgc 1440
gcccgcgtcc cttctccct ctccagcctc ggggctgtcc gcggggggac ggctgccttc 1500

gggggggacg gggcaggcg gggttcggt tctggcgtgt gaccggcggc tctagagcct 1560
ctgctaacca tgttcatgcc ttcttctttt tctacagct cctgggcaac gtgctggtta 1620
ttgtgctgtc tcatcatttt ggcaaagaat t 1651

<210> 78
<211> 546
<212> DNA
<213> Artificial sequence
<220>
<223> EF1-alpha Promoter
<400> 78

aaggatctgc gatcgctccg gtgcccgtca gtgggcagag cgcacatcgc ccacagtcgc 60
cgagaagttg gggggagggg tcggcaattg aacgggtgcc tagagaaggt ggcgcggggt 120
aaactgggaa agtgatgtcg tgtactggct ccgccttttt cccgagggtg ggggagaacc 180

gtatataagt gcagtagtcg ccgtgaacgt tctttttcgc aacgggtttg ccgccagaac 240
acagctgaag cttcgagggg ctgcgcatctc tctttcacgc gccgcgcgcc ctacctgagg 300
ccgccatcca cgccggttga gtgcggttct gccgcctccc gcctgtggtg cctcctgaac 360
tgcgtccgcc gtctaggtaa gtttaaagct caggctcaga cgggccttt gtccggcgct 420
cccttggagc ctacctagac tcagccggct ctccacgctt tgcctgaccc tgcttgcctca 480
actctacgtc ttgttttctg tttctgttct gcgccgttac agatccaagc tgtgaccggc 540
gcctac 546

<210> 79
<211> 317
<212> DNA
<213> Artificial sequence
<220>
<223> SV-40 Promoter
<400> 79

ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60
agtcagcaac cagggtgtga aagtcgccag gtcgccagc aggcagaagt atgcaaagca 120
tgcatctcaa ttagtcagca accatagtcc cggccctaac tccgcccac cggccctaa 180

ctccgcccag ttccgcccat tctccgcccc atggctgact aatTTTTTTT atttatgcag 240
aggccgaggc cgctcggcc tctgagctat tccagaagta gtgaggaggc ttttttgag 300
gcctaggctt ttgcaaa 317

<210> 80

<211> 725

<212> DNA

<213> Artificial sequence

<220>

<223> EalbAATp

<400> 80

gacgcgatg ctctctaga ctcgacgct gtctctagat tacactacac attctgcaag 60
catagcacag agcaatgttc tactttaatt actttcattt tcttgatcc tcacagccta 120
gaaaataacc tgcgttacag catccactca gtatcccttg agcatgaggt gacactactt 180
aacatagggg cgagatggta ctttgtgtct cctgctctgt cagcagggca cagtacttgc 240
tgataccagg gaatgtttgt tcttaataac catcattccg gacgtgtttg ccttggccag 300
ttttccatgt acatgcagaa agaagtttgg actgatcaat acagtcctct gcctttaag 360

caataggaaa aggccaactt gtctacgttt agtatgtggc tgtagatctg tacccgccac 420
cccctccacc ttggacacag gacgtgtgg tttctgagcc aggtacaatg actccttcg 480
gtaagtgcag tggaagctgt aactgcca ggcaaagcgt ccgggcagcg taggcgggcg 540
actcagatcc cagccagtgg acttagcccc tgtttgctcc tccgataact ggggtgacct 600
tggttaatat tcaccagcag cctccccctg tgcctctctg gatccactgc ttaaatacgg 660
acgaggacag ggccctgtct cctcagcttc aggcaccacc actgacctgg gacagtgaag 720
cggcc 725

<210> 81

<211> 51

<212> DNA

<213> Artificial sequence

<220>

<223> Splice Acceptor

<400> 81

gataggcacc tattggtctt actgacatcc actttgcctt tctctccaca g 51

<210> 82

<211> 86

<212> DNA

<213> artificial sequence

<220>

<223> synthetic ama2 intron 26

<400> 82

gtaagttagt cattgtttgg tgcaaagata ccaatcaatg gttttgcatt cagttttgtc	60
atagtgattt ctcttcttgt tttcag	86

<210> 83

<211> 131

<212> DNA

<213> artificial sequence

<220>

>

<223> SV40-PolyA

<400> 83

aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca	60
aaataaagcat ttttttcaact gcattctagt tgtggtttgt ccaaactcat caatgtatct	120
tatcatgtct g	131

<210> 84

<211> 263

<212> DNA

<213> artificial sequence

<220>

<223> BHG-PolyA

<400> 84

gtgtcaccta aatgctagag ctgctgac agcctcgact gtgccttcta gttgccagcc	60
atctgttgtt tgccctccc ccgtgccttc cttgaccctg gaaggcgcca ctccactgt	120
cctttcctaa taaatgagg aaattgcac gcattgtctg agtaggtgtc attctattct	180

ggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata gcaggcatgc	240
tggggatgcg gtgggctcta tgg	263

<210> 85

<211> 49

<212> DNA

<213> artificial sequence

<220>

<223> HSV-Tk-PolyA

<400> 85

cggcaataaa aagacagaat aaaacgcacg ggtgttgggt cgtttgttc 49

<210> 86

<211> 232

<212> DNA

<213> artificial sequence

<220>

<223> Insulator a

<400> 86

gagggacagc cccccccaa agccccagg gatgtaatta cgtccctccc ccgctagggg 60

gcagcagcga gccgcccggg gctccgctcc ggtccggcgc tcccccgca tccccgagcc 120

ggcagcgtgc ggggacagcc cgggcacggg gaaggtggca cgggatcgct ttcctctgaa 180

cgcttctcgc tgctctttga gcctgcagac acctgggggg atacggggaa aa 232

<210> 87

<211> 231

<212> DNA

<213> artificial sequence

<220>

<223> insulator b

<400> 87

ttttcccggt atccccccag gtgtctgcag gctcaaagag cagcgagaag cgttcagagg 60

aaagcgatcc cgtgccacct tcccgtgcc cgggctgtcc ccgcacgtg ccggtcggg 120

gatgcggggg gagcgccgga ccggagcgga gccccgggcg gctcgctgct gccccctagc 180

gggggaggga cgtaattaca tccttggggg ctttgggggg gggctgtccc t 231

<210> 88

<211> 63

<212> DNA

<213> artificial sequence

<220>

<223> ITR-5'

<400> 88
ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60
atg 63
<210> 89
<211> 35
<212> DNA
<213> artificial sequence
<220>
<223> ITR-3'
<400> 89
catgcgtcaa ttttacgcag actatctttc taggg 35
<210> 90
<211> 20
<212> DNA
<213> artificial sequence
<220>
<223> gRNA-Rosa26-1
<
400> 90
agtctttcta gaagatgggc 20
<210> 91
<211> 20
<212> DNA
<213> artificial sequence
<220>
<223> gRNA-Rosa26-2
<400> 91
cagtctttct agaagatggg 20
<210> 92
<211> 20
<212> DNA
<213> artificial sequence
<220>
<223> gRNA-Albumin-1
<400> 92

caatctttaa atatgttg 20

<210> 93

<211> 20

<212> DNA

<213> artificial sequence

<220>

<

223> gRNA-Rosa26-4

<400> 93

actccagtct ttctagaaga 20

<210> 94

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> gRNA-Albumin-1

<400> 94

caatctttaa atatgttg 20

<210> 95

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> gRNA-Albumin-2

<400> 95

gttacaggaa aatctgaagg 20

<210> 96

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> gRNA-Albumin-3

<400> 96

taactttgag ttagcagag 20

<210> 97

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> gRNA-Albumin-4

<400> 97

tatagcatgg tcgagcaggc

20

<210> 98

<211> 3117

<212> PRT

<213> mus musculus

<400> 98

Met Pro Ala Ala Thr Ala Gly Ile Leu Leu Leu Leu Leu Gly Thr

1 5 10 15

Leu Glu Gly Ser Gln Thr Gln Arg Arg Gln Ser Gln Ala His Gln Gln

20 25 30

Arg Gly Leu Phe Pro Ala Val Leu Asn Leu Ala Ser Asn Ala Leu Ile

35 40 45

Thr Thr Asn Ala Thr Cys Gly Glu Lys Gly Pro Glu Met Tyr Cys Lys

50 55 60

Leu Val Glu His Val Pro Gly Gln Pro Val Arg Asn Pro Gln Cys Arg

65 70 75 80

Ile Cys Asn Gln Asn Ser Ser Asn Pro Tyr Gln Arg His Pro Ile Thr

85 90 95

Asn Ala Ile Asp Gly Lys Asn Thr Trp Trp Gln Ser Pro Ser Ile Lys

100 105 110

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

115 120 125

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

130 135 140

Pro Gly Asn Trp Ile Leu Glu Arg Ser Leu Asp Asp Val Glu Tyr Lys

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

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

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

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

Pro Asn Tyr Pro Arg Pro Cys Gln Pro Cys His Cys Asp Pro Thr Gly

405 410 415

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

420 425 430

Leu Lys Pro Gly Ser Cys His Cys Lys Thr Gly Phe Gly Gly Val Asn

435 440 445

Cys Asp Arg Cys Val Arg Gly Tyr His Gly Tyr Pro Asp Cys Gln Pro

450 455 460

Cys Asn Cys Ser Gly Leu Gly Ser Thr Asn Glu Asp Pro Cys Val Gly

465 470 475 480

Pro Cys Ser Cys Lys Glu Asn Val Glu Gly Glu Asp Cys Ser Arg Cys

485 490 495

Lys Ser Gly Phe Phe Asn Leu Gln Glu Asp Asn Gln Lys Gly Cys Glu

500 505 510

Glu Cys Phe Cys Ser Gly Val Ser Asn Arg Cys Gln Ser Ser Tyr Trp

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

Gly Tyr Tyr Trp Ser Ala Pro Pro Tyr Leu Gly Asn Arg Leu Pro

580 585 590

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

595 600 605

Glu Glu Asp Asp Thr Glu Lys Ile Leu Gln Leu Met Ile Ile Phe Glu

610 615 620

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

625 630 635 640

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

645 650 655
Thr Ile His Gly Thr Asn Leu Pro Val Thr Arg Lys Asp Phe Met Ile

660 665 670
Val Leu Thr Asn Leu Glu Arg Val Leu Met Gln Ile Thr Tyr Asn Leu

675 680 685
Gly Met Asp Ala Ile Phe Arg Leu Ser Ser Val Asn Leu Glu Ser Ala

690 695 700
Val Pro Tyr Pro Thr Asp Arg Arg Ile Ala Thr Asp Val Glu Val Cys

705 710 715 720
Gln Cys Pro Pro Gly Tyr Ser Gly Ser Ser Cys Glu Thr Cys Trp Pro

725 730 735
Arg His Arg Arg Val Asn Gly Thr Ile Phe Gly Gly Ile Cys Glu Pro

740 745 750
Cys Gln Cys Phe Ala His Ala Glu Ala Cys Asp Asp Ile Thr Gly Glu

755 760 765
Cys Leu Asn Cys Lys Asp His Thr Gly Gly Pro Tyr Cys Asn Glu Cys

770 775 780
Leu Pro Gly Phe Tyr Gly Asp Pro Thr Arg Gly Ser Pro Glu Asp Cys

785 790 795 800
Gln Pro Cys Ala Cys Pro Leu Asn Ile Pro Ser Asn Asn Phe Ser Pro

805 810 815
Thr Cys His Leu Asp Arg Ser Leu Gly Leu Ile Cys Asp Glu Cys Pro

820 825 830
Ile Gly Tyr Thr Gly Pro Arg Cys Glu Arg Cys Ala Glu Gly Tyr Phe

835 840 845
Gly Gln Pro Ser Ile Pro Gly Gly Ser Cys Gln Pro Cys Gln Cys Asn

850 855 860
Asp Asn Leu Asp Tyr Ser Ile Pro Gly Ser Cys Asp Ser Leu Ser Gly

865 870 875 880
Ser Cys Leu Ile Cys Lys Pro Gly Thr Thr Gly Arg Tyr Cys Glu Leu

885 890 895

Cys Ala Asp Gly Tyr Phe Gly Asp Ala Val Asn Ala Lys Asn Cys Gln

900 905 910

Pro Cys Arg Cys Asn Ile Asn Gly Ser Phe Ser Glu Ile Cys His Thr

915 920 925

Arg Thr Gly Gln Cys Glu Cys Arg Pro Asn Val Gln Gly Arg His Cys

930 935 940

Asp Glu Cys Lys Pro Glu Thr Phe Gly Leu Gln Leu Gly Arg Gly Cys

945 950 955 960

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

965 970 975

Ala Ser Gly Gln Cys Trp Cys Gln Pro Gly Val Ala Gly Lys Lys Cys

980 985 990

Asp Arg Cys Ala His Gly Tyr Phe Asn Phe Gln Glu Gly Gly Cys Ile

995 1000 1005

Ala Cys Asp Cys Ser His Leu Gly Asn Asn Cys Asp Pro Lys Thr Gly

1010 1015 1020

Gln Cys Ile Cys Pro Pro Asn Thr Thr Gly Glu Lys Cys Ser Glu Cys

1025 1030 1035 1040

Leu Pro Asn Thr Trp Gly His Ser Ile Val Thr Gly Cys Lys Val Cys

1045 1050 1055

Asn Cys Ser Thr Val Gly Ser Leu Ala Ser Gln Cys Asn Val Asn Thr

1060 1065 1070

Gly Gln Cys Ser Cys His Pro Lys Phe Ser Gly Met Lys Cys Ser Glu

1075 1080 1085

Cys Ser Arg Gly His Trp Asn Tyr Pro Leu Cys Thr Leu Cys Asp Cys

1090 1095 1100

Phe Leu Pro Gly Thr Asp Ala Thr Thr Cys Asp Leu Glu Thr Arg Lys

1105 1110 1115 1120

Cys Ser Cys Ser Asp Gln Thr Gly Gln Cys Ser Cys Lys Val Asn Val

1125 1130 1135

Glu Gly Val His Cys Asp Arg Cys Arg Pro Gly Lys Phe Gly Leu Asp

1140 1145 1150
Ala Lys Asn Pro Leu Gly Cys Ser Ser Cys Tyr Cys Phe Gly Val Thr

1155 1160 1165
Ser Gln Cys Ser Glu Ala Lys Gly Leu Ile Arg Thr Trp Val Thr Leu

1170 1175 1180
Ser Asp Glu Gln Thr Ile Leu Pro Leu Val Asp Glu Ala Leu Gln His

1185 1190 1195 1200
Thr Thr Thr Lys Gly Ile Ala Phe Gln Lys Pro Glu Ile Val Ala Lys

1205 1210 1215
Met Asp Glu Val Arg Gln Glu Leu His Leu Glu Pro Phe Tyr Trp Lys

1220 1225 1230
Leu Pro Gln Gln Phe Glu Gly Lys Lys Leu Met Ala Tyr Gly Gly Lys

1235 1240 1245
Leu Lys Tyr Ala Ile Tyr Phe Glu Ala Arg Asp Glu Thr Gly Phe Ala

1250 1255 1260
Thr Tyr Lys Pro Gln Val Ile Ile Arg Gly Gly Thr Pro Thr His Ala

1265 1270 1275 1280
Arg Ile Ile Thr Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr

1285 1290 1295
Arg His Glu Ile Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp

1300 1305 1310
Asp Pro Arg Ile Ser Arg Thr Val Thr Arg Glu Asp Phe Leu Asp Ile

1315 1320 1325
Leu Tyr Asp Ile His Tyr Ile Leu Ile Lys Ala Thr Tyr Gly Asn Val

1330 1335 1340
Val Arg Gln Ser Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Pro

1345 1350 1355 1360
Gly His Val Leu Ala Gly Ser Pro Pro Ala His Leu Ile Glu Arg Cys

1365 1370 1375
Asp Cys Pro Pro Gly Tyr Ser Gly Leu Ser Cys Glu Thr Cys Ala Pro

1380 1385 1390

Gly Phe Tyr Arg Leu Arg Ser Glu Pro Gly Gly Arg Thr Pro Gly Pro

1395 1400 1405

Thr Leu Gly Thr Cys Val Pro Cys Gln Cys Asn Gly His Ser Ser Gln

1410 1415 1420

Cys Asp Pro Glu Thr Ser Val Cys Gln Asn Cys Gln His His Thr Ala

1425 1430 1435 1440

Gly Asp Phe Cys Glu Arg Cys Ala Leu Gly Tyr Tyr Gly Ile Val Arg

1445 1450 1455

Gly Leu Pro Asn Asp Cys Gln Pro Cys Ala Cys Pro Leu Ile Ser Pro

1460 1465 1470

Ser Asn Asn Phe Ser Pro Ser Cys Val Leu Glu Gly Leu Glu Asp Tyr

1475 1480 1485

Arg Cys Thr Ala Cys Pro Arg Gly Tyr Glu Gly Gln Tyr Cys Glu Arg

1490 1495 1500

Cys Ala Pro Gly Tyr Thr Gly Ser Pro Ser Ser Pro Gly Gly Ser Cys

1505 1510 1515 1520

Gln Glu Cys Glu Cys Asp Pro Tyr Gly Ser Leu Pro Val Pro Cys Asp

1525 1530 1535

Arg Val Thr Gly Leu Cys Thr Cys Arg Pro Gly Ala Thr Gly Arg Lys

1540 1545 1550

Cys Asp Gly Cys Glu His Trp His Ala Arg Glu Gly Ala Glu Cys Val

1555 1560 1565

Phe Cys Gly Asp Glu Cys Thr Gly Leu Leu Leu Gly Asp Leu Ala Arg

1570 1575 1580

Leu Glu Gln Met Thr Met Asn Ile Asn Leu Thr Gly Pro Leu Pro Ala

1585 1590 1595 1600

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

1605 1610 1615

His Leu Leu Ser Pro Gln Arg Ala Pro Glu Arg Leu Ile Gln Leu Ala

1620 1625 1630

Glu Gly Asn Val Asn Thr Leu Val Met Glu Thr Asn Glu Leu Leu Thr

1635	1640	1645	
Arg Ala Thr Lys Val Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala			
1650	1655	1660	
Glu Arg Thr Asn Ser Arg Ala Glu Ser Leu Glu Glu Phe Ile Lys Gly			
1665	1670	1675	1680
Leu Val Gln Asp Ala Glu Ala Ile Asn Glu Lys Ala Val Gln Leu Asn			
	1685	1690	1695
Glu Thr Leu Gly Asn Gln Asp Lys Thr Ala Glu Arg Asn Leu Glu Glu			
1700	1705	1710	
Leu Gln Lys Glu Ile Asp Arg Met Leu Lys Glu Leu Arg Ser Lys Asp			
1715	1720	1725	
Leu Gln Thr Gln Lys Glu Val Ala Glu Asp Glu Leu Val Ala Ala Glu			
1730	1735	1740	
Gly Leu Leu Lys Arg Val Asn Lys Leu Phe Gly Glu Pro Arg Ala Gln			
1745	1750	1755	1760
Asn Glu Asp Met Glu Lys Asp Leu Gln Gln Lys Leu Ala Glu Tyr Lys			
1765	1770	1775	
Asn Lys Leu Asp Asp Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys			
1780	1785	1790	
Thr Arg Asp Ala Asn Arg Leu Ser Ala Ala Asn Gln Lys Asn Met Thr			
1795	1800	1805	
Ile Leu Glu Thr Lys Lys Glu Ala Ile Glu Gly Ser Lys Arg Gln Ile			
1810	1815	1820	
Glu Asn Thr Leu Lys Glu Gly Asn Asp Ile Leu Asp Glu Ala Asn Arg			
1825	1830	1835	1840
Leu Leu Gly Glu Ile Asn Ser Val Ile Asp Tyr Val Asp Asp Ile Lys			
1845	1850	1855	
Thr Lys Leu Pro Pro Met Ser Glu Glu Leu Ser Asp Lys Ile Asp Asp			
1860	1865	1870	
Leu Ala Gln Glu Ile Lys Asp Arg Arg Leu Ala Glu Lys Val Phe Gln			
1875	1880	1885	

Ala Glu Ser His Ala Ala Gln Leu Asn Asp Ser Ser Ala Val Leu Asp
1890 1895 1900
Gly Ile Leu Asp Glu Ala Lys Asn Ile Ser Phe Asn Ala Thr Ala Ala
1905 1910 1915 1920
Phe Arg Ala Tyr Ser Asn Ile Lys Asp Tyr Ile Asp Glu Ala Glu Lys
1925 1930 1935
Val Ala Arg Glu Ala Lys Glu Leu Ala Gln Gly Ala Thr Lys Leu Ala
1940 1945 1950
Thr Ser Pro Gln Gly Leu Leu Lys Glu Asp Ala Lys Gly Ser Leu Gln
1955 1960 1965
Lys Ser Phe Arg Ile Leu Asn Glu Ala Lys Lys Leu Ala Asn Asp Val
1970 1975 1980
Lys Gly Asn His Asn Asp Leu Asn Asp Leu Lys Thr Arg Leu Glu Thr
1985 1990 1995 2000
Ala Asp Leu Arg Asn Ser Gly Leu Leu Gly Ala Leu Asn Asp Thr Met
2005 2010 2015
Asp Lys Leu Ser Ala Ile Thr Asn Asp Thr Ala Ala Lys Leu Gln Ala
2020 2025 2030
Val Lys Glu Lys Ala Arg Glu Ala Asn Asp Thr Ala Lys Ala Val Leu
2035 2040 2045
Ala Gln Val Lys Asp Leu His Gln Asn Leu Asp Gly Leu Lys Gln Asn
2050 2055 2060
Tyr Asn Lys Leu Ala Asp Ser Val Ala Lys Thr Asn Ala Val Val Lys
2065 2070 2075 2080
Asp Pro Ser Lys Asn Lys Ile Ile Ala Asp Ala Gly Thr Ser Val Arg
2085 2090 2095
Asn Leu Glu Gln Glu Ala Asp Arg Leu Ile Asp Lys Leu Lys Pro Ile
2100 2105 2110
Lys Glu Leu Glu Asp Asn Leu Lys Lys Asn Ile Ser Glu Ile Lys Glu
2115 2120 2125
Leu Ile Asn Gln Ala Arg Lys Gln Ala Asn Ser Ile Lys Val Ser Val

2130	2135	2140	
Ser Ser Gly Gly Asp Cys Val Arg Thr Tyr Arg Pro Glu Ile Lys Lys			
2145	2150	2155	2160
Gly Ser Tyr Asn Asn Ile Val Val His Val Lys Thr Ala Val Ala Asp			
	2165	2170	2175
Asn Leu Leu Phe Tyr Leu Gly Ser Ala Lys Phe Ile Asp Phe Leu Ala			
	2180	2185	2190
Ile Glu Met Arg Lys Gly Lys Val Ser Phe Leu Trp Asp Val Gly Ser			
	2195	2200	2205
Gly Val Gly Arg Val Glu Tyr Pro Asp Leu Thr Ile Asp Asp Ser Tyr			
	2210	2215	2220
Trp Tyr Arg Ile Glu Ala Ser Arg Thr Gly Arg Asn Gly Ser Ile Ser			
2225	2230	2235	2240
Val Arg Ala Leu Asp Gly Pro Lys Ala Ser Met Val Pro Ser Thr Tyr			
	2245	2250	2255
His Ser Val Ser Pro Pro Gly Tyr Thr Ile Leu Asp Val Asp Ala Asn			
	2260	2265	2270
Ala Met Leu Phe Val Gly Gly Leu Thr Gly Lys Ile Lys Lys Ala Asp			
	2275	2280	2285
Ala Val Arg Val Ile Thr Phe Thr Gly Cys Met Gly Glu Thr Tyr Phe			
	2290	2295	2300
Asp Asn Lys Pro Ile Gly Leu Trp Asn Phe Arg Glu Lys Glu Gly Asp			
2305	2310	2315	2320
Cys Lys Gly Cys Thr Val Ser Pro Gln Val Glu Asp Ser Glu Gly Thr			
	2325	2330	2335
Ile Gln Phe Asp Gly Glu Gly Tyr Ala Leu Val Ser Arg Pro Ile Arg			
	2340	2345	2350
Trp Tyr Pro Asn Ile Ser Thr Val Met Phe Lys Phe Arg Thr Phe Ser			
	2355	2360	2365
Ser Ser Ala Leu Leu Met Tyr Leu Ala Thr Arg Asp Leu Lys Asp Phe			
	2370	2375	2380

Met Ser Val Glu Leu Ser Asp Gly His Val Lys Val Ser Tyr Asp Leu

2385 2390 2395 2400

Gly Ser Gly Met Thr Ser Val Val Ser Asn Gln Asn His Asn Asp Gly

2405 2410 2415

Lys Trp Lys Ala Phe Thr Leu Ser Arg Ile Gln Lys Gln Ala Asn Ile

2420 2425 2430

Ser Ile Val Asp Ile Asp Ser Asn Gln Glu Glu Asn Val Ala Thr Ser

2435 2440 2445

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

2450 2455 2460

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

2465 2470 2475 2480

Pro Glu Val Asn Val Lys Lys Tyr Ser Gly Cys Leu Lys Asp Ile Glu

2485 2490 2495

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

2500 2505 2510

Val Thr Lys Gly Cys Ser Leu Glu Asn Val Tyr Thr Val Ser Phe Pro

2515 2520 2525

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

2530 2535 2540

Glu Ile Asn Leu Ser Phe Ser Thr Arg Asn Glu Ser Gly Ile Ile Leu

2545 2550 2555 2560

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

2565 2570 2575

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

2580 2585 2590

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

2595 2600 2605

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

2610 2615 2620

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

2625 2630 2635 2640
 His Met Gln Asn Leu Thr Glu Glu Gln Pro Ile Glu Val Lys Lys Leu

 2645 2650 2655
 Phe Val Gly Gly Ala Pro Pro Glu Phe Gln Pro Ser Pro Leu Arg Asn
 2660 2665 2670
 Ile Pro Ala Phe Gln Gly Cys Val Trp Asn Leu Val Ile Asn Ser Ile
 2675 2680 2685
 Pro Met Asp Phe Ala Gln Pro Ile Ala Phe Lys Asn Ala Asp Ile Gly
 2690 2695 2700
 Arg Cys Thr Tyr Gln Lys Pro Arg Glu Asp Glu Ser Glu Ala Val Pro

 2705 2710 2715 2720
 Ala Glu Val Ile Val Gln Pro Gln Pro Val Pro Thr Pro Ala Phe Pro
 2725 2730 2735
 Phe Pro Ala Pro Thr Met Val His Gly Pro Cys Val Ala Glu Ser Glu
 2740 2745 2750
 Pro Ala Leu Leu Thr Gly Ser Lys Gln Phe Gly Leu Ser Arg Asn Ser
 2755 2760 2765
 His Ile Ala Ile Ala Phe Asp Asp Thr Lys Val Lys Asn Arg Leu Thr

 2770 2775 2780
 Ile Glu Leu Glu Val Arg Thr Glu Ala Glu Ser Gly Leu Leu Phe Tyr
 2785 2790 2795 2800
 Met Ala Arg Ile Asn His Ala Asp Phe Ala Thr Val Gln Leu Arg Asn
 2805 2810 2815
 Gly Phe Pro Tyr Phe Ser Tyr Asp Leu Gly Ser Gly Asp Thr Ser Thr
 2820 2825 2830
 Met Ile Pro Thr Lys Ile Asn Asp Gly Gln Trp His Lys Ile Lys Ile

 2835 2840 2845
 Val Arg Val Lys Gln Glu Gly Ile Leu Tyr Val Asp Asp Ala Ser Ser
 2850 2855 2860
 Gln Thr Ile Ser Pro Lys Lys Ala Asp Ile Leu Asp Val Val Gly Ile
 2865 2870 2875 2880

Leu Tyr Val Gly Gly Leu Pro Ile Asn Tyr Thr Thr Arg Arg Ile Gly

2885 2890 2895

Pro Val Thr Tyr Ser Leu Asp Gly Cys Val Arg Asn Leu His Met Glu

2900 2905 2910

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

2915 2920 2925

Thr Cys Phe Ala Asn Ala Glu Ser Gly Thr Tyr Phe Asp Gly Thr Gly

2930 2935 2940

Phe Ala Lys Ala Val Gly Gly Phe Lys Val Gly Leu Asp Leu Leu Val

2945 2950 2955 2960

Glu Phe Glu Phe Arg Thr Thr Arg Pro Thr Gly Val Leu Leu Gly Val

2965 2970 2975

Ser Ser Gln Lys Met Asp Gly Met Gly Ile Glu Met Ile Asp Glu Lys

2980 2985 2990

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

2995 3000 3005

Asp Ala Gly Ile Pro Gly His Met Cys Asn Gly Gln Trp His Lys Val

3010 3015 3020

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

3025 3030 3035 3040

Gln Val Asp Ala Gln Ser Pro Asn Ser Ala Ser Thr Ser Ala Asp Thr

3045 3050 3055

Asn Asp Pro Val Phe Val Gly Gly Phe Pro Gly Gly Leu Asn Gln Phe

3060 3065 3070

Gly Leu Thr Thr Asn Ile Arg Phe Arg Gly Cys Ile Arg Ser Leu Lys

3075 3080 3085

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

3090 3095 3100

Leu Glu Leu Arg Gly Val Gln Pro Val Ser Cys Pro Thr

3105 3110 3115

<210> 99

<211> 9353

<212> DNA

<213> mus musculus

<400> 99

atgcctgcgg ccaccgccgg gatcctcttg ctctgtctct tggggacgct cgaaggtcc	60
cagactcagc ggcgacagtc ccaagcgcat caacagagag gtttatttcc tgctgtcctg	120
aatcttgctt cgaatgcact catcacaacc aatgtacat gtggggaaaa aggacccgag	180
atgtactgca agttggtgga acatgtcccc gggcagcctg tgaggaaccc tcagtgccga	240
atctgcaatc agaacagcag caatccatac cagaggcacc cgattacgaa tgctattgat	300
ggcaagaaca catggtggca gagtcccagt atcaagaatg gagtggata ccattatgtg	360
acaattactc tggatttaca gcaggtgttc cagattgcct acgtaattgt gaaggcagcc	420
aattcccctc ggcttgaaa ctggattttg gaacgttccc tggatgacgt ggagtacaaa	480
ccctggcagt atcatgcgtt gacagacacg gagtgcctga ccctctaca tatctatccc	540
cgcactggac caccatccta cgccaagat gatgaggtca tctgcacttc atttatttcg	600
aagatccacc ctttagaaaa tggagagatt cacatttctt tgatcaatgg gagaccaagt	660
gctgatgacc cctccccga actcctggaa ttcacctctg ctgctacat tcgcctgaga	720
tttcagagga tccgcacctt gaatgcagac ttgatgatgt ttgtcaca agaccccaga	780
gaaatcgatc cattgtcac acgaagatat tactattctg tcaaggatat ttcagttggc	840
gggatgtgca tctgttatgg tcatgcccg gcttgtccac ttgacctgc aacaaataaa	900
tcacgtgtg agtgtgaaca taacacctgt ggggaaagct gtgacagggt ctgtccagga	960
ttccatcaga agccttggag agctggaacc ttctcacca agtctgagtg tgaagcatgc	1020
aattgtcacg gaaaagctga ggaatgctat tatgatgaaa ctgttgctag cagaaatcta	1080
agtttaata tacatgggaa gtacatcgga ggggtgtgt gcatcaactg cacacataac	1140
acggctggga taaattgtga gacatgtgtt gatggattct tcagaccaa aggggtatca	1200
ccaaattatc caagaccatg ccagccatgt cactgtgatc caactggctc ccttagtgaa	1260
gtctgtgtca aagatgagaa atatgccag cgagggtga aacctggatc ctgtcactgc	1320
aaaactggct ttggaggcgt gaactgtgat cgctgtgtca ggggttacca tggttacca	1380
gactgccaac ccgttaactg tagtggcttg gggagcaca acgaggacc ttgcgttggg	1440
ccctgtagct gtaaggagaa tgttgaaggt gaagactgta gtcgttgcaa atctggtttc	1500
ttcaacttgc aagaagataa tcagaaaggc tgtgaggagt gtttctgttc aggagtatca	1560
aacagatgtc agagttccta ctggacctat gggaatatc aagacatgcg tggttggtat	1620

ctcacagacc tccttgccg cattcggatg gctccccagc ttgataaccc tgactcacct	1680
cagcagatca gcatcagtaa ctctgaggcc cggaaatccc tgcttgatgg ttactactgg	1740
agtgcaccgc ctccatatct gggaaacaga cttccagctg ttggaggaca gttgtcattt	1800
accatctcat atgacctcga agaagaggaa gacgatacag aaaaaatcct tcagctgatg	1860
attatctttg agggaaatga cttagaatc agcacagcgt ataaggaggt gtacttagag	1920
ccatctgaag aacacattga ggaggtgtca ctcaaagaag aggcctttac tatacatgga	1980
acaaatttgc cagtcactag aaaagatttc atgattgttc tcacaaattt ggagagagtc	2040
cttatgcaaa tcacatacaa cttagggatg gacgccatct tcaggctgag ttctgtcaat	2100
cttgaatctg ctgtccctta tcctactgat agacgtattg caactgatgt ggaagtttgc	2160
cagtgtccac ctgggtacag tggcagctct tgtgaaacat gttggcctag gcaccgaaga	2220
gttaacggca ccatttttgg tggcatttgt gaacctgtc agtgctttgc tcatgcagaa	2280
gcctgtgatg acatcacagg agaatgtctg aactgtaagg atcacacagg tgggccgtac	2340
tgcaatgaat gtctccctgg tttctatggt gatcctactc gaggaagccc tgaagactgt	2400
cagccctgtg cctgtccact caatatccca tcaataaact ttagtccaac atgccattta	2460
gaccggagtc tgggattgat ctgtgacgag tgtcctattg ggtacacagg accgcgtgt	2520
gagaggtgtg cagaaggcta ttttggacaa cttccatac ctggaggatc atgtcagcca	2580
tgccaatgca atgacaacct tgactactcc atccctggca gctgtgacag cctgtctggc	2640
tcctgtctga ttgtgaacc aggtacaaca ggccggtact gtgagctctg tgctgatggg	2700
tattttggag acgcggttaa tgcaaagaac tgtcaaccat gccgttgtaa tatcaatggc	2760
tccttctcag agattttgca cactagaact gggcaatgtg agtgcagacc caatgtgcag	2820
gggcggcact gtgacgagtg taagcctgaa acctttggcc tgcaactggg aaggggttgt	2880
ctgccctgca actgcaattc ttttgggtct aagtcctttg actgtgaagc aagtgggcag	2940
tgctgggtgc agcctggagt agcaggggaag aaatgtgacc gttgtgccca tggctacttc	3000
aacttccaag aaggaggctg catagcttgt gactgttctc atctgggcaa caactgtgac	3060
ccaaaaactg gccaatgcat ttgccaccc aataccactg gagaaaagtg ttctgagtgt	3120
cttccaaca cctggggica cagcattgtc accggctgta aggtttgtaa ctgcagcact	3180
gtggggtcct tggtttctca gtgcaatgta aacacgggcc agtgcagctg tcatccaaaa	3240
ttctctggta tgaatgtct agagtgcagc cgaggtcact ggaactatcc tctctgcact	3300
ctatgtgact gtttcttcc agggacagat gccacgactt gtgatctgga gactaggaaa	3360
tgctcctgta gtgatcaaac tggacagtgc agctgtaagg tgaatgtgga aggcgtccac	3420
tgtgacaggt gccggcctgg caaatttggga ctagatgcca agaaccact tggctgcagc	3480

agctgctact gctttggagt tactagtcaa tgctctgaag caaaggggct gatccgtacg	3540
tgggtgactt tgagtgatga acagaccatt ctacctctgg tggatgaggc cctgcagcac	3600
acgactacca aaggcattgc ttccagaaa ccagagattg ttgcaaagat ggatgaagtc	3660
aggcaagagc tccatttggga acctttttac tggaaactcc cacaacaatt tgaaggga	3720
aagtgtatgg cttatgttgg caaactcaag tatgccatct attttgaggc tcgggatgag	3780
acaggctttg ccacataaa acctcaagtt atcattcgag gtggaactcc tactcatgct	3840
agaattatta ccagacacat ggctgcccct ctcatgtggc agttgacacg gcatgaaata	3900
gaaatgacag agaaagaatg gaaatattat ggtgatgatc ctggaatcag tagaactgtg	3960
acccgtgaag acttcttggga tatactatat gatattcact atatccttat caaggctact	4020
tatggaacg ttgtgagaca aagccgcatt tctgaaatct ccatggaagt agccgaacca	4080
ggacatgtat tagcaggag cccaccagca cacttgatag aaagatgcga ttgccctcct	4140
ggctattctg gcttgtcttg tgagacgtgt gcaccaggat tttatcgact tcgttctgaa	4200
ccagggtggc gcactcctgg accaacctta ggaacctgtg ttcctgcca atgtaatgga	4260
cacagcagtc agtgtgatcc tgagacctca gtatgccaga attgtcagca tcacactgct	4320
ggtgacttct gtgagcgtg tgcccttggc tactatggaa tcgtcagggg attgccaaat	4380
gactgtcaac catgtgcttg tctctgatt tcgccagca acaatttcag cccctcttgt	4440
gtattggaag gtctggaaga ttaccgttgc accgcctgcc caaggggcta tgaaggacag	4500
tactgtgaaa ggtgtgcccc aggtatact ggcagcccaa gcagccccgg aggtcctgc	4560
caagaatgtg agtgtgacct ttatggctcc ctaccggttc cctgtgaccg ggtcacggga	4620
ctctgcacgt gccgccctgg agccacagga aggaagtgtg atggctgcga gcactggcat	4680
gcacgcgagg gtgcagagtg tgtcttttgt ggagacgagt gtacaggcct tcttcttgg	4740
gacctggctc gtctagagca gatgacctg aacatcaacc tcacgggccc actgctgct	4800
ccatataaaa ttctgtatgg tcttgaaaat acaactcagg aactcaagca cctgctatca	4860
ccgcaacggg caccagagag gctcattcag ttggcagagg gcaacgtgaa cacacttgtg	4920
atggaacaa atgagctgct aaccagagca accaaagtga cagcagatgg tgagcaaaca	4980
ggacaagatg ctgagaggac caactccaga gcagaatcct tggaagaatt cattaaagg	5040
cttgtccagg atgtgaagc cataaatgaa aaagctgtac aactaaatga aaccttagga	5100
aatcaagata agacagcaga gagaaacttg gaggagcttc aaaaggaaat cgaccgatg	5160
ctgaaggaac tgagaagtaa agatcttcaa acacagaagg aagttgctga ggatgagctc	5220

gtggcagcag aaggccttct gaagagagta aacaagctgt ttggagagcc cagagcccag 5280

aatgaagata tggaagagga tctccagcag aaactggcag agtacaagaa caaacttgat 5340

gatgcttggg atctattgag agaagccact gataaaactc gagatgctaa tcgtttgtct 5400

gctgccaatc aaaaaaacat gaccatactg gagacaaaga aggaggctat tgaaggtagc 5460

aaacgacaaa tagagaacac tttaaaggaa ggcaatgaca tccttgatga agccaatcga 5520

ctcttaggtg aaatcaactc agtcatagat tatgtcgacg acattaaaac taagttgcca 5580

ccaatgtccg aggagctgag tgacaaaata gatgacctcg cccaggaaat aaaggacaga 5640

aggcttgctg agaaggtgtt ccaggctgag agccatgctg ctcagctgaa cgactcgtct 5700

gctgtacttg atggaatcct ggatgaggct aagaacatct ctticaatgc cacggcagcc 5760

ttcagagctt acagtaatat taaagactac attgatgaag ctgagaaaagt ggccagagaa 5820

gccaaagagc ttgccaaggg ggctacaaaa ctggcaacaa gtctcaggg cttattaaaa 5880

gaagatgcca aaggctccct tcagaaaagc ttcaggatcc tcaatgaagc caagaagcta 5940

gcaaacgatg tgaagggaaa tcacaatgat ctaaatgacc tgaaaaccag gttagaaact 6000

gctgacctta gaaacagtgg acttctagga gctctaaatg acaccatgga caagttatca 6060

gccattacaa atgacacggc tgctaaactg caggctgtca aagagaaagc cagagaagcc 6120

aatgacacag caaaagctgt cctggcccag gttaaggacc tgcatcagaa cctagatggc 6180

ctgaagcaaa actacaataa actggcagac agcgtggcca aaacgaacgc tgtggtgaaa 6240

gaccttcca aaaaacaaat cattgcagat gcaggcactt ccgtgagaaa tctagaacag 6300

gaagctgacc ggctaatega caaactcaag cccatcaagg agcttgagga caacctaaag 6360

aaaaacattt ctgaaataaa ggaactgac aaccaagctc ggaacaagc taactctatc 6420

aaagtatctg tttcttcggg aggtgactgt gttcgaacat acaggccaga aatcaagaaa 6480

ggaagctaca ataacatcgt tgtccatgtc aagaccgctg ttgccgacaa cctccttttt 6540

tatcttggaa gtgccaaatt tattgacttt cttgctatag aaatgcgcaa aggcaaagtc 6600

agcttctctt gggatgttgg ctctggagt ttggcgagtag agtatccaga cttgaccatc 6660

gacgactcct attggtaccg tattgaagca tcaagaacgg gaagaaatgg atctatttct 6720

gtgagagctt tagatggacc caaagccagt atggtaccca gcacctacca ttcagtgtct 6780

cctcccgggt acactatcct agatgtggat gcaaatgcaa tgctgtttgt tgggtggcctg 6840

accggaaaaa taaagaaggc cgatgctgta cgtgtgatca ccttcaccgg ctgtatggga 6900

gaaacatact ttgacaacaa acctataggt ttatggaact tccgggagaa agaaggcgac 6960

tgtaagggat gtactgtcag cccacaagtg gaagatagtg aggggactat tcagtttgat	7020
ggtgaaggct atgcattagt gagccgcctt atccgctggt accccaacat ctccacagtc	7080
atgttcaagt tccggacatt ttcatcaagt gctctcctga tgtatcttgc cacacgagac	7140
ctgaaagatt tcatgagtg agagctcagt gatggacatg tgaaagtcag ctatgacctg	7200
ggctcaggaa tgacttcctg tgtcagcaat caaaaccata atgatgggaa atggaaagca	7260
ttcacgctgt cgcggattca gaaacaagcc aacatatcga ttgtcgacat cgattctaac	7320
caggaggaga atgtagctac ttcatcttct ggaaacaact ttggtcttga cttgaaagca	7380
gatgacaaaa tatattttgg tggcctgcc aactctgagaa acttgagtat gaaagcaagg	7440
ccagaagtca atgtgaagaa atactccggc tgcctcaaag atattgaaat ttcaagaaca	7500
ccttacaata tactcagcag ccctgattat gttgggtgtga ccaaaggctg ttcactggag	7560
aatgtttata cagtttagtt ccccaagcct ggttttgtgg agcttgccgc tgtgtctatt	7620
gatgttggaa cagaaatcaa tctgtccttt agtaccagga acgagtctgg gatcattctc	7680
ttgggaagtg gagggacact cacaccacc aggagaaaac ggagacaaac cacacaggct	7740
tattatgcc aattctcaa caagggccgc ttggaagtgc atctctctc ggggacacgg	7800
acaatgagga aaattgtcat caaacggag ccaaatgtt ttcatgatgg gagagaacat	7860
tctgtccacg tagaaagaac cagaggcatc ttcactgttc aaattgatga agacagaaga	7920
catatgcaa acctgacaga ggaacagccc atcgaagtga aaaagctctt tgtcgggggt	7980
gctcctcctg aatttcagcc ctcccactc aggaatatc cgccctttca aggctgtgtg	8040
tggaaccttg ttattaactc catcccatg gactttgcgc agcctatagc cttcaaaaat	8100
gccgacattg gccgctgtac ctatcaaaag ccccggaag atgagagtga agcagttcca	8160
gctgaagtta ttgtccagcc tcagccggtg cccaccctg ccttcccttt cccggcccc	8220
accatggtgc atggcccttg tgttcagaa tcagaaccag ctcttctgac agggagcaag	8280
cagtttgggc ttccagaaa cagccacatt gcaattgctt ttgatgacac caaagttaa	8340
aatgcctca ccattgagct ggaggtacga actgaagctg aatcaggctt gctcttctac	8400
atggctcgga tcaatcatgc tgattttgct actgttcagc tgaggaatgg cttcccgta	8460
ttcagttatg atttggggag tggggacaca agcaccatga tccccaaaa aatcaacgat	8520
ggccagtggc acaagattaa gattgtgaga gtgaagcagg agggaattct ttatgtggat	8580
gatgcctcca gccaaacat cagtcceaag aaagctgaca tcctggatgt cgtggggatt	8640
ctgtatgtcg gtggattgcc gatcaactat accacacgca gaattggtcc agtgacttac	8700
agcctggatg gctgtgttag gaattcttac atggaacaag ccctgttga tctggaccag	8760
cctacctca gctttcacgt tgggacatgc ttgcgaatg cagagagtgg gacttacttt	8820

gatggaaccg gttttgctaa agcagttggt gggttcaaag ttggattgga ctttcttgtg 8880
 gaatttgaat tccgtaccac aagacccact ggggtcctcc tgggagtcag cagtcagaag 8940
 atggatggaa tgggtattga aatgatcgac gagaagctta tgttccacgt ggataatggc 9000
 gctggccgat tcactgcaat ctatgatgct gggatcccag gccacatgtg caatggacag 9060

tggcataaag tcactgcca gaagatcaaa aaccgtcttg agctggtggt agatgggaac 9120
 caggtggatg cccagagccc aaactcagca tcgacatcag ctgatacaaa cgacctgtt 9180
 tttgttggcg gtttccagg tggcctcaat cagtttggcc tgaccaccaa cattaggttc 9240
 cgaggtgca tccgatctct gaagctcacc aaaggcactg gcaaaccgct ggaggttaat 9300
 ttgccaagg ccctggaact gaggggtgtt caacctgtat catgcccagac tac 9353

<210> 100

<211> 21

<212> DNA

<213> artificial sequence

<220>

<223> SYBR-Mm_Tfrc-F qPCR

<400> 100

tgggcactag attgatacc t 21

<210> 101

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> SYBR-Mm_Tfrc-R qPCR

<400> 101

atgagctgac cagccacttc 20

<210> 102

<211> 19

<212> DNA

<213> artificial sequence

<220>

<223> SYBR-RFP-F qPCR

<400> 102

atggccagct ccgaggatg 19

<210> 103
 <211> 22
 <212> DNA
 <213> artificial sequence
 <220>
 <223> SYBR-RFP-R qPCR
 <400> 103
 gaactgaggg ctcagaatat cc 22

<210> 104
 <211> 20
 <212> DNA
 <213> artificial sequence
 <220>
 <223> SYBR-hLama F qPCR
 <400> 104
 tagaacatgt ccctgggcag 20

<210> 105
 <211> 20
 <212> DNA
 <213> artificial sequence
 <220>
 <223> SYBR-hLama r qPCR
 <400> 105
 atacgcgatc tggaacacct 20

<210> 106
 <211> 20
 <212> DNA
 <213> artificial sequence
 <220>
 <223> SYBR-Mm-GAPDH-F qPCR
 <400> 106
 cgtcccgtag acaaaatggt 20

<210> 107

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> SYBR-Mm-GAPDH-R qPCR

<400> 107

tggaagatgg tgatgggctt

20

<210> 108

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac R275A/N347S/K375A/D450N/S592G

<400> 108

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

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

545 550 555 560

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 Gly

580 585 590

Cys Phe

<210> 109

<211> 594

<212

> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac N347S/D450N/T560A/S573A/F594L

<400> 109

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

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> 110

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac N347A/D450N

<400> 110

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> 111

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac R275A/N347S/R372A/D450N/T560A/F594L

<400> 111

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> 112

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R202K/R275A/N347S/R372A/D450N/T560A/F594L

<400> 112

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> 113

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R275A/R277A/N347S/R372A/D450N/T560A/S564P/F594L

<400> 113

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> 114

<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> 114

Gly Gly Gly Gly Ser

1 5

<210> 115

<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> 115

Glu Ala Ala Ala Lys

1 5

<210> 116

<211> 3210

<212> PRT

<213> Artificial Sequence

<220>

<223> laminin subunit α2 isoform?X1

<400> 116

Met Pro Gly Ala Ala Gly Val Leu Leu Leu Leu Leu Ser Gly Gly

1 5 10 15
Leu Gly Gly Val Gln Ala Gln Arg Pro Gln Gln Gln Arg Gln Ser Gln

20 25 30
Ala His Gln Gln Arg Gly Leu Phe Pro Ala Val Leu Asn Leu Ala Ser

35 40 45
Asn Ala Leu Ile Thr Thr Asn Ala Thr Cys Gly Glu Lys Gly Pro Glu

50 55 60
Met Tyr Cys Lys Leu Val Glu His Val Pro Gly Gln Pro Val Arg Asn

65 70 75 80
Pro Gln Cys Arg Ile Cys Asn Gln Asn Ser Ser Asn Pro Asn Gln Arg

85 90 95
His Pro Ile Thr Asn Ala Ile Asp Gly Lys Asn Thr Trp Trp Gln Ser

100 105 110
Pro Ser Ile Lys Asn Gly Ile Glu Tyr His Tyr Val Thr Ile Thr Leu

115 120 125
Asp Leu Gln Gln Val Phe Gln Ile Ala Tyr Val Ile Val Lys Ala Ala

130 135 140
Asn Ser Pro Arg Pro Gly Asn Trp Ile Leu Glu Arg Ser Leu Asp Asp

145 150 155 160
Val Glu Tyr Lys Pro Trp Gln Tyr His Ala Val Thr Asp Thr Glu Cys

165 170 175
Leu Thr Leu Tyr Asn Ile Tyr Pro Arg Thr Gly Pro Pro Ser Tyr Ala

180 185 190
Lys Asp Asp Glu Val Ile Cys Thr Ser Phe Tyr Ser Lys Ile His Pro

195 200 205
Leu Glu Asn Gly Glu Ile His Ile Ser Leu Ile Asn Gly Arg Pro Ser

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

 260 265 270
 Arg Tyr Tyr Tyr Ser Val Lys Asp Ile Ser Val Gly Gly Met Cys Ile
 275 280 285
 Cys Tyr Gly His Ala Arg Ala Cys Pro Leu Asp Pro Ala Thr Asn Lys
 290 295 300
 Ser Arg Cys Glu Cys Glu His Asn Thr Cys Gly Asp Ser Cys Asp Gln
 305 310 315 320
 Cys Cys Pro Gly Phe His Gln Lys Pro Trp Arg Ala Gly Thr Phe Leu

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

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

 450 455 460

Asp Cys Lys Ala Cys Asn Cys Ser Gly Leu Gly Ser Lys Asn Glu Asp
 465 470 475 480
 Pro Cys Phe Gly Pro Cys Ile Cys Lys Glu Asn Val Glu Gly Gly Asp
 485 490 495
 Cys Ser Arg Cys Lys Ser Gly Phe Phe Asn Leu Gln Glu Asp Asn Trp
 500 505 510
 Lys Gly Cys Asp Glu Cys Phe Cys Ser Gly Val Ser Asn Arg Cys Gln
 515 520 525
 Ser Ser Tyr Trp Thr Tyr Gly Lys Ile Gln Asp Met Ser Gly Trp Tyr
 530 535 540
 Leu Thr Asp Leu Pro Gly Arg Ile Arg Val Ala Pro Gln Gln Asp Asp
 545 550 555 560
 Leu Asp Ser Pro Gln Gln Ile Ser Ile Ser Asn Ala Glu Ala Arg Gln
 565 570 575
 Ala Leu Pro His Ser Tyr Tyr Trp Ser Ala Pro Ala Pro Tyr Leu Gly
 580 585 590
 Asn Lys Leu Pro Ala Val Gly Gly Gln Leu Thr Phe Thr Ile Ser Tyr
 595 600 605
 Asp Leu Glu Glu Glu Glu Glu Asp Thr Glu Arg Val Leu Gln Leu Met
 610 615 620
 Ile Ile Leu Glu Gly Asn Asp Leu Ser Ile Ser Thr Ala Gln Asp Glu
 625 630 635 640
 Val Tyr Leu His Pro Ser Glu Glu His Thr Asn Val Leu Leu Leu Lys
 645 650 655
 Glu Glu Ser Phe Thr Ile His Gly Thr His Phe Pro Val Arg Arg Lys
 660 665 670
 Glu Phe Met Thr Val Leu Ala Asn Leu Lys Arg Val Leu Leu Gln Ile
 675 680 685
 Thr Tyr Ser Phe Gly Met Asp Ala Ile Phe Arg Leu Ser Ser Val Asn
 690 695 700
 Leu Glu Ser Ala Val Ser Tyr Pro Thr Asp Gly Ser Ile Ala Ala Ala

705 710 715 720
 Val Glu Val Cys Gln Cys Pro Pro Gly Tyr Thr Gly Ser Ser Cys Glu
 725 730 735
 Ser Cys Trp Pro Arg His Arg Arg Val Asn Gly Thr Ile Phe Gly Gly
 740 745 750
 Ile Cys Glu Pro Cys Gln Cys Phe Gly His Ala Glu Ser Cys Asp Asp
 755 760 765
 Val Thr Gly Glu Cys Leu Asn Cys Lys Asp His Thr Gly Gly Pro Tyr

 770 775 780
 Cys Asp Lys Cys Leu Pro Gly Phe Tyr Gly Glu Pro Thr Lys Gly Thr
 785 790 795 800
 Ser Glu Asp Cys Gln Pro Cys Ala Cys Pro Leu Asn Ile Pro Ser Asn
 805 810 815
 Asn Phe Ser Pro Thr Cys His Leu Asp Arg Ser Leu Gly Leu Ile Cys
 820 825 830
 Asp Gly Cys Pro Val Gly Tyr Thr Gly Pro Arg Cys Glu Arg Cys Ala

 835 840 845
 Glu Gly Tyr Phe Gly Gln Pro Ser Val Pro Gly Gly Ser Cys Gln Pro
 850 855 860
 Cys Gln Cys Asn Asp Asn Leu Asp Phe Ser Ile Pro Gly Ser Cys Asp
 865 870 875 880
 Ser Leu Ser Gly Ser Cys Leu Ile Cys Lys Pro Gly Thr Thr Gly Arg
 885 890 895
 Tyr Cys Glu Leu Cys Ala Asp Gly Tyr Phe Gly Asp Ala Val Asp Ala

 900 905 910
 Lys Asn Cys Gln Pro Cys Arg Cys Asn Ala Gly Gly Ser Phe Ser Glu
 915 920 925
 Val Cys His Ser Gln Thr Gly Gln Cys Glu Cys Arg Ala Asn Val Gln
 930 935 940
 Gly Gln Arg Cys Asp Lys Cys Lys Pro Asn Met Trp Arg Asp Pro Glu
 945 950 955 960

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

965 970 975

Pro Gln Cys Asp Ile Thr Gly Arg Cys Val Cys Lys Ser Gly Phe Val

980 985 990

Gly Lys Gln Cys Asn Leu Gly Arg Gln Val His Gln Gln Glu Glu Gln

995 1000 1005

Pro Arg Arg Ala Gln Arg Val Leu Gly Ser Pro Gln Arg Trp Ala Ile

1010 1015 1020

Gly Ser Ser Ser Gly Cys Pro Arg Gly Ala Tyr Arg Ala Pro Ala Pro

1025 1030 1035 1040

Ala Gly Thr Phe Gly Leu Gln Ser Ala Arg Gly Cys Val Pro Cys Asn

1045 1050 1055

Cys Asn Ser Phe Gly Ser Lys Ser Phe Asp Cys Glu Glu Ser Gly Gln

1060 1065 1070

Cys Trp Cys Gln Pro Gly Val Thr Gly Lys Lys Cys Asp Arg Cys Ala

1075 1080 1085

His Gly Tyr Phe Asn Phe Gln Glu Gly Gly Cys Thr Ala Cys Glu Cys

1090 1095 1100

Ser His Leu Gly Asn Asn Cys Asp Pro Lys Thr Gly Arg Cys Ile Cys

1105 1110 1115 1120

Pro Pro Asn Thr Ile Gly Glu Lys Cys Ser Lys Cys Ala Pro Asn Thr

1125 1130 1135

Trp Gly His Ser Ile Thr Thr Gly Cys Lys Ala Cys Asn Cys Ser Thr

1140 1145 1150

Val Gly Ser Leu Asp Phe Gln Cys Asn Val Asn Thr Gly Gln Cys Asn

1155 1160 1165

Cys His Pro Lys Phe Ser Gly Ala Lys Cys Thr Glu Cys Ser Arg Gly

1170 1175 1180

His Trp Asn Tyr Pro Arg Cys Asn Leu Cys Asp Cys Phe Leu Pro Gly

1185 1190 1195 1200

Thr Asp Ala Thr Thr Cys Asp Ser Glu Thr Lys Lys Cys Ser Cys Ser

1205 1210 1215
Asp Gln Thr Gly Gln Cys Thr Cys Lys Val Asn Val Glu Gly Ile His

1220 1225 1230
Cys Asp Arg Cys Arg Pro Gly Lys Phe Gly Leu Asp Ala Lys Asn Pro

1235 1240 1245
Leu Gly Cys Ser Ser Cys Tyr Cys Phe Gly Thr Thr Thr Gln Cys Ser

1250 1255 1260
Glu Ala Lys Gly Leu Ile Arg Thr Trp Val Thr Leu Lys Ala Glu Gln

1265 1270 1275 1280
Thr Ile Leu Pro Leu Val Asp Glu Ala Leu Gln His Thr Thr Thr Lys

1285 1290 1295
Gly Ile Val Phe Gln His Pro Glu Ile Val Ala His Met Asp Leu Met

1300 1305 1310
Arg Glu Asp Leu His Leu Glu Pro Phe Tyr Trp Lys Leu Pro Glu Gln

1315 1320 1325
Phe Glu Gly Lys Lys Leu Met Ala Tyr Gly Gly Lys Leu Lys Tyr Ala

1330 1335 1340
Ile Tyr Phe Glu Ala Arg Glu Glu Thr Gly Phe Ser Thr Tyr Asn Pro

1345 1350 1355 1360
Gln Val Ile Ile Arg Gly Gly Thr Pro Thr His Ala Arg Ile Ile Val

1365 1370 1375
Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr Arg His Glu Ile

1380 1385 1390
Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp Asp Pro Arg Val

1395 1400 1405
His Arg Thr Val Thr Arg Glu Asp Phe Leu Asp Ile Leu Tyr Asp Ile

1410 1415 1420
His Tyr Ile Leu Ile Lys Ala Thr Tyr Gly Asn Phe Met Arg Gln Ser

1425 1430 1435 1440
Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Gln Gly Arg Gly Thr

1445 1450 1455

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

1460 1465 1470

Gly Tyr Ser Gly Leu Ser Cys Glu Ala Cys Leu Pro Gly Phe Tyr Arg

1475 1480 1485

Leu Arg Ser Gln Pro Gly Gly Arg Thr Pro Gly Pro Thr Leu Gly Thr

1490 1495 1500

Cys Val Pro Cys Gln Cys Asn Gly His Ser Ser Leu Cys Asp Pro Glu

1505 1510 1515 1520

Thr Ser Ile Cys Gln Asn Cys Gln His His Thr Ala Gly Asp Phe Cys

1525 1530 1535

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

1540 1545 1550

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

1555 1560 1565

Ser Pro Ser Cys Val Ala Glu Gly Leu Asp Asp Tyr Arg Cys Thr Ala

1570 1575 1580

Cys Pro Arg Gly Tyr Glu Gly Gln Tyr Cys Glu Arg Cys Ala Pro Gly

1585 1590 1595 1600

Tyr Thr Gly Ser Pro Gly Asn Pro Gly Gly Ser Cys Gln Glu Cys Glu

1605 1610 1615

Cys Asp Pro Tyr Gly Ser Leu Pro Val Pro Cys Asp Pro Val Thr Gly

1620 1625 1630

Phe Cys Thr Cys Arg Pro Gly Ala Thr Gly Arg Lys Cys Asp Gly Cys

1635 1640 1645

Lys His Trp His Ala Arg Glu Gly Trp Glu Cys Val Phe Cys Gly Asp

1650 1655 1660

Glu Cys Thr Gly Leu Leu Leu Gly Asp Leu Ala Arg Leu Glu Gln Met

1665 1670 1675 1680

Val Met Ser Ile Asn Leu Thr Gly Pro Leu Pro Ala Pro Tyr Lys Met

1685 1690 1695

Leu Tyr Gly Leu Glu Asn Met Thr Gln Glu Leu Lys His Leu Leu Ser

1700 1705 1710
 Pro Gln Arg Ala Pro Glu Arg Leu Ile Gln Leu Ala Glu Gly Asn Leu
 1715 1720 1725
 Asn Thr Leu Val Thr Glu Met Asn Glu Leu Leu Thr Arg Ala Thr Lys

 1730 1735 1740
 Val Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala Glu Arg Thr Asn
 1745 1750 1755 1760
 Thr Arg Ala Lys Ser Leu Gly Glu Phe Ile Lys Glu Leu Ala Arg Asp
 1765 1770 1775
 Ala Glu Ala Val Asn Glu Lys Ala Ile Lys Leu Asn Glu Thr Leu Gly
 1780 1785 1790
 Thr Arg Asp Glu Ala Phe Glu Arg Asn Leu Glu Gly Leu Gln Lys Glu

 1795 1800 1805
 Ile Asp Gln Met Ile Lys Glu Leu Arg Arg Lys Asn Leu Glu Thr Gln
 1810 1815 1820
 Lys Glu Ile Ala Glu Asp Glu Leu Val Ala Ala Glu Ala Leu Leu Lys
 1825 1830 1835 1840
 Lys Val Lys Lys Leu Phe Gly Glu Ser Arg Gly Glu Asn Glu Glu Met
 1845 1850 1855
 Glu Lys Asp Leu Arg Glu Lys Leu Ala Asp Tyr Lys Asn Lys Val Asp

 1860 1865 1870
 Asp Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys Ile Arg Glu Ala
 1875 1880 1885
 Asn Arg Leu Phe Ala Val Asn Gln Lys Asn Met Thr Ala Leu Glu Lys
 1890 1895 1900
 Lys Lys Glu Ala Val Glu Ser Gly Lys Arg Gln Ile Glu Asn Thr Leu
 1905 1910 1915 1920
 Lys Glu Gly Asn Asp Ile Leu Asp Glu Ala Asn Arg Leu Ala Asp Glu

 1925 1930 1935
 Ile Asn Ser Ile Ile Asp Tyr Val Glu Asp Ile Gln Thr Lys Leu Pro
 1940 1945 1950

Pro Met Ser Glu Glu Leu Asn Asp Lys Ile Asp Asp Leu Ser Gln Glu

1955 1960 1965

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

1970 1975 1980

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

1985 1990 1995 2000

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

2005 2010 2015

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

2020 2025 2030

Ala Lys Asp Leu Ala His Glu Ala Thr Lys Leu Ala Thr Gly Pro Arg

2035 2040 2045

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

2050 2055 2060

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

2065 2070 2075 2080

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

2085 2090 2095

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

2100 2105 2110

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

2115 2120 2125

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

2130 2135 2140

Glu Leu His Gln Asn Leu Asp Gly Leu Lys Lys Asn Tyr Asn Lys Leu

2145 2150 2155 2160

Ala Asp Ser Val Ala Lys Thr Asn Ala Val Val Lys Asp Pro Ser Lys

2165 2170 2175

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

2180 2185 2190

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

2195	2200	2205	
Asp Asn Leu Lys Lys Asn Ile Ser Glu Ile Lys Glu Leu Ile Asn Gln			
2210	2215	2220	
Ala Arg Lys Gln Ala Asn Ser Ile Lys Val Ser Val Ser Ser Gly Gly			
2225	2230	2235	2240
Asp Cys Ile Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly Ser Tyr Asn			
	2245	2250	2255
Asn Ile Val Val Asn Val Lys Thr Ala Val Ala Asp Asn Leu Leu Phe			
2260	2265	2270	
Tyr Leu Gly Ser Ala Lys Phe Ile Asp Phe Leu Ala Ile Glu Met Arg			
2275	2280	2285	
Lys Gly Lys Val Ser Phe Leu Trp Asp Val Gly Ser Gly Val Gly Arg			
2290	2295	2300	
Val Glu Tyr Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp Tyr Arg Ile			
2305	2310	2315	2320
Val Ala Ser Arg Thr Gly Arg Asn Gly Thr Ile Ser Val Arg Ala Leu			
2325	2330	2335	
Asp Gly Pro Lys Ala Ser Ile Val Pro Ser Thr His His Ser Thr Ser			
2340	2345	2350	
Pro Pro Gly Tyr Thr Ile Leu Asp Val Asp Ala Asn Ala Met Leu Phe			
2355	2360	2365	
Val Gly Gly Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala Val Arg Val			
2370	2375	2380	
Ile Thr Phe Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp Asn Lys Pro			
2385	2390	2395	2400
Ile Gly Leu Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys Lys Gly Cys			
2405	2410	2415	
Thr Val Ser Pro Gln Val Glu Asp Ser Glu Gly Thr Ile Gln Phe Asp			
2420	2425	2430	
Gly Glu Gly Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp Tyr Pro Asn			
2435	2440	2445	

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

2450 2455 2460

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

2465 2470 2475 2480

Leu Thr Asp Gly His Ile Lys Val Ser Tyr Asp Leu Gly Ser Gly Met

2485 2490 2495

Ala Ser Val Val Ser Asn Gln Asn His Asn Asp Gly Lys Trp Lys Ser

2500 2505 2510

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

2515 2520 2525

Ile Asp Thr Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser Ser Gly Asn

2530 2535 2540

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

2545 2550 2555 2560

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

2565 2570 2575

Leu Lys Lys Tyr Ser Gly Cys Leu Lys Asp Ile Glu Ile Ser Arg Thr

2580 2585 2590

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

2595 2600 2605

Cys Ser Leu Glu Asn Val Tyr Thr Val Ser Phe Pro Lys Pro Gly Phe

2610 2615 2620

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

2625 2630 2635 2640

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

2645 2650 2655

Gly Thr Pro Ala Pro Pro Arg Arg Lys Arg Arg Gln Thr Gly Gln Ala

2660 2665 2670

Tyr Tyr Ala Ile Leu Leu Asn Arg Gly Arg Leu Glu Val His Leu Ser

2675 2680 2685

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

2690 2695 2700
 Leu Phe His Asp Gly Arg Glu His Ser Val His Val Glu Arg Thr Arg
 2705 2710 2715 2720
 Gly Ile Phe Thr Val Gln Val Asp Glu Asn Arg Arg Tyr Met Gln Asn
 2725 2730 2735
 Leu Thr Val Glu Gln Pro Ile Glu Val Lys Lys Leu Phe Val Gly Gly
 2740 2745 2750
 Ala Pro Pro Glu Phe Gln Pro Ser Pro Leu Arg Asn Ile Pro Pro Phe

 2755 2760 2765
 Glu Gly Cys Ile Trp Asn Leu Val Ile Asn Ser Val Pro Met Asp Phe
 2770 2775 2780
 Ala Arg Pro Val Ser Phe Lys Asn Ala Asp Ile Gly Arg Cys Ala His
 2785 2790 2795 2800
 Gln Lys Leu Arg Glu Asp Glu Asp Gly Ala Ala Pro Ala Glu Ile Val
 2805 2810 2815
 Ile Gln Pro Glu Pro Val Pro Thr Pro Ala Phe Pro Thr Pro Thr Pro

 2820 2825 2830
 Val Leu Thr His Gly Pro Cys Ala Ala Glu Ser Glu Pro Ala Leu Leu
 2835 2840 2845
 Ile Gly Ser Lys Gln Phe Gly Leu Ser Arg Asn Ser His Ile Ala Ile
 2850 2855 2860
 Ala Phe Asp Asp Thr Lys Val Lys Asn Arg Leu Thr Ile Glu Leu Glu
 2865 2870 2875 2880
 Val Arg Thr Glu Ala Glu Ser Gly Leu Leu Phe Tyr Met Ala Arg Ile

 2885 2890 2895
 Asn His Ala Asp Phe Ala Thr Val Gln Leu Arg Asn Gly Leu Pro Tyr
 2900 2905 2910
 Phe Ser Tyr Asp Leu Gly Ser Gly Asp Thr His Thr Met Ile Pro Thr
 2915 2920 2925
 Lys Ile Asn Asp Gly Gln Trp His Lys Ile Lys Ile Met Arg Ser Lys
 2930 2935 2940

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

2945 2950 2955 2960

Pro Lys Lys Ala Asp Ile Leu Asp Val Val Gly Met Leu Tyr Val Gly

2965 2970 2975

Gly Leu Pro Ile Asn Tyr Thr Thr Arg Arg Ile Gly Pro Val Thr Tyr

2980 2985 2990

Ser Ile Asp Gly Cys Val Arg Asn Leu His Met Ala Glu Ala Pro Ala

2995 3000 3005

Asp Leu Glu Gln Pro Thr Ser Ser Phe His Val Gly Thr Cys Phe Ala

3010 3015 3020

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

3025 3030 3035 3040

Val Gly Gly Phe Lys Val Gly Leu Asp Leu Leu Val Glu Phe Glu Phe

3045 3050 3055

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

3060 3065 3070

Met Asp Gly Met Gly Ile Glu Met Ile Asp Glu Lys Leu Met Phe His

3075 3080 3085

Val Asp Asn Gly Ala Gly Arg Phe Thr Ala Val Tyr Asp Ala Gly Val

3090 3095 3100

Pro Gly His Leu Cys Asp Gly Gln Trp His Lys Val Thr Ala Asn Lys

3105 3110 3115 3120

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

3125 3130 3135

Gln Ser Pro Asn Pro Ala Ser Thr Ser Ala Asp Thr Asn Asp Pro Val

3140 3145 3150

Phe Val Gly Gly Phe Pro Asp Asp Leu Lys Gln Phe Gly Leu Thr Thr

3155 3160 3165

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

3170 3175 3180

Thr Gly Lys Pro Leu Glu Val Asn Phe Ala Lys Ala Leu Glu Leu Arg

3185 3190 3195 3200
 Gly Val Gln Pro Val Ser Cys Pro Ala Asn
 3205 3210

<210> 117
 <211> 3207
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> laminin subunit α2 isoform?X2
 <400> 117

Met Pro Gly Ala Ala Gly Val Leu Leu Leu Leu Leu Ser Gly Gly
 1 5 10 15
 Leu Gly Gly Val Gln Ala Gln Arg Pro Gln Gln Gln Arg Gln Ser Gln
 20 25 30
 Ala His Gln Gln Arg Gly Leu Phe Pro Ala Val Leu Asn Leu Ala Ser
 35 40 45

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

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

				165				170				175			
Leu	Thr	Leu	Tyr	Asn	Ile	Tyr	Pro	Arg	Thr	Gly	Pro	Pro	Ser	Tyr	Ala
				180				185				190			
Lys	Asp	Asp	Glu	Val	Ile	Cys	Thr	Ser	Phe	Tyr	Ser	Lys	Ile	His	Pro
				195				200				205			
Leu	Glu	Asn	Gly	Glu	Ile	His	Ile	Ser	Leu	Ile	Asn	Gly	Arg	Pro	Ser
				210				215				220			
Ala	Asp	Asp	Pro	Ser	Pro	Glu	Leu	Leu	Glu	Phe	Thr	Ser	Ala	Arg	Tyr
225				230				235				240			
Ile	Arg	Leu	Arg	Phe	Gln	Arg	Ile	Arg	Thr	Leu	Asn	Ala	Asp	Leu	Met
				245				250				255			
Met	Phe	Ala	His	Lys	Asp	Pro	Arg	Glu	Ile	Asp	Pro	Ile	Val	Thr	Arg
				260				265				270			
Arg	Tyr	Tyr	Tyr	Ser	Val	Lys	Asp	Ile	Ser	Val	Gly	Gly	Met	Cys	Ile
				275				280				285			
Cys	Tyr	Gly	His	Ala	Arg	Ala	Cys	Pro	Leu	Asp	Pro	Ala	Thr	Asn	Lys
290				295				300							
Ser	Arg	Cys	Glu	Cys	Glu	His	Asn	Thr	Cys	Gly	Asp	Ser	Cys	Asp	Gln
305				310				315				320			
Cys	Cys	Pro	Gly	Phe	His	Gln	Lys	Pro	Trp	Arg	Ala	Gly	Thr	Phe	Leu
				325				330				335			
Thr	Lys	Thr	Glu	Cys	Glu	Ala	Cys	Asn	Cys	His	Gly	Lys	Ala	Glu	Glu
				340				345				350			
Cys	Tyr	Tyr	Asp	Glu	Asn	Val	Ala	Arg	Arg	Asn	Leu	Ser	Leu	Asn	Ile
				355				360				365			
Arg	Gly	Lys	Tyr	Ile	Gly	Gly	Gly	Val	Cys	Ile	Asn	Cys	Thr	Gln	Asn
370				375				380							
Thr	Ala	Gly	Ile	Asn	Cys	Glu	Thr	Cys	Thr	Asp	Gly	Phe	Phe	Arg	Pro
385				390				395				400			
Lys	Gly	Val	Ser	Pro	Asn	Tyr	Pro	Arg	Pro	Cys	Gln	Pro	Cys	His	Cys
				405				410				415			

Asp Pro Ile Gly Ser Leu Asn Glu Val Cys Val Lys Asp Glu Lys His
420 425 430

Ala Arg Arg Gly Leu Ala Pro Gly Ser Cys His Cys Lys Thr Gly Phe
435 440 445

Gly Gly Val Ser Cys Asp Arg Cys Ala Arg Gly Tyr Thr Gly Tyr Pro
450 455 460

Asp Cys Lys Ala Cys Asn Cys Ser Gly Leu Gly Ser Lys Asn Glu Asp
465 470 475 480

Pro Cys Phe Gly Pro Cys Ile Cys Lys Glu Asn Val Glu Gly Gly Asp
485 490 495

Cys Ser Arg Cys Lys Ser Gly Phe Phe Asn Leu Gln Glu Asp Asn Trp
500 505 510

Lys Gly Cys Asp Glu Cys Phe Cys Ser Gly Val Ser Asn Arg Cys Gln
515 520 525

Ser Ser Tyr Trp Thr Tyr Gly Lys Ile Gln Asp Met Ser Gly Trp Tyr
530 535 540

Leu Thr Asp Leu Pro Gly Arg Ile Arg Val Ala Pro Gln Gln Asp Asp
545 550 555 560

Leu Asp Ser Pro Gln Gln Ile Ser Ile Ser Asn Ala Glu Ala Arg Gln
565 570 575

Ala Leu Pro His Ser Tyr Tyr Trp Ser Ala Pro Ala Pro Tyr Leu Gly
580 585 590

Asn Lys Leu Pro Ala Val Gly Gly Gln Leu Thr Phe Thr Ile Ser Tyr
595 600 605

Asp Leu Glu Glu Glu Glu Glu Asp Thr Glu Arg Val Leu Gln Leu Met
610 615 620

Ile Ile Leu Glu Gly Asn Asp Leu Ser Ile Ser Thr Ala Gln Asp Glu
625 630 635 640

Val Tyr Leu His Pro Ser Glu Glu His Thr Asn Val Leu Leu Leu Lys
645 650 655

Glu Glu Ser Phe Thr Ile His Gly Thr His Phe Pro Val Arg Arg Lys

660 665 670
 Glu Phe Met Thr Val Leu Ala Asn Leu Lys Arg Val Leu Leu Gln Ile
 675 680 685

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

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

 Asn Phe Ser Pro Thr Cys His Leu Asp Arg Ser Leu Gly Leu Ile Cys
 820 825 830
 Asp Gly Cys Pro Val Gly Tyr Thr Gly Pro Arg Cys Glu Arg Cys Ala
 835 840 845
 Glu Gly Tyr Phe Gly Gln Pro Ser Val Pro Gly Gly Ser Cys Gln Pro
 850 855 860
 Cys Gln Cys Asn Asp Asn Leu Asp Phe Ser Ile Pro Gly Ser Cys Asp
 865 870 875 880

 Ser Leu Ser Gly Ser Cys Leu Ile Cys Lys Pro Gly Thr Thr Gly Arg
 885 890 895
 Tyr Cys Glu Leu Cys Ala Asp Gly Tyr Phe Gly Asp Ala Val Asp Ala
 900 905 910

Lys Asn Cys Gln Pro Cys Arg Cys Asn Ala Gly Gly Ser Phe Ser Glu

915 920 925

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

930 935 940

Gly Gln Arg Cys Asp Lys Cys Lys Pro Asn Met Trp Arg Asp Pro Glu

945 950 955 960

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

965 970 975

Pro Gln Cys Asp Ile Thr Gly Arg Cys Val Cys Lys Ser Gly Phe Val

980 985 990

Gly Lys Gln Cys Asn Leu Gly Arg Gln Val His Gln Gln Glu Glu Gln

995 1000 1005

Pro Arg Arg Ala Gln Arg Val Leu Gly Ser Pro Gln Arg Trp Ala Ile

1010 1015 1020

Gly Ser Ser Ser Gly Cys Pro Arg Gly Ala Tyr Arg Ala Pro Ala Pro

1025 1030 1035 1040

Ala Gly Thr Phe Gly Leu Gln Ser Ala Arg Gly Cys Val Pro Cys Asn

1045 1050 1055

Cys Asn Ser Phe Gly Ser Lys Ser Phe Asp Cys Glu Glu Ser Gly Gln

1060 1065 1070

Cys Trp Cys Gln Pro Gly Val Thr Gly Lys Lys Cys Asp Arg Cys Ala

1075 1080 1085

His Gly Tyr Phe Asn Phe Gln Glu Gly Gly Cys Thr Ala Cys Glu Cys

1090 1095 1100

Ser His Leu Gly Asn Asn Cys Asp Pro Lys Thr Gly Arg Cys Ile Cys

1105 1110 1115 1120

Pro Pro Asn Thr Ile Gly Glu Lys Cys Ser Lys Cys Ala Pro Asn Thr

1125 1130 1135

Trp Gly His Ser Ile Thr Thr Gly Cys Lys Ala Cys Asn Cys Ser Thr

1140 1145 1150

Val Gly Ser Leu Asp Phe Gln Cys Asn Val Asn Thr Gly Gln Cys Asn

1155	1160	1165	
Cys His Pro Lys Phe Ser Gly Ala Lys Cys Thr Glu Cys Ser Arg Gly			
1170	1175	1180	
His Trp Asn Tyr Pro Arg Cys Asn Leu Cys Asp Cys Phe Leu Pro Gly			
1185	1190	1195	1200
Thr Asp Ala Thr Thr Cys Asp Ser Glu Thr Lys Lys Cys Ser Cys Ser			
1205	1210	1215	
Asp Gln Thr Gly Gln Cys Thr Cys Lys Val Asn Val Glu Gly Ile His			
1220	1225	1230	
Cys Asp Arg Cys Arg Pro Gly Lys Phe Gly Leu Asp Ala Lys Asn Pro			
1235	1240	1245	
Leu Gly Cys Ser Ser Cys Tyr Cys Phe Gly Thr Thr Thr Gln Cys Ser			
1250	1255	1260	
Glu Ala Lys Gly Leu Ile Arg Thr Trp Val Thr Leu Lys Ala Glu Gln			
1265	1270	1275	1280
Thr Ile Leu Pro Leu Val Asp Glu Ala Leu Gln His Thr Thr Thr Lys			
1285	1290	1295	
Gly Ile Val Phe Gln His Pro Glu Ile Val Ala His Met Asp Leu Met			
1300	1305	1310	
Arg Glu Asp Leu His Leu Glu Pro Phe Tyr Trp Lys Leu Pro Glu Gln			
1315	1320	1325	
Phe Glu Gly Lys Lys Leu Met Ala Tyr Gly Gly Lys Leu Lys Tyr Ala			
1330	1335	1340	
Ile Tyr Phe Glu Ala Arg Glu Glu Thr Gly Phe Ser Thr Tyr Asn Pro			
1345	1350	1355	1360
Gln Val Ile Ile Arg Gly Gly Thr Pro Thr His Ala Arg Ile Ile Val			
1365	1370	1375	
Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr Arg His Glu Ile			
1380	1385	1390	
Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp Asp Pro Arg Val			
1395	1400	1405	

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

1410 1415 1420

His Tyr Ile Leu Ile Lys Ala Thr Tyr Gly Asn Phe Met Arg Gln Ser

1425 1430 1435 1440

Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Gln Gly Arg Gly Thr

1445 1450 1455

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

1460 1465 1470

Gly Tyr Ser Gly Leu Ser Cys Glu Ala Cys Leu Pro Gly Phe Tyr Arg

1475 1480 1485

Leu Arg Ser Gln Pro Gly Gly Arg Thr Pro Gly Pro Thr Leu Gly Thr

1490 1495 1500

Cys Val Pro Cys Gln Cys Asn Gly His Ser Ser Leu Cys Asp Pro Glu

1505 1510 1515 1520

Thr Ser Ile Cys Gln Asn Cys Gln His His Thr Ala Gly Asp Phe Cys

1525 1530 1535

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

1540 1545 1550

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

1555 1560 1565

Ser Pro Ser Cys Val Ala Glu Gly Leu Asp Asp Tyr Arg Cys Thr Ala

1570 1575 1580

Cys Pro Arg Gly Tyr Glu Gly Gln Tyr Cys Glu Arg Cys Ala Pro Gly

1585 1590 1595 1600

Tyr Thr Gly Ser Pro Gly Asn Pro Gly Ser Cys Gln Glu Cys Glu Cys

1605 1610 1615

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

1620 1625 1630

Cys Thr Cys Arg Pro Gly Ala Thr Gly Arg Lys Cys Asp Gly Cys Lys

1635 1640 1645

His Trp His Ala Arg Glu Gly Trp Glu Cys Val Phe Cys Gly Asp Glu

1650	1655	1660	
Cys Thr Gly Leu Leu Leu Gly Asp Leu Ala Arg Leu Glu Gln Met Val			
1665	1670	1675	1680
Met Ser Ile Asn Leu Thr Gly Pro Leu Pro Ala Pro Tyr Lys Met Leu			
	1685	1690	1695
Tyr Gly Leu Glu Asn Met Thr Gln Glu Leu Lys His Leu Leu Ser Pro			
	1700	1705	1710
Gln Arg Ala Pro Glu Arg Leu Ile Gln Leu Ala Glu Gly Asn Leu Asn			
	1715	1720	1725
Thr Leu Val Thr Glu Met Asn Glu Leu Leu Thr Arg Ala Thr Lys Val			
	1730	1735	1740
Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala Glu Arg Thr Asn Thr			
1745	1750	1755	1760
Arg Ala Lys Ser Leu Gly Glu Phe Ile Lys Glu Leu Ala Arg Asp Ala			
	1765	1770	1775
Glu Ala Val Asn Glu Lys Ala Ile Lys Leu Asn Glu Thr Leu Gly Thr			
	1780	1785	1790
Arg Asp Glu Ala Phe Glu Arg Asn Leu Glu Gly Leu Gln Lys Glu Ile			
	1795	1800	1805
Asp Gln Met Ile Lys Glu Leu Arg Arg Lys Asn Leu Glu Thr Gln Lys			
	1810	1815	1820
Glu Ile Ala Glu Asp Glu Leu Val Ala Ala Glu Ala Leu Leu Lys Lys			
1825	1830	1835	1840
Val Lys Lys Leu Phe Gly Glu Ser Arg Gly Glu Asn Glu Glu Met Glu			
	1845	1850	1855
Lys Asp Leu Arg Glu Lys Leu Ala Asp Tyr Lys Asn Lys Val Asp Asp			
	1860	1865	1870
Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys Ile Arg Glu Ala Asn			
	1875	1880	1885
Arg Leu Phe Ala Val Asn Gln Lys Asn Met Thr Ala Leu Glu Lys Lys			
	1890	1895	1900

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

Lys Asp Arg Lys Leu Ala Glu Lys Val Ser Gln Ala Glu Ser His Ala
 1970 1975 1980
 Ala Gln Leu Asn Asp Ser Ser Ala Val Leu Asp Gly Ile Leu Asp Glu
 1985 1990 1995 2000
 Ala Lys Asn Ile Ser Phe Asn Ala Thr Ala Ala Phe Lys Ala Tyr Ser
 2005 2010 2015
 Asn Ile Lys Asp Tyr Ile Asp Glu Ala Glu Lys Val Ala Lys Glu Ala
 2020 2025 2030

Lys Asp Leu Ala His Glu Ala Thr Lys Leu Ala Thr Gly Pro Arg Gly
 2035 2040 2045
 Leu Leu Lys Glu Asp Ala Lys Gly Cys Leu Gln Lys Ser Phe Arg Ile
 2050 2055 2060
 Leu Asn Glu Ala Lys Lys Leu Ala Asn Asp Val Lys Glu Asn Glu Asp
 2065 2070 2075 2080
 His Leu Asn Gly Leu Lys Thr Arg Ile Glu Asn Ala Asp Ala Arg Asn
 2085 2090 2095

Gly Asp Leu Leu Arg Thr Leu Asn Asp Thr Leu Gly Lys Leu Ser Ala
 2100 2105 2110
 Ile Pro Asn Asp Thr Ala Ala Lys Leu Gln Ala Val Lys Asp Lys Ala
 2115 2120 2125
 Arg Gln Ala Asn Asp Thr Ala Lys Asp Val Leu Ala Gln Ile Thr Glu
 2130 2135 2140
 Leu His Gln Asn Leu Asp Gly Leu Lys Lys Asn Tyr Asn Lys Leu Ala

2145 2150 2155 2160

 Asp Ser Val Ala Lys Thr Asn Ala Val Val Lys Asp Pro Ser Lys Asn
 2165 2170 2175
 Ile Ala Asp Ala Asp Ala Thr Val Lys Asn Leu Glu Gln Glu Ala Asp
 2180 2185 2190
 Arg Leu Ile Asp Lys Leu Lys Pro Ile Lys Glu Leu Glu Asp Asn Leu
 2195 2200 2205
 Lys Lys Asn Ile Ser Glu Ile Lys Glu Leu Ile Asn Gln Ala Arg Lys
 2210 2215 2220

 Gln Ala Asn Ser Ile Lys Val Ser Val Ser Ser Gly Gly Asp Cys Ile
 2225 2230 2235 2240
 Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly Ser Tyr Asn Asn Ile Val
 2245 2250 2255
 Val Asn Val Lys Thr Ala Val Ala Asp Asn Leu Leu Phe Tyr Leu Gly
 2260 2265 2270
 Ser Ala Lys Phe Ile Asp Phe Leu Ala Ile Glu Met Arg Lys Gly Lys
 2275 2280 2285

 Val Ser Phe Leu Trp Asp Val Gly Ser Gly Val Gly Arg Val Glu Tyr
 2290 2295 2300
 Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp Tyr Arg Ile Val Ala Ser
 2305 2310 2315 2320
 Arg Thr Gly Arg Asn Gly Thr Ile Ser Val Arg Ala Leu Asp Gly Pro
 2325 2330 2335
 Lys Ala Ser Ile Val Pro Ser Thr His His Ser Thr Ser Pro Pro Gly
 2340 2345 2350

 Tyr Thr Ile Leu Asp Val Asp Ala Asn Ala Met Leu Phe Val Gly Gly
 2355 2360 2365
 Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala Val Arg Val Ile Thr Phe
 2370 2375 2380
 Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp Asn Lys Pro Ile Gly Leu
 2385 2390 2395 2400

Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys Lys Gly Cys Thr Val Ser
2405 2410 2415

Pro Gln Val Glu Asp Ser Glu Gly Thr Ile Gln Phe Asp Gly Glu Gly
2420 2425 2430

Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp Tyr Pro Asn Ile Ser Thr
2435 2440 2445

Val Met Phe Lys Phe Arg Thr Phe Ser Ser Ser Ala Leu Leu Met Tyr
2450 2455 2460

Leu Ala Thr Arg Asp Leu Arg Asp Phe Met Ser Val Glu Leu Thr Asp
2465 2470 2475 2480

Gly His Ile Lys Val Ser Tyr Asp Leu Gly Ser Gly Met Ala Ser Val
2485 2490 2495

Val Ser Asn Gln Asn His Asn Asp Gly Lys Trp Lys Ser Phe Thr Leu
2500 2505 2510

Ser Arg Ile Gln Lys Gln Ala Asn Ile Ser Ile Val Asp Ile Asp Thr
2515 2520 2525

Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser Ser Gly Asn Asn Phe Gly
2530 2535 2540

Leu Asp Leu Lys Ala Asp Asp Lys Ile Tyr Phe Gly Gly Leu Pro Thr
2545 2550 2555 2560

Leu Arg Asn Leu Ser Met Lys Ala Arg Pro Glu Val Asn Leu Lys Lys
2565 2570 2575

Tyr Ser Gly Cys Leu Lys Asp Ile Glu Ile Ser Arg Thr Pro Tyr Asn
2580 2585 2590

Ile Leu Ser Ser Pro Asp Tyr Val Gly Val Thr Lys Gly Cys Ser Leu
2595 2600 2605

Glu Asn Val Tyr Thr Val Ser Phe Pro Lys Pro Gly Phe Val Glu Leu
2610 2615 2620

Ser Pro Val Pro Ile Asp Val Gly Thr Glu Ile Asn Leu Ser Phe Ser
2625 2630 2635 2640

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

2645 2650 2655
Ala Pro Pro Arg Arg Lys Arg Arg Gln Thr Gly Gln Ala Tyr Tyr Ala
2660 2665 2670

Ile Leu Leu Asn Arg Gly Arg Leu Glu Val His Leu Ser Thr Gly Ala
2675 2680 2685

Arg Thr Met Arg Lys Ile Val Ile Arg Pro Glu Pro Asn Leu Phe His
2690 2695 2700

Asp Gly Arg Glu His Ser Val His Val Glu Arg Thr Arg Gly Ile Phe
2705 2710 2715 2720

Thr Val Gln Val Asp Glu Asn Arg Arg Tyr Met Gln Asn Leu Thr Val
2725 2730 2735

Glu Gln Pro Ile Glu Val Lys Lys Leu Phe Val Gly Gly Ala Pro Pro
2740 2745 2750

Glu Phe Gln Pro Ser Pro Leu Arg Asn Ile Pro Pro Phe Glu Gly Cys
2755 2760 2765

Ile Trp Asn Leu Val Ile Asn Ser Val Pro Met Asp Phe Ala Arg Pro
2770 2775 2780

Val Ser Phe Lys Asn Ala Asp Ile Gly Arg Cys Ala His Gln Lys Leu
2785 2790 2795 2800

Arg Glu Asp Glu Asp Gly Ala Ala Pro Ala Glu Ile Val Ile Gln Pro
2805 2810 2815

Glu Pro Val Pro Thr Pro Ala Phe Pro Thr Pro Thr Pro Val Leu Thr
2820 2825 2830

His Gly Pro Cys Ala Ala Glu Ser Glu Pro Ala Leu Leu Ile Gly Ser
2835 2840 2845

Lys Gln Phe Gly Leu Ser Arg Asn Ser His Ile Ala Ile Ala Phe Asp
2850 2855 2860

Asp Thr Lys Val Lys Asn Arg Leu Thr Ile Glu Leu Glu Val Arg Thr
2865 2870 2875 2880

Glu Ala Glu Ser Gly Leu Leu Phe Tyr Met Ala Arg Ile Asn His Ala
2885 2890 2895

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

2900 2905 2910

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

2915 2920 2925

Asp Gly Gln Trp His Lys Ile Lys Ile Met Arg Ser Lys Gln Glu Gly

2930 2935 2940

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

2945 2950 2955 2960

Ala Asp Ile Leu Asp Val Val Gly Met Leu Tyr Val Gly Gly Leu Pro

2965 2970 2975

Ile Asn Tyr Thr Thr Arg Arg Ile Gly Pro Val Thr Tyr Ser Ile Asp

2980 2985 2990

Gly Cys Val Arg Asn Leu His Met Ala Glu Ala Pro Ala Asp Leu Glu

2995 3000 3005

Gln Pro Thr Ser Ser Phe His Val Gly Thr Cys Phe Ala Asn Ala Gln

3010 3015 3020

Arg Gly Thr Tyr Phe Asp Gly Thr Gly Phe Ala Lys Ala Val Gly Gly

3025 3030 3035 3040

Phe Lys Val Gly Leu Asp Leu Leu Val Glu Phe Glu Phe Arg Thr Thr

3045 3050 3055

Thr Thr Thr Gly Val Leu Leu Gly Ile Ser Ser Gln Lys Met Asp Gly

3060 3065 3070

Met Gly Ile Glu Met Ile Asp Glu Lys Leu Met Phe His Val Asp Asn

3075 3080 3085

Gly Ala Gly Arg Phe Thr Ala Val Tyr Asp Ala Gly Val Pro Gly His

3090 3095 3100

Leu Cys Asp Gly Gln Trp His Lys Val Thr Ala Asn Lys Ile Lys His

3105 3110 3115 3120

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

3125 3130 3135

Asn Pro Ala Ser Thr Ser Ala Asp Thr Asn Asp Pro Val Phe Val Gly

3140 3145 3150
 Gly Phe Pro Asp Asp Leu Lys Gln Phe Gly Leu Thr Thr Ser Ile Pro
 3155 3160 3165
 Phe Arg Gly Cys Ile Arg Ser Leu Lys Leu Thr Lys Gly Thr Gly Lys
 3170 3175 3180

Pro Leu Glu Val Asn Phe Ala Lys Ala Leu Glu Leu Arg Gly Val Gln
 3185 3190 3195 3200
 Pro Val Ser Cys Pro Ala Asn

3205

<210> 118

<211> 3206

<212> PRT

<213> Artificial Sequence

<220>

<223> laminin subunit α2 isoform?X3

<400> 118

Met Pro Gly Ala Ala Gly Val Leu Leu Leu Leu Leu Ser Gly Gly
 1 5 10 15
 Leu Gly Gly Val Gln Ala Gln Arg Pro Gln Gln Gln Arg Gln Ser Gln

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

85 90 95
 His Pro Ile Thr Asn Ala Ile Asp Gly Lys Asn Thr Trp Trp Gln Ser
 100 105 110
 Pro Ser Ile Lys Asn Gly Ile Glu Tyr His Tyr Val Thr Ile Thr Leu
 115 120 125

Asp Leu Gln Gln Val Phe Gln Ile Ala Tyr Val Ile Val Lys Ala Ala
 130 135 140
 Asn Ser Pro Arg Pro Gly Asn Trp Ile Leu Glu Arg Ser Leu Asp Asp

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

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

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

 340 345 350
 Cys Tyr Tyr Asp Glu Asn Val Ala Arg Arg Asn Leu Ser Leu Asn Ile
 355 360 365
 Arg Gly Lys Tyr Ile Gly Gly Gly Val Cys Ile Asn Cys Thr Gln Asn

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

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

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

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

 595 600 605
 Asp Leu Glu Glu Glu Glu Asp Thr Glu Arg Val Leu Gln Leu Met
 610 615 620

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

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

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

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

 850 855 860
 Cys Gln Cys Asn Asp Asn Leu Asp Phe Ser Ile Pro Gly Ser Cys Asp

865 870 875 880
 Ser Leu Ser Gly Ser Cys Leu Ile Cys Lys Pro Gly Thr Thr Gly Arg
 885 890 895
 Tyr Cys Glu Leu Cys Ala Asp Gly Tyr Phe Gly Asp Ala Val Asp Ala
 900 905 910
 Lys Asn Cys Gln Pro Cys Arg Cys Asn Ala Gly Gly Ser Phe Ser Glu

 915 920 925
 Val Cys His Ser Gln Thr Gly Gln Cys Glu Cys Arg Ala Asn Val Gln
 930 935 940
 Gly Gln Arg Cys Asp Lys Cys Lys Pro Asn Met Trp Arg Asp Pro Glu
 945 950 955 960
 Lys Arg Phe Cys Val Leu Cys Asp Cys Asp Pro Val Gly Ser Val Ser
 965 970 975
 Pro Gln Cys Asp Ile Thr Gly Arg Cys Val Cys Lys Ser Gly Phe Val

 980 985 990
 Gly Lys Gln Cys Asn Leu Gly Arg Gln Val His Gln Gln Glu Glu Gln
 995 1000 1005
 Pro Arg Arg Ala Gln Arg Val Leu Gly Ser Pro Gln Arg Trp Ala Ile
 1010 1015 1020
 Gly Ser Ser Ser Gly Cys Pro Arg Gly Ala Tyr Arg Ala Pro Ala Pro
 1025 1030 1035 1040
 Ala Gly Thr Phe Gly Leu Gln Ser Ala Arg Gly Cys Val Pro Cys Asn

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

 1105 1110 1115 1120

Pro Pro Asn Thr Ile Gly Glu Lys Cys Ser Lys Cys Ala Pro Asn Thr

1125 1130 1135

Trp Gly His Ser Ile Thr Thr Gly Cys Lys Ala Cys Asn Cys Ser Thr

1140 1145 1150

Val Gly Ser Leu Asp Phe Gln Cys Asn Val Asn Thr Gly Gln Cys Asn

1155 1160 1165

Cys His Pro Lys Phe Ser Gly Ala Lys Cys Thr Glu Cys Ser Arg Gly

1170 1175 1180

His Trp Asn Tyr Pro Arg Cys Asn Leu Cys Asp Cys Phe Leu Pro Gly

1185 1190 1195 1200

Thr Asp Ala Thr Thr Cys Asp Ser Glu Thr Lys Lys Cys Ser Cys Ser

1205 1210 1215

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

1220 1225 1230

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

1235 1240 1245

Leu Gly Cys Ser Ser Cys Tyr Cys Phe Gly Thr Thr Thr Gln Cys Ser

1250 1255 1260

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

1265 1270 1275 1280

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

1285 1290 1295

Gly Ile Val Phe Gln His Pro Glu Ile Val Ala His Met Asp Leu Met

1300 1305 1310

Arg Glu Asp Leu His Leu Glu Pro Phe Tyr Trp Lys Leu Pro Glu Gln

1315 1320 1325

Phe Glu Gly Lys Lys Leu Met Ala Tyr Gly Gly Lys Leu Lys Tyr Ala

1330 1335 1340

Ile Tyr Phe Glu Ala Arg Glu Glu Thr Gly Phe Ser Thr Tyr Asn Pro

1345 1350 1355 1360

Gln Val Ile Ile Arg Gly Gly Thr Pro Thr His Ala Arg Ile Ile Val

1365	1370	1375	
Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr Arg His Glu Ile			
1380	1385	1390	
Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp Asp Pro Arg Val			
1395	1400	1405	
His Arg Thr Val Thr Arg Glu Asp Phe Leu Asp Ile Leu Tyr Asp Ile			
1410	1415	1420	
His Tyr Ile Leu Ile Lys Ala Thr Tyr Gly Asn Phe Met Arg Gln Ser			
1425	1430	1435	1440
Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Gln Gly Arg Gly Thr			
1445	1450	1455	
Thr Met Thr Pro Pro Ala Asp Leu Ile Glu Lys Cys Asp Cys Pro Leu			
1460	1465	1470	
Gly Tyr Ser Gly Leu Ser Cys Glu Ala Cys Leu Pro Gly Phe Tyr Arg			
1475	1480	1485	
Leu Arg Ser Gln Pro Gly Gly Arg Thr Pro Gly Pro Thr Leu Gly Thr			
1490	1495	1500	
Cys Val Pro Cys Gln Cys Asn Gly His Ser Ser Leu Cys Asp Pro Glu			
1505	1510	1515	1520
Thr Ser Ile Cys Gln Asn Cys Gln His His Thr Ala Gly Asp Phe Cys			
1525	1530	1535	
Glu Arg Cys Ala Leu Gly Tyr Tyr Gly Ile Val Lys Gly Leu Pro Asn			
1540	1545	1550	
Asp Cys Gln Gln Cys Ala Cys Pro Leu Ile Ser Ser Ser Asn Asn Phe			
1555	1560	1565	
Ser Pro Ser Cys Val Ala Glu Gly Leu Asp Asp Tyr Arg Cys Thr Ala			
1570	1575	1580	
Cys Pro Arg Gly Tyr Glu Gly Gln Tyr Cys Glu Arg Cys Ala Pro Gly			
1585	1590	1595	1600
Tyr Thr Gly Ser Pro Gly Asn Pro Gly Gly Ser Cys Gln Glu Cys Glu			
1605	1610	1615	

Cys Asp Pro Tyr Gly Ser Leu Pro Val Pro Cys Asp Pro Val Thr Gly

1620 1625 1630

Phe Cys Thr Cys Arg Pro Gly Ala Thr Gly Arg Lys Cys Asp Gly Cys

1635 1640 1645

Lys His Trp His Ala Arg Glu Gly Trp Glu Cys Val Phe Cys Gly Asp

1650 1655 1660

Glu Cys Thr Gly Leu Leu Leu Gly Asp Leu Ala Arg Leu Glu Gln Met

1665 1670 1675 1680

Val Met Ser Ile Asn Leu Thr Gly Pro Leu Pro Ala Pro Tyr Lys Met

1685 1690 1695

Leu Tyr Gly Leu Glu Asn Met Thr Gln Glu Leu Lys His Leu Leu Ser

1700 1705 1710

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

1715 1720 1725

Asn Thr Leu Val Thr Glu Met Asn Glu Leu Leu Thr Arg Ala Thr Lys

1730 1735 1740

Val Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala Glu Arg Thr Asn

1745 1750 1755 1760

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

1765 1770 1775

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

1780 1785 1790

Thr Arg Asp Glu Ala Phe Glu Arg Asn Leu Glu Gly Leu Gln Lys Glu

1795 1800 1805

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

1810 1815 1820

Lys Glu Ile Ala Glu Asp Glu Leu Val Ala Ala Glu Ala Leu Leu Lys

1825 1830 1835 1840

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

1845 1850 1855

Glu Lys Asp Leu Arg Glu Lys Leu Ala Asp Tyr Lys Asn Lys Val Asp

1860 1865 1870
 Asp Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys Ile Arg Glu Ala

 1875 1880 1885
 Asn Arg Leu Phe Ala Val Asn Gln Lys Asn Met Thr Ala Leu Glu Lys
 1890 1895 1900
 Lys Lys Glu Ala Val Glu Ser Gly Lys Arg Gln Ile Glu Asn Thr Leu
 1905 1910 1915 1920
 Lys Glu Gly Asn Asp Ile Leu Asp Glu Ala Asn Arg Leu Ala Asp Glu
 1925 1930 1935
 Ile Asn Ser Ile Ile Asp Tyr Val Glu Asp Ile Gln Thr Lys Leu Pro

 1940 1945 1950
 Pro Met Ser Glu Glu Leu Asn Asp Lys Ile Asp Asp Leu Ser Gln Glu
 1955 1960 1965
 Ile Lys Asp Arg Lys Leu Ala Glu Lys Val Ser Gln Ala Glu Ser His
 1970 1975 1980
 Ala Ala Gln Leu Asn Asp Ser Ser Ala Val Leu Asp Gly Ile Leu Asp
 1985 1990 1995 2000
 Glu Ala Lys Asn Ile Ser Phe Asn Ala Thr Ala Ala Phe Lys Ala Tyr

 2005 2010 2015
 Ser Asn Ile Lys Asp Tyr Ile Asp Glu Ala Glu Lys Val Ala Lys Glu
 2020 2025 2030
 Ala Lys Asp Leu Ala His Glu Ala Thr Lys Leu Ala Thr Gly Pro Arg
 2035 2040 2045
 Gly Leu Leu Lys Glu Asp Ala Lys Gly Cys Leu Gln Lys Ser Phe Arg
 2050 2055 2060
 Ile Leu Asn Glu Ala Lys Lys Leu Ala Asn Asp Val Lys Glu Asn Glu

 2065 2070 2075 2080
 Asp His Leu Asn Gly Leu Lys Thr Arg Ile Glu Asn Ala Asp Ala Arg
 2085 2090 2095
 Asn Gly Asp Leu Leu Arg Thr Leu Asn Asp Thr Leu Gly Lys Leu Ser
 2100 2105 2110

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

2115 2120 2125

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

2130 2135 2140

Glu Leu His Gln Asn Leu Asp Gly Leu Lys Lys Asn Tyr Asn Lys Leu

2145 2150 2155 2160

Ala Asp Ser Val Ala Lys Thr Asn Ala Val Val Lys Asp Pro Ser Lys

2165 2170 2175

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

2180 2185 2190

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

2195 2200 2205

Asp Asn Leu Lys Lys Asn Ile Ser Glu Ile Lys Glu Leu Ile Asn Gln

2210 2215 2220

Ala Arg Lys Gln Ala Asn Ser Ile Lys Val Ser Val Ser Ser Gly Gly

2225 2230 2235 2240

Asp Cys Ile Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly Ser Tyr Asn

2245 2250 2255

Asn Ile Val Val Asn Val Lys Thr Ala Val Ala Asp Asn Leu Leu Phe

2260 2265 2270

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

2275 2280 2285

Lys Gly Lys Val Ser Phe Leu Trp Asp Val Gly Ser Gly Val Gly Arg

2290 2295 2300

Val Glu Tyr Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp Tyr Arg Ile

2305 2310 2315 2320

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

2325 2330 2335

Asp Gly Pro Lys Ala Ser Ile Val Pro Ser Thr His His Ser Thr Ser

2340 2345 2350

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

2355 2360 2365
 Val Gly Gly Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala Val Arg Val
 2370 2375 2380
 Ile Thr Phe Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp Asn Lys Pro

 2385 2390 2395 2400
 Ile Gly Leu Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys Lys Gly Cys
 2405 2410 2415
 Thr Val Ser Pro Gln Val Glu Asp Ser Glu Gly Thr Ile Gln Phe Asp
 2420 2425 2430
 Gly Glu Gly Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp Tyr Pro Asn
 2435 2440 2445
 Ile Ser Thr Val Met Phe Lys Phe Arg Thr Phe Ser Ser Ser Ala Leu

 2450 2455 2460
 Leu Met Tyr Leu Ala Thr Arg Asp Leu Arg Asp Phe Met Ser Val Glu
 2465 2470 2475 2480
 Leu Thr Asp Gly His Ile Lys Val Ser Tyr Asp Leu Gly Ser Gly Met
 2485 2490 2495
 Ala Ser Val Val Ser Asn Gln Asn His Asn Asp Gly Lys Trp Lys Ser
 2500 2505 2510
 Phe Thr Leu Ser Arg Ile Gln Lys Gln Ala Asn Ile Ser Ile Val Asp

 2515 2520 2525
 Ile Asp Thr Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser Ser Gly Asn
 2530 2535 2540
 Asn Phe Gly Leu Asp Leu Lys Ala Asp Asp Lys Ile Tyr Phe Gly Gly
 2545 2550 2555 2560
 Leu Pro Thr Leu Arg Asn Leu Arg Pro Glu Val Asn Leu Lys Lys Tyr
 2565 2570 2575
 Ser Gly Cys Leu Lys Asp Ile Glu Ile Ser Arg Thr Pro Tyr Asn Ile

 2580 2585 2590
 Leu Ser Ser Pro Asp Tyr Val Gly Val Thr Lys Gly Cys Ser Leu Glu
 2595 2600 2605

Asn Val Tyr Thr Val Ser Phe Pro Lys Pro Gly Phe Val Glu Leu Ser
 2610 2615 2620
 Pro Val Pro Ile Asp Val Gly Thr Glu Ile Asn Leu Ser Phe Ser Thr
 2625 2630 2635 2640
 Lys Asn Glu Ser Gly Ile Ile Leu Leu Gly Ser Gly Gly Thr Pro Ala
 2645 2650 2655
 Pro Pro Arg Arg Lys Arg Arg Gln Thr Gly Gln Ala Tyr Tyr Ala Ile
 2660 2665 2670
 Leu Leu Asn Arg Gly Arg Leu Glu Val His Leu Ser Thr Gly Ala Arg
 2675 2680 2685
 Thr Met Arg Lys Ile Val Ile Arg Pro Glu Pro Asn Leu Phe His Asp
 2690 2695 2700
 Gly Arg Glu His Ser Val His Val Glu Arg Thr Arg Gly Ile Phe Thr
 2705 2710 2715 2720
 Val Gln Val Asp Glu Asn Arg Arg Tyr Met Gln Asn Leu Thr Val Glu
 2725 2730 2735
 Gln Pro Ile Glu Val Lys Lys Leu Phe Val Gly Gly Ala Pro Pro Glu
 2740 2745 2750
 Phe Gln Pro Ser Pro Leu Arg Asn Ile Pro Pro Phe Glu Gly Cys Ile
 2755 2760 2765
 Trp Asn Leu Val Ile Asn Ser Val Pro Met Asp Phe Ala Arg Pro Val
 2770 2775 2780
 Ser Phe Lys Asn Ala Asp Ile Gly Arg Cys Ala His Gln Lys Leu Arg
 2785 2790 2795 2800
 Glu Asp Glu Asp Gly Ala Ala Pro Ala Glu Ile Val Ile Gln Pro Glu
 2805 2810 2815
 Pro Val Pro Thr Pro Ala Phe Pro Thr Pro Thr Pro Val Leu Thr His
 2820 2825 2830
 Gly Pro Cys Ala Ala Glu Ser Glu Pro Ala Leu Leu Ile Gly Ser Lys
 2835 2840 2845
 Gln Phe Gly Leu Ser Arg Asn Ser His Ile Ala Ile Ala Phe Asp Asp

2850 2855 2860
 Thr Lys Val Lys Asn Arg Leu Thr Ile Glu Leu Glu Val Arg Thr Glu
 2865 2870 2875 2880
 Ala Glu Ser Gly Leu Leu Phe Tyr Met Ala Arg Ile Asn His Ala Asp
 2885 2890 2895
 Phe Ala Thr Val Gln Leu Arg Asn Gly Leu Pro Tyr Phe Ser Tyr Asp

 2900 2905 2910
 Leu Gly Ser Gly Asp Thr His Thr Met Ile Pro Thr Lys Ile Asn Asp
 2915 2920 2925
 Gly Gln Trp His Lys Ile Lys Ile Met Arg Ser Lys Gln Glu Gly Ile
 2930 2935 2940
 Leu Tyr Val Asp Gly Ala Ser Asn Arg Thr Ile Ser Pro Lys Lys Ala
 2945 2950 2955 2960
 Asp Ile Leu Asp Val Val Gly Met Leu Tyr Val Gly Gly Leu Pro Ile

 2965 2970 2975
 Asn Tyr Thr Thr Arg Arg Ile Gly Pro Val Thr Tyr Ser Ile Asp Gly
 2980 2985 2990
 Cys Val Arg Asn Leu His Met Ala Glu Ala Pro Ala Asp Leu Glu Gln
 2995 3000 3005
 Pro Thr Ser Ser Phe His Val Gly Thr Cys Phe Ala Asn Ala Gln Arg
 3010 3015 3020
 Gly Thr Tyr Phe Asp Gly Thr Gly Phe Ala Lys Ala Val Gly Gly Phe

 3025 3030 3035 3040
 Lys Val Gly Leu Asp Leu Leu Val Glu Phe Glu Phe Arg Thr Thr Thr
 3045 3050 3055
 Thr Thr Gly Val Leu Leu Gly Ile Ser Ser Gln Lys Met Asp Gly Met
 3060 3065 3070
 Gly Ile Glu Met Ile Asp Glu Lys Leu Met Phe His Val Asp Asn Gly
 3075 3080 3085
 Ala Gly Arg Phe Thr Ala Val Tyr Asp Ala Gly Val Pro Gly His Leu

 3090 3095 3100

Cys Asp Gly Gln Trp His Lys Val Thr Ala Asn Lys Ile Lys His Arg
 3105 3110 3115 3120
 Ile Glu Leu Thr Val Asp Gly Asn Gln Val Glu Ala Gln Ser Pro Asn
 3125 3130 3135
 Pro Ala Ser Thr Ser Ala Asp Thr Asn Asp Pro Val Phe Val Gly Gly
 3140 3145 3150
 Phe Pro Asp Asp Leu Lys Gln Phe Gly Leu Thr Thr Ser Ile Pro Phe

 3155 3160 3165
 Arg Gly Cys Ile Arg Ser Leu Lys Leu Thr Lys Gly Thr Gly Lys Pro
 3170 3175 3180
 Leu Glu Val Asn Phe Ala Lys Ala Leu Glu Leu Arg Gly Val Gln Pro
 3185 3190 3195 3200
 Val Ser Cys Pro Ala Asn
 3205
 <210> 119
 <211> 3212
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> laminin subunit α2 isoform?X4
 <400> 119
 Met Gln Trp Cys Phe Ser Val Leu Met Trp Arg Ser Gln Glu Thr Ser

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

 65 70 75 80
 Arg Asn Pro Gln Cys Arg Ile Cys Asn Gln Asn Ser Ser Asn Pro Asn

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

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

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

 260 265 270
 Thr Arg Arg Tyr Tyr Tyr Ser Val Lys Asp Ile Ser Val Gly Gly Met
 275 280 285
 Cys Ile Cys Tyr Gly His Ala Arg Ala Cys Pro Leu Asp Pro Ala Thr
 290 295 300
 Asn Lys Ser Arg Cys Glu Cys Glu His Asn Thr Cys Gly Asp Ser Cys
 305 310 315 320
 Asp Gln Cys Cys Pro Gly Phe His Gln Lys Pro Trp Arg Ala Gly Thr

 325 330 335

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

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

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

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

580 585 590
 Leu Gly Asn Lys Leu Pro Ala Val Gly Gly Gln Leu Thr Phe Thr Ile
 595 600 605
 Ser Tyr Asp Leu Glu Glu Glu Glu Asp Thr Glu Arg Val Leu Gln
 610 615 620
 Leu Met Ile Ile Leu Glu Gly Asn Asp Leu Ser Ile Ser Thr Ala Gln
 625 630 635 640
 Asp Glu Val Tyr Leu His Pro Ser Glu Glu His Thr Asn Val Leu Leu

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

 705 710 715 720
 Ala Ala Val Glu Val Cys Gln Cys Pro Pro Gly Tyr Thr Gly Ser Ser
 725 730 735
 Cys Glu Ser Cys Trp Pro Arg His Arg Arg Val Asn Gly Thr Ile Phe
 740 745 750
 Gly Gly Ile Cys Glu Pro Cys Gln Cys Phe Gly His Ala Glu Ser Cys
 755 760 765
 Asp Asp Val Thr Gly Glu Cys Leu Asn Cys Lys Asp His Thr Gly Gly

 770 775 780
 Pro Tyr Cys Asp Lys Cys Leu Pro Gly Phe Tyr Gly Glu Pro Thr Lys
 785 790 795 800
 Gly Thr Ser Glu Asp Cys Gln Pro Cys Ala Cys Pro Leu Asn Ile Pro
 805 810 815
 Ser Asn Asn Phe Ser Pro Thr Cys His Leu Asp Arg Ser Leu Gly Leu
 820 825 830

Ile Cys Asp Gly Cys Pro Val Gly Tyr Thr Gly Pro Arg Cys Glu Arg

835 840 845
Cys Ala Glu Gly Tyr Phe Gly Gln Pro Ser Val Pro Gly Gly Ser Cys

850 855 860
Gln Pro Cys Gln Cys Asn Asp Asn Leu Asp Phe Ser Ile Pro Gly Ser
865 870 875 880

Cys Asp Ser Leu Ser Gly Ser Cys Leu Ile Cys Lys Pro Gly Thr Thr
885 890 895
Gly Arg Tyr Cys Glu Leu Cys Ala Asp Gly Tyr Phe Gly Asp Ala Val

900 905 910
Asp Ala Lys Asn Cys Gln Pro Cys Arg Cys Asn Ala Gly Gly Ser Phe
915 920 925

Ser Glu Val Cys His Ser Gln Thr Gly Gln Cys Glu Cys Arg Ala Asn
930 935 940

Val Gln Gly Gln Arg Cys Asp Lys Cys Lys Pro Asn Met Trp Arg Asp
945 950 955 960
Pro Glu Lys Arg Phe Cys Val Leu Cys Asp Cys Asp Pro Val Gly Ser

965 970 975
Val Ser Pro Gln Cys Asp Ile Thr Gly Arg Cys Val Cys Lys Ser Gly
980 985 990

Phe Val Gly Lys Gln Cys Asn Leu Gly Arg Gln Val His Gln Gln Glu
995 1000 1005

Glu Gln Pro Arg Arg Ala Gln Arg Val Leu Gly Ser Pro Gln Arg Trp
1010 1015 1020

Ala Ile Gly Ser Ser Ser Gly Cys Pro Arg Gly Ala Tyr Arg Ala Pro

1025 1030 1035 1040

Ala Pro Ala Gly Thr Phe Gly Leu Gln Ser Ala Arg Gly Cys Val Pro
1045 1050 1055

Cys Asn Cys Asn Ser Phe Gly Ser Lys Ser Phe Asp Cys Glu Glu Ser
1060 1065 1070

Gly Gln Cys Trp Cys Gln Pro Gly Val Thr Gly Lys Lys Cys Asp Arg

1075 1080 1085
Cys Ala His Gly Tyr Phe Asn Phe Gln Glu Gly Gly Cys Thr Ala Cys

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

1155 1160 1165
Cys Asn Cys His Pro Lys Phe Ser Gly Ala Lys Cys Thr Glu Cys Ser
1170 1175 1180
Arg Gly His Trp Asn Tyr Pro Arg Cys Asn Leu Cys Asp Cys Phe Leu
1185 1190 1195 1200
Pro Gly Thr Asp Ala Thr Thr Cys Asp Ser Glu Thr Lys Lys Cys Ser
1205 1210 1215
Cys Ser Asp Gln Thr Gly Gln Cys Thr Cys Lys Val Asn Val Glu Gly

1220 1225 1230
Ile His Cys Asp Arg Cys Arg Pro Gly Lys Phe Gly Leu Asp Ala Lys
1235 1240 1245
Asn Pro Leu Gly Cys Ser Ser Cys Tyr Cys Phe Gly Thr Thr Thr Gln
1250 1255 1260
Cys Ser Glu Ala Lys Gly Leu Ile Arg Thr Trp Val Thr Leu Lys Ala
1265 1270 1275 1280
Glu Gln Thr Ile Leu Pro Leu Val Asp Glu Ala Leu Gln His Thr Thr

1285 1290 1295
Thr Lys Gly Ile Val Phe Gln His Pro Glu Ile Val Ala His Met Asp
1300 1305 1310
Leu Met Arg Glu Asp Leu His Leu Glu Pro Phe Tyr Trp Lys Leu Pro
1315 1320 1325

Glu Gln Phe Glu Gly Lys Lys Leu Met Ala Tyr Gly Gly Lys Leu Lys

1330 1335 1340

Tyr Ala Ile Tyr Phe Glu Ala Arg Glu Glu Thr Gly Phe Ser Thr Tyr

1345 1350 1355 1360

Asn Pro Gln Val Ile Ile Arg Gly Gly Thr Pro Thr His Ala Arg Ile

1365 1370 1375

Ile Val Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr Arg His

1380 1385 1390

Glu Ile Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp Asp Pro

1395 1400 1405

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

1410 1415 1420

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

1425 1430 1435 1440

Gln Ser Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Gln Gly Arg

1445 1450 1455

Gly Thr Thr Met Thr Pro Pro Ala Asp Leu Ile Glu Lys Cys Asp Cys

1460 1465 1470

Pro Leu Gly Tyr Ser Gly Leu Ser Cys Glu Ala Cys Leu Pro Gly Phe

1475 1480 1485

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

1490 1495 1500

Gly Thr Cys Val Pro Cys Gln Cys Asn Gly His Ser Ser Leu Cys Asp

1505 1510 1515 1520

Pro Glu Thr Ser Ile Cys Gln Asn Cys Gln His His Thr Ala Gly Asp

1525 1530 1535

Phe Cys Glu Arg Cys Ala Leu Gly Tyr Tyr Gly Ile Val Lys Gly Leu

1540 1545 1550

Pro Asn Asp Cys Gln Gln Cys Ala Cys Pro Leu Ile Ser Ser Ser Asn

1555 1560 1565

Asn Phe Ser Pro Ser Cys Val Ala Glu Gly Leu Asp Asp Tyr Arg Cys

1570 1575 1580
 Thr Ala Cys Pro Arg Gly Tyr Glu Gly Gln Tyr Cys Glu Arg Cys Ala
 1585 1590 1595 1600
 Pro Gly Tyr Thr Gly Ser Pro Gly Asn Pro Gly Gly Ser Cys Gln Glu

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

 1665 1670 1675 1680
 Gln Met Val Met Ser Ile Asn Leu Thr Gly Pro Leu Pro Ala Pro Tyr
 1685 1690 1695
 Lys Met Leu Tyr Gly Leu Glu Asn Met Thr Gln Glu Leu Lys His Leu
 1700 1705 1710
 Leu Ser Pro Gln Arg Ala Pro Glu Arg Leu Ile Gln Leu Ala Glu Gly
 1715 1720 1725
 Asn Leu Asn Thr Leu Val Thr Glu Met Asn Glu Leu Leu Thr Arg Ala

 1730 1735 1740
 Thr Lys Val Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala Glu Arg
 1745 1750 1755 1760
 Thr Asn Thr Arg Ala Lys Ser Leu Gly Glu Phe Ile Lys Glu Leu Ala
 1765 1770 1775
 Arg Asp Ala Glu Ala Val Asn Glu Lys Ala Ile Lys Leu Asn Glu Thr
 1780 1785 1790
 Leu Gly Thr Arg Asp Glu Ala Phe Glu Arg Asn Leu Glu Gly Leu Gln

 1795 1800 1805
 Lys Glu Ile Asp Gln Met Ile Lys Glu Leu Arg Arg Lys Asn Leu Glu
 1810 1815 1820

Thr Gln Lys Glu Ile Ala Glu Asp Glu Leu Val Ala Ala Glu Ala Leu
 1825 1830 1835 1840
 Leu Lys Lys Val Lys Lys Leu Phe Gly Glu Ser Arg Gly Glu Asn Glu
 1845 1850 1855
 Glu Met Glu Lys Asp Leu Arg Glu Lys Leu Ala Asp Tyr Lys Asn Lys

 1860 1865 1870
 Val Asp Asp Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys Ile Arg
 1875 1880 1885
 Glu Ala Asn Arg Leu Phe Ala Val Asn Gln Lys Asn Met Thr Ala Leu
 1890 1895 1900
 Glu Lys Lys Lys Glu Ala Val Glu Ser Gly Lys Arg Gln Ile Glu Asn
 1905 1910 1915 1920
 Thr Leu Lys Glu Gly Asn Asp Ile Leu Asp Glu Ala Asn Arg Leu Ala

 1925 1930 1935
 Asp Glu Ile Asn Ser Ile Ile Asp Tyr Val Glu Asp Ile Gln Thr Lys
 1940 1945 1950
 Leu Pro Pro Met Ser Glu Glu Leu Asn Asp Lys Ile Asp Asp Leu Ser
 1955 1960 1965
 Gln Glu Ile Lys Asp Arg Lys Leu Ala Glu Lys Val Ser Gln Ala Glu
 1970 1975 1980
 Ser His Ala Ala Gln Leu Asn Asp Ser Ser Ala Val Leu Asp Gly Ile

 1985 1990 1995 2000
 Leu Asp Glu Ala Lys Asn Ile Ser Phe Asn Ala Thr Ala Ala Phe Lys
 2005 2010 2015
 Ala Tyr Ser Asn Ile Lys Asp Tyr Ile Asp Glu Ala Glu Lys Val Ala
 2020 2025 2030
 Lys Glu Ala Lys Asp Leu Ala His Glu Ala Thr Lys Leu Ala Thr Gly
 2035 2040 2045
 Pro Arg Gly Leu Leu Lys Glu Asp Ala Lys Gly Cys Leu Gln Lys Ser

 2050 2055 2060
 Phe Arg Ile Leu Asn Glu Ala Lys Lys Leu Ala Asn Asp Val Lys Glu

2065 2070 2075 2080
 Asn Glu Asp His Leu Asn Gly Leu Lys Thr Arg Ile Glu Asn Ala Asp
 2085 2090 2095
 Ala Arg Asn Gly Asp Leu Leu Arg Thr Leu Asn Asp Thr Leu Gly Lys
 2100 2105 2110
 Leu Ser Ala Ile Pro Asn Asp Thr Ala Ala Lys Leu Gln Ala Val Lys

 2115 2120 2125
 Asp Lys Ala Arg Gln Ala Asn Asp Thr Ala Lys Asp Val Leu Ala Gln
 2130 2135 2140
 Ile Thr Glu Leu His Gln Asn Leu Asp Gly Leu Lys Lys Asn Tyr Asn
 2145 2150 2155 2160
 Lys Leu Ala Asp Ser Val Ala Lys Thr Asn Ala Val Val Lys Asp Pro
 2165 2170 2175
 Ser Lys Asn Lys Ile Ile Ala Asp Ala Asp Ala Thr Val Lys Asn Leu

 2180 2185 2190
 Glu Gln Glu Ala Asp Arg Leu Ile Asp Lys Leu Lys Pro Ile Lys Glu
 2195 2200 2205
 Leu Glu Asp Asn Leu Lys Lys Asn Ile Ser Glu Ile Lys Glu Leu Ile
 2210 2215 2220
 Asn Gln Ala Arg Lys Gln Ala Asn Ser Ile Lys Val Ser Val Ser Ser
 2225 2230 2235 2240
 Gly Gly Asp Cys Ile Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly Ser

 2245 2250 2255
 Tyr Asn Asn Ile Val Val Asn Val Lys Thr Ala Val Ala Asp Asn Leu
 2260 2265 2270
 Leu Phe Tyr Leu Gly Ser Ala Lys Phe Ile Asp Phe Leu Ala Ile Glu
 2275 2280 2285
 Met Arg Lys Gly Lys Val Ser Phe Leu Trp Asp Val Gly Ser Gly Val
 2290 2295 2300
 Gly Arg Val Glu Tyr Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp Tyr

 2305 2310 2315 2320

Arg Ile Val Ala Ser Arg Thr Gly Arg Asn Gly Thr Ile Ser Val Arg

2325 2330 2335

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

2340 2345 2350

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

2355 2360 2365

Leu Phe Val Gly Gly Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala Val

2370 2375 2380

Arg Val Ile Thr Phe Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp Asn

2385 2390 2395 2400

Lys Pro Ile Gly Leu Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys Lys

2405 2410 2415

Gly Cys Thr Val Ser Pro Gln Val Glu Asp Ser Glu Gly Thr Ile Gln

2420 2425 2430

Phe Asp Gly Glu Gly Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp Tyr

2435 2440 2445

Pro Asn Ile Ser Thr Val Met Phe Lys Phe Arg Thr Phe Ser Ser Ser

2450 2455 2460

Ala Leu Leu Met Tyr Leu Ala Thr Arg Asp Leu Arg Asp Phe Met Ser

2465 2470 2475 2480

Val Glu Leu Thr Asp Gly His Ile Lys Val Ser Tyr Asp Leu Gly Ser

2485 2490 2495

Gly Met Ala Ser Val Val Ser Asn Gln Asn His Asn Asp Gly Lys Trp

2500 2505 2510

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

2515 2520 2525

Val Asp Ile Asp Thr Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser Ser

2530 2535 2540

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

2545 2550 2555 2560

Gly Gly Leu Pro Thr Leu Arg Asn Leu Ser Met Lys Ala Arg Pro Glu

2565	2570	2575	
Val Asn Leu Lys Lys Tyr Ser Gly Cys Leu Lys Asp Ile Glu Ile Ser			
2580	2585	2590	
Arg Thr Pro Tyr Asn Ile Leu Ser Ser Pro Asp Tyr Val Gly Val Thr			
2595	2600	2605	
Lys Gly Cys Ser Leu Glu Asn Val Tyr Thr Val Ser Phe Pro Lys Pro			
2610	2615	2620	
Gly Phe Val Glu Leu Ser Pro Val Pro Ile Asp Val Gly Thr Glu Ile			
2625	2630	2635	2640
Asn Leu Ser Phe Ser Thr Lys Asn Glu Ser Gly Ile Ile Leu Leu Gly			
2645	2650	2655	
Ser Gly Gly Thr Pro Ala Pro Pro Arg Arg Lys Arg Arg Gln Thr Gly			
2660	2665	2670	
Gln Ala Tyr Tyr Ala Ile Leu Leu Asn Arg Gly Arg Leu Glu Val His			
2675	2680	2685	
Leu Ser Thr Gly Ala Arg Thr Met Arg Lys Ile Val Ile Arg Pro Glu			
2690	2695	2700	
Pro Asn Leu Phe His Asp Gly Arg Glu His Ser Val His Val Glu Arg			
2705	2710	2715	2720
Thr Arg Gly Ile Phe Thr Val Gln Val Asp Glu Asn Arg Arg Tyr Met			
2725	2730	2735	
Gln Asn Leu Thr Val Glu Gln Pro Ile Glu Val Lys Lys Leu Phe Val			
2740	2745	2750	
Gly Gly Ala Pro Pro Glu Phe Gln Pro Ser Pro Leu Arg Asn Ile Pro			
2755	2760	2765	
Pro Phe Glu Gly Cys Ile Trp Asn Leu Val Ile Asn Ser Val Pro Met			
2770	2775	2780	
Asp Phe Ala Arg Pro Val Ser Phe Lys Asn Ala Asp Ile Gly Arg Cys			
2785	2790	2795	2800
Ala His Gln Lys Leu Arg Glu Asp Glu Asp Gly Ala Ala Pro Ala Glu			
2805	2810	2815	

Ile Val Ile Gln Pro Glu Pro Val Pro Thr Pro Ala Phe Pro Thr Pro

2820 2825 2830

Thr Pro Val Leu Thr His Gly Pro Cys Ala Ala Glu Ser Glu Pro Ala

2835 2840 2845

Leu Leu Ile Gly Ser Lys Gln Phe Gly Leu Ser Arg Asn Ser His Ile

2850 2855 2860

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

2865 2870 2875 2880

Leu Glu Val Arg Thr Glu Ala Glu Ser Gly Leu Leu Phe Tyr Met Ala

2885 2890 2895

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

2900 2905 2910

Pro Tyr Phe Ser Tyr Asp Leu Gly Ser Gly Asp Thr His Thr Met Ile

2915 2920 2925

Pro Thr Lys Ile Asn Asp Gly Gln Trp His Lys Ile Lys Ile Met Arg

2930 2935 2940

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

2945 2950 2955 2960

Ile Ser Pro Lys Lys Ala Asp Ile Leu Asp Val Val Gly Met Leu Tyr

2965 2970 2975

Val Gly Gly Leu Pro Ile Asn Tyr Thr Thr Arg Arg Ile Gly Pro Val

2980 2985 2990

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

2995 3000 3005

Pro Ala Asp Leu Glu Gln Pro Thr Ser Ser Phe His Val Gly Thr Cys

3010 3015 3020

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

3025 3030 3035 3040

Lys Ala Val Gly Gly Phe Lys Val Gly Leu Asp Leu Leu Val Glu Phe

3045 3050 3055

Glu Phe Arg Thr Thr Thr Thr Gly Val Leu Leu Gly Ile Ser Ser

3060 3065 3070
Gln Lys Met Asp Gly Met Gly Ile Glu Met Ile Asp Glu Lys Leu Met

3075 3080 3085
Phe His Val Asp Asn Gly Ala Gly Arg Phe Thr Ala Val Tyr Asp Ala

3090 3095 3100

Gly Val Pro Gly His Leu Cys Asp Gly Gln Trp His Lys Val Thr Ala

3105 3110 3115 3120

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

3125 3130 3135

Glu Ala Gln Ser Pro Asn Pro Ala Ser Thr Ser Ala Asp Thr Asn Asp

3140 3145 3150

Pro Val Phe Val Gly Gly Phe Pro Asp Asp Leu Lys Gln Phe Gly Leu

3155 3160 3165

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

3170 3175 3180

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

3185 3190 3195 3200

Leu Arg Gly Val Gln Pro Val Ser Cys Pro Ala Asn

3205 3210

<210> 120

<211> 2587

<212> PRT

<213> Artificial Sequence

<220>

<223> laminin subunit α2 isoform?X5

<400> 120

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

1 5 10 15

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

20 25 30

Lys Glu Glu Ser Phe Thr Ile His Gly Thr His Phe Pro Val Arg Arg

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

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

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

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

Ala Lys Asn Cys Gln Pro Cys Arg Cys Asn Ala Gly Gly Ser Phe Ser
290 295 300

Glu Val Cys His Ser Gln Thr Gly Gln Cys Glu Cys Arg Ala Asn Val
305 310 315 320

Gln Gly Gln Arg Cys Asp Lys Cys Lys Pro Asn Met Trp Arg Asp Pro
325 330 335

Glu Lys Arg Phe Cys Val Leu Cys Asp Cys Asp Pro Val Gly Ser Val
340 345 350

Ser Pro Gln Cys Asp Ile Thr Gly Arg Cys Val Cys Lys Ser Gly Phe
355 360 365

Val Gly Lys Gln Cys Asn Leu Gly Arg Gln Val His Gln Gln Glu Glu
370 375 380

Gln Pro Arg Arg Ala Gln Arg Val Leu Gly Ser Pro Gln Arg Trp Ala
385 390 395 400

Ile Gly Ser Ser Ser Gly Cys Pro Arg Gly Ala Tyr Arg Ala Pro Ala
405 410 415

Pro Ala Gly Thr Phe Gly Leu Gln Ser Ala Arg Gly Cys Val Pro Cys
420 425 430

Asn Cys Asn Ser Phe Gly Ser Lys Ser Phe Asp Cys Glu Glu Ser Gly
435 440 445

Gln Cys Trp Cys Gln Pro Gly Val Thr Gly Lys Lys Cys Asp Arg Cys
450 455 460

Ala His Gly Tyr Phe Asn Phe Gln Glu Gly Gly Cys Thr Ala Cys Glu
465 470 475 480

Cys Ser His Leu Gly Asn Asn Cys Asp Pro Lys Thr Gly Arg Cys Ile
485 490 495

Cys Pro Pro Asn Thr Ile Gly Glu Lys Cys Ser Lys Cys Ala Pro Asn
500 505 510

Thr Trp Gly His Ser Ile Thr Thr Gly Cys Lys Ala Cys Asn Cys Ser
515 520 525

Thr Val Gly Ser Leu Asp Phe Gln Cys Asn Val Asn Thr Gly Gln Cys

530 535 540
 Asn Cys His Pro Lys Phe Ser Gly Ala Lys Cys Thr Glu Cys Ser Arg
 545 550 555 560

 Gly His Trp Asn Tyr Pro Arg Cys Asn Leu Cys Asp Cys Phe Leu Pro
 565 570 575
 Gly Thr Asp Ala Thr Thr Cys Asp Ser Glu Thr Lys Lys Cys Ser Cys
 580 585 590
 Ser Asp Gln Thr Gly Gln Cys Thr Cys Lys Val Asn Val Glu Gly Ile
 595 600 605
 His Cys Asp Arg Cys Arg Pro Gly Lys Phe Gly Leu Asp Ala Lys Asn
 610 615 620

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

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

 Val Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr Arg His Glu
 755 760 765
 Ile Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp Asp Pro Arg
 770 775 780

Val His Arg Thr Val Thr Arg Glu Asp Phe Leu Asp Ile Leu Tyr Asp
 785 790 795 800
 Ile His Tyr Ile Leu Ile Lys Ala Thr Tyr Gly Asn Phe Met Arg Gln
 805 810 815

Ser Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Gln Gly Arg Gly
 820 825 830
 Thr Thr Met Thr Pro Pro Ala Asp Leu Ile Glu Lys Cys Asp Cys Pro
 835 840 845
 Leu Gly Tyr Ser Gly Leu Ser Cys Glu Ala Cys Leu Pro Gly Phe Tyr
 850 855 860
 Arg Leu Arg Ser Gln Pro Gly Gly Arg Thr Pro Gly Pro Thr Leu Gly
 865 870 875 880

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

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

Gly Phe Cys Thr Cys Arg Pro Gly Ala Thr Gly Arg Lys Cys Asp Gly
 1010 1015 1020
 Cys Lys His Trp His Ala Arg Glu Gly Trp Glu Cys Val Phe Cys Gly

1025	1030	1035	1040
Asp Glu Cys Thr Gly Leu Leu Leu Gly Asp Leu Ala Arg Leu Glu Gln			
	1045	1050	1055
Met Val Met Ser Ile Asn Leu Thr Gly Pro Leu Pro Ala Pro Tyr Lys			
	1060	1065	1070
Met Leu Tyr Gly Leu Glu Asn Met Thr Gln Glu Leu Lys His Leu Leu			
	1075	1080	1085
Ser Pro Gln Arg Ala Pro Glu Arg Leu Ile Gln Leu Ala Glu Gly Asn			
	1090	1095	1100
Leu Asn Thr Leu Val Thr Glu Met Asn Glu Leu Leu Thr Arg Ala Thr			
1105	1110	1115	1120
Lys Val Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala Glu Arg Thr			
	1125	1130	1135
Asn Thr Arg Ala Lys Ser Leu Gly Glu Phe Ile Lys Glu Leu Ala Arg			
	1140	1145	1150
Asp Ala Glu Ala Val Asn Glu Lys Ala Ile Lys Leu Asn Glu Thr Leu			
	1155	1160	1165
Gly Thr Arg Asp Glu Ala Phe Glu Arg Asn Leu Glu Gly Leu Gln Lys			
	1170	1175	1180
Glu Ile Asp Gln Met Ile Lys Glu Leu Arg Arg Lys Asn Leu Glu Thr			
1185	1190	1195	1200
Gln Lys Glu Ile Ala Glu Asp Glu Leu Val Ala Ala Glu Ala Leu Leu			
	1205	1210	1215
Lys Lys Val Lys Lys Leu Phe Gly Glu Ser Arg Gly Glu Asn Glu Glu			
	1220	1225	1230
Met Glu Lys Asp Leu Arg Glu Lys Leu Ala Asp Tyr Lys Asn Lys Val			
	1235	1240	1245
Asp Asp Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys Ile Arg Glu			
	1250	1255	1260
Ala Asn Arg Leu Phe Ala Val Asn Gln Lys Asn Met Thr Ala Leu Glu			
1265	1270	1275	1280

Lys Lys Lys Glu Ala Val Glu Ser Gly Lys Arg Gln Ile Glu Asn Thr

1285 1290 1295

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

1300 1305 1310

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

1315 1320 1325

Pro Pro Met Ser Glu Glu Leu Asn Asp Lys Ile Asp Asp Leu Ser Gln

1330 1335 1340

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

1345 1350 1355 1360

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

1365 1370 1375

Asp Glu Ala Lys Asn Ile Ser Phe Asn Ala Thr Ala Ala Phe Lys Ala

1380 1385 1390

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

1395 1400 1405

Glu Ala Lys Asp Leu Ala His Glu Ala Thr Lys Leu Ala Thr Gly Pro

1410 1415 1420

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

1425 1430 1435 1440

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

1445 1450 1455

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

1460 1465 1470

Arg Asn Gly Asp Leu Leu Arg Thr Leu Asn Asp Thr Leu Gly Lys Leu

1475 1480 1485

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

1490 1495 1500

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

1505 1510 1515 1520

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

1525	1530	1535
Leu Ala Asp Ser Val Ala Lys Thr Asn Ala Val Val Lys Asp Pro Ser		
1540	1545	1550
Lys Asn Lys Ile Ile Ala Asp Ala Asp Ala Thr Val Lys Asn Leu Glu		
1555	1560	1565
Gln Glu Ala Asp Arg Leu Ile Asp Lys Leu Lys Pro Ile Lys Glu Leu		
1570	1575	1580
Glu Asp Asn Leu Lys Lys Asn Ile Ser Glu Ile Lys Glu Leu Ile Asn		
1585	1590	1595
Gln Ala Arg Lys Gln Ala Asn Ser Ile Lys Val Ser Val Ser Ser Gly		
1605	1610	1615
Gly Asp Cys Ile Arg Thr Tyr Lys Pro Glu Ile Lys Lys Gly Ser Tyr		
1620	1625	1630
Asn Asn Ile Val Val Asn Val Lys Thr Ala Val Ala Asp Asn Leu Leu		
1635	1640	1645
Phe Tyr Leu Gly Ser Ala Lys Phe Ile Asp Phe Leu Ala Ile Glu Met		
1650	1655	1660
Arg Lys Gly Lys Val Ser Phe Leu Trp Asp Val Gly Ser Gly Val Gly		
1665	1670	1675
Arg Val Glu Tyr Pro Asp Leu Thr Ile Asp Asp Ser Tyr Trp Tyr Arg		
1685	1690	1695
Ile Val Ala Ser Arg Thr Gly Arg Asn Gly Thr Ile Ser Val Arg Ala		
1700	1705	1710
Leu Asp Gly Pro Lys Ala Ser Ile Val Pro Ser Thr His His Ser Thr		
1715	1720	1725
Ser Pro Pro Gly Tyr Thr Ile Leu Asp Val Asp Ala Asn Ala Met Leu		
1730	1735	1740
Phe Val Gly Gly Leu Thr Gly Lys Leu Lys Lys Ala Asp Ala Val Arg		
1745	1750	1755
Val Ile Thr Phe Thr Gly Cys Met Gly Glu Thr Tyr Phe Asp Asn Lys		
1765	1770	1775

Pro Ile Gly Leu Trp Asn Phe Arg Glu Lys Glu Gly Asp Cys Lys Gly

1780 1785 1790

Cys Thr Val Ser Pro Gln Val Glu Asp Ser Glu Gly Thr Ile Gln Phe

1795 1800 1805

Asp Gly Glu Gly Tyr Ala Leu Val Ser Arg Pro Ile Arg Trp Tyr Pro

1810 1815 1820

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

1825 1830 1835 1840

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

1845 1850 1855

Glu Leu Thr Asp Gly His Ile Lys Val Ser Tyr Asp Leu Gly Ser Gly

1860 1865 1870

Met Ala Ser Val Val Ser Asn Gln Asn His Asn Asp Gly Lys Trp Lys

1875 1880 1885

Ser Phe Thr Leu Ser Arg Ile Gln Lys Gln Ala Asn Ile Ser Ile Val

1890 1895 1900

Asp Ile Asp Thr Asn Gln Glu Glu Asn Ile Ala Thr Ser Ser Ser Gly

1905 1910 1915 1920

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

1925 1930 1935

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

1940 1945 1950

Asn Leu Lys Lys Tyr Ser Gly Cys Leu Lys Asp Ile Glu Ile Ser Arg

1955 1960 1965

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

1970 1975 1980

Gly Cys Ser Leu Glu Asn Val Tyr Thr Val Ser Phe Pro Lys Pro Gly

1985 1990 1995 2000

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

2005 2010 2015

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

2020	2025	2030	
Gly Gly Thr Pro Ala Pro Pro Arg Arg Lys Arg Arg Gln Thr Gly Gln			
2035	2040	2045	
Ala Tyr Tyr Ala Ile Leu Leu Asn Arg Gly Arg Leu Glu Val His Leu			
2050	2055	2060	
Ser Thr Gly Ala Arg Thr Met Arg Lys Ile Val Ile Arg Pro Glu Pro			
2065	2070	2075	2080
Asn Leu Phe His Asp Gly Arg Glu His Ser Val His Val Glu Arg Thr			
2085	2090	2095	
Arg Gly Ile Phe Thr Val Gln Val Asp Glu Asn Arg Arg Tyr Met Gln			
2100	2105	2110	
Asn Leu Thr Val Glu Gln Pro Ile Glu Val Lys Lys Leu Phe Val Gly			
2115	2120	2125	
Gly Ala Pro Pro Glu Phe Gln Pro Ser Pro Leu Arg Asn Ile Pro Pro			
2130	2135	2140	
Phe Glu Gly Cys Ile Trp Asn Leu Val Ile Asn Ser Val Pro Met Asp			
2145	2150	2155	2160
Phe Ala Arg Pro Val Ser Phe Lys Asn Ala Asp Ile Gly Arg Cys Ala			
2165	2170	2175	
His Gln Lys Leu Arg Glu Asp Glu Asp Gly Ala Ala Pro Ala Glu Ile			
2180	2185	2190	
Val Ile Gln Pro Glu Pro Val Pro Thr Pro Ala Phe Pro Thr Pro Thr			
2195	2200	2205	
Pro Val Leu Thr His Gly Pro Cys Ala Ala Glu Ser Glu Pro Ala Leu			
2210	2215	2220	
Leu Ile Gly Ser Lys Gln Phe Gly Leu Ser Arg Asn Ser His Ile Ala			
2225	2230	2235	2240
Ile Ala Phe Asp Asp Thr Lys Val Lys Asn Arg Leu Thr Ile Glu Leu			
2245	2250	2255	
Glu Val Arg Thr Glu Ala Glu Ser Gly Leu Leu Phe Tyr Met Ala Arg			
2260	2265	2270	

Ile Asn His Ala Asp Phe Ala Thr Val Gln Leu Arg Asn Gly Leu Pro
2275 2280 2285

Tyr Phe Ser Tyr Asp Leu Gly Ser Gly Asp Thr His Thr Met Ile Pro
2290 2295 2300

Thr Lys Ile Asn Asp Gly Gln Trp His Lys Ile Lys Ile Met Arg Ser
2305 2310 2315 2320

Lys Gln Glu Gly Ile Leu Tyr Val Asp Gly Ala Ser Asn Arg Thr Ile
2325 2330 2335

Ser Pro Lys Lys Ala Asp Ile Leu Asp Val Val Gly Met Leu Tyr Val
2340 2345 2350

Gly Gly Leu Pro Ile Asn Tyr Thr Thr Arg Arg Ile Gly Pro Val Thr
2355 2360 2365

Tyr Ser Ile Asp Gly Cys Val Arg Asn Leu His Met Ala Glu Ala Pro
2370 2375 2380

Ala Asp Leu Glu Gln Pro Thr Ser Ser Phe His Val Gly Thr Cys Phe
2385 2390 2395 2400

Ala Asn Ala Gln Arg Gly Thr Tyr Phe Asp Gly Thr Gly Phe Ala Lys
2405 2410 2415

Ala Val Gly Gly Phe Lys Val Gly Leu Asp Leu Leu Val Glu Phe Glu
2420 2425 2430

Phe Arg Thr Thr Thr Thr Thr Gly Val Leu Leu Gly Ile Ser Ser Gln
2435 2440 2445

Lys Met Asp Gly Met Gly Ile Glu Met Ile Asp Glu Lys Leu Met Phe
2450 2455 2460

His Val Asp Asn Gly Ala Gly Arg Phe Thr Ala Val Tyr Asp Ala Gly
2465 2470 2475 2480

Val Pro Gly His Leu Cys Asp Gly Gln Trp His Lys Val Thr Ala Asn
2485 2490 2495

Lys Ile Lys His Arg Ile Glu Leu Thr Val Asp Gly Asn Gln Val Glu
2500 2505 2510

Ala Gln Ser Pro Asn Pro Ala Ser Thr Ser Ala Asp Thr Asn Asp Pro

2515 2520 2525
Val Phe Val Gly Gly Phe Pro Asp Asp Leu Lys Gln Phe Gly Leu Thr
2530 2535 2540

Thr Ser Ile Pro Phe Arg Gly Cys Ile Arg Ser Leu Lys Leu Thr Lys
2545 2550 2555 2560

Gly Thr Gly Lys Pro Leu Glu Val Asn Phe Ala Lys Ala Leu Glu Leu
2565 2570 2575

Arg Gly Val Gln Pro Val Ser Cys Pro Ala Asn
2580 2585

<210> 121

<211> 1916

<212> PRT

<213> Artificial Sequence

<220>

<223> laminin subunit α2 isoform?X6

<400> 121

Met Pro Gly Ala Ala Gly Val Leu Leu Leu Leu Leu Ser Gly Gly

1 5 10 15
Leu Gly Gly Val Gln Ala Gln Arg Pro Gln Gln Gln Arg Gln Ser Gln

20 25 30
Ala His Gln Gln Arg Gly Leu Phe Pro Ala Val Leu Asn Leu Ala Ser

35 40 45
Asn Ala Leu Ile Thr Thr Asn Ala Thr Cys Gly Glu Lys Gly Pro Glu

50 55 60
Met Tyr Cys Lys Leu Val Glu His Val Pro Gly Gln Pro Val Arg Asn

65 70 75 80
Pro Gln Cys Arg Ile Cys Asn Gln Asn Ser Ser Asn Pro Asn Gln Arg

85 90 95
His Pro Ile Thr Asn Ala Ile Asp Gly Lys Asn Thr Trp Trp Gln Ser

100 105 110
Pro Ser Ile Lys Asn Gly Ile Glu Tyr His Tyr Val Thr Ile Thr Leu

115 120 125

Asp Leu Gln Gln Val Phe Gln Ile Ala Tyr Val Ile Val Lys Ala Ala

130

135

140

Asn Ser Pro Arg Pro Gly Asn Trp Ile Leu Glu Arg Ser Leu Asp Asp

145

150

155

160

Val Glu Tyr Lys Pro Trp Gln Tyr His Ala Val Thr Asp Thr Glu Cys

165

170

175

Leu Thr Leu Tyr Asn Ile Tyr Pro Arg Thr Gly Pro Pro Ser Tyr Ala

180

185

190

Lys Asp Asp Glu Val Ile Cys Thr Ser Phe Tyr Ser Lys Ile His Pro

195

200

205

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

210

215

220

Ala Asp Asp Pro Ser Pro Glu Leu Leu Glu Phe Thr Ser Ala Arg Tyr

225

230

235

240

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

245

250

255

Met Phe Ala His Lys Asp Pro Arg Glu Ile Asp Pro Ile Val Thr Arg

260

265

270

Arg Tyr Tyr Tyr Ser Val Lys Asp Ile Ser Val Gly Gly Met Cys Ile

275

280

285

Cys Tyr Gly His Ala Arg Ala Cys Pro Leu Asp Pro Ala Thr Asn Lys

290

295

300

Ser Arg Cys Glu Cys Glu His Asn Thr Cys Gly Asp Ser Cys Asp Gln

305

310

315

320

Cys Cys Pro Gly Phe His Gln Lys Pro Trp Arg Ala Gly Thr Phe Leu

325

330

335

Thr Lys Thr Glu Cys Glu Ala Cys Asn Cys His Gly Lys Ala Glu Glu

340

345

350

Cys Tyr Tyr Asp Glu Asn Val Ala Arg Arg Asn Leu Ser Leu Asn Ile

355

360

365

Arg Gly Lys Tyr Ile Gly Gly Gly Val Cys Ile Asn Cys Thr Gln Asn

370 375 380
 Thr Ala Gly Ile Asn Cys Glu Thr Cys Thr Asp Gly Phe Phe Arg Pro

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

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

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

 580 585 590
 Asn Lys Leu Pro Ala Val Gly Gly Gln Leu Thr Phe Thr Ile Ser Tyr
 595 600 605
 Asp Leu Glu Glu Glu Glu Asp Thr Glu Arg Val Leu Gln Leu Met
 610 615 620

Ile Ile Leu Glu Gly Asn Asp Leu Ser Ile Ser Thr Ala Gln Asp Glu
625 630 635 640
Val Tyr Leu His Pro Ser Glu Glu His Thr Asn Val Leu Leu Leu Lys
645 650 655
Glu Glu Ser Phe Thr Ile His Gly Thr His Phe Pro Val Arg Arg Lys
660 665 670
Glu Phe Met Thr Val Leu Ala Asn Leu Lys Arg Val Leu Leu Gln Ile
675 680 685
Thr Tyr Ser Phe Gly Met Asp Ala Ile Phe Arg Leu Ser Ser Val Asn
690 695 700
Leu Glu Ser Ala Val Ser Tyr Pro Thr Asp Gly Ser Ile Ala Ala Ala
705 710 715 720
Val Glu Val Cys Gln Cys Pro Pro Gly Tyr Thr Gly Ser Ser Cys Glu
725 730 735
Ser Cys Trp Pro Arg His Arg Arg Val Asn Gly Thr Ile Phe Gly Gly
740 745 750
Ile Cys Glu Pro Cys Gln Cys Phe Gly His Ala Glu Ser Cys Asp Asp
755 760 765
Val Thr Gly Glu Cys Leu Asn Cys Lys Asp His Thr Gly Gly Pro Tyr
770 775 780
Cys Asp Lys Cys Leu Pro Gly Phe Tyr Gly Glu Pro Thr Lys Gly Thr
785 790 795 800
Ser Glu Asp Cys Gln Pro Cys Ala Cys Pro Leu Asn Ile Pro Ser Asn
805 810 815
Asn Phe Ser Pro Thr Cys His Leu Asp Arg Ser Leu Gly Leu Ile Cys
820 825 830
Asp Gly Cys Pro Val Gly Tyr Thr Gly Pro Arg Cys Glu Arg Cys Ala
835 840 845
Glu Gly Tyr Phe Gly Gln Pro Ser Val Pro Gly Gly Ser Cys Gln Pro
850 855 860
Cys Gln Cys Asn Asp Asn Leu Asp Phe Ser Ile Pro Gly Ser Cys Asp

865 870 875 880
 Ser Leu Ser Gly Ser Cys Leu Ile Cys Lys Pro Gly Thr Thr Gly Arg
 885 890 895
 Tyr Cys Glu Leu Cys Ala Asp Gly Tyr Phe Gly Asp Ala Val Asp Ala

 900 905 910
 Lys Asn Cys Gln Pro Cys Arg Cys Asn Ala Gly Gly Ser Phe Ser Glu
 915 920 925
 Val Cys His Ser Gln Thr Gly Gln Cys Glu Cys Arg Ala Asn Val Gln
 930 935 940
 Gly Gln Arg Cys Asp Lys Cys Lys Pro Asn Met Trp Arg Asp Pro Glu
 945 950 955 960
 Lys Arg Phe Cys Val Leu Cys Asp Cys Asp Pro Val Gly Ser Val Ser

 965 970 975
 Pro Gln Cys Asp Ile Thr Gly Arg Cys Val Cys Lys Ser Gly Phe Val
 980 985 990
 Gly Lys Gln Cys Asn Leu Gly Arg Gln Val His Gln Gln Glu Glu Gln
 995 1000 1005
 Pro Arg Arg Ala Gln Arg Val Leu Gly Ser Pro Gln Arg Trp Ala Ile
 1010 1015 1020
 Gly Ser Ser Ser Gly Cys Pro Arg Gly Ala Tyr Arg Ala Pro Ala Pro

1025 1030 1035 1040
 Ala Gly Thr Phe Gly Leu Gln Ser Ala Arg Gly Cys Val Pro Cys Asn
 1045 1050 1055
 Cys Asn Ser Phe Gly Ser Lys Ser Phe Asp Cys Glu Glu Ser Gly Gln
 1060 1065 1070
 Cys Trp Cys Gln Pro Gly Val Thr Gly Lys Lys Cys Asp Arg Cys Ala
 1075 1080 1085
 His Gly Tyr Phe Asn Phe Gln Glu Gly Gly Cys Thr Ala Cys Glu Cys

 1090 1095 1100
 Ser His Leu Gly Asn Asn Cys Asp Pro Lys Thr Gly Arg Cys Ile Cys
 1105 1110 1115 1120

Pro Pro Asn Thr Ile Gly Glu Lys Cys Ser Lys Cys Ala Pro Asn Thr

1125 1130 1135

Trp Gly His Ser Ile Thr Thr Gly Cys Lys Ala Cys Asn Cys Ser Thr

1140 1145 1150

Val Gly Ser Leu Asp Phe Gln Cys Asn Val Asn Thr Gly Gln Cys Asn

1155 1160 1165

Cys His Pro Lys Phe Ser Gly Ala Lys Cys Thr Glu Cys Ser Arg Gly

1170 1175 1180

His Trp Asn Tyr Pro Arg Cys Asn Leu Cys Asp Cys Phe Leu Pro Gly

1185 1190 1195 1200

Thr Asp Ala Thr Thr Cys Asp Ser Glu Thr Lys Lys Cys Ser Cys Ser

1205 1210 1215

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

1220 1225 1230

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

1235 1240 1245

Leu Gly Cys Ser Ser Cys Tyr Cys Phe Gly Thr Thr Thr Gln Cys Ser

1250 1255 1260

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

1265 1270 1275 1280

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

1285 1290 1295

Gly Ile Val Phe Gln His Pro Glu Ile Val Ala His Met Asp Leu Met

1300 1305 1310

Arg Glu Asp Leu His Leu Glu Pro Phe Tyr Trp Lys Leu Pro Glu Gln

1315 1320 1325

Phe Glu Gly Lys Lys Leu Met Ala Tyr Gly Gly Lys Leu Lys Tyr Ala

1330 1335 1340

Ile Tyr Phe Glu Ala Arg Glu Glu Thr Gly Phe Ser Thr Tyr Asn Pro

1345 1350 1355 1360

Gln Val Ile Ile Arg Gly Gly Thr Pro Thr His Ala Arg Ile Ile Val

1365	1370	1375	
Arg His Met Ala Ala Pro Leu Ile Gly Gln Leu Thr Arg His Glu Ile			
1380	1385	1390	
Glu Met Thr Glu Lys Glu Trp Lys Tyr Tyr Gly Asp Asp Pro Arg Val			
1395	1400	1405	
His Arg Thr Val Thr Arg Glu Asp Phe Leu Asp Ile Leu Tyr Asp Ile			
1410	1415	1420	
His Tyr Ile Leu Ile Lys Ala Thr Tyr Gly Asn Phe Met Arg Gln Ser			
1425	1430	1435	1440
Arg Ile Ser Glu Ile Ser Met Glu Val Ala Glu Gln Gly Arg Gly Thr			
1445	1450	1455	
Thr Met Thr Pro Pro Ala Asp Leu Ile Glu Lys Cys Asp Cys Pro Leu			
1460	1465	1470	
Gly Tyr Ser Gly Leu Ser Cys Glu Ala Cys Leu Pro Gly Phe Tyr Arg			
1475	1480	1485	
Leu Arg Ser Gln Pro Gly Gly Arg Thr Pro Gly Pro Thr Leu Gly Thr			
1490	1495	1500	
Cys Val Pro Cys Gln Cys Asn Gly His Ser Ser Leu Cys Asp Pro Glu			
1505	1510	1515	1520
Thr Ser Ile Cys Gln Asn Cys Gln His His Thr Ala Gly Asp Phe Cys			
1525	1530	1535	
Glu Arg Cys Ala Leu Gly Tyr Tyr Gly Ile Val Lys Gly Leu Pro Asn			
1540	1545	1550	
Asp Cys Gln Gln Cys Ala Cys Pro Leu Ile Ser Ser Ser Asn Asn Phe			
1555	1560	1565	
Ser Pro Ser Cys Val Ala Glu Gly Leu Asp Asp Tyr Arg Cys Thr Ala			
1570	1575	1580	
Cys Pro Arg Gly Tyr Glu Gly Gln Tyr Cys Glu Arg Cys Ala Pro Gly			
1585	1590	1595	1600
Tyr Thr Gly Ser Pro Gly Asn Pro Gly Gly Ser Cys Gln Glu Cys Glu			
1605	1610	1615	

Cys Asp Pro Tyr Gly Ser Leu Pro Val Pro Cys Asp Pro Val Thr Gly
 1620 1625 1630
 Phe Cys Thr Cys Arg Pro Gly Ala Thr Gly Arg Lys Cys Asp Gly Cys
 1635 1640 1645
 Lys His Trp His Ala Arg Glu Gly Trp Glu Cys Val Phe Cys Gly Asp
 1650 1655 1660
 Glu Cys Thr Gly Leu Leu Leu Gly Asp Leu Ala Arg Leu Glu Gln Met

 1665 1670 1675 1680
 Val Met Ser Ile Asn Leu Thr Gly Pro Leu Pro Ala Pro Tyr Lys Met
 1685 1690 1695
 Leu Tyr Gly Leu Glu Asn Met Thr Gln Glu Leu Lys His Leu Leu Ser
 1700 1705 1710
 Pro Gln Arg Ala Pro Glu Arg Leu Ile Gln Leu Ala Glu Gly Asn Leu
 1715 1720 1725
 Asn Thr Leu Val Thr Glu Met Asn Glu Leu Leu Thr Arg Ala Thr Lys

 1730 1735 1740
 Val Thr Ala Asp Gly Glu Gln Thr Gly Gln Asp Ala Glu Arg Thr Asn
 1745 1750 1755 1760
 Thr Arg Ala Lys Ser Leu Gly Glu Phe Ile Lys Glu Leu Ala Arg Asp
 1765 1770 1775
 Ala Glu Ala Val Asn Glu Lys Ala Ile Lys Leu Asn Glu Thr Leu Gly
 1780 1785 1790
 Thr Arg Asp Glu Ala Phe Glu Arg Asn Leu Glu Gly Leu Gln Lys Glu

 1795 1800 1805
 Ile Asp Gln Met Ile Lys Glu Leu Arg Arg Lys Asn Leu Glu Thr Gln
 1810 1815 1820
 Lys Glu Ile Ala Glu Asp Glu Leu Val Ala Ala Glu Ala Leu Leu Lys
 1825 1830 1835 1840
 Lys Val Lys Lys Leu Phe Gly Glu Ser Arg Gly Glu Asn Glu Glu Met
 1845 1850 1855
 Glu Lys Asp Leu Arg Glu Lys Leu Ala Asp Tyr Lys Asn Lys Val Asp

1860 1865 1870
 Asp Ala Trp Asp Leu Leu Arg Glu Ala Thr Asp Lys Ile Arg Glu Ala
 1875 1880 1885
 Asn Arg Leu Phe Ala Val Asn Gln Lys Asn Met Thr Ala Leu Glu Gln
 1890 1895 1900
 Leu Pro Ala Lys Val Ile Lys Thr Asn Gln Ser Ile
 1905 1910 1915
 <210> 122
 <211> 594
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Modified Hyperactive Piggybac triple mutant

D450N/R372A/K375A/R376A

<400> 122
 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> 123

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R245A/N347S/R372A/D450N/T560A/S564P/S573A/S592G

<400> 123

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> 124

<211> 594

<212>

> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

R277A/G325A/N347A/K375A/D450N/T560A/S564P/S573A/S592G/F594L

<400> 124

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> 125

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

V34M/R275A/G325A/N347S/S351A/R372A/K375A/D450N/T560A/S564P

<400> 125

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> 126

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

G325A/N347S/K375A/D450N/S573A/M589V/S592G

<400> 126

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> 127

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac S230N/R277A/N347S/K375A/D450N

<400> 127

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> 128

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac T43I/R372A/K375A/A411T/D450N

<400> 128

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> 129

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

G325A/N347S/S351A/K375A/D450N/S573A/M589V/S592G

<400> 129

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 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
385 390 395 400
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 Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
 465 Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
 485 Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Gly Leu Thr Ser

500 Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
 515 Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
 530 Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
 545 Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Ser Ala Ala Cys Lys Lys

565 Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Val Cys Gln Gly
 580 Cys Phe

<210> 130

<211> 594

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified hyperactive PiggyBac

Y177H/R275A/G325A/K375A/D450N/T560A/S564P/S592G

<400> 130

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> 131

<211> 390

<212> DNA

<213> Artificial Sequence

<220>

<223> MCP N55K

<400> 131

atggcttcaa actttactca gttcgtgctc gtggacaatg gtgggacagg ggatgtgaca 60
 gtggctcctt ctaatttcgc taatgggggtg gcagagtgga tcagctccaa ctcacggagc 120
 caggcctaca aggtgacatg cagcgtcagg cagtctagtg cccagaagag aaagtatacc 180
 atcaaggtgg aggtccccaa agtggctacc cagacagtgg gcggagtcga actgcctgtc 240
 gccgcttgga ggtcctacct gaacatggag ctactatcc caattttcgc taccaattct 300
 gactgtgaac tcacgtgaa ggcaatgcag gggtcctca aagacggtaa tcctatccct 360

 tccgccatcg ccgctaactc aggtatctac 390

<210> 132

<211> 130

<212> PRT

<213> Artificial Sequence

<220>

<223> MCP N55K

<400> 132

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

Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala Asn Ser Gly
115 120 125

Ile Tyr

130

<210> 133

<211> 162

<212> DNA

<213> Artificial Sequence

<220>

<223> gRNA-MS2(tetraloop)-AAVS1-3 spacer

<400> 133

ggggccacta gggacaggat gtttagagc taggccaaca tgaggatcac ccatgtctgc 60

agggcctagc aagttaaat aaggctagtc cgttatcaac ttggccaaca tgaggatcac 120

ccatgtctgc agggccaagt ggcaccgagt cgtgtctttt tt 162

<210> 134

<211> 162

<212> RNA

<213> Artificial Sequence

<220>

<223> gRNA-MS2(tetraloop)-AAVS1-3 spacer

<400> 134

ggggccacua gggacaggau guuuuagagc uaggccaaca ugaggaucaac ccaugucugc	60
agggccuagc aaguuaaaau aaggcuaguc cguuaucaac uaggccaaca ugaggaucaac	120
ccaugucugc agggccaagu ggcaccgagu cggugcuuuu uu	162