

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0005179

(43) 공개일자 2021년01월13일

(51) 국제특허분류(Int. Cl.)

C12N 15/90 (2006.01) A61K 48/00 (2006.01)

A61P 7/00 (2006.01) C07K 14/805 (2006.01)

C12N 15/113 (2010.01) C12N 9/22 (2006.01)

(52) CPC특허분류

C12N 15/907 (2013.01)

A61K 48/00 (2013.01)

(21) 출원번호 10-2020-7034099

(22) 출원일자(국제) 2019년04월24일

심사청구일자 없음

(85) 번역문제출일자 2020년11월26일

(86) 국제출원번호 PCT/US2019/028861

(87) 국제공개번호 WO 2019/209914

국제공개일자 2019년10월31일

(30) 우선권주장

62/663,553 2018년04월27일 미국(US)

62/820,521 2019년03월19일 미국(US)

(71) 출원인

시애틀 칠드런즈 호스피탈 디/비/에이 시애틀 칠드런즈 리서치 인스티튜트

미국 98105 워싱턴 시애틀 4800 샌드 포인트 웨이 노스이스트

(72) 발명자

롤링스, 데이비드 제이.

미국 98101 워싱턴주 시애틀 나인스 애브뉴 1900 시애틀 칠드런즈 호스피탈 (디비에이 시애틀 칠드런즈 리서치 인스티튜트) 내

파다비, 소우마

미국 98101 워싱턴주 시애틀 나인스 애브뉴 1900 시애틀 칠드런즈 호스피탈 (디비에이 시애틀 칠드런즈 리서치 인스티튜트) 내

(뒷면에 계속)

(74) 대리인

양영준, 김영

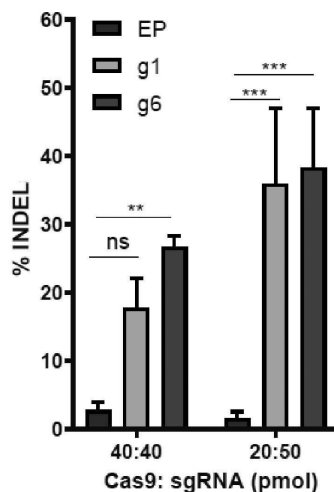
전체 청구항 수 : 총 101 항

(54) 발명의 명칭 헤모글로빈-관련 돌연변이를 편집하기 위한 상동성-지정 복구 주형 설계 및 전달

(57) 요약

본원에 제공된 방법 및 조성물의 일부 실시양태는 헤모글로빈 유전자좌, 예컨대 겸상 적혈구 돌연변이를 포함하는 헤모글로빈-관련 돌연변이의 변형에 관한 것이다. 일부 실시양태는 겸상 적혈구 돌연변이의 부위에 포스포디에스테르 DNA 가닥 파단의 도입을 통한 겸상 적혈구 돌연변이의 변형에 관한 것이다.

대표도 - 도1b



(52) CPC특허분류

A61P 7/00 (2018.01)
C07K 14/805 (2013.01)
C12N 15/113 (2013.01)
C12N 9/22 (2013.01)
C12N 2310/20 (2017.05)
C12N 2320/32 (2013.01)
C12N 2320/34 (2013.01)
C12N 2330/51 (2013.01)
C12N 2750/14143 (2013.01)

(72) 발명자

샤렌베르크, 앤드류 엠.

미국 98101 워싱턴주 시애틀 나인스 애브뉴 1900
시애틀 칠드런즈 호스피탈 (디비에이 시애틀 칠드
런즈 리서치 인스티튜트) 내

자코비, 카일

미국 98101 워싱턴주 시애틀 나인스 애브뉴 1900
시애틀 칠드런즈 호스피탈 (디비에이 시애틀 칠드
런즈 리서치 인스티튜트) 내

명세서

청구범위

청구항 1

(i) 가이드 RNA (gRNA)를 코딩하는 폴리뉴클레오타이드를 세포에 도입하는 단계, 및
(ii) 주형 폴리뉴클레오타이드를 세포에 도입하는 단계
를 포함하는, 세포에서 HBB 유전자를 편집하는 방법.

청구항 2

제1항에 있어서, gRNA가 서열식별번호(SEQ ID NO):01-06 중 어느 하나의 뉴클레오타이드 서열과 적어도 95% 동일성을 갖는 핵산을 포함하는 것인 방법.

청구항 3

제1항 또는 제2항에 있어서, gRNA가 서열식별번호:07-12 중 어느 하나의 뉴클레오타이드 서열과 적어도 95% 동일성을 갖는 핵산을 포함하는 것인 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, gRNA가 서열식별번호:01-06 중 어느 하나의 뉴클레오타이드 서열을 포함하는 것인 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, gRNA가 서열식별번호:01의 뉴클레오타이드 서열을 포함하는 것인 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, gRNA가 서열식별번호:07의 뉴클레오타이드 서열을 포함하는 것인 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, gRNA를 코딩하는 폴리뉴클레오타이드를 세포에 도입하는 단계가 세포를 CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드를 포함하는 리보핵단백질 (RNP)과 접촉시키는 것을 포함하는 것인 방법.

청구항 8

제7항에 있어서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드가 0.1:1 내지 1:10의 비율을 갖는 것인 방법.

청구항 9

제7항에 있어서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드가 1:1 내지 1:5의 비율을 갖는 것인 방법.

청구항 10

제7항에 있어서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드가 약 1:2.5의 비율을 갖는 것인 방법.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 주형 폴리뉴클레오타이드가 HBB 유전자의 적어도 일부 또는 그의 상보체를 코딩하는 것인 방법.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 주형 폴리뉴클레오티드가 야생형 HBB 유전자의 적어도 일부 또는 그의 상보체를 코딩하는 것인 방법.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, HBB 유전자의 적어도 일부가 HBB 유전자의 엑손 1을 포함하는 것인 방법.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 바이러스 벡터가 주형 폴리뉴클레오티드를 포함하는 것인 방법.

청구항 15

제14항에 있어서, 벡터가 아데노-연관 바이러스 (AAV) 벡터인 방법.

청구항 16

제14항 또는 제15항에 있어서, 벡터가 자기-상보적 AAV (scAAV) 벡터인 방법.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 주형 폴리뉴클레오티드가 HBB 유전자의 적어도 약 4 kb를 포함하는 것인 방법.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, 단일-가닥 공여자 올리고뉴클레오티드 (ssODN)가 주형 폴리뉴클레오티드를 포함하는 것인 방법.

청구항 19

제18항에 있어서, ssODN이 서열식별번호:64-72 중 어느 하나의 뉴클레오티드 서열과 적어도 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 것인 방법.

청구항 20

제18항 또는 제19항에 있어서, ssODN이 서열식별번호:64-72 중 어느 하나의 뉴클레오티드 서열을 포함하는 것인 방법.

청구항 21

제1항 내지 제20항 중 어느 한 항에 있어서, 이중-가닥 과단이 HBB 유전자의 엑손 1에서 생성되는 것인 방법.

청구항 22

제21항에 있어서, 이중-가닥 과단이 HBB 유전자의 엑손 1에서 6번째 코돈에 인접하게 생성되는 것인 방법.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 단계 (i)이 단계 (ii) 전에 수행되는 것인 방법.

청구항 24

제1항 내지 제22항 중 어느 한 항에 있어서, 단계 (i) 및 (ii)가 동시에 수행되는 것인 방법.

청구항 25

제1항 내지 제24항 중 어느 한 항에 있어서, 단계 (i) 및/또는 (ii)가 뉴클레오펙션을 수행하는 것을 포함하는 것인 방법.

청구항 26

제25항에 있어서, 뉴클레오펙션을 수행하는 것이 론자(LONZA) 시스템의 사용을 포함하는 것인 방법.

청구항 27

제26항에 있어서, 시스템이 구형과 펄스의 사용을 포함하는 것인 방법.

청구항 28

제1항 내지 제27항 중 어느 한 항에 있어서, 단계 (i) 및/또는 (ii)가 약 200,000개 세포/20 μ l 뉴클레오펙션 반응물을 접촉시키는 것을 포함하고, 뉴클레오펙션 반응물이 gRNA 및/또는 주형 폴리뉴클레오티드를 포함하는 것인 방법.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서, 세포가 포유동물인 방법.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 세포가 인간인 방법.

청구항 31

제1항 내지 제30항 중 어느 한 항에 있어서, 세포가 1차 세포인 방법.

청구항 32

제1항 내지 제31항 중 어느 한 항에 있어서, 세포가 조혈 줄기 세포 (HSC)인 방법.

청구항 33

제1항 내지 제32항 중 어느 한 항에 있어서, 세포가 T 세포 또는 B 세포인 방법.

청구항 34

제1항 내지 제33항 중 어느 한 항에 있어서, 세포가 CD34+ 세포인 방법.

청구항 35

제1항 내지 제34항 중 어느 한 항에 있어서, HBB 유전자가 서열식별번호:37의 뉴클레오티드 서열과 적어도 95% 동일성을 갖는 것인 방법.

청구항 36

제1항 내지 제35항 중 어느 한 항에 따라 제조된 세포를 이를 필요로 하는 대상체에게 투여하는 것을 포함하는, 대상체에서 겸상 적혈구 빈혈을 치료하거나 호전시키는 방법.

청구항 37

제36항에 있어서, 대상체가 인간인 방법.

청구항 38

HBB 유전자의 적어도 일부를 코딩하는 제1 서열;

가이드 RNA 절단 부위를 코딩하는 제2 서열; 및

뉴클레아제 결합 부위를 코딩하는 제3 서열

을 포함하는, HBB 유전자의 상동성 지정 복구 (HDR)를 위한 핵산.

청구항 39

제38항에 있어서, HBB 유전자의 적어도 일부가 HBB 유전자의 엑손 1을 포함하는 것인 핵산.

청구항 40

제38항 또는 제39항에 있어서, HBB 유전자가 서열식별번호: 37의 뉴클레오타이드 서열과 적어도 95% 동일성을 갖는 핵산을 포함하는 것인 핵산.

청구항 41

제38항 내지 제40항 중 어느 한 항에 있어서, 제2 서열이 서열식별번호:01-06 중 어느 하나에 기재된 뉴클레오타이드 서열과 적어도 95% 동일성을 갖는 핵산을 포함하는 것인 핵산.

청구항 42

제38항 내지 제41항 중 어느 한 항에 있어서, 제2 서열이 서열식별번호:01-06 중 어느 하나에 기재된 뉴클레오타이드 서열을 포함하는 것인 핵산.

청구항 43

제38항 내지 제42항 중 어느 한 항에 있어서, 제2 서열이 서열식별번호:07-12 중 어느 하나에 기재된 뉴클레오타이드 서열을 포함하는 것인 핵산.

청구항 44

제38항 내지 제43항 중 어느 한 항에 있어서, 제2 서열이 서열식별번호:01의 뉴클레오타이드 서열을 포함하는 것인 핵산.

청구항 45

제38항 내지 제44항 중 어느 한 항에 있어서, 하나 이상의 뉴클레아제 결합 부위가 정방향 및 역방향 전사 활성 인자-유사 이펙터 뉴클레아제 (TALEN) 결합 부위를 포함하는 것인 핵산.

청구항 46

제38항 내지 제45항 중 어느 한 항에 있어서, 하나 이상의 뉴클레아제 결합 부위가 클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR) 연관 단백질 9 (Cas9) 결합 부위인 핵산.

청구항 47

제38항 내지 제46항 중 어느 한 항에 있어서, 하나 이상의 인핸서 요소를 추가로 포함하는 핵산.

청구항 48

제38항 내지 제47항 중 어느 한 항에 있어서, 상동성 아암 서열을 추가로 포함하는 핵산.

청구항 49

제38항 내지 제48항 중 어느 한 항에 있어서, 프로모터를 코딩하는 핵산 서열을 추가로 포함하는 핵산.

청구항 50

제38항 내지 제49항 중 어느 한 항의 핵산을 포함하는 세포.

청구항 51

제50항에 있어서, 포유동물인 세포.

청구항 52

제50항 또는 제51항에 있어서, 인간인 세포.

청구항 53

제50항 내지 제52항 중 어느 한 항에 있어서, 1차 세포인 세포.

청구항 54

제50항 내지 제53항 중 어느 한 항에 있어서, 조혈 줄기 세포 (HSC)인 세포.

청구항 55

제50항 내지 제54항 중 어느 한 항에 있어서, T 세포 또는 B 세포인 세포.

청구항 56

제50항 내지 제55항 중 어느 한 항에 있어서, CD34+ 세포인 세포.

청구항 57

제50항 내지 제56항 중 어느 한 항에 있어서, 생체외인 세포.

청구항 58

제38항 내지 제49항 중 어느 한 항의 핵산을 포함하는 벡터.

청구항 59

제58항에 있어서, 아데노-연관 바이러스 벡터 (AAV)인 벡터.

청구항 60

제58항 또는 제59항에 있어서, 자기-상보적 AAV (scAAV)인 벡터.

청구항 61

제59항 내지 제60항 중 어느 한 항의 벡터 및 뉴클레아제를 코딩하는 핵산을 포함하는, 세포에서 HBB 유전자의 상동성 지정 복구 (HDR)를 위한 시스템.

청구항 62

제61항에 있어서, 뉴클레아제가 TALEN 뉴클레아제인 시스템.

청구항 63

제61항에 있어서, 뉴클레아제가 CAS 뉴클레아제인 시스템.

청구항 64

제61항 내지 제63항 중 어느 한 항에 있어서, 벡터 및 핵산이 세포로의 공동전달을 위해 구성된 것인 시스템.

청구항 65

제64항에 있어서, 세포로의 공동전달이 내인성 HBB 유전자좌를 변형시키는 것인 시스템.

청구항 66

제61항 내지 제65항 중 어느 한 항에 있어서, 세포가 포유동물인 시스템.

청구항 67

제61항 내지 제66항 중 어느 한 항에 있어서, 세포가 인간인 시스템.

청구항 68

제61항 내지 제67항 중 어느 한 항에 있어서, 세포가 1차 세포인 시스템.

청구항 69

제61항 내지 제68항 중 어느 한 항에 있어서, 세포가 조혈 줄기 세포 (HSC)인 시스템.

청구항 70

제61항 내지 제69항 중 어느 한 항에 있어서, 세포가 T 세포 또는 B 세포인 시스템.

청구항 71

제61항 내지 제70항 중 어느 한 항에 있어서, 세포가 CD34+ 세포인 시스템.

청구항 72

제61항 내지 제71항 중 어느 한 항에 있어서, HBB 유전자가 서열식별번호:37의 뉴클레오타이드 서열과 적어도 95% 동일성을 갖는 것인 시스템.

청구항 73

제50항 내지 제57항 중 어느 한 항의 세포 또는 제58항 내지 제60항 중 어느 한 항의 벡터를 대상체에게 투여하는 단계; 및

뉴클레아제를 대상체에게 투여하는 단계

를 포함하는, HBB 유전자의 HDR 촉진을 필요로 하는 대상체에서 HBB 유전자의 HDR을 촉진하는 방법.

청구항 74

제73항에 있어서, 뉴클레아제가 TALEN 뉴클레아제인 방법.

청구항 75

제73항에 있어서, 뉴클레아제가 CAS 뉴클레아제인 방법.

청구항 76

제73항 내지 제75항 중 어느 한 항에 있어서, 뉴클레아제가 세포 또는 벡터와 함께 대상체에게 공동투여되는 것인 방법.

청구항 77

제73항 내지 제76항 중 어느 한 항에 있어서, 투여가 입양 세포 전달을 포함하는 것인 방법.

청구항 78

제73항 내지 제77항 중 어느 한 항에 있어서, 세포가 인간 세포인 방법.

청구항 79

제73항 내지 제78항 중 어느 한 항에 있어서, 세포가 1차 세포인 방법.

청구항 80

제73항 내지 제79항 중 어느 한 항에 있어서, 세포가 자가 세포인 방법.

청구항 81

제73항 내지 제80항 중 어느 한 항에 있어서, 세포가 T 세포인 방법.

청구항 82

제73항 내지 제81항 중 어느 한 항에 있어서, 세포가 HSC인 방법.

청구항 83

제73항 내지 제82항 중 어느 한 항에 있어서, 세포가 $CD34^+$ HSC인 방법.

청구항 84

제73항 내지 제83항 중 어느 한 항에 있어서, 대상체가 겸상 적혈구 질환을 앓는 것인 방법.

청구항 85

제73항 내지 제84항 중 어느 한 항에 있어서, HDR을 촉진하는 것이 HBB 유전자에 대한 하나 이상의 편집을 유발하는 것인 방법.

청구항 86

제85항에 있어서, HBB 유전자에 대한 하나 이상의 편집이 겸상 적혈구 돌연변이에 대한 보정을 포함하는 것인 방법.

청구항 87

제86항에 있어서, 겸상 적혈구 돌연변이가 E7V 돌연변이를 포함하는 것인 방법.

청구항 88

제50항 내지 제57항 중 어느 한 항의 세포 또는 제58항 내지 제60항 중 어느 한 항의 벡터를 대상체에게 투여하는 단계;

뉴클레아제를 대상체에게 투여하는 단계; 및

임의로 겸상 적혈구 질환 (SCD) 또는 SCD와 연관된 질환 증상을 위한 요법을 받는 것으로부터 이익을 얻을 수 있는 것으로 대상체를 확인 또는 선택하고/거나, 임의로 상기 대상체에서 SCD 진행의 개선 또는 SCD와 연관된 질환 증상의 개선을 측정하는 단계

를 포함하는, SCD 또는 SCD와 연관된 질환 증상의 치료, 억제 또는 호전을 필요로 하는 대상체에서 SCD 또는 SCD와 연관된 질환 증상을 치료, 억제 또는 호전시키는 방법.

청구항 89

제88항에 있어서, 뉴클레아제가 TALEN 뉴클레아제인 방법.

청구항 90

제88항에 있어서, 뉴클레아제가 CRISPR/Cas 뉴클레아제인 방법.

청구항 91

제88항 내지 제90항 중 어느 한 항에 있어서, 뉴클레아제가 세포 또는 벡터와 함께 대상체에게 공동투여되는 것인 방법.

청구항 92

제88항 내지 제91항 중 어느 한 항에 있어서, 세포가 대상체로부터의 것이고, 세포가 제*항 내지 제*항 중 어느 한 항의 핵산을 세포에 도입함으로써 유전자 변형되는 것인 방법.

청구항 93

제88항 내지 제92항 중 어느 한 항에 있어서, 투여가 입양 세포 전달에 의해 수행되는 것인 방법.

청구항 94

제88항 내지 제93항 중 어느 한 항에 있어서, 세포가 인간 세포인 방법.

청구항 95

제88항 내지 제94항 중 어느 한 항에 있어서, 세포가 1차 세포인 방법.

청구항 96

제88항 내지 제95항 중 어느 한 항에 있어서, 세포가 자가 세포인 방법.

청구항 97

제88항 내지 제96항 중 어느 한 항에 있어서, 세포가 T 세포인 방법.

청구항 98

제88항 내지 제97항 중 어느 한 항에 있어서, 세포가 HSC인 방법.

청구항 99

제88항 내지 제98항 중 어느 한 항에 있어서, 세포가 $CD34^+$ HSC인 방법.

청구항 100

제88항 내지 제99항 중 어느 한 항에 있어서, 세포를 대상체의 골수에 생착시키는 것을 추가로 포함하는 방법.

청구항 101

제100항에 있어서, 세포가 대상체로부터의 것이고, 세포가 골수와 동일한 대상체로부터의 것인 방법.

발명의 설명

기술 분야

[0001] <관련 출원>

[0002] 본 출원은 발명의 명칭이 "헤모글로빈-관련 돌연변이를 편집하기 위한 상동성-지정 복구 주형 설계 및 전달"인 2019년 3월 19일에 출원된 미국 가출원 번호 62/820521 및 발명의 명칭이 "헤모글로빈-관련 돌연변이를 편집하기 위한 상동성-지정 복구 주형 설계 및 전달"인 2018년 4월 27일에 출원된 미국 가출원 번호 62/663553을 우선권 주장하며, 이들의 개시내용은 그 전문이 본원에 참조로 각각 포함된다.

[0003] <서열 목록에 대한 참조>

[0004] 본 출원은 전자 형식의 서열 목록과 함께 출원된다. 서열 목록은 2019년 4월 15일에 생성된 파일명 SCRI194WOSEQLISTING의 파일로서 제공되며, 이는 크기가 대략 159 Kb이다. 서열 목록의 전자 형식의 정보는 그 전문이 본원에 참조로 포함된다.

[0005] <발명의 기술분야>

[0006] 본원에 제공된 방법 및 조성물의 일부 실시양태는 헤모글로빈 유전자좌, 예컨대 겸상 적혈구 돌연변이를 포함하는 헤모글로빈-관련 돌연변이의 변형에 관한 것이다. 일부 실시양태는 겸상 적혈구 돌연변이의 부위에 포스포디에스테르 DNA 가닥 파단의 도입을 통한 겸상 적혈구 돌연변이의 변형에 관한 것이다.

배경 기술

[0007] <발명의 배경>

[0008] 겸상 적혈구 질환 (SCD)은 혈액 장애, 예컨대 겸상 적혈구 빈혈을 포함한다. 일부 경우에, SCD는 적혈구에서 발견되는 산소-운반 단백질 헤모글로빈의 이상을 초래한다. 이는 단단한 겸상-유사 형상 및/또는 빈혈을 포함하는 적혈구를 초래할 수 있다.

[0009] 엔도뉴클레아제-기반 시스템은 빠르게 유의한 유전자 편집 도구가 되었다. 유전자 편집을 위한 엔도뉴클레아제-기반 접근법의 예는 제한 없이, 아연 핑거 뉴클레아제 (ZFN), TAL 이펙터 뉴클레아제 (TALEN), 메가뉴클레아제 (예컨대 MegaTAL) 및 CRISPR/Cas9를 포함하는 시스템을 포함한다. SCD를 억제 또는 치료하기 위한 더 많은 접근법에 대한 필요성은 명백하다.

발명의 내용

- [0010] <요약>
- [0011] 본원에 제공된 방법 및 조성물의 실시양태는 헤모글로빈 유전자좌, 예컨대 겸상 적혈구 돌연변이를 포함하는 헤모글로빈-관련 돌연변이의 변형에 관한 것이다. 일부 실시양태는 HBB 유전자의 상동성 지정 복구 (HDR)를 위한 핵산에 관한 것이다.
- [0012] 일부 실시양태는 (i) 가이드 RNA (gRNA)를 코딩하는 폴리뉴클레오티드를 세포에 도입하는 단계, 및 (ii) 주형 폴리뉴클레오티드를 세포에 도입하는 단계를 포함하는, 세포에서 HBB 유전자를 편집하는 방법을 포함한다.
- [0013] 일부 실시양태에서, gRNA는 서열식별번호(SEQ ID NO):01-06 중 어느 하나의 뉴클레오티드 서열과 적어도 95% 동일성을 갖는 핵산을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:07-12 중 어느 하나의 뉴클레오티드 서열과 적어도 95% 동일성을 갖는 핵산을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:01-06 중 어느 하나의 뉴클레오티드 서열을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:01의 뉴클레오티드 서열을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:07의 뉴클레오티드 서열을 포함한다.
- [0014] 일부 실시양태에서, gRNA를 코딩하는 폴리뉴클레오티드를 세포에 도입하는 단계는 세포를 CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오티드를 포함하는 리보핵단백질 (RNP)과 접촉시키는 것을 포함한다. 일부 실시양태에서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오티드는 0.1:1 내지 1:10의 비율을 갖는다. 일부 실시양태에서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오티드는 1:1 내지 1:5의 비율을 갖는다. 일부 실시양태에서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오티드는 약 1:2.5의 비율을 갖는다.
- [0015] 일부 실시양태에서, 주형 폴리뉴클레오티드는 HBB 유전자의 적어도 일부 또는 그의 상보체를 코딩한다. 일부 실시양태에서, 주형 폴리뉴클레오티드는 야생형 HBB 유전자의 적어도 일부 또는 그의 상보체를 코딩한다. 일부 실시양태에서, HBB 유전자의 적어도 일부는 HBB 유전자의 엑손 1을 포함한다.
- [0016] 일부 실시양태에서, 바이러스 벡터는 주형 폴리뉴클레오티드를 포함한다. 일부 실시양태에서, 벡터는 아테노-연관 바이러스 (AAV) 벡터이다. 일부 실시양태에서, 벡터는 자기-상보적 AAV (scAAV) 벡터이다. 일부 실시양태에서, 주형 폴리뉴클레오티드는 HBB 유전자의 적어도 약 4 kb를 포함한다.
- [0017] 일부 실시양태에서, 단일-가닥 공여자 올리고뉴클레오티드 (ssODN)는 주형 폴리뉴클레오티드를 포함한다. 일부 실시양태에서, ssODN은 서열식별번호:64-72 중 어느 하나의 뉴클레오티드 서열과 적어도 95% 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 실시양태에서, ssODN은 서열식별번호:64-72 중 어느 하나의 뉴클레오티드 서열을 포함한다.
- [0018] 일부 실시양태에서, 이중-가닥 파단은 HBB 유전자의 엑손 1에서 생성된다. 일부 실시양태에서, 이중-가닥 파단은 HBB 유전자의 엑손 1에서 6번째 코돈에 인접하게 생성된다.
- [0019] 일부 실시양태에서, 단계 (i)은 단계 (ii) 전에 수행된다. 일부 실시양태에서, 단계 (i) 및 (ii)는 동시에 수행된다. 일부 실시양태에서, 단계 (i) 및/또는 (ii)는 뉴클레오펙션을 수행하는 것을 포함한다. 일부 실시양태에서, 뉴클레오펙션을 수행하는 것은 룬자(LONZA) 시스템의 사용을 포함한다. 일부 실시양태에서, 시스템은 구형과 펄스의 사용을 포함한다. 일부 실시양태에서, 단계 (i) 및/또는 (ii)는 약 200,000개 세포/20 μ l 뉴클레오펙션 반응물을 접촉시키는 것을 포함하고, 뉴클레오펙션 반응물은 gRNA 및/또는 주형 폴리뉴클레오티드를 포함한다.
- [0020] 일부 실시양태에서, 세포는 포유동물이다. 일부 실시양태에서, 세포는 인간이다. 일부 실시양태에서, 세포는 1차 세포이다. 일부 실시양태에서, 세포는 조혈 줄기 세포 (HSC)이다. 일부 실시양태에서, 세포는 T 세포 또는 B 세포이다. 일부 실시양태에서, 세포는 CD34⁺ 세포이다.
- [0021] 일부 실시양태에서, HBB 유전자는 서열식별번호:37의 뉴클레오티드 서열과 적어도 95% 동일성을 갖는다.
- [0022] 일부 실시양태에서, 핵산은 다음 중 하나 이상을 포함한다: HBB 유전자를 코딩하는 제1 서열; 하나 이상의 가이드 RNA 절단 부위를 코딩하는 제2 서열; 및 하나 이상의 뉴클레아제 결합 부위를 코딩하는 제3 서열.
- [0023] 일부 실시양태에서, HBB 유전자는 서열식별번호: 37에 기재된 핵산 서열을 포함한다. 일부 실시양태에서, 제2 서열은 서열식별번호: 1에 기재된 핵산 서열을 포함한다. 일부 실시양태에서, 하나 이상의 뉴클레아제 결합 부위는 정방향 및 역방향 전사 활성인자-유사 이펙터 뉴클레아제 (TALEN) 결합 부위를 포함한다. 일부 실시양태

에서, 하나 이상의 핵산 결합 부위는 클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR) 연관 단백질 9 (Cas9) 결합 부위이다. 일부 실시양태는 하나 이상의 인핸서 요소를 포함한다. 일부 실시양태는 상동성 아암 서열을 포함한다. 일부 실시양태는 프로모터를 코딩하는 핵산 서열을 포함한다.

[0024] 일부 실시양태는 세포에서 HBB 단백질 발현의 HDR을 촉진하기 위한 벡터에 관한 것이다. 일부 실시양태에서, 벡터는 다음 중 하나 이상을 포함한다: HBB 유전자를 코딩하는 제1 서열; 하나 이상의 가이드 RNA 절단 부위를 코딩하는 제2 서열; 및 하나 이상의 뉴클레아제 결합 부위를 코딩하는 제3 서열.

[0025] 일부 실시양태에서, HBB 유전자는 서열식별번호: 37에 기재된 핵산 서열을 포함한다. 일부 실시양태에서, 제2 서열은 서열식별번호: 1에 기재된 핵산 서열을 포함한다. 일부 실시양태에서, 하나 이상의 뉴클레아제 결합 부위는 정방향 및 역방향 전사 활성인자-유사 이펙터 뉴클레아제 (TALEN) 결합 부위를 포함한다. 일부 실시양태에서, 하나 이상의 핵산 결합 부위는 클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR) 연관 단백질 9 (Cas9) 결합 부위이다. 일부 실시양태는 하나 이상의 인핸서 요소를 포함한다. 일부 실시양태에서, 벡터는 아데노-연관 바이러스 벡터 (AAV)이다. 일부 실시양태에서, 벡터는 자기-상보적 AAV (scAAV)이다. 일부 실시양태에서, 세포는 인간 세포이다. 일부 실시양태에서, 세포는 1차 세포이다. 일부 실시양태에서, 세포는 자가 세포이다. 일부 실시양태에서, 세포는 T 세포이다. 일부 실시양태에서, 세포는 조혈 줄기 세포 (HSC)이다. 일부 실시양태에서, 세포는 CD34⁺ HSC이다. 일부 실시양태는 세포에서 CD40L 단백질 발현의 HDR을 촉진하기 위한 시스템에 관한 것이다. 일부 실시양태에서, 시스템은 상기 단락 중 하나 이상의 임의의 벡터, 및 뉴클레아제를 코딩하는 핵산을 포함한다. 일부 실시양태에서, 뉴클레아제는 TALEN 뉴클레아제이다. 일부 실시양태에서, 뉴클레아제는 Cas 뉴클레아제이다. 일부 실시양태에서, 벡터 및 핵산은 세포로의 공동전달을 위해 구성된다. 일부 실시양태에서, 세포로의 공동전달은 내인성 HBB 유전자좌를 변형시킨다. 일부 실시양태에서, 세포는 1차 인간 조혈 세포이다.

[0026] 일부 실시양태는 HBB를 발현하기 위한 세포에 관한 것이다. 일부 실시양태에서, 세포는 핵산을 포함한다. 일부 실시양태에서, 핵산은 다음 중 하나 이상을 포함한다: HBB 유전자를 코딩하는 제1 서열; 프로모터를 코딩하는 제2 서열; 하나 이상의 가이드 RNA 절단 부위를 코딩하는 제3 서열; 및 하나 이상의 뉴클레아제 결합 부위를 코딩하는 제4 서열.

[0027] 일부 실시양태에서, 핵산은 벡터 내에 있다. 일부 실시양태에서, 벡터는 AAV이다. 일부 실시양태에서, AAV는 scAAV이다. 일부 실시양태에서, 세포는 인간 세포이다. 일부 실시양태에서, 세포는 1차 세포이다. 일부 실시양태에서, 세포는 자가 세포이다. 일부 실시양태에서, 세포는 T 세포이다. 일부 실시양태에서, 세포는 HSC이다. 일부 실시양태에서, 세포는 CD34⁺ HSC이다.

[0028] 일부 실시양태는 HBB 유전자의 HDR 촉진을 필요로 하는 대상체에서 HBB 유전자의 HDR을 촉진하는 방법에 관한 것이다. 일부 실시양태에서, 방법은 다음 중 하나 이상을 포함한다: 상기 단락 중 하나 이상의 임의의 세포 또는 벡터를 대상체에게 투여하는 단계; 및 뉴클레아제를 대상체에게 투여하는 단계.

[0029] 일부 실시양태에서 뉴클레아제는 TALEN 뉴클레아제이다. 일부 실시양태에서, 뉴클레아제는 Cas 뉴클레아제이다. 일부 실시양태에서, 뉴클레아제는 세포 또는 벡터와 함께 대상체에게 공동투여된다. 일부 실시양태에서, 세포는 대상체로부터의 것이고, 여기서 세포는 상기 단락 중 하나 이상의 핵산 또는 벡터를 세포에 도입함으로써 유전자 변형된다. 일부 실시양태에서, 투여는 입양 세포 전달에 의해 수행된다. 일부 실시양태에서, 세포는 인간 세포이다. 일부 실시양태에서, 세포는 1차 세포이다. 일부 실시양태에서, 세포는 자가 세포이다. 일부 실시양태에서, 세포는 T 세포이다. 일부 실시양태에서, 세포는 HSC이다. 일부 실시양태에서, 세포는 CD34⁺ HSC이다. 일부 실시양태에서, 대상체는 겸상 적혈구 질환을 앓는다. 일부 실시양태에서, HDR을 촉진하는 것은 HBB 유전자에 대한 하나 이상의 편집을 유발한다. 일부 실시양태에서, HBB 유전자에 대한 하나 이상의 편집은 겸상 적혈구 돌연변이에 대한 보정을 포함한다. 일부 실시양태에서, 겸상 적혈구 돌연변이는 E7V 돌연변이를 포함한다.

[0030] 일부 실시양태는 겸상 적혈구 질환 (SCD) 또는 SCD와 연관된 질환 증상의 치료, 억제 또는 호전을 필요로 하는 대상체에서 겸상 적혈구 질환 (SCD) 또는 SCD와 연관된 질환 증상을 치료, 억제 또는 호전시키는 방법에 관한 것이다. 일부 실시양태에서, 방법은 다음 중 하나 이상을 포함한다: 상기 단락 중 어느 하나 이상의 세포 또는 벡터를 대상체에게 투여하는 단계; 뉴클레아제를 대상체에게 투여하는 단계; 및 임의로 SCD 또는 SCD와 연관된 질환 증상을 위한 요법을 받는 것으로부터 이익을 얻을 수 있는 것으로 대상체를 확인 또는 선택하고/거나, 임

의로 상기 대상체에서 SCD 진행의 개선 또는 SCD와 연관된 질환 증상의 개선을 측정하는 단계.

[0031]

일부 실시양태에서, 뉴클레아제는 TALEN 뉴클레아제이다. 일부 실시양태에서, 뉴클레아제는 CRISPR/Cas 뉴클레아제이다. 일부 실시양태에서, 뉴클레아제는 세포 또는 벡터와 함께 대상체에게 공동투여된다. 일부 실시양태에서, 세포는 대상체로부터의 것이고, 여기서 세포는 상기 단락 중 어느 하나 이상의 핵산 또는 벡터를 세포에 도입함으로써 유전자 변형된다. 일부 실시양태에서, 투여는 입양 세포 전달에 의해 수행된다. 일부 실시양태에서, 세포는 인간 세포이다. 일부 실시양태에서, 세포는 1차 세포이다. 일부 실시양태에서, 세포는 자가 세포이다. 일부 실시양태에서, 세포는 T 세포이다. 일부 실시양태에서, 세포는 HSC이다. 일부 실시양태에서, 세포는 CD34⁺ HSC이다. 일부 실시양태는 세포를 대상체의 골수에 정착시키는 것을 포함한다. 일부 실시양태에서, 세포는 대상체로부터의 것이고, 세포는 골수와 동일한 대상체로부터의 것이다.

도면의 간단한 설명

[0032]

도 1a, 도 1b 및 도 1c는 HBB 유전자좌에서 뉴클레아제에 의한 효율적인 편집을 나타내는 데이터를 나타내는 그래픽 표시를 도시한다. 각각의 도 1a, 도 1b 및 도 1c는 뉴클레아제에 대한 반응으로 % INDEL을 도시하는 그래프이다.

도 2a, 도 2b, 도 2c, 도 2d, 도 2e 및 도 2f는 결실 복구 주형 1242, 1243, 1244, 1245의 rAAV6 전달을 시험하기 위한 설계를 나타내는 그래픽 표시, 및 관련 데이터를 도시한다.

도 3a, 도 3b, 도 3c, 도 3d, 도 3e 및 도 3f는 비결실 복구 주형 1289, 1290의 rAAV6 전달을 시험하기 위한 설계의 도시, 및 관련 데이터를 포함한다.

도 4a, 도 4b, 도 4c 및 도 4d는 인간 코돈-최적화 겸상 도입 카세트 1246, 1247, 1248, 1249의 rAAV6 전달을 시험하기 위한 설계의 도시, 및 관련 데이터를 포함한다.

도 5a, 도 5b, 도 5c, 도 5d, 도 5e 및 도 5f는 겸상 돌연변이 도입 (GTC) 1314의 rAAV6 전달을 시험하기 위한 설계의 도시, 및 관련 데이터를 포함한다.

도 6a, 도 6b 및 도 6c는 겸상 도입 복구 주형 1321의 rAAV6 전달을 시험하기 위한 설계의 도시, 및 관련 데이터를 포함한다.

도 7a, 도 7b, 도 7c 및 도 7d는 겸상 돌연변이 (GTC 변화)를 도입하는 ssODN을 시험하기 위한 설계의 도시, 및 관련 데이터를 포함한다.

도 8a, 도 8b, 도 8c, 도 8d, 도 8e 및 도 8f는 겸상 보정 (CCC GAA 변화)을 위한 ssODN을 시험하기 위한 설계의 도시, 및 관련 데이터를 포함한다.

도 9a, 도 9b 및 도 9c는 12주에 W41 마우스의 골수에서 편집된 인간 세포의 생착을 나타내는 그래픽 데이터를 포함한다.

도 10a는 sgRNA 및 TALEN 결합 부위의 위치를 나타내는 게놈 HBB 유전자의 개략도를 도시한다. GAG (코돈 6)로부터 GTC 또는 GTG로의 뉴클레오타이드 치환은 아미노산을 글루타메이트로부터 발린으로 변화시키고, SCD를 유발한다.

도 10b는 TIDE/ICE에 의해 측정된 HBB 유전자에서 DSB를 생성하는 TALEN mRNA (T) 또는 후보 sgRNA (g1-g6, RNP로서 전달됨)의 스크리닝을 도시한다 (공여자 n = 2-3).

도 10c는 mPBSC에서 편집 효율을 최대화하는 cas9: sgRNA 비율의 최적화를 도시한다. NHEJ 비율은 TIDE/ICE (1:1의 Cas9:sgRNA 비율 (각각 40 pmol), 공여자 n = 2 또는 1:2.5의 비율 (20 pmol의 Cas9 및 50 pmol의 sgRNA, 공여자 n = 15)에 의해 분석되었다.

도 10d는 RNP로서 전달된 sgRNA-g1을 사용하여 mPBSC에서 MiSeq 분석에 의한 HBB에서 온-타겟 파괴 및 HBD에서 가능한 오프-타겟 파괴의 평가를 도시한다 (공여자 n = 7).

도 10e는 NEON 전기천공 시스템 (공여자 n = 15) 또는 룬자(Lonza) 뉴클레오펙션 시스템 (공여자 n = 3)을 사용하여 RNP로서 전달된 sgRNA-g1의 편집 효율을 도시한다. 모든 막대 그래프는 평균 ± SD를 나타낸다. * p < 0.05 ** p < 0.01 *** p < 0.001 ****, p < 0.0001. p-값은 던넛(Dunnett)의 다중 비교와 이원 ANOVA에 의해 각 샘플 평균을 각 대조군 샘플 평균과 비교함으로써 계산하였다.

도 10f는 CCTop 알고리즘에 의해 예측된 상위 5개의 오프-타겟 유전자의 오프-타겟 분석을 도시한다. 겔은 T7 엔도뉴클레아제 검정에 의해 평가된 모의-처리 (M) 및 sgRNA-g1 RNP-처리 (RNP) 샘플로부터 증폭된 상위 5개의 오프-타겟 유전자의 앰플리콘을 나타낸다. (i) OT1: DENND3 (레인 1-2), (ii) OT2: MIR7974 (레인 3-4), (iii) OT3: LINC01206 (레인 5-6) (iv) OT4: HBD (레인 7-8) (v) OT5: TULP4 (레인 9-10) (vi) 표적 부위: HBB (레인 11-12). 별표 (*)는 절단된 밴드를 나타낸다. #은 어떤 잠재적 절단 단편 (TULP4의 경우 313 bp 및 143 bp)과도 매칭하지 않는 고스트 밴드를 나타낸다.

도 10g는 상위 5개의 오프-타겟 유전자 (i) OT1: DENND3 (ii) OT2: MIR7679 (iii) OT3: LINC01206 (iv) OT4: HBD (v) OT5: TULP4 (vi) 표적 부위: HBB (n = 2 실험)의 TIDE/ICE 시퀀싱 분석을 도시한다.

도 11a는 HDR에 의해 코돈 6에서 겸상 돌연변이를 도입하는 GTC (E6V) 또는 코돈 최적화 SNP 변화를 도입하는 GAA (E6optE)를 구동하도록 설계된 rAAV6 카세트의 개략도를 도시한다.

도 11b는 RNP 및 rAAV6 전달을 사용한 유전자-편집에 이어 mPBSC에서 적혈구계 분화를 시험하기 위한 실험 타임 라인을 도시한다.

도 11c는 RNP 단독으로 전기천공, rAAV6 공여자 주형 단독으로 형질도입, 또는 표시된 농도에서 RNP 및 GTC (E6V) rAAV6 공여자 주형의 공동전달 후, WT (%), ddPCR에 의해 측정된 HDR (%) 및 TIDE/ICE 시퀀싱에 의해 측정된 NHEJ (%)를 각각 도시한다 (공여자 n = 4).

도 11d는 RNP 단독, rAAV6 단독으로 처리된 세포, 또는 RNP + GTC (E6V) rAAV6 처리된 세포에서 β -글로빈 발현을 측정하기 위한 적혈구계 세포의 RP-HPLC 분석을 도시한다 (공여자 n = 7).

도 11e는 RNP 단독으로 전기천공, rAAV6 공여자 주형 단독으로 형질도입, 또는 표시된 농도에서 RNP 및 GAA (E6optE) rAAV6 공여자 주형의 공동전달 후, WT (%), ddPCR에 의해 측정된 HDR (%) 및 TIDE/ICE 시퀀싱에 의해 측정된 NHEJ (%)를 각각 도시한다 (1% rAAV6; 공여자 n = 3).

도 11f는 RNP 단독, rAAV6 단독으로 처리된 세포, 또는 RNP + GAA (E6optE) rAAV6 처리된 세포에서 β -글로빈 발현을 측정하기 위한 적혈구계 세포의 RP-HPLC 분석을 도시한다 (공여자 n = 3). (β A = 성인 글로빈, β S = 겸상 글로빈, γ G = 감마 2, γ A = 감마 1). 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001. p-값은 던넛의 다중 비교와 이원 ANOVA에 의해 NHEJ (%), HDR (%), WT (%) 또는 글로빈 하위유형 (%)의 각 샘플 평균을 모의 샘플의 각 NHEJ (%), HDR (%), WT (%) 또는 글로빈 하위유형 (%)과 비교함으로써 계산하였다.

도 11g는 Neon 전기천공 시스템 (n = 3) 또는 론자 뉴클레오택션 시스템 (n = 1)을 사용하여 RNP 및 GTC (E6V) rAAV6의 공동전달 후 각각 ddPCR 및 TIDE/ICE 시퀀싱에 의해 측정된 HDR 및 NHEJ 결과를 도시한다.

도 11h는 RNP로 편집되고 neon 전기천공을 사용하여 GTC (E6V) rAAV6 (n = 5)으로 형질도입된 샘플의 콜로니 시퀀싱을 도시한다.

도 11i는 전기천공 및 GTC (E6V) 또는 GAA (E6optE) rAAV6 형질도입 후 2일차에 mPBSC의 생존율을 도시한다.

도 11j는 rAAV6 단독 및 RNP + GTC (E6V) rAAV6으로 처리된 세포에서 시험관내 글로빈 사량체를 결정하기 위한 적혈구계 세포의 IEC를 도시한다 (HbF: 태아, HbA: 성인, HbA2: 미성년자, HbS: 겸상). 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. n은 개별 실험의 수를 나타낸다. * p < 0.05, ** p < 0.01, *** p < 0.001, ****, p < 0.0001. p-값은 던넛의 다중 비교와 이원 ANOVA에 의해 각 샘플 평균을 각 대조군 샘플 평균과 비교함으로써 계산하였다.

도 11k는 편집되고 분화된 적혈구계 세포의 RP-HPLC 분석을 도시한다. 겸상 글로빈 발현을 구동하는 참조, 모의, rAAV6 단독 및 RNP + GTC (E6V) rAAV6 (3%) 형질도입된 세포의 RP-HPLC 크로마토그램 트레이스 (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1). 수직 숫자는 HPLC 용출 시간이다. 아래쪽 트레이스는 겸상 글로빈 발현을 나타낸다 (적색 화살표).

도 11l은 sgRNA-g1 RNP로 시험된 rAAV6으로서 전달된 복합체 cDNA 카세트의 개략도를 도시한다. 1321은 적혈구계 인헨서와 함께 GTC 변화 (E6V 아미노산 변화)를 구동하는 HBG1 Δ 13 프로모터; HPFH-2 및 HS-40을 갖는다. MND-GFP는 HDR에 대한 대리 역할을 하며, SV40 폴리아데닐화 서열과 역방향 배향을 갖는다. 1322는 HBB 프로모터를 제거하기 위한 결실 (Δ -127, -71)을 갖는 것을 제외하고 1321과 동일한 설계를 갖는다. 실험 설정은 도 11b와 유사하였다.

도 11m은 전기천공 및 RNP-단독, rAAV6-단독의 형질도입 및 RNP + rAAV6 형질도입 후 14일에 유동 세포계측법에 의해 측정된 GFP 발현을 도시한다 (1321 공여자 $n = 2$, 1322 공여자 $n = 1$).

도 11n은 RNP 단독, rAAV6 단독으로 처리된 세포, 또는 RNP + GTC (E6V) 1321/1322 rAAV6 처리된 세포에서 β -글로빈 발현을 측정하기 위한 적혈구계 세포의 RP-HPLC 분석을 도시한다. 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. n 은 개별 실험의 수를 나타낸다. (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1).

도 11o는 편집되고 분화된 적혈구계 세포의 RP-HPLC 분석을 도시한다. 성인 글로빈 발현을 구동하는 모의, rAAV6 단독, RNP 단독 및 RNP + GAA (E6optE) rAAV6의 RP-HPLC 크로마토그램 트레이스 (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1). 수직 숫자는 HPLC 용출 시간이다. 아래쪽 트레이스는 성인 글로빈 발현의 복원을 나타낸다.

도 12a는 HDR에 의해 코돈 6에서 겸상 돌연변이를 도입하는 GTC (E6V) 또는 코돈 최적화 SNP 변화를 도입하는 GAA (E6optE)를 구동하도록 설계된 ssODN 카세트의 개략도를 도시한다.

도 12b는 RNP 및 ssODN 전달을 사용한 유전자-편집에 이어 mPBSC에서 적혈구계 분화를 시험하기 위한 실험 타임 라인을 도시한다.

도 12c는 RNP 단독으로 전기천공 또는 표시된 농도에서 RNP 및 GTC (E6V) ssODN 공여자 주형의 공동전달 후 WT (%), ddPCR에 의해 측정된 HDR (%) 및 TIDE/ICE 시퀀싱에 의해 측정된 NHEJ (%)를 각각 도시한다 (50 pmol의 ssODN; 공여자 $n = 5$).

도 12d는 RNP 단독으로 처리된 세포 또는 RNP + GTC (E6V) ssODN 처리된 세포에서 β -글로빈 발현을 측정하기 위한 적혈구계 세포의 RP-HPLC 분석을 도시한다 (50 pmol의 ssODN; 공여자 $n = 5$).

도 12e는 RNP 단독으로 전기천공 또는 표시된 농도에서 RNP 및 GAA (E6optE) ssODN의 공동전달 후 WT (%), ddPCR에 의해 측정된 HDR (%) 및 TIDE/ICE 시퀀싱에 의해 측정된 NHEJ (%)를 각각 도시한다 (50 pmol의 ssODN; 공여자 $n = 8$).

도 12f는 RNP 단독으로 처리된 세포 또는 RNP + GAA (E6optE) ssODN 처리된 세포에서 β -글로빈 발현을 측정하기 위한 적혈구계 세포의 RP-HPLC 분석을 도시한다 (50 pmol의 ssODN; 공여자 $n = 6$). (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1). 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. p -값은 던넛의 다중 비교와 이원 ANOVA에 의해 NHEJ (%), HDR (%), WT (%) 또는 글로빈 하위유형 (%)의 각 샘플 평균을 모의 샘플의 각 NHEJ (%), HDR (%), WT (%) 또는 β 글로빈 하위유형 (%)과 비교함으로써 계산하였다.

도 12g는 HDR에 의해 코돈 6에서 겸상 돌연변이를 도입하는 GTC 또는 GTG (E6V) ssODN 또는 코돈 최적화 SNP 변화를 도입하는 GAA (E6optE) ssODN으로 전기천공 후 2일차에 CD34+ mPBSC의 생존율을 도시한다.

도 12h는 RNP 단독으로 전기천공 또는 표시된 농도에서 RNP 및 공여자 GTG (E6V) ssODN의 공동전달 후 WT (%), ddPCR에 의해 측정된 HDR (%) 및 TIDE/ICE 시퀀싱에 의해 측정된 NHEJ를 각각 도시한다 (50 pmol ssODN, 공여자 $n = 3$).

도 12i는 RNP로 편집되고 Neon 전기천공 시스템으로 시험된 GTG (E6V) ssODN 및 GAA (E6optE) ssODN으로 변형된 샘플의 콜로니 시퀀싱을 도시한다 (공여자 $n = 3$).

도 12j는 GTG (E6V) ssODN 전달을 갖는 편집된 세포에서 글로빈 발현을 결정하기 위한 적혈구계 세포의 RP-HPLC 분석을 도시한다 (공여자 $n = 3$). (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1).

도 12k는 RNP 단독으로 전기천공 또는 Neon 전기천공 시스템 (50 pmol ssODN, $n = 2$) 또는 론자 뉴클레오택션 시스템 (50 pmol ssODN, $n = 3$)을 사용하여 표시된 농도에서 RNP 및 GAA (E6optE) ssODN의 공동전달 후 WT (%), ddPCR에 의해 측정된 HDR (%) 및 TIDE/ICE 시퀀싱에 의해 측정된 NHEJ (%)를 각각 도시한다. 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. n 은 개별 실험의 수를 나타낸다. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

도 12l은 편집되고 분화된 적혈구계 세포의 RP-HPLC 분석을 도시한다. 겸상 글로빈 발현을 구동하는 모의, RNP 단독 및 RNP + GTC (E6V) ssODN 및 rAAV6 공여자 주형의 RP-HPLC 크로마토그램 트레이스 (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1). 수직 숫자는 HPLC 용출 시간이다. 아래쪽 트레이스는 겸상 글

로빈 발현을 나타낸다 (화살표).

도 12m은 편집되고 분화된 적혈구계 세포의 RP-HPLC 분석을 도시한다. 겸상 글로빈 발현을 구동하는 참조, 모의, RNP 단독 및 RNP + GTG (E6V) ssODN의 RP-HPLC 크로마토그램 트레이스 (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1). 수직 숫자는 HPLC 용출 시간이다. 아래쪽 트레이스는 겸상 글로빈 발현을 나타낸다 (화살표).

도 12n은 편집되고 분화된 적혈구계 세포의 RP-HPLC 분석을 도시한다. 성인 글로빈 발현을 구동하는 모의, RNP 단독 및 RNP + GAA (E6V) ssODN 및 rAAV6 공여자 주형의 RP-HPLC 크로마토그램 트레이스 (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1). 수직 숫자는 HPLC 용출 시간이다. 아래쪽 트레이스는 성인 글로빈 발현의 복원을 나타낸다.

도 13a는 GTC (E6V) rAAV6 (n = 6) vs ssODN (GTC (E6V, n = 8), GTG (E6V, n = 3), 또는 GAA (E6V) ssODN 사용)으로 처리된 세포에서 MiSeq 분석에 의한 HDR vs. NHEJ 편집의 정량화를 도시한다.

도 13b는 rAAV6 또는 ssODN 전달로 HDR의 촉진 후 세포에 존재하는 잔류 indel을 RNP-매개 편집 단독을 비교하는 MiSeq에 의한 indel 스펙트럼 분석을 도시한다 (공여자 n = 6).

도 13c는 다음 샘플에서 측정된 다양한 유전자 편집 결과 WT, NHEJ (삽입, 치환, 결실) 및 HDR을 도시한다: 모의, RNP 단독, rAAV6과 RNP 및 ssODN과 RNP의 공동전달. 분석된 샘플은 편집 후 14일차에 분석된 이식전 투입 샘플이었다. (n = 4). 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001. p-값은 던넛의 다중 비교와 이원 ANOVA에 의해 NHEJ (%), HDR (%)의 각 샘플 평균을 모의 샘플의 각 NHEJ (%), HDR (%)과 비교함으로써 계산하였다.

도 13d는 시험관내 편집된 샘플 및 생체내 BM 샘플로부터의 정렬된 쌍형성된 말단 판독의 수를 도시한다. 각 점은 고유한 샘플을 나타낸다.

도 13e는 모의, RNP 단독, rAAV6과 RNP 및 ssODN과 RNP의 공동전달에서 관찰된 우세한 NHEJ 사건으로부터의 컨센서스 서열을 도시한다.

도 13f는 rAAV6 또는 ssODN 전달로 HDR의 촉진 후 MiSeq 분석에 의한 시험관내 및 생체내 프레임 이동 돌연변이 %의 정량화를 도시한다.

도 14a는 mPBSC에서 시험관내 및 NBSGW 마우스 모델에서 생체내 GTC (E6V) rAAV6 또는 ssODN 처리된 세포를 사용한 유전자-편집을 시험하기 위한 실험 타임라인을 도시한다.

도 14b는 FSC/SSC 및 단일 세포에 기초한 게이팅을 갖는 BM 및 비장에서 인간 세포 (hCD45+) 키메리즘을 도시한다.

도 14c는 BM hCD45+ 집단 내의 인간 CD19+ 및 CD33+ 서브세트를 도시한다.

도 14d는 mCD45- 집단에 게이팅된 BM에서 인간 CD235+ 세포를 도시한다. BM 세포를 적혈구계 분화 배지에서 14일 동안 생체외에서 배양하고, CD235+ (생체외)를 유동 세포계측법에 의해 측정하였다.

도 14e는 BM hCD45+ 집단 내의 인간 CD34+ 및 CD34⁺CD38^{lo} 세포의 비율을 도시한다.

도 14f는 GTC (E6V) rAAV6 또는 ssODN 처리된 투입 세포 내에서 ddPCR에 의해 결정된 HDR 비율 (14일차, n = 4); 및 이식 후 3주 (21일차; n = 2) 및 이식 후 12-14주 (84-96일차; 모의: n = 8, RNP + rAAV6: n = 17, RNP + ssODN: n = 18)를 도시한다.

도 14g는 GTC (E6V) rAAV6 또는 ssODN 처리된 투입 세포에 대한 TIDE/ICE 시퀀싱에 의해 결정된 NHEJ 비율 (14일차), 이식 후 3주 (21일차) 및 이식 후 12-14주 (84-96일차)를 도시한다.

도 14h는 표시된 시점에서 GTC (E6V) rAAV6 또는 ssODN 처리된 세포에 대한 MiSeq 분석에 의해 결정된 HDR 비율을 도시한다.

도 14i는 표시된 시점에서 GTC (E6V) rAAV6 또는 ssODN 처리된 세포에 대한 MiSeq 분석에 의해 결정된 NHEJ 비율을 도시한다. n은 샘플 또는 동물을 나타낸다. 투입 n = 4, 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. ns: 유의하지 않음. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001. p-값은 던넛의 다중 비교와 이원 ANOVA에 의해 NHEJ (%), HDR (%) 및 WT (%)의 각 샘플 평균을 모의 샘플의 각 NHEJ (%), HDR (%) 및 WT

(%)와 비교함으로써 계산하였다.

도 14j는 hCD45+ 집단으로부터 게이팅된 비장에서의 인간 CD19+ 및 CD33+ 집단을 도시한다.

도 14k는 비-CD19+ 및 비-CD33+ 세포로부터 게이팅된 BM 및 비장에서의 인간 CD3+ 집단을 도시한다.

도 14l은 GTC (E6V) 공여자로 편집된 HSC로 이식된 NBSGW 수용자 마우스의 BM 내의 인간 세포 (hCD45+)의 대표적인 유동 플롯을 도시한다. 유동 플롯은 (i) 모의-편집된, (ii) rAAV6-편집된 및 (iii) ssODN-편집된 세포 수용자의 BM 내의 CD19+, CD33+ 및 CD235+ 세포를 포함하는 다중-계통 생착을 나타낸다. 게이팅 전략: 살아있는, 단일 세포, hCD45+ > CD19+ CD33+. 적혈구계 세포는 mCD45- 세포 상에서 게이팅되었다.

도 14m (왼쪽 패널)은 (i) 모의-편집된, (ii) rAAV6-편집된 또는 (iii) ssODN-편집된 (GTC (E6V)로 변형된) 집단을 나타내는 이식전 CD34+ 및 CD34+ CD38lo 세포의 대표적인 유동 플롯을 도시한다. 게이팅 전략: 살아있는, 단일 세포, hCD45+ > CD34+ CD38+ > CD90+ CD133+. 도 14n (오른쪽 패널)은 CD133+ CD90+ 이중 양성 세포에 의해 확인된 바와 같이 LT-HSC에 대해 농축된 집단을 확인하기 위한 추가 마커를 사용한 CD34+CD38lo 세포의 대표적인 유동 플롯을 도시한다.

도 14n은 (i) 모의-편집된, (ii) rAAV6-편집된 또는 (iii) ssODN-편집된 세포 (GTC (E6V) 공여자 구축물)로 이식된 NBSGW 마우스의 BM으로부터 CD34+ 및 CD34+CD38lo 구획의 대표적인 유동 플롯을 도시한다. 게이팅 전략: 살아있는, 단일 세포, hCD45+ > CD34+ CD38+.

도 14o는 GTC (E6V) rAAV6 또는 ssODN 전달을 사용한 CD34+ mPBSC의 유전자 편집 후 적혈구계 배양에서 β -글로빈 하위유형을 측정하기 위한 RP-HPLC 분석을 도시한다.

도 14p는 수확 후 2주 동안 적혈구계 배양 조건에서 생체외에서 확장된, 모의 (n = 2), GTC (E6V)-편집된 rAAV6 (n = 4) 또는 ssODN (n = 3) 변형된 세포로 이식된 수용자 마우스로부터 12-14주에 단리된 BM 세포를 도시한다. 발현된 글로빈 하위유형을 측정하기 위해 RP-HPLC 분석을 수행하였다 (α = 알파, β A = 성인, β S = 겸상, γ G = 감마 2, γ A = 감마 1).

도 14q는 유전자 편집 후 발현된 글로빈 사량체를 결정하기 위한 단일 BFU-E 콜로니 (메토컬트(methocult) 배양물로부터 생성됨)의 이온 교환 HPLC를 도시한다.

도 14r은 유전자 편집된 세포에서 발현된 글로빈 사량체를 측정하기 위한 단일 BFU-E 콜로니의 이온 교환 HPLC의 요약을 도시한다 (HbF: 태아, HbA: 성인, HbA2: 미성년자, HbS: 겸상). 모든 막대 그래프는 평균 $\pm$ SD를 나타낸다. n은 개별 동물의 수를 나타낸다. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001, p-값은 던넛의 다중 비교와 이원 ANOVA에 의해 각 샘플 평균을 모의 또는 대조군 샘플의 각 샘플 평균과 비교함으로써 계산하였다.

도 14s는 편집되고 분화된 적혈구계 콜로니의 IEC 분석을 도시한다. 글로빈 사량체를 측정하기 위한 이식체 4로부터의 단일 BFU-E 콜로니의 이온 교환 HPLC의 예 (HbF: 태아, HbA: 성인, HbA2: 미성년자, HbS: 겸상). 아래쪽 트레이스는 생착된, GTC (E6V) ssODN-편집된 HSC로부터 유래된 단일 콜로니에서 겸상 글로빈 발현 (적색 화살표)을 나타낸다. 수직 숫자는 HPLC 용출 시간이다.

도 14t는 편집되고 분화된 적혈구계 콜로니의 IEC 분석을 도시한다. 글로빈 사량체를 결정하기 위한 이식체 3으로부터의 단일 BFU-E 콜로니의 이온 교환 HPLC (HbF: 태아, HbA: 성인, HbA2: 미성년자, HbS: 겸상). 아래쪽 트레이스는 생착된 GTC (E6V) ssODN-편집된 HSC로부터 유래된 콜로니에서 겸상 글로빈 발현 (적색 화살표)을 나타낸다. 수직 숫자는 HPLC 용출 시간이다.

도 15a는 SCGM 또는 SFEM-II 배지를 사용하여 CM149 또는 ER100 론자 뉴클레오펙션 방법에 대해, 세포 계수에 의한 편집 후 생존율 백분율, ddPCR에 의한 % HDR, ICE에 의한 % NHEJ 및 유동 세포계측법에 의해 측정된 LT-HSC 구획 (게이팅 전략: 살아있는, 단일 세포, CD34⁺ CD38^{lo} > CD90⁺ CD133⁺)의 그래프를 도시한다.

도 15b는 SCGM 또는 SFEM-II 배지를 사용하여 CM149 또는 ER100 론자 뉴클레오펙션 방법에 대해, rAAV6 또는 ssODN 공여자 주형을 사용한 편집 및 HDR 후 베타-유사 글로빈 백분율의 그래프를 도시한다.

도 16a는 뉴클레오펙션시에 다양한 세포 밀도에서 SFEM-II 배지를 사용하여 CM149 또는 ER100 론자 뉴클레오펙션 방법에 대해 생존율 백분율에 대한 그래프를 도시한다.

도 16b는 뉴클레오펙션시에 다양한 세포 밀도에서 SFEM-II 배지를 사용하여 CM149 또는 ER100 론자 뉴클레오펙

선 방법에 대해 HDR에 대한 그래프를 도시한다.

도 16c는 뉴클레오펙션시에 다양한 세포 밀도에서 SFEM-II 배지를 사용하여 CM149 또는 ER100 론자 뉴클레오펙션 방법에 대해 NHEJ에 대한 그래프를 도시한다.

도 17a는 다양한 론자 뉴클레오펙션 프로토콜에 대해 EP, RNP, 또는 RNP 및 ssODN으로 처리된 세포에 대한 편집 후 2-14차에 대한 생존율 백분율의 그래프를 도시한다.

도 17b는 다양한 론자 뉴클레오펙션 프로토콜에 대해 EP, RNP, 또는 RNP 및 ssODN으로 처리된 세포에 대한 HDR, NHEJ 및 베타-유사 글로빈 발현 백분율의 그래프를 도시한다.

도 17c는 다양한 론자 뉴클레오펙션 프로토콜을 거친 세포에 대한 생존율 및 HDR의 비교를 도시한다.

도 18은 DU100 또는 CX100 론자 뉴클레오펙션 프로토콜을 거친 세포에 대한 생존율, HDR 및 NHEJ 백분율의 그래프를 도시한다.

도 19는 E6V (GTC) 변화 및 EoptE (GAA) 변화 둘 모두에 대한 대표적인 모의, AAV, RNP, RNP + AAV 및 RNP + ssODN 샘플에 대한 ddPCR 검정 결과를 나타낸다.

도 20a는 ddPCR 데이터로부터 결정된 HDR 백분율의 그래프를 도시한다.

도 20b는 ICE 알고리즘 데이터로부터 결정된 HDR 백분율의 그래프를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0033] <상세한 설명>

[0034] 본원에 개시된 조성물 및 방법의 일부 실시양태는 헤모글로빈-관련 돌연변이의 편집에 관한 것이다. 이러한 일부 실시양태는 겸상 적혈구 돌연변이의 부위에서 포스포디에스테르 DNA 가닥 파단의 도입을 통한 겸상 적혈구 돌연변이의 게내 편집을 포함한다.

[0035] 겸상 적혈구 질환은 성인 글로빈 (β A)의 소수성을 증가시키고 중합에 민감하게 만드는 단일 뉴클레오티드 전환에 의해 유발된다. SCD를 갖는 환자는 이환율이 증가되고 수명이 단축되어 수혈에 의존하는 경우가 많다. 건강한 공여자로부터의 HLA-일치 동종이계 이식을 통해 치유 치료를 달성할 수 있지만, HLA-일치 공여자의 이용가능성은 제한되며, 이식편-대-숙주 질환 (GvHD)의 가능성 및 고강도 골수절제 조절에 따른 단기간 및 장기간 영향에 의해 결과가 복잡해진다. 자가 줄기 세포에서 유전자 편집은 HLA-일치 공여자 이용가능성의 한계를 우회하고, 자기-재생 줄기 세포에서 질환-유발 돌연변이를 직접 보정할 수 있다. 또한, 성공적인 표적화된-유전자 편집의 확률은 초기 바이러스 벡터에 의해 제기되는 무작위 통합의 역사적 위험을 완화할 것이다.

[0036] 유전자 편집은 세포 DNA 복구 기구에 의해 이음매 없는 복구, 오류를 범하기 쉬운(error-prone) 비상동성 말단 연결 (NHEJ), 또는 DNA 공여자 주형의 존재하에 정확한 HDR로서 분해되는 이중-가닥 파단 (DSB)을 생성하는 부위-특이적 엔도뉴클레아제를 포함한다. 이들 복구 결과는 세포 주기 단계에 의해 현저하게 영향을 받는다. G0/G1 단계의 휴지 세포에서의 DSB는 주로 NHEJ로서 분해되는 반면, HDR에 의한 분해는 S/G2 단계로의 진입을 필요로 한다. 이들 복구 결과는 상호 배타적이므로, 개별 HSC 내에 및 HSC 집단에 걸쳐 전체 결과에 대해 경쟁한다.

[0037] HBB 유전자의 엑손 1에서 SCD 단일 뉴클레오티드 돌연변이는 인테그라제-결손 렌티바이러스 벡터 (IDLV), rAAV6 및 ssODN을 포함하는 DNA 복구 주형의 공동전달을 위한 여러 대안적인 방법과 조합하여 아연 핑거 뉴클레아제 (ZFN) mRNA, TALEN 및 CRISPR/Cas9를 포함하는 디자인어 뉴클레아제를 사용하는 상동성-지정 복구에 의해 보정될 수 있다 (문헌 [Hoban, MD, et al. (2015). Blood 125: 2597-2604]; [DeWitt, MA, et al (2017) Methods 121-122: 9-15]; [Dever, DP, et al. (2016) Nature 539: 384-389]; 및 [Hoban, MD, et al. (2016) Mol Ther 24: 1561-1569.].17-20], 이들은 각각 그 전문이 본원에 참조로 명시적으로 포함됨). 이들 접근법 중, rAAV6 또는 단일-가닥 올리고데옥시뉴클레오티드 (ssODN)를 사용하는 것이 가장 효율적인 공여자 주형 전달 플랫폼을 구성한다. 그러나, 정확한 HDR vs. NHEJ의 빈도를 포함한 총 편집 결과는 rAAV6 및 ssODN 공여자 주형 전달 방법에 대해 동시에 비교되지 않았다. 또한, 임상적으로 관련성이 있고 치료적이기 위해서는, 고충실도 HDR 결과가 DSB를 부적절하게 복구하고 게놈 불안정성을 유발하는 오류를 범하기 쉬운(error prone) NHEJ를 비례적으로 초과해야 한다.

[0038] (i) HDR 및 NHEJ 결과의 비율; (ii) 편집 후 HSC 구획의 완전성 및 장기간 생착 잠재력의 보존; 및 (iii) 편집

된 세포의 종방향 지속성의 변경에서 공여자 주형 전달의 역할을 더 잘 이해하기 위해, 성인 CD34+ mPBSC에서 시험관내 및 생체내 다양한 공여자 주형 전달 방법이 본원에 개시된 바와 같이 평가되었다. 이들 연구는 건강한 공여자 mPBSC에 도입된 겸상 돌연변이 (GTC 또는 GTG; 글루타메이트에서 발린으로의 변화를 코딩함 (E6V) 또는 침묵 변화 (GAA; 글루타메이트에서 글루타메이트로 코딩함 (E6optE))를 포함하였다. HBB의 엑손 1의 RNP-매개 파괴 및 대안적인 공여자 주형 전달 후, 유전자 편집 결과는 ddPCR을 통한 분자 분석 및 겸상 글로빈의 유도 (β S; GTC 또는 GTG의 경우; E6V 변화) 또는 성인 글로빈의 복원 (GAA 변화의 경우; E6optE)을 통한 글로빈 발현을 기능적 결과로서 사용하여 평가하였다. 이들 접근법을 사용하여, 대안적인 전달 플랫폼의 결과를 직접 비교하였다. 시험관내 연구는 NHEJ보다 비례적으로 더 큰 HDR을 유발하는 rAAV6 전달에 대한 우월성을 입증한 반면, ssODN 공여자 주형 전달은 HDR보다 유의하게 더 많은 NHEJ를 도입하였다. 동시에, GTC 변화 (E6V)를 함유하는 이식된 HDR-편집된 HSC 세포의 생착 및 지속성의 종방향 평가를 수행하였다. 시험관내 발견과는 대조적으로, ssODN 공여자 주형에 의해 변형된 훨씬 더 큰 백분율의 세포가 NBSGW 수용자 마우스의 골수 (BM)에서 12-14 주에 지속되었다. 종합하면, 발견은 HDR-기반 유전자 편집을 위한 대안적인 방법의 중요한 기능적 평가를 제공한다.

[0039] 본원에 제공된 유전자-편집 시스템 및 방법은 제한 없이 유전자 파괴 및/또는 유전자 표적화를 포함하는 임의의 뉴클레아제-기반 유전자 편집 접근법에 적용될 수 있다. 예를 들어, 본 개시내용의 측면은 CRISPR/Cas9-기반 유전자 편집에 관한 것이다. 일부 대안에서, 유전자 편집의 Cas9 뉴클레아제-매개 향상이 제공된다. 일부 실시양태에서, 뉴클레아제-기반 유전자 편집 시스템 및 방법이 제공된다. 유전자 편집을 위한 뉴클레아제-기반 접근법의 예는 뉴클레아제, 예컨대, 제한 없이, ZFN, TALEN, 메가뉴클레아제 (예를 들어, MegaTAL) 또는 CRISPR/Cas9를 포함하는 시스템을 포함한다.

[0040] 일부 대안에서, 뉴클레아제는 게놈에서 원하는 위치에 특이적 이중 가닥 파단을 도입함으로써 표적화된 게놈 변형을 수행하고, 세포 복구 메커니즘을 활용하여 상동성 재조합 및 비상동성 말단-연결 메커니즘에 의해 유도된 파단을 복구한다. 여러 조작된 뉴클레아제를 사용할 수 있다. 제한이 아닌 예로서, 뉴클레아제는 아연 핑거 뉴클레아제 (ZFN), 전사 활성인자-유사 이펙터 뉴클레아제 (TALEN), CRISPR/Cas 시스템, RNA 가이드된 엔도뉴클레아제 또는 조작된 메가뉴클레아제 재조합된 귀소 엔도뉴클레아제를 포함할 수 있다. 표적화된 유전자 파괴는 연구, 치료, 농업 및 산업 용도에 광범위하게 적용된다. 표적화된 유전자 파괴를 생성하기 위한 한 가지 전략은 부위-특이적 엔도뉴클레아제에 의해 유발된 이중-가닥 DNA 파단의 생성을 통한 것이다.

[0041] 일부 대안에서, CRISPR/Cas9는 다양한 세포 유형에서 효율적으로 가이드 RNA의 발현을 가능하게 한다. 가이드 RNA를 발현하기 위한 시스템의 예는 아데노-연관 바이러스 벡터 (AAV)의 사용을 기반으로 한다. AAV 벡터는 광범위한 1차 세포를 형질도입할 수 있다.

[0042] 일부 대안에서, Cas9-기반 접근법은 아데노-연관 바이러스 벡터 (AAV)가 Cas9 표적화에 필요한 가이드 RNA를 발현하는데 사용되는 경우 최소 독성으로 유전자 편집 효율을 향상시킨다.

[0043] <정의>

[0044] 본원에 사용된 바와 같은 "하나"는 하나 또는 하나 초과를 의미할 수 있다.

[0045] 본원에 사용된 바와 같은 용어 "약"은 값이 값을 결정하는데 사용되는 방법에 대한 오차의 내재 변동, 또는 실험 간에 존재하는 변동을 포함함을 나타낸다.

[0046] "핵산" 또는 "핵산 분자"는 폴리뉴클레오티드, 예컨대 데옥시리보핵산 (DNA) 또는 리보핵산 (RNA), 올리고뉴클레오티드, 중합효소 연쇄 반응 (PCR)에 의해 생성된 단편, 또는 라이게이션, 절단, 엔도뉴클레아제 작용 및 엑소뉴클레아제 작용 중 임의의 것에 의해 생성된 단편을 지칭한다. 핵산 분자는 자연-발생 뉴클레오티드 (예컨대 DNA 또는 RNA), 또는 자연-발생 뉴클레오티드의 유사체 (예를 들어, 자연-발생 뉴클레오티드의 거울상이성질체 형태), 또는 둘 모두의 조합인 단량체로 구성될 수 있다. 변형된 뉴클레오티드는 당 모이어티 및/또는 피리미딘 또는 퓨린 염기 모이어티의 변경을 가질 수 있다. 당 변형은 예를 들어, 하나 이상의 히드록실기를 할로젠, 알킬기, 아민 또는 아지도기로 대체하는 것을 포함하거나, 당은 에테르 또는 에스테르로 관능화될 수 있다. 더욱이, 전체 당 모이어티는 입체적이고 전자적으로 유사한 구조, 예컨대 아자-당 및/또는 카보시클릭 당 유사체로 대체될 수 있다. 염기 모이어티의 변형의 예는 알킬화된 퓨린 또는 피리미딘, 아실화된 퓨린 또는 피리미딘, 또는 다른 널리 공지된 헤테로시클릭 치환체를 포함한다. 핵산 단량체는 포스포디에스테르 결합 또는 이러한 연결의 유사체에 의해 연결될 수 있다. 포스포디에스테르 연결의 유사체는 포스포로티오에이트, 포스포로디티오에이트, 포스포셀레노에이트, 포스포로디셀레노에이트, 포스포로아닐로티오에이트, 포스포아닐리데

이트 또는 포스포르아미데이트를 포함한다. 용어 "핵산 분자"는 또한 소위 "펩티드 핵산"을 포함하며, 이는 폴리아미드 백본에 부착된 자연-발생 또는 변형된 핵산 염기를 포함한다. 핵산은 단일 가닥 또는 이중 가닥일 수 있다.

[0047] "코딩하는(coding for)" 또는 "코딩하는(encoding)"이 본원에 사용되며, 다른 거대분자의 합성을 위한 주형으로 작용하는 폴리뉴클레오타이드, 예컨대 유전자, cDNA 또는 mRNA에서 뉴클레오타이드의 특이적 서열, 예컨대 정의된 아미노산 서열의 특성을 지칭한다. 그러므로, 해당 유전자에 상응하는 mRNA의 전사 및 번역이 세포 또는 다른 생물학적 시스템에서 단백질을 생성하는 경우 유전자는 단백질을 코딩한다.

[0048] 일부 대안에서, CRISPR/Cas9 시스템의 기본 구성분은 표적 유전자, 가이드 RNA, 및 Cas9 엔도뉴클레아제, 그 의 유도체 또는 단편을 포함한다. 일부 대안에서, 유전자 편집을 위해 CRISPR/Cas9를 적용하는 한 측면은 가이드 RNA를 다양한 세포 유형에 효율적으로 전달하기 위한 시스템에 대한 필요성이다. 이는 예를 들어 핵산으로서 시험관내 생성된 가이드 RNA의 전달을 포함할 수 있다 (시험관내 전사 또는 화학적 합성에 의해 생성된 가이드 RNA). 일부 대안에서 가이드 RNA를 코딩하는 핵산은 변형된 염기, 예컨대 2'-O-메틸 염기의 혼입에 의해 뉴클레아제 내성이 된다.

[0049] 본원에 기재된 변형된 염기 중 하나 이상을 함유할 수 있는, 본원에 기재된 대안에 유용한 예시적인 가이드 RNA는 서열식별번호: 1, 서열식별번호: 2, 서열식별번호: 3, 서열식별번호: 4, 서열식별번호: 5 또는 서열식별번호: 6에 의해 코딩된 서열로 제공된다. 일부 대안에서, 가이드 RNA를 발현하기 위한 중요한 시스템은 AAV 벡터가 광범위한 1차 세포를 형질도입할 수 있기 때문에 아데노-연관 바이러스 (AAV) 벡터의 사용을 기반으로 한다. 일부 대안에서, AAV 벡터는 감염을 유발하지 않으며 게놈에 통합하는 것으로 공지되지 않았다. 따라서, 일부 대안에서, AAV 벡터의 사용은 안전하고 효능있다는 이점을 갖는다.

[0050] 용어 "상보적"은 상보적 서열이 참조 폴리뉴클레오타이드 서열의 전체 또는 하나 이상의 부분과 상동성이 있음을 의미한다. 예시를 위해, 뉴클레오타이드 서열 "CATTAG"는 참조 서열 "CATTAG"에 상응하고, 참조 서열 "GTAATC"에 상보적이다.

[0051] "프로모터"는 구조적 유전자의 전사를 지시하는 뉴클레오타이드 서열이다. 일부 대안에서, 프로모터는 구조적 유전자의 전사 시작 부위에 근접한 유전자의 5' 비-코딩 영역에 위치한다. 전사 개시에서 기능하는 프로모터 내의 서열 요소는 종종 컨센서스 뉴클레오타이드 서열을 특징으로 한다. 이들 프로모터 요소는 RNA 폴리머라제 결합 부위, TATA 서열, CAAT 서열, 분화-특이적 요소 (DSE; McGehee et al., Mol. Endocrinol. 7:551 (1993); 그 전문이 본원에 참조로 명시적으로 포함됨), 시클릭 AMP 반응 요소 (CRE), 혈청 반응 요소 (SRE; Treisman, Seminars in Cancer Biol. 1:47 (1990); 그 전문이 본원에 참조로 명시적으로 포함됨), 글루코코르티코이드 반응 요소 (GRE), 및 다른 전사 인자에 대한 결합 부위, 예컨대 CRE/ATF (O'Reilly et al., J. Biol. Chem. 267:19938 (1992)), AP2 (Ye et al., J. Biol. Chem. 269:25728 (1994)), SP1, cAMP 반응 요소 결합 단백질 (CREB; Loeken, Gene Expr. 3:253 (1993)) 및 팔랑체 인자 (일반적으로, 문헌 [Watson et al., eds., Molecular Biology of the Gene, 4th ed. (The Benjamin/Cummings Publishing Company, Inc. 1987)], 및 [Lemaigre and Rousseau, Biochem. J. 303:1 (1994)] 참조; 모든 참고문헌은 그 전문이 본원에 참조로 명시적으로 포함됨)를 포함한다. 본원에 사용된 바와 같이, 프로모터는 구성적으로 활성, 억압성 또는 유도성일 수 있다. 프로모터가 유도성 프로모터인 경우, 유도제에 대한 반응으로 전사 속도가 증가한다. 대조적으로, 프로모터가 구성적 프로모터인 경우 전사 속도는 유도제에 의해 조절되지 않는다. 억압성 프로모터가 또한 공지되어 있다. 일부 대안에서, 조절성 요소는 비번역된 영역일 수 있다. 일부 대안에서, 비번역된 영역은 5' 비번역된 영역이다. 일부 대안에서, 비번역된 영역은 3' 비번역된 영역이다. 일부 대안에서, 5' 또는 3' 비번역된 영역이 사용된다. 일부 대안에서, 5' 및 3' 비번역된 영역 둘 모두가 사용된다. 관련 기술분야의 통상의 기술자는 본원의 대안에서 사용되는 바와 같은 비번역된 영역의 의미를 이해할 것이다.

[0052] "조절성 요소"는 코어 프로모터의 활성을 조정하는 뉴클레오타이드 서열이다. 예를 들어, 조절성 요소는 특정 세포, 조직, 또는 소기관에서 배타적으로 또는 우선적으로 전사를 가능하게 하는 세포 인자와 결합하는 뉴클레오타이드 서열을 함유할 수 있다. 이들 유형의 조절성 요소는 "세포-특이적", "조직-특이적" 또는 "소기관-특이적" 방식으로 발현되는 유전자와 정상적으로 연관된다. 일부 대안에서, 세포에서 적어도 하나의 표적 유전자를 편집하기 위한 시스템이 제공되며, 여기서 시스템은 CRISPR 가이드 RNA를 코딩하는 제1 핵산 서열을 포함하며, 여기서 CRISPR 가이드 RNA는 세포에서 적어도 하나의 표적 유전자에 상보적이고, 여기서 상기 제1 핵산 서열은 벡터에 존재하고; 상기 시스템은 또한 Cas9 단백질을 코딩하는 제2 핵산 서열, 제1 아데노바이러스 단백질을 코딩하는 제3 핵산 서열, 및 제2 아데노바이러스 단백질을 코딩하는 제4 핵산 서열을 포함한다. 일부 대안에서, 제

1, 제2, 제3 및 제4 핵산 서열은 진핵생물 세포, 예컨대 인간 세포에서 작동가능한 조절성 요소에 연결된다.

[0053] "폴리펩티드"는 자연적으로 또는 합성적으로 생성되는지 여부에 관계없이 펩티드 결합에 의해 연결된 아미노산 잔기의 중합체이다. 약 10개 미만의 아미노산 잔기의 폴리펩티드는 일반적으로 "펩티드"로 지칭된다. 폴리펩티드는 단백질로서 간주될 수 있다.

[0054] "단백질"은 하나 이상의 폴리펩티드 쇄를 포함하는 거대분자이다. 단백질은 또한 비-펩티드 구성성분, 예컨대 탄수화물 기를 포함할 수 있다. 탄수화물 및 다른 비-펩티드 치환체는 단백질이 생성되는 세포에 의해 단백질에 첨가될 수 있으며, 세포 유형에 따라 달라질 것이다. 단백질은 그의 아미노산 백본 구조의 관점에서 본원에서 정의되며; 치환체, 예컨대 탄수화물 기는 일반적으로 특정되지 않지만, 그럼에도 불구하고 존재할 수 있다. 일부 대안에서, 세포에서 적어도 하나의 표적 유전자를 편집하기 위한 시스템이 제공되며, 여기서 방법은 CRISPR 가이드 RNA를 코딩하는 제1 핵산 서열을 포함하며, 여기서 CRISPR 가이드 RNA는 세포에서 적어도 하나의 표적 유전자에 상보적이고, 여기서 상기 제1 핵산 서열은 백터에 존재하고; 상기 시스템은 또한 Cas9 단백질을 코딩하는 제2 핵산 서열, 제1 아데노바이러스 단백질을 코딩하는 제3 핵산 서열, 및 제2 아데노바이러스 단백질을 코딩하는 제4 핵산 서열을 포함한다.

[0055] 용어 "숙주 세포"는 본 대안에 따른 Cas9-mRNA/AAV-가이드 RNA가 도입된 세포, 뿐만 아니라 본원의 시스템이 제공된 세포를 의미한다. 숙주 세포는 원핵생물 세포 또는 진핵생물 세포일 수 있다. 원핵생물 숙주 세포의 예는 이. 콜라이(*E. coli*), 질소 고정 박테리아, 스탕필로코쿠스 아우레우스(*Staphylococcus aureus*), 스탕필로코쿠스 알부스(*Staphylococcus albus*), 락토바실러스 아시도필루스(*Lactobacillus acidophilus*), 바실러스 안트라시스(*Bacillus anthracis*), 바실러스 서브틸리스(*Bacillus subtilis*), 바실러스 투링기엔시스(*Bacillus thuringiensis*), 클로스트리디움 테타니(*Clostridium tetani*), 클로스트리디움 보툴리눔(*Clostridium botulinum*), 스트렙토코쿠스 뮤탄스(*Streptococcus mutans*), 스트렙토코쿠스 뉴모니아에(*Streptococcus pneumoniae*), 마이코플라스마(*mycoplasma*) 또는 시아노박테리아(*cyanobacteria*)를 포함하나 이에 제한되지는 않는다. 진핵생물 숙주 세포의 예는 원생동물, 진균, 조류, 식물, 곤충, 양서류, 새 및/또는 포유동물 세포를 포함하나 이에 제한되지는 않는다. 일부 대안에서, 세포에서 적어도 하나의 표적 유전자를 편집하기 위한 시스템이 제공되며, 여기서 세포는 진핵생물 세포이다. 일부 대안에서, 세포는 포유동물 세포이다. 일부 대안에서, 세포는 인간 세포이다. 일부 대안에서, 세포는 1차 세포이다. 일부 대안에서, 세포는 형질전환된 세포가 아니다. 일부 대안에서, 세포는 1차 림프구이다. 일부 대안에서, 세포는 1차 림프구, CD34+ 줄기 세포, 간세포, 심근세포, 뉴런, 아교 세포, 근육 세포 또는 장 세포이다.

[0056] 용어 "엔도뉴클레아제"는 폴리뉴클레오티드 쇄 내의 포스포디에스테르 결합을 절단하는 효소를 지칭한다. 폴리뉴클레오티드는 이중-가닥 DNA (dsDNA), 단일-가닥 DNA (ssDNA), RNA, DNA 및/또는 RNA의 이중-가닥 하이브리드, 및/또는 합성 DNA (예를 들어, A, C, G 및 T 이외의 염기를 함유함)일 수 있다. 엔도뉴클레아제는 폴리뉴클레오티드를 대칭적으로 절단하여, "평활" 말단을 남기거나, 직접 대향하지 않는 위치에서 "점착성 말단"으로 지칭될 수 있는 오버행을 생성할 수 있다. 본원에 기재된 방법 및 조성물은 엔도뉴클레아제에 의해 생성된 절단 부위에 적용될 수 있다. 시스템의 일부 대안에서, 시스템은 엔도뉴클레아제, 예컨대 Cas9, TALEN 또는 MegaTAL, 또는 엔도뉴클레아제의 도메인, 예를 들어, Cas9, TALEN 또는 MegaTAL, 또는 그의 하나 이상의 부분을 포함하는 융합 단백질을 코딩하는 핵산을 추가로 제공할 수 있다. 이들 예는 제한적인 것으로 의도되지 않으며, 다른 엔도뉴클레아제 및 다른 엔도뉴클레아제를 포함하는 시스템 및 방법의 대안 및 이들 예시적인 대안의 변동 및 변형이 과도한 실험 없이 가능하다. 이러한 모든 변동 및 변형은 본 개시내용의 범위 내에 있다.

[0057] 용어 "전사 활성인자-유사 이펙터 뉴클레아제" 또는 "TAL 이펙터 뉴클레아제" (TALEN)는 뉴클레아제 도메인에 융합된 TAL-이펙터 도메인을 포함하는 뉴클레아제를 지칭한다. 식물 병원체 산토모나스(*Xanthomonas*)로부터 단리된 TAL-이펙터 DNA 결합 도메인이 설명되었다 (문헌 [Boch et al., (2009) Science 29 Oct. 2009 (10.1126/science.117881)] 및 [Moscou and Bogdanove, (2009) Science 29 Oct. 2009 (10.1126/science.1178817)] 참조; 두 참고문헌은 그 전문이 본원에 참조로 명시적으로 포함됨). 이들 DNA 결합 도메인은 원하는 표적에 결합하도록 조작되고 뉴클레아제 도메인, 예컨대 FokI 뉴클레아제 도메인에 융합되어 TAL 이펙터 도메인-뉴클레아제 융합 단백질을 유도할 수 있다. 본원에 기재된 방법 및 시스템은 TAL 이펙터 뉴클레아제에 의해 생성된 절단 부위에 적용될 수 있다. 본원에 제공된 시스템의 일부 대안에서, 시스템은 TALEN 뉴클레아제 또는 TALEN 뉴클레아제를 코딩하는 백터 또는 핵산을 추가로 포함할 수 있다. 본원에 제공된 방법의 일부 대안에서, 방법은 뉴클레아제, 예컨대 TALEN 뉴클레아제를 제공하는 것을 추가로 포함할 수 있다.

[0058] 일부 대안에서, TALEN은 Tal 이펙터 DNA 결합 도메인을 DNA 절단 도메인에 융합함으로써 생성된 인공 제한 효소

이다. Tal 이펙터는 높은 친화성으로 하나의 뉴클레오티드에 결합할 수 있는 고도로 상동성인 34개 아미노산 모듈로 이루어진 박테리아 DNA-결합 단백질일 수 있다. 반복-가변 디-뉴클레오티드로 지칭되는 TALEN 모듈의 가변 제12 및 제13 아미노산은 염기 특이성을 부여하고 (즉, NN→G/A, NI→A, NG→T, NK→G, HD→C, 및 NS→A/T/C/G), 뉴클레오티드 서열을 표적화할 수 있는 TALEN 어레이는 개별 모듈을 어셈블리함으로써 생성될 수 있다. 아미노산 서열 및 DNA 인식 사이의 관계는 상관 반복 가변 이잔기(Repeat Variable Di-residue)(RVD)와 접촉하는 반복 세그먼트의 조합을 선택함으로써 특이적 DNA 결합 도메인의 조작용을 허용하였다. TALEN은 관심있는 세포에서 이중-가닥 파단 (DSB)을 유도함으로써 게놈을 편집하는데 사용될 수 있으며, 여기서 세포는 여러 유형의 복구 메커니즘으로 반응할 수 있다.

[0059] MegaTAL은 2개의 별개의 부류의 DNA 표적화 효소의 조합으로부터 유래된다. 메가뉴클레아제 (귀소 엔도뉴클레아제라고도 함)는 동일한 도메인에서 DNA 인식 및 뉴클레아제 기능 둘 모두의 이점을 갖는 단일 펩티드 쇄이다. 본원에 제공된 시스템의 일부 대안에서, 시스템은 MegaTAL 뉴클레아제 또는 MegaTAL 뉴클레아제를 코딩하는 벡터 또는 핵산을 추가로 포함할 수 있다. 본원에 제공된 방법의 일부 대안에서, 방법은 MegaTAL 뉴클레아제 또는 MegaTAL 뉴클레아제를 코딩하는 벡터 또는 핵산을 제공하는 것을 추가로 포함할 수 있다.

[0060] 아연 핑거 단백질 (ZFP)은 진핵생물 DNA 결합 단백질이다. 예를 들어, 게놈 편집을 위한 가장 일반적인 ZFP 모티프는 Cys2-His2 핑거이며, 각 유형은 뉴클레오티드 트리플렛에 특이적이다. 인공 ZFP 도메인은 개별 아연 핑거의 어셈블리에 의해 통상적으로 9-18nt 길이인 특이적 DNA 서열을 표적화하도록 생성될 수 있다. 아연 핑거 뉴클레아제 (ZFN)는 인간에서 다양한 세포 유형에서 표적화된 게놈 조작용을 수행하기 위한 강력한 도구이다. ZFN은 비특이적 엔도뉴클레아제 도메인에 연결된 조작된 DNA-결합 아연 핑거 도메인으로 이루어지고, 상동성 및 비상동성 재조합 둘 모두를 자극하는 이중-가닥 파단 (DSB)을 도입할 수 있으며, 그 후 이는 게놈 조작용을 수행하도록 활용될 수 있다. 이와 같이, ZFP는 연구 및 유전자 요법 적용 둘 모두에서 잠재력을 갖는다.

[0061] 클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR)는 염기 서열의 짧은 반복부를 함유할 수 있는 DNA 유전자좌를 포함하며, 여기서 각 반복 다음에 바이러스 노출로부터 스페이서 DNA의 짧은 세그먼트가 이어진다. CRISPR 영역은 CRISPR과 관련된 단백질을 코딩하는 cas 유전자와 연관될 수 있다. CRISPR/Cas 시스템은 외래 유전적 요소, 예컨대 플라스미드 및 파지에 대한 내성을 부여하고 후천 면역 형태를 제공하는 원핵생물 면역 체계이다. CRISPR 스페이서는 진핵생물 유기체에서 RNAi와 유사한 방식으로 이들 외인성 유전적 요소를 인식하고 절단한다. 게놈 편집 메커니즘으로서, RNA 가이드된 엔도뉴클레아제, Cas 단백질 및 적절한 가이드 RNA는 세포로 전달될 수 있으며, 유기체의 게놈은 원하는 위치에서 절단될 수 있다. CRISPR은 유전자를 표적화/변형시키기 위한 효율적인 메커니즘이며, 메커니즘은 관련 기술분야의 통상의 기술자에게 공지되어 있다.

[0062] Cas9 (CRISPR 연관 단백질 9)는 다른 박테리아 중에서 스트렙토코쿠스 피오게네스(*Streptococcus pyogenes*)에서 CRISPR (클러스터링된 규칙적 산재성 회문식 반복부) 적응 면역 체계와 연관된 RNA-가이드된 DNA 엔도뉴클레아제 효소이다. 에스. 피오게네스는 Cas9를 사용하여, 외래 DNA, 예컨대 침입 박테리오파지 DNA 또는 플라스미드 DNA를 기억하고 나중에 심문하고 절단한다. Cas9는 외래 DNA를 풀고 가이드 RNA의 20개 염기 쌍 스페이서 영역에 상보적인지 확인하여 이 심문을 수행한다. DNA 기질이 가이드 RNA에 상보적인 경우, Cas9는 침입 DNA를 절단한다.

[0063] CRISPR (클러스터링된 규칙적 간격의 짧은 회문식 반복부)은 염기 서열의 짧은 반복부를 함유하는 원핵생물 DNA의 세그먼트이다. 각 반복 다음에 박테리아 바이러스 또는 플라스미드에 대한 이전 노출로부터 "스페이서 DNA"의 짧은 세그먼트가 이어진다. CRISPR/Cas 시스템은 생명 나무 전체에 걸쳐 종에서 유전자 편집 (특정 유전자의 서열의 첨가, 파괴 또는 변화) 및 유전자 조절에 사용되어 왔다. Cas9 단백질, 그의 유도체 또는 단편 및 적절한 가이드 RNA를 세포에 전달함으로써, 유기체의 게놈은 임의의 원하는 위치에서 절단될 수 있다. CRISPR을 사용하여 전체 집단의 게놈을 변경할 수 있는 RNA-가이드된 유전자 드라이브를 구축할 수 있다. 일부 대안에서, 세포에서 적어도 하나의 표적 유전자를 편집하기 위한 시스템이 제공되며, 여기서 방법은 CRISPR 가이드 RNA를 코딩하는 제1 핵산 서열, Cas9 단백질, 그의 유도체 또는 단편을 코딩하는 제2 핵산 서열, 제1 아데노바이러스 단백질을 코딩하는 제3 핵산 서열 및 제2 아데노바이러스 단백질을 코딩하는 제4 핵산 서열을 포함하며, 여기서 CRISPR 가이드 RNA는 세포에서 적어도 하나의 표적 유전자에 상보적이고, 여기서 상기 제1 핵산 서열은 벡터에 존재한다.

[0064] 일부 대안에서, 화학적으로 변형된 가이드 RNA의 사용이 고려된다. 화학적으로-변형된 가이드 RNA는 인간 1차 세포에서 CRISPR-Cas 게놈 편집에 사용되어 왔다 (Hendel, A. et al., Nat Biotechnol. 2015 Sep; 33(9):985-9). 가이드 RNA의 화학적 변형은 뉴클레아제 내성을 부여하는 변형을 포함할 수 있다. 뉴클레아제는 엔도뉴클

레아제, 또는 엑소뉴클레아제, 또는 둘 모두일 수 있다. 일부 화학적 변형은 제한 없이 2'-플루오로, 2'-O-메틸, 포스포로티오에이트 디티올 3'-3' 말단 연결, 2-아미노-dA, 5-메틸-dC, C-5 프로피닐-C, 또는 C-5 프로피닐-U, 모르폴리노를 포함한다. 이들 예는 제한적인 것으로 의도되지 않으며, 다른 화학적 변형 및 이들 예 시적인 대안의 변동 및 변형이 또한 고려된다.

[0065] 용어 "절단"은 폴리뉴클레오타이드의 공유결합 백본의 파손을 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 모두가 가능하고, 이중-가닥 절단은 2개의 별개의 단일-가닥 절단 사건의 결과로 발생할 수 있다. 이중 가닥 DNA, RNA, 또는 DNA/RNA 하이브리드 절단은 평할 말단 또는 엇갈린 말단의 생성을 유발할 수 있다.

[0066] 본원에 사용된 바와 같은 용어 "대상체"는 비인간 영장류 및 인간을 포함하는 동물계의 모든 구성원을 포함한다. 일부 대안에서, 세포에서 적어도 하나의 표적 유전자를 편집하기 위한 시스템이 제공되며, 여기서 방법은 CRISPR 가이드 RNA를 코딩하는 제1 핵산 서열, Cas9 단백질, 그의 유도체 또는 단편을 코딩하는 제2 핵산 서열, 제1 아데노바이러스 단백질을 코딩하는 제3 핵산 서열 및 제2 아데노바이러스 단백질을 코딩하는 제4 핵산 서열을 포함하며, 여기서 CRISPR 가이드 RNA는 세포에서 적어도 하나의 표적 유전자에 상보적이고, 여기서 상기 제1 핵산 서열은 벡터에 존재한다. 일부 대안에서, 편집된 유전자를 포함하는 세포는 필요로 하는 대상체에게 전달된다.

[0067] 회귀-절단 엔도뉴클레아제에 의해 도입된 표적화된 DNA 이중-가닥 파단은 비상동성 말단 봉합 DNA 복구 경로를 사용함으로써 다양한 세포 유형에서 유전자 파괴 적용에 활용될 수 있다. 그러나, 엔도뉴클레아제는 종종 정확한 복구가 필요한 화학적으로 깨끗한 파단을 생성하여, 표적화된 유전자 파괴의 효율을 제한한다. 본원에 기재된 여러 대안은 엔도뉴클레아제-유도된 부위-특이적 DNA 이중-가닥 파단의 부정확한 복구에 의해 야기되는 표적화된 유전자 파괴의 속도를 개선시키는 방법에 관한 것이다. 일부 대안에서, 시스템은 표적화된 유전자 파괴의 속도를 향상시키기 위해 말단-프로세싱 효소와 커플링된 부위 특이적 엔도뉴클레아제를 추가로 포함할 수 있다. 커플링은 예를 들어, 물리적, 공간적 및/또는 시간적일 수 있다.

[0068] 어떤 특정 이론에도 얽매이지 않지만, "오류를 범하기 쉬운" 비상동성 말단-연결 (NHEJ)에 의한 이중-가닥 DNA 파단의 분해는 표적화된 파괴 및 유전적 녹아웃을 생성하는데 활용될 수 있으며, 이는 NHEJ 프로세스가 파단 부위에 삽입 및 결실을 초래할 수 있기 때문이다. NHEJ는 여러 하위 경로에 의해 매개되며, 각각은 별개의 돌연변이 결과를 갖는다. 고전적인 NHEJ 경로 (cNHEJ)는 KU/DNA-PKcs/Lig4/XRCC4 복합체를 포함하며, 최소 프로세싱으로 말단을 함께 역으로 라이게이션한다. 디자인어 엔도뉴클레아제 플랫폼 (아연-핑거 뉴클레아제 (ZFN), TAL 이펙터 뉴클레아제 (TALEN) 및 귀소 엔도뉴클레아제 (HE))에 의해 생성된 DNA 파단은 모두 라이게이션 전에 프로세싱을 필요로 하지 않는 화학적으로 깨끗한 상용성 오버행 파단을 남기기 때문에, 이들은 cNHEJ 경로에 의한 정확한 복구를 위한 우수한 기질이다. 파단을 해결하기 위한 고전적인 NHEJ 경로의 부재 또는 실패시, 대안적인 NHEJ 경로 (altNHEJ)가 치환할 수 있으나: 이들 경로는 훨씬 더 돌연변이유발성이다.

[0069] 어떤 특정 이론에도 얽매이지 않지만, 말단-프로세싱 효소에 의한 DNA 이중-가닥 파단의 변형은 복구를 altNHEJ 경로를 향해 편향시킬 수 있다. 또한, 말단-프로세싱 효소의 상이한 서브세트는 상이한 메커니즘에 의해 파괴를 향상시킬 수 있다. 예를 들어, 3' 오버행에 노출된 포스포디에스테르 결합을 특이적으로 가수분해하는 엑소뉴클레아제인 Trex2는 파단 부위에서 복구를 돌연변이유발성 결실을 향해 편향시킨다. 대조적으로, 비-주형 폴리머라제인 말단 데옥시뉴클레오타이드 트랜스퍼라제 (TdT)는 라이게이션 전에 DNA 말단을 변경하기 위해 뉴클레오타이드 첨가의 첨가를 촉진함으로써 파단 부위에서 복구를 돌연변이유발성 삽입을 향해 편향시킬 것으로 예상된다. 따라서, 관련 기술분야의 통상의 기술자는 본원에 제공된 임의의 시스템 또는 방법으로 원하는 조작 결과를 제공하기 위해 상이한 활성을 갖는 말단-프로세싱 효소를 사용할 수 있다. 또한, 관련 기술분야의 통상의 기술자는 최대 또는 고유한 유형의 효과를 달성하기 위해 상이한 말단-프로세싱 효소 간의 상승작용을 사용할 수 있다.

[0070] 본원에 기재된 엔도뉴클레아제, 말단-프로세싱 효소 및 융합 단백질을 코딩하는 다양한 RNA 분자는 선택된 단백질 또는 펩티드를 세포에 제공하기 위해 구축될 수 있다. 엔도뉴클레아제, 말단-프로세싱 효소 및 융합 단백질을 코딩하는 RNA 분자는 관련 기술분야에 공지된 바와 같이 선택된 숙주 세포에서 발현을 최적화하기 위해 상이한 코돈을 함유하도록 변형될 수 있다. 일부 대안에서, RNA는 50, 100, 150, 200, 250, 300, 350, 400, 450 또는 500개의 공유결합으로 연결된 아데노신 잔기의 폴리(A) 테일, 또는 상기 언급된 값 중 임의의 2개에 의해 정의된 범위 내의 잔기의 양을 포함할 수 있다.

- [0071] 시스템의 여러 대안은 부위-특이적 엔도뉴클레아제 및 말단-프로세싱 효소의 동시 발현을 위한 벡터 또는 핵산을 추가로 포함하여, 표적화된 유전자 파괴의 효율을 최대 대략 70배만큼 개선시켜, 72시간 미만에 표적 부위를 함유하는 세포의 집단의 100%에서 돌연변이유발성 결과를 본질적으로 고정시킨다.
- [0072] 발현 벡터
- [0073] 일부 대안에서, 발현 구축물은 관련 기술분야에 공지된 방법을 사용하여 설계될 수 있다. 핵산 발현 벡터의 예는 재조합 바이러스, 렌티바이러스, 아데노바이러스, 플라스미드, 박테리아 인공 염색체, 효모 인공 염색체, 인간 인공 염색체, 미니서클 DNA, 에피솜, cDNA, RNA, 또는 PCR 생성물을 포함하나 이에 제한되지는 않는다. 일부 대안에서, 핵산 발현 벡터는 단일 펩티드 (예를 들어, 엔도뉴클레아제, 말단-프로세싱 효소, 또는 엔도뉴클레아제 및 말단-프로세싱 활성을 갖는 융합 단백질)를 코딩한다. 일부 대안에서, 핵산 발현 벡터는 단일 폴리시스트론 발현 카세트에서 하나 이상의 엔도뉴클레아제 및 하나 이상의 말단-프로세싱 효소를 코딩한다. 시스템의 일부 대안에서, 하나 이상의 엔도뉴클레아제 및 하나 이상의 말단-프로세싱 효소가 제공되며, 여기서 이들은 2A 펩티드 서열 또는 "자가절단" 또는 자기-절단 서열에 의해 서로 연결된다. 일부 대안에서, 핵산 발현 벡터는 DNA 발현 벡터이다. 일부 대안에서, 핵산 발현 벡터는 RNA 발현 벡터이다. 일부 대안에서, 발현 벡터는 바이러스 벡터이다. 본원에 제공된 시스템의 일부 대안에서, 바이러스 벡터는 아데노-연관 바이러스 (AAV) 벡터이다.
- [0074] 일부 대안에서, 핵산 발현 벡터는 엔도뉴클레아제(들), 말단-프로세싱 효소(들), 및/또는 엔도뉴클레아제 및 말단-프로세싱 활성을 갖는 융합 단백질(들)을 선택 마커와 함께 수용하고 발현한 숙주 세포의 확인 또는 선택을 용이하게 하는 하나 이상의 선택 마커를 추가로 포함한다. 선택 마커의 예는 형광 단백질을 코딩하는 유전자, 예를 들어, EGFP, DS-Red, YFP, 또는 CFP; 선택 작용제에 대한 내성을 부여하는 단백질을 코딩하는 유전자, 예를 들어, PuroR 유전자, ZeoR 유전자, HygroR 유전자, neoR 유전자, 또는 블라스티시딘 내성 유전자를 포함하나 이에 제한되지는 않는다. 일부 경우에, 선택 마커는 형광 리포터 및 선택 마커를 포함한다.
- [0075] 일부 대안에서, DNA 발현 벡터는 하나 이상의 엔도뉴클레아제(들), 말단-프로세싱 효소(들), 및/또는 엔도뉴클레아제 및 말단-프로세싱 활성을 갖는 융합 단백질(들)의 발현을 구동할 수 있는 프로모터를 포함한다. 프로모터의 예는 레트로바이러스 LTR 요소; 구성적 프로모터, 예컨대 CMV, HSV1-TK, SV40, EF-1 α , 또는 β -액틴; 유도성 프로모터, 예컨대 Tet-오퍼레이터 요소를 함유하는 것; 및/또는 조직 특이적 프로모터를 포함하나 이에 제한되지는 않는다. 적합한 박테리아 및 진핵생물 프로모터는 관련 기술분야에 널리 공지되어 있고, 예를 들어 문헌 [Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd ed. 2001)]; [Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990)]; 및 [Current Protocols in Molecular Biology (2010)]에 기재되어 있으며, 참조문헌들은 그 전문이 본원에 참조로 명시적으로 포함된다. 식물 프로모터의 비제한적인 예는 에이. 탈리아나(*A. thaliana*) 유비퀴틴-3 (ubi-3)으로부터 유래된 프로모터 서열을 포함한다.
- [0076] 일부 대안에서, 하나 이상의 엔도뉴클레아제, 말단-프로세싱 효소, 및/또는 엔도뉴클레아제 및 말단-프로세싱 활성 또는 엑소뉴클레아제 활성을 갖는 융합 단백질을 코딩하는 핵산은 본원에 제공된 시스템의 벡터 및 핵산과 함께 진핵생물 세포로의 형질전환을 위해 벡터로 클로닝된다. 일부 대안에서, 상이한 엔도뉴클레아제 및 말단-프로세싱 효소를 코딩하는 핵산은 동일한 벡터로 클로닝된다. 이러한 경우에, 상이한 엔도뉴클레아제 및 말단-프로세싱 효소를 코딩하는 핵산은 T2A, 자기-절단 서열, 프로테아제 절단 부위, 또는 IRES 서열에 의해 임의로 분리될 수 있다. 벡터는 원핵생물 벡터, 예를 들어, 플라스미드, 또는 서플 벡터, 곤충 벡터, 또는 본원에 기재된 식물 벡터를 포함하는 진핵생물 벡터일 수 있다. 뉴클레아제 및 융합 단백질의 발현은 구성적 프로모터 또는 유도성 프로모터의 제어하에 있을 수 있다. 일부 대안에서, 벡터는 Cas9, 그의 유도체 또는 단편을 코딩하는 핵산 서열을 포함한다. 일부 대안에서, 벡터는 Trex를 코딩하는 핵산 서열을 포함한다. 일부 대안에서, 벡터에서 유전자 및/또는 핵산은 포유동물 세포, 예컨대 인간 세포에서의 발현을 위해 코돈 최적화된다. 일부 대안에서, 벡터는 mRNA이다. 일부 대안에서, 벡터는 Cas9 단백질, 그의 유도체 또는 단편을 코딩하는 mRNA이다. 일부 대안에서, Cas9 단백질, 그의 유도체 또는 단편을 코딩하는 핵산은 진핵생물 세포, 예컨대 인간 세포에서의 발현을 위해 코돈 최적화된다. 일부 대안에서, Cas9 단백질, 그의 유도체 또는 단편은 에스. 피오게네스로부터 유래되거나, 다른 유기체로부터의 다른 Cas9 단백질로 만든 컨센서스 서열이다.
- [0077] 엔도뉴클레아제 및/또는 말단-프로세싱 활성을 갖는 폴리펩티드 및/또는 엔도뉴클레아제 및/또는 말단-프로세싱 활성을 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 숙주 세포에 도입하는 것은 본원에 기재된 바와 같이 또는 관련 기술분야의 통상의 기술자에게 공지된 바와 같이 핵산 또는 단백질 전달을 위한 임의의 적합한 방법을 사용할 수 있다. 본원에 기재된 폴리펩티드 및 폴리뉴클레오티드는 시험관내에서 배양된 세포로, 뿐만 아니

라 계내에서 조직 및 전체 유기체로 전달될 수 있다. 본 대안의 폴리펩티드 및 폴리뉴클레오티드를 숙주 세포에 도입하는 것은 화학적으로, 생물학적으로 또는 기계적으로 성취될 수 있다. 이는 전기천공, 초음파천공, 유전자 총의 사용, 지질형질감염, 인산칼슘 형질감염, 덴드리머의 사용, 미세주사, 폴리브렌, 원형질체 융합, 아데노바이러스, AAV 또는 레트로바이러스 벡터를 포함하는 바이러스 벡터, 또는 그룹 II 리보자임의 사용을 포함할 수 있으나 이에 제한되지는 않는다.

[0078] AAV 벡터에 대한 면역 반응

[0079] 아데노-연관 바이러스 (AAV) 벡터는 유전적 질환의 유전자 요법-기반 치료에 사용될 수 있다. 그러나, AAV 벡터에 대한 면역 반응의 생성은 벡터의 치료 효능을 약화시킬 수 있다. 유사하게, CRISPR/Cas9-기반 (또는 하나 이상의 다른 뉴클레아제-기반) 게놈 편집에 사용된 AAV 벡터에 대한 면역 반응의 생성은 유전자 표적화의 효능을 약화시킬 수 있다.

[0080] 일부 대안에서, CRISPR/Cas9-기반 (및/또는 하나 이상의 다른 뉴클레아제-기반) 게놈 편집에 사용되는 AAV 벡터는 감소된 면역원성을 보유할 것으로 고려된다. 일부 대안에서, CRISPR/Cas9-기반 (및/또는 하나 이상의 다른 뉴클레아제-기반) 게놈 편집에 사용되는 AAV 벡터는 면역원성을 보유하지 않을 것으로 고려된다. 일부 대안에서, 감소된 면역원성 때문에, AAV 벡터에 대한 내성이 발생할 가능성이 최소화될 것이다. 일부 대안에서, 면역원성 결여 때문에, AAV 벡터에 대한 내성이 발생할 가능성은 감소되거나 존재하지 않을 것이다.

[0081] 유기체

[0082] 본원에 기재된 대안은 유전자, 특히 예를 들어, 헤모글로빈 또는 헤모글로빈-관련 유전자를 편집하고자 하는 임의의 진핵생물 유기체에 적용가능할 수 있다. 진핵생물 유기체의 예는 조류, 식물, 동물 (예를 들어, 포유동물, 예컨대 마우스, 래트, 영장류, 돼지, 소, 양, 토끼, 개, 고양이 또는 말 등), 어류, 또는 곤충을 포함하나 이에 제한되지는 않는다. 일부 대안에서, 유기체로부터 단리된 세포는 본원에 기재된 바와 같이 유전자 변형된다. 일부 대안에서, 변형된 세포는 생식적으로 성숙한 유기체로 발달한다. 진핵생물 (예를 들어, 조류, 효모, 식물, 진균, 어류, 새 또는 포유동물 세포) 세포가 사용될 수 있다. 하나 이상의 추가 유전자 변형을 함유하는 유기체로부터의 세포가 또한 사용될 수 있다.

[0083] 포유동물 세포의 예는 관심있는 유기체의 임의의 세포 또는 세포주, 예를 들어 난자, 체세포, K562 세포, CHO (차이니즈 햄스터 난소) 세포, HEP-G2 세포, BaF-3 세포, 슈나이더(Schneider) 세포, COS 세포 (SV40 T-항원을 발현하는 원숭이 신장 세포), CV-1 세포, HuTu80 세포, NTERA2 세포, NB4 세포, HL-60 세포 또는 HeLa 세포, 293 세포 또는 골수종 세포, 예컨대 SP2 또는 NS0를 포함한다. 배아 및 성인 줄기 세포와 마찬가지로 말초 혈액 단핵세포 (PBMC) 또는 T-세포가 또한 사용될 수 있다. 예를 들어, 사용될 수 있는 줄기 세포는 배아 줄기 세포 (ES), 유도된 만능 줄기 세포 (iPSC), 중간엽 줄기 세포, 조혈 줄기 세포, 근육 줄기 세포, 피부 줄기 세포, 지방 유래 줄기 세포, 또는 뉴런 줄기 세포를 포함한다. 일부 대안에서, 세포에서 적어도 하나의 표적 유전자를 편집하기 위한 시스템이 제공되며, 여기서 시스템은 CRISPR 가이드 RNA를 코딩하는 제1 핵산 서열을 포함하며, 여기서 CRISPR 가이드 RNA는 세포에서 적어도 하나의 표적 유전자에 상보적이고, 여기서 상기 제1 핵산 서열은 벡터에 존재하고, 여기서 상기 시스템은 Cas9 단백질을, 그의 유도체 또는 단백을 코딩하는 제2 핵산 서열, 제1 아데노바이러스 단백질을 코딩하는 제3 핵산 서열 및 제2 아데노바이러스 단백질을 코딩하는 제4 핵산 서열을 추가로 포함한다. 일부 대안에서, 세포는 진핵생물 세포이다. 일부 대안에서, 세포는 포유동물 세포, 예컨대 인간 세포이다. 일부 대안에서, 세포는 1차 세포이다. 일부 대안에서 세포는 형질전환된 세포가 아니다. 일부 대안에서, 세포는 1차 림프구, CD34+ 줄기 세포, 간세포, 심근세포, 뉴런, 아교 세포, 근육 세포 또는 장 세포이다.

[0084] 본원에 기재된 바와 같은 "조혈 줄기 세포" 또는 "HSC"는 골수성 세포 예컨대, 예를 들어, 대식세포, 단핵구, 대식세포, 호중구, 호염기구, 호산구, 적혈구, 거핵구/혈소판, 수지상 세포 또는 림프계 계통 (예컨대, 예를 들어, T-세포, B-세포 또는 NK-세포)을 생성할 수 있는 전구체 세포이다. 일부 대안에서, HSC는 세 가지 부류의 줄기 세포가 존재하는 이질적인 집단을 가지며, 이는 혈액내 림프성 대 골수성 자손의 비율 (L/M)로 구별된다.

[0085] 제약 투여

[0086] 본원에 제공된 시스템 또는 방법에 의해 제조된 세포는 DNA 서열의 표적화된 절단 및 치료적 또는 예방적 적용, 예를 들어, 헤모글로빈-관련 질환, 예컨대 겸상 적혈구 질환 또는 베타 지중해빈혈의 치료, 억제 또는 호전을 위해 환자에게 직접 투여될 수 있다. 일부 대안에서, 세포는 본원에 제공된 조성물, 시스템 또는 방법에 의해 제조된다. 일부 대안에서, 조성물이 제공되며, 여기서 조성물은 세포를 포함한다. 일부 대안에서, 본원에 기

재된 조성물은 질환을 치료, 예방, 호전 또는 억제하는 방법 또는 질환과 연관된 질환 상태 또는 증상을 호전시키는 방법에 사용될 수 있다. 일부 대안에서, 세포 또는 조성물은 유전적 질환을 치료, 예방, 호전 또는 억제하기 위해 투여된다.

- [0087] 세포를 포함하는 조성물은 임의의 적합한 방식으로, 및 일부 대안에서 제약상 허용되는 담체와 함께 투여된다. 이러한 단백질 또는 폴리뉴클레오티드를 투여하는 적합한 방법이 이용가능하고 관련 기술분야의 통상의 기술자에게 널리 공지되어 있으며, 특정 조성물을 투여하기 위해 하나 초과와 경로의 사용될 수 있지만, 특정 경로는 종종 또 다른 경로보다 더 즉각적이고 더 효과적인 반응을 제공할 수 있다.
- [0088] 제약상 허용되는 담체는 부분적으로 투여될 특정 조성물, 뿐만 아니라 조성물 투여에 사용되는 특정 방법에 의해 결정된다. 따라서, 이용가능한 제약 조성물의 다양한 적합한 제형이 있다 (예를 들어, 문헌 [Remington's Pharmaceutical Sciences] 참조).
- [0089] 비경구 투여, 예컨대, 예를 들어, 정맥내, 근육내, 피내 및 피하 경로에 의한 투여에 적합한 제형은 항산화제, 완충제, 정균제, 또는 제형을 의도된 수용자의 혈액과 등장성으로 만드는 용질을 함유할 수 있는 수성 및 비수성, 등장성 멸균 주사 용액, 또는 현탁화제, 가용화제, 증점제, 안정화제 또는 보존제를 포함할 수 있는 수성 또는 비수성 멸균 현탁액을 포함한다. 개시된 조성물은 예를 들어, 정맥내 주입에 의해, 경구로, 국소로, 복강내로, 방광내로 또는 경막내로 투여될 수 있다. 화합물의 제형은 단위-용량 또는 다중-용량 밀봉된 컨테이너, 예컨대 앰플 또는 바이알에 제공될 수 있다. 주사 용액 및 현탁액은 멸균 분말, 과립 또는 정제로부터 제조될 수 있다.
- [0090] 일부 대안에서, 비경구, 피하, 관절내, 기관지내, 복강내(intraabdominal), 관절낭내, 연골내, 강내, 복내(intracelial), 소뇌내, 뇌실내, 결장내, 자궁경관내, 위내, 간내, 심근내, 골내, 골반내, 심낭내, 복강내, 흉막내, 전립선내, 폐내, 직장내, 신장내, 망막내, 척수내, 활액낭내, 흉곽내, 자궁내, 방광내, 병변내, 볼루스, 질, 직장, 협측, 설하, 비강내, 또는 경피 투여 경로 중 하나 이상이 고려된다. 일부 대안에서, 투여될 조성물은 상기 언급된 경로 중 하나 이상을 통한 전달을 위해 제제화될 수 있다.
- [0091] 핵산 조성물
- [0092] 일부 대안은 HBB 유전자를 편집하기 위한 조성물에 관한 것이다. 일부 대안에서, 조성물은 핵산을 포함한다. 일부 대안에서, 핵산은 단일 가이드 RNA (sgRNA), 예컨대 서열식별번호: 1, 2, 3, 4, 5 또는 6 중 어느 하나에 의해 코딩되는 것을 포함한다. 일부 대안에서, sgRNA를 코딩하는 뉴클레오티드 서열은 서열식별번호: 1, 2, 3, 4, 5 또는 6의 서열과 적어도 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나, 또는 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일하다.
- [0093] 일부 대안에서, 핵산은 가이드 RNA (gRNA) 스캐폴드와 조합된 sgRNA, 예컨대 서열식별번호: 7, 8, 9, 10, 11 또는 12 중 어느 하나에 의해 코딩되는 것을 포함한다. 일부 대안에서, sgRNA를 코딩하는 뉴클레오티드 서열은 서열식별번호: 7, 8, 9, 10, 11 또는 12의 서열과 적어도 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나, 또는 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일하다.
- [0094] 일부 대안에서, 핵산은 서열식별번호: 13-18 중 어느 하나에 의해 코딩된 프로토스페이서 인접한 모티프 (PAM) 서열을 포함한다.
- [0095] 일부 대안은 겸상 돌연변이, 예컨대 SCL-g1 (본원에서 "SCL-1" 또는 "g1"이라고도 함)에서 직접 DNA를 절단하는 sgRNA를 포함한다. 예를 들어, Cas9 뉴클레아제와 함께, CRISPR 대안은 비가변 스캐폴드 서열의 표적 DNA 5'에 특이적인 대략 20개-염기 서열을 함유하는 sgRNA의 도입을 포함할 수 있다. sgRNA는 프로모터의 제어하에 sgRNA-코딩 서열을 갖는 플라스미드로 세포를 형질전환시킴으로써 또는 RNA로서 전달될 수 있다.
- [0096] 일부 대안에서, 핵산은 결실 복구 주형 또는 비결실 복구 주형, 예컨대 AAV에 의해 전달될 주형을 포함한다. 일부 대안에서, 주형은 서열식별번호: 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 21 중 하나 이상을 포함하거나; 또는 서열식별번호: 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 21 중 어느 하나의 서열과 적어도 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일한 서열을 포함한다. 일부 대안에서, 복구 주형의 뉴클레오티드 서열은 서열식별번호: 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 또는 36 중 어느 하나의 서열과 적어도 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일하다. 주형

은 상동성-지정 복구 (HDR)를 최대화하기 위해 조절인자 및/또는 인헨서를 포함할 수 있다.

- [0097] 일부 대안에서, 핵산은 TALEN, 예컨대 서열식별번호: 22 또는 서열식별번호: 23에 코딩된 것을 포함한다. 일부 대안에서, TALEN의 뉴클레오타이드 서열은 서열식별번호: 22 또는 서열식별번호: 23에 따른 서열 또는 이에 의해 코딩된 서열과 적어도 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일하다.
- [0098] 일부 대안에서, 핵산은 단일-가닥 공여자 올리고뉴클레오타이드 (ssODN)를 포함한다. 일부 실시양태에서, ssODN은 서열식별번호: 19, 20 또는 21 중 하나 이상을 포함한다. 일부 대안에서, ssODN의 뉴클레오타이드 서열은 서열식별번호: 19, 20 또는 21 중 어느 하나에 따른 서열 또는 이에 의해 코딩된 서열과 적어도 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일하다.
- [0099] HBB 유전자를 편집하는 특정 방법
- [0100] 본원에 제공된 방법 및 조성물의 일부 실시양태는 세포에서 HBB 유전자를 편집하는 방법을 포함한다. 이러한 일부 실시양태에서, 편집은 HDR을 포함할 수 있다. 일부 실시양태는 (i) 가이드 RNA (gRNA)를 코딩하는 폴리뉴클레오타이드를 세포에 도입하거나, TALEN을 코딩하는 폴리뉴클레오타이드를 세포에 도입하는 단계; 및 (ii) 주형 폴리뉴클레오타이드를 세포에 도입하는 단계를 포함한다.
- [0101] 일부 실시양태에서, gRNA는 서열식별번호:01-06 중 어느 하나의 뉴클레오타이드 서열과 적어도 약 85%, 90% 또는 95% 동일성을 갖는 핵산을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:07-12 중 어느 하나의 뉴클레오타이드 서열과 적어도 약 85%, 90% 또는 95% 동일성을 갖는 핵산을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:01-06 중 어느 하나의 뉴클레오타이드 서열을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:01의 뉴클레오타이드 서열을 포함한다. 일부 실시양태에서, gRNA는 서열식별번호:07의 뉴클레오타이드 서열을 포함한다.
- [0102] 일부 실시양태에서, gRNA를 코딩하는 폴리뉴클레오타이드를 세포에 도입하는 단계는 세포를 CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드를 포함하는 리보핵단백질 (RNP)과 접촉시키는 것을 포함한다. 일부 실시양태에서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드는 0.1:1 내지 1:10, 또는 1:1 내지 1:5의 비율을 갖는다. 일부 실시양태에서, CAS9 단백질 및 gRNA를 코딩하는 폴리뉴클레오타이드는 약 1:2.5의 비율을 갖는다.
- [0103] 일부 실시양태에서, 주형 폴리뉴클레오타이드는 HBB 유전자의 적어도 일부 또는 그의 상보체를 코딩한다. 일부 실시양태에서, 주형 폴리뉴클레오타이드는 야생형 HBB 유전자의 적어도 일부 또는 그의 상보체를 코딩한다. 일부 실시양태에서, HBB 유전자의 적어도 일부는 HBB 유전자의 엑손 1을 포함한다.
- [0104] 일부 실시양태에서, 바이러스 벡터는 주형 폴리뉴클레오타이드를 포함한다. 일부 실시양태에서, 벡터는 아테노-연관 바이러스 (AAV) 벡터이다. 일부 실시양태에서, 벡터는 자기-상보적 AAV (scAAV) 벡터이다. 일부 실시양태에서, 주형 폴리뉴클레오타이드는 HBB 유전자의 적어도 약 4 kb를 포함한다.
- [0105] 일부 실시양태에서, 단일-가닥 공여자 올리고뉴클레오타이드 (ssODN)는 주형 폴리뉴클레오타이드를 포함한다.
- [0106] 일부 실시양태에서, ssODN은 서열식별번호:64-72 중 어느 하나의 뉴클레오타이드 서열과 적어도 80%, 85%, 90% 또는 95% 동일성을 갖는 뉴클레오타이드 서열을 포함한다. 일부 실시양태에서, ssODN은 서열식별번호:64-72 중 어느 하나의 뉴클레오타이드 서열을 포함한다.
- [0107] 일부 실시양태에서, 이중-가닥 파단은 HBB 유전자의 엑손 1에서 생성된다. 일부 실시양태에서, 이중-가닥 파단은 HBB 유전자의 엑손 1에서 6번째 코돈에 인접하게 생성된다.
- [0108] 일부 실시양태에서, 단계 (i)은 단계 (ii) 전에 수행된다. 일부 실시양태에서, 단계 (i) 및 (ii)는 동시에 수행된다. 일부 실시양태에서, 단계 (i) 및/또는 (ii)는 뉴클레오펙션을 수행하는 것을 포함한다. 일부 실시양태에서, 뉴클레오펙션을 수행하는 것은 분자 시스템의 사용을 포함한다. 일부 실시양태에서, 시스템은 구형과 필스의 사용을 포함한다. 일부 실시양태에서, 단계 (i) 및/또는 (ii)는 약 200,000개 세포/20 μ l 뉴클레오펙션 반응물을 접촉시키는 것을 포함하고, 뉴클레오펙션 반응물은 gRNA 및/또는 주형 폴리뉴클레오타이드를 포함한다.
- [0109] 일부 실시양태에서, 세포는 포유동물이다. 일부 실시양태에서, 세포는 인간이다. 일부 실시양태에서, 세포는 1차 세포이다. 일부 실시양태에서, 세포는 조혈 줄기 세포 (HSC)이다. 일부 실시양태에서, 세포는 T 세포 또는 B 세포이다. 일부 실시양태에서, 세포는 CD34+ 세포이다.

- [0110] 일부 실시양태에서, HBB 유전자는 서열식별번호:37의 뉴클레오타이드 서열과 적어도 95% 동일성을 갖는다.
- [0111] 요법 방법
- [0112] 겸상 적혈구 질환 (SCD)은 6번째 아미노산 (E6V)에서 글루탐산의 발린으로의 치환을 유발하는 HBB 유전자의 엑손 1에서 단일 뉴클레오타이드 전환에 의해 유발된다. 이러한 변화는 성인 글로빈 (β^A)의 소수성을 증가시키고 중합에 민감하게 하여 적혈구의 특징적인 겸상 패틴을 유발한다. 겸상 환자는 이환율이 증가되고 수명이 단축되어 수혈에 의존적인 상태를 유지한다. 공여자 주형 (재조합 아데노-연관 바이러스 (rAAV) 또는 ssODN)의 존재하에 뉴클레아제를 사용한 유전자 편집은 돌연변이를 수정하고 세포 복구 기구에 의한 주형-구동 복구를 구동할 수 있다. 최적의 이점을 위해, SCD에서 임상 유전자 편집은 NHEJ를 통한 표적 HBB 뉴클레아제-구동 유전자 파괴를 공동으로 제한하면서 효율적인 공여자-지정 뉴클레오타이드 변화를 유발할 것이다.
- [0113] 따라서, 본원에 제공된 일부 대안은 치료적 게놈 편집 접근법을 사용하여 SCD를 치료, 호전, 억제 또는 개선시키는 것에 관한 것이다. 일부 대안에서, HSPEC에서 내인성 프로모터 및 인헨서의 제어하에 무손상 HBB cDNA를 도입하기 위한 시스템 및 방법이 제공된다. 일부 대안에서, 본원에 기재된 시스템 및 방법은 HBB에서 번역학적 및 기능적 결함을 구제하고 치유 요법을 제공한다.
- [0114] 일부 대안은 HBB 유전자를 편집하는 방법에 관한 것이다. 예를 들어, 방법은 HBB 유전자를 포함하는 세포를 제공하는 것을 포함할 수 있다. 일부 대안에서, 방법은 본원에 기재된 핵산 조성물 중 하나 이상, 예컨대 서열식별번호: 1-36 중 하나 이상에 따른 서열 또는 이에 의해 코딩된 서열, 또는 예를 들어, 서열식별번호: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 또는 36 중 어느 하나에 따른 서열 또는 이에 의해 코딩된 서열과 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하거나 또는 상기 언급된 백분율 중 임의의 2개에 의해 정의된 범위 내에서 동일한 서열을 세포에 제공하는 것을 포함한다.
- [0115] 일부 대안에서, 세포는 포유동물 세포, 인간 세포, 1차 세포, 림프구, CD34+ 줄기 세포, 간세포, 심근세포, 뉴런, 아교 세포, 근육 세포 또는 장 세포, 또는 본원에 기재된 임의의 세포이다. 일부 대안에서, 핵산은 세포를 바이러스 벡터로 형질도입하거나 세포를 바이러스로 감염시킴으로써 세포에 제공된다. 일부 대안에서, 바이러스 벡터는 아데노-연관 바이러스 (AAV) 벡터, 예컨대 재조합 AAV이다. AAV는 다중 혈청형의 하나 또는 혼합물, 예컨대 혈청형 6일 수 있다. 예를 들어, 재조합 혈청형 6 AAV (rAAV6)가 일부 대안에서 사용된다.
- [0116] 일부 대안에서, 핵산은 숙주 세포, 예를 들어 진핵생물 세포, 예컨대 인간 세포에서의 발현을 위해 코돈-최적화된다. 일부 대안은 유전자-편집 단백질, 예컨대 Cas9 단백질을 코딩하는 제2 핵산을 세포에 제공하는 것을 포함한다. 일부 대안에서, 제2 핵산은 제1 핵산, 예컨대 하나 이상의 AAV 유전자를 코딩하는 핵산과는 별도의 핵산이나, 제1 핵산과 조합될 수 있다. 일부 대안에서, 핵산을 제공하는 것은 HBB 유전자에 대한 하나 이상의 편집, 예컨대, 예를 들어, 겸상 적혈구 돌연변이에 대한 보정을 유발한다. 일부 대안에서, 겸상 적혈구 돌연변이에 대한 보정은 E7V 돌연변이의 보정을 포함한다. 일부 대안에서, 핵산을 제공하는 것은 HBB 유전자의 엑손 1에서 파손된 포스포디에스테르 결합을 초래한다.
- [0117] 일부 대안에서, 방법은 세포를 대상체의 골수에 생착시키는 것을 추가로 포함한다. 일부 대안에서, 세포는 대상체로부터의 것이고, 세포는 골수와 동일한 대상체로부터의 것이다. 다시 말해서, 생착될 세포는 대상체의 세포 또는 골수와 동종일 수 있다. 일부 대안에서, 세포는 대상체의 세포 또는 골수와 동종이게이다.
- [0118] 특정 서열
- [0119] 일부 실시양태는 서열식별번호: 1-36으로부터의 하나 이상의 서열을 포함한다. 표 1은 표 2에 추가로 기재된 서열식별번호: 22-36과 비교하여 상대적으로 짧은 서열식별번호: 1-21의 서열을 포함한다. 서열식별번호: 1-6은 sgRNA 표적 서열이다. 서열식별번호: 1 및 6은 안티센스 가닥 서열을 포함하고, 서열식별번호: 2-5는 센스 가닥 서열을 포함한다. 서열식별번호: 1-36은 5'에서 3' 방향이다. 서열식별번호: 36은 서열식별번호: 35와 유사하지만, HBB 프로모터의 직접 반복을 갖지 않으므로 작은 결실을 갖는다. 이 rAAV6 주형 (서열식별번호: 36)은 HBB 유전자좌로 E7V 돌연변이를 구동하는 HBG1 프로모터를 구동하고, 천연 인트론 1을 유지한다. MND-GFP는 서열식별번호: 36에서 프로모터 간섭을 방지하기 위해 역방향 배향에 있다. 표 2는 서열식별번호: 24-36에 관한 추가 정보를 제공한다. 서열식별번호: 22는 15,662의 서열 길이 및 49 KB의 크기를 갖는 DNA 분자인 반면, 서열식별번호: 23은 15,866의 서열 길이 및 54 KB의 크기를 갖는 DNA 분자이다. 서열식별번호: 22 및 23에 대한 벡터는 pWNY2.0로부터 상용성인 pEVL300 noBsmBI GG이다.

[0120] <표 1>

SEQ ID NO:	식별자(들)	서열
1	g1 또는 SCL-1	GTAACGGCAGACTTCTCCTC
2	g2 또는 SCL-2	GTCTGCCGTTACTGCCCTGT
3	g3 또는 SCL-3	TCTGCCGTTACTGCCCTGT
4	g4 또는 SCL-4	AGTCTGCCGTTACTGCCCTG
5	g5 또는 SCL-5	AAGGTGAACGTGGATGAAGT
6	g6 또는 SCL-6	CTTGCCCCACAGGGCAGTAA
7	SCL-1 + gRNA 스캐폴드	<u>GTAACGGCAGACTTCTCCTC</u> GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCAGTCGGTGCTTTT
8	SCL-2 + gRNA 스캐폴드	<u>GTCTGCCGTTACTGCCCTGT</u> GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCAGTCGGTGCTTTT
9	SCL-3 + gRNA 스캐폴드	<u>TCTGCCGTTACTGCCCTGT</u> GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCAGTCGGTGCTTTT
10	SCL-4 + gRNA 스캐폴드	<u>AGTCTGCCGTTACTGCCCTG</u> TTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCAGTCGGTGCTTTT
11	SCL-5 + gRNA 스캐폴드	<u>AAGGTGAACGTGGATGAAGT</u> GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCAGTCGGTGCTTTT
12	SCL-6 + gRNA 스캐폴드	<u>CTTGCCCCACAGGGCAGTAA</u> GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCAGTCGGTGCTTTT
13	SCL-1 PAM 서열	AGG

[0121]

SEQ ID NO:	식별자(들)	서열
14	SCL-2 PAM 서열	GGG
15	SCL-3 PAM 서열	GGG
16	SCL-4 PAM 서열	TGG
17	SCL-5 PAM 서열	TGG
18	SCL-6 PAM 서열	CGG
19	ssODN- V7E:GTC	TCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG ACACAACCTGTGTTCACTAGCAACCTCAAACAGACAC CATGGTGCATCTGACTCCTGTCGAGAAGTCTGCCGT ACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTT GGTGGTGAGGCCCTGGGCAGGT
20	ssODN- V7E:GTG	TCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG ACACAACCTGTGTTCACTAGCAACCTCAAACAGACAC CATGGTGCATCTGACTCCTGTGGAGAAGTCTGCCGT ACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTT GGTGGTGAGGCCCTGGGCAGGT
21	ssODN CCCGAA	TCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG ACACAACCTGTGTTCACTAGCAACCTCAAACAGACAC CATGGTGCATCTGACTCCCGAAGAGAAGTCTGCCGT TACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGT TGGTGGTGAGGCCCTGGGCAGGT

[0122]

[0123] <표 2>

SEQ ID NO:	대안적인 명칭 또는 기재어	서열 길이	크기
22	FrankenpEVL_SCL L Talen		
23	FrankenpEVL_SCL R Talen		
24	AMS#1242 pAAV HBB(400)>synTron.HBBopt(T87Q).wPRE- O.BGHpA;MND>GFP::P2A::Ex2	7365	40 KB
25	AMS#1243 pAAV HBB(400)>synTron.HBBopt(T87Q).wPRE- O.BGHpA;HPFH2.MND>GFP::P2A::Ex2	8087	42 KB
26	AMS#1244 pAAV HBB(400)>synTron.HBBopt(T87Q).wPRE3.US E-SV40pA;HPFH2.MND>GFP::P2A::Ex2	7712	49 KB
27	AMS#1245 pAAV HBB(400)>HBBopt(T87Q).wPRE- O.BGHpA.HPFH2;MND>GFP::P2A::Ex2	7925	57 KB
28	AMS#1246 pAAV HBB(600)>HBBopt(E7V) 겸상 비복구	5216	31 KB
29	AMS#1247 pAAV HBB(400)>HBBopt(E7V) 겸상 비복구	4816	23 KB
30	AMS#1248 pAAV HBB(200)>HBBopt(E7V) 겸상 비복구	4416	28 KB
31	AMS#1249 pAAV HBB(50)>HBBopt(E7V) 겸상 비복구	4116	22 KB
32	AMS#1289 pAAV HBB(400)d0>HBBopt(T87Q).wPRE- O.BGHpA.HPFH2;MND>GFP.SV40pA	7995	78 KB
33	AMS#1290 pAAV HBB(400)d14>HBBopt(T87Q).wPRE- O.BGHpA.HPFH2;MND>GFP.SV40pA	7995	79 KB
34	AMS#1314 pAAV HBB(2.1kb).SCmutation	8036	64 KB

[0124]

SEQ ID NO:	대안적인 명칭 또는 기재어	서열 길이	크기
35	AMS#1321 pAAV HBB(800).d53,71 [MND>GFP.SV40pA];HBG1d13p>HBB(E7V)o pt	8207	95 KB
36	AMS#1322 pAAV HBB(7,800).d-127,71 [MND>GFP.SV40pA];HBG1d13p>HBB(E7V)o pt	8009	90 KB

[0125]

[0126] <실시예>

[0127] 실시예 1-β-글로빈 (HBB) 유전자좌에서 HDR을 위한 rAAV6 및 ssODN의 비교

[0128] Crispr/Cas9 리보핵단백질 (RNP)을 사용하여 인간 동원된 말초 혈액 CD34+ 세포 (hPBSC)에 E6V 겸상 돌연변이를 도입함으로써 SCD 보정에서 rAAV6 및 ssODN 전달의 영향 및 임상적 관련성을 평가하였다. 두 가지 공여자 전달 전략을 사용하였다: 2.2 kb 상동성 아암 (HA)을 갖는 rAAV6 (AMS#1314) 및 168개 뉴클레오티드를 포함하는 ssODN (E7V-GTC 및 E7V-GTG 변화 및 CCCGAA 변화를 갖는 V7E). 잔류 NHEJ 비율과 비교하여 HDR의 효율을 Crispr/Cas9 RNP 생성된 이중-가닥 파단에 따라 평가하였다 (3% rAAV6, AMS#1314 및 12.5, 25, 50, 또는 100 pmol의 ssODN).

- [0129] 인간 동원된 말초 혈액 CD34+ 세포를 시토카인을 함유하는 SCGM 배지에서 48시간 동안 해동시켰다: 100 ng/ml의 SCF, IL-6, Flt-3L, TPO. 세포를 해동 48시간 후 전기천공하고, rAAV6 전달을 위해 바이러스-함유 회수 배지 (3% rAAV6, AMS#1314)에 첨가하였다. GTC 또는 GTG 변화 또는 CCCGAA 변화를 갖는 V7E를 갖는 12.5, 25, 50, 또는 100 pmol의 E7V ssODN을 함유하는 용량 적정의 ssODN을 RNP와 함께 전기천공한 후, 회수 배지에 첨가하였다. 18-24시간 후, 세포를 1% Pen/Strep, 20 ng/mL hSCF, 1 ng/mL hIL-3, 2 IU/mL EPO 및 20% 열-불활성화된 FBS와 함께 IMDM을 함유하는 분화 배지로 옮겼다. 세포를 14일 동안 분화시키고, 적혈구계 세포를 다양한 글로빈 하위유형에 대해 분석하였다.
- [0130] 시험관내 시험시, 편집 후 10일차에 HDR:NHEJ의 비율은 rAAV6 (AMS#1314)을 사용하여 시험된 5명의 CD34+ 공여자에 걸쳐 31%:18%였고, ssODN E7V GTG를 사용하여 시험된 3명의 CD34+ 공여자에 걸쳐 13%:30%였다 (각각 대략 2:1 vs. 1:2의 비율). 편집된 CD34+ 세포를 2주 동안 적혈구계 세포로 분화시키고, 적혈구계 전구체 중 β^S (겸상 글로빈)의 양을 RP-HPLC를 사용하여 정량화하였다. β^S 의 양은 각각 rAAV6 주형 전달을 사용한 경우 18-27% (n=5) 및 ssODN 전달을 사용한 경우 0-9% (n=3)였다. 그러므로, rAAV6 또는 ssODN E7V GTG는 HBB 유전자좌 내에 표적화된 뉴클레오타이드 변화를 성공적으로 도입하였다. rAAV6(AMS#1314)은 HBB 유전자좌 내에 표적화된 뉴클레오타이드 변화를 도입하는 우수한 방법이었다. 이들 발견은 또한 온-타겟 유전자 파괴를 제한하면서 뉴클레오타이드 변화를 도입하도록 설계된 임상 유전자 편집에 대한 잠재력의 평가에서 지표로서 HDR:NHEJ 비율 및 치료 단백질 수준 둘 다를 측정하는 이점을 강조한다.
- [0131] 실시예 2-HBB 유전자의 엑손-1 내의 겸상 돌연변이를 편집하고 보정하기 위한 HDR 주형 설계 및 전달
- [0132] 겸상 돌연변이를 편집하고 보정하기 위해, HBB 유전자의 E7V 돌연변이에서 편집하는 뉴클레아제를 개발하였다. HBB 유전자의 엑손 1에서 편집하기 위해 화학적으로 변형된 단일 가이드 RNA (sgRNA)의 TALEN 뿐만 아니라 Crispr/Cas9 리보핵단백질-매개 전달을 최적화하였다. 데이터는 K562 및 인간 조혈 줄기 세포 (CD34+)에서 겸상 유전자좌에서 효율적인 편집을 나타내었다. 고유한 조절성 요소를 갖는 신규 HDR 복구 주형의 rAAV6-매개 전달을 위해 다양한 복구 주형 아키텍처를 설계하였다. 항-겸상화 (T87Q) 글로빈 카세트, 겸상 글로빈 도입 및 겸상 보정 카세트를 HBB 유전자좌에서 시험하였다. 이들 주형의 설계는 고유하였다. HBB 유전자좌에서 효율적이고 임상적으로 관련된 상동성-의존적 복구 (HDR) 비율이 달성되었다.
- [0133] HBB 유전자좌에서 HDR을 구동하도록 복구 주형의 ssODN 전달을 또한 설계하고 최적화하였다. 겸상 돌연변이를 도입하고 겸상 돌연변이를 보정하는 ssODN을 설계하고 시험하였다. 겸상 유전자좌에서 단일- 또는 이중-대립유전자 HDR 통합이 달성되었다. 또한, 전달 방식에 관계없이 통합된 주형으로부터 임상적으로-관련된 글로빈 발현이 관찰되었다.
- [0134] RNP 및 TALEN을 스크리닝하여, 인간 HBB 유전자의 엑손 1에서 이중 가닥 (ds) 파단을 생성하였다. 도 1a는 RNP 및 TALEN으로서 전달된 다양한 sgRNA를 비교하여 뉴클레아제 효율을 나타낸다. 가장 높은 % INDEL을 유발한 뉴클레아제는 g4, g5 및 g6이었다. Cas9:sgRNA 비율은 편집 효율을 최대화하도록 구성되었으며, Cas9:sgRNA의 20:50 비율은 40:40의 비율보다 더 높은 % INDEL을 유발하였다 (도 1b). 도 1c에 나타낸 바와 같이, RNP로서 전달된 sgRNA 1 및 6을 사용하여 상이한 CD34+ 공여자에 걸쳐 편집 효율을 평가하였다. 결과는 HBB 유전자좌에서 효율적인 편집이 뉴클레아제로 달성되었음을 나타낸다.
- [0135] 조절인자 및 인헨서는 항-겸상화 글로빈 (β^{T87Q})을 도입하는 결실 주형 1242-1245를 사용하여 HDR을 최대화하도록 구성되었다. 도 2a는 결실 주형 1242-1245의 개략도를 포함하고, 각각에 사용된 요소를 나타낸다. 도 2b에 나타낸 바와 같이, 전기천공 및 AAV6 형질도입 후 2일차에 CD34+ 세포의 생존율을 결정하였다. 결실 복구 주형을 사용한 형질도입은 약 60% 생존율을 유발하는 경향이 있었다. 도 2c는 결실 주형 1242-1245와 함께 RNP로서 전달된 3개의 상이한 가이드와 HDR의 비교를 나타낸다. 도 2d에 나타낸 바와 같이, RNP 및 결실 주형을 사용한 ddPCR에 기초한 상대 HDR %를 결실 주형 1242-1245로 결정하였다. 적혈구계 분화 배지에서 2주 동안 분화된 세포로부터 β -글로빈 발현을 관찰하여 TIDE 시퀀싱 분석 (도 2e) 및 HPLC 분석 (도 2f)을 사용하여 편집 효율을 또한 결정하였다. 결과는 항-겸상화 글로빈을 갖는 결실 복구 주형이 rAAV6에 의해 세포에 효과적이고 효율적으로 전달되고 세포의 계승으로 통합되었음을 나타낸다.
- [0136] 비결실 주형을 사용한 HDR은 또한 항-겸상화 글로빈 β^{T87Q} 를 도입하도록 구성되었다. 도 3a는 비결실 주형 1289-1290의 개략도를 포함하고, 각각에 사용된 요소를 나타낸다. 도 3b에 나타낸 바와 같이, 전기천공 및 AAV6 형질도입 후 2일차에 CD34+ 세포의 생존율을 결정하였으며, 또한 약 60%, 및 약 30% 내지 약 90% 범위 내

에 있는 경향이 있었다. 도 3c는 RNP 및 비결실 주형; 1289-1290을 사용한 유동 세포계측법에 기초한 HDR %를 나타낸다. 도 3d에 나타낸 바와 같이, RNP 및 비결실 주형을 사용한 ddPCR에 기초한 상대 HDR %를 주형 1289 및 1290에 대해 결정하였다. 벌크 GFP+ 집단 및 GFP+/- 분류된 세포에서 적혈구계 분화 배지에서 2주 동안 분화된 세포로부터 β-글로빈 발현을 관찰하여 TIDE 시퀀싱 분석 (도 3e) 및 HPLC 분석 (도 3f)을 사용하여 편집 효율을 또한 결정하였다. 결과는 항-겸상화 글로빈을 갖는 비결실 복구 주형이 rAAV6에 의해 세포에 효과적이고 효율적으로 전달되었고, HBB 유전자좌에 대한 변화가 효과적으로 이루어졌음을 나타낸다.

[0137] 겸상 돌연변이를 도입하는 다양한 HR 아암 길이를 갖는 인간 코돈-최적화 주형을 사용한 HDR을 개발하였다. 도 4a는 인간 코돈-최적화 주형 1246-1249의 개략도를 포함한다. 도 4b에 나타낸 바와 같이, 전기천공 및 AAV6 형질도입 후 2일차에 CD34+ 세포의 생존율을 결정하였다. 도 4c는 RNP 및 인간 코돈-최적화 주형 1246-1249를 사용한 ddPCR에 기초한 절대 HDR %를 나타낸다. 도 4d에 나타낸 바와 같이, 적혈구계 분화 배지에서 2주 동안 분화된 세포로부터 β-글로빈 발현을 관찰하여 HPLC 분석을 수행하였다. 결과는 항-겸상화 글로빈을 갖는 인간 코돈-최적화 주형이 rAAV6에 의해 세포에 효과적이고 효율적으로 전달되었고, HBB 유전자좌에 대한 변화가 효과적으로 이루어졌음을 나타낸다.

[0138] 주형 1314를 사용한 HDR은 겸상 돌연변이를 도입하는데 효과적이었다. 도 5a는 주형 1314의 개략도를 포함한다. 도 5b에 나타낸 바와 같이, 전기천공 및 AAV6 형질도입 후 2일차에 CD34+ 세포의 생존율을 결정하였다. 도 5c는 5명의 상이한 공여자로부터의 RNP 및 주형 1314로 편집된 샘플의 콜로니 시퀀싱의 결과를 나타낸다. 적혈구계 분화 배지에서 2주 동안 분화된 세포로부터 β-글로빈 발현을 관찰하여 4명의 공여자의 TIDE 시퀀싱 (도 5d) 및 HPLC 분석 (도 5e)을 사용하여 편집 효율을 결정하였다. 14일차 분화된 HDR 샘플로부터 다양한 글로빈 하위유형을 나타내는 크로마토그램 (도 5f)이 생성되었다. 결과는 rAAV6에 의해 전달된 주형의 전달 후 겸상 돌연변이 도입이 달성되었음을 나타낸다.

[0139] 비결실 주형을 사용한 HDR은 겸상 돌연변이를 도입하도록 구성되었다. 도 6a는 비결실 주형, 1321의 개략도를 포함한다. 도 6b에 나타낸 바와 같이, 전기천공 및 AAV6 형질도입 후 2일차에 CD34+ 세포의 생존율을 결정하였다. 주형 1321을 사용하여 편집된 샘플에서 HDR 사건이 측정되었다 (도 6c). 결과는 겸상 돌연변이 도입이 rAAV6에 의해 전달된 비결실 주형의 전달 후 달성되었음을 나타낸다.

[0140] ssODN을 사용한 HDR은 겸상 돌연변이를 도입하도록 구성되었다. 도 7a는 ssODN E7V의 개략도를 포함한다. 도 7b에 나타낸 바와 같이, 100, 50, 25 또는 12.5 pmol GTC를 사용한 전기천공 후 2일차에 CD34+ 세포의 생존율을 결정하였다. 이들 동일한 용량은 도 7c 및 7d에서 또한 사용되었다. E7V ssODN의 용량 적정으로 편집된 샘플에서 ddPCR에 의해 HDR 및 NHEJ를 측정하였다 (도 7c). 도 7d에 나타낸 바와 같이, 적혈구계 세포에서 발현된 다양한 글로빈 하위유형의 HPLC 분석을 수행하였다. 결과는 겸상 돌연변이가 ssODN의 전달 후 달성되었음을 나타낸다.

[0141] ssODN을 사용한 HDR은 겸상 돌연변이 (CCC GAA)를 보정하도록 구성되었다. 도 8a는 ssODN V7E의 개략도를 포함한다. 도 8b에 나타낸 바와 같이, 전기천공 후 2일차에 CD34+ 세포의 생존율을 측정하였다. 편집 후 3명의 상이한 공여자의 TIDE 시퀀싱에 의해 INDEL을 평가하였다 (도 8c). 도 8d에 나타낸 바와 같이, V7E ssODN의 용량 적정으로 편집된 샘플에서 ddPCR에 의해 HDR 및 NHEJ를 측정하였다. 도 8e에 나타낸 바와 같이, V7E ssODN으로 편집된 샘플에서 콜로니 시퀀싱에 의해 HDR 및 NHEJ를 또한 측정하였다. 적혈구계 세포에서 다양한 글로빈 하위유형의 HPLC 분석을 수행하였다 (도 8f). 결과는 ssODN이 전달 및 겸상 보정에 효과적이었음을 나타낸다.

[0142] 편집된 세포를 W41 SCID 마우스에 생착시켰다. 도 9a는 골수에서 12주에 인간 CD45+ 생착의 결과를 나타낸다. 생착 방법은 6-8-주령 W41 마우스를 25 mg/kg의 부설판으로 처리하는 것을 포함하였다. 2×10^6 개 인간 세포를 꼬리-정맥 주사에 의해 전달하였다. 마우스를 12주 동안 모니터링하고, 전체 인간 키메라즘, 다중-계통 생착 및 적혈구계 재구성을 희생 시점에 측정하였다. % HDR을 인간-특이적 프라이머를 사용한 ddPCR에 의해 측정하고, % Indel을 TIDE 시퀀싱에 의해 측정하였다.

[0143] 도 9b에 나타낸 바와 같이, INDEL을 생착된 인간 세포에서 TIDE 시퀀싱에 의해 측정하였다. 도 9c에 나타낸 바와 같이, HDR을 12주에 골수에서 ddPCR에 의해 측정하였다. 결과는 편집된 세포가 대상체의 골수에 생착되어 비-겸상 혈액 세포를 생성하였음을 나타낸다.

[0144] 본원에 기재된 바와 같이, 겸상 돌연변이를 보정하기 위해 HBB 유전자좌에 삽입될 수 있는 HDR 주형을 설계하기 위해 다양한 전략을 사용하였다. 주형은 다음 세 그룹을 포함하였다: 그룹 1: 다양한 인핸서, 인트론, 프로모터, 폴리A 테일, 다양한 상동성 아암 길이, 및/또는 T87Q 항-겸상화 글로빈을 HBB 유전자에 삽입하는 결실 및

비결실 카세트를 갖는 rAAV6-기반 HDR 주형. 이러한 주형은 HDR의 증거를 제공하고, 복구 주형을 인간 세포에 삽입함으로써 구동된 고유한 항-겸상화 (T87Q) 글로빈 발현을 나타내었다. 그룹 2: 다양한 인핸서, 인트론, 프로모터, 폴리A 테일, 다양한 상동성 아암 길이, 및/또는 HBB 유전자에 겸상 돌연변이를 삽입하거나 겸상 돌연변이를 보정하는 결실 및 비결실 카세트를 갖는 rAAV6-기반 HDR 주형. 이러한 주형은 HDR의 증거를 제공하고, 복구 주형을 인간 세포에 삽입함으로써 구동된 고유한 겸상 글로빈 발현 (HbS) 또는 성인 헤모글로빈 (HbA)을 나타내었다. 그룹 3: HBB 유전자에 겸상 돌연변이의 삽입을 구동하거나 겸상 보정을 구동하는 ssODN-기반 주형. 데이터는 HDR의 증거를 제공하고, ssODN을 인간 세포에 삽입함으로써 구동된 고유한 겸상 글로빈 발현 또는 성인 헤모글로빈 (HbA)을 나타내었다.

- [0145] 본원에서 설계되고 시험된 신규 주형의 예는 다음을 포함한다: 본원에 기재된 rAAV6 HDR 복구 주형 설계; 1242, 1243, 1244, 1245, 1246, 1247, 1248, 1249, 1289, 1290, 1314, 1321, 1322를 포함하는 rAAV6으로서 전달된 HDR 복구 주형; HBB 유전자의 엑손1의 코돈 7에서 GTG 변화 뿐만 아니라 GTC 변화를 구동하는 겸상 돌연변이를 도입하는 ssODN 설계 (GTC 및 GTG 변화를 갖는 E7V 카세트 둘 모두는 고유하고, 이전에 보고된 적이 없음); 및 겸상 돌연변이를 보정하고 HBB 유전자의 코돈 6 및 7에서 CCC GAA 변화를 구동하는 ssODN 설계.
- [0146] 겸상 돌연변이로부터 17개 염기 떨어져 있는 sgRNA (SCL-g6)를 사용하였다. HDR 주형을 scAAV6으로서 전달하고, HBB 유전자에서 통합하고 항-겸상화 HBB cDNA (HbAS3)를 도입하는 긴 cDNA 카세트를 전달하였다. cDNA를 유전자 시작에 삽입하여, 내인성 프로모터/인핸서 기능을 보존하였다. 대안적으로, 이전에 시험된 HDR 주형은 코돈 7에서 GTC 변화와 함께 다양한 코돈에서 5개의 코돈-최적화 뉴클레오티드 서열 변화를 갖는 2.2 kb HR 아암을 갖는 rAAV6 E6V 공여자를 사용하였다 (gTCgagaagtctgcAgtCactgcTctAtggggGaaA; 서열식별번호:38). 이들 주형을 RNP로서 전달된 SCL 가이드 6과 함께 작동하도록 설계하였다.
- [0147] 이전에 시도된 ssODN 주형은 HBB 유전자의 엑손 1 내의 코돈 7에서 E7V GTA (Dewitt et al.) 또는 V7E GAA 변화를 도입하고, 겸상 돌연변이로부터 17개 염기인 SCL-g6 sgRNA와 함께 작동한다. 반면에, 본원에 기재된 rAAV6 공여자 카세트는 겸상 돌연변이에서 특이적으로 절단하는 신규 가이드, SCL-g1을 사용한다. 본원에 기재된 신규 rAAV6 및 ssODN 복구 주형은 임상적으로 관련된 예상외로 높은 수준의 HDR을 구동하는 비결실 HDR 사건을 생성한다. 일부 대안에서, 이들 신규 공여자 주형은 (a) 항-겸상화 T87Q 글로빈을 삽입하거나, (b) 겸상 돌연변이를 도입하거나, (c) 인간 코돈-최적화 겸상 보정을 전달한다. 본원에 기재된 신규 rAAV6 카세트는 프로모터, 인핸서, 폴리A 테일 및 조절성 요소의 고유한 조합을 사용하여 글로빈 발현을 최대화한다.
- [0148] 일부 실시양태에서 기재된 접근법의 이점은 돌연변이의 부위에서의 편집이 기능적 결과를 개선시킬 수 있다는 점이다. 일부 실시양태에서, sgRNA는 겸상 돌연변이에서 절단된다. 일부 실시양태에서, 상동성 아암에 대한 절단 부위에 대한 근접성은 돌연변이의 전환의 개선을 허용한다. 일부 실시양태에서, rAAV6 주형 또는 ssODN 주형은 겸상 돌연변이에서 편집하는 SCL-g1과 함께 작동하도록 특이적으로 생성되어 겸상 돌연변이의 고도로 효율적인 보정을 허용한다.
- [0149] 일부 실시양태에서, 주형 설계는 HBB 유전자좌에서 HDR을 최대화한다. 일부 실시양태에서, HBB 유전자좌는 또한 일부 실시양태에서 주요 치료적 이점을 제공하는 본원에 나타난 이중-대립유전자 통합에 대한 예상외의 성향을 갖는다.
- [0150] 일부 실시양태에서, 글로빈 발현은 조절성 요소의 선택을 통해 최대화된다. 예를 들어, 특정 조절성 요소의 선택은 T87Q 항-겸상화 글로빈 또는 성인 글로빈의 발현을 개선 또는 증가시킬 수 있다. 일부 실시양태에서, 다음 요소 중 하나 이상은 더 높은 수준의 HDR의 구동을 조정하고, 글로빈 발현을 증가시킨다: SV40 폴리A 테일, HPFH-2 인핸서, 및/또는 wPRE-3 요소.
- [0151] 일부 실시양태에서, 인간 코돈 최적화 겸상 보정 카세트의 전달은 겸상 환자에서 기능적 HbA 헤모글로빈을 복원하는데 도움이 된다. 일부 실시양태에서, 근접한 위치에 천연 인트론 1을 갖는 것은 최대 글로빈 발현을 허용한다.
- [0152] 실시예 3-대안적인 공여자 주형 전달 방법을 사용한 HSC에서 HBB 유전자에서 상동성-지정 복구의 생체내 결과
- [0153] 실험 프로토콜
- [0154] rAAV6 생산: rAAV6 스톱을 생산하였다. rAAV6 벡터, 혈청형 헬퍼 및 HgT1-아데노 헬퍼 플라스미드를 HEK293T 세포로 형질감염시켰다. 세포를 48시간에 수확하고 용해시키고 벤조나제로 처리하였다. 아이오덕사놀 밀도 구배를 사용하여 재조합 rAAV6 계층을 갖는 비리온을 정제하였다. ITR 특이적 프라이머 및 프로브를 사용하여 rAAV6 계층의 qPCR-기반 역가를 결정하였다. 1%, 2% 및 3%의 배양 부피를 사용하여 rAAV6을 mPBSC에 형질도입

하였다.

- [0155] CD34⁺ 조혈 줄기 세포: 동결된 mPBSC를 프레드 허친슨 암 연구기관의 혈액학 우수 협력 센터(Cooperative Center for Excellence in Hematology at Fred Hutchinson Cancer Research Institute)(미국 워싱턴주 시애틀)에서 구입하였다.
- [0156] sgRNA 및 TALEN 설계: CRISPR 설계 도구 (<http://crispr.mit.edu/> 및 <http://crispor.tefor.net/>)를 사용하여 겸상 돌연변이에 가깝게 절단될 것으로 예측된 가이드를 설계하였다. 모든 가이드를 처음 3개의 5' 및 3' 말단 잔기에서 3' 포스포로티오에이트 뉴클레오티드간 연결을 갖는 화학적으로 변형된 2'-O-메틸 유사체로서 합성하였다 (신테고 인크.(Synthego Inc.), 미국 캘리포니아주). 겸상 돌연변이에서 절단하는 TALEN을 골든 게이트(Golden Gate) 클로닝 전략으로 어셈블리하였다. TALEN mRNA를 이전에 공개된 프로토콜 (Grier, AE, et al., (2016). Mol Ther Nucleic Acids 5: e306; 그 전문이 본원에 참조로 명시적으로 포함됨)에 기초하여 생산하였다.
- [0157] 전기천공, 세포의 형질도입 및 적혈구계 분화 배양: Alt-R S.p Cas9 뉴클레아제 3NLS 단백질을 모든 연구에서 사용하였다 (인테그레이티드 DNA 테크놀로지스 인크.(Integrated DNA Technologies Inc.), 미국 아이오와주 코랄빌). CD34⁺ 세포를 각각 100 ng/ml의 FLT-3 리간드, TPO, hSCF 및 IL-6 (페프로텍(Peprotech), 미국 뉴저지주 로키힐)을 갖는 SCGM 배지 (셀제닉스(CellGenix), 미국 뉴햄프셔주)에서 배양하였다. 1300 V, 20 밀리초 및 1 펄스에서 NEON 전기천공 시스템 (써모피셔 사이언티픽(ThermoFischer Scientific), 미국 매사추세츠주 월섬) 또는 론자 4-D 뉴클레오펙터 (론자(Lonza), 스위스 바젤, CM149 프로토콜)를 사용하여 해동 48시간 후 세포를 전기천공하였다. 20 pmol의 Cas9 및 50 pmol의 sgRNA (2 x 10⁵ 개 세포당, 1:2.5의 비율의 Cas9: sgRNA)를 혼합함으로써 전기천공 또는 뉴클레오펙션 직전에 Cas9 RNP를 만들었다. RNP 혼합물을 신선하게 만들고, 실온에서 15분 동안 인큐베이션하였다. ssODN 공여자 주형을 2 x 10⁵ 개 세포마다 100, 50, 25, 12.5 pmol로 사용하고, 전기천공 또는 뉴클레오펙션 직전에 RNP의 혼합물에 첨가하였다. 전기천공 또는 뉴클레오펙션 후 세포를 ssODN 처리된 및 대조군 세포에 대해 시토카인을 갖는 rAAV6 함유 SCGM 배지 (1%, 2% 또는 3% 배양 부피; 3% GTC rAAV6, 대략 4500-5100의 MOI, 1% GAA rAAV6, 대략 2190의 MOI) 또는 시토카인을 갖는 단순 SCGM 배지에 첨가하였다. 세포를 배지에서 밤새 37°C에서 18시간 동안 인큐베이션하였다. 18시간 후, 세포를 1 ng/ml hIL-3, 2 IU/ml EPO, 20 ng/ml h-SCF, 20% 열-불활성화된 FBS 및 1% pen/strep을 갖는 IMDM 배지를 함유하는 조직 배양 비-처리 플레이트로 옮겼다. (피셔 사이언티픽(Fisher Scientific)(미국 뉴햄프셔 햄프턴) 및 페프로텍(미국 뉴저지주 로키힐)). 증식 스트레스 또는 과밀로 인한 태아 헤모글로빈 유도를 최소화하기 위해 세포 밀도를 5 x 10⁵ 내지 1 x 10⁶ 개 세포/ml 사이에서 유지하였다. CD235 발현을 BV421-표지된 글리코포린-A 항체 (BD, 562938)를 사용하여 유동 세포계측법에 의해 14일차에 모니터링하였다.
- [0158] ddPCR을 사용하여 rAAV6 및 ssODN으로 HDR 사건 측정: gDNA를 DNeasy 혈액 및 조직 키트 (퀴아젠(Qiagen), 미국 메릴랜드주 저먼타운)로 추출하고 RNase-처리하였다. 100 ng의 gDNA를 6 단위의 ECORV-HF (뉴 잉글랜드 바이오랩스(New England Biolabs), 미국 매사추세츠주 Ipswich)로 37°C에서 15분 동안 처리하고, 앰플리콘 영역의 외부의 gDNA를 절단하였다. ddPCR 정방향 및 역방향 프라이머 (ddPCR F/R)를 사용하여 210 bp 앰플리콘을 증폭시켰다. WT-HEX 및 HDR-FAM 프로브 실행과 함께 이중 프로브 검정으로서 검정을 설계하고, 프로브용 ddPCR 수퍼믹스 (dUTP 없음, 바이오-래드(BIO-Rad))를 사용하여 동일한 ddPCR F/R 프라이머와 함께 별도의 웰에서 병렬로 참조-HEX 프로브를 실행하였다. 표 3 및 표 4는 프라이머 및 프로브를 나열한다.

[0159] <표 3>

프라이머	정방향 (SEQ ID NO.)	역방향 (SEQ ID NO.)
HBB-1250	AGGCTTTTGTGTTCCCCAGA (SEQ ID NO:39)	AGCCTTCACCTTAGGGTTGC (SEQ ID NO:40)
SCL-386	GGGTGGCCAATCTACTCCC (SEQ ID NO:41)	CCTCTGGGTCCAAGGGTAGA (SEQ ID NO:42)
ddPCR	CATAAAAGTCAGGGCAGAG (SEQ ID NO:43)	GTCTCCTTAAACCTGTCTTG (SEQ ID NO:44)
LINC01206	CAAAAAGCAAAATTTGGGGATA (SEQ ID NO:45)	CTTTTAGCCCAGTGCCAGAC (SEQ ID NO:46)
MIR7974	ATCAGCCCCCTCTTTCTGGAT (SEQ ID NO:47)	AGTGCAGTGGTGCCATCATA (SEQ ID NO:48)
HBD	CAGATCCCCAAAGGACTCAA (SEQ ID NO:49)	GCGGTGGGGAGATATGTAGA (SEQ ID NO:50)
TULP4	CACGCCAGGATGTAAGCTCT (SEQ ID NO:51)	TCTGAGGCAAAAGTGCAAGA (SEQ ID NO:52)
DENND3	GGGGGTTTCTATCCCTCACT (SEQ ID NO:53)	CAAGAGGGTCAGGTTGAGGA (SEQ ID NO:54)
HBB-Miseq (어댑터 포함)	TCGTCGGCAGCGTCAGATGTGT ATAAGAGACAGGGTTGGCCAA TCTACTCCC (SEQ ID NO:55)	GTCTCGTGGGCTCGGAGATGTG TATAAGAGACAGCCTCTGGGTC CAAGGGTAGA (SEQ ID NO:56)
HBD-Miseq (어댑터 포함)	TCGTCGGCAGCGTCAGATGTGT ATAAGAGACAGCACAACTAAT GAAACCCTGCT (SEQ ID NO:57)	GTCTCGTGGGCTCGGAGATGTG TATAAGAGACAGTCTACACATG CCCAGTTTCCA (SEQ ID NO:58)

[0160]

[0161] <표 4>

프로브	서열 (SEQ ID NO.)
GTC HDR FAM	CTCCTGTCGAGAAGTCTGC (SEQ ID NO:59)
GAA HDR FAM	CTCCCGAAGAGAAGTCTGC (SEQ ID NO:60)
GTG HDR FAM	CTCCTGTGGAGAAGTCTGC (SEQ ID NO:61)
GAG WT HEX	TGACTCCTGTCGAGAAGT (SEQ ID NO:62)
REF HEX	GTTCACTAGCAACCTCAAACAGACACC (SEQ ID NO:63)

[0162]

[0163] 액적을 생성하고 바이오-래드(BIO-RAD) 서모사이클러에서 증폭시켰다 (95℃: 5분, 94℃: 30초, 56℃: 1분, 72℃: 1분, 단계 2로 이동: 49회 사이클, 98℃: 10분, 12℃: ∞). 바이오-래드 QX200 기계 (바이오-래드(BIO-RAD), 미국 캘리포니아주 허큘리스)에서 FAM 및 HEX 형광 세기를 측정하였다. HDR (%) 사건 (HDR-FAM⁺) 및 WT (WT-HEX⁺) 사건을 참조 유전자 (REF-HEX⁺, 표 3)에 대한 보정 후 계산하였다.

[0164] INDEL 빈도 측정: 전기천공 후 10일차로부터 gDNA를 사용하여, 정방향 및 역방향 프라이머로 절단 부위 주변의 1250 bp 앰플리콘을 증폭시켰다 (HBB-F/R-1250, 표 3). 뉴클레오스핀(NucleoSpin) 겔 및 PCR 클린-업 키트 (마커리 네이겔(Machery Nagel), 미국 펜실베이니아주 베들레헴)를 사용하여 PCR 생성물을 세정하고, 시퀀싱 프라이머 (SCL-F/R-386, 표 3)로 생어(Sanger) 시퀀싱을 수행하였다. 편집 후 INDEL을 측정하기 위해 TIDE/ICE 알고리즘을 사용하여 서열을 분석하였다.

[0165] MiSeq 분석: MiSeq 프라이머 (표 3)와 PrimeSTAR GXL DNA 폴리머라제 (다카라(TaKaRa), 일본 쿠사츠)를 사용하

여 200 ng의 gDNA로부터 HBB (386 bp) 및 HBD (301 bp) 유전자-특이적 앰플리콘을 증폭시켰다. 프라이머는 앰플리콘에 오버행 어댑터 서열을 첨가하였다. 넥스테라(Nextera) 96-인덱스 키트 (FC-121-1012, 일루미나 (Illumina), 미국 캘리포니아주 샌디에고)를 사용하여, 각 샘플에 5' 및 3' 고유 인덱스를 첨가하였다. 샘플을 에이겐코트(Agencourt) AMPure XP (백크만 쿨터(Beckman Coulter), 미국 캘리포니아주 브레아)로 정제하고, 밴드를 아가로스/PAGE 겔에서 확인하였다. 샘플을 측정하고 풀팅하여 라이브러리들을 만들고, 품질 제어를 큐비트(Qubit) (쓰모피서 사이언티픽, 미국 매사추세츠주 월섬)에서 수행하고, MiSeq 500 CycleV2 키트 (일루미나, 미국 캘리포니아주 샌디에고)에서 분석하였다. 크리스프레스2(Crispresso2) 알고리즘을 사용하여 데이터를 마이닝하였다. HBB 분석을 온-타겟 유전자 변형에 사용하고, HBD를 오프-타겟 분석에 사용하였다.

[0166] NBSGW 마우스에서의 생학 연구: NOD, B6, SCID Il2r γ $^{-/-}$ Kit(W41/W41) (NBSGW) 마우스를 잭슨 래보러토리즈 (Jackson Laboratories)에서 구입하고, 지정된 무병원체 시설에서 유지시켰다. 모든 동물 연구를 실험 동물 관리 평가 인증 협회(Association for Assessment and Accreditation of Laboratory Animal Care) 표준에 따라 수행하고, SCRI 동물실험윤리위원회(Institutional Animal Care and Use Committee)의 승인을 받았다.

[0167] 6-7주령 NBSGW 마우스를 편집된 세포의 이식 24시간 전에 부설판 (셀렉캠(Selleckchem))으로 처리하였다. 2 x 10⁶ 개의 편집된 세포를 편집 24시간 후에 꼬리 정맥으로 주입하였다. 동물을 규칙적으로 모니터링하였다. 이들 동물로부터의 BM 및 비장을 전달 3주 및 12-14주 후 수확하고, 세포를 인간 키메라즘 hCD45⁺, mCD45⁺ 및 CD19⁺, CD33⁺, CD235⁺, CD3⁺, CD34⁺, CD38⁺ 세포의 다중-계통 생학에 대해 분석하였다. BM 세포로부터의 gDNA를 수확하고, ddPCR에 의해 분석하여 HDR (%) 및 WT (%)를 결정하였다. Indel을 TIDE/ICE 시퀀싱에 의해 분석하였다. BM 세포를 수확 후 2주 동안 적혈구계 분화 배지에서 배양하였다. 생체의 분화 배양으로부터의 세포를 유동 세포계측법에 의해 CD235⁺ 발현에 대해 측정하였다. 세포를 또한 펠렛화하고 세척하고 수확 2주 후 RP-HPLC에 의해 분석하여 글로빈 발현을 조사하였다. BM 세포 (30,000개 세포/플레이트/3 ml의 메토컬트)를 메토컬트 완전 배지 (스텝셀 테크놀로지스(STEMCELL technologies), 캐나다 밴쿠버)에 첨가하고, CFU 분석을 위해 플레이트팅하였다. 단일 BFU-E 콜로니를 수확 14일 후 골라내서, 물에서 용해시키고, 글로빈 발현에 대해 IEC에 의해 분석하였다.

[0168] 통계적 분석: 실험으로부터 수집된 데이터를 던넛의 다중 비교 시험과 이원 ANOVA 분석을 사용하여 그래프패드(GraphPad) 프리즘(Prism) 7에서 분석하였다. 그룹에 걸친 모든 샘플을 대조군 또는 모의 처리된 세포와 비교하여 유의성을 평가하였다. ns: 유의하지 않음, * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

[0169] 적혈구계 세포 용해: 분화 배지에서 14일 동안 배양된 적혈구계 세포를 수집하고, PBS에서 세척하여 오염 단백질을 제거하였다. HPLC 등급 물에서 세포의 저장성 용해를 수행하였다. 용혈물의 상등액을 20,000 g에서 30분 동안 4℃에서 원심분리하고, 1 - 10 μ g의 단백질을 컬럼에 주사하였다.

[0170] 적혈구계 세포의 RP-HPLC 분석: 적혈구계 분화 후, 글로빈 하위유형의 발현을 에어리스(Aeris) 3.6 μ m 와이드포어(Widepore) C4 250 x 4.6 mm 컬럼 (페노메넥스)을 사용하여 시마즈(Shimadzu) 프로미넌스(Prominence) UFLC 크로마토그래프에서 RP-HPLC에 의해 평가하였다. 사용된 이동상은 다음과 같았다: A: 물 0.1% TFA (트리플루오로아세트산), B: 0.8 ml/분의 유속에서 아세트오니트릴 0.08% TFA. 39%로부터 50% B까지의 구배를 75-분 타이밍된 프로그램을 통해 실행하였다. 컬럼 오븐 온도는 40℃이고, 샘플 트레이를 4℃에서 유지시켰다. 220 nm에서 피크가 검출되었다. 다양한 글로빈 피크의 용출 시간을 비교하기 위해 참조를 실행하였다.

[0171] 적혈구계 세포의 IEC: PBS 세척 후 세포를 다음 이동상을 사용하여 PolyCATA 200 x 2.1 mm 5 μ m 1000 Å (PolyC#202CT0510)에서 분석하였다: A상: 아세트산을 사용하여 pH 6.5로 조정된 HPLC 등급 물 중 Tris 40 mM, KCN 3 mM, B상: 아세트산을 사용하여 pH 6.5로 조정된 HPLC 등급 물, NaCl 0.2 M 중 Tris 40 mM, KCN 3 mM. 타이밍된 24-분 프로그램을 사용하여 0.3 mL/분의 유속으로 2% 내지 100% B 구배를 생성하였다. 컬럼 오븐 온도는 30℃이고, 샘플 트레이를 4℃에서 유지시켰다. 418 nm에서 피크가 검출되었다. 글로빈 사량체의 용출 시간을 비교하기 위해 참조를 실행하였다.

[0172] 콜로니 시퀀싱: GXL DNA 폴리머라제 (다카라 바이오(Takara Bio))를 사용하여 50 ng의 gDNA로부터 HBB-1250 정방향 및 역방향 프라이머 (표 3)로 절단 부위 주변의 1250 bp 앰플리콘을 증폭시켰다. PCR 생성물을 뉴클레오스핀 겔 및 PCR 클린-업 키트 (마커리-네이젤, 미국 펜실베이니아주 베들레헴)를 사용하여 정제하고, 제로 블런트(Zero Blunt) TOPO PCR 클로닝 벡터 (피서 사이언티픽, 미국 뉴햄프셔 햄프턴)로 서브클로닝하고, TOP10 감응성 세포 (피서 사이언티픽, 미국 뉴햄프셔 햄프턴)로 형질전환시켰다. 카나마이신-내성 콜로니를 골라내고, SCL-386 프라이머로 시퀀싱하였다. 서열이 WT, NHEJ 또는 HDR 결과인지 결정하기 위해 개별 서열을

분석하였다.

[0173] T7-엔도뉴클레아제 검정: GXL DNA 폴리머라제 (다카라 바이오)를 사용하여 HBB-1250 프라이머를 사용하여 총 gDNA로부터 뉴클레아제 절단 부위 주변의 1250 bp 영역을 증폭시켰다. PCR 생성물을 뉴클레오스핀 겔 및 PCR 클린-업 키트 (마커리-네이겔, 미국 펜실베이니아주 베들레헴)를 사용하여 정제하였다. 400 ng의 PCR 생성물을 19 ul 반응 부피에서 1x 완충액 2 (뉴 잉글랜드 바이오랩스, 미국 매사추세츠주 입스위치)에서 변성시키고 재어닐링하였다. 샘플을 T7 엔도뉴클레아제 I (뉴 잉글랜드 바이오랩스, 미국 매사추세츠주 입스위치)로 처리하고, 37°C에서 15분 동안 인큐베이션한 후, 1% 아가로스 겔에 로딩하고, 이미지화하였다.

[0174] 유동 세포계측법 및 분석: 유동 세포계측법 분석을 LSR II 유동 세포계측기 (BD 바이오사이언시스(BD Biosciences))에서 수행하고, 플로우조(FlowJo) 소프트웨어 (트리스타(TreeStar))를 사용하여 데이터 분석을 수행하였다. 살아있는 세포에 상응하는 FSC/SSC 집단에 게이트를 그리고, FSC-A/FSC-W를 사용하여 일중선을 그렸다.

[0175] ssODN 설계: 5' 및 3' 말단 상의 2개의 말단 뉴클레오티드에서 포스포로티오에이트 연결을 사용하여 IDT (울트라머(Ultramer)® DNA 올리고뉴클레오티드)에 의해 단일 가닥 올리고뉴클레오티드 (ssODN)를 상업적으로 합성하였다. 표 5는 HDR에 사용된 ssODN 서열을 나열한다.

[0176] <표 5>

ssODN 서열	서열 (SEQ ID NO.)
E6V GAG > GTC	TCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGACACAACGTGTTTCACTAGCAA (SEQ ID NO:64)
	CCTCAAACAGACACCATGGTGCATCTGACTCCTGTCGAGAAGTC TGCCGTTACTGCCCT (SEQ ID NO:65)
	GTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGT (SEQ ID NO:66)
E6V GAG > CCC GAA	TCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGACACAACGTGTTTCACTAGCAA (SEQ ID NO:67)
	CCTCAAACAGACACCATGGTGCATCTGACTCCCGAAGAGAAGTC TGCCGTTACTGCCCT (SEQ ID NO:68)
	GTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGT (SEQ ID NO:69)
E6V GAG > GTG	TCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGACACAACGTGTTTCACTAGCAA (SEQ ID NO:70)
	CCTCAAACAGACACCATGGTGCATCTGACTCCTGTGGAGAAGTC TGCCGTTACTGCCCT (SEQ ID NO:71)
	GTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGT (SEQ ID NO:72)
밑줄 친 서열은 야생형 서열로부터의 변화이다	

[0177]

[0178] CD34 세포에서 뉴클레아제 효율

[0179] 뉴클레아제 스크린을 수행하여, HBB 유전자의 엑손 1 내에 이중-가닥 파단 (DSB)을 효율적으로 생성하는 방법을 확인하였다 (도 10a). RNP 복합체 또는 TALEN-기반 뉴클레아제로서 전달된 대안적인 sgRNA의 절단 효율을 CD34+ mPBSC에서 평가하였다. 초기 뉴클레아제 스크린에서, 일련의 후보 sgRNA의 RNP 전달을 1:1의 Cas9:sgRNA 비율에서 시험하고, 가이드 4 (g4), g5, g6, g1을 DSB 생성에 가장 효율적인 것으로 확인하였다 (도 10b). 이들 발견에 기초하여, SCD 돌연변이의 부위인 코돈 6에 인접한 DSB를 생성하여 sgRNA-g1을 최적화하였다. sgRNA-g6 (G10)을 또한 병렬로 광범위하게 시험하였다. neon 전기천공 시스템으로 Cas9: sgRNA 1:2.5의 비율에서 두 가이드를 시험한 결과, sgRNA-g1에 대한 총 편집 비율이 2배가 되었다 (g1: 17.8 ± 4.4%에서 35.2 ± 10.6%로 증가됨; g6: 26.7 ± 1.6%에서 38.3 ± 8.7%로 증가됨, 도 10c). MiSeq 분석에 의한 sgRNA-g1에 대한 온-타겟 HBB 파괴는 38.7 ± 12.2% (n = 7 공여자)였고, 오프-타겟 HBD 파괴는 0.129 ± 0.01% (도 10d)

었다. sgRNA-g1에 대해 CCTop 22에 의해 예측된 상위 5개의 오프-타겟 유전자는 T7 엔도뉴클레아제 검정 (도 10f) 및 TIDE 시퀀싱 (도 10g)에 의해 Indel을 나타내지 않았다. 전체 편집 비율은 전기천공 시스템 (Neon; $35.2 \pm 10.6\%$, $n = 15$ 공여자, 도 10e)과 비교하여 뉴클레오펙션 시스템 (론자; $86 \pm 2.6\%$, $n = 3$ 공여자)의 사용으로 대략 2.5배만큼 증가하였다.

[0180] rAAV6 공여자 주형 전달을 통한 GTC 변화의 도입

[0181] HBB 유전자의 엑손 1의 코돈 6에서 GTC (E6V를 코딩함) 또는 침묵 변화 GAA (E6optE를 코딩함)를 도입하도록 설계된 2.2 kb 상동성 아암으로 rAAV6 벡터를 구축하였다. 설계는 전사 및 번역을 최대화하기 위해 인트론 1 및 천연 프로모터/인핸서 영역을 보존하는데 중점을 두었다 (도 11a). 실험 타임라인은 도 11b에 나타난다. RNP-매개 절단 후 3% GTC (E6V를 코딩함) rAAV6 공여자 주형을 시험한 결과, HDR 비율은 $37.5 \pm 15\%$ 이고 잔류 NHEJ 비율은 $12.7 \pm 5.3\%$ 였다 (도 11c). 전기천공 및 뉴클레오펙션 시스템 둘 모두를 사용한 GTC (E6V를 코딩함) rAAV6의 시험은 총 편집 비율의 증가가 HDR 및 잔류 NHEJ 둘 모두의 비율의 증가를 초래하는 것으로 나타났다 (도 11g). RNP가 GTC rAAV6과 함께 공동전달될 때 측정된 HDR 및 잔류 NHEJ 비율을 콜로니 시퀀싱에 의해 추가로 검증하였다 ($30.8 \pm 6.3\%$ HDR, $17.9 \pm 7.2\%$ NHEJ, $n = 5$ 공여자, 도 11h). 3% 배양 부피의 GTC rAAV6을 사용한 전기천공 및 형질도입 후 세포 생존율은 평균 79.2% 였다 (도 11i). 역상 HPLC (RP-HPLC)에 의해 분화된 적혈구계 전구체에서 글로빈 하위유형을 측정하였다. 단독 편집 및 GTC rAAV6의 존재하에 편집은 β A의 유의한 감소 (82% 내지 $54 \pm 16\%$, neon을 사용함) 및 γ A (HBG1) 및 γ G (HBG2)의 2배 증가를 유발하였다. RNP 및 3% GTC rAAV6의 공동전달은 $28.4 \pm 9.8\%$ β S 발현을 유발하였다 ($n = 6$ 공여자, 도 11d). RNP 및 GTC rAAV6 (E6V를 코딩함)의 공동전달 후 생성된 적혈구계 세포에서 글로빈 사량체를 이온 교환 크로마토그래피 (IEC)에 의해 측정하였다. RNP 및 3% GTC rAAV6의 공동전달로 HbA의 감소 및 HbS 사량체의 용량-의존적 증가 (15.8%)가 관찰되었다 (도 S2D). RP-HPLC 분석에 의한 GTC (E6V를 코딩함) rAAV6 처리된 세포의 샘플 크로마토그램은 38.6% β s 피크의 존재를 확인하였다 (도 11k).

[0182] 이들 뉴클레오티드 변화 연구와 병행하여, 대안적인 조직-특이적 인핸서 및 조절인자와 연관하여 GFP 발현 카세트를 도입하도록 설계된 더 복잡한 rAAV6 공여자 주형 구축물에 대한 광범위한 시험을 또한 수행하였다 (도 11l). 이들 후자의 rAAV6 공여자 주형은 유의한 HDR 비율을 초래하였지만 (1321: 14.4% 및 1322: 18.4% , 도 11m), 원하는 글로빈 발현은 절충되었다 (β s = 1.35 내지 2.6% , 도 11n).

[0183] rAAV6 공여자 주형 전달을 통한 GAA SNP 변화의 도입

[0184] 정상 세포에서 겸상 돌연변이의 도입은 환자 세포에서 돌연변이를 복귀시킬 잠재력을 평가하지 않으며, 또한 편집된 적혈구계 선조의 적합성을 변경할 수 있다. SCD 대상체로부터의 HSC를 사용하는 연구 대신에, 침묵 SNP 변화 (GAA; E6optE를 코딩함)의 도입을 시험하였다. RNP와 함께 공동전달된 GAA (E6optE를 코딩함) rAAV6 공여자 주형 (1% 배양 부피)을 시험한 결과, HDR 비율은 $37.5 \pm 6\%$ 이고 NHEJ 비율은 $43.7 \pm 11.5\%$ 였다 (도 11e). 중요하게는, RNP 및 GTC (E6V를 코딩함) rAAV6 카세트의 공동전달과 비교하여 RNP 및 GAA (E6optE를 코딩함) rAAV6의 공동전달을 사용한 총 NHEJ 사건의 증가는 뉴클레오펙션 시스템을 사용한 전체 편집 비율의 증가를 반영할 가능성이 높다. GTC (E6V를 코딩함) rAAV6 편집에 대한 데이터는 Neon ($n = 3$) 및 론자 ($n = 1$) 둘 모두를 사용한 실험을 포함하였지만, GAA (E6optE를 코딩함) rAAV6 공동전달을 론자 ($n = 3$)에서만 배타적으로 시험하였다. 1% GAA (E6optE를 코딩함) rAAV6을 사용한 전기천공 및 형질도입 후 세포 생존율은 평균 60% 였다 (도 11i). RP-HPLC 분석은 β A 수준의 현저한 감소 (대조군 세포에서 82% 로부터 RNP 편집된 세포에서 $16.9 \pm 15\%$) 및 RNP-편집된 샘플에서 β A (HBG1) 및 γ G (HBG2)의 3배 증가를 확인하였다. 대조적으로, RNP 및 1% GAA (E6optE를 코딩함) rAAV6의 공동전달은 β A 수준의 덜 확고한 감소 ($54.2 \pm 10\%$ β A 발현; $n = 3$, 도 11f) 및 β A (HBG1) 및 γ G (HBG2)의 덜 현저한 증가를 유발하였다. RNP 및 GAA (E6optE를 코딩함) rAAV6의 공동전달 후 β A 발현의 유지는 AAV-매개 HDR에 기인할 수 있다. 이러한 결론과 일치하여, RNP 및 GAA (E6optE를 코딩함) rAAV6의 공동전달로 처리된 세포의 RP-HPLC 분석은 HDR 후 64.7% β A를 나타내었다 (도 11o). 종합하면, 본 발명자들의 발견은 환자 세포에서 겸상 돌연변이를 복귀시키도록 설계된 침묵 돌연변이 또는 겸상 돌연변이 도입을 유도하는, HBB의 엑손 1에서 높은-수준의 HDR을 촉진하는 RNP 및 rAAV6 공동전달의 능력을 입증하였다.

[0185] ssODN 공여자 주형 전달을 통한 GTC 변화의 도입

[0186] mPBSC에서 rAAV6을 사용하여 달성된 동일한 뉴클레오티드 변화를 도입하는 RNP 및 ssODN의 공동전달의 효율을 평가하는 연구를 수행하였다. GTG (E6V를 코딩함), GTC (E6V를 코딩함) 또는 GAA (E6optE를 코딩함) 뉴클레오티드 변화를 생성하도록 대안적인 168 bp ssODN을 설계하였다 (도 12a). 실험 타임라인은 도 12b에 나타난다.

시험된 ssODN의 농도가 증가함에 따라 세포독성의 용량-의존적 증가가 있었다 (도 12g). RNP 및 50 pmol의 ssODN의 공동전달 후 HDR 유전자 전환 비율은 GTC ODN의 경우 $11.9 \pm 3.4\%$ 및 GTG ODN의 경우 $17 \pm 4.3\%$ 였고, 잔류 NHEJ는 각각 $17.4 \pm 17.5\%$ 및 $20.0 \pm 1.7\%$ 였다 (도 12c, 도 12h). 50 pmol의 GTG ssODN (E6V를 코딩함)에 대한 HDR 및 NHEJ 비율을 콜로니 시퀀싱에 의해 추가로 검증하였으며, 각각 $12.6 \pm 8.8\%$ 및 $30.1 \pm 12.4\%$ 였다 (도 12i). 글로빈 하위유형을 RP-HPLC에 의해 평가하였다. RNP-매개 파괴 후 β A의 유의한 감소 및 1.5배 증가 β A (HBG1) 및 γ G (HBG2)가 관찰되었다. RNP 및 다양한 농도의 ssODN의 공동전달은 50 pmol의 GTC (E6V를 코딩함) ssODN (도 12d) 및 GTG ssODN (도 12j)으로 최적의 β S 발현과 함께 겸상 글로빈 발현의 용량-의존적 증가를 초래하였다. GTC (E6V를 코딩함) ssODN을 사용한 편집은 5.2% β S 발현을 초래하고 (50 pmol, n = 5, 도 3d), GTG (E6V를 코딩함) ssODN은 5.3% β S 발현을 각각 초래하였다 (50 pmol, n = 3) (도 12j). 이들 평균과 일치하여, 분화된 적혈구계 세포로부터 유도된 샘플 크로마토그램은 GTC (E6V를 코딩함) ssODN의 경우 8.9% β S 발현 (도 12l) 및 GTG (E6V를 코딩함) ssODN의 경우 9.2% β S (도 12m)를 나타내었다. GTC ssODN vs. rAAV6을 사용하여 동일한 공여자로부터 mPBSC에서 편집의 직접 비교는 각각 8.9% vs. 24.5% β S 발현을 나타내었다 (도 12l).

[0187] ssODN 공여자 주형 전달을 통한 GAA SNP 변화의 도입

[0188] rAAV6 공여자 주형을 사용한 연구와 일치하여, 대안적인 침묵 SNP 변화 (GAA; E6optE를 코딩함)의 도입을 시험하였다. RNP 및 50 pmol의 GAA (E6optE를 코딩함) ssODN의 공동전달은 $44 \pm 13.8\%$ 의 잔류 NHEJ 비율과 함께 $24.5 \pm 7.6\%$ 의 HDR 유전자 전환 비율을 초래하였다 (도 12e). ssODN의 농도가 증가함에 따라, HDR의 용량-의존적 증가 및 NHEJ의 상응하는 감소가 Neon 및 론자 시스템 둘 모두에서 관찰되었다 (도 12k). Neon 전기천공 시스템의 사용 후 편집 결과를 콜로니 시퀀싱에 의해 또한 검증하였으며, $10.6 \pm 2.8\%$ 의 HDR 비율 및 $35.5 \pm 8.6\%$ 의 잔류 NHEJ 비율을 나타내었다 (도 12i). 분화된 적혈구계 펠렛에서 글로빈 하위유형을 측정하고, β A의 유의한 감소 (25.7%, n = 6 공여자) 및 β A (HBG1) 및 γ G (HBG2)의 1.5 내지 3배 증가가 RNP-매개 파괴와 함께 관찰되었다. 대조적으로, GAA (E6optE를 코딩함) ssODN 공여자 주형의 사용은 58.4%에서 β A 발현을 유지하였다 (n = 6 공여자, 도 12f). 편집된 분화된 적혈구계 세포에서 글로빈 하위유형을 나타내는 샘플 크로마토그램은 RNP 및 GAA (E6optE를 코딩함) ssODN의 공동전달 후 RNP 파괴된 샘플에서 0% HbA로부터 75.6%로의 증가를 입증하였다 (도 12n). RNP 및 HDR 공여자 주형의 공동전달을 사용하여 편집된 mPBSC와 RNP-단독 편집된 세포의 직접 비교는 GAA ssODN 및 rAAV6 공여자 주형의 경우 각각 0%로부터 75.6% 및 64.7%로의 HbA 발현의 증가를 나타내었다 (도 12n).

[0189] MiSeq 분석에 의한 ssODN 및 rAAV6 공여자 주형 전달 방법의 비교

[0190] 본 발명자들의 대안적인 플랫폼을 사용하여 달성된 유전자 편집 효율을 추가로 평가하기 위해, MiSeq 분석을 사용하여 편집 결과를 검증하였다. 모든 ssODN 공여자 주형 vs GTC (E6V를 코딩함) rAAV6 공여자 주형과 관련하여 RNP 공동전달을 위해 Neon 전기천공 시스템을 사용하여 달성된 HDR 및 NHEJ 비율을 평가하였다. 각 시험관 내 (및 생체내) 샘플로부터 평균 113,000개의 쌍별 정렬된 판독을 얻었다 (도 13d). 데이터는 시험관내에서 rAAV6 공여자 주형이 NHEJ보다 더 높은 수준의 HDR을 구동하였고 (GTC rAAV6: 27.8% HDR, 16% NHEJ), ssODN 전달이 HDR보다 더 높은 수준의 NHEJ를 구동하였음 (GTC ssODN: 14.3% HDR 및 19.6% NHEJ)을 입증하였다 (도 13a, 도 13b, 도 13c). 생성된 indel 스펙트럼을 분석하여, RNP 단독은 2% 삽입과 함께 60.4% 결실 (주로 -3, -1, -5, -6 및 -12 bp 결실)을 초래하였다. RNP와 공여자 주형의 공동전달은 indel 스펙트럼을 주로 -3 및 -1 bp 결실로 감소시켰다 (도 13b, 도 13e). 야생형 (WT), 결실을 갖는 NHEJ 및 HDR 대립유전자가 rAAV6-편집된 및 ssODN-편집된 샘플에서 관찰되었다 (도 13c). 크리스프레소 27 분석은 rAAV6 공여자 주형 전달이 ssODN 공여자 주형 전달 (시험관내 12.2% 및 5.3% 생체내; 도 13f)과 비교하여 더 적은 프레임 이동 돌연변이 (8.6% 시험관내 및 1.4% 생체내)를 초래함을 확인하였다.

[0191] 생체내 HDR-편집된 세포의 지속적인 생착에 대한 ssODN vs. rAAV6 전달의 영향

[0192] HDR-편집된 CD34+ 세포의 장기간 생착 잠재력의 변경에서 대안적인 공여자 주형 플랫폼의 역할을 이해하기 위해, GTC (E6V를 코딩함) 변화를 도입하도록 편집된 건강한 대조군 mPBSC를, 인간 적혈구계 구획의 발달을 허용하는 면역결핍 균주인 부설판 컨디션닝된 (12.5 내지 25 mg/kg) NBSGW 수용자 마우스에 이식하였다. 각 플랫폼으로 편집된 동일한 공여자로부터 유래된 세포 (2×10^6 개 세포)를 전기천공 후 1일차에 이식하였다. 이식된 동물을 시간 경과에 따라 평가하고, BM 및 비장에서 인간 세포 생착에 대해 3 및 12-14주에 평가하였다 (도 14a).

[0193] 인간 키메리즘은 모의- 또는 ssODN-편집된 세포의 수용자에 필적하였다. 대조적으로, hCD45+ 세포 생착의 유의

한 감소가 rAAV6-편집된 세포의 수용자에서 관찰되었다 (도 14b). 또한, CD19+ B 세포의 비율은 rAAV6-편집된 그룹에서 완만하게 증가되었으며, 이는 더 분화된 자손을 향한 왜곡을 시사한다 (도 14c). 골수성 (CD33+), T 세포 (CD3+) 및 적혈구계 (CD235+) 세포를 포함하는 다른 계통을 코호트에 걸쳐 동등하게 표시하였다 (도 14c, 도 14d, 도 14j, 도 14k, 도 14l). BM으로부터 단리된 세포를 수확 후 2주 동안 적혈구계 분화 배지에서 배양하여 CD235+ 세포의 확장을 허용하였다 (수확에서 4.01%의 생체의 배양에서 27.6%로의 증가; 도 14d). 이식 전 및 이식 후 편집된 공여자 세포의 대표적인 유동 플롯은 다음을 포함하는 원시 HSC 하위집단의 동등한 비율을 나타내었다: CD34+; CD34+CD38lo, 및 CD34+CD38loCD133+CD90+ 세포 (도 14e, 도 14m, 도 14n). 4개의 이식에 걸친 투입 HDR 비율 (배양에서 14일차)은 rAAV6 및 ssODN 전달 방법의 경우 각각 $24.28 \pm 7.5\%$ 및 $17.5 \pm 6\%$ 였다. 이식 후 3주에 BM에서 HDR 편집된 세포는 rAAV6 및 ssODN 전달의 경우 각각 $13.58 \pm 0.16\%$ 및 $15.19 \pm 2.8\%$ ($n = 2$)로 감소하였다. 12-14주에, HDR 비율은 rAAV6 공여자 주형 편집된 세포의 수용자에서 $0.66 \pm 0.66\%$ ($n = 17$)로 급격히 감소하였다. 두드러지게는, HDR 비율이 또한 감소하였지만, ssODN 공여자 주형 편집된 세포의 수용자에서 $4.136 \pm 2.1\%$ ($n = 18$)로 훨씬 더 적은 정도로 감소하였다 (도 14f). 투입 NHEJ는 rAAV6 및 ssODN 공여자 주형 전달 방법의 경우 각각 $7 \pm 1.4\%$ 및 $13.5 \pm 3.7\%$ 였고, 이식 후 3주에 변화없이 유지되었고 (rAAV6: $9 \pm 3\%$, ssODN: $12.3 \pm 2.1\%$), 12-14주에 감소하였다 (NHEJ rAAV6: $1.3 \pm 0.85\%$, ssODN: $5 \pm 2.7\%$, 도 14g). 생체내 BM에서의 HDR 및 NHEJ 비율을 MiSeq 분석에 의해 검증하였다 (HDR: rAAV6: $0.65 \pm 0.65\%$, ssODN: $3.84 \pm 2.1\%$; NHEJ: rAAV6: $2.5 \pm 2.5\%$ 및 ssODN: $9.9 \pm 5.3\%$, 도 14h, 도 14i).

[0194] 각 초기 편집된 세포 집단의 서브셋을 적혈구계 배양 조건에서 시험관내에서 유지시키고, 글로빈 하위유형에 대해 분석하였다. rAAV6-편집된 세포는 $16.4 \pm 6.8\%$ 를 나타내고, ssODN-편집된 세포는 $12.42 \pm 4.4\%$ β S 발현을 나타내었다 (도 14o). HPLC에 의해 분석된 생체의 BM 배양물은 ssODN-편집된 그룹에서 3.8% β S ($n = 3$ 동물)를 나타내었다. 대조적으로, β S는 rAAV6 ($n = 4$ 동물) 또는 모의-편집된 샘플 ($n = 2$ 동물, 도 14p)에서 검출되지 않았다. 69개 BFU-E 콜로니의 HPLC는 β S를 발현하는 ssODN-편집된 그룹으로부터 유래된 3/35 콜로니가 평균 5.13% β S 발현을 유발함을 나타내었다. 대조적으로, β S 발현은 rAAV6-편집된 ($n = 26$ 콜로니) 또는 모의-편집된 콜로니 ($n = 8$ 콜로니, 도 14q, 도 14r)에서 검출되지 않았다. ssODN-편집된 그룹으로부터 유래된 단일 적혈구계 콜로니의 크로마토그램은 38.7% , 84.5% 및 56.3% 의 β S 발현 수준을 나타내었다 (도 14s, 도 14t). 모의 샘플에 대한 단일 콜로니의 HPLC 프로파일은 97% HbA를 함유한 반면, 편집된 그룹은 HbA의 감소 및 HbF (rAAV6: 17.4% , ssODN: 17.9%) 및/또는 HbS (도 14q, 도 14r, 도 14s, 도 14t)의 증가를 가졌다. 종합하면, 이들 연구는 ssODN-변형된 세포가 생체내에서 rAAV6-변형된 세포를 능가하여 HDR-편집된 세포의 더 높은 지속적인 생장 및 겸상 글로빈 발현 모두를 유발함을 입증하였다.

[0195] DNA 공여자 주형의 전달은 인간 조혈 줄기 세포에서 표적화된 유전자 절단 후 정확한 유전자 보정을 달성하는데 유용하다. HDR 대 NHEJ의 전체 비율은 겸상 적혈구 질환에서 유전자 보정의 잠재적 임상적 이점에 영향을 미쳤다. 본원에 개시된 연구에서, 시험관내 초기 유전자 전환 사건을 달성하기 위한 전달 방법 뿐만 아니라 생존, 줄기-유사 잠재력 및 생체내 편집된 세포의 지속적인 생장에 대한 영향에 대해 대안적인 공여자 주형의 역할을 평가하였다. 조합된 데이터는 복잡성을 입증하였고, SCD에서 장기간 임상 유전자 보정을 달성하는데 있어 몇 가지 문제를 해결하였다. HSC 생존율의 큰 차이가 관찰되지 않았지만, ssODN, rAAV6 공여자 주형과 비교하여 rAAV6을 사용한 시험관내 표현형 또는 확장은 일관적으로 더 높은 HDR: NHEJ 비율을 매개하는 것으로 나타났다. 대조적으로, 이식 실험에서, ssODN 공여자 주형으로 편집된 HSC를 사용하여 훨씬 더 높은 수준의 지속적인 HDR이 달성되었다.

[0196] 겸상 돌연변이 부위 주변의 53 bp 영역에 걸쳐있는 후보 TALEN 쌍 및 다중 후보 가이드 RNA의 초기 스크리닝을 수행하였다. 본원에 나타난 바와 같이, sgRNA-g1은 겸상 돌연변이 부위에 바로 인접한 DSB를 효율적으로 생성하였으며 (21-22 bp 사이), 따라서 돌연변이 부위로부터 16 bp 떨어진 DSB를 생성한 sgRNA-g6 (G10)보다 더 유용한 가이드로서 선택되었다. 1:2.5의 비율의 Cas9: sgRNA의 사용은 입증가능한 오프-타겟 효과 없이 인간 mPBSC에서 최고 수준의 편집을 촉진시켰다 (도 10d, 도 10f, 도 10g). 중요하게는, 총 편집 비율은 뉴클레오택션 시스템을 사용할 때 2배 초과 증가하였으며 (도 10e), 시험관내 잔류 indel 뿐만 아니라 HDR 둘 모두가 증가하였다 (도 11g, 도 12g). sgRNA-g1을 함유하는 RNP의 전달 후, HBB의 엑손 1의 6번째 코돈에서 1개 또는 2개의 뉴클레오타이드 변화를 구동하는 일련의 ssODN에 대한 대안적인 rAAV6 카세트의 능력을 시험하였다. 시험관내 연구는 rAAV6이 NHEJ보다 더 큰 비율의 HDR을 촉진함을 입증하였다 (GTC rAAV6: $37.5 \pm 15\%$ HDR 및 $12.7 \pm 5.3\%$ NHEJ, 도 11c). 대조적으로, ssODN 전달은 HDR보다 더 많은 NHEJ를 구동하였다 (GTC ssODN: $11.9 \pm 3.4\%$ HDR 및 $17.4 \pm 17.5\%$ NHEJ, 도 12c, 도 13a).

- [0197] 실시예 4-CM149 및 ER100 론자 방법의 비교
- [0198] CM149 및 ER100 론자 뉴클레오펙션 방법을 RNP 편집에 이어 rAAV6 또는 ssODN 공여자 주형 전달과 비교하였다. 생존율 (MUSE 세포 계수기를 사용하여 평가됨), HDR (ddPCR에 의해 평가됨) 및 NHEJ (ICE에 의해 평가됨) 결과를, SCGM (1백만개 세포/ml) 또는 SFEM-II 배지 (250,000개 세포/ml)에서 배양된 동원된 CD34+ HSC를 사용하여 론자 프로그램 CM149 및 ER100으로 비교하였다.
- [0199] RNP 단독 및 HDR 후 잔류 indel에 대한 NHEJ 파괴 비율은 시험된 두 배지에서 2개의 론자 방법에 대해 필적하였다. 대조적으로, CM149 방법을 사용하여 RNP 및 50 pmol 및 25 pmol 농도의 ssODN을 사용한 HDR-편집은 ER100 방법보다 편집 후 2일차에 더 많은 생존 세포를 초래하였다 (도 15a). 40% HDR이 SFEM-II 배지에서 RNP + 50 pmol의 ssODN으로 달성되었다. ER100은 50 pmol 및 25 pmol의 ssODN으로 CM149보다 더 높은 HDR 비율을 달성한 반면, ER100을 사용하면 생존율 및 세포 수가 극적으로 감소되었으며, 따라서 바람직하지 않았다. 저밀도 배양을 사용한 SFEM-II 배지는 SCGM 배지보다 LT-HSC (CD34+ CD38 Lo) 구획 (CD34+ CD38Lo)에서 더 많은 세포를 보존하였으며, 따라서 생체내 생착에 바람직할 가능성이 높았다. 두 방법은 시험된 두 배지에서 대략 30% β A 성인 글로빈 발현을 구동하였다 (도 15b). 두 방법은 편집 후 (2일차) 70% 미만의 생존율을 초래하였으며, 따라서 두 론자 방법은 장기간 생착에 차선택일 가능성이 높았다.
- [0200] 실시예 5-뉴클레오펙션시 세포 밀도의 비교
- [0201] 최적의 HDR, NHEJ 및 생존율을 구동하는데 있어서 세포 밀도의 역할을 평가하기 위해 뉴클레오펙션시 세포 밀도를 비교하였다. 세포수가 RNP 편집 및 rAAV6 및 ssODN 공여자 주형 전달 후 생존율, HDR 및 NHEJ 결과에서 역할을 갖는지 이해하기 위해, 20 μ l 론자 반응 중 세포 밀도 (20 μ l 뉴클레오펙션 반응 당 200,000/400,000/600,000개 세포)는 다양하였다.
- [0202] 200,000개 세포/20 μ l 뉴클레오펙션 반응은 시험된 두 방법 (CM149 및 ER100)으로 뉴클레오펙션 후 2일차에 편집 후 CD34+ 세포의 생존율과 관련하여 나머지 세포 밀도를 증가하였다 (도 16a). 200,000개 세포/20 μ l 뉴클레오펙션 반응은, 400,000개 세포/20 μ l 반응이 가장 잘 작동하는 ER100으로 시험된 RNP + 50 pmol의 ssODN을 제외하고, 편집 후 14일차에 측정된 바와 같이 HDR과 관련하여 나머지 세포 밀도를 증가하였다 (도 16b). 세포 밀도는 NHEJ 비율에 영향을 미치지 않았다 (도 16c). 총 NHEJ 비율 및 HDR 후 잔류 NHEJ 비율은 시험된 두 방법으로 다양한 밀도에 걸쳐 필적하였다.
- [0203] 실시예 6-세포 생존율 및 HDR 비율의 평가
- [0204] RNP 및 ssODN 전달과 관련하여 대안적인 론자 프로그램을 사용하여 세포 생존율 및 HDR 비율을 평가하였다.
- [0205] CX100은 편집 후 2일차에 70% 생존율을 달성하고 14일차에 27% HDR을 달성하였다. E0100, DU100 및 DZ100은 40-50% HDR을 가졌으나, 편집 후 2일차에 오직 20-45% 생존율을 가졌다 (도 17a). 생존율과 관련하여: ER100 < DZ100 < DU100 < E0100 < CM149 < CX100. HDR과 관련하여 ER100 < CM149 < CX100 < DZ100 < E0100 < DU100 (도 17b). 높은 생존율 및 높은 HDR이 바람직하였고, 그러므로 CX100 프로그램은 장기간 HSC에서 HDR을 달성하기 위해 RNP 및 ssODN 전달과 조합하기에 가장 바람직한 플랫폼이었다 (도 17c).
- [0206] 실시예 7-ssODN의 평가
- [0207] 생존율 및 HDR을 최대화하기 위해 CX100 및 DU100을 사용한 ssODN (pmol 적정)의 평가를 수행하였다. 2일차에 세포 생존율 및 HDR 비율을 최대화하는 조건을 찾기 위해 최고 성능의 론자 뉴클레오펙션 방법과 관련하여 대안적인 용량의 ssODN을 평가하였다. 론자 DU100 또는 CX100 프로그램을 사용하여 RNP와 함께 ssODN을 100 pmol, 50 pmol 및 25 pmol에서 시험하였다.
- [0208] CX100은 DU100보다 더 큰 생존율을 유발하였다. 50 pmol의 CX100 + RNP는 80%의 생존율 및 30%의 HDR을 가졌다. RNP 파괴는 두 방법에서 동일하였다. 잔류 NHEJ는 DU100과 비교하여 CX100을 사용한 RNP + 50 pmol의 ssODN에서 더 높았다 (도 18).
- [0209] 실시예 4-7은 다양한 론자 뉴클레오펙션 방법을 사용하여 생존율을 보존하면서 HDR을 최대화하는 조건을 예시한다. 전반적으로, SFEM-II 배지는 LT-HSC 구획에 더 많은 세포가 보존되었기 때문에 SCGM 배지보다 더 바람직하였다. 론자 CX100 프로그램은 30%의 HDR 결과를 구동하면서 편집된 세포의 80% 생존율을 보존하였다. 그러므로, 연구는 다음을 포함하여 ssODN 및 rAAV6을 사용한 HBB의 HDR 편집의 임상 번역에 가장 적합한 제한된 범위의 조건을 나타내었다: (a) 250,000 내지 1000,000개 세포/ml를 사용하는 저밀도 배양 조건에서 SFEM-II 배지의 사용; (b) SFEM-II 배지에서 100 ng/ml의 IL-6, TPO, FLT-3L, SCF의 시토카인의 사용; (c) 다음을 사용하여 최

적의 결과를 얻는 대안적인 론자 프로그램의 사용: (i) HDR 최대화를 위한 ER100< CM149< CX100< DZ100< EO100< DU100; (ii) 생존율 최대화를 위한 ER100< DZ100< DU100< EO100< CM149< CX100; (iii) 1:1, 1:2.5, 1:5 비율의 RNP의 사용 (20 pmol 내지 40 pmol의 Cas9 사용); (iv) 200,000개 세포에 대한 10 pmol 내지 100 pmol 농도의 ssODN의 사용; (v) 1-3% 배양 부피, 바이러스 역가를 기준으로 MOI 2000 내지 6000에서 rAAV6의 사용.

[0210] 실시예 8-생체내 및 시험관내 HDR 및 야생형 결과의 검출

[0211] HDR 사건 및 비편집된 (야생형) 사건 둘 모두를 측정하기 위해 ddPCR 검정을 사용하였다. HDR FAM 프로브가 결합하거나 야생형 HEX 프로브가 게놈 DNA (gDNA)에 결합하는 상호 배타적인 검정으로서 검정을 개발하였다. FAM 및 HEX 프로브를 함께 혼합하고, 동일한 웰에서 결합 부위에 대해 경쟁하도록 허용하였다. 동일한 앰플리콘 내의 모든 gDNA에 결합하는 참조-HEX 프로브를 내부 참조로 사용하고 병렬로 실행하였다. 계산은 다음과 같았다:

[0212] $\% \text{ HDR} = (\% \text{ FAM+})/(\text{Ref HEX+})$

[0213] $\% \text{ WT} = (\% \text{ HEX+})/(\text{Ref HEX+})$

[0214] 도 19는 E6V (GTC) 변화 및 EoptE (GAA) 변화 둘 모두에 대한 대표적인 모의, AAV, RNP, RNP + AAV 및 RNP + ssODN 샘플에 대한 ddPCR 검정 결과를 나타낸다. 모든 FAM+ 사건은 HDR 사건을 뚜렷하게 나타내고, HEX+ 사건은 야생형 사건을 나타내었다.

[0215] 실시예 9-ddPCR에 의한 HDR 및 ICE에 의한 NHEJ의 검증

[0216] 백분율 (%) HDR을 ddPCR 데이터로부터 계산하고 (도 20a), % NHEJ를 ICE 알고리즘 데이터로부터 계산하였다 (도 20b). % NHEJ를 결정하기 위해 ICE 분석을 사용하였다. 녹아웃 비율 (NHEJ) 뿐만 아니라 녹인 비율 (HDR) 둘 모두를 결정하기 위해 ICE 알고리즘을 사용할 수 있다. indel 및 HDR을 뚜렷하게 구별하기 위해 가이드 서열 및 공여자 주형 서열을 제공하였다. 분석은 indel 및 HDR이 올바르게 확인되었는지 보장하기 위해 입증된 서열 트레이스를 전달하였다. HDR 및 NHEJ 데이터를 하나의 그래프로 조합하여, 총 편집 비율 (총 편집 = % NHEJ + % HDR)을 나타낼 뿐만 아니라 HDR의 증가에 따라 NHEJ의 감소가 있음을 나타낸다. ICE 및 ddPCR 데이터를 확인한 크리스프레소 알고리즘을 사용하여 MiSeq 분석에 의해 HDR 및 NHEJ 데이터를 추가로 검증하고, RNP 전달과 함께 대부분의 NHEJ 사건이 주로 결실임을 나타내었다.

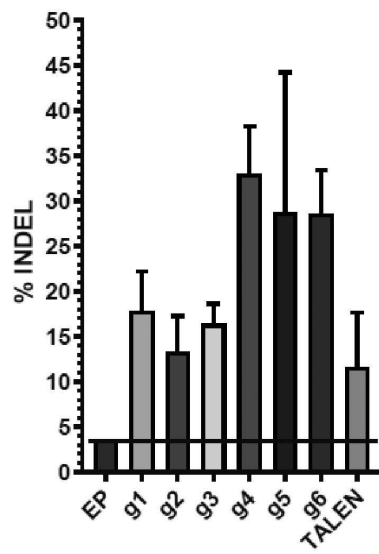
[0217] 본원에 사용된 바와 같은 용어 "포함하는(comprising)"은 "포함하는(including)", "함유하는(containing)" 또는 "특징으로 하는"과 동의어이며, 포괄적이거나 개방형이고, 추가의 인용되지 않은 요소 또는 방법 단계를 배제하지 않는다.

[0218] 상기 설명은 본 발명의 여러 방법 및 재료를 개시한다. 본 발명은 방법 및 재료의 변형, 뿐만 아니라 제작 방법 및 장비의 변경에 민감하다. 이러한 변형은 본 개시내용 또는 본원에 개시된 본 발명의 실시를 고려하면 관련 기술분야의 통상의 기술자에게 명백하게 될 것이다. 결과적으로, 본 발명은 본원에 개시된 특정 대안으로 제한되는 것이 아니라, 본 발명의 진정한 범위 및 취지 내에 있는 모든 변형 및 대안을 포함하는 것으로 의도된다.

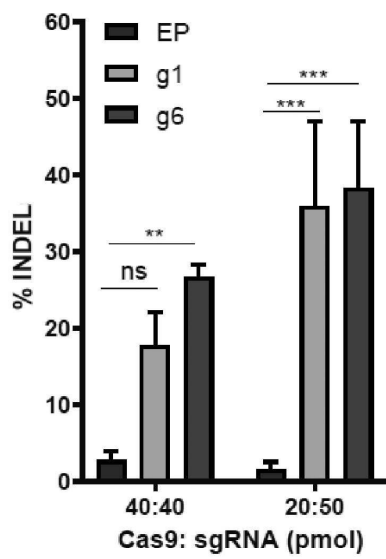
[0219] 공개 및 미공개 출원, 특허 및 문헌 참조를 포함하나 이에 제한되지는 않는 본원에 인용된 모든 참고문헌은 이들의 전문이 본원에 참조로 포함되고, 본원에서 본 명세서의 일부가 된다. 참조로 포함된 간행물 및 특허 또는 특허 출원이 명세서에 포함된 개시내용과 모순되는 한, 명세서는 임의의 이러한 모순되는 자료를 대체하고/거나 우선하도록 의도된다.

도면

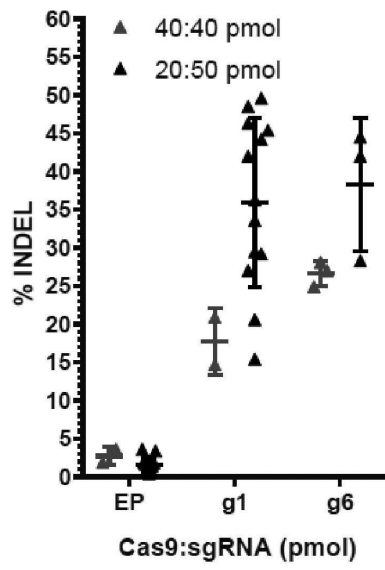
도면1a



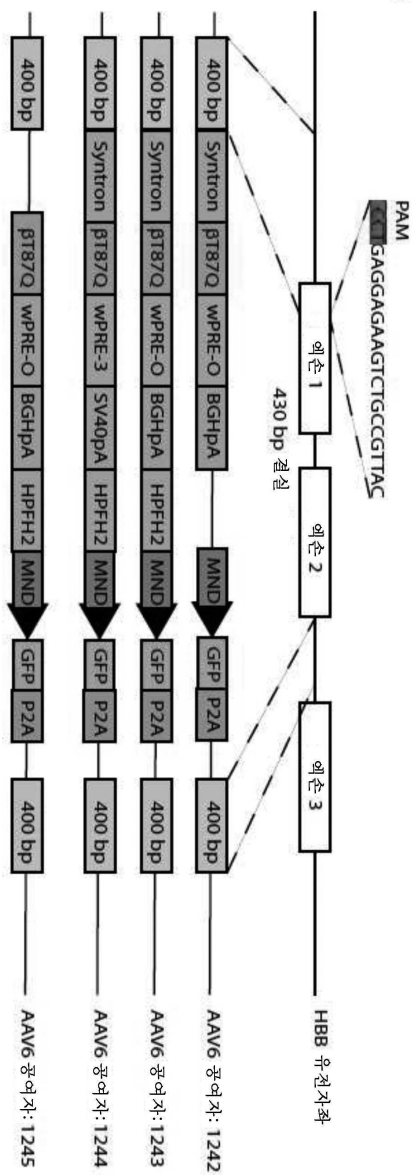
도면1b



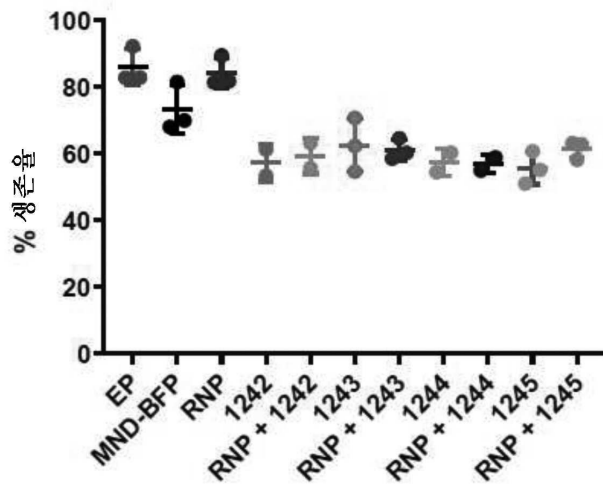
도면1c



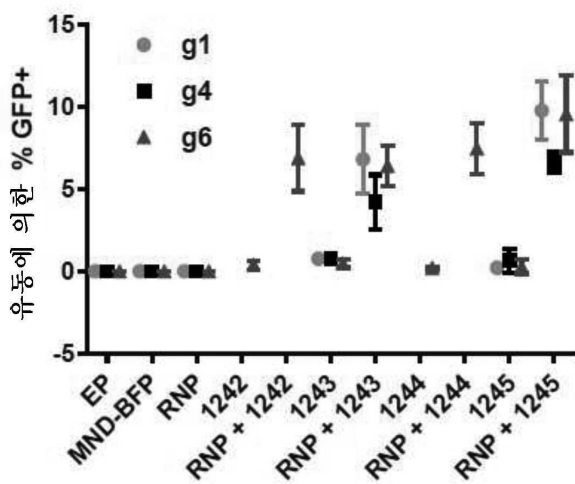
도면2a



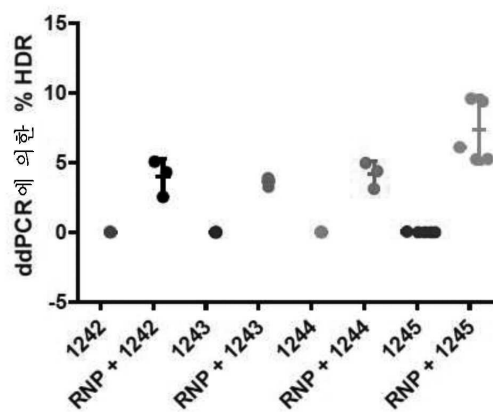
도면2b



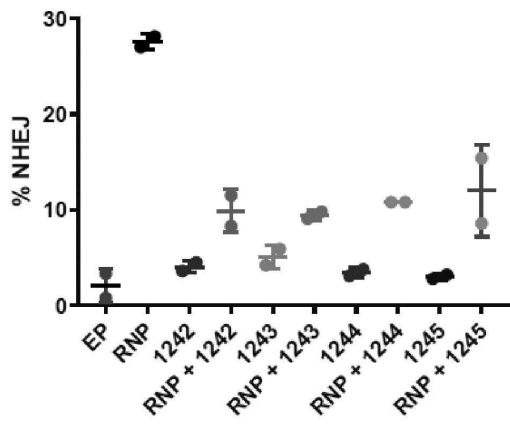
도면2c



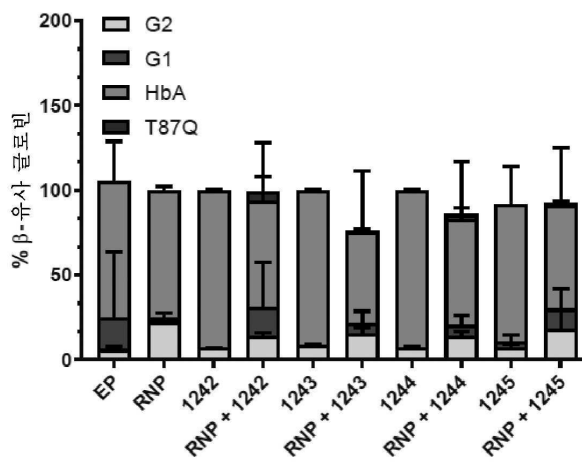
도면2d



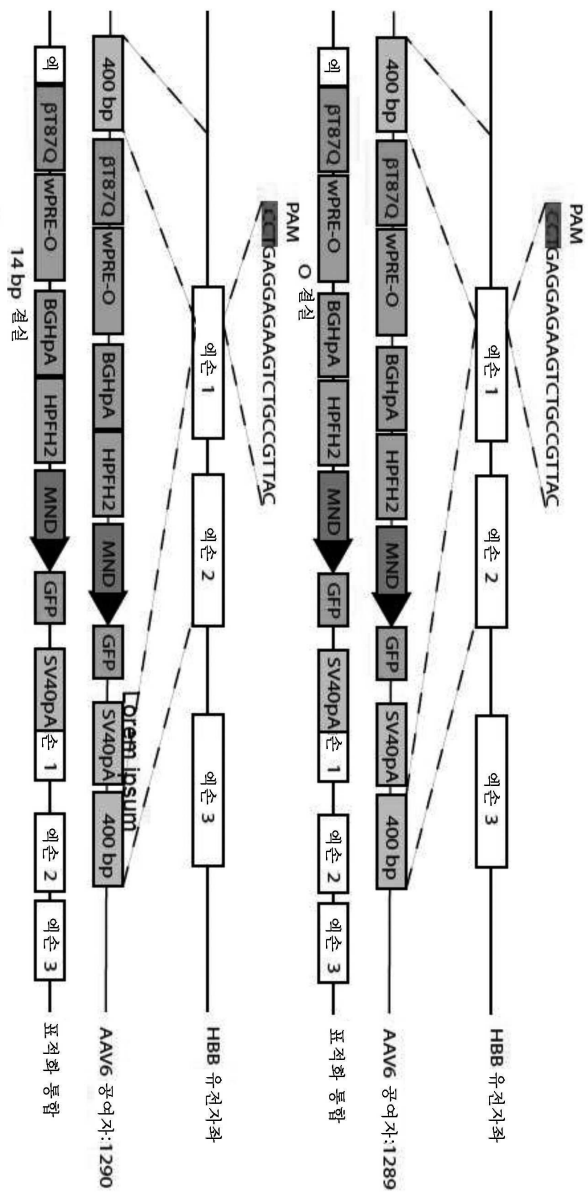
도면2e



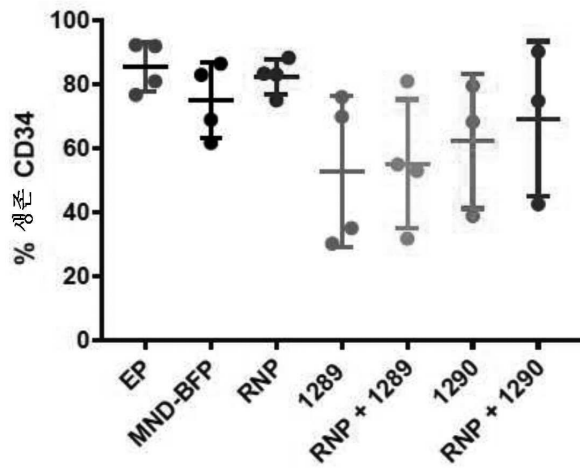
도면2f



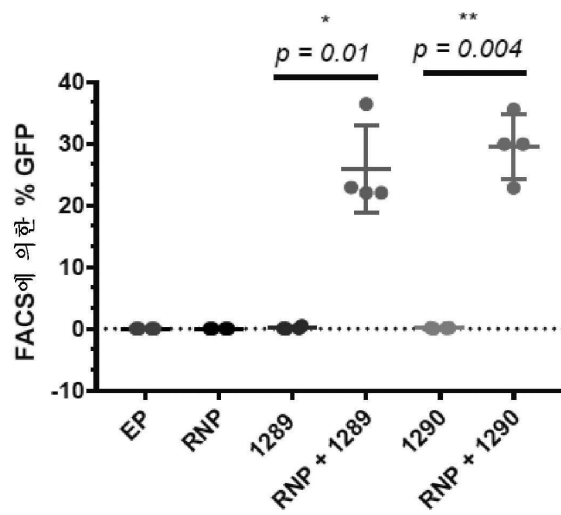
도면3a



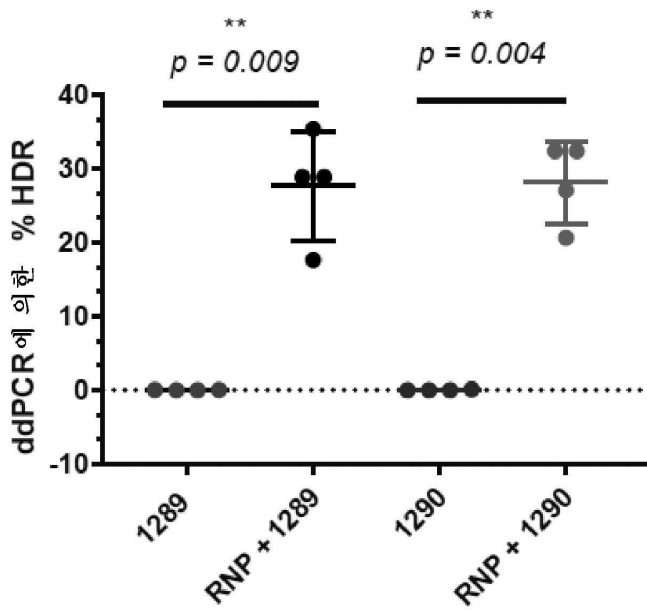
도면3b



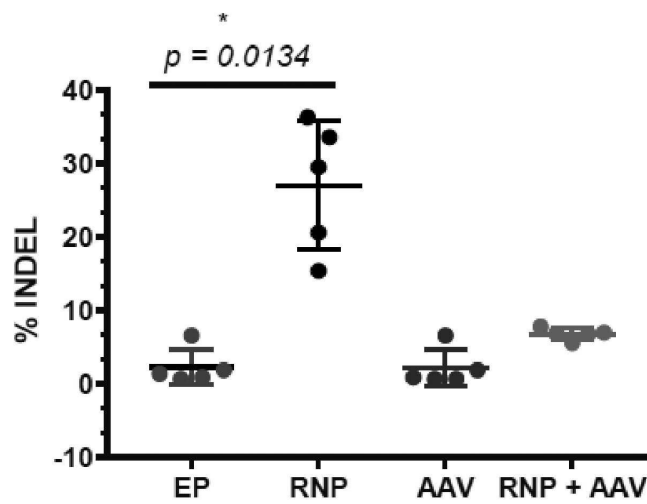
도면3c



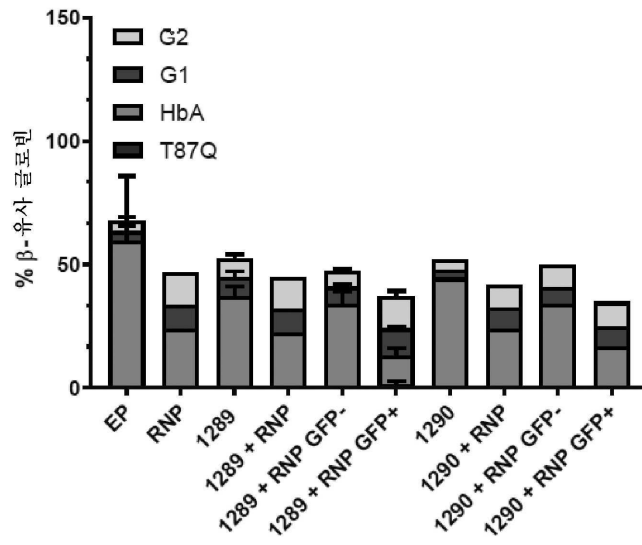
도면3d



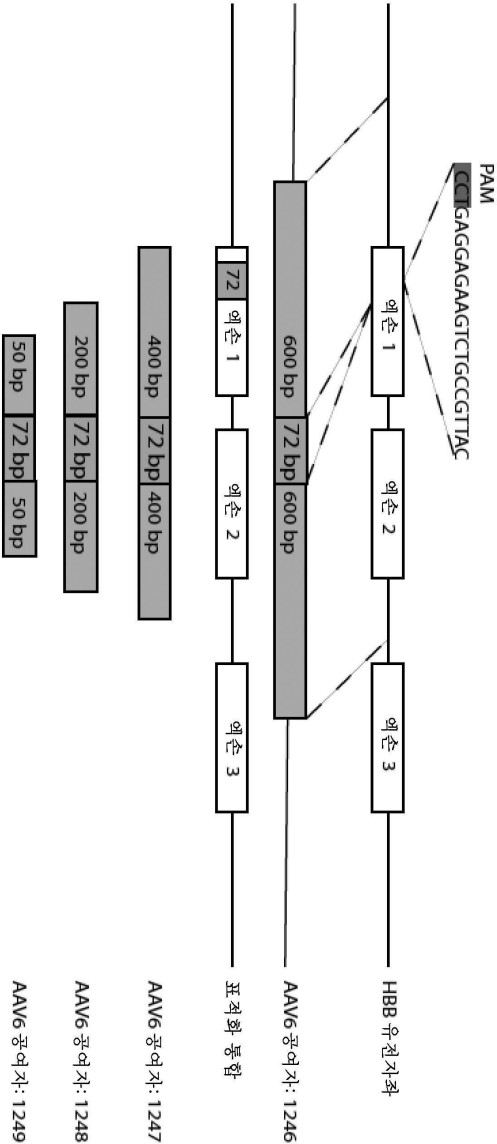
도면3e



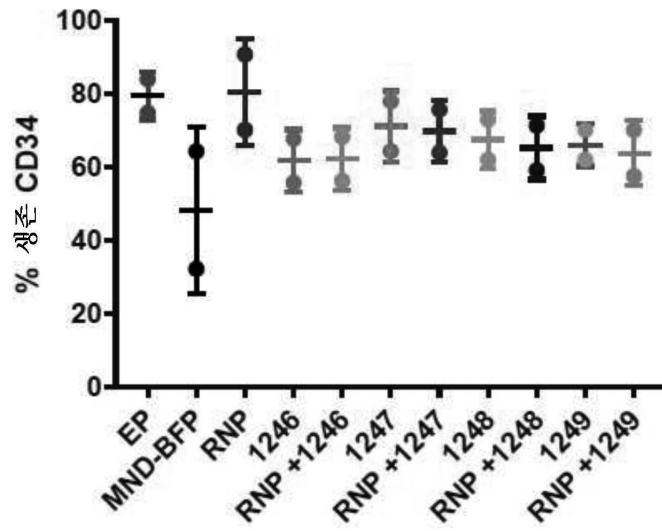
도면3f



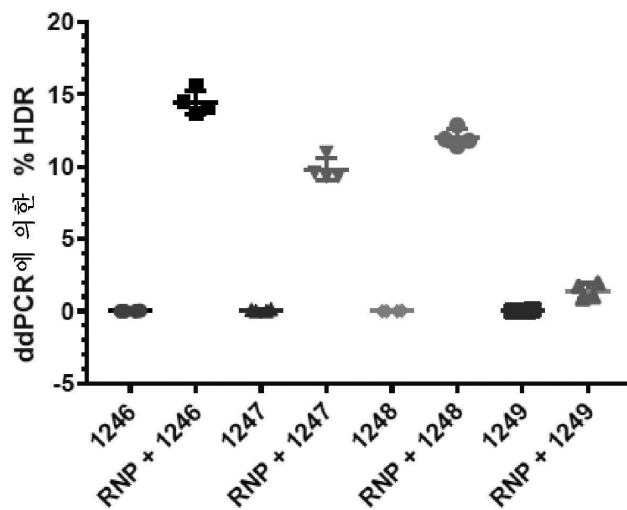
도면4a



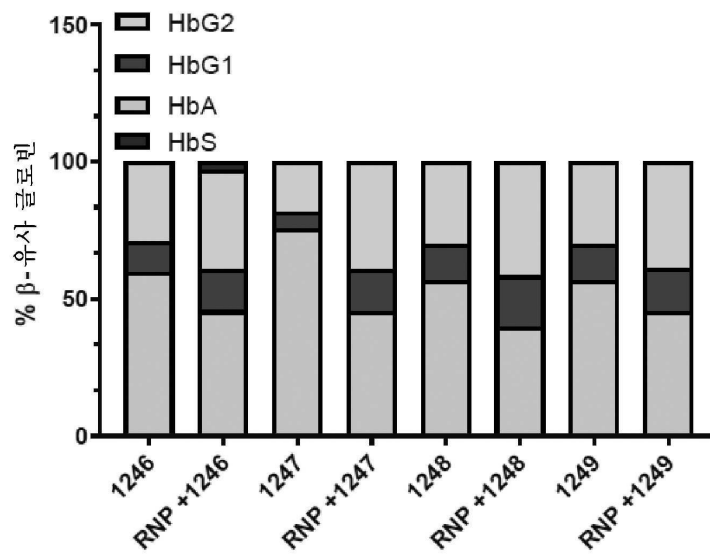
도면4b



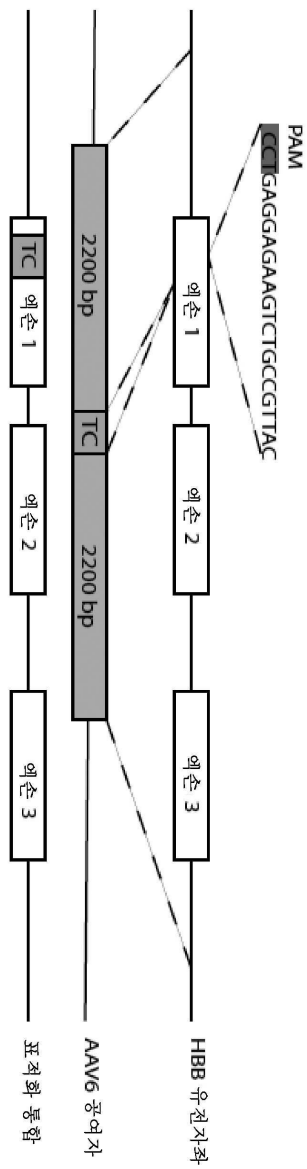
도면4c



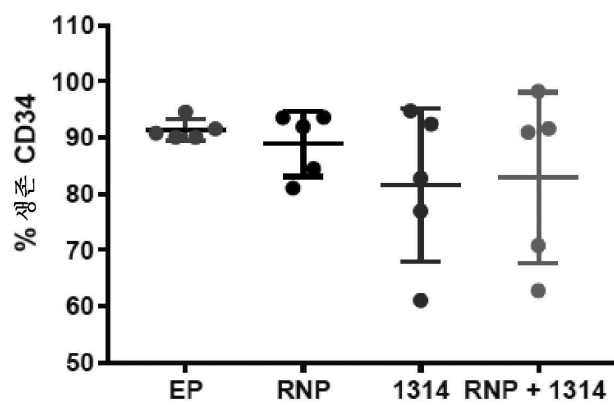
도면4d



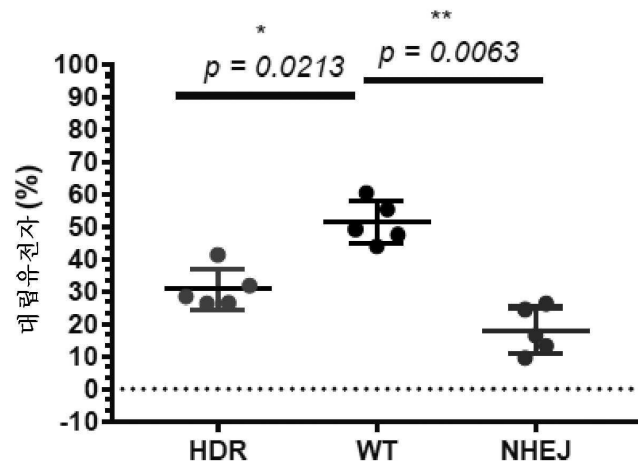
도면5a



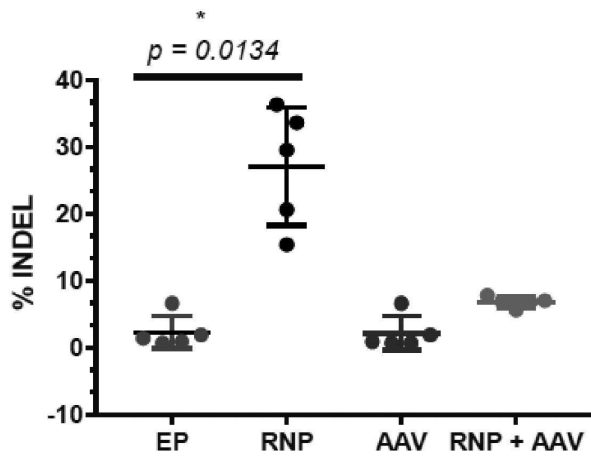
도면5b



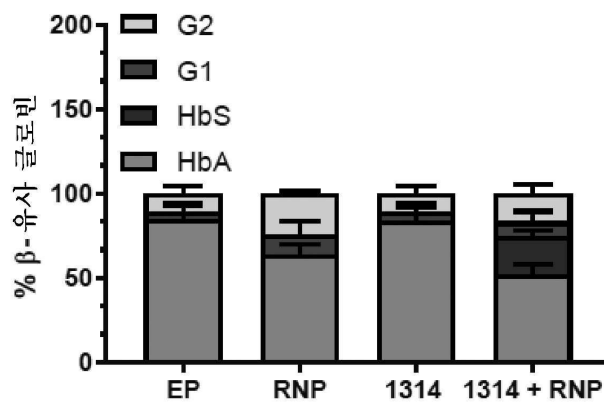
도면5c



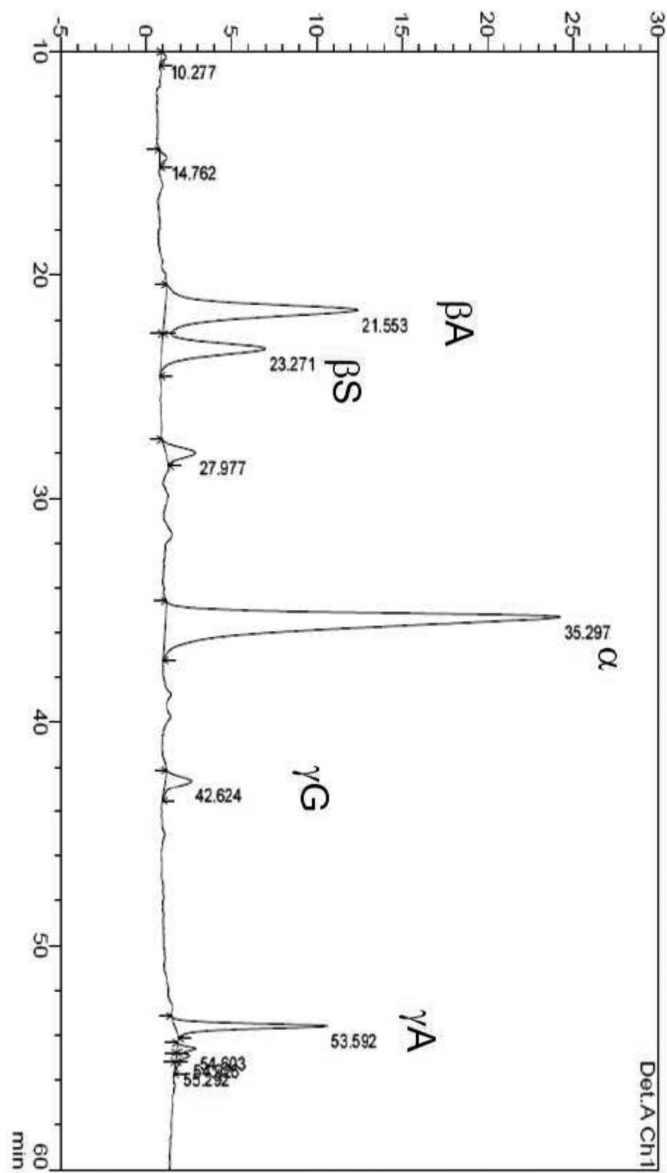
도면5d



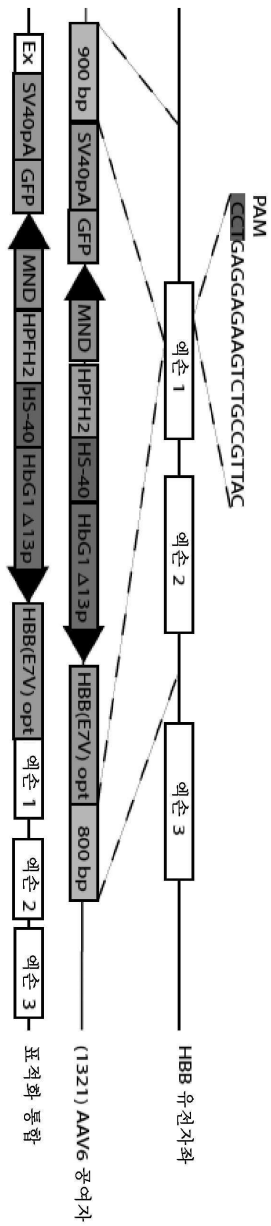
도면5e



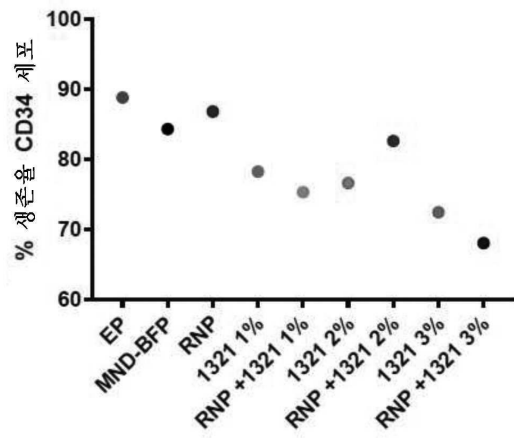
도면5f



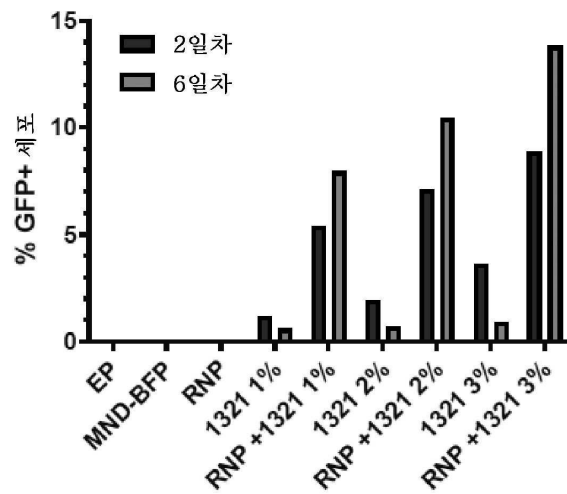
도면6a



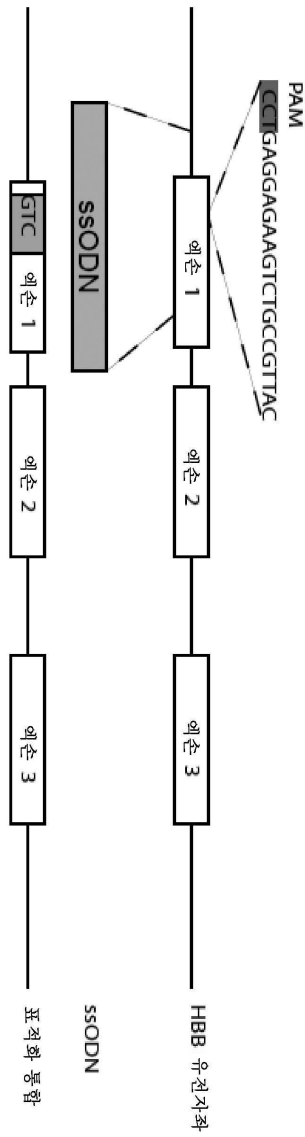
도면6b



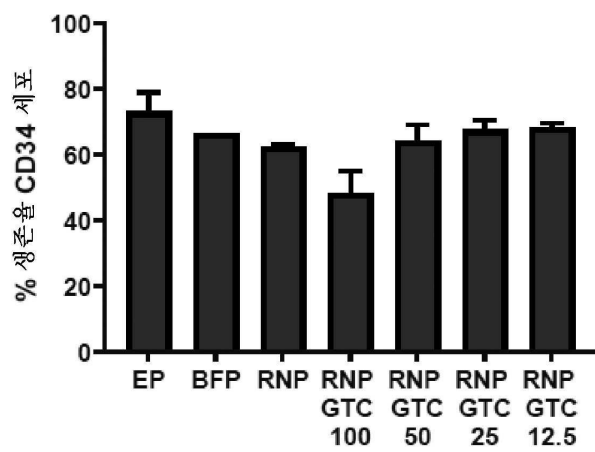
도면6c



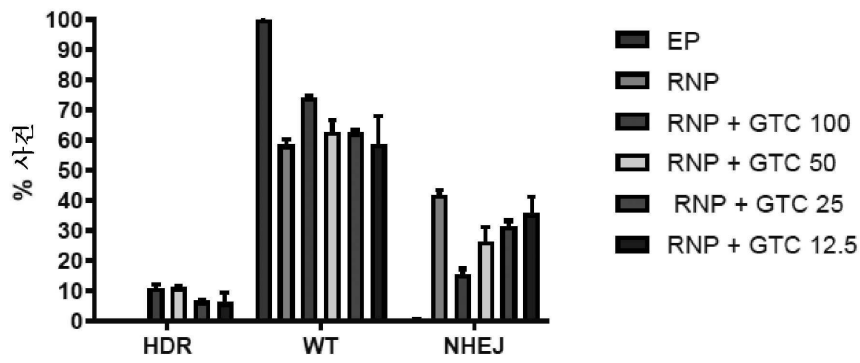
도면7a



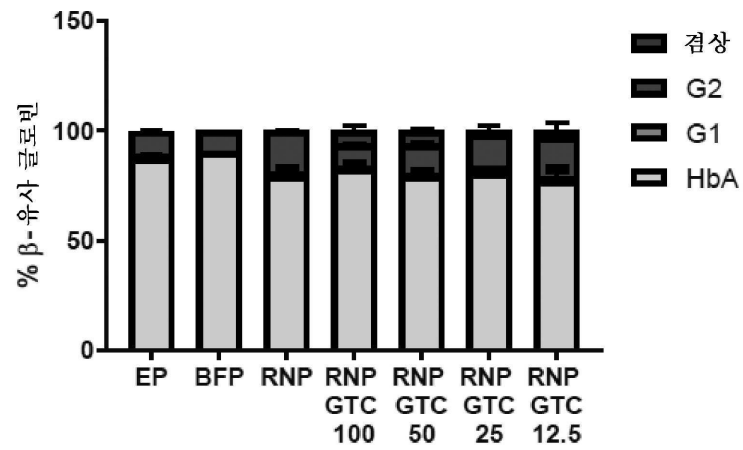
도면7b



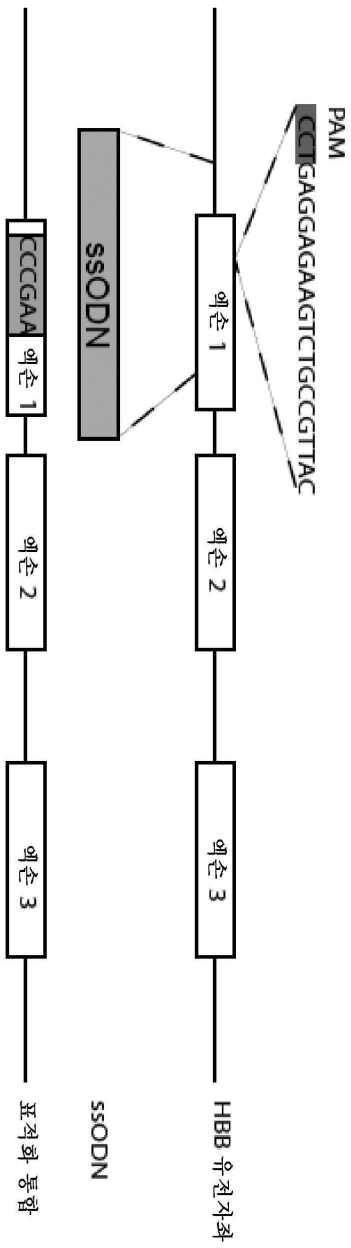
도면7c



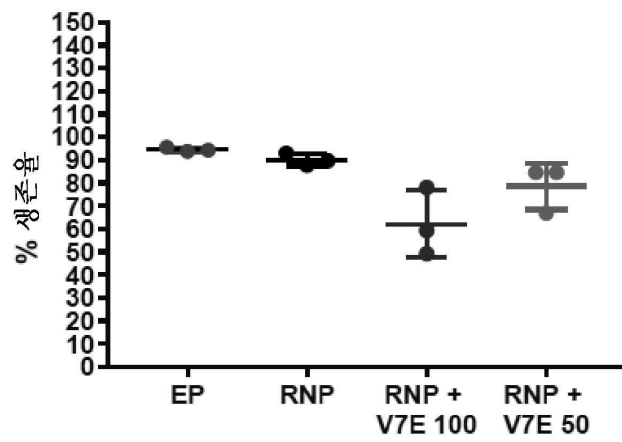
도면7d



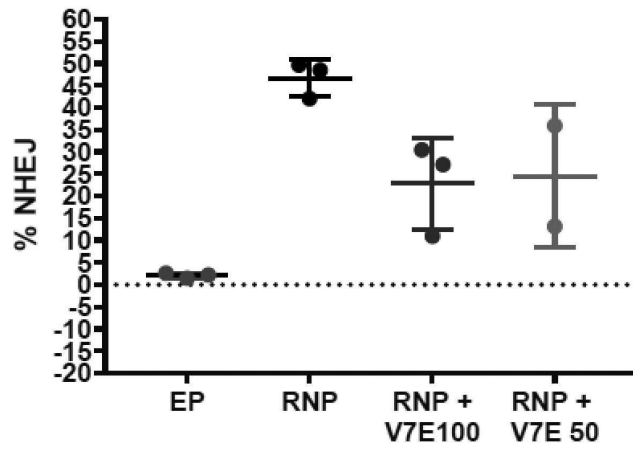
도면8a



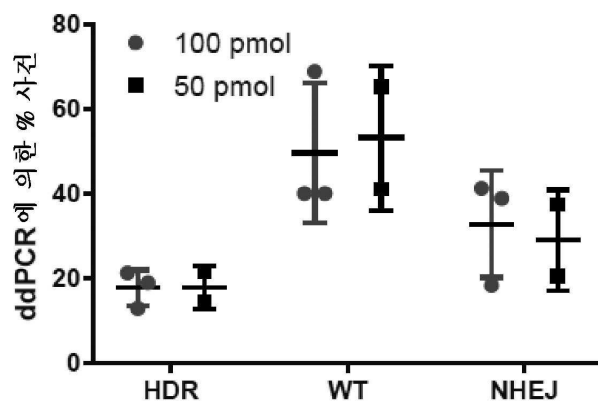
도면8b



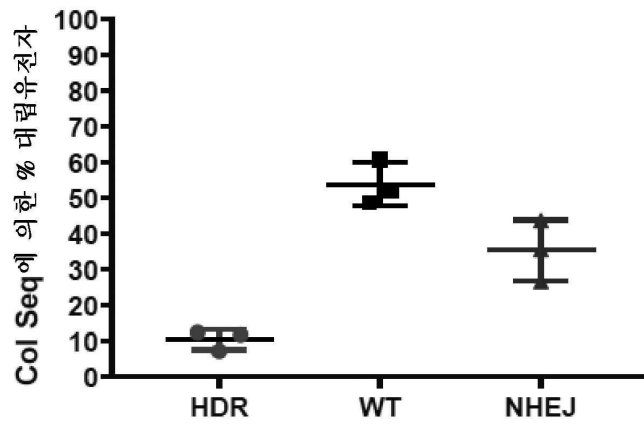
도면8c



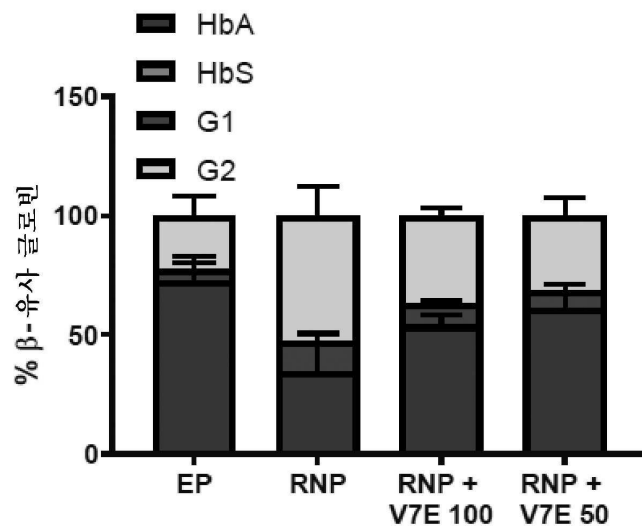
도면8d



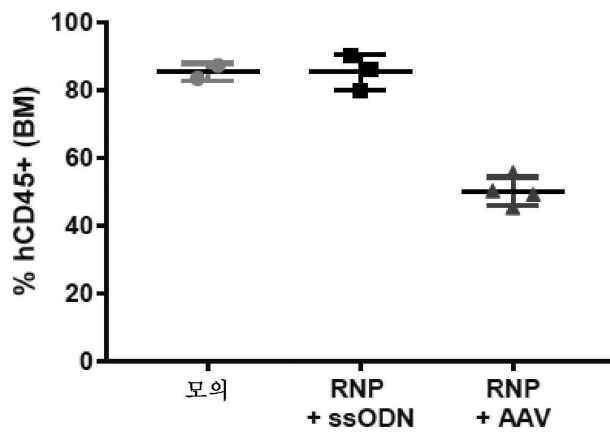
도면8e



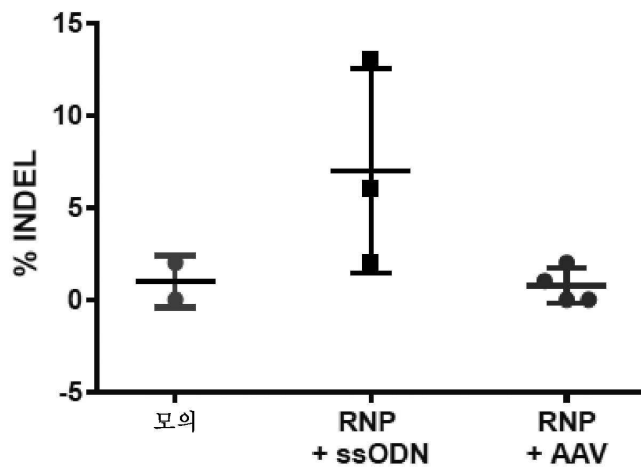
도면8f



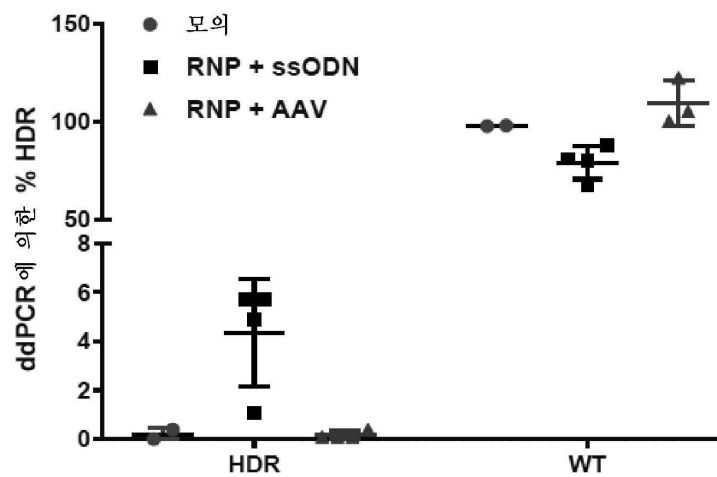
도면9a



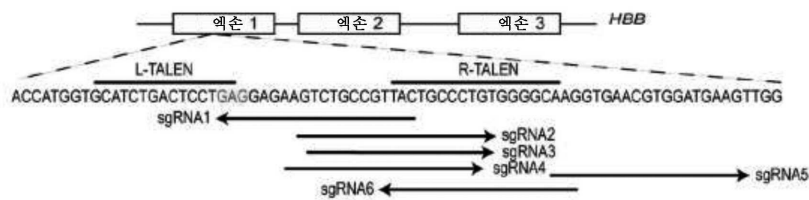
도면9b



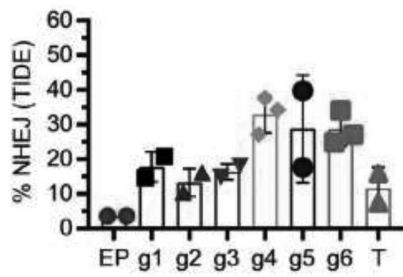
도면9c



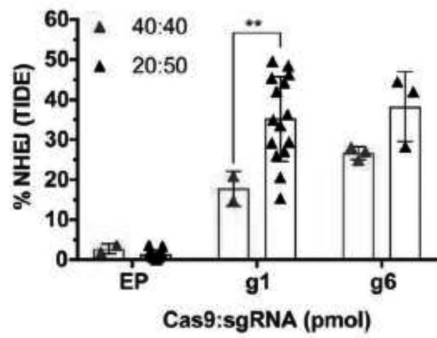
도면10a



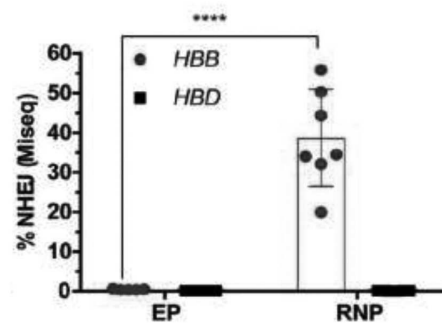
도면10b



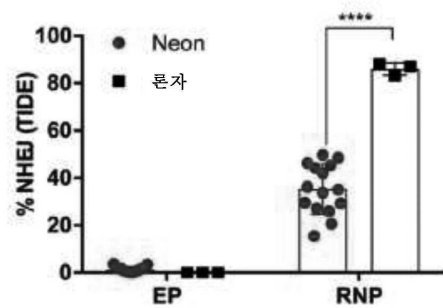
도면10c



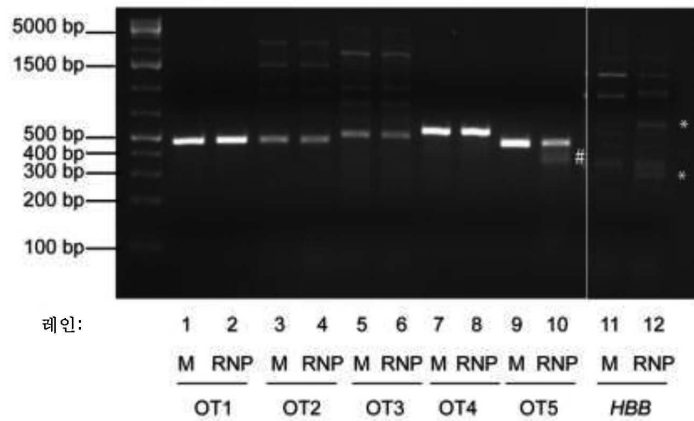
도면10d



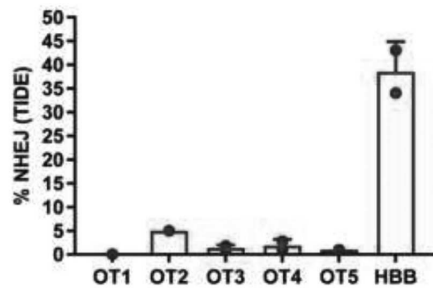
도면10e



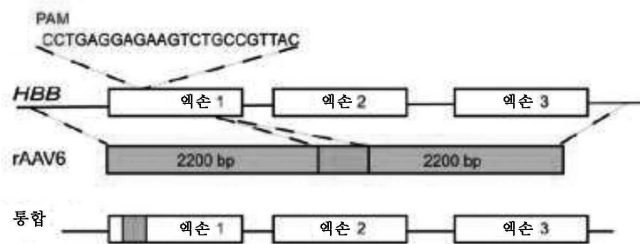
도면10f



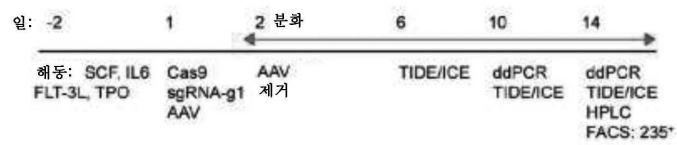
도면10g



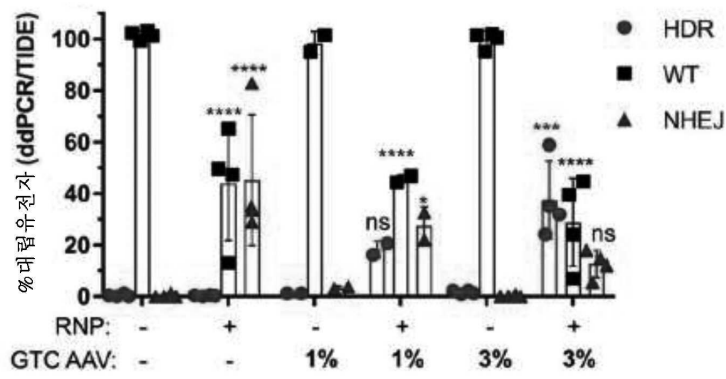
도면11a



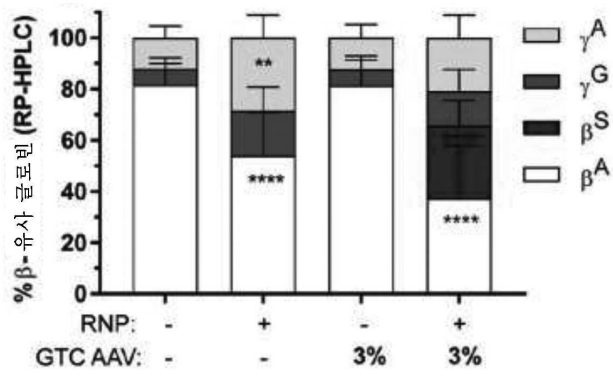
도면11b



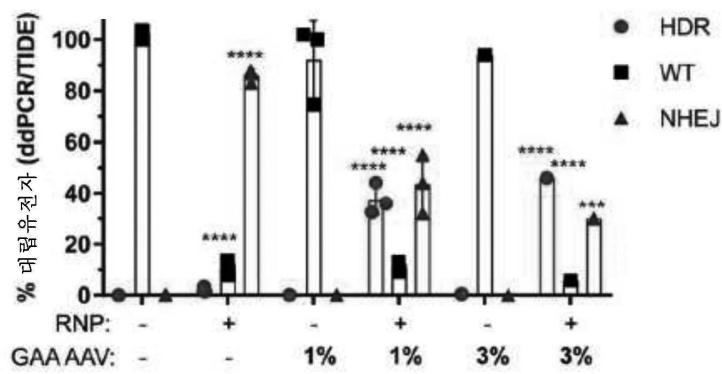
도면11c



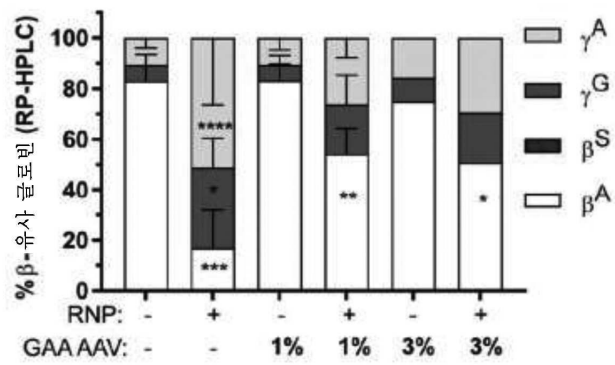
도면11d



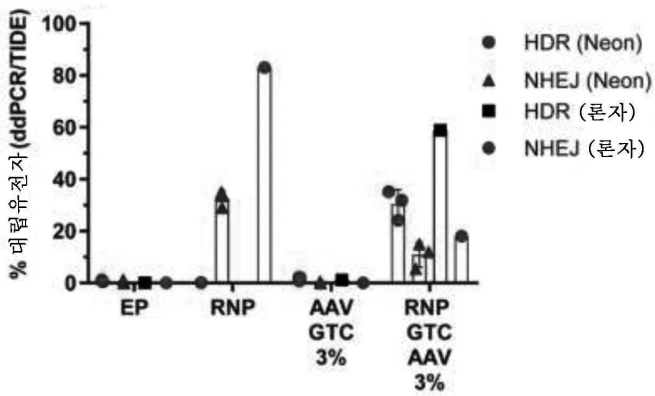
도면11e



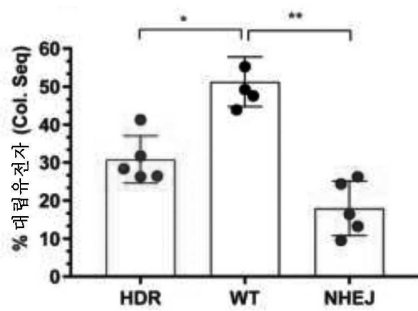
도면11f



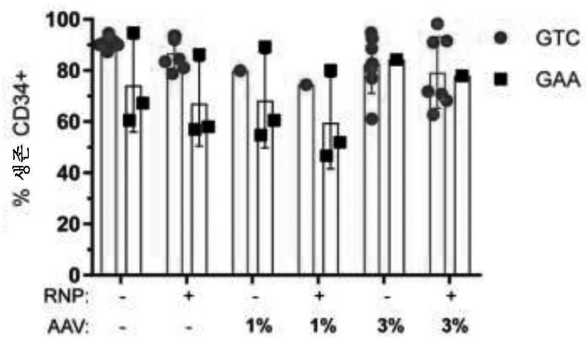
도면11g



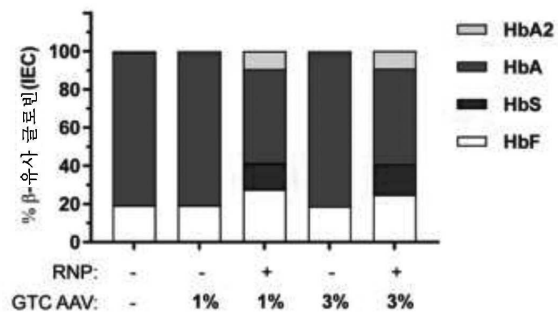
도면11h



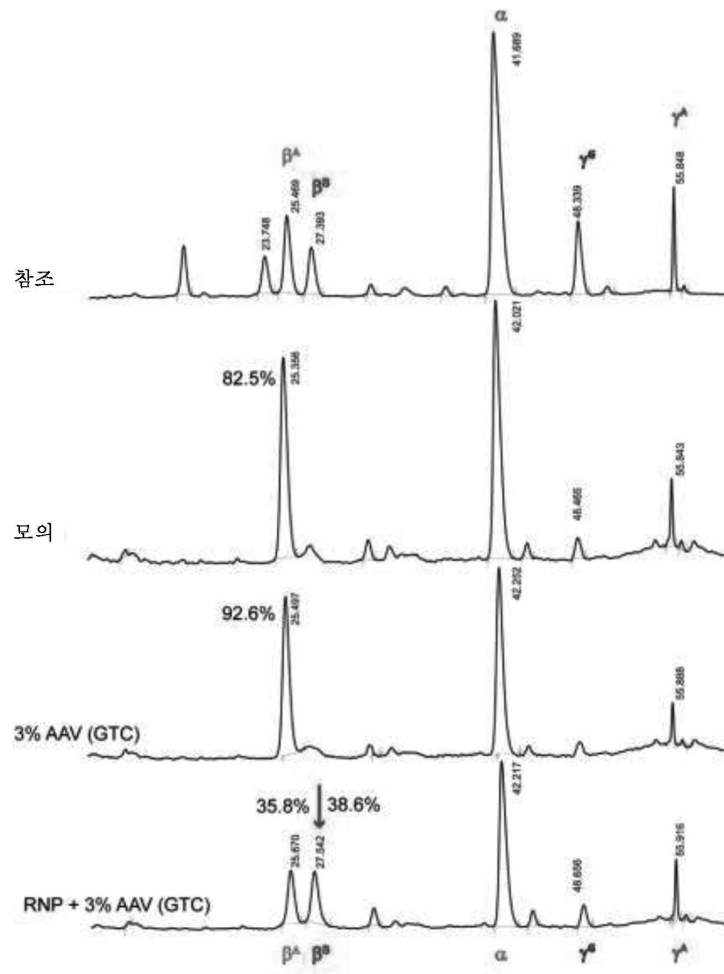
도면11i



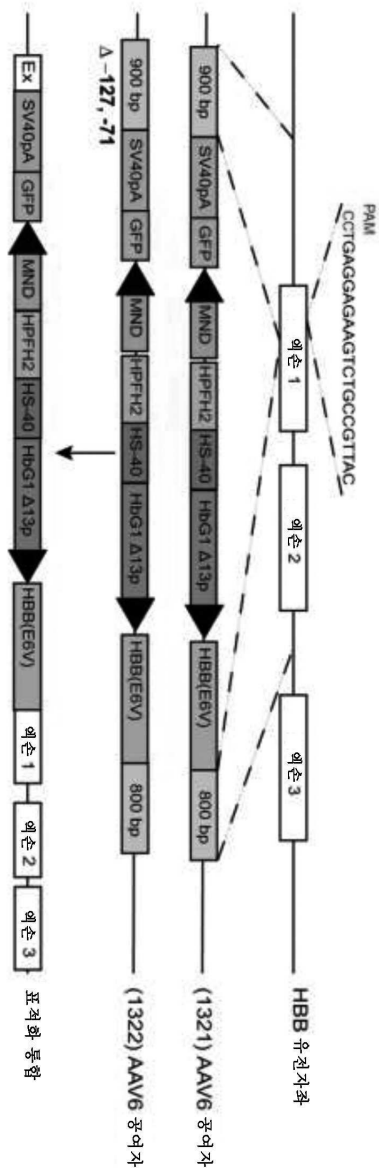
도면11j



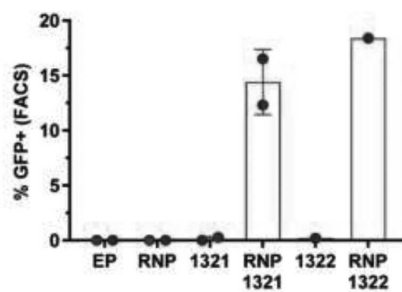
도면11k



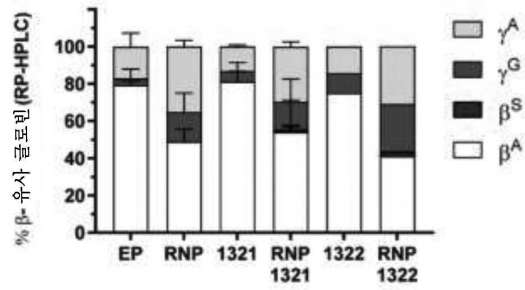
도면11l



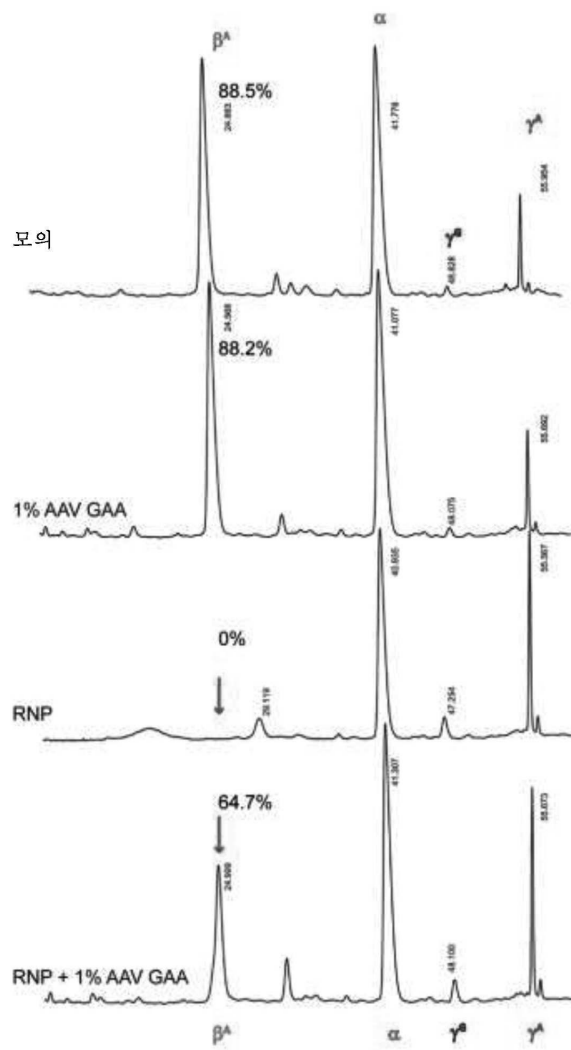
도면11m



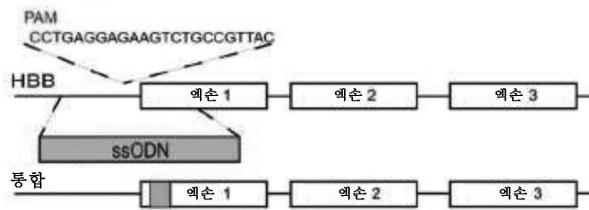
도면11n



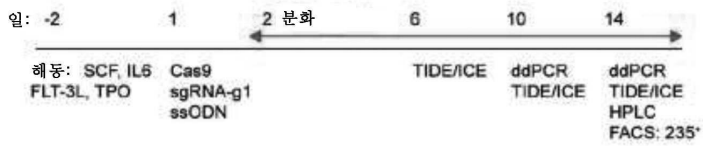
도면11o



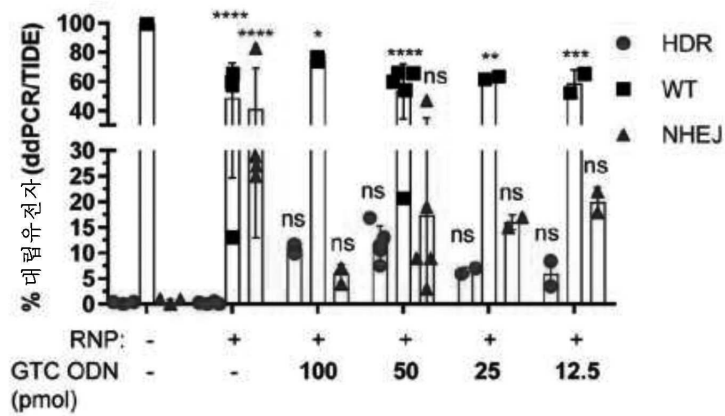
도면12a



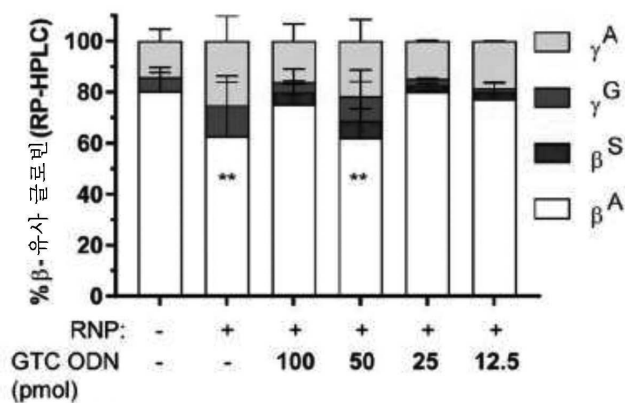
도면12b



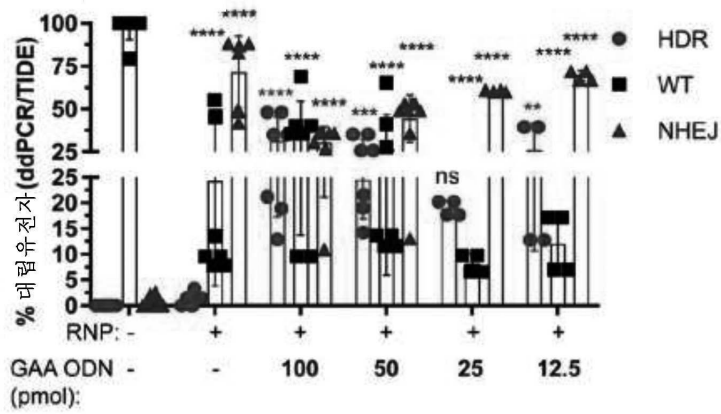
도면12c



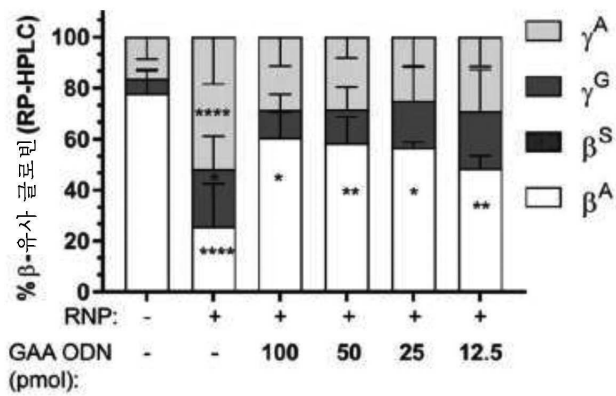
도면12d



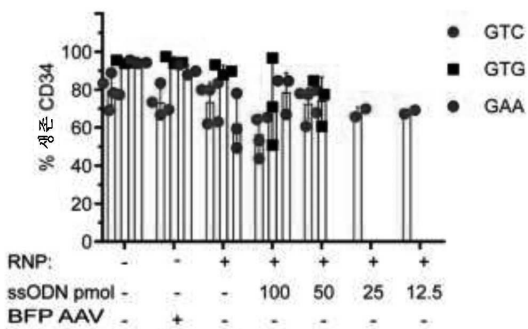
도면12e



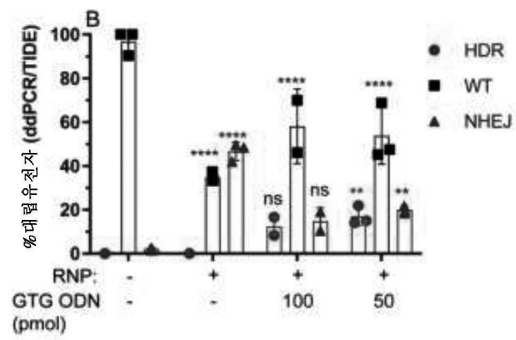
도면12f



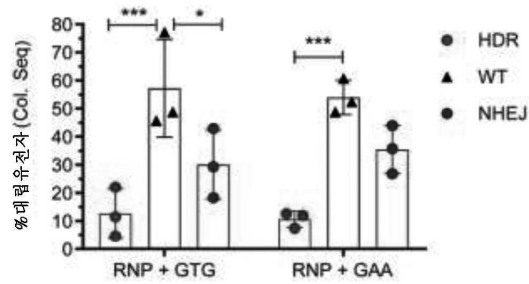
도면12g



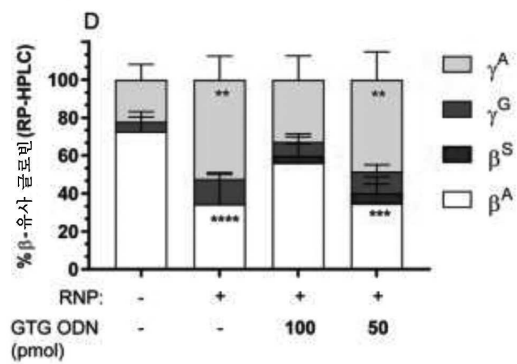
도면12h



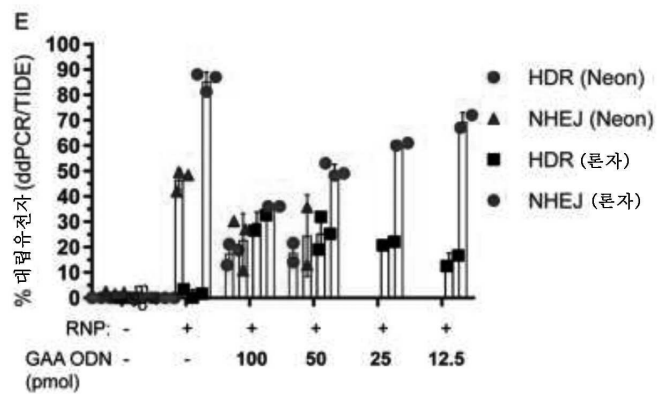
도면12i



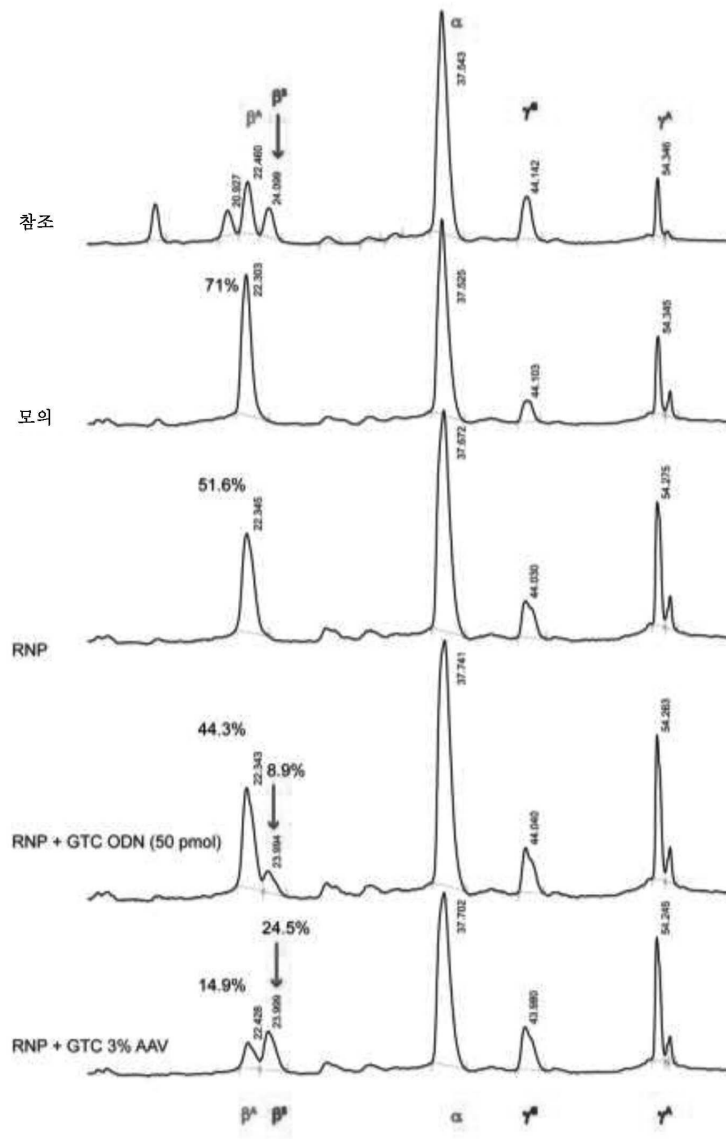
도면12j



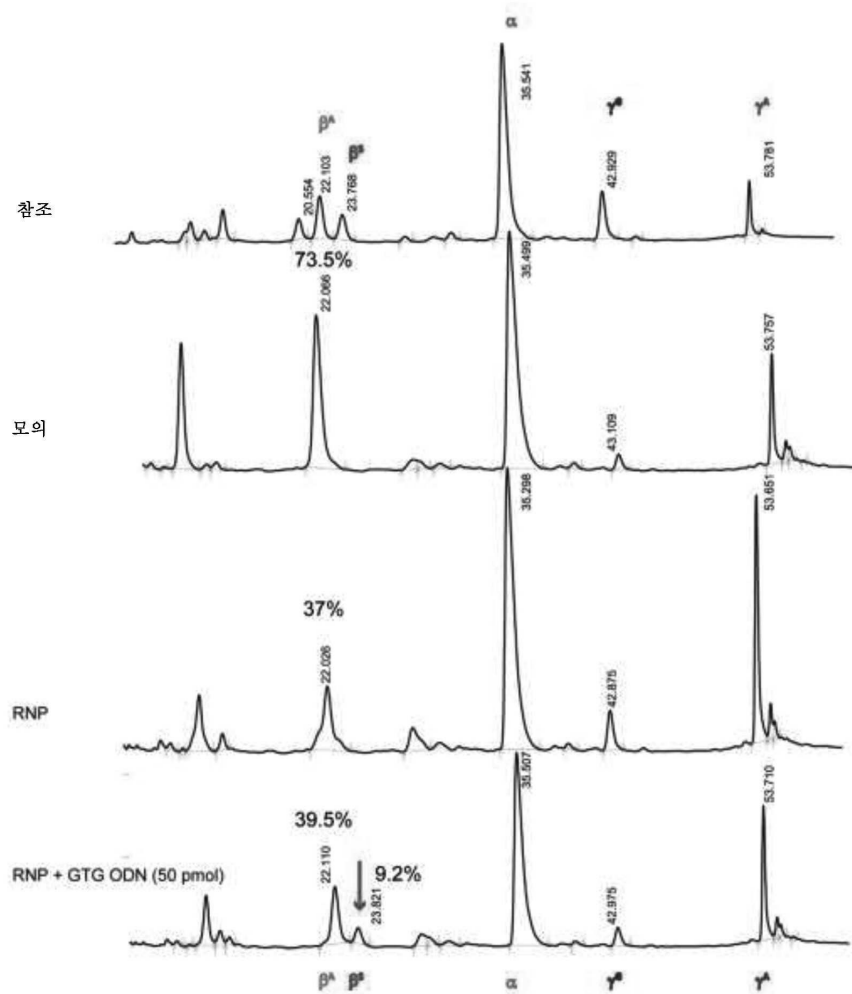
도면12k



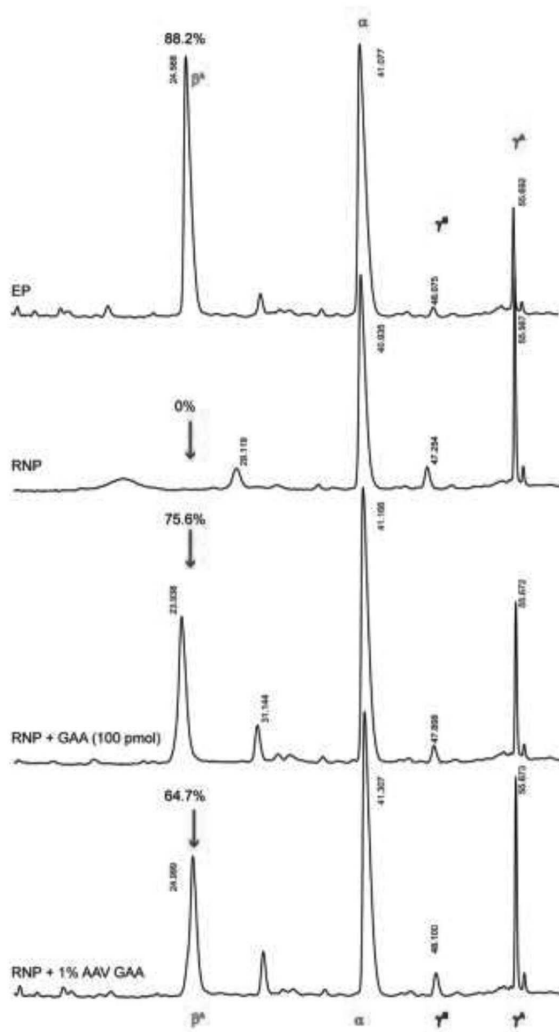
도면121



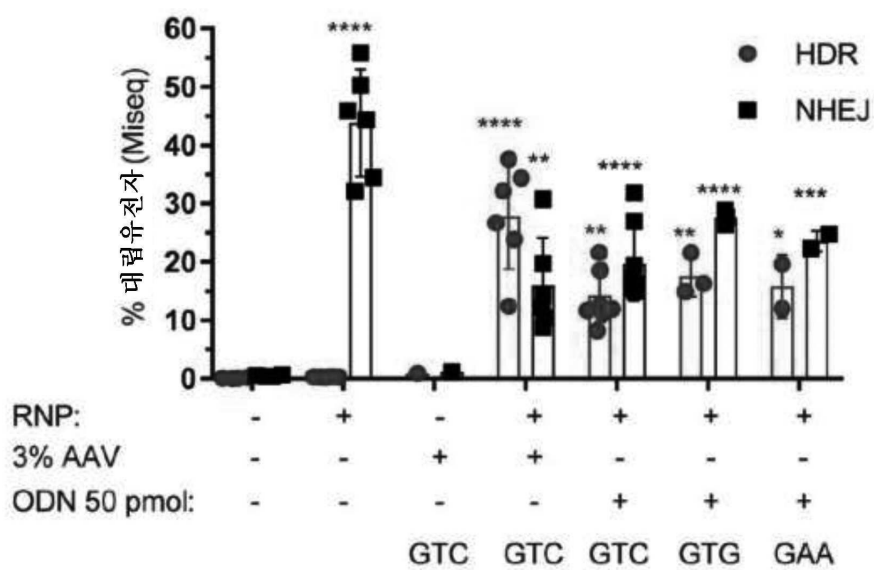
도면 12m



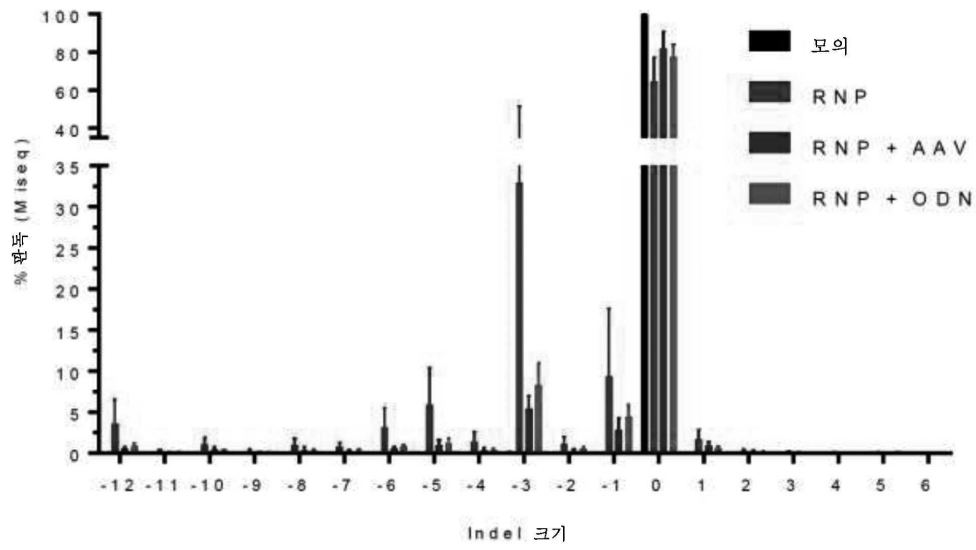
도면12n



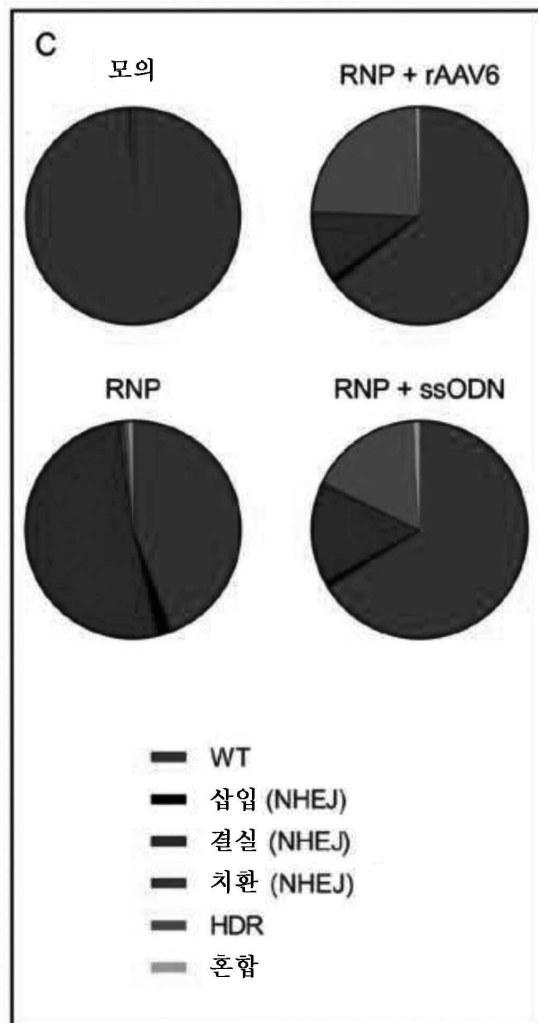
도면13a



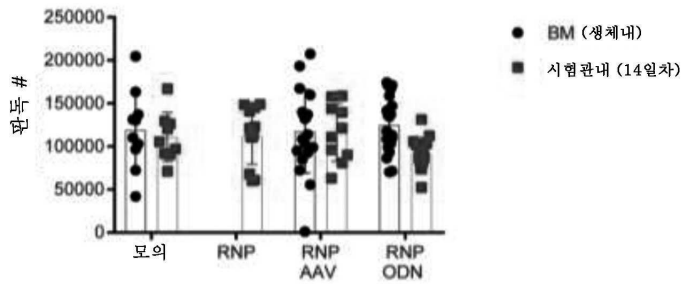
도면13b



도면13c



도면13d



도면13e

모의

CAACCTCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG

RNP

CAACCTCAAACAGACACCATGGTGCACCTGACTCCT---GAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG
CAACCTCAAACAGACACCATGGTGCACCTGACTCCT----GAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG
CAACCTCAAACAGACACCATGGTGCACCTGACTC-----CTGCCGTTACTGCCCTGTGGGGCAAGGTG
CAACCTCAAACAGACACCATGGTGCACCTGACTCCTG-----AGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG

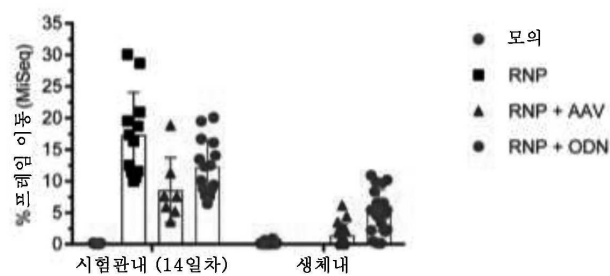
RNP + AAV

CAACCTCAAACAGACACCATGGTGCACCTGACTCCT---GAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG
CAACCTCAAACAGACACCATGGTGCACCTGACTCCTGA-GAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG

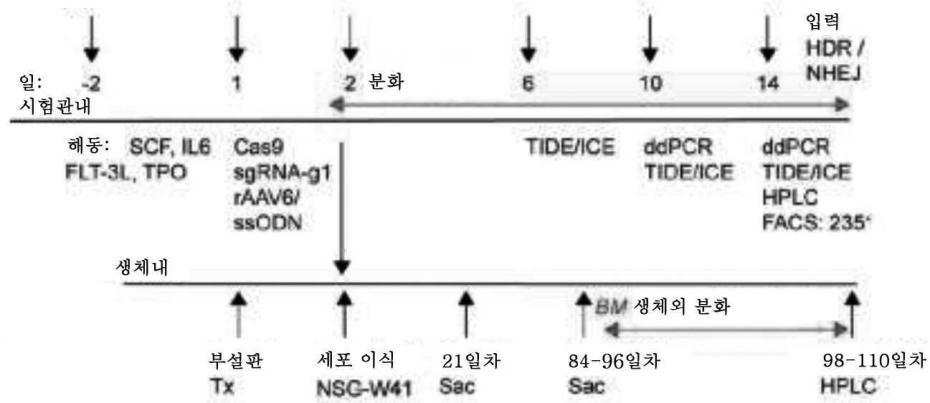
RNP + ODN

CAACCTCAAACAGACACCATGGTGCACCTGACTCCT---GAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG
CAACCTCAAACAGACACCATGGTGCACCTGACTCCTGA-GAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTG

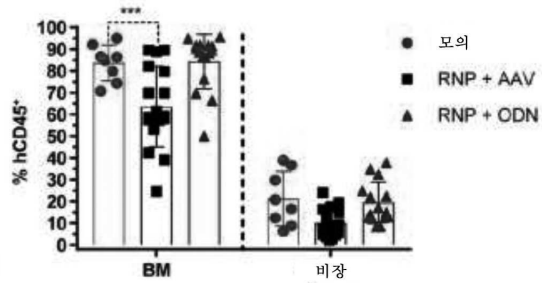
도면13f



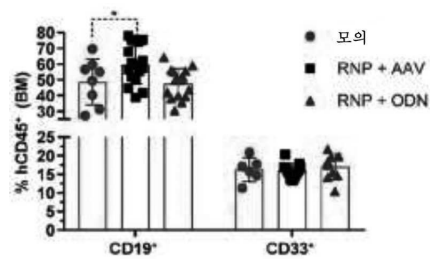
도면14a



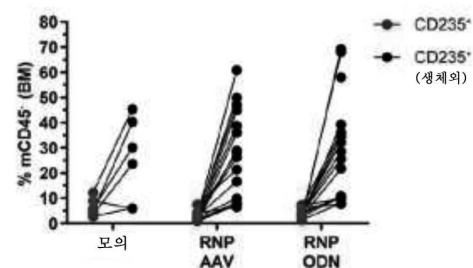
도면14b



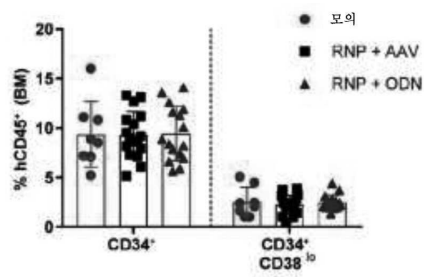
도면14c



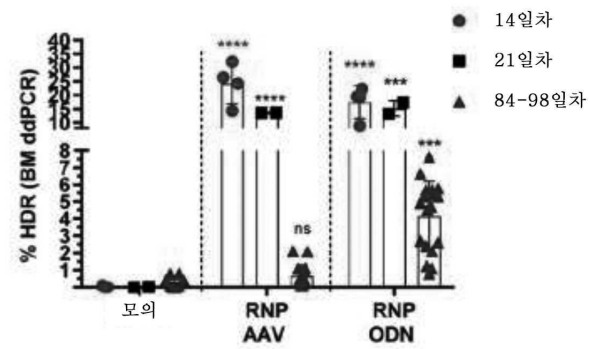
도면14d



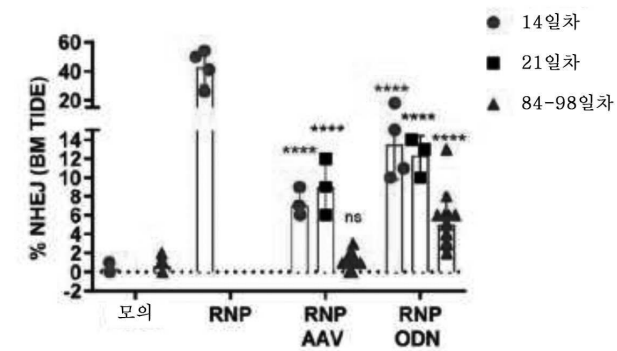
도면14e



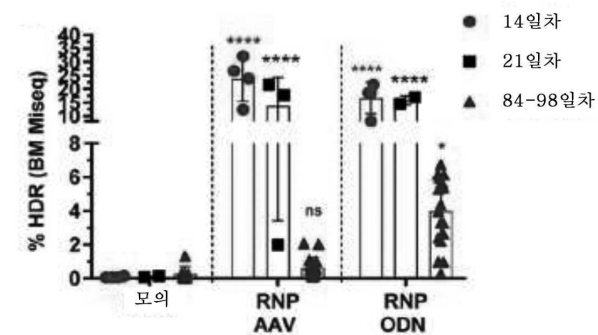
도면14f



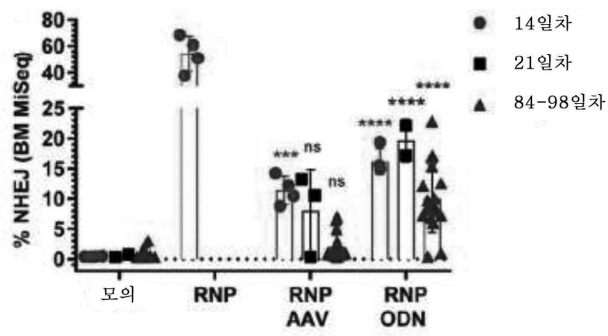
도면14g



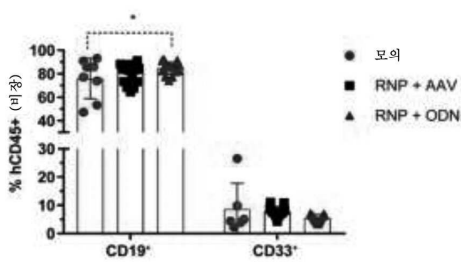
도면14h



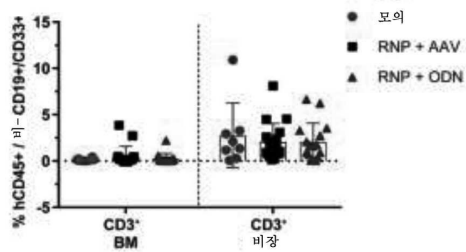
도면14i



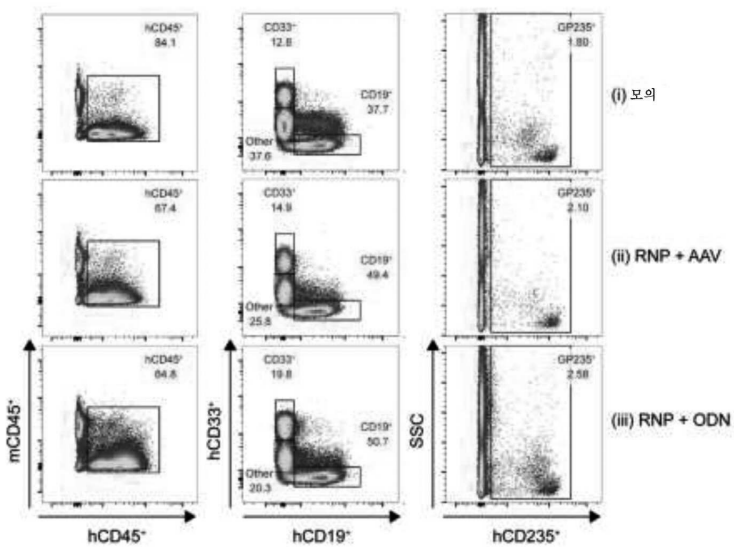
도면14j



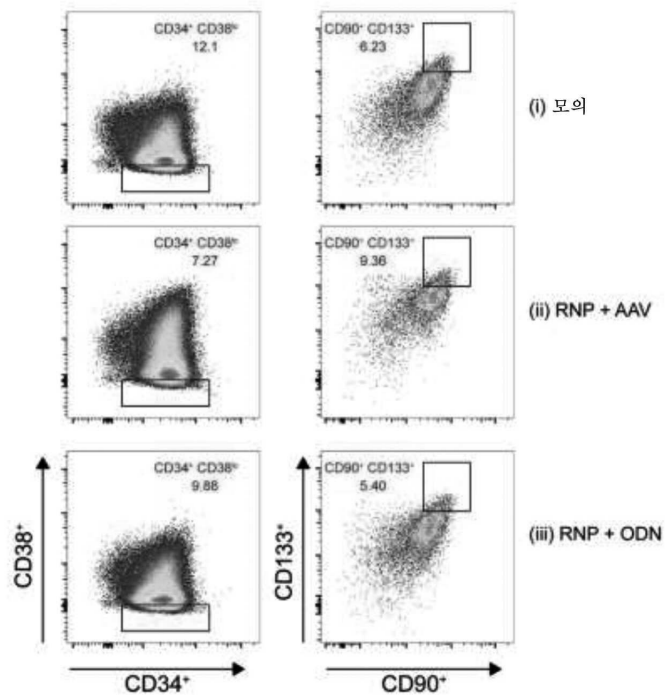
도면14k



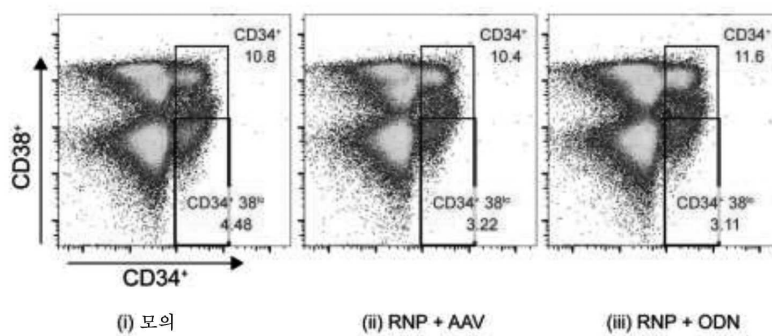
도면14l



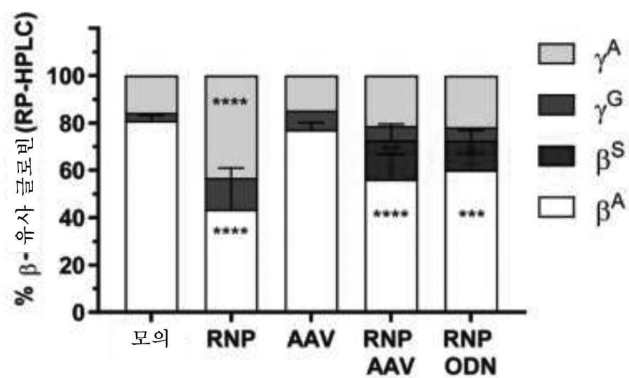
도면14m



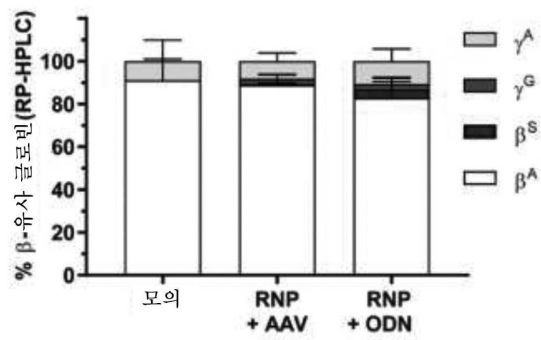
도면14n



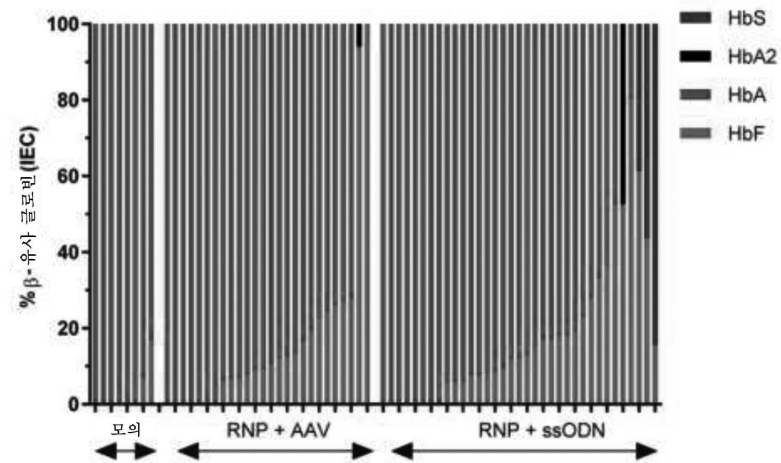
도면14o



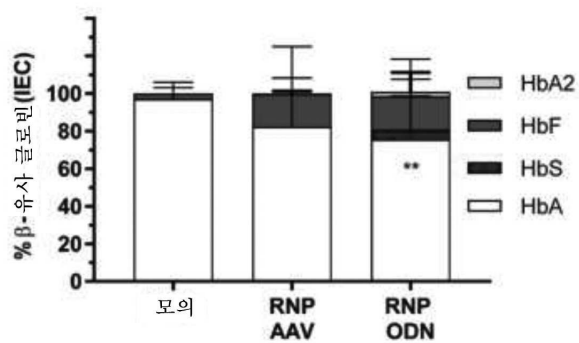
도면14p



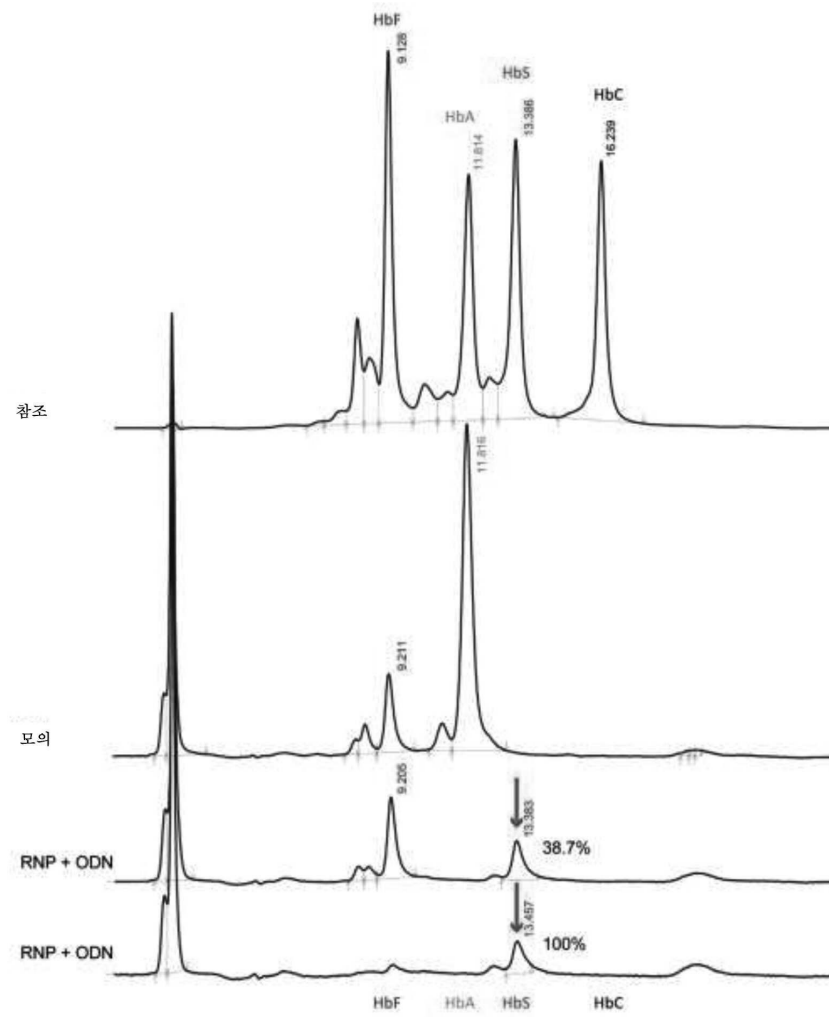
도면14q



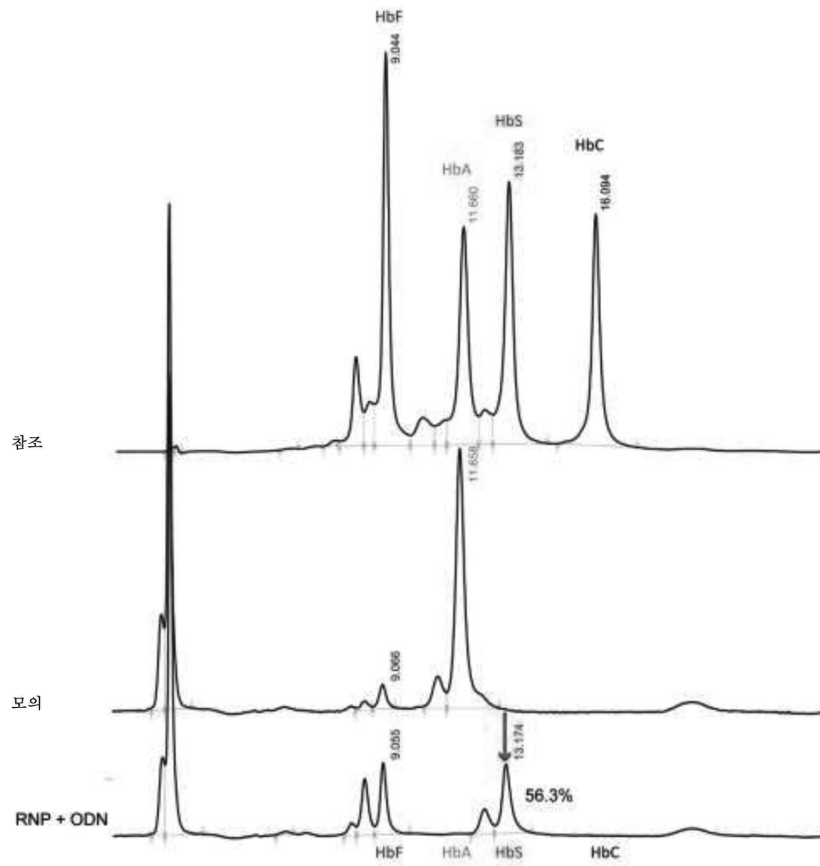
도면14r



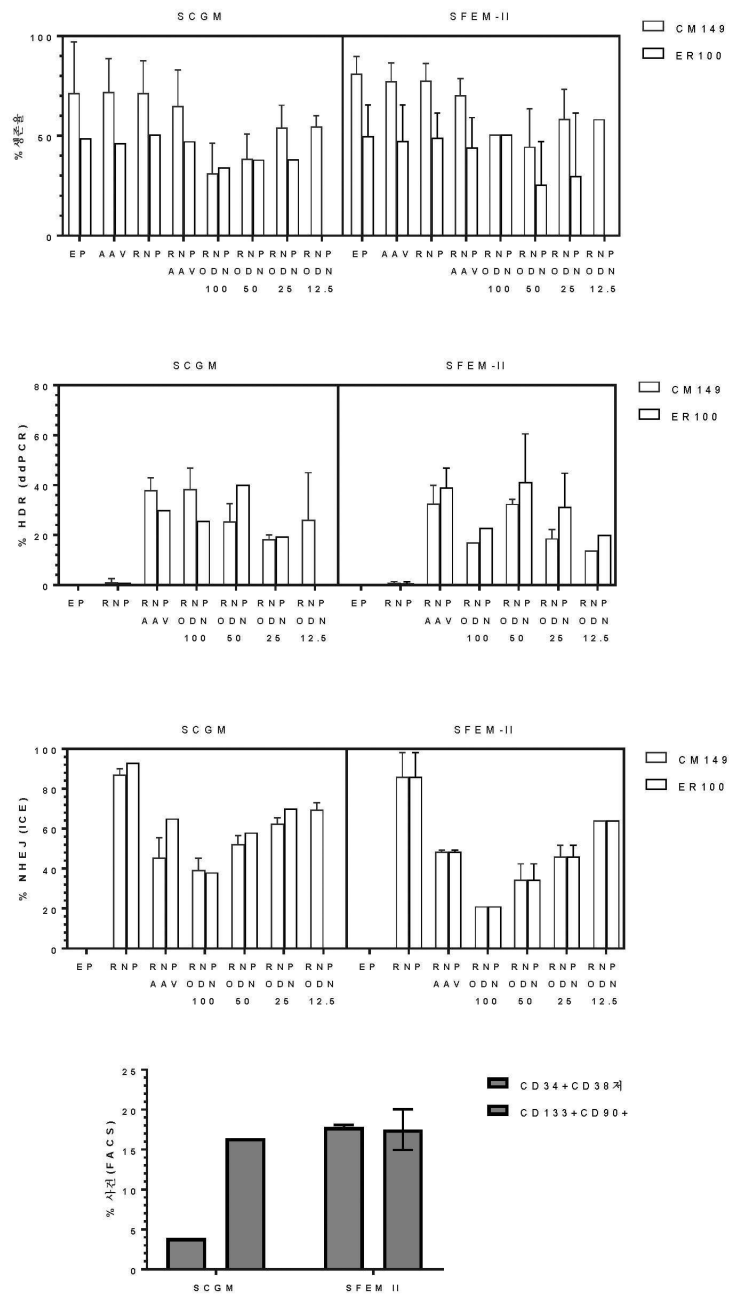
도면14s



도면14t

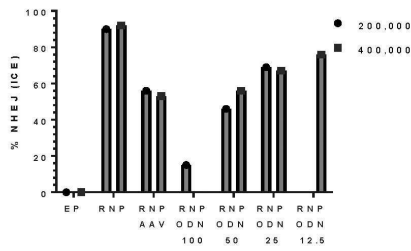


도면15a

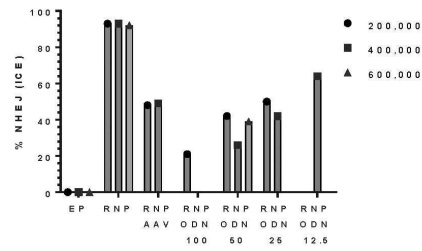


도면16c

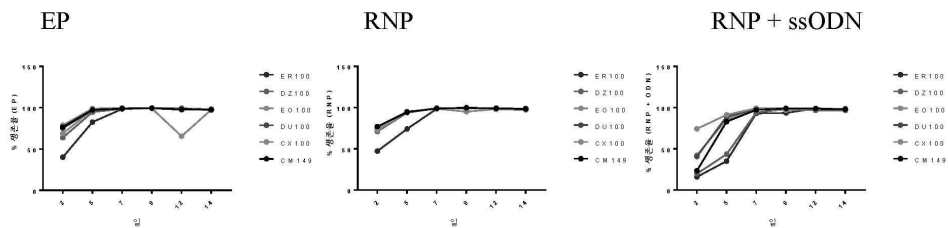
SFEM-II 배지: NHEJ (ICE)
CM149



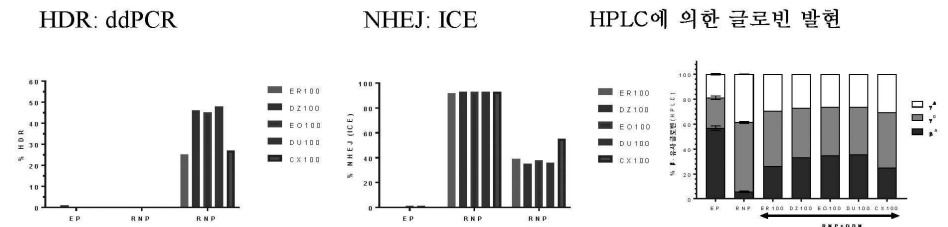
ER100



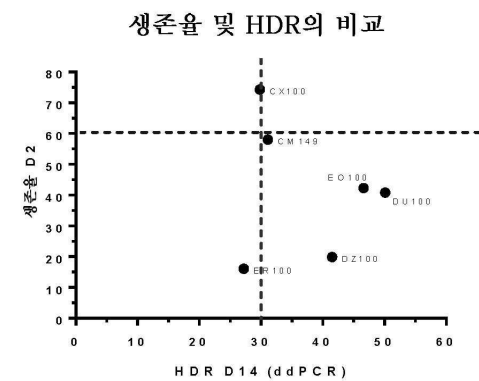
도면17a



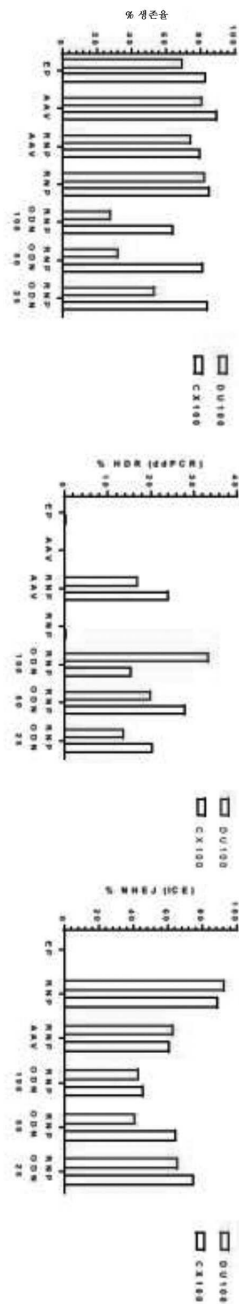
도면17b



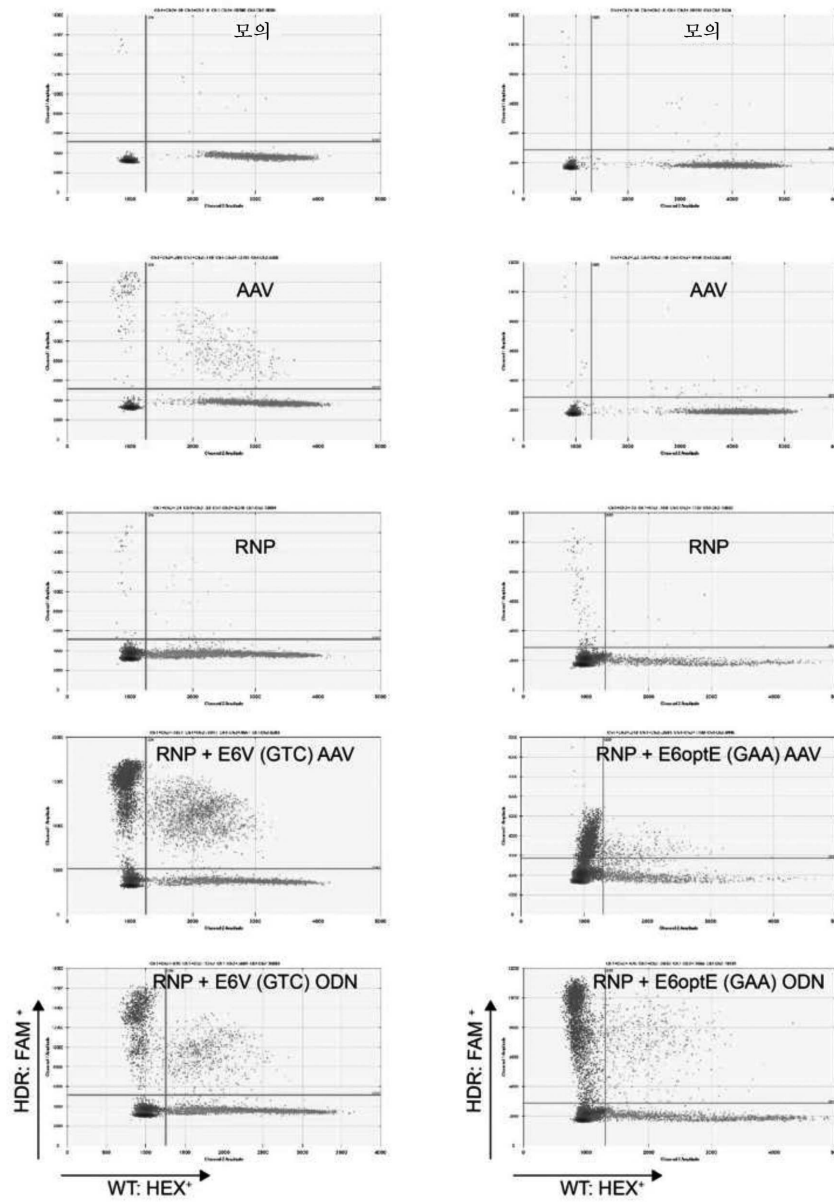
도면17c



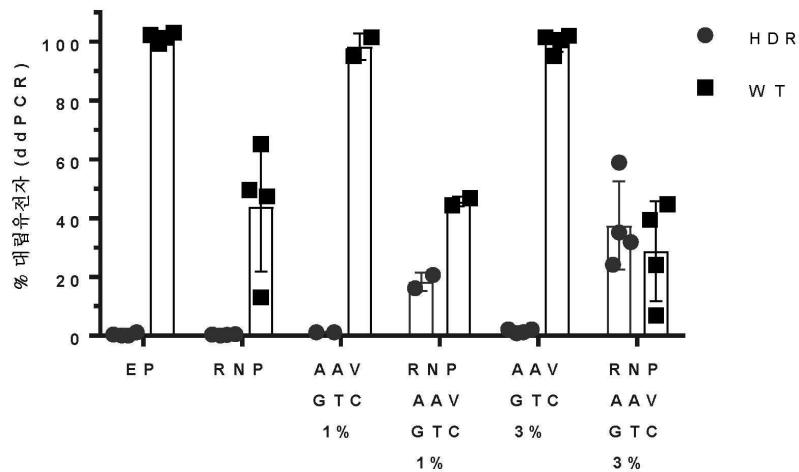
도면 18



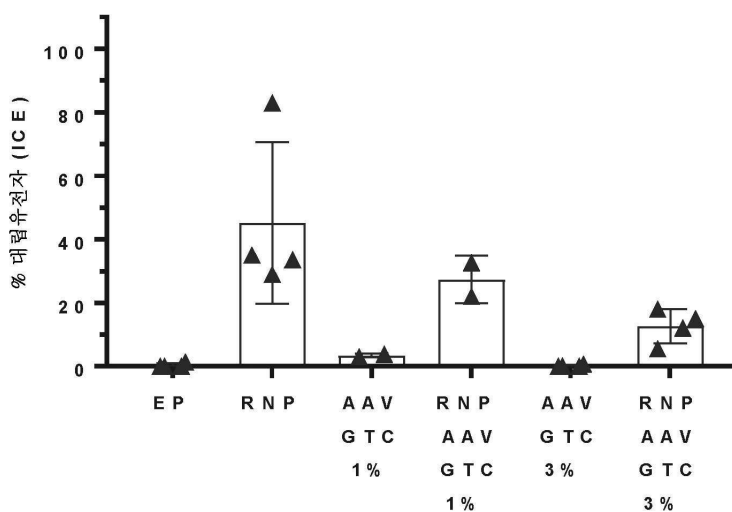
도면19



도면20a



도면20b



서열 목록

SEQUENCE LISTING

<110> David J. Rawlings

Sowmya Pattabhi

Andrew M. Scharenberg

Kyle Jacoby

Christopher Thomas Lux

<120> HOMOLOGY-DIRECTED REPAIR TEMPLATE DESIGN

AND DELIVERY TO EDIT HEMOGLOBIN-RELATED MUTATIONS

<130> SCRI.194WO

<150> 62663553

<151> 2018-04-27

<150> 62820521

<151> 2019-03-19

<160> 81

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> g1 or SCL-1

<400> 1

gtaacggcag acttctctc

20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> g2 or SCL-2

<400> 2

gtctgccgtt actgccctgt

20

<210> 3

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> g3 or SCL-3

<400> 3

tctgccgtta ctgccctgt

19

<210> 4

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> g4 or SCL-4

<400> 4

agtctgccgt tactgccctg

20

<210> 5

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> g5 or SCL-5

<400> 5

aaggtgaacg tggatgaagt

20

<210> 6

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> g6 or SCL-6

<400> 6

cttgccccac agggcagtaa

20

<210> 7

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-1 + gRNA scaffold

<400> 7

gtaacggcag acttctcctc gtttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 100

<210> 8

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-2 + gRNA scaffold

<400> 8

gtctgccgtt actgccctgt gtttagagc tagaaatagc aagttaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 100

<210> 9

<211> 99

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-3 + gRNA scaffold

<400> 9

tctgccgtta ctgccctgtg ttttagagc agaaatagca agttaaata aaggctagtc 60

gttatcaact tgaaaaagt gcaccgagtc ggtgctttt 99

<210> 10

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-4 + gRNA scaffold

<400> 10

agtcgccgt tactgccctg gtttagagc tagaaatagc aagttaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 100

<210> 11

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-5 + gRNA scaffold

<400> 11

aaggtgaacg tggatgaagt gttttagagc tagaaatagc aagttaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 100

<210> 12

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-6 + gRNA scaffold

<400> 12

cttgccccac agggcagtaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 100

<210> 13

<211> 3

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-1 PAM sequence

<400> 13

agg 3

<210> 14

<211> 3

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-2 PAM sequence

<400> 14

ggg

3

<210> 15

<211> 3

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-3 PAM sequence

<400> 15

ggg

3

<210> 16

<211> 3

<212> DNA

<213> Artificial Sequence

<220>

>

<223> SCL-4 PAM sequence

<400> 16

tgg

3

<210> 17

<211> 3

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-5 PAM sequence

<400> 17

tgg

3

<210> 18

<211> 3

<212> DNA

<213> Artificial Sequence

<220>

<223> SCL-6 PAM sequence

<400> 18

cgg

3

<210> 19

<211> 168

<212> DNA

<213> Artificial Sequence

<220>

<223> ssODN-V7E:GTC

<400> 19

tcagggcaga gccatctatt gcttacattt gcttctgaca caactgtgtt cactagcaac 60
ctcaaacaga caccatgggtg catctgactc ctgtcgagaa gtctgccgtt actgccctgt 120
ggggcaaggt gaacgtggat gaagttgggt gtgaggccct gggcaggt 168

<210> 20

<211> 168

<212> DNA

<213> Artificial Sequence

<220>

<223> ssODN-V7E:GTG

<400> 20

tcagggcaga gccatctatt gcttacattt gcttctgaca caactgtgtt cactagcaac 60
ctcaaacaga caccatgggtg catctgactc ctgtggagaa gtctgccgtt actgccctgt 120
ggggcaaggt gaacgtggat gaagttgggt gtgaggccct gggcaggt 168

<210> 21

<211> 168

<212> DNA

<213> Artificial Sequence

<220>

<223> ssODN

CCCGAA

<400> 21

tcagggcaga gccatctatt gcttacattt gcttctgaca caactgtgtt cactagcaac 60
ctcaaacaga caccatgggtg catctgactc ccgaagagaa gtctgccgtt actgccctgt 120
ggggcaaggt gaacgtggat gaagttgggtg gtgaggccct gggcaggt 168

<210> 22

<211> 15662

<212> DNA

<213> Artificial Sequence

<220>

<223> FrankenpEVL_SCL L Talen

<400> 22

gcgtataatg gactattgtg tgctgataag gagaacataa gcgcagaaca atatgtatct 60
attccggtgt tigtgttcctt tggtattctg ctattatgtt ctcttatagt gtgacgaaag 120
cagcataatt aatcgccact tgttctttga ttgtgttacg atatccagag acttagaaac 180
gggggaaccg ggatgagcaa ggtaaaaatc ggtgagtga tcaacacgct tgtgaatgag 240
gtagaggcaa ttgatgcctc agaccgcca caaggcgaca aaacgaagag aattaaagcc 300
gcagccgcac ggtataagaa cgcgttatit aatgataaaa gaaagttccg tgggaaagga 360
ttgcagaaaa gaataaccgc gaatactttt aacgcctata tgagcagggc aagaaagcgg 420

tttgatgata aattacatca tagctttgat aaaaatatta ataaattatc ggaaaagtat 480
cctctttaca gcgaagaatt atcttcatgg ctttctatgc ctacggctaa tattcgccag 540
cacatgtcat cgttacaatc taaattgaaa gaaataatgc cgcttgccga agagttatca 600
aatgtaagaa taggctctaa aggcagtgtg gcaaaaatag caagactaat aaaaaatat 660
ccagattgga gttttgctct tagtgattta aacagtgtatg attggaagga gcgccgtgac 720
tatctttata agttattcca acaaggctct gcgttggttag aagaactaca ccagctcaag 780
gtcaaccatg aggttctgta ccattctgcag ctaagccctg cggagcgtac atctatacag 840
caacgatggg ccgatgttct gcgcgagaag aagcgtaatg ttgtggttat tgactacca 900

acatacatgc agtctatcta tgatatittg aataatcctg cgactttatt tagtttaaac 960
actcgttctg gaatggcacc ttggccctt gctctggctg cggtatcagg gcgaagaatg 1020
attgagataa tgtttcaggg tgaatttgcc gtttcaggaa agtatacgtt taatttctca 1080
gggcaagcta aaaaacgctc tgaagataaa agcgtaacca gaacgattta tactttatgc 1140

gaagcaaaat tattcgttga attattaaca gaattgcgtt cttgctctgc tgcattgat 1200
 ttcgatgagg ttgttaaagg atatggaaag gatgatacaa ggtctgagaa cggcaggata 1260
 aatgctatatt tagcaaaagc atttaaccct tgggttaaat catttttcgg cgatgaccgt 1320
 cgtgtttata aagatagccg cgctatttac gctcgcatcg cttatgagat gttcttcgcg 1380

gtcgatccac ggtggaaaaa cgtcgacgag gatgtgttct tcatggagat tctcggacac 1440
 gacgatgaga acaccagct gcactataag cagttcaagc tggccaactt ctccagaacc 1500
 tggcgacctg aagttagggga tgaaaacacc aggctgggtg ctctgcagaa actggacgat 1560
 gaaatgccag gctttgccag aggtgacgct ggcgctccgtc tgcataaac cgtaagcag 1620
 ctggtggagc aggaccatc agcaaaaata accaacagca ctctccgggc ctttaaat 1680
 agcccgacga tgattagccg gtacctggag ttgcccgtg atgcattggg gcagttcgtt 1740
 ggcgagaacg ggcagtggca gctgaagata gagacacctg caatcgtcct gcctgatgaa 1800
 gaatccgttg aaaccatcga cgaaccggat gatgagtcct aagacgacga gctggatgaa 1860

gatgaaattg agctcgacga gggtagggc gatgaaccaa ccgaagagga agggccagaa 1920
 gaacatcagc caactgctt aaaacccgtc ttcaagcctg caaaaaataa cggggacgga 1980
 acgtacaaga tagagttaga atacgatgga aagcattatg cctggtccgg ccccgccgat 2040
 agccctatgg ccgcaatcg atccgatgg gaaacgtact acagctaaaa gaaaagccac 2100
 cgggtgtaat cggtaggctt ttatttagg cctgtcccta cccatccct gcaagggacg 2160
 gaaggattag gcggaactg cagctgcaac tacggacatc gccgtcccga ctgcaggac 2220
 ttccccgct aaagcgggc ttaaattcgg gctggccaac cctattttc tgcaatcgt 2280
 ggcatgtta gtttcgtga tagcgtttcc agcttttcaa tggccagctc aaaatgtgt 2340

ggcagcacct tctccagttc cgtatcaata tcggtgatcg gcagctctcc acaagacata 2400
 ctccggcgac cgccacgaac tacatcgcg agcagctccc gttcgtagac acgatgttg 2460
 cccagagccg tttctgcagc cgttaataatc cggcgagct cggcgatgat tgccgggaga 2520
 tcatccacgg ttattgggtt cggtagtggg ttctgcagg cgcggcgag agccatccag 2580
 acgccgctaa cccatgcgtt acggtactga aaactttgtg ctatgtcgtt tatcaggccc 2640
 cgaagtctt cttctgccg ccagtcagc ggttcaccgg cgttcttagg ctccagctcg 2700
 acaaaagcat actcgccgtt ttccggata gctggcagaa cctcgttcgt caccacttg 2760
 cggaaccgcc aggtgtcgt cccctgttc accgcgtcg ggacggag gattatggtg 2820

tagaggccag attccgatac cacatttact tcctggcca tccgatcaag tttttgtgc 2880
 tcggttaaac cgagggtcaa ttttcatca tgatccagct tacgcaatgc atcagaagg 2940
 ttggctatat tcaatgcagc acagatatcc agcgccaca accacgggtc accaccgaca 3000

agaaccaccc gtatagggtg gctttcctga aatgaaaaga cggagagagc cttcattgcg 3060
 cctccccgga tttcagctgc tcagaaaggg acaggagagc gccgcgagct tctgctgta 3120
 gtctgcgcgc gacctgcaga agttccgcag cttcctgcaa atacagcgtg gcctcataac 3180
 tggagatagt gcggtgagca gagcccacaa gcgcttcaac ctgcagcagg cgttcctcaa 3240
 tcgtcttcag caggccctgg gcgtttaact gaatctggtt catgcgatca cctcgtgac 3300

cgggataccg gctgacagaa cgaggacaaa acggctggcg aactggcgac gagcttctcg 3360
 ctccgatgat gcagtggtag aaaggcgggt gatatgggat tttttgtccg tgcggacgac 3420
 agctgcaaat ttgaatttga acatggtatg cattcctatc ttgtataggg tgctaccacc 3480
 agagttgaga atctctatag gggtagtagc ccagacaggg ttctcaacac cggtaacaaga 3540
 agaaaccggc ccaaccgaag ttggcccat ctgagccacc ataattcagg tatgcgcaga 3600
 tttaacacac aaaaaaacac gctggcgcgt gttgtgcgct tcttgcatt cggggttag 3660
 aggcccggt gcagattttg ctgcagcggg gtaactctac cgccaaagca gaacgcacgt 3720
 caataattta ggtggatatt ttaccccggt accagtcacg tgcacaggtg tttttatagt 3780

ttgctttact gactgatcag aacctgatca gttattggag tccgtaatc ttattgatga 3840
 ccgcagccac cttagatgtt gtctcaaac ccatacggcc acgaatgagc cactggaacg 3900
 gaatagtcag caggtacagc ggaacgaacc acaaacggtt cagacgctgc cagaacgtcg 3960
 catcacgacg ttccatccat tcggtattgt cgacgacctg gtaagcgtat tgcctggcg 4020
 tttttgctgc ttccgagtag caatctctt caccacaaag aaagtactt atctgcttc 4080
 agttttcgaa cctttctct ttgagccgt tttccagctc attcctccac aaaacaggca 4140
 cccatctct gcgataaatc atgattattt gtcctttaaa taaggctgta gaactgcaa 4200
 atcgtctctg ttacatgct gtacgtatg gcgtagcaaa ttgccgttcc atccctgtaa 4260

tccaccttct ttggaaagat cgtccttgac ctacacgaaga actttatcca atagccctgc 4320
 ggcacaagaa attgcctgct ctggatcagc aaattcatat tgattaatag gtgattgcca 4380
 cacacaaaa acaggaatca tcttttcggc taaacgcctc tctgttctt tcttaatctc 4440
 aagttgtaag cggaccagct caccatccat catTTTTTgt agatcatgcg ccactattca 4500
 cccccactgg ccatcagcaa ataaagcttc atactcggac accggcaggc ggcttcacg 4560
 gattgaaagg tcaagccaac cagtcacaga tgggtcagcc ttatccgatt ctccaccg 4620
 ttctgcagct gtagcaacca ggcatctac cgccttcatt tagtcttctg tacggaacca 4680
 gccgtagtta atgccacat cagtaactgc ccaggccatc ttttctctt cggcctcaat 4740

agcccgatg cggttatcgc acagctcgcg acagtacttc agctgttcgt aatccagttg 4800
 cttcaggaac tctggtgctg acgtcatagt ggcttcacct tataggcttt tagaagcgcc 4860

ctggcttcgt ctgtgtggtc ttccatgctc ttatcgctgg caatgcagca ataaactccc 4920
 tcactatctg agaaccctgt catccgaatg atcgtgaatg gaagttcccg gccagtttta 4980
 taatcgctat agcttgtcgc gtcgtggctg accttgacca cataagggtc gtagccctcc 5040
 acgatgacaa ggcatgcccg ttgttttccc attaccctc cggttatac gccacggctt 5100
 gccgctggct tagaaacgt ttcagcagcc ttatttcgcg tactgatagc aggtccataa 5160
 attcggtcat gtacagcgag gcgaacgttc tcgcatgct ggccactggc cacaggcgta 5220

 ccgcctccat ttcggttctt ggcaacgcgt tctccgcca cgcctccgt accgccaccg 5280
 ggatagcctc cagtgcctgg ataattactg attgtggggc gtcggaacg tgctctgttt 5340
 tggatcgagg gttaccatgt atatctatat ttagatccaa atcgcatcc acttcgatgg 5400
 tggttttttc caccttacgt gcgtgaattg ataaaccggc ctcgcgcgcg ttctccacga 5460
 tattcatgag gaactcgacc gagtcgggt caatggaacg catcgtgggg cgtgcatcgc 5520
 catctctggc gcgtctggtc ttactggata gcccataga ctccaggatg cctatgcaga 5580
 ggtctgcagg cgctttcttc ttgcctttct ctgtgttgaa gccgccgatg cgtaaacgt 5640
 tgttttagcag atcgcgccgt tccggcgtga gcaggttata tctggcgctt ttgagggcgt 5700

 ccatgtctgc ttcaccttcc agggtttttg gatcgatacc gcagtcgagg aagtactgct 5760
 gcagcgtcgc cgatttgagg gtgtagaaac cacgcatgcc tatctcaaca gcaggggtcg 5820
 atttcactcg gtaatcggtt atggccggga atttagcctg gaactctgcg tcggcctgtt 5880
 cccgctcat ggccgtatg acgaactgct gccatcttcc ggcaacgcga taagcgtagg 5940
 taaagtgaat caacgttct tcacggtaaa ggcgacgggc ggttatctca tccagctgca 6000
 tggtttcaaa caggcgact ttttcaggc cgccgtcgaa atagaatttt aacgccacct 6060
 cgtcgacatc cagctgcagc tccttttga tgtccagcg gaccagctgg gcctgtcat 6120
 ccagggacag ggtgcgtttt tttatcaact catcgtgttc ggctgtgtca ggagtatcga 6180

 cactcagggt gcgtccata agctgtcaa agaccagttc acgggcttct ttacgtaaat 6240
 ccttaccgat gctgtttgca agcgctcgg tgccatagg cgcgacctga tagccatcat 6300
 catgcatgat gcaaatcatg ttgctggcat aatcatttct ggccgatgcc tcgagcgagg 6360
 cggttttaat ttgagctgc atgaatgaag agttagccac gccagtgaa attcggtcac 6420
 cgtcaaagac aacgtctgtc agcagcccgg agtggccagc cgtttcgagc aaggcctgag 6480
 cgtaggcgag ttgatTTTT tccgatcgg tttcacgttt accggaagc ttgtcgaaac 6540
 cgataatgta ttctgagct gtacggtcgc ggcgagcat ctggatggcg tcgctgggga 6600
 ccacttcgcc gcagaacatg ccgaaatggc ggtggaagtg tttctctca atcgatacac 6660

 ctgaagatat cgacgggctg tagatgaggc cgtcatattt tttaccatc actttaggct 6720

ggttggtgaa atcgtcgact tccttctcct gtttgttttt ctggttaacg cagagaaact 6780
 ttttgcagg gaactgtagt ctcagctgca tggtaacgtc ttcggcgaac gtcgaactgt 6840
 cgggtggccag catgattcgt tcgccgcgtt gcactgcagc gataacctcg gtcattgatcc 6900
 gatttttctc ggtataaaat acgcggatag gcttgttggg ttcgcggttg cgaacgtcga 6960
 ccgggagttc aatcacgtga atttgcagcc aggcaggtag gccagctcc tcgcgtcgt 7020
 tcatgccag ttcagccagg tcaacaagca gatcgttggc atcgcatcc accataatgg 7080
 catgctcttc agtacgcgc agcgcgtcga taagcgtgtt gaatagcct accgggtttt 7140

ccatgcacg cccggccaga atggcacgca ggccctgtgt tgcttcatcg aagccgaaga 7200
 agtcatgctg gcgcatcagc ggttgcagc agcctttaag tatggagtgt atgcaaatag 7260
 tcagcttgtt ggcatatggc gccatttctt gatagccggg atcctgataa tgcagaatgt 7320
 cggttttcgc gcctttccct tcggtcatca tttcatgcag gccgcctatc agggatagc 7380
 ggtgcgcgac ggaaacgcca cgcgtggact gcagcatcag tggacgcagg aggcctgtcg 7440
 atttaccga ccccatcccg gcgcggacaa taacgatgcc ctgcagctgt gcggcgtatg 7500
 tcatcacctc atcggtcac ctcggaggttt caaacgttt gtaagtgtg tgtgacggcg 7560
 gaaggttcgg gttagtgatg cgttactga acgaacgtga tgtttgcgcg gcacggcatt 7620

tgcgattcaa ccggcgcgta atgtgatctt taacggtacc gttataaatt tctgcgatac 7680
 ccataccccg cagcgtgctg ctgaaaaggc gcataagtcc tttcgggctg tttggtaccg 7740
 ggcatgtcag catgccaata tcaacggcgc gaagcagttc tttggcaaaa gtgcgtctgt 7800
 tcagacgcgg gagagtacgc agcttattca gcgtgatcga caacagatcg gttgcacggc 7860
 tcagatgatt tctcgttaac tggcgcgca cttccttcag ccctctcagg ctgtgcaggt 7920
 cgttaaaatc gctgcattcc agctcagggt catcctcaaa agttgggtta acacatttga 7980
 cgccggaaaa cttctccatg atgtcgaatc cgggtcggag gcctgtgttg ctttttctt 8040
 cagctgagga ttgcggtcg ttatcgagag cgcaagtgt ttgcgcagcc gggtacatgt 8100

tcaccagctg ctcgacaacg tgaatcatgt tgttagcgga aaccgcaatg actaccgct 8160
 caaagcgttt tttcgggtcg tttctggctg ccagccagat ggatgccccg gtggcgaaac 8220
 cctctgcagt cgcaattttt tgcgccccct gcaggtcgcc aataacaaag catgcaccga 8280
 cgaaatcacc gttagtgtg gcgctggctt ggaacttgcc accattcaga tcgatacgtt 8340
 gccagccaac aatccgcccc tcttttcttc cgtccagggt ggacagaggt atcgccatgt 8400
 aagttgttgg tcacggctc catttcgcac tgcgtgact ggtcacgcga cgtatatcac 8460
 aagcgccaaa tacgtcacga attccctttt ttaccgcata aggccaggag ccattctcag 8520
 ctggcgaatg ttcccaggcg cgatggaaag ccaaccatcc aagcaggcgt tcctgctcca 8580

tctgattgtt ttttaaatca ttaacgcgtt gttgttcagc tccgaggcgg cgtgcttcag 8640
 cctggcgctc catgcgtgca cgttcttctt ccggctgagc gaccacggtc gcaccattcc 8700
 gttgctgttc acggcgatac tccgaaaaca ggaatgaaaa gccactccag gagccagcgt 8760
 catgcgcttt ttcaacgaag ttaacgaaag gataactgat gccatccttg ctctgctcaa 8820
 ggctgaata gatttcaca cggcctttaa ggctcttctg cagagcttcc ggggaggaat 8880
 tattgtaggt ggtatagcgc tctacaccac cgcgcggatt gagctgaatc ttatcagcac 8940
 acgcaggcca gttgataccg gccatcttcg ccagctcagt cagctcatca cgtgccgcgt 9000
 caagcagtga aaacggatcg ctgccaaagc gctccgcgta gaattcttgt aaggctcattt 9060

tttagccttt ccatgcgaat tagcatTTTT tccgggttgaa aaaatccgca ggagcagcca 9120
 caataaacgc actatcttcc tgaaggacgt atctgcgtta tcgtggctac ttcctgaaaa 9180
 aggcccgagt ttccgactc ggtttttttt tcgtcttttt tccgctgcta cggctctggtt 9240
 caaccccgac aaagtataga tccgattaaa ccagaattat agtcagcaat aaacctgtt 9300
 attgtatcat ctacctcaa ccatgaacga ttgatcgtta ccgactactt ggtgcacaaa 9360
 ttgaagatca cttttatcat ggataaccg ttgagagtta gcactatcaa ggtagtaatg 9420
 ctgctcgtca taacgggcta atcgttgaat tggatctcgc ccgttattat caciaaccag 9480
 tacatcctca cccggtacaa gcgtaagtga agaatcgacc aggataacat ctcccgctg 9540

gtagtttcgc tgaatctggt tcccagcgt cagtgcgtaa acggtgttcc gttgactcac 9600
 gaacggcagg aatcgctctg tggtggcagg ttctccaggc tgccagtctc tatccgttcc 9660
 tgtctctgtc gtaccaataa caggaacgcg gctcggatca gattcagtgc catacagtat 9720
 ccattgcacg ggcttacgca ggcatTTTT cagcgatagc ccgatctcca gcgacggcat 9780
 cacgtcgcca cgttctaagt ttggacgcc cggaagagag attctacag cttctgccac 9840
 ttgcttcagc gtcagtttca gctctaaacg gcgtgcttcc agtcgttcgc ctctgtttt 9900
 cataccctta atcataaatg atctctttat agctggctat aatttttata aattatacct 9960
 agctttaatt ttcatttatt gattataata atcccatga aaccgaaga acttgtgcgc 10020

catttcggcg atgtggaaaa agcagcgggt ggctggggcg tgacacccgg cgcagtctat 10080
 caatggctgc aagctgggga gattccacct ctacgacaaa gcgatataga ggtccgtacc 10140
 gcgtacaaat taaagatga ttccacctct cagcgcatgg gtaaggaagg gcataacaag 10200
 gggatcctct agacgcagaa aggccacccc gaaggtgagc cagtgtgatt acatttgccg 10260
 cctaactgtg gccagtccag ttacgtgga gtcactagtgc cggccgcgac aacttgtcta 10320
 gggcccaatg gccatacac ttagtgtaat acgactcact ataggagag cggccgcttt 10380
 ttcagcaaga ttaagccgcc accatggcgc cgcggcctcc taagaagaag cggaaagtgc 10440

aattcgtgga tctgcgaaca ctgggctata gccagcagca gcaggagaag atcaaaccce 10500
aggtgaggtc cacagtcgca cagcaccatg aagccctggt gggccacggg ttcactcacg 10560
ctcatattgt cgcactgtct cagcatccag ccgctctggg aaccgtggca gtcacatacc 10620
agcacatcat tactgccctg cccgaggctc cccatgaaga catcgtggga gtcggcaaac 10680
agtggagcgg cgcacgggcc ctggaggctc tgctgaccga cgcaggggaa ctgagaggac 10740
ccccctcgca gctggatata gggcagctgg tgaagattgc taagagggga ggggtgacag 10800
caatggaagc cgtccacgca agcaggaacg cactgacagg ggccccctg aacctgactc 10860

cggaccaagt ggtggctatc gccagcaacc acggcggcaa gcaagcgctc gaaacgggtc 10920
agcggctgtt gccggtgctg tgccaggacc atggcctgac tccggaccaa gtggtggcta 10980
tcgccagcca cgatggcggc aagcaagcgc tcgaaacggt gcagcggctg ttgccggtgc 11040
tgtgccagga ccatggcctg accccggacc aagtgggtgc tatcgccagc aacattggcg 11100
gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag gaccatggcc 11160
tgactccgga ccaagtgggt gctatcgcca gccacgatgg cggcaagcaa gcgctcgaaa 11220
cgggtgcagc gctgttgccg gtgctgtgcc aggacatgg cctgactccg gaccaagtgg 11280

tggttatcgc cagccacgat ggcggcaagc aagcgctcga aacggtgcag cggctgttgc 11340
cgggtgctgtg ccaggaccat ggctgaccc cggaccaagt ggtggctatc gccagcaacg 11400
gtggtggaaa acaggccctt gaaacgggtc agcggctgtt gccggtgctg tgccaggacc 11460
atggcctgac tccggaccaa gtggtggcta tcgccagcaa ccacggcggc aagcaagcgc 11520
tcgaaacggt gcagcggctg ttgccggtgc tgtgccagga ccatggcctg accccggacc 11580
aagtgggtgc tatcgccagc aacattggcg gcaagcaagc gctcgaaacg gtgcagcggc 11640
tgttgccggt gctgtgccag gaccatggcc tgactccgga ccaagtgggt gctatcgcca 11700

gccacgatgg cggcaagcaa gcgctcgaaa cgggtgcagc gctgttgccg gtgctgtgcc 11760
aggacatgg cctgaccccg gaccaagtgg tggttatcgc cagcaacggt ggcggcaagc 11820
aagcgctcga aacggtgcag cggctgttgc cgggtgctgtg ccaggaccat ggctgaccc 11880
cggaccaagt ggtggctatc gccagccacg atggcggcaa gcaagcgctc gaaacgggtc 11940
agcggctgtt gccggtgctg tgccaggacc atggcctgac tccggaccaa gtggtggcta 12000
tcgccagcca cgatggcggc aagcaagcgc tcgaaacggt gcagcggctg ttgccggtgc 12060
tgtgccagga ccatggcctg accccggacc aagtgggtgc tatcgccagc aacggtggcg 12120

gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag gaccatggcc 12180
tgactccgga ccaagtgggt gctatcgcca gcaaccacgg cggcaagcaa gcgctcgaaa 12240
cgggtgcagc gctgttgccg gtgctgtgcc aggacatgg cctgaccccg gaccaagtgg 12300

tggctatcgc cagcaacatt ggcggcaagc aagcgctcga aagcattgtg gcccagctga 12360
 gccggcctga tccggcggtg gccgcgttga ccaacgacca cctggtcgct ctggcttgcc 12420
 tgggaggacg ccttgcctatg gacgctgtga agaaaggact gccccacgca cccgaactga 12480
 ttagacgggt gaaccggaga atcggcgaga gaacatccca tagggtggca atctctagaa 12540

ctcagctggt caagagtga ctggaggaaa agaaatcaga gctgcgccac aagctgaaat 12600
 acgtgcctca tgagtatatc gaactgatcg agattgctcg caattcaacc caggaccgga 12660
 tcctggaaat gaaagtgatg gagttcttta tgaaagtcta cggatatcgg gggaaacacc 12720
 tgggaggag cagaaagcca gatggggcca tctacacagt gggatcccc atcgactatg 12780
 gcgtgattgt cgatactaaa gcctacagcg gaggtataa cctgcctatc ggccaggctg 12840
 acgagatgca gagatactg gaggaaaacc agaccgcaa taagcatatt aacccaatg 12900
 aatggtggaa agtgtatcct agctccgtca cagagttcaa gtttctgttc gtgagcggac 12960

actttaagg caactacaaa gcacagctga ctaggctgaa tcatatcacc aactgcaatg 13020
 gagccgtgct gtctgtcgag gaactgctga tcgggggaga gatgattaag gctggcacac 13080
 tgactctgga ggaagtgagg cgcaagttca acaatgggga aatcaacttc taacctgcag 13140
 gatgataagc tagccccggg cgtacggaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13200
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13260
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13320
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13380

aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aacgagacct 13440
 tagggccatt agacttgaag tcaagcggcc gcttacaact ggaccttgct ggtacataga 13500
 actgattaac tgaccattta aatcatacca acatgggtcaa ataaacgaa aggctcagtc 13560
 gaaagactgg gcctttcgtt ttaatctgat cggcacgtaa gaggttccaa ctttcacat 13620
 aatgaaataa gatcactacc gggcgtatct tgagttatcg agattttcag gagctaagga 13680
 agctaaaatg agccatatc aacgggaaac gtcttgctcg aggccgcgat taaattccaa 13740
 catggatgct gatttatatg ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc 13800

gacaatctat cgattgtatg ggaagcccg tgcgccagag ttgtttctga aacatggcaa 13860
 aggtagcgtt gccaatgatg ttacagatga gatggtcagg ctaaactggc tgacggaatt 13920
 tatgcctctt ccgaccatca agcattttat ccgtactcct gatgatgcat ggttactcac 13980
 cactgcgatc ccagggaata cagcattcca ggtattagaa gaatatcctg attcaggtga 14040
 aaatatgtt gatgcgctgg cagtgttct gcgccgttg cattcgattc ctgtttgtaa 14100
 ttgtcctttt aacggcgatc gcgtatttcg tctggctcag gcgcaatcac gaatgaataa 14160

cggttttggtt ggtagcgagt attttgatga cgagcgtaat ggctggcctg ttgaacaagt 14220

ctggaaagaa atgcataaac ttttgccatt ctaccggat tcagtcgtca ctcatggtga 14280

ttttcactt gataacctta tttttgacga ggggaaatta ataggttgta ttgatgttgg 14340

acgagtcgga atcgagacc gataccagga tcttgccatc ctatggaact gcctcgggtga 14400

gttttctcct tcattacaga aacggctttt tcaaaaatat ggtattgata atcctgatat 14460

gaataaattg cagtttctact tgatgctcga tgagtttttc taacctaggt gacagaagtc 14520

aaaagcctcc ggtcggaggc ttttgacttt ctgctagatc tgtttcaatg cggatgaagg 14580

ccaggcagct ggggattatg tccagaccg gccagcatgt tggttttatc gcatattcag 14640

cgttgtcgcg ttiaccagg taaaatggaa gcagtgtatc gtcgtcgtga atgtgcaaat 14700

caggaacgta accgtggtac atagatgcag tccttgcgg gtcgttcct tcaacagta 14760

ggacgcggtg ccttgcaag gctaaccatt gcgcctggtg tactgcagat gaggttttat 14820

aaacccctcc ctgtgtgac ataacggaaa gtacaaccgg gtttttatcg tcaggtcttt 14880

ggtttgggtt accaaacaca ctccgcatat ggctaatttg gtcaattgtg tagccagcgc 14940

gacgttctac tcggccctc atctcaaat caggagccgg tagacgacca gctttttcgc 15000

catctctgat agcctgcgtt gttacgccga tcaggtctgc aacttctgtt atacccagc 15060

ggcgagtaat acgacgcgt tccgggctgt catgccgaa ctgtgcgatg gcaatagcgc 15120

gcgtcatttc ctgaccgca ttgatacagt ctttcagca attaatcaac gacatcctgt 15180

ttcctctcaa acatgccctt atctttgtgt ttttcatcat actttacgtt tttaaagca 15240

agcaacataa aaaaagcaaa gtgacttaga aaacgcaaag ttaaggttca aatcaatttt 15300

ttgatgcgt acagaagcta tttagcttca tctaagcga acggtattac ttacgttgg 15360

atatttaaaa ctaacttaa tgattttaaa tgataataaa tcataccaat tgctatcaaa 15420

agttaagca acatgctgat tttcacgtg tttatacact ttgaggcatc tetatctctt 15480

ctgtctctat attgaaacac aatcaaagaa catcaatcca tgtgacatcc cccactatct 15540

aagaacacca taacagaaca caacatagga atgcaacatt aatgtatcaa taattcggaa 15600

catatgcact atatcatatc tcaattacgg aacatatcag cacacaattg cccattatac 15660

gc 15662

<210> 23

<211> 15866

<212> DNA

<213> Artificial Sequence

<220>

<223> FrankenpEVL_SCL R Talen

<400> 23

gcgtataatg gactattgtg tgctgataag gagaacataa gcgcagaaca atatgtatct 60
attccggtgt tgtgttcctt tgttattctg ctattatgtt ctcttatagt gtgacgaaag 120

cagcataatt aatcgccact tgtttcttga ttgtgttacg atatccagag acttagaaac 180
gggggaaccg ggatgagcaa ggtaaaaatc ggtgagttga tcaacacgct tgtgaatgag 240
gtagaggcaa ttgatgcctc agaccgccca caaggcgaca aaacgaagag aattaaagcc 300
gcagccgcac ggtataagaa cgcgttatit aatgataaaa gaaagtccg tgggaaagga 360
ttgcagaaaa gaataaccgc gaatactttt aacgcctata tgagcagggc aagaaagcgg 420
tttgatgata aattacatca tagctttgat aaaaatatta ataaattatc ggaaaagtat 480
ctcttttaca gcgaagaatt atcttcattg ctttctatgc ctacggctaa tattcgccag 540
cacatgtcat cgttacaatc taaattgaaa gaaataatgc cgcttgccga agagttatca 600

aatgtaagaa taggctctaa aggcagtgat gcaaaaatag caagactaat aaaaaatat 660
ccagattgga gttttgctct tagtgattta aacagtgatg attggaagga gcgccgtgac 720
tatctttata agttattcca acaaggctct gcgttggttag aagaactaca ccagctcaag 780
gtcaaccatg aggttctgta ccatctgcag ctaagccctg cggagcgtac atctatacag 840
caacgatggg ccgatgttct gcgcgagaag aagcgtaatg ttgtggttat tgactacca 900
acatacatgc agtctatcta tgatattttg aataatcctg cgactttatt tagtttaaac 960
actcgttctg gaatggcacc tttggccttt gctctggctg cggatcagg gcgaagaatg 1020
attgagataa tgtttcaggg tgaatttggc gtttcaggaa agtatacggt taattttca 1080

gggcaagcta aaaaacgctc tgaagataaa agcgtaacca gaacgattta tactttatgc 1140
gaagcaaaat tttcgttga attattaaca gaattgcgtt cttgctctgc tgcattctgat 1200
ttcgatgagg ttgttaaagg atatggaaag gatgatacaa ggtctgagaa cggcaggata 1260
aatgctatit tagcaaaagc atttaacctt tgggttaaatt catttttcgg cgatgaccgt 1320
cgtgtttata aagatagccg cgctatttac gctcgcatcg cttatgagat gttcttcgc 1380
gtcgatccac ggtggaaaaa cgtcgacgag gatgtgttct tcatggagat tctcggacac 1440
gacgatgaga acaccagct gcactataag cagttcaagc tggccaactt ctccagaacc 1500
tggcgacctg aagttgggga tgaaaacacc aggctggtgg ctctgcagaa actggacgat 1560

gaaatgccag gctttgccag aggtgacgtt ggcgctccgtc tgcatgaaac cgttaagcag 1620
ctggtggagc aggacccatc agcaaaaata accaacagca ctctccgggc ctttaattt 1680

agccccacga tgattagccg gtacctggag ttgcccgtg atgcattggg gcagttcgtt 1740
ggcgagaacg ggcagtggca gctgaagata gagacacctg caatcgtcct gcctgatgaa 1800
gaatccgttg aaaccatcga cgaaccgat gatgagtcct aagacgacga gctggatgaa 1860
gatgaaattg agctcgacga ggggtggcggc gatgaaccaa ccgaagagga agggccagaa 1920
gaacatcagc caactgctct aaaacccgtc ttcaagcctg caaaaaataa cggggacgga 1980
acgtacaaga tagagtttga atacgatgga aagcattatg cctggtccgg ccccgcgat 2040

agccctatgg ccgcaatgcg atccgcatgg gaaacgtact acagctaaaa gaaaagccac 2100
cgggtgttaat cgggtggcttt ttatttgagg cctgtcccta cccatcccct gcaagggacg 2160
gaaggattag gcggaactg cagctgcaac tacggacatc gccgtccga ctgcaggac 2220
ttccccgct aaagcggggc ttaaatcgg gctggccaac cctattttc tgcaatcgt 2280
ggcgatgtta gtttcgtgga tagcgtttcc agcttttcaa tggccagctc aaaatgtgct 2340
ggcagcacct tctccagttc cgtatcaata tcggtgatcg gcagctctcc acaagacata 2400
ctccggcgac cgccacgaac tacatcgcg agcagctccc gttcgtagac acgcatgtt 2460
cccagagcgg tttctgcagc cgttaatac cggcgagct cggcgatgat tgccgggaga 2520

tcatccacgg ttattgggtt cgggtgatggg ttctgcagg cgcgcgagg agccatccag 2580
acgccgctaa cccatgcgtt acggtactga aaactttgtg ctatgtcgtt tatcaggccc 2640
cgaagtctt tttctgccg ccagtcaggt ggttcaccgg cgttcttagg ctgagctcg 2700
acaaaagcat actcgccgtt tttccggata gctggcagaa cctcgttcgt caccacttg 2760
cggaaccgcc aggtgtcgt cccctgttt accgcgtcg ggcagcggag gattatggtg 2820
tagaggccag attccgatac cacatttact tccctggcca tccgatcaag tttttgtgcc 2880
tcggttaaac cgaggttcaa ttttcatca tgatccagct tacgcaatgc atcagaagg 2940
ttggctatat tcaatgcagc acagatatcc agcgccacaa accacgggtc accaccgaca 3000

agaaccacc gtatagggtg gctttcctga aatgaaaaga cggagagagc cttcattgcg 3060
cctccccga tttcagctgc tcagaaagg acaggagca gccgcgagct tctgcgtga 3120
gttcgcgcg gacctgcaga agttccgcag cttcctgcaa atacagcgtg gcctcataac 3180
tggagatagt gcggtgagca gagcccacaa gcgcttcaac ctgcagcagg cgttcctcaa 3240
tcgtcttcag caggccctgg gcgtttaact gaatctggtt catgcgatca cctcgtgac 3300
cgggatcagg gctgacagaa cgaggacaaa acggtggcg aactggcgac gagcttctcg 3360
ctcgatgat gcagtgggtg aaaggcgtg gatatgggat ttttgtccg tgcggacgac 3420
agctgcaaat ttgaattga acatggtatg cattcctatc ttgtataggg tgctaccacc 3480

agagttgaga atctctatag ggggtgtagc ccagacaggg ttctcaacac cggtaaca 3540

agaaaccggc ccaaccgaag ttggcccat ctgagccacc ataattcagg tatgcgcaga 3600
 tttaacacac aaaaaaacac gctggcgcgt gttgtgcgt tcttgcatt cggggttag 3660
 aggcccggt gcagatttg ctgcagcggg gtaactctac cgccaaagca gaacgcacgt 3720
 caataattta ggiggatatt ttacccgtg accagtcacg tgcacagggtg tttttatagt 3780
 ttgctttact gactgatcag aacctgatca gttattggag tccgtaatc ttattgatga 3840
 ccgcagccac cttagatgtt gtetcaaacc ccatacggcc acgaatgagc cactggaacg 3900
 gaatagtcag caggtacagc ggaacgaacc acaaacggtt cagacgctgc cagaacgtcg 3960

catcacgacg ttccatccat tcggtattgt cgacgacctg gtaagcgtat tgtcctggcg 4020
 tttttgctgc ttccgagtag caatcctctt caccacaaag aaagtactt atctgcttc 4080
 agttttcgaa cccttctctt ttgagccgct tttccagctc attcctccac aaaacaggca 4140
 cccatcctct gcgataaate atgattatit gtcttttaaa taaggctgta gaactgcaa 4200
 atcgctctcg ttcacatgct gtacgtagat gcgtagcaaa ttgccgttcc atccctgtaa 4260
 tccaccttct ttggaaagat cgtccttgac ctacgaaga actttatcca atagccctgc 4320
 ggcaacgaa attgcctgct ctggatcagc aaattcatat tgattaatag gtgattgcca 4380
 cacacaaaa acaggaatca tcttttcggc taaacgcctc tctgttctt tcttaatctc 4440

aagttgtaag cggaccagct caccatccat catTTTTTgt agatcatgcg ccactattca 4500
 cccccactgg ccatcagcaa ataaagctt atactcggac accggcagcg ggcttccacg 4560
 gattgaaagg tcaagccaa cagtcagga tgggtcagcc ttatccgatt ctcccaccg 4620
 ttctgcagct gtagcaacca ggcatctac cgccttcag tagtcttctg tacggaacca 4680
 gccgtagtta atgccacat cagtaactgc ccaggccatc ttttctctt cggcctcaat 4740
 agcccgatg cggttatcg acagctcgc acagtactt agctgttcgt aatccagttg 4800
 cttcaggaac tctggtgtcg acgtcatagt ggcttcacct tataggcttt tagaagcgcc 4860
 ctggcttcgt ctgtgtggtc ttccatgctc ttatcgctgg caatgcagca ataaactccc 4920

tcactatctg agaaccggt catccgaatg atcgtgaatg gaagtcccg gccagtttta 4980
 taatcgctat agcttgcgc gtcgtggctg acctgacca cataagggtc gtagecctec 5040
 acgatgaaa ggcatcccg ttgttttccc attaccctc cggttatata gccacggctt 5100
 gccgtggct tagaaacgt ttcagcagcc ttatttcgcg tactgatagc aggtccataa 5160
 attcggctat gtacagcgag gcgaacgtt tcgcgatgct ggccactggc cacaggcgta 5220
 ccgcctccat ttcggttgc ggcaacgcgt tctccgcca cgcctcgggt accgccaccg 5280
 ggatagcctc cagtgcctgg ataattactg attgtggggc gtcggaacg tgctctgttt 5340
 tggatcgagg gttaccatgt atatctatat ttagatccaa atcgcatcc acttcgatgg 5400

tggttttttc caccttacgt gcgtgaattg ataaaccggc ctgcggcgc tctccacga 5460
tattcatgag gaactcgacc gagtccgggt caatggaacg catcgtgggg cgtgcatcgc 5520
catctctggc gcgtctggtc ttactggata gcccataga ctccaggatg cctatgcaga 5580
ggctctcagg cgctttcttc ttgcctttct ctgtgttgaa gccgccgatg cgtaaaacgt 5640
tgttttagcag atcgcgccgt tccggcgtga gcaggttatc tctggcgcgt ttgagggcgt 5700
ccatgtctgc ttcaccttcc agggtttttg gatcgatacc gcagtcgcgg aagtactgct 5760
gcagcgtcgc cgatttgagg gtgtagaaac cacgcatgcc tatctcaaca gcaggggtcg 5820
atttcactcg gtaatcggtt atggccggga atttagcctg gaactctgcg tcggcctgtt 5880

cccgcgtcat ggccgtagtg acgaactgct gccatcttcc ggcaacgcga taagcgtagg 5940
taaagtgaat caacgcttct tcacggtaa ggcgacgggc ggttatctca tccagctgca 6000
tggtttcaaa caggcgact ttttcaggc cgcgctcga atagaatttt aacgccacct 6060
cgtcgacatc cagctgcagc tccttttcga tgtcccagcg gaccagctgg gcctgctcat 6120
ccagggacag ggtgcgtttt tttatcaact catcgtgttc ggcttggtca ggagtatcga 6180
cactcagggtg gcgtccata agctgtccta agaccagttc acgggcttct ttacgtaaat 6240
ccttaccgat gcgttttga agcgcgtcgg tggccatagg cgcgacctga tagccatcat 6300
catgcatgat gcaaatcatg ttgctggcat aatcatttct ggccgatgcc tcgagcgcgg 6360

cggttttaat tttagctgc atgaatgaag agttagccac gccgagtga attcggtcac 6420
cgtcaaagac aacgtctgtc agcagcccgg agtggccagc cgtttcgagc aaggcctgcg 6480
cgtaggcgcg ttgattttt tccggatcgg tttcacgttt accgcgaagc ttgtcgaaac 6540
cgataatgta ttctgagct gtacggtcgc ggcgagcat ctggatggcg tcgctgggga 6600
ccacttcgcc gcagaacatg ccgaaatggc ggtggaagtg tttctctca atcgatacac 6660
ctgaagatat cgacgggctg tagatgaggc cgtcatattt tttcaccatc actttaggct 6720
ggttggtgaa atcgtcgact tccttctcct gtttgtttt ctggttaacg cagagaaact 6780
ttttgtcagg gaactgtagt ctacgtgca tggtaacgtc ttcggcgaac gtcgaactgt 6840

cggtggccag catgattcgt tcgccgcgtt gcactgcagc gataacctcg gtcatgatec 6900
gatttttctc ggtataaaat acgcggatag gcttgttggc ttcgcggttg cgaacgtcga 6960
ccgggagttc aatcacgtga atttcagcc aggcaggtag gccagctcc tcgcgtcgt 7020
tcacgccag ttcagccagg tcaacaagca gatcgttggc atcgcatcc accataatgg 7080
catgctcttc agtacgcgcc agcgcgtcga taagcgtgtt gaatacgcct accgggtttt 7140
ccatgcacg cccggccaga atggcacgca ggccctgtgt tgcttcatcg aagccgaaga 7200
agtcatgctg gcgcatcagc ggttgcagc agcctttaag tatggagttg atgcaaatag 7260

tcagcttgtt ggcatatggc gccatttcct gatagccggg atcctgataa tgcagaatgt 7320

cggctttcgc gcctttccct tcggatcatca ttcatgcag gccgcctatc agggatacgc 7380
 ggtgcgcgac ggaaacgccg cgcgtggact gcagcatcag tggacgcagg aggcctgtcg 7440
 atttaccga ccccatcccg gcgcggacaa taacgatgcc ctgcagctgt gcggcgtatg 7500
 tcatcacctc atcggatcac ctggaggttt caaacgttt gtaagtgatg tgtgacgggc 7560
 gaaggttcgg gttggtgatg cgttcactga acgaacgtga tgtttgcgcg gcacggcatt 7620
 tgcgattcaa ccggcgcgta atgtgatctt taacgtacc gttataaatt tctgcgatac 7680
 ccataccccg cagcgtgctg ctgaaaaggc gcataagttc tttcgggctg tttggtaccg 7740
 ggcatgtcag catgccaata tcaacggcgc gaagcagttc tttggcaaaa gtgcgtctgt 7800

tcagacgcgg gagagtacgc agcttattca gcgtgatcga caacagatcg gttgcacggc 7860
 tcagatgatt tctcgttaac tggcgcgaga ctctcttcag ccctctcagg ctgtgcaggt 7920
 cgttaaaatc gctgcattcc agctcagggt catcctcaaa agttgggtaa acacatttga 7980
 cgccggaaaa ctctccatg atgtcgaatc cgtgctggag gcctgtgttg ccttttcctt 8040
 cagctgagga tttgcggtcg ttatcgagag cgcaagtgat ttgcgcagcc gggtagatgt 8100
 tcaccagctg ctgcacaac tgaatcatgt tgttagcgga aaccgcaatg actaccgctg 8160
 caaagcgttt tttcgggtcg tttctggtcg ccagccagat ggatgccccg gtggcgaaac 8220
 cctctgcagt cgcaattttt tgcgccccct gcaggtcgcc aataacaaag catgcaccga 8280

cgaaatcacc gttagtatg gcgctggtct ggaacttgcc accattcaga tcgatacgtt 8340
 gccagccaac aatccgcccc tcttttcttc cgtccagggt ggacagaggt atcgccatgt 8400
 aagttgttgg tccacggctc catttcgcac tgtcgtgact ggtcacgcga cgtatatcac 8460
 aagcgccaaa tacgtcacga attccctttt ttaccgcata aggccaggag ccattctcag 8520
 ctggcgaatg ttcccaggcg cgatggaaag ccaaccatcc aagcaggcgt tctgtctcca 8580
 tctgattgtt ttttaaatca ttaacgcgtt gttgttcagc tcggaggcgg cgtgcttcag 8640
 cctggcgctc catgcgtgca cgttcttctt ccggctgagc gaccacggtc gcaccattcc 8700
 gttgctgttc acggcgatac tccgaaaaca ggaatgaaaa gccactccag gagccagcgt 8760

catgcgcttt ttcaacgaag ttaacgaaag gataactgat gccatccttg ctctgtctca 8820
 ggctgaata gatttcaca cggcctttta ggctcttctg cagagcttcc ggggaggaat 8880
 tattgtaggt ggiatagcgc tctacaccac cgcgcggatt gagctgaatc ttatcagcac 8940
 acgcaggcca gttgataccg gccatcttcg ccagctcagt cagctcatca cgtgccgcgt 9000
 caagcagtga aaacggatcg ctgccaaagc gctccgcgta gaattcttgt aaggctattt 9060
 tttagccttt ccatgcgaat tagcattttt tcgggttgaa aaaatccgca ggagcagcca 9120

caataaacgc aciatctttc tgaaggacgt atctgcgtta tcgtggctac ttcctgaaaa 9180
aggcccagat ttgccgactc ggtttttttt tcgtcttttt tcggctgcta cggctctggtt 9240

caaccccgac aaagtataga tcggattaaa ccagaattat agtcagcaat aaacctgtt 9300
attgtatcat ctacctcaa ccatgaacga ttgatcgta ccgactactt ggtgcacaaa 9360
ttgaagatca cttttatcat ggataacccg ttgagagtta gcactatcaa ggtagtaatg 9420
ctgctcgta taacgggcta atcgttgaat tggatctcg ccgttattat cacaaaccag 9480
tacatcctca cccggtacaa gcgtaagtga agaatcgacc aggataacat ctcccggctg 9540
gtagtctgc tgaatctggt tcccagcgt cagtgcgtaa acggtgttcc gttgactcac 9600
gaacggcagg aatcgctctg tgttggcagg ttctccaggc tgccagtctc tatccggtcc 9660
tgtctctgtc gtaccaataa caggaacgcg gtctggatca gattcagtgc catacagtat 9720

ccattgcacg ggcttacgca ggcatcttgc cagcgatagc ccgatctcca gcgacggcat 9780
cacgtcgcca cgttctaagt ttggacgcc cggaagagag attcctacag cttctgccac 9840
ttgcttcagc gtcagtttca gctctaaacg gcgtgctttc agtcgttcgc ctctgtttt 9900
cataccctta atcataaatg atctctttat agctggetat aatttttata aattatacct 9960
agctttaatt ttacttatt gattataata atcccatga aaccgaaga acttgtgcgc 10020
catttcggcg atgtggaaaa agcagcgggt ggcggtggcg tgacacccgg cgcagtctat 10080
caatggctgc aagctgggga gattccacct ctacgacaaa gcgatataga ggtccgtacc 10140

gcgtacaaat taaagagiga tttcacctct cagcgcatgg gtaaggaagg gcataacaag 10200
gggatcctct agacgcagaa aggccaccc gaaggtgagc cagtgtgatt acatttgagg 10260
cctaactgtg gccagtccag ttacgtgga gtcactagtgc cggccgcgac aacttgtcta 10320
gggccaatg gccatacac ttagtgtaat acgactcact ataggagag cggccgcttt 10380
ttcagcaaga ttaagccgcc accatggcgc cgcggcctcc taagaagaag cggaaagtcg 10440
aattcgtgga tctgcgaaca ctgggctata gccagcagca gcaggagaag atcaaacca 10500
aggtgaggtc cacagtcgca cagcaccatg aagccctggt gggccacggg ttactcacg 10560

ctcatattgt cgcactgtct cagcatccag ccgctctggg aacctggca gtcacatacc 10620
agcacatcat tactgcctg cccaggcta ccatgaaga catcgtggga gtcggcaaac 10680
agtggagcgg cgcacgggcc ctggaggctc tctgaccga cgcaggggaa ctgagaggac 10740
cccctctgca gctggatata gggcagctgg tgaagattgc taagagggga ggggtgacag 10800
caatggaagc cgtccacgca agcaggaacg cactgacagg ggccccctg aacctgacce 10860
cggaccaagt ggtggctatc gccagcaacg gtggcgga gcaagcgtc gaaacggtgc 10920

agcggctgtt gccggtgctg tgccaggacc atggcctgac cccggaccaa gtggtggcta 10980

tcgccagcaa ccacggcggc aagcaagcgc tcgaaacggt gcagcggctg ttgccggtgc 11040

tgtgccagga ccatggcctg actccggacc aagtgggtgc tatgccagc cacgatggcg 11100

gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag gaccatggcc 11160

tgactccgga ccaagtgggt gctatcgcca gccacgatgg cggcaagcaa gcgctcgaaa 11220

cgggtgcagcg gctgttgccg gtgctgtgcc aggacatgg cctgactccg gaccaagtgg 11280

tggctatcgc cagccacgat ggcggaagc aagcgctcga aacggtgcag cggctgttgc 11340

cgggtgctgtg ccaggacat ggctgactc cggaccaagt ggtggctatc gccagccacg 11400

atggtggaaa acaggccctt gaaacggtgc agcggctgtt gccggtgctg tgccaggacc 11460

atggcctgac cccggaccaa gtggtggcta tcgccagcaa cattggcggc aagcaagcgc 11520

tcgaaacggt gcagcggctg ttgccggtgc tgtgccagga ccatggcctg actccggacc 11580

aagtgggtgc tatgccagc cacgatggcg gcaagcaagc gctcgaaacg gtgcagcggc 11640

tgttgccggt gctgtgccag gaccatggcc tgaccccgga ccaagtgggt gctatcgcca 11700

gcaacattgg cggcaagcaa gcgctcgaaa cgggtgcagcg gctgttgccg gtgctgtgcc 11760

aggacatgg cctgaccccg gaccaagtgg ttgctatcgc cagcaaccac ggcggaagc 11820

aagcgctcga aacggtgcag cggctgttgc cgggtgctgt ccaggacat ggctgaccc 11880

cggaccaagt ggtggctatc gccagcaacc acggcgga gcaagcgctc gaaacggtgc 11940

agcggctgtt gccggtgctg tgccaggacc atggcctgac cccggaccaa gtggtggcta 12000

tcgccagcaa ccacggcggc aagcaagcgc tcgaaacggt gcagcggctg ttgccggtgc 12060

tgtgccagga ccatggcctg actccggacc aagtgggtgc tatgccagc cacgatggcg 12120

gcaagcaagc gctcgaaacg gtgcagcggc tgttgccggt gctgtgccag gaccatggcc 12180

tgaccccgga ccaagtgggt gctatcgcca gcaacattgg cggcaagcaa gcgctcgaaa 12240

cgggtgcagcg gctgttgccg gtgctgtgcc aggacatgg cctgaccccg gaccaagtgg 12300

tggctatcgc cagcaaccac ggcggaagc aagcgctcga aacggtgcag cggctgttgc 12360

cgggtgctgtg ccaggacat ggctgaccc cggaccaagt ggtggctatc gccagcaacg 12420

gtggcgga gcaagcgctc gaaacggtgc agcggctgtt gccggtgctg tgccaggacc 12480

atggcctgac cccggaccaa gtggtggcta tcgccagcaa cattggcggc aagcaagcgc 12540

tcgaaagcat tggggccag ctgagccggc ctgatccggc gttggcccg ttgaccaacg 12600

accacctgtt cgctctggt tgcctgggag gacgcctgc tatggacgt gtgaagaaag 12660

gactgcccc cgaccccgaa ctgattagac gggatgaaccg gagaatcggc gagagaacat 12720
 cccatagggt ggcaatctct agaactcagc tggatcaagag tgaactggag gaaaagaaat 12780
 cagagctgcg ccacaagctg aaatacgtgc ctcatgagta tatcgaactg atcgagattg 12840
 ctgcaattc aaccaggagc cggatcctgg aaatgaaagt gatggagtgc tttatgaaag 12900
 tctacggata tcgggggaaa caccctggag ggagcagaaa gccagatggg gccatctaca 12960
 cagtgggagc ccccatcgac tatggcgtga ttgtcgatac taaagcctac agcggaggct 13020
 ataacctgcc tatcgccag gctgacgaga tgcagagata cgtggaggaa aaccagaccc 13080

gcaataagca tattaacccc aatgaatggt ggaaagtgt tctagctcc gtcacagagt 13140
 tcaagtttct gtctgtgagc ggacacttta agggcaacta caaagcacag ctgactaggc 13200
 tgaatcatat caccaactgc aatggagccg tgctgtctgt cgaggaactg ctgatcgggg 13260
 gagagatgat taaggctggc acactgactc tggaggaagt gaggcgcaag ttcaacaatg 13320
 gggaaatcaa ctctaacct gcagatgat aagctagccc cgggcgtacg gaaaaaaaa 13380
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13440
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13500

aaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13560
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 13620
 aaaaaaaaa aaaaaacgag accttagggc cattagactt gaagtcaagc ggccgcttac 13680
 aactggacct tgctggtaca tagaactgat taactgacca tttaatcat accaactagg 13740
 tcaataaaaa cgaaaggctc agtcgaaaga ctgggccttt cgttttaac tgatcggcac 13800
 gtaagagggt ccaactttca ccataatgaa ataagatcac taccgggctg attttgagtt 13860
 atcgagattt tcaggagcta aggaagctaa aatgagccat attcaacggg aaacgtcttg 13920

ctcaggccg cgattaaatt ccaacatgga tctgattta tatgggtata aatgggctcg 13980
 cgataatgtc gggcaatcag gtgcgacaat ctatcgattg tatgggaagc ccgatgcgc 14040
 agagttgttt ctgaaacatg gcaaaggtag cgttgccaat gatgttacag atgagatggt 14100
 caggctaaac tggtgacgg aatttatgcc tcttcgacc atcaagcatt ttatccgtac 14160
 tctgatgat gcatggttac tcaccactgc gatcccaggg aaacagcat tccaggtatt 14220
 agaagaatat cctgattcag gtgaaatat tgtgatgag ctggcagtgt tctgagccg 14280
 gtgcatcgc attcctgttt gtaattgtcc ttttaacggc gatcgctat ttcgtctggc 14340

tcaggcgcaa tcacgaatga ataacggttt ggttggtgag agtgattttg atgacgagcg 14400
 taatggctgg cctgttgaa aagtcaggaa agaatgcat aaactttgc cattctcacc 14460
 ggattcagtc gtcactcatg gtgatttctc acttgataac cttattttg acgaggggaa 14520

attaataggt tgiattgatg ttggacgagt cggaatcgca gaccgatacc aggatcttgc 14580
 catcctatgg aactgcctcg gtgagttttc tccttcatta cagaaacggc tttttcaaaa 14640
 atatggtatt gataatcctg atatgaataa attgcagttt cacttgatgc tcgatgagtt 14700
 tttctaacct aggtgacaga agtcaaaaagc ctccggtcgg aggcttttga ctttctgcta 14760

gatctgtttc aatgcggtga agggccaggc agctggggat tatgtccaga cccggccagc 14820
 atgttggttt tatcgcatat tcagcgttgt cgcgtttacc caggtaaaat ggaagcagtg 14880
 tatcgtctgc gtgaatgtgc aaatcaggaa cgtaaccgtg gtacatagat gcagtcctt 14940
 gcgggtcgtt ccttcaacg agtaggacgc ggtgcccttg caaggctaac cattgcgcct 15000
 ggtgtactgc agatgaggtt ttataaaccc ctcccttggtg tgacataacg gaaagtacaa 15060
 ccgggttttt atcgtcaggt ctttggtttg ggttaccaa cactccgc atatggctaa 15120
 tttggtcaat tgtgtagcca gcgcgacgtt ctactcgcc cctcatctca aaatcaggag 15180

ccggtagacg accagctttt tccgcatctc tgatagcctg cgggtgttacg ccgatcaggt 15240
 ctgcaacttc tgttataccc cagcggcgag taatacgacg cgcttcggg ctgtcatcgc 15300
 cgaactgtgc gatggcaata gcgcgctca tttctgacc gcgattgata cagtctttca 15360
 gcaaattaat taacgacatc ctgtttctc tcaaacaatgc ccttatcttt gtgtttttca 15420
 tcatacttta cgtttttaaa gcaaagcaac ataaaaaag caaagtgact tagaaaacgc 15480
 aaagttaagg ttcaaatcaa tttttgatg cgctacagaa gctatttagc ttcactaag 15540
 cgcaacggta ttacttacgt tggatatattt aaaacctaac ttaatgattt taaatgataa 15600

taaatcatac caattgctat caaaagttaa gcgaacatgc tgattttcac gctgtttata 15660
 cactttgagg catctctatc tcttctgtct ctatatgaa acacaatcaa agaacatcaa 15720
 tccatgtgac atccccact atctaagaac accataacag aacacaacat aggaatgcaa 15780
 cattaatgta tcaataatc ggaacatatg cactatatca tatctcaatt acggaacata 15840
 tcagcacaca attgccatt atacgc 15866

<210> 24

<211> 7365

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1242

<400> 24

cagctgcgcg ctgcctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60

tttggtcgcc cgccctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120

actaggggtt cctttagatt aatgattaac ccgcatgct acttatctac acgcgtagat 180

ctagtgcac aacttcttat ttgtgtaata agaaaattgg gaaaacgac ttcaatatgc 240

ttaccaagct gtgattccaa atattacgta aatacattg caaaggagga tgtttttagt 300

agcaatttgt actgatgga tggggccaag agatatact tagagggagg gctgagggtt 360

tgaagtcgaa ctctaagcc agtgccagaa gagccaagga caggtacggc tgtcatcact 420

tagacctcac cctgtggagc cacacctag ggttggccaa tctactcca ggagcagga 480

gggcaggagc cagggtctgg cataaaagtc agggcagagc catctattgc ttacatttgc 540

ttctgacaca actgtgttca ctagcaacct caaacagaca ccaggtgagt taaacccatg 600

agagagaata acagaatgc gagtgatggg ccagtttaagc gtagatggct aattagtcca 660

gacaaatgta aaatgccaac accgtctgta aagaaaccta actgacctc ttcctttgtc 720

ctgtcttctt cacaggccgc caccatggc catcttacac cggaggagaa gtccgtgta 780

acggcactgt gggggaaagt taatgtcgat gaagtcggcg gtgaagcact cggcagggtt 840

ctggtagtgt acccgtggac acaacgattc ttgaaagt tctgggacct gtccacacc 900

gatgctgtga tgggtaatcc aaaagtaaaa gcacacggca agaaagtcct cggcgcttt 960

agtgatggc tggccattt ggataactt aagggtacat tcgcgcagct ttccgaactc 1020

cactgtgaca agttgcact agatccagaa aacttccggc ttctgggcaa tgtgcttga 1080

tgcgttctgg ctccacattt tgggaaggag ttacccac ccgtgcaagc ggcttacaa 1140

aaagtggcag caggagtggc taatgccctt gcacataaat atactaagg taccgagcat 1200

cttaccgcca ttatttccca tttttgttct gtttttcttg atttgggtat acatttaaat 1260

gttaataaaa caaatggc gggcaatcat ttacattttt agggatatgt aattactagt 1320

tcagggtgat tgccacaaga caaacatgtt aagaaactt cccgttattt acgtctgtt 1380

cctgttaatc aacctctgga ttacaaaatt tgtgaaagat tgactgatat tcttaactat 1440

gttgctcctt ttacgtgtg tggatatgct gctttatagc ctctgtatct agctattgct 1500

tccegtacgg ctttcgtttt ctctccttg tataaatcct ggttgctgtc tcttttagag 1560

gagttgtggc ccgtgtccg tcaactggc gtggtgtgct ctgtgtttgc tgacgcaacc 1620

cccactggct ggggcattgc caccacctgt caactcctt ctgggacttt cgctttcccc 1680

ctcccgatcg ccacggcaga actcatgcc gcctgccttg cccgtgctg gacaggggct 1740

aggttctgg gcaactgataa ttccgtggc ttgtctgtgc cttctagtgc ccagccatct 1800

gttgtttgcc cttccccgt gccttccttg accctggaag gtgccactcc cactgtcctt 1860

tcctaataaa atgaggaaat tgcacgcat tgcctgagta ggtgtcattc tattctgggg 1920
 ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg 1980

gatgcggtgg gctctatggc ccgcgggaac agagaaacag gagaatatgg gccaaacagg 2040
 atatctgttg taagcagttc ctgccccggc tcagggccaa gaacagttgg aacagcagaa 2100
 tatgggccaa acaggatata tgtggttaagc agttcctgcc ccggctcagg gccagaaca 2160
 gatggcccc agatgcggtc ccgcctcag cagtttctag agaaccatca gatgtttcca 2220
 ggggtcccca aggacctgaa atgacctgt gccttatttg aactaacaa tcagttcgt 2280
 tctcgttct gttcgcgcgc ttctgtccc cgagctctat ataagcagag ctggtttagt 2340
 gaacctcag atcgctgga gacgcatcc acgctgttt gacttcata gaaggcgcc 2400
 gcgcgccac catggtgagc aaggcgagg agctgttcac cggggtggtg cccatcctgg 2460

tcgagctgga cggcgacgta aacggccaca agttcagct gtccggcgag ggcgaggcg 2520
 atgccaccta cggcaagctg accctgaagt tcactgcac caccggcaag ctgccgtgc 2580
 cctggccac cctcgtgacc accctgacct acggcgtgca gtgcttcagc cgtaccccg 2640
 accacatgaa gcagcacgac ttcttcaagt ccgcatgcc cgaaggctac gtccaggagc 2700
 gcaccatctt cttaaggac gacggcaact acaagaccg cgccgagtg aagttcagg 2760
 gcgacacct ggtgaaccgc atcgagctga agggcatcga cttcaaggag gacggcaaca 2820
 tcctggggca caagctggag tacaactaca acagccaca cgtctatac atggccgaca 2880
 agcagaagaa cggcatcaag gtgaacttca agatccgcca caacatcgag gacggcagcg 2940

tgcagctgc cgaccactac cagcagaaca ccccatcgg cgacggcccc gtgctgctgc 3000
 ccgacaacca ctacctgagc acccagtcg cctgagcaa agacccaac gagaagcgcg 3060
 atcacatggt cctgctggag ttctgaccg ccgccgggat cactctcgcc atggacgagc 3120
 tgtacaaggg aagcggagct actaacttca gcctgctgaa gcaggctgga gacgtggagg 3180
 agaaccttg acctacctgc aggcctgaga acttcagggt gactctatgg gacgcttgat 3240
 gttttcttc cctttcttt ctatggttaa gttcatgtca taggaagggg ataagtaaca 3300
 gggtagctt tagaatggga aacagacgaa tgattgcac agtgtggaag tctcaggatc 3360
 gttttagttt cttttatttg ctgttcataa caattgttt ctttgttta attcttgctt 3420

tctttttttt tcttctccgc aatttttact attatactta atgccttaac attgtgtata 3480
 acaaaaggaa atatctctga gatacttaa gtaacttaa aaaaaacttt acacagtctg 3540
 cctagtacat tactatttgg aatatatgtg tgcttatttg catattcata atctccctac 3600
 ttgtcgacg tagataagta gcatggcggg ttaatcatia actacaagga acccctagt 3660
 atggagtgg cactccctc tctgcgcgt cgctcgctca ctgaggccgg gcgaccaaag 3720

gtcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc gcgccagctg 3780
gcgtaatagc gaagaggccc gcaccgatcg cctttcccaa cagttgcga gcctgaatgg 3840
cgaatggcga ttccgttgca atggctggcg gtaatatgt tctggatatt accagcaagg 3900

ccgatagttt gagttcttct actcaggcaa gtgatgttat tactaatcaa agaagtattg 3960
cgacaacggt taatttgcgt gatggacaga ctcttttact cgggtggcctc actgattata 4020
aaaacacttc tcaggattct ggcgtagcgt tctgtctaa aatcccttta atcggcctcc 4080
tgtttagctc ccgctctgat tctaacgagg aaagcacgtt atacgtgctc gtcaaagcaa 4140
ccatagtacg cgccctgtag cggcgcatga agcgcggcgg gtgtggtggt tacgcgcagc 4200
gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt cctttccttt 4260
ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc 4320
cgatttagtg ctttacggca cctcgacccc aaaaaacttg attaggtga tggttcacgt 4380

agtgggcat cgccctgata gacggttttt cgcccttga cgttggagtc cacgttcttt 4440
aatagtgac tctgttcca aactggaaca aactcaacc ctatctcgtt ctattctttt 4500
gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa 4560
aaatttaacg cgaattttaa caaaatatta acgtttacaa tttaaatatt tgcttataca 4620
atcttcctgt ttttggggct tttctgatta tcaaccgggg tacatatgat tgacatgcta 4680
gttttacgat taccgttcat cgattctctt gtttgcctca gactctcagg caatgacctg 4740
atagcctttg tagagacctc tcaaaaatag ctaccctctc cggcatgaat ttatcagcta 4800
gaacggttga atatcatatt gatggtgatt tgactgtctc cggcctttct caccgtttg 4860

aatctttacc tacacattac tcaggcattg catttaaaat atatgagggt tctaaaaatt 4920
tttatccttg cgttgaaata aaggcttctc ccgcaaaagt attacagggt cataatgttt 4980
ttggtacaac cgatttagct ttatgctctg aggcctttatt gcttaatttt gctaattctt 5040
tgccctgcct gtatgattta ttggatgttg gaatcgctg atcggttatt ttctccttac 5100
gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc 5160
cgcatagtta agccagcccc gacacccgcc aacacccgct gacgcgcctt gacgggcttg 5220
tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca 5280
gaggttttca ccgtcatcac cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt 5340

tttatagggt aatgtcatga taataatggt ttcttagacg tcagggtggca cttttcgggg 5400
aaatgtgcgc ggaaccctta tttgtttatt ttctaaata cattcaaata tgtatccgct 5460
catgagacaa taacctgat aaatgcttca ataatatga aaaaggaaga gtatgagtat 5520
tcaacatttc cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc 5580

tcacccagaa acgctgggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg 5640
 ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg 5700
 ttttccaatg atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga 5760
 cgccgggcaa gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta 5820

ctcaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc 5880
 tgccataacc atgagtgata aactgcggc caacttactt ctgacaacga tcggaggacc 5940
 gaaggagcta accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg 6000
 ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc 6060
 aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca 6120
 acaattaata gactggatgg aggcgggataa agttgcagga ccacttctgc gctcggccct 6180
 tccggtcggc tggtttatgt ctgataaatc tggagccggt gagcgtgggt ctgcggtat 6240
 cattgcagca ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg 6300

gagtcaggca actatggatg aacgaaatag acagatcgct gagatagggt cctcactgat 6360
 taagcattgg taactgtcag accaagttaa ctcatatata ctttagattg atttaaaact 6420
 tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca tgacaaaat 6480
 cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc 6540
 ttcttgagat ctttttttc tgcgcgtaat ctgctgcttg caaacaacaa aaccaccgt 6600
 accagcgggt gtttgtttgc cggatcaaga gctaccaact cttttccga aggttaactgg 6660
 cttcagcaga gcgcagatac caaatactgt ctttctagt tagccgtagt taggccacca 6720
 cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc 6780

tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga 6840
 taaggcgag cggtcgggt gaacgggggg ttctgtcaca cagcccagct tggagcgaac 6900
 gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttccga 6960
 agggagaaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag 7020
 ggagcttcca gggggaacg cctgttatct ttatagtcct gtcgggttcc gccacctctg 7080
 acttgagcgt cgatttttgt gatgtctgtc aggggggagg agcctatgga aaaacgccag 7140
 caacgcggcc tttttacgt tcttgccctt ttgtggcct tttgtcaca tgttctttcc 7200
 tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc 7260

tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc 7320
 aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat taatg 7365

<210> 25

<211> 8087

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1243

<400> 25

cagctgcgcg ctgcctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 ttgtgtcgcg cgccctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
 actaggggtt cctgtagtat aatgattaac ccgccatgct acttatctac acgcgtatag 180
 ctagtgcacg aacttcttat ttgtgtaata agaaaattgg gaaaacgacg ttcaatatgc 240

ttaccaagct gtgattccaa atattacgta aatacacttg caaaggagga tgtttttagt 300
 agcaatttgt actgatggta tggggccaag agatatact tagaggagg gctgagggtt 360
 tgaagtccaa ctctaagcc agtgccagaa gagccaagga caggtagcgc tgcacacact 420
 tagacctcac cctgtggagc cacaccctag ggttggccaa tctactccca ggagcaggga 480
 gggcaggagc cagggtctgg cataaaagtc agggcagagc catctattgc ttacatttgc 540
 ttctgacaca actgtgttca ctagcaacct caaacagaca ccaggtgagt taaacccatg 600
 agagagaata acagaactgc gagtgatggg ccagtttaagc gtagatggct aattagtcca 660
 gacaaatgta aaatgccaac accgtctgta aagaaaccta actgacccg ttccctttgtc 720

ctgtcttctt cacaggccgc caccatggc catcttacac cggaggagaa gtccgctgta 780
 acggcactgt ggggaaagt taatgtcgat gaagtcggcg gtgaagcact cggcagggtt 840
 ctggtagtgt acccgtggac acaacgattc ttgaaagt ttggggacct gtccacaccc 900
 gatgctgtga tgggtaatcc aaaagtaaaa gcacacggca agaaagtcct cggcgcgttt 960
 agtgatggtc tggccattt ggataacttg aagggtacat tcgcgcagct ttccgaactc 1020
 cactgtgaca agttgcacgt agatccagaa aacttccggc ttctgggcaa tgtgcttgta 1080
 tgcgttcttg ctaccattt tgggaaggag ttacccac ccgtgcaagc ggcttaccaa 1140
 aaagtggcgc caggagtggc taatgccctt gcacataaat atactaagg taccgagcat 1200

cttaccgcca ttattccca tatttgttct gtttttcttg atttgggtat acatttaaat 1260
 gtaataaaa caaatgggtg gggcaatcat ttacattttt agggatatgt aattactagt 1320
 tcagggtgat tgccacaaga caaacatgtt aagaaacttt cccgttatct acgtctgttt 1380
 cctgttaatc aacctctgga ttacaaaatt tgtgaaagat tgactgatat tcttaactat 1440
 gttgctcctt ttacgctgtg tggatatgct gctttatagc ctctgtatct agctattgct 1500

tcccgtacgg ctttcgtttt ctcttccttg tataaatcct ggttgctgtc tcttttagag 1560
gagttgtggc ccgttgtccg tcaacgtggc gtggtgtgct ctgtgtttgc tgacgcaacc 1620
cccactggct ggggcattgc caccacctgt caactccttt ctgggacttt cgctttcccc 1680

ctcccgatcg ccacggcaga actcatcgcc gcctgccttg cccgctgctg gacaggggct 1740
aggttgttgg gactgataa ttccgtggtg ttgtctgtgc cttctagtgt ccagccatct 1800
gttgtttgcc cctccccctg gccttccttg accctggaag gtgccactcc cactgtcctt 1860
tcctaataaa atgaggaaat tgcacgcat tgtctgagta ggtgtcattc tattctgggg 1920
ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg 1980
gatgcggtgg gctctatggc ttaattaacg agatcgagac catcctggct aacacagtga 2040
aaccctgtct ctactaaaaa aatacaaaaa attagccggg cttggtggcg ggtgcctgta 2100
gtcccagcta ctatggagcg tgaggcggga gaatggcgtg aacgcggggg gcggagcttg 2160

cagtgagcag agatcagggg ccactgcact ccagcctggg cgacagagag agactctgtc 2220
tcaaaaaaaaa gaaaaaaaaa atttagtaga ctactaaaaa aaatccagag atagttattg 2280
atgcatatgt aaaagtcttc caatatctac aagtacaatg aaaaaaaaaa aaccttgaat 2340
taagttaga actcattgac aatgtttcaa aggatgtgag ggataaacta aaatttgggc 2400
agtacatgct gttcctgtgt acttggaca gagggagaaa atctgggctg gaaatattgt 2460
tataggagtt agcacatgaa ggtgacaact aaattatttg gagtagatgg agtcaccagc 2520
acatgtgaat agttttagaa tgaatgacc caagatagaa ctttgagag ccccaaatt 2580
taaataaaat cagtataaga gaagaggaag aaaccaaag gtatactagt ctaaattgtt 2640

tcttagtgac aaaagaataa cctgaatatt agattagctg cctatatgct ctctgaatca 2700
atttcattca acatgcaaca gtccgcggga acagagaaac aggagaatat gggccaaaca 2760
ggatactgt ggtgaagcagt tcctgccccg gtcagggcc aagaacagtt ggaacagcag 2820
aatatgggcc aaacaggata tctgtggtta gcagttcctg ccccggtca gggccaagaa 2880
cagatggtec ccagatgcgg tcccgcctc agcagtttct agagaacat cagatgtttc 2940
cagggtgccc caaggacctg aaatgacct gtgccttatt tgaactaacc aatcagttcg 3000
cttctcgctt ctgttcgcgc gcttctgtc cccgagctct atataagcag agctcgttta 3060
gtgaaccgtc agatgcctg gagacgcat ccacgtgtt ttgacttcca tagaaggcgg 3120

ccgcgccgcc accatggtga gcaagggcga ggagctgttc accggggtgg tgcccatcct 3180
ggtcgagctg gacggcgacg taaacggcca caagttcagc gtgtccggcg agggcgaggg 3240
cgatgccacc tacggcaage tgacctgaa gttcatctgc accaccgga agctgccccg 3300
gccctggccc accctcgtga ccacctgac ctacggcgtg cagtgttca gccgtaccc 3360

cgaccacatg aagcagcacg actttcttcaa gtccgccatg cccgaaggct acgtccagga 3420
 gcgcaccatc ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg tgaagttcga 3480
 gggcgacacc ctggtgaacc gcatcgagct gaagggcatc gacttcaagg aggacggcaa 3540
 catcctgggg cacaagctgg agtacaacta caacagccac aacgtctata tcatggccga 3600

caagcagaag aacggcatca aggtgaactt caagatccgc cacaacatcg aggacggcag 3660
 cgtgcagctc gccgaccact accagcagaa caccacctc ggcgacggcc ccgtgctgct 3720
 gcccacaac cactacctga gcaccagtc cgccctgagc aaagaccca acgagaagcg 3780
 cgatcacatg gtctgtctgg agttcgtgac cgcgccggg atcactctcg gcatggacga 3840
 gctgtacaag ggaagcggag ctactaactt cagcctgctg aagcaggctg gagacgtgga 3900
 ggagaacctt ggacctacct gcaggcctga gaacttcagg gtgagtctat gggacgcttg 3960
 atgttttctt tcccttctt ttctatggtt aagttcatgt cataggaagg ggataagtaa 4020
 cagggtacag ttiagaatgg gaaacagacg aatgattgca tcagtgtgga agtctcagga 4080

tcgttttagt ttcttttatt tgctgttcat aacaattggt ttcttttgtt taattcttgc 4140
 ttctttttt ttctttctcc gcaattttta ctattatact taatgcctta acattgtgta 4200
 taacaaaagg aaatatctct gagatacatt aagtaactta aaaaaaact ttacacagtc 4260
 tgccatgtac attactattt ggaatatatg tgtgcttatt tgcatttca taatctccct 4320
 actttgtcga cgtagataag tagcatggcg ggtaatcat taactacaag gaacccttag 4380
 tgatggagtt ggccactccc tctctgcgcg ctgctcgtc cactgaggcc gggcgaccaa 4440
 aggtcgcccg acgcccgggc ttgcccggg cgccctcagt gagcgagcga gcgcgccagc 4500
 tggcgtaata gcgaagaggg ccgcaccgat cgcccttccc aacagttagc cagcctgaat 4560

ggcgaatggc gattccgttg caatggctgg cgtaaatatt gttctggata ttaccagcaa 4620
 ggccgatagt ttgagttctt ctactcaggc aagtgatgtt attactaatc aaagaagtat 4680
 tgcgacaacg gtttaatttg gtgatggaca gactcttcta ctcggtggcc tcaatgatta 4740
 taaaaaacct tctcaggatt ctggcgctacc gttctgtct aaaatccctt taatcggcct 4800
 cctgttttag tcccgtcttg attctaacga ggaaagcacg ttatacgtgc tcgtcaaagc 4860
 aaccatagta cgcgccctgt agcgccgcat taagcgcggc ggggtgtgtg gttacgcga 4920
 gcgtgaccgc tacacttgcc agcgccctag cgccgctcc ttctgcttc ttccttct 4980
 ttctcgccac gttcgccggc ttccccgtc aagctctaaa tcgggggctc ctttagggt 5040

tccgatttag tgctttacgg cactcgacc caaaaaact tgattagggt gatggttcac 5100
 gtagtgggc atcgccctga tagacggtt ttgccttt gacgttgag tccacgttct 5160
 ttaatagtgg actctgttc caaactggaa caacactcaa ccctatctcg gtctattctt 5220

ttgatttata agggattttg ccgatttcgg cctattgggt aaaaaatgag ctgatttaac 5280
 aaaaatttaa cggaattttt aacaaaatat taacgtttac aatttaaata ttgtcttata 5340
 caatcttctt gtttttgggg cttttctgat tatcaaccgg ggtacatatg attgacatgc 5400
 tagttttacg attaccgttc atcgatttctc ttgtttgtctc cagactctca ggcaatgacc 5460
 tgatagcctt ttagagagacc tctcaaaaat agctaccctc tccggcatga atttatcagc 5520

tagaacgggt gaatatcata ttgatgggtga ttgactgtc tccggccttt ctcaccggtt 5580
 tgaatcttta cctacacatt actcaggcat tgcatttaaa atatatgagg gttctaaaaa 5640
 tttttatcct tgcgttgaaa taaaggcttc tccgcacaaa gtattacagg gtcataatgt 5700
 ttttggtaca accgatttag ctttatgtctc tgaggcttta ttgcttaatt ttgctaattc 5760
 ttgaccttgc ctgtatgatt tattggatgt tggaaatgcc tgatgcggta ttttctcctt 5820
 acgcatctgt gcggtatttc acaccgcata tgggtgcactc tcagtacaat ctgctctgat 5880
 gccgcatagt taagccagcc ccgacaccgc ccaacaccgc ctgacgcgcc ctgacgggct 5940
 tgtctgtctc cgcatccgc ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt 6000

cagaggtttt caccgtcatc accgaaacgc gcgagacgaa agggcctcgt gatacccta 6060
 tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcagggtg cacttttcgg 6120
 ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg 6180
 ctcatgagac aataacctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt 6240
 attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cttttgcct tctgttttt 6300
 gctcaccag aaacgttgtt gaaagtaaaa gatgctgaag atcagttggg tgcacgagt 6360
 ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg ccccgaagaa 6420
 cgttttccaa tgatgagcac ttttaaagt ctgctatgtg gcgcggtatt atcccgatt 6480

gacgcggggc aagagcaact cggctgcgcg atacactatt ctcagaatga cttggttgag 6540
 tactcaccag tcacagaaaa gcactttacg gatggcatga cagtaagaga attatgcagt 6600
 gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac gatcggagga 6660
 ccgaaggagc taaccgcttt ttgacacaac atgggggagc atgtaactcg cttgatcgt 6720
 tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta 6780
 gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg 6840
 caacaattaa tagactggat ggaggcggat aaagttgcag gaccatttct gcgctcggcc 6900
 cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt 6960

atcattgcag cactggggcc agatggtaag cctccccgta tcgtagttat ctacacgagc 7020
 gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg 7080

attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat tgatttataaa 7140
cttcattttt aatttataag gatctaggtg aagatccttt ttgataatct catgaccaa 7200
atccctaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga 7260
tcttcttgag atcctttttt tctgcgcgtc atctgctgct tgcaaataaa aaaaccaccg 7320
ctaccagcgg tggttttgtt gccgcatcaa gagctaccaa ctctttttcc gaaggtaact 7380
ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta gttaggccac 7440

cacttcaaga actctgtagc accgcctaca tacctcgtc tgctaatect gttaccagt 7500
gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg 7560
gataaggcgc agcggctcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga 7620
acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc 7680
gaaggagaga aggcggacag gtatccgta agcggcaggg tcggaacagg agagcgcacg 7740
agggagcttc cagggggaaa cgcctgggtat ctttatagtc ctgctgggtt tcgccacctc 7800
tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc 7860
agcaacgcgg cttttttacg gttctctggc ttttctggc cttttgtca catgttcttt 7920

cttgcgttat cccctgattc tgttgataac cgtattaccg cttttgagt agctgatacc 7980
gctcgcgcga gccgaacgac cgagcgcagc gattcagtga gcgaggaagc ggaagagcgc 8040
ccaatagca aaccgcctct cccgcgcgt tggccgattc attaatg 8087

<210> 26

<211> 7712

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1244

<400> 26

cagctgcgcg ctgctcgtc cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
tttggtcgcc cggcctcagt gagcgcgca gcgcgcagag agggagtggc caactccatc 120
actaggggtt cttgtagtt aatgattaac ccgcatgct acttatctac acgcgtagat 180

ctagtgcac aacttcttat ttgtgtaata agaaaattgg gaaaacgac ttcaatatgc 240
ttaccaagct gtgattccaa atattacgta aatacattg caaaggagga tgtttttagt 300
agcaatttgt actgatggtg tggggccaag agatatact tagagggagg gctgagggtt 360
tgaagtcaa ctctaagcc agtgccagaa gagccaagga caggtacggc tgatcatcact 420

tagacctcac ccgtgggagc cacaccctag ggctggccaa tctactccca ggagcaggga 480
gggcaggagc cagggcctggg cataaaagtc agggcagagc catctattgc ttacatttgc 540
ttctgacaca actgtgttca ctagcaacct caaacagaca ccaggtgagt taaacccatg 600
agagagaata acagaactgc gagtgatggg ccagttaagc gtagatggct aattagtcca 660

gacaaatgta aaatgccaac accgtctgta aagaaaccta actgatectc ttcctttgtc 720
ctgtcttctt cacaggccgc caccatggtc catcttacac cggaggagaa gtccgctgta 780
acggcactgt gggggaaagt taatgtcgat gaagtcggcg gtgaagcact cggcaggttg 840
ctggtagtgt acccgtggac acaacgattc ttgaaagt tctgggacac gtccacacc 900
gatgctgtga tgggtaatcc aaaagtaaaa gcacacggca agaaagtcct cggcgcgttt 960
agtgatggtc tggccattt ggataacttg aagggtacat tgcgcagct ttccgaactc 1020
cactgtgaca agttgcacgt agatccagaa aacttccggc ttctgggcaa tgtgcttgta 1080
tgcgttctgg ctccattt tgggaaggag ttacccac ccgtgcaagc ggcttaccaa 1140

aaagtggctg caggagtggc taatgccctt gcacataaat atcactaagg taccgataat 1200
caacctctgg attacaaaa ttgtgaaaga ttgactggta ttcttaacta tgttgcct 1260
tttacctat gtggatacg tctttaatg cctttgtatc atgctattgc ttcccgatg 1320
gctttcattt tctcctcct gtataaatcc tggtagttc ttgccacggc ggaactcacc 1380
gccgcctgcc ttgcccgtg ctggacaggg gctcggtgt tgggcactga caattccgtg 1440
gtgtttattt gtgaaattg tgatgctatt gctttattg taaccattct agctttattt 1500
gtgaaattg tgatgctatt gctttattg taaccattat aagctgcaat aaacaagtta 1560
acaacaaca ttgcattcat ttatgtttc aggttcaggg ggagatgtgg gaggtttttt 1620

aaagcttaat taacgagatc gagaccatcc tgctaacac agtgaaaccc cgtctctact 1680
aaaaaaatc aaaaaattag cgggcttgg tggcgggtgc ctgtagtccc agctactatg 1740
gaggctgagg cgggagaatg gcgtgaacgc gggggcgga gcttgcagt agcagagatc 1800
aggggccact gcactccagc ctgggcgaca gagagagact ctgtctcaa aaaaagaaaa 1860
aaaaaattha gtagactagc taaaaaatc cagagatagt tattgatgca tatgtaaaag 1920
tcttccaata ttacaagta caatgaaaa aaaataacct tgaattaagt gtagaactca 1980
ttgacaatgt ttcaaggat gtgagggata aactaaaatt tgggcagtac atgctgttcc 2040
tgtgtacttg gaacagagg agaaaatctg ggctggaaat attgttatag gagtttagcac 2100

atgaagtgta caactaaatt atttgagta gatggagtca ccagcacatg tgaatagttt 2160
tagaatgaaa tgaccaaga tagaactttg gagagccccc aaatttaaat aaaatcagta 2220
taagagaaga ggaagaaacc aaatggtata ctagtctaaa ttgtttctta gtgacaaaag 2280

aataacctga atattagatt agctgcctat atgctctctg aatcaatttc attcaacatg 2340
 caacagtcgg cggaacaga gaaacaggag aatatgggcc aaacaggata tctgtggtaa 2400
 gcagtctctg ccccggtcga gggccaagaa cagttggaac agcagaatat gggccaaaca 2460
 ggatactctg ggtaagcagt tctgccccg gctcagggcc aagaacagat ggtccccaga 2520
 tgcggccccg cctcagcag tttctagaga accatcagat gtttccaggg tgccccaagg 2580

acctgaaatg accctgtgcc ttatttgaac taaccaatca gttcgcttct cgcttctgtt 2640
 cgcgcgcttc tgctccccga gctctatata agcagagctc gtttagtgaa ccgtcagatc 2700
 gcctggagac gccatccacg ctgttttgac ttccatagaa ggcgcccgcg ccgccacat 2760
 ggtgagcaag ggcgaggagc tggttaccgg ggtgggtccc atcctggtcg agctggacgg 2820
 cgacgtaaac ggccacaagt tcagcgtgtc cgcgaggggc gagggcgatg ccacctacgg 2880
 caagctgacc ctgaagtcca tctgcaccac cggcaagctg cccgtgccct ggcccacct 2940
 cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccgccacc acatgaagca 3000
 gcacgacttc ttcaagtcgg ccatgcccga aggctacgtc caggagcgca ccattctctt 3060

caaggacgac ggcaactaca agaccgcgc cgaggtgaag ttcgaggcg acacctggt 3120
 gaaccgcatc gagctgaagg gcatcgactt caaggaggac ggcaacatcc tggggcacia 3180
 gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc agaagaacgg 3240
 catcaagggt aacttcaaga tccgccacia catcgaggac ggcagcgtgc agctcgccga 3300
 ccactaccag cagaaccccc ccatcggcga cggccccgtg ctgctgccg acaaccacta 3360
 cctgagcacc cagtcgcccc tgagcaaaga cccaacgag aagcgcatc acatggtcct 3420
 gctggagttc gtgaccgccg ccgggatcac tctcgcatg gacgagctgt acaagggaag 3480
 cggagctact aacttcagcc tgctgaagca ggctggagac gtggaggaga accctggacc 3540

tacctgcagg cctgagaact tcagggtgag tctatgggac gcttgatgtt ttctttcccc 3600
 ttcttttcta tggtaagt ctatgcatag gaaggggata agtaacaggg tacagtttag 3660
 aatgggaaac agacgaatga ttgcatcagt gtggaagtct caggatcgtt ttagtttctt 3720
 ttatttgctg ttcataaaca ttgttttctt ttgtttaatt cttgctttct tttttttct 3780
 tctccgaat ttttactatt atactaatg ccttaacatt gtgtataaca aaaggaaata 3840
 tctctgagat acattaagta acttaaaaaa aaactttaca cagtctgcct agtacattac 3900
 tatttggat atagtgtgc ttatttgcac attcataatc tccctacttt gtcgacgtag 3960
 ataagtagca tggcgggtta atcataact acaagggaacc cctagtgtg gagttggcca 4020

ctccctctct gcgcgctcgc tcgctcactg aggccgggag accaaaggtc gcccgcgc 4080
 cgggctttgc ccggcgggcc tcagtgcagc agcgagcgcg ccagctggcg taatagcgaa 4140

gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc tgaatggcga atggcgattc 4200
 cgttgcaatg gctggcggtg atattgttct ggatattacc agcaaggccg atagtttgag 4260
 ttcttctact caggcaagtg atgttattac taatcaaaga agtattgcga caacggttaa 4320
 ttitgctgat ggacagactc ttttactcgg tggcctcact gattataaaa acacttctca 4380
 ggattctggc gtaccgttcc tgtctaaaat ccccttaate ggcctcctgt ttagctcccg 4440
 ctctgattct aacgaggaaa gcacgttata cgtgctcgtc aaagcaacca tagtacgcgc 4500

cctgtagcgg cgcatthaagc gcggcggttg tggtaggtac gcgcagcgtg accgctacac 4560
 ttgccagcgc ctagcgccc gctccttctg ctttcttccc ttcctttctc gccacgttcg 4620
 ccggctttcc ccgtcaagct ctaaactcggg ggctcccttt agggttccga tttagtgtt 4680
 tacggcacct cgaccccaaa aaacttgatt agggatgatg ttcacgtagt ggcccatcgc 4740
 cctgatagac ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct 4800
 tgttccaaac tggaaacaaca ctcaacccta tctcggctca tttctttgat ttataaggga 4860
 ttttgccgat ttccgcctat tggttaaaaa atgagctgat ttaacaaaaa ttaacgcga 4920
 attttaaca aatattaacg tttaacaattt aaatatttgc ttataacaac ttcctgtttt 4980

tggggctttt ctgattatca accggggtac atatgattga catgctagtt ttacgattac 5040
 cgttcacga ttctcttggt tgcctcagac tctcaggcaa tgacctgata gcctttgtag 5100
 agacctctca aaaaatagcta ccctctccgg catgaattta tcagctagaa cggttgaata 5160
 tcatattgat ggtgatttga ctgtctccgg cctttctcac ccgtttgaat cttacctac 5220
 acattactca ggcattgcat ttaaaatata tgagggttct aaaaattttt atccttgcgt 5280
 tgaaataaag gcttctcccg caaaagtatt acagggtcat aatgtttttg gtacaaccga 5340
 tttagcttta tgctctgagg ctttattgct taattttgct aattctttgc ctgacctgta 5400
 tgatttattg gatgttgga tgcctgatg cggatatttc tccttacgca tctgtcgggt 5460

atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc 5520
 cagccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct gctcccgga 5580
 tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag gttttcacgc 5640
 tcatcaccga aacgcgcgag acgaaagggc ctctgtatgc gcctattttt atagggtaat 5700
 gtcagtataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga 5760
 accctatatt gttattttt ctaaatacat tcaaatatgt atccgctcat gagacaataa 5820
 ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca acatttccgt 5880
 gtgcacctta ttcccttttt tgcggcattt tgccttcttg tttttgtca cccagaaacg 5940

ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta catcgaactg 6000

gatctcaaca gcggttaagat ccttgagagt ttctgccccg aagaacgttt tccaatgatg 6060
agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc cgggcaagag 6120
caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc accagtcaca 6180
gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgtctgc cataaccatg 6240
agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa ggagctaacc 6300
gcttttttgc acaacatggg ggatcatgta actcgcttgc atcgttggga accggagctg 6360
aatgaagcca taccaaagca cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg 6420

ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca attaatagac 6480
tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc ggctggctgg 6540
tttattgctg ataaatctgg agccggtagc cgtgggtctc gcggtatcat tgcagcactg 6600
gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag tcaggcaact 6660
atggatgaac gaaatagaca gatcgtgagc ataggtgcct cactgattaa gcattggtaa 6720
ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca tttttaattt 6780
aaaaggatct aggtgaagat cttttttgat aatctcatga ccaaaatccc ttaacgtgag 6840
ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc ttgagatcct 6900

ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc agcgggtggtt 6960
tgtttgccgg atcaagagct accaactctt ttccgaagg taactggctt cagcagagcg 7020
cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt caagaactct 7080
gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc tgccagtggc 7140
gataagtctg gtcttaccgg gttggactca agacgatagt taccggataa ggcgagcgg 7200
tcgggctgaa cgggggggtc gtgcacacag ccagcttgg agcgaacgac ctacaccgaa 7260
ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg 7320
gacaggtatc cggtaaagcg cagggtcggg acaggagagc gcacgaggga gcttccaggg 7380

ggaaacgcct ggtatcttta tagtctgtc gggtttcgcc acctctgact tgagcgtcga 7440
tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa gcggccttt 7500
ttacggttcc tggccttttg ctggcctttt gctcacatgt tcttctctgc gttatccct 7560
gattctgtgg ataaccgtat taccgctttt gagtgaactg ataccgctcg ccgagccga 7620
acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgcccaat acgcaaaccg 7680
cctctccccg cgcgttggcc gattcattaa tg 7712

<210> 27

<211> 7925

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1245

<400> 27

```

cagctgcgcg ctgcctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
actaggggtt cttgtagt ttatgattaac ccgcatgct acttatctac acgcgtagat 180
ctagtgcac cacttcttat ttgtgtaata agaaaattgg gaaaacgac ttcaaatatgc 240
ttaccaagct gtgattccaa atattacgta aatacacttg caaaggagga tgtttttagt 300
agcaatttgt actgatggta tggggccaag agatatact tagagggagg gctgagggtt 360
tgaagtcгаа ctctaagcc agtgccagaa gagccaagga caggtacggc tgtcatcact 420

tagacctcac cctgtggagc cacaccctag ggttggccaa tctactcca ggagcaggga 480
gggcaggagc cagggtcggg cataaaagtc agggcagagc catctattgc ttacatttgc 540
ttctgacaca actgtgttca ctagcaacct caaacagaca ccatggtcca tcttacaccg 600
gaggagaagt ccgctgtaac ggcactgtgg gggaaaagta atgtcgatga agtcggcggg 660
gaagcactcg gcaggttgct ggtagtgtac ccgtggacac aacgattctt tgaaagtttc 720
ggggacctgt ccacaccga tgcgtgtgat ggtaatccaa aagtaaaagc acacggcaag 780
aaagtcctcg gcgcgtttag tgatggtctg gccatttgg ataacttgaa gggtacattc 840
gcgcagcttt ccgaactcca ctgtgacaag ttgcacgtag atccagaaaa ctccggctt 900

ctgggcaatg tgcttgtatg cgttctggct caccattttg ggaaggagtt taccacacc 960
gtgcaagcgg cttacaaaaa agtggctgca ggagtggcta atgcccttgc acataaatat 1020
cactaaggta ccgagcatct taccgccatt tattccata tttgttctgt tttctttgat 1080
ttgggtatac atttaaatgt taataaaaca aaatgggtgg gcaatcattt acatttttag 1140
ggatatgtaa ttactagttc aggtgtattg ccacaagaca aacatgttaa gaaactttcc 1200
cgttatttac gctctgttcc tgtaaatcaa cctctggatt acaaaatttg tgaaagattg 1260
actgatattc ttaactatgt tgcctctttt acgctgtgtg gatagctgc tttatagcct 1320
ctgtatctag ctattgcttc ccgtacggct ttcgttttct cctccttgta taaatcctgg 1380

ttgctgtctc ttttagagga gttgtggccc gttgtccgtc aacgtggcgt ggtgtgctct 1440
gtgtttgctg acgcaacccc cactggctgg ggcattgcca ccacctgtca actcctttct 1500
gggactttcg ctttccccct ccgcatcgcc acggcagaac tcacgcgcgc ctgccttgcc 1560

```

cgctgctgga caggggctag gttgctgggc actgataatt ccgtgggtgtt gtctgtgcct 1620
tctagtgtcc agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaaggt 1680
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgattg tctgagtagg 1740
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac 1800
aatagcaggc atgctgggga tgcggtgggc tctatggctt aattaacgag atcgagacca 1860

tcctggctaa cacagtgaat ccccgctctt actaaaaaaa tacaaaaaat tagccgggct 1920
tggtggcggg tgccgttagt cccagctact atggaggctg aggcgggaga atggcgtgaa 1980
cgcggggggc ggagcttgca gtgagcagag atcaggggcc actgcactcc agcctgggag 2040
acagagagag actctgtctc aaaaaaaga aaaaaaaat ttagtagact agctaaaaaa 2100
atccagagat agttattgat gcatatgtaa aagtcttcca atatttaca gtacaatgaa 2160
aaaaaataa cttgaatta agtgtagaac tcattgacaa tgtttcaaag gatgtgaggg 2220
ataaactaaa atttgggcag tacatgctgt tcctgtgtac ttggaacaga gggagaaaat 2280
ctgggctgga aatattgtta taggagttag cacatgaagg tgacaactaa attatttga 2340

gtagatggag tcaccagcac atgtgaatag ttttagaatg aaatgacca agatagaact 2400
ttggagagcc ccaaattta aataaaatca gtataagaga agaggaagaa accaaatggt 2460
atactagtct aaattgtttc ttagtgacaa aagaataacc tgaatattag attagctgcc 2520
tatatgctct ctgaatcaat ttatttcaac atgcaacagt ccgcggaac agagaaacag 2580
gagaatatgg gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggccaa 2640
gaacagttgg aacagcagaa tatgggccaa acaggatata tgtgtaagc agttcctgcc 2700
ccggctcagg gccaaagaac gatggtcccc agatgcggtc ccgcctcag cagtttctag 2760
agaaccatca gatgtttcca ggggtgcccc aggacctgaa atgacctgt gccttatttg 2820

aactaacaa tcagttcgct tctcgttctt gttcgcgcg cttctgtccc cgagctctat 2880
ataagcagag ctggtttagt gaaccgtcag atcgctgga gacgccatcc acgtgtttt 2940
gatttcata gaaggcgcc gcgcccac catggtgagc aaggcgagg agctgttcac 3000
cggggtggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 3060
gtccggcgag ggcgaggcg atgccacct cggcaagctg accctgaagt tcatctgac 3120
caccggaag ctgccccgtc cctggccac cctcgtgacc accctgacct acggcgtgca 3180
gtgcttcagc cgtaccccc accacatgaa gcagcacgac ttcttcaagt ccgcatgcc 3240
cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagaccg 3300

cgccgaggtg aagttcgagg gcgacacct ggtgaaccgc atcgagctga agggcatcga 3360
cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccaca 3420

cgctctatc atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 3480
 caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca ccccatcg 3540
 cgacggcccc gtgctgctgc ccgacaacca ctacctgagc acccagtcgg cctgagcaa 3600
 agacccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg ccgccgggat 3660
 cactctcggc atggacgagc tgtacaaggg aagcggagct actaacttca gcctgctgaa 3720
 gcaggctgga gacgtggagg agaaccctgg acctacctgc aggcctgaga acttcagggt 3780

gagtctatgg gacgcttgat gttttctttc cccttctttt ctatgggtta gttcatgtca 3840
 taggaagggg ataagtaaca gggtagctt tagaatggga aacagacgaa tgattgcac 3900
 agtgtggaag ttcaggatc gttttagttt cttttatttg ctgttcataa caattgtttt 3960
 cttttgttta attcttgctt tctttttttt tcttctccgc aatttttact attatactta 4020
 atgccttaac attgtgtata acaaaaggaa atatctctga gatacattaa gtaacttaaa 4080
 aaaaaacttt acacagctcg cctagtacat tactatttgg aatatatgtg tgcttatttg 4140
 catattcata atctccctac tttgtcgacg tagataagta gcatggcggg ttaatcatta 4200
 actacaagga acccctagtg atggagtggg ccactccctc tctgcgcgct cgctcgtca 4260

ctgaggccgg gcgaccaaag gtgcgccgac gcccgggctt tgcccgggcg gcctcagtga 4320
 gcgagcgagc gcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttccaa 4380
 cagttgcga gcctgaatgg cgaatggcga ttccgttgca atggctggcg gtaatatgt 4440
 tctggatatt accagcaagg ccgatagttt gagttcttct actcaggcaa gtgatgttat 4500
 tactaatcaa agaagtattg cgacaacggt taatttgcgt gatggacaga ctcttttact 4560
 cgggtgcctc actgattata aaaacacttc tcaggattct ggcgtagcgt tcctgtctaa 4620
 aatcccttta atcgccctcc tgtttagctc ccgctctgat tctaacagg aaagcacgtt 4680
 atacgtgctc gtcaaagcaa ccatagtacg cgccctgtag cggcgcatca agcgcggcgg 4740

gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgtcctt 4800
 tcgttttctt cccttctttt ctgccacgt tcgccggtt tccccgtaa gctctaaatc 4860
 gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg 4920
 attagggtga tggttcacgt agtgggcat cgccctgata gacgttttt cgcccttga 4980
 cgttgagtc cactttcttt aatagtggac tcttggtcca aactggaaca aactcaacc 5040
 ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa 5100
 aaaatgagct gatttaacaa aaatttaacg cgaatttta caaaatatta acgtttacaa 5160
 tttaaatatt tgcttatata atcttctgt ttttggggct tttctgatta tcaaccgggg 5220

tacatatgat tgacatgcta gttttacgat taccgttcat cgattctctt gtttgcctca 5280

gactctcagg caatgacctg atagcctttg tagagacctc tcaaaaatag ctaccctctc 5340
 cggcatgaat ttatcagcta gaacggttga atatcatatt gatggtgatt tgactgtctc 5400
 cggcctttct caccgcttg aatctttacc tacacattac tcaggcattg catttaaaat 5460
 atatgagggt tctaaaaatt ttatccttg cgttgaaata aaggcttctc ccgcaaaagt 5520
 attacagggt cataatgttt ttggtacaac cgatttagct ttatgctctg aggcctttatt 5580
 gcttaatfff gctaattctt tgccttgctt gtatgattta ttggatgttg gaatcgcttg 5640
 atcggttatt ttctccttac gcactgtgct ggtatttcac accgcatatg gtgcactctc 5700

agtacaatct gctctgatgc cgcatagtta agccagcccc gacacccgcc aacacccgct 5760
 gacgcgccct gacgggcttg tctgtctccg gcatccgctt acagacaagc tgtgaccgtc 5820
 tccgggagct gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gagacgaaag 5880
 ggcctcgtga tacgcctatt ttataggtt aatgtcatga taataatggt ttcttagacg 5940
 tcagggtgca cttttcgggg aaatgtgcgc ggaacccta tttgtttatt ttctaaata 6000
 cattcaaata tgtatccgct catgagacaa taacctgat aaatgcttca ataatttga 6060
 aaaaggaaga gtatgagat tcaacatttc cgtgtcgccc ttattccctt ttttgcgca 6120
 ttttgccttc ctgtttttgc taccacagaa acgctggtga aagtaaaaga tgctgaagat 6180

cagttgggtg cagcagtggt ttacatcgaa ctggatctca acagcggtaa gatccttgag 6240
 agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagtctt gctatgtggc 6300
 gcggtattat cccgtattga cgccgggcaa gagcaactcg gtcgcccgat aactattctt 6360
 cagaatgact tgggtgagta ctcaccagtc acagaaaagc atcttacgga tggcatgaca 6420
 gtaagagaat tatgcagtgc tgccataacc atgagtata aactgcggc caacttactt 6480
 ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat gggggatcat 6540
 gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccatacaaaa cgacgagcgt 6600
 gacaccacga tgcctgtagc aatggcaaca acgttgcgca aactattaac tggcgaacta 6660

cttactctag cttcccggca acaattaata gactggatgg aggcggataa agttgcagga 6720
 ccacttctgc gctcggccct tccggtggtc tggtttattg ctgataaatc tggagccggt 6780
 gagcgtgggt ctcgcggtat cattgcagca ctggggccag atggtaagcc ctcccgtatc 6840
 gtagttatct acacgacggg gactcaggca actatggatg aacgaaatag acagatcgct 6900
 gagatagggt cctcactgat taagcattgg taactgtcag accaagtta ctcatatata 6960
 ctttagattg atttaaaact tcatTTTTaa tttaaaagga tctaggtgaa gatccttttt 7020
 gataatctca tgaccaaaa cctttaacgt gagttttcgt tccactgagc gtcagacccc 7080
 gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg 7140

caaacaaaaa aaccaccgct accagcgggtg gtttgtttgc cggatcaaga gctaccaact 7200
 ctttttcga aggttaactgg cttcagcaga ggcagatac caaatactgt ctttctagt 7260
 tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata cctcgtcttg 7320
 ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac 7380
 tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacggggggg ttcgtgcaca 7440
 cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga 7500
 gaaagcgcca cgcttccga agggagaaaag gcggacaggt atccggtaag cggcagggtc 7560
 ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct 7620

gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc aggggggcgg 7680
 agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct 7740
 tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc 7800
 tttgagttag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc 7860
 gaggaagcgg aagagcgccc aatacgcaaa ccgctctcc ccgcgcgttg gccgattcat 7920
 taatg 7925

<210> 28

<211> 5216

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1246

<400> 28

cagctgcgcg ctgcctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 tttggtgcc cggcctcagt gagegagcga ggcgcagag agggagtggc caactccatc 120
 actaggggtt cttgtagt aatgattaac ccgcatgct acttatctac acgcgtatag 180
 ctatatacac acatatatat atatatTTTT tcttttctta ccagaagggt ttaatccaaa 240
 taaggagaag atatgcttag aaccgaggta gagttttcat ccattctgtc ctgtaagtat 300
 tttgcatatt ctggagacgc aggaagagat ccatctacat atcccaaagc tgaattatgg 360
 tagacaaaac tcttccactt ttagtgcac aacttcttat ttgtgtaata agaaaattgg 420

gaaaacgac ttcaatatgc ttaccaagct gtgattccaa atattacgta aatacacttg 480
 caaaggagga tgtttttagt agcaatttgt actgatggta tggggccaag agatatatct 540
 tagaggagg gctgagggtt tgaagtccaa ctccaaagcc agtgccagaa gagccaagga 600

caggtacggc tgcacact tagacctcac cctgtggagc cacaccctag ggttgccaa 660
 tctactccca ggagcaggga gggcaggagc cagggctggg cataaaagtc agggcagagc 720
 catctattgc ttacatttgc ttctgacaca actgtgttca ctagcaacct caaacagaca 780
 ccatgggtcca tcttacaccg gtggagaaat ccgctgtaac ggcactttgg gggaaagtta 840
 atgtcgatga ggtcgggtgt gaggccctgg gcaggttgg atcaaggta caagacaggt 900

 ttaaggagac caatagaaac tgggcatgtg gagacagaga agactcttgg gtttctgata 960
 ggcactgact ctctctgcct attggtctat tttccaccc ttaggtgct ggtggtctac 1020
 ccttggacc agaggttctt tgagtccttt ggggatctgt ccactctga tctgttatg 1080
 ggcaacccta aggtgaaggc tcatggcaag aaagtgtcg gtgcctttag tgatggcctg 1140
 gctcacctgg acaacctcaa gggcaccttt gccacactga gtgagctgca ctgtgacaag 1200
 ctgcacgtgg atctgagaa cttcaggggtg agtctatggg acgcttgatg ttttctttcc 1260
 ctttctttt taigttaag ttcatgtcat aggaaggga taagtaacag ggtacagttt 1320
 agaatgggaa acagacgaat gattgcatca gtgtggaagt ctcaggatcg ttttagtttc 1380

 ttttatttgc tgtcataac aattgttttc tttgtttta ttcttcttt ctttttttt 1440
 ctctccgca attgtcgac gtagataagt agcatggcgg gtaatcatt aactacaagg 1500
 aaccctagt gatggagtgt gccactcct ctctgcgcgc tcgtcgtc actgaggccg 1560
 ggcgacaaa ggtgcgccga cgcccgggct ttgcccgggc ggctcagtg agcgagcgag 1620
 cgcccgact ggcgtaatat cgaaggaggc cgcaccgatc gcccttcca acagtgcgc 1680
 agcctgaatg gcgaatggcg attccgttgc aatggctggc ggtaatatg ttctggatat 1740
 taccagcaag gccgatagtt tgagtcttc tactcaggca agtgatgta ttactaatca 1800
 aagaagtatt gcgacaacgg ttaatttgcg tgatggacag actcttttac tcggtggcct 1860

 cactgattat aaaaacactt ctcaggattc tggcgtagcg ttctgtcta aaatcccttt 1920
 aatcgccctc ctgtttagct cccgtctga ttctaacgag gaaagcacgt tatacgtgct 1980
 cgtcaaagca accatagtag gcgccctgta gcggcgcat aagcgcgcg ggtgtggtg 2040
 ttacgcgag cgtgaccgt acacttgeca gcgccctage gccgctcct ttcgtttct 2100
 tcccttctt tctgccacg ttcccggtt tccccgtca agctctaat cgggggctcc 2160
 ctttagggtt ccgatttagt gctttacggc acctgaccc caaaaaactt gattaggggtg 2220
 atggttcacg tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttgaggt 2280
 ccacgttctt taatagtga ctctgttcc aaactggaac aacactcaac cctatctcgg 2340

 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2400
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca atttaatat 2460

ttgcttatac aatcttcctg tttttggggc ttttctgatt atcaaccggg gtacatatga 2520
 ttgacatgct agttttacga ttaccgttca tgcattctct tgtttgtccc agactctcag 2580
 gcaatgacct gatagccttt gtagagacct ctcaaaaata gctaccctct ccggcatgaa 2640
 tttatcagct agaacgggtg aatatcatat tgatgggatg ttgactgtct ccggcctttc 2700
 tcaccgcttt gaatctttac ctacacatta ctacggcatt gcatttaaaa tatatgaggg 2760
 ttctaaaaat ttttatcctt gcgttgaaat aaaggcttct cccgcaaaag tattacaggg 2820

 tcataatggt ttigtacaa ccgatttagc tttatgctct gaggccttat tgcttaattt 2880
 tgctaattct ttgccttgcc tgtatgattt attggatggt ggaatgcct gatgcggtat 2940
 tttctcctta cgcactgtg cggtatttca caccgcatat ggtgcactct cagtacaatc 3000
 tgctctgatg ccgcatagtt aagccagccc cgacaccgc caacaccgc tgacgcgcc 3060
 tgacgggctt gtctgtccc ggcacccgt tacagacaag ctgtgaccgt ctccgggagc 3120
 tgcattgtgc agaggctttc accgtcatca ccgaaacgc cgagacgaaa gggcctcgtg 3180
 ataccctat tttataggt taatgtcatg ataataatgg tttcttagac gtcaggtggc 3240
 acttttcggg gaaatgtgcg cggaaccct atttgtttat tttctaaat acattcaaat 3300

 atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag 3360
 agtatgagta ttcaacattt ccgtgtgcc ctattccct tttttgcgc attttgcctt 3420
 cctgtttttg ctaccacaga aacgtgggtg aaagtaaaag atgtgaaga tcagttgggt 3480
 gcacgagtgg gttacatga actggatctc aacagcggta agatccttga gatttttcgc 3540
 cccgaagaac gttttccat gatgagcact tttaaagtgc tgctatgtgg cgcggtatta 3600
 tcccgtattg acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac 3660
 ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa 3720
 ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg 3780

 atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc 3840
 ctgtatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg 3900
 atgcctgtag caatggcaac aacgttgcc aaactattaa ctggcgaact acttactcta 3960
 gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg accacttctg 4020
 cgctcggccc ttccgctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg 4080
 tctcgggta tcattgcagc actggggcca gatggttaagc cctcccgat cgtagtattc 4140
 tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt 4200
 gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat acttttagatt 4260

 gatttaaaac ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc 4320

atgaccaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag 4380
 atcaaaggat ctctttgaga tctttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa 4440
 aaaccaccgc taccagcggg ggtttgtttg ccggatcaag agctaccaac tctttttccg 4500
 aaggttaactg gcttcagcag agcgcagata ccaaatactg tctttctagt gtagccgtag 4560
 ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaattcctg 4620
 ttaccagtgg ctgctgccag tggcgataag tctgtcttta ccgggttgga ctcaagacga 4680
 tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc 4740

ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc 4800
 acgcttcccg aaggagagaaa ggccgacagg tatccggtaa gcggcagggt cggaacagga 4860
 gagcgacga gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt 4920
 cgccacctct gacttgagcg tcatTTTTTg tgatgctcgt caggggggcg gacccatgg 4980
 aaaaacgcca gcaacgcggc ctttttacgg ttccctggcct tttgctggcc ttttgctcac 5040
 atgttctttc ctgcgttacc ccctgattct gtggataacc gtattaccgc ctttgagtga 5100
 gctgataccg ctgcgccgag ccgaacgacc gagcgagcag agtcagttag cgaggaagcg 5160
 gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatg 5216

<210> 29

<211> 4816

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1247

<400> 29

cagctgcgcg ctgctcgtc cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 ttgtgtgcc cgccctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
 actaggggtt cctttagtct aatgattaac ccgcatgct acttatctac acgcgtatag 180
 ctagtgcac aacttcttat ttgtgtaata agaaaattgg gaaaacgac ttcaatatgc 240
 ttaccaagct gtgattccaa atattacgta aatacacttg caaaggagga tgtttttagt 300
 agcaatttgt actgatggtg tggggccaag agatatatct tagaggagg gctgagggtt 360

tgaagtccaa ctctaagcc agtgccagaa gagccaagga caggtacggc tgtcatcact 420
 tagacctcac cctgtggagc cacaccctag ggttggccaa tctactcca ggagcaggga 480
 gggcaggagc cagggtggg cataaaagtc agggcagagc catctattgc ttacatttgc 540

ttctgacaca acigtgttca ctagcaacct caaacagaca ccatgggtcca tcttacaccg 600
 gtggagaaat ccgtgtaac ggcactttgg gggaaagtta atgtcgatga ggtcgggtgg 660
 gaggccctgg gcaggttgg atcaaggta caagacaggt ttaaggagac caatagaaac 720
 tgggcatgtg gagacagaga agactcttgg gtttctgata ggcactgact ctctctgcct 780
 attggtctat tttccaccc ttaggtgct ggtggtctac ccttggaccc agaggttctt 840

 tgagtctttt ggggatctgt ccactcctga tctgtttatg ggcaacccta aggtgaaggc 900
 tcatggcaag aaagtgtcgt gtgccttttag tgatggcctg gctcacctgg acaacctcaa 960
 gggcaccttt gccacactga gtgagctgca ctgtgacaag ctgcacgtgg atcctgagaa 1020
 cttcaggggtg agtctatggg acgcttgaig ttttgtcgac gtagataagt agcatggcgg 1080
 gttaatcatt aactacaagg aaccctagt gatggagtgt gccactcct ctctgcgcgc 1140
 tcgtcgtc actgaggcgg ggcgacaaa ggtcgccga cgccgggct ttgcccgggc 1200
 ggctcagtg agcgagcgag cgcgccagct ggcgtaatag cgaagaggcc cgcaccgac 1260
 gcccttccca acagtgcgc agcctgaatg gcgaatggcg attccgttgc aatggctggc 1320

 ggtaatatgt ttctggatat taccagcaag gccgatagtt tgagttcttc tactcagga 1380
 agtgatgtta ttactaatca aagaagtatt gcgacaacgg ttaatttgcg tgatggacag 1440
 actcttttac tcggtggcct cactgattat aaaaacactt ctcaggattc tggcgtaccg 1500
 ttctgtcta aaatcccttt aatcggcctc ctgttttagct cccgctctga ttctaacgag 1560
 gaaagcacgt tatcgtgct cgtcaaagca accatagtag gcgcctgta gcggcgcatt 1620
 aagcgcgcg ggtgtggtgg ttacgcgcag cgtgaccgt acacttgcca gcgccttagc 1680
 gccgctcct ttcgctttct tcccttcctt tctgccacg ttcgccggct tccccgtca 1740
 agctctaaat cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgacc 1800

 caaaaaactt gattaggggtg atggttcacg tagtgggcca tcgccctgat agacggtttt 1860
 tcgccctttg acgttggagt ccacgttctt taatagtga ctcttgttcc aaactggaac 1920
 aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc 1980
 ctattggta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt 2040
 aacgtttaca atttaaatat ttgcttatac aatcttctg ttttggggc tttctgatt 2100
 atcaaccggg gtacatatga ttgacatgct agttttacga ttaccgttca tcgattctct 2160
 tgtttgctcc agactctcag gcaatgacct gatagccttt gtagagacct ctcaaaaata 2220
 gctaccctct ccggcatgaa tttatcagct agaacggttg aatatcatat tgatggtgat 2280

 ttgactgtct ccggccttct tcaccggttt gaatctttac ctacacatta ctcaggcatt 2340
 gcatttaaaa tatatgaggg ttctaataat tttatcctt gcgttgaaat aaaggcttct 2400

cccgc aaaag tattacaggg tcataatgtt ttgtgtacaa ccgatttagc tttatgctct 2460
gaggctttat tgcttaattt tgctaattct ttgccttgcc tgtatgattt attggatgtt 2520
ggaatcgct gatgcggtat tttctcctta cgcactctgtg cggatattca caccgcata 2580
ggtgcactct cagtacaatc tgctctgatg ccgcatagtt aagccagccc cgacacccgc 2640
caacacccgc tgacgcgccc tgacgggctt gtctgtctcc ggcatccgct tacagacaag 2700
ctgtgaccgt ctccgggagc tgcatgtgtc agaggttttc accgtcatca ccgaaacgcg 2760

cgagacgaaa gggcctcgtg atacgcctat ttttataggt taatgtcatg ataataatgg 2820
tttcttagac gtcagggtgc acttttcggg gaaatgtgcg cggaaccctt attgtttat 2880
ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc 2940
aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc ctatttcctt 3000
tttttgcggc attttgcctt cctgtttttg ctacccagaa aacgtgtgtg aaagtaaaag 3060
atgctgaaga tcagttaggt gcacagatgg gttacatcga actggatctc aacagcggta 3120
agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc 3180
tgctatgtgg cgcggtatta tcccgtattg acgcgggga agagcaactc ggtcgccgca 3240

tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg 3300
atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg 3360
ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca 3420
tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa gccatacaa 3480
acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgccg aaactattaa 3540
ctggcgaact acttactcta gcttcccggc aacaattaat agactggatg gaggcggata 3600
aagttgcagg accacttctg cgctcgcccc ttccggctgg ctggtttatt gctgataaat 3660
ctggagccgg tgagcgtggg tctcgggta tcattgcagc actggggcca gatggttaagc 3720

cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata 3780
gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt 3840
actcatatat actttagatt gatttaaaac ttcattttta atttaaaagg atctagggtga 3900
agatcctttt tgataatctc atgacaaaaa tccttaacg tgagttttcg ttccactgag 3960
cgtagacccc cgtagaaaag atcaaaggat cttcttgaga tctttttttt ctgcgcgtaa 4020
tctgtgtctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccgatcaag 4080
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgagata ccaaatactg 4140
tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat 4200

acctcgtctt gctaactctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta 4260

ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg 4320
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc 4380
gtgagctatg agaaagcgcc acgcttcccg aaggagaaaa ggcgacagc tatccggtaa 4440
gcggcagggt cggaacagga gagcgacga gggagcttcc agggggaaac gcctggtatc 4500
tttatagtcc tgcgggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt 4560
caggggggag gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttcttgccct 4620
tttgctggcc tttgctcac atgttcttct ctgcgttctc cctgattct gtggataacc 4680

gtattaccgc ctttgagtg gctgataccg ctgccgcag ccgaacgacc gagcgagcg 4740
agtcagttag cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt 4800
ggccgattca ttaatg 4816

<210> 30

<211> 4416

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1248

<400> 30

cagctgcgcg ctgcctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
tttggtgcc cgccctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
actaggggtt ccttgtagtt aatgattaac ccgcatgct acttatctac acgcgtagat 180

cttgccagaa gagccaagga caggtacggc tgtcatcact tagacctcac cctgtggagc 240
cacaccttag ggttgccaa tctactcca ggagcaggga gggcaggagc cagggtggg 300
cataaaagtc agggcagagc catctattgc ttacatttgc ttctgacaca actgtgttca 360
ctagcaacct caaacagaca ccatgggtcca tcttacaccg gtggagaaat ccgctgtaac 420
ggcacttttg gggaaagtta atgtcgatga ggtcgggtgt gaggccctgg gcaggttgg 480
atcaaggtta caagacaggt ttaaggagac caatagaaac tgggcatgtg gagacagaga 540
agactcttgg gtttctgata ggcactgact ctctctgcct attggtctat ttccccacc 600
ttaggtgctt ggtgtctac ccttggaacc agaggttctt tgagtccttt gggggtcgac 660

gtagataagt agcatggcgg gttaatcatt aactacaagg aaccctagt gatggagtgt 720
gccactccct ctctgcgcgc tcgctcgtc actgaggccg ggcgacaaa ggtcgccga 780
cgccgggct ttgccgggc gccctcagt agcgagcgag cgcgccagct ggcgtaatag 840

cgaagaggcc cgcaccgata gcccttccca acagttgcgc agcctgaatg gcgaatggcg 900
 attccgttgc aatggctggc ggtaatatg ttctggatat taccagcaag gccgatagtt 960
 tgagtcttc tactcaggca agtgatgta ttactaatca aagaagtatt gcgacaacgg 1020
 ttaatttgcg tgatggacag actcttttac tcggtggcct cactgattat aaaaacactt 1080
 ctccaggattc tggcgtaccg ttctgtcta aaatcccttt aatcggcctc ctgttttagct 1140

 cccgctctga ttctaacgag gaaagcacgt tatacgtgct cgtcaaagca accatagtag 1200
 gcgccctgta gcggcgcat aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct 1260
 acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttctt tctcgccacg 1320
 ttgcgcggct tccccgta agctctaaat cgggggctcc ctttagggtt ccgatttagt 1380
 gctttacggc acctcgacc caaaaaactt gattaggtg atggttcacg tagtgggcca 1440
 tcgccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt taatagtga 1500
 ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt tgatttataa 1560
 gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca aaaatttaac 1620

 gcgaatttta acaaaatatt aacgtttaca atttaaatat ttgcttatac aatcttctg 1680
 tttttggggc ttttctgatt atcaaccggg gtacatatga ttgacatgct agttttacga 1740
 ttaccgttca tcgattctct tgtttgtcc agactctcag gcaatgacct gatagccttt 1800
 gtagagacct ctcaaaaata gctaccctct ccggcatgaa tttatcagct agaacggttg 1860
 aatatcatat tgatggtgat ttgactgtct ccggcctttc tcaccggtt gaacttttac 1920
 ctacacatta ctccaggcatt gcatttaaaa tatatgaggg ttctaaaaat ttttatcctt 1980
 gcgttgaaat aaaggcttct cccgcaaaag tattacaggg tcataatgtt tttggtacaa 2040
 ccgatttagc tttatgtctt gaggtttt atgttaattt tgctaattct ttgccttgcc 2100

 tgatgattt attgcatgtt ggaatcgct gatgcggtat tttctctta cgcactctgt 2160
 cggatattca caccgcatat ggtgcactct cagtacaatc tgctctgat ccgcatagtt 2220
 aagccagccc cgacacccgc caacacccgc tgacgcgccc tgacgggctt gtctgtctcc 2280
 ggcatccgct tacagacaag ctgtgaccgt ctccgggagc tgcatgtgtc agaggttttc 2340
 accgtcatca ccgaaacgcg cgagacgaaa gggcctcgtg atacccctat ttttataggt 2400
 taatgtcatg ataataatg tttcttagac gtcaggtggc acttttcggg gaaatgtgcg 2460
 cggaacccct attgtttat tttctaaat acattcaaat atgtatccgc tcatgagaca 2520
 ataaccctga taaatgctc aataatattg aaaaaggaag agtatgagta ttcaacattt 2580

 ccgtgtcgcc ctattccct ttttgcggc attttgcctt cctgttttg ctcaccaga 2640
 aacgttggtg aaagtaaaag atgtgaaga tcagttgggt gcacgagtg gttacatcga 2700

actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat 2760
gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtattg acgccgggca 2820
agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt 2880
cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 2940
catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 3000
aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcggt gggaaccgga 3060

gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac 3120
aacgttgcgc aaactattaa ctggcgaact acttactcta gtttcccggc aacaattaat 3180
agactggatg gaggcggata aagttagcagg accacttctg cgctcggccc ttccggctgg 3240
ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggtg tcattgcagc 3300
actggggcca gatggttaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc 3360
aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 3420
gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac ttcattttta 3480
atttaaaagg atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg 3540

tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga 3600
tccttttttt ctgcgcgtaa tctgtctgtt gcaaacaaaa aaaccaccgc taccagcggg 3660
ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg gtttcagcag 3720
agcgcagata ccaataactg tccttctagt gtagccgtag ttaggccacc acttcaagaa 3780
ctctgtagca ccgcctacat acctcgtctt gctaactctg ttaccagtgg ctgctgccag 3840
tggcgataag tcgtgtctta ccgggttggg ctcaagacga tagttaccgg ataaggcgca 3900
gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac 3960
cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagagaa 4020

ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc 4080
aggggggaaac gcctggtatc tttatagtcc tctcgggttt cgccacctct gacttgagcg 4140
tcgatttttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc 4200
ctttttacgg ttcttggcct tttgctggcc ttttctcac atgttctttc ctgcgttacc 4260
ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgccag 4320
ccgaacgacc gagcgacgag agtcagttag cgaggaagcg gaagagcgcc caatacgcaa 4380
accgcctctc cccgcgcgtt ggccgattca ttaatg 4416

<211> 4116

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1249

<400> 31

cagctgcgcg ctgcctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
 actaggggtt ccttgtagtt aatgattaac ccgcatgct acttatctac acgcgtagat 180
 ctacatttgc ttctgacaca actgtgttca ctagcaacct caaacagaca ccatgggtcca 240
 tcttacaccg gtggagaaat ccgctgtaac ggcactttgg gggaaagtta atgtcgatga 300
 ggtcgggtgtt gaggccctgg gcaggttgggt atcaaggtta caagacaggt ttaagtcgac 360

gtagataagt agcatggcgg gttaatcatt aactacaagg aaccctagt gatggagtgt 420
 gccactccct ctctgcgcgc tcgctcgctc actgaggccg ggcgacaaa ggtcgcccga 480
 cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag cgcgccagct ggcgtaatat 540
 cgaagaggcc cgcaccgatc gcccttccca acagttgcgc agcctgaatg gcgaatggcg 600
 attccgttgc aatggctggc ggtaatatgt ttctggatat taccagcaag gccgatagtt 660
 tgagtcttct tactcaggca agtgatgtta ttactaatca aagaagtatt gcgacaacgg 720
 ttaatttgcg tgatggacag actcttttac tcggtggcct cactgattat aaaaacactt 780
 ctcaggattc tggcgtaccg ttctgtctta aaatcccttt aatcggcctc ctgttttagct 840

cccgtcttga ttctaacgag gaaagcacgt tatacgtgct cgtcaaagca accatagtac 900
 gcgccttgta gcggcgcatc aagcgcgcg ggtgtggtgg ttacgcgcag cgtgaccgct 960
 acacttgcca gcgccttagc gcccgctcct ttcgctttct tcccttcctt tctcgccacg 1020
 ttcgcccggct ttccccgtca agctctaaat cgggggctcc ctttaggggt ccgatttagt 1080
 gctttacggc acctcgacc caaaaaactt gattaggggt atggttcacg tagtgggcca 1140
 tcgccttgat agacggtttt tcgccccttg acgttggagt ccacgttctt taatagtga 1200
 ctcttggtcc aaactggaac aacactcaac cctatctcgg tctattcttt tgatttataa 1260
 gggattttgc cgatttcggc ctattgggtta aaaaatgagc tgatttaaca aaaatttaac 1320

gcgaatttta acaaaatatt aacgtttaca atttaaatat ttgcttatac aatcttcctg 1380
 tttttggggc ttttctgatt atcaaccggg gtacatatga ttgacatgct agttttacga 1440
 ttaccgttca tcgattctct tgtttgctcc agactctcag gcaatgacct gatagccttt 1500
 gtagagacct ctcaaaaata gctaccctct ccggcatgaa tttatcagct agaacggttg 1560

aatatcatat tgatggatgat ttgactgtct ccggcctttc tcacccgttt gaatctttac 1620
ctacacatta ctacaggcatt gcatttaaaa tatatgaggg ttctaaaaat ttttatcctt 1680
gcgttgaaat aaaggcttct cccgcaaaag tattacaggg tcataatgtt ttgtgtacaa 1740
ccgatttagc tttatgtctt gaggccttat tgcttaattt tgctaattct ttgccttgcc 1800

tgtatgattt attggatgtt ggaatcgctt gatgcggtat tttctcctta cgcactgttg 1860
cggtatttca caccgcatat ggtgcactct cagtacaatc tgctctgatg ccgcatagtt 1920
aagccagccc cgacacccgc caacacccgc tgacgcgccc tgacgggctt gtctgtctcc 1980
ggcatccgct tacagacaag ctgtgaccgt ctccgggagc tgcatgtgtc agaggttttc 2040
accgtcatca ccgaaacgcg cgagacgaaa gggcctcgtg atacgcctat ttttataggt 2100
taatgtcatg ataataatgg tttcttagac gtcagggtggc acttttcggg gaaatgtgcg 2160
cggaaccctt atttgtttt ttttctaaat acattcaaat atgtatccgc tcatgagaca 2220
ataaccctga taaatgtctt aataatattg aaaaaggaag agtatgagta ttcaacattt 2280

ccgtgtcgcc cttattccct tttttgcggc attttgcctt cctgtttttg ctcaccaga 2340
aacgttgggt aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga 2400
actggatctc aacagcggtt agatccttga gagttttcgc cccgaagaac gttttccaat 2460
gatgagcact tttaaagtgc tgctatgtgg cgcggtatta tcccgtattg acgccgggca 2520
agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt 2580
cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 2640
catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 2700
aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga 2760

gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac 2820
aacgttgcgc aaactattaa ctggcgaaact acttactcta gcttcccggc aacaattaat 2880
agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg 2940
ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc 3000
actggggcca gatggttaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc 3060
aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 3120
gtlaactgtc gaccaagttt actcatatat acttttagatt gattttaaac ttcattttta 3180
attttaaagg atctaggtag agatcctttt tgataatctc atgacaaaaa tcccttaacg 3240

tgagttttcg ttccactgag cgtcagacc cgtagaaaag atcaaaggat cttcttgaga 3300
tccttttttt ctgcggttaa tctgtctgtt gcaaacaaaa aaaccaccgc taccagcgtt 3360
ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag 3420

agcgcagata ccaataactg tccttctagt gtagccgtag ttaggccacc acttcaagaa 3480
ctctgtagca ccgcctacat acctcgtctt gctaactctg ttaccagtgg ctgctgccag 3540
tggcgataag tcgtgtctta ccgggttggg ctcaagacga tagttaccgg ataaggcgca 3600
gcggtcgggc tgaacggggg gtctgtgcac acagcccagc ttggagcgaa cgacctacac 3660
cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagagaaa 3720

ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgacga gggagcttcc 3780
aggggggaaac gcctggatc tttatagtcc tctcgggttt cgccacctct gacttgagcg 3840
tcgatttttg tgatgctcgt caggggggcg gacccatagg aaaaacgcca gcaacgcggc 3900
ctttttacgg ttcttggcct tttgttggcc ttttgtcac atgttcttcc ctgcgttacc 3960
ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgccgag 4020
ccgaacgacc gagcgacgag agtcagttag cgaggaagcg gaagagcgcc caatacgcaa 4080
accgcctctc cccgcgcgtt ggccgattca ttaatg 4116

<210> 32

<211> 7995

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1289

<400> 32

cagctgcgcg ctgctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
actaggggtt cctttagatt aatgattaac ccgcatgct acttatctac acgcgtatag 180
ctagtgcac aacttcttat ttgtgtaata agaaaattgg gaaaacgac ttcaatatgc 240
ttaccaagct gtgattccaa atattacgta aatacattg caaaggagga tgtttttagt 300
agcaatttgt actgatggtg tggggccaag agatatact tagagggagg gctgagggtt 360

tgaagtccaa ctctaagcc agtgccagaa gagccaagga caggtacggc tgtcatcact 420
tagacctcac cctgtggagc cacaccctag ggttggccaa tctactccca ggagcaggga 480
gggcaggagc cagggtggg cataaaagtc agggcagagc catctattgc ttacatttgc 540
ttctgacaca actgtgttca ctagcaacct caaacagaca ccatggtgca tctgactcct 600
gaggagaagt ccgctgtaac ggcactgttg gggaaagtta atgtcgatga agtcggcggg 660
gaagcactcg gcaggttctt ggtagtgtac ccgtggacac aacgattctt tgaaagtttc 720

ggggacctgt ccacaccga tgctgtgatg ggtaatccaa aagtaaaagc acacggcaag 780
 aaagtccctcg gcgcgttttag tgatggctctg gccatttgg ataacttgaa gggtacattc 840

 gcgcagcttt ccgaactcca ctgtgacaag ttgcacgtag atccagaaaa cticcggctt 900
 ctgggcaatg tgcttgatg cgttctggct caccattttg ggaaggagt taccacacc 960
 gtgcaagcgg cttacaaaa agtggctgca ggagtggcta atgcccttgc acataaatat 1020
 cactaaggta ccgagcatct taccgccatt tatteccata tttgttctgt ttttcttgat 1080
 ttgggtatac atttaaatgt taataaaaca aaatgggtggg gcaatcattt acatttttag 1140
 ggatatgtaa ttactagtct aggtgtattg ccacaagaca aacatgttaa gaaactttcc 1200
 cgttatttac gcctgttcc tgttaataca cctctggatt acaaaatttg tgaaagattg 1260
 actgatattc ttaactatgt tgctcctttt acgctgtgtg gatatgctgc tttatagcct 1320

 ctgtatctag ctattgcttc ccgtacggct ttctgtttct cctccttgta taaatcctgg 1380
 ttgtgtcttc ttttagagga gttgtggccc gttgtccgtc aacgtggcgt ggtgtgctct 1440
 gtgtttgctg acgcaacccc cactggctgg ggcatgcca ccacctgtca actcctttct 1500
 gggactttcg ctttccccct ccgcatcgcc acggcagaac tcatgccgc ctgccttgc 1560
 cgctgtgga caggggctag gttgtgggc actgataatt ccgtgggtgt gtctgtgcct 1620
 tctagtggc agccatctgt tgtttgccc tccccgtgc cttccttgac cctggaaggt 1680
 gccactccca ctgtccttc ctaataaaat gaggaaattg catcgattg tctgagtagg 1740
 tgtcattcta ttctgggggg tgggtgggg caggacagca agggggagga ttgggaagac 1800

 aatagcaggc atgctgggga tgcgtgggc tctatggctt aattaacgag atcgagacca 1860
 tcctggctaa cacagtgaac cccgtctct actaaaaaa tacaaaaat tagccgggct 1920
 tgggtggcggg tgcctgtagt ccgactact atggagctg aggcgggaga atggcgtgaa 1980
 cgcggggggc ggagcttgca gtgagcagag atcaggggcc actgcactcc agcctgggcg 2040
 acagagagag actctgtctc aaaaaaaga aaaaaaaat ttagtagact agctaaaaa 2100
 atccagagat agttattgat gcataatgaa aagtcttcca atatttaca gtacaatgaa 2160
 aaaaaataa ccttgaatta agtgtagaac tcattgacaa tgtttcaaag gatgtgaggg 2220
 ataaactaaa atttgggcag tacatgctgt tcctgtgtac ttggaacaga gggagaaaat 2280

 ctgggctgga aatattgta taggagttag cacatgaagg tgacaactaa attatttga 2340
 gtagatggag tcaccagcac atgtgaatag ttttagaatg aaatgacca agatagaact 2400
 ttggagagcc ccaaattta aataaaatca gtataagaga agaggaagaa accaaatggt 2460
 atactagtct aaatgtttc ttagtgacaa aagaataacc tgaatattag attagctgcc 2520
 tatatgctct ctgaatcaat ttcatcaac atgcaacagt ccgcgggaac agagaacag 2580

gagaatatgg gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggccaa 2640
gaacagttgg aacagcagaa tatgggccaa acaggatatc tgtggttaagc agttcctgcc 2700
ccggctcagg gccaaagaaca gatggtcccc agatgcggtc ccgccctcag cagtttctag 2760

agaacatca gatgtttcca ggggtcccca aggacctgaa atgacctgt gccttatattg 2820
aactaacaa tcagttcgt tctcgttct gtctgcgcgc ttctgtctcc cgagctctat 2880
ataagcagag ctcgtttagt gaaccgtcag atcgctgga gacgcatcc acgtgtttt 2940
gacttcata gaaggcgcc gcgccccac catggtgagc aaggcgagg agctgttcac 3000
cggggtggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 3060
gtccggcgag ggcgaggcg atgccacct cggcaagctg accctgaagt tcatctgcac 3120
caccggcaag ctgcccgtgc cctggccac cctcgtgacc accctgacct acggcgtgca 3180
gtgcttcagc cgctaccccg accacatgaa gcagcacgac ttcttcaagt ccgcatgcc 3240

cgaaggctac gtccaggagc gcaccatctt ctcaaggac gacggcaact acaagaccg 3300
cgccgaggtg aagttcagg gcgacacct ggtgaaccgc atcgagctga agggcatcga 3360
ctcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccaaa 3420
cgtctatatac atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgca 3480
caacatcgag gacggcagcg tgcagctgc cgaccactac cagcagaaca ccccatcgg 3540
cgacggcccc gtgtgtgtgc ccgacaacca ctacctgagc acccagtcg cctgagcaa 3600
agacccaac gagaagcgc atcacatggt cctgctggag ttcgtgaccg ccgccgggat 3660
cactctcggc atggacgagc tgtacaagta agctttattt gtgaaattg tgatgctatt 3720

gctttatttg taaccattat aagctgcaat aaacaagtta acaacaaca ttgcattcat 3780
tttatgttc aggttcagg ggagatgtgg gaggttttt aaagccctgc aggaggagaa 3840
gtctgccgtt actgcctgt ggggcaaggt gaacgtggat gaagtgtgtg gtgaggccct 3900
gggcaggttg gtatcaaggt tacaagacag gtttaaggag accaatagaa actgggcatg 3960
tgagacaga gaagactctt gggtttctga taggcactga ctctctctgc ctattggtct 4020
atttccac ccttagctg ctggtgtct acccttgac ccagaggttc tttgagtcct 4080
ttggggatct gtccactct gatgtgtta tggcaaccc taaggtgaag gctcatggca 4140
agaaagtgt cgggtgcctt agtgatggc tggtcacct ggacaacct aaggcacct 4200

ttgccacact gagtgagctg cactgtgaca agcgtcgac tagataagta gcatggcggg 4260
ttaatcatta actacaagga acccctagt atggagtgg ccactccctc tctgcgcgt 4320
cgctcgctca ctgaggccgg gcgaccaaag gtgcgccgac gccggggtt tgccgggcg 4380
gcctcagtga gcgagcagc gcgccagct gcgtaatagc gaagaggccc gcaccgatcg 4440

cccttcccaa cagttagcgca gcctgaatgg cgaatggcga ttccgttgca atggctggcg 4500
glaatatgtt tctggatatt accagcaagg ccgatatgtt gagttcttct actcaggcaa 4560
gtgatgttat tactaatcaa agaagtattg cgacaacggt taatttgcgt gatggacaga 4620
ctcttttact cggtaggcctc actgattata aaaacacttc tcaggattct ggcgtagcgt 4680

tcctgtctaa aatcccttta atcggcctcc tgtttagctc ccgctctgat tctaacgagg 4740
aaagcacgtt atacgtgctc gtcaaagcaa ccatagtacg cgcctcttag cggcgcat 4800
agcgcggcgg gtgtgggtgt tacgcgcagc gtgaccgcta cacttgccag cgccttagcg 4860
cccgtctctt tcgctttctt ccttctcttt ctgcacagct tcgcccgttt tccccgtcaa 4920
gtcttaaate gggggctccc tttagggctc cgatttagtg ctttacggca cctcgacccc 4980
aaaaaacttg attaggggtga tgggtcacgt agtgggcat cgcctgata gacgggtttt 5040
cgcccttga cgttggagtc cagttctttt aatagtggac tcttggtcca aactggaaca 5100
acactcaacc ctatctcgtt ctattctttt gatttataag ggattttgcc gatttcggcc 5160

tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattttaa caaaatatta 5220
acgtttacaa tttaaatatt tgcttataca atcttctgt ttttggggtt tttctgatta 5280
tcaaccgggg tacatatgat tgacatgcta gttttacgat taccgttcat cgattctctt 5340
gtttgtcca gactctcagg caatgacctg atagccttg tagagacctc tcaaaaatag 5400
ctacctctc cggcatgaat ttatcagcta gaacgggtga atatcatatt gatggtgatt 5460
tgactgtctc cggcctttct caccgtttg aatctttacc tacacattac tcaggcattg 5520
catttaaaat atatgagggt tctaaaaatt ttatccttg cgttgaaata aaggcttctc 5580
ccgcaaaagt attacagggt cataatgttt ttgtacaac cgatttagct ttatgtctgt 5640

aggctttatt gcttaatttt gctaattctt tgccttgctt gtatgattta ttggatgttg 5700
gaatgcctg atcggtattt ttctcttac gcatctgtgc ggtatttcac accgcatatg 5760
gtgactctc agtacaatct gctctgatgc cgcatagtta agccagcccc gacaccgccc 5820
aacaccgct gacgcgcctt gacgggcttg tctgctcccg gcatccgtt acagacaagc 5880
tgtgaccgtc tccgggagct gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc 5940
gagacgaaag ggctctgta tacgcctatt ttataggtt aatgcatga taataatggt 6000
ttcttagacg tcagggtgga cttttcgggg aaatgtgcgc ggaacccta tttgtttatt 6060
tttctaaata cattcaata tgatccgct catgagacaa taaccctgat aaatgcttca 6120

ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgcct ttattccctt 6180
ttttcgga ttttgcctc ctgtttttgc tcaccagaa acgtgggtga aagtaaaaga 6240
tgctgaagat cagtgggtg cagagtggg ttacatcgaa ctggatctca acagcgtaa 6300

gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagtctt 6360
gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg gtcgccgcat 6420
acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga 6480
tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata acactgcggc 6540
caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 6600

gggggatcat gtaactcgcc ttgatcggtg ggaaccggag ctgaatgaag ccataccaaa 6660
cgacgagcgt gacaccagga tgccgttagc aatggcaaca acgttgcgca aactattaac 6720
tggcgaacta cttactctag cttcccgga acaattaata gactggatgg aggcggataa 6780
agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc 6840
tggagccggt gagcgtgggt ctgcgggtat cattgcagca ctggggccag atggtaagcc 6900
ctcccgatc gtagttatct acacgacggg gactcaggca actatggatg aacgaaatag 6960
acagatcgt gagatagggt cctcactgat taagcattgg taactgtcag accaagtta 7020
ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa 7080

gatccttttt gataatctca tgacaaaaat cccttaacgt gagttttcgt tccactgagc 7140
gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat ctttttttc tgcgcgtaat 7200
ctgctgcttg caaacaaaaa aaccaccgt accagcgggt gtttgtttgc cggatcaaga 7260
gtaccaact cttttccga aggttaactgg cttcagcaga gcgcagatac caaatactgt 7320
ccttctagt tagccgtagt taggccacca cttcaagaac tctglagcac gcctacata 7380
cctcgtctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac 7440
cgggttggac tcaagacgat agttaccgga taaggcgag cggtcgggt gaacgggggg 7500
ttcgtgcaca cagcccagct tggagcgaac gacctacac gaactgagat acctacagc 7560

tgagctatga gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccgtaag 7620
cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaacg cctggtatct 7680
ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 7740
agggggcgag agcctatgga aaaacgccag caacgcggcc tttttacgt tcttggcctt 7800
ttgtggcct ttgtctcaca tgtttttcc tgcgttatcc cctgattctg tggataaccg 7860
tattaccgc tttgagttag ctgataccgc tcgccgagc cgaacgaccg agcgagcga 7920
gtcagtgagc gaggaagcgg aagagcgccc aatagcaaa ccgcctctcc ccgcgcttg 7980
gccgattcat taatg 7995

<211> 7995

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1290

<400> 33

cagctgcgcg ctgcctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 ttgtgtgcc cgccctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
 actaggggtt ccttgtagtt aatgattaac ccgcatgct acttatctac acgcgtagat 180
 ctagtgcac aacttcttat ttgtgtaata agaaaattgg gaaaacgac ttcaatatgc 240
 ttaccaagct gtgattccaa atattacgta aatacattg caaaggagga tgtttttagt 300
 agcaatttgt actgatgga tggggccaag agatatatct tagaggagg gctgagggtt 360

tgaagtccaa ctctaagcc agtgccagaa gagccaagga caggtacggc tgcatcact 420
 tagacctcac cctgtggagc cacaccctag ggttggccaa tctactcca ggagcaggga 480
 gggcaggagc cagggtctgg cataaaagtc agggcagagc catctattgc ttacatttgc 540
 ttctgacaca actgtgttca ctagcaacct caaacagaca ccatggtgca tctgactcct 600
 gaggagaagt ccgtgtaac ggcactgtgg gggaaagtta atgtcgatga agtcggcgg 660
 gaagcactcg gcaggttgct ggtagtgtac ccgtggacac aacgattctt tgaaagtttc 720
 ggggacctgt ccacaccga tgctgtgatg ggtaatccaa aagtaaaagc acacggcaag 780
 aaagtccctg gcgcgtttag tgatggtctg gccatttgg ataacttgaa gggtacattc 840

gcgcagcttt ccgaactcca ctgtgacaag ttgcacgtag atccagaaaa ctccggctt 900
 ctgggcaatg tgcttgtatg cgttctggct caccattttg ggaaggagt ttacccacc 960
 gtgcaagcgg cttacaaaa agtggctgca ggagtggcta atgcccttgc acataaatat 1020
 cactaagga ccgagcatct taccgccatt tattccata ttgttctgt ttttcttgat 1080
 ttgggtatac atttaaatgt taataaaaca aaatgggtgg gcaatcattt acatttttag 1140
 ggatatgtaa ttactagtcc aggtgtattg ccacaagaca aacatgttaa gaaactttcc 1200
 cgttatattac gctctgttcc tgtaaatcaa cctctggatt acaaaatttg tgaaagattg 1260
 actgatattc ttaactatgt tgctcctttt acgctgtgtg gatagctgc tttatagcct 1320

ctgtatctag ctattgcttc ccgtacggct ttctgtttct cctccttgta taaatcctgg 1380
 ttgtgtcttc ttttagagga gttgtggccc gttgtccgtc aacgtggcgt ggtgtgctct 1440
 gtgtttgctg acgcaacccc cactggctgg ggcaattgcca ccactgtca actcctttct 1500
 gggactttcg ctttcccct cccgatcgcc acggcagaac tcatcgccgc ctgccttggc 1560

cgctgctgga caggggctag gttgctgggc actgataatt ccgtgggtgtt gtctgtgcct 1620
tctagtgtcc agccatctgt tgtttgcccc tccccctgac cttccttgac cctggaaggt 1680
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgattg tctgagtagg 1740
tgtcattcta ttctgggggg tgggggtgggg caggacagca agggggagga ttgggaagac 1800

aatagcaggc atgctgggga tgcggtgggc tctatggctt aattaacgag atcgagacca 1860
tcctggctaa cacagtgaac ccccgctctt actaaaaaaa tacaaaaaat tagccgggct 1920
tggtggcggg tgctgtagt cccagctact atggaggctg aggcgggaga atggcgtgaa 1980
cgccgggggc ggagcttgca gtgagcagag atcaggggcc actgcactcc agcctgggag 2040
acagagagag actctgtctc aaaaaaaga aaaaaaaat ttagtagact agctaaaaaa 2100
atccagagat agttattgat gcatatgtaa aagtcttcca atatttaca gtacaatgaa 2160
aaaaaataa ccttgaatta agtgtagaac tcattgacaa tgtttcaaag gatgtgaggg 2220
ataaactaaa atttgggcag tacatgctgt tcctgtgtac ttggaacaga gggagaaaat 2280

ctgggctgga aatattgtta taggagttag cacatgaagg tgacaactaa attatttga 2340
gtagatggag tcaccagcac atgtgaatag ttttagaatg aaatgacca agatagaact 2400
ttggagagcc cccaaattta aataaaatca gtataagaga agaggaagaa accaaatggt 2460
atactagtct aaattgtttc ttagtgacaa aagaataacc tgaatattag attagctgcc 2520
tatatgctct ctgaatcaat ttcatccaac atgcaacagt ccgcggaac agagaaacag 2580
gagaatatgg gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggccaa 2640
gaacagttag aacagcagaa tatgggccaa acaggatata tgtgtaagc agttcctgcc 2700
ccggctcagg gccaaagaac gatggtcccc agatgcggtc ccgcctcag cagtttctag 2760

agaaccatca gatgtttcca ggggtcccca aggacctgaa atgacctgt gccttatttg 2820
aactaacca tcagttcgct tctcgcttct gttcgcgcg cttctgctcc cgagctctat 2880
ataagcagag ctgctttagt gaaccgtcag atcgctgga gacgccatcc acgtgtttt 2940
gacttcata gaaggcgcc gcgccccac catggtgagc aaggcgagg agctgttcac 3000
cggggtgggt cccatcctgg tcgagctgga cggcgacgta aacggccaca agttcagcgt 3060
gtccggcgag ggcgaggggc atgccacct cggcaagctg accctgaagt tcactgcac 3120
caccggaag ctgccccg cctggccac cctcgtgacc accctgacct acggcgtgca 3180
gtgcttcagc cgtaccccc accacatgaa gcagcacgac ttcttcaagt ccgcatgcc 3240

cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagaccg 3300
cgccgaggtg aagtctgagg gcgacacct ggtgaaccgc atcgagctga agggcatcga 3360
cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca acagccaca 3420

cgtctatatac atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca 3480
 caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca ccccatcg 3540
 cgacggcccc gtgctgctgc ccgacaacca ctacctgagc acccagtcgg cctgagcaa 3600
 agacccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg ccgccgggat 3660
 cactctcggc atggacgagc tgtacaagta agctttatct gtgaaatttg tgatgctatt 3720

gctttatctg taaccattat aagctgcaat aaacaagtta acaacaacaa ttgcattcat 3780
 tttatgtttc aggttcaggg ggagatgtgg gaggtttttt aaagccctgc aggaggagaa 3840
 gtctgaggtt actgccctgt ggggcaaggt gaacgtggat gaagtgtgtg gtgaggccct 3900
 gggcagggtt gtatcaaggt tacaagacag gtttaaggag accaatagaa actgggcatg 3960
 tggagacaga gaagactctt gggtttctga taggcactga ctctctctgc ctattggtct 4020
 attttccac ccttaggctg ctggtggtct acccttggac ccagaggttc tttgagtcct 4080
 ttggggatct gtccactcct gatgctgtta tgggcaaccc taagtggaag gctcatggca 4140
 agaaagtgtc cgggtgcctt agtgatggcc tggtcacct ggacaacctc aagggcacct 4200

ttgccacact gactgagctg cactgtgaca agcgtcgacg tagataagta gcatggcggg 4260
 ttaatcatta actacaagga acccctagtg atggagtgtg ccactccctc tctgcgcgt 4320
 cgctcgctca ctgaggccgg gcgaccaaag gtcgcccagc gcccggtt tgccccggcg 4380
 gcctcagtga gcgagcgagc gcgcagctg gcgtaatagc gaagaggccc gcaccgatcg 4440
 ccttcccaa cagttgcga gcctgaatgg cgaatggcga ttccgttgca atggctggcg 4500
 gtaatatgtt tctggatatt accagcaagg ccgatatgtt gattttctt actcaggcaa 4560
 gtgatgttat tactaatcaa agaagtattg cgacaacggt taatttgctg gatggacaga 4620
 ctcttttact cgggtgcctc actgattata aaaacacttc tcaggattct ggcgtaccgt 4680

tcctgtctaa aatcccttta atcgccctcc tgtttagctc ccgtctgat tctaacgagg 4740
 aaagcacgtt atacgtgctc gtcaaagcaa ccatagtacg cgccctgtag cggcgctta 4800
 agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg 4860
 cccgtcctt tcgttttctt ccttccctt ctcgccagct tcgccggtt tccccgtcaa 4920
 gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc 4980
 aaaaaacttg attagggtga tggttcacgt agtgggcat cgccctgata gacggttttt 5040
 cgccctttga cgttgagtc cacgttctt aatagtgga cttgtttca aactggaaca 5100
 acactcaacc ctatctcggc ctattctttt gatttataag ggattttgcc gatttcggcc 5160

tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattttta caaatatta 5220
 acgtttacaa tttaaatatt tgcttataca atcttctgt tttggggct tttctgatta 5280

tcaaccgggg tacatatgat tgacatgcta gttttacgat taccgttcat cgattctctt 5340
gtttgtccca gactctcagg caatgacctg atagcctttg tagagacctc tcaaaaatag 5400
ctacctctc cgcatgaat ttatcagcta gaacggttga atatcatatt gatggtgatt 5460
tgactgtctc cgccctttct caccgctttg aatctttacc tacacattac tcaggcattg 5520
catttaaaat atatgagggt tctaaaaatt ttatccttg cgttgaaata aaggcttctc 5580
ccgcaaaagt attacagggt cataatgttt ttggtacaac cgatttagct ttatgctctg 5640

aggctttatt gcttaatttt gctaattctt tgccttgctt gtatgattta ttggatgttg 5700
gaatcgctg atcggttatt ttctccttac gcatctgtgc ggtatttcac accgcatatg 5760
gtgactctc agtacaatct gctctgatgc cgcatagtta agccagcccc gacaccgccc 5820
aacaccgct gacgcgccct gacgggcttg tctgctccc gcatccgctt acagacaagc 5880
tgtgaccgtc tccgggagct gcatgtgtca gaggttttca ccgtcatcac cgaacgcgc 5940
gagacgaaag ggcctctga tacgcctatt ttataggtt aatgtcatga taataatggt 6000
ttcttagacg tcagggtgca cttttcgggg aaatgtgcgc ggaacccta tttgtttatt 6060
tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca 6120

ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt 6180
ttttgcggca ttttgcttc ctgtttttgc tcaccagaa acgctggtga aagtaaaaga 6240
tgctgaagat cagtgggtg cagcagtggtt ttacatcgaa ctggatctca acagcggtaa 6300
gatccttgag agttttgcc ccgaagaacg tttccaatg atgagcactt ttaaagtctt 6360
gctatgtggc gcggtattat cccgtattga cgcgggcaa gagcaactcg gtcgccgcat 6420
acactattct cagaatgact tggttgagta ctaccagtc acagaaaagc atcttacgga 6480
tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata acactgcggc 6540
caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 6600

gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa 6660
cgacgagcgt gacaccagca tgcctgtagc aatggcaaca acgttgcgca aactattaac 6720
tggcgaacta cttactctag ctcccgga acaattaata gactggatgg aggcggataa 6780
agttgcagga ccacttctgc gctcgccct tccggctggc tggtttattg ctgataaatc 6840
tggagccggt gagcgtgggt ctgcggtat cattgcagca ctggggccag atggtaagcc 6900
ctcccgatc gtagttatct acacgacggg gagtgcaggca actatggatg aacgaaatag 6960
acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta 7020
ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa 7080

gatccttttt gataatctca tgacaaaaat cccttaacgt gagttttcgt tccactgagc 7140

gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat ctttttttc tgcgcgtaat 7200
 ctgctgcttg caaacaaaaa aaccaccgct accagcgggtg gtttgtttgc cggatcaaga 7260
 gctaccaact ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt 7320
 ccttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata 7380
 cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac 7440
 cgggttggac tcaagacgat agttaccgga taaggcgag cggtcgggct gaacgggggg 7500
 ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg 7560

tgagctatga gaaagcgcca cgcttcccga agggagaaaag gcggacaggt atccggttaag 7620
 cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaaac cctggatatc 7680
 ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 7740
 agggggggcg agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt 7800
 ttgctggcct ttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg 7860
 tattaccgcc ttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgacgca 7920
 gtcagtgcg gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg 7980
 gccgattcat taatg 7995

<210> 34

<211> 8036

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1314

<400> 34

cagctgcgcg ctgctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 tttagtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
 actaggggtt cctgcggccg cgattcaaac ttccgcagaa cactttattt cacatataca 180
 tgcctcttat atcagggatg tgaacagggt tcttgaacac tgtctaaatc taaaacaatg 240
 ctaatgcagg tttaaattta ataaaaataa atccaaaatc taacagccaa gtcaaatctg 300
 catgttttaa catttaaat attttaaaga cgtcttttcc caggattcaa catgtgaaat 360

cttttctcag ggatacacgt gtgcctagat cctcattgct ttagtttttt acagaggaat 420
 gaatataaaa agaaaatact taaattttat ccctcttacc tctataatca tacataggca 480
 taatttttta acctaggctc cagatagcca tagaagaacc aaacactttc tgcgtgtgtg 540

agaataatca gaggtagatt ttttcacaag tacctgatga ggggttagac aggtagaaaa 600
 agtgagagat ctctatttat ttagcaataa tagagaaagc atttaagaga ataaagcaat 660
 ggaaataaga aatttataaa tttctttctg ataactagaa atagaggatc cagttttctt 720
 tggtaacct aaattttatt tcatatttat gttttatatt atttttattt atttttattt 780
 gtgtaatcgt agtttcagag tgtagagct gaaaggaaga agtaggagaa acatgcaaag 840

taaaagtata acactttcct tactaaaccg acatgggttt ccaggtaggg gcaggattca 900
 ggatgactga cagggccctt aggggaacact gagaccctac gctgacctca taaatgcttg 960
 ctacctttgc tgttttaatt acatctttta atagcaggaa gcagaactct gcacttcaa 1020
 agtttttctt cacttgagga gtttaatttag tacaaggga aaaagtacag ggggatggga 1080
 gaaaggcgt cagtttgga agctatagag aaagaagagt aaattttagt aaaggagggt 1140
 taaacaaaca aaatataaag agaaatagga acttgaatca aggaatgat tttaaaacgc 1200
 agtattctta gtggactaga ggaaaaaat aatctgagcc aagtagaaga cttttcccc 1260
 tcctaccctt actttctaag tcacagaggc tttttgtcc ccagacact cttgcagatt 1320

agtcaggca gaaacagtta gatgtccca gttaacctcc tatttgacac cactgattac 1380
 cccattgata gtcacacttt ggggttgaag tgacttttta tttatttgta ttttgactg 1440
 cattaagagg tctctagttt tttatctctt gtttccaaa acctaataag taactaatgc 1500
 acagagcaca ttgatttgta tttattctat ttttagacat aatttattag catgcatgag 1560
 caaattaaga aaaacaaca caaatgaatg catatatatg tatatgtatg tgtgtatata 1620
 tacacacata tatatatata tttttcttt tcttaccaga aggttttaac ccaaataagg 1680
 agaagatatg cttagaaccg aggtagagtt ttcattccatt ctgtcctgta agtattttgc 1740
 atattctgga gacgcaggaa gagatccatc tacatatccc aaagctgaat tatggtagac 1800

aaaactcttc cacttttagt gcatcaactt cttatttttg taataagaaa attgggaaaa 1860
 cgatcttcaa tatgcttacc aagctgtgat tccaaatatt acgtaaatac acttgcaaag 1920
 gaggatgttt ttagtagcaa tttgtactga tggtagggg ccaagagata tatcttagag 1980
 ggagggtga gggtttgaag tccaactcct aagccagtgc cagaagagcc aaggacaggt 2040
 acggctgtca tcacttagac ctcacctgt ggagccacac ctaggggttg gccaatctac 2100
 tcccaggagc agggaggga ggagccagg ctgggcataa aagtcagggc agagccatct 2160
 attgcttaca ttgcttctg acacaactgt gttcactagc aacctcaaac agacaccatg 2220
 gtgcatctga ctctgtcga gaagtctgcc gttactgccc tgtggggcaa ggtgaacgtg 2280

gatgaagttg gtggtgagc cctgggcagg ttggtatcaa ggttacaaga caggtttaag 2340
 gagaccaata gaaactgggc atgtggagac agagaagact cttgggtttc tgataggcac 2400

tgactctctc tgcctattgg tctatittcc cacccttagg ctgctgggtg tctacccttg 2460
gaccagagg ttctttgagt cctttgggga tctgtccact cctgatgctg ttatgggcaa 2520
ccctaagggtg aaggctcatg gcaagaaagt gctcgggtgcc tttagttagt gcctggctca 2580
cctggacaac ctcaaggga cctttgccac actgagttag ctgcactgtg acaagctgca 2640
cgtggatcct gagaacttca gggtagtct atgggacgt tgatgttttc tttcccttc 2700
ttttctatgg ttaagttcat gtcataggaa ggggataagt aacagggtac agtttagaat 2760

gggaaacaga cgaatgattg catcagtgtg gaagtctcag gatcgtttta gtttctttta 2820
tttgctgttc ataacaattg ttttcttttg ttttaattctt gctttctttt ttttcttct 2880
ccgcaatfff tactattata cttaatgcct taacattgtg tataacaaaa ggaaatatct 2940
ctgagatata ttaagtaact taaaaaaaaa ctttacacag tctgcctagt acattactat 3000
ttggaatata tgtgtgctta ttgcatatt cataatctcc ctactttatt ttcttttatt 3060
tttaattgat acataatcat tatacatatt tatgggttaa agtgtaatgt tttaatatgt 3120
gtacacatat tgaccaaac agggttaatt tgcatttgta attttaaaaa atgctttctt 3180
cttttaatat actttttgt ttatcttatt tctaactt tcctaactct ctttctttca 3240

gggcaataat gatacaatgt atcatgcctc ttgcacat tctaaagaat aacagtata 3300
atttctgggt taaggcaata gcaatatctc tgcataaaa ttttctgca tataaattgt 3360
aactgatgta agaggtttca tattgctaag agcagctaca atccagctac cattctgctt 3420
ttattttatg gtgggataa ggctggatta ttctgagtc aagctaggcc cttttgctaa 3480
tcatgttcat acctcttacc ttctccac agctcctggg caacgtgctg gtctgtgtgc 3540
tgccccatca cttggcaaa gaattcacc caccagtga ggctgcctat cagaaagtgg 3600
tggtggtgt ggctaagcc ctggccaca agtatcacta agctcgctt cttgtgtcc 3660
aatttctatt aaaggttct ttgtcccta agtccaacta ctaactggg ggatattatg 3720

aagggccttg agcatctgga ttctgcctaa taaaaacat ttattttcat tgcaatgatg 3780
tatttaaatt atttctgaat attttactaa aaagggaatg tgggaggtca gtgcatttaa 3840
aacataaaga aatgaagagc tagttcaaac cttgggaaaa tacactatat cttaactcc 3900
atgaaagaag gtgaggctgc aaacagctaa tgcacattgg caacagcccc tgatgcata 3960
gccttattca tccctcagaa aaggattcaa gtagaggctt gatttggagg ttaaagtttt 4020
gctatgctgt attttacatt acttattgtt ttagctgtcc tcatgaatgt ctttacta 4080
cccatttct taccctgcat ctctcagct tgactccact cagtctctt gcttagagat 4140
accaccttc cctgaagtg ttcttccat gttttacggc gagatggttt ctcctgcct 4200

ggccactcag ccttagttgt ctctgtgtc ttatagaggt ctactgaag aaggaaaaac 4260

aggggtcatg gtttgactgt cctgtgagcc cttcttccct gatcggatcc cctgcaggag 4320
gaacccttag tgatggagtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc 4380
gggcgaccaa aggtcgcccg acgcccgggc ttgcccggg cggcctcagt gagcgagcga 4440
gcgcgcagct ggcgtaatag cgaagaggcc cgcaccgatc gcccttccca acagttgcgc 4500
agcctgaatg gcgaatggcg attccgttgc aatggctggc ggtaatatg ttctggatat 4560
taccagcaag gccgatagtt tgagtcttc tactcaggca agtgatgtta ttactaatca 4620
aagaagtatt gcgacaacgg ttaatttgcg tgatggacag actcttttac tcggtggcct 4680

cactgattat aaaaacactt ctcaggattc tggcgtaccg ttctgtctta aaatcccttt 4740
aatcggcctc ctgttttagt cccgctctga ttctaacgag gaaagcacgt tatacgtgct 4800
cgtcaaagca accatagtac gcgccctgta gcggcgcatc aagcgcggcg ggtgtggtgg 4860
ttacgcgcag cgtgaccgt acacttgcca gcgccctagc gcccgctcct ttcgctttct 4920
tcccttccct tctcgccacg ttccgcccgt ttccccgtca agctctaaat cgggggctcc 4980
ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaaactt gattagggtg 5040
atggttcacg tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttggagt 5100
ccacgttctt taatagtga ctctgttcc aaactggaac aacactcaac cctatctcgg 5160

tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 5220
tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca atttaaatat 5280
ttgcttatac aatcttctg tttttggggc ttttctgatt atcaaccggg gtacatatga 5340
ttgacatgct agttttacga ttaccgttca tcgattctct tgtttgctcc agactctcag 5400
gcaatgacct gatagccttt gtagagacct ctcaaaaata gctaccctct ccggcatgaa 5460
tttatcagct agaacggttg aatatcatat tgatggtgat ttgactgtct ccggcctttc 5520
tcaccgcttt gaatctttac ctacacatta ctacggcatt gcatttaaaa tatatgaggg 5580
ttctaaaaat ttttatcctt gcgttgaaat aaaggcttct cccgcaaaag tattacaggg 5640

tcataatggt tttggtacaa ccgatttagc tttatgctct gaggtttat tgcttaattt 5700
tgctaattct ttgccttgcc tgtatgattt attggatgtt ggaatgcct gatgcggtat 5760
tttctccta cgcatctgtg cggtatttca caccgcatat ggtgcactct cagtacaatc 5820
tgctctgatg ccgcatagtt aagccagccc cgacacccgc caacacccgc tgacgcgccc 5880
tgacgggctt gtctgtccc ggcatccgt tacagacaag ctgtgaccgt ctccgggagc 5940
tgcatgtgtc agaggttttc accgtcatca ccgaaacgcg cgagacgaaa gggcctcgtg 6000
atacgcctat ttttataggt taatgtcatg ataataatgg tttcttagac gtcaggtggc 6060
acttttcggg gaaatgtgcg cggaaccctt atttgtttat ttttctaaat acattcaaat 6120

atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag 6180
 agtatgagta ttcaacattt ccgtgtcgcc ctattccct ttttgcggc attttgcctt 6240
 cctgtttttg ctcaccaga aacgttggtg aaagtaaaag atgctgaaga tcagttgggt 6300
 gcacgagtg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc 6360
 cccgaagaac gttttccaat gatgagcact tttaaagtgc tgctatgtgg cgcggtatta 6420
 tcccgtattg acgccgggca agagcaactc ggtcgccgca tacactatc tcagaatgac 6480
 ttggttgagt atcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa 6540
 ttatgcagt ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg 6600

atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc 6660
 ctgatcggtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg 6720
 atgcctgtag caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta 6780
 gcttcccgcc aacaattaat agactggatg gaggcggata aagttgcagg accacttctg 6840
 cgctcgccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg 6900
 tctcgcggtg tcatlgcagc actggggcca gatggtgaag cctcccgat cgtagtattc 6960
 tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt 7020
 gcctcactga ttaagcattg gtaactgtca gaccaagttt actcataat acttttagatt 7080

gatttaaaac ttcatTTTTA atttaaaagg atctaggtga agatcctttt tgataatctc 7140
 atgacaaaaa tcccttaacg tgagttttcg ttccactgag cgtcagacc cgtagaaaag 7200
 atcaaaggat cttcttgaga tcttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa 7260
 aaaccaccgc taccagcggg ggtttgtttg ccggatcaag agctaccaac tctttttccg 7320
 aaggttaactg gcttcagcag agcgcagata ccaatactg tcttctagt gtacccgtag 7380
 ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaactctg 7440
 ttaccagtgg ctgctgccag tggcgataag tctgtcttta ccgggttgga ctcaagacga 7500
 tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc 7560

ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc 7620
 acgttccccg aaggagagaa ggccgacagg tatccggtaa gcgcgagggt cggaacagga 7680
 gagcgcagca gggagcttcc aggggggaaac gcctgggtatc tttatagtcc tgtcgggttt 7740
 cgccacctct gacttgagcg tcgatttttg tgatgctcgt cagggggcg gagcctatgg 7800
 aaaaacgcca gcaacgcggc ctttttacgg ttcttgccct tttgctggcc ttttgctcac 7860
 atgttctttc ctgcgttate cctgtattct gtggataacc gtattaccgc ctttgagtga 7920
 gctgataccg ctgcgccag ccgaacgacc gagcgcagcg agtcagttag cgaggaagcg 7980

gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatg 8036

<210> 35

<211> 8207

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1321

<400> 35

cagctgcgcg ctgcctcgct cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 ttgtgtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
 actaggggtt cctacgcgta gatctagtta gatgtcccca gttaacctcc tatattgacac 180
 cactgattac cccattgata gtcacacttt gggttgtaag tgacttttta tttatttgta 240
 tttttgactg cattaagagg tctctagttt tttatctctt gtttcccaaa acctaataag 300
 taactaatgc acagagcaca ttgatttgta tttattctat ttttagacat aatttattag 360

catgcatgag caaattaaga aaaacaacaa caaatgaatg catatatatg tatatgtatg 420
 tgtgtatata tacacacata tatatatata ttttttcttt tcttaccaga aggttttaac 480
 ccaaataagg agaagatatg cttagaaccg aggtagagtt ttcattccatt ctgtcctgta 540
 agtatattgc atattctgga gacgcaggaa gagatccatc tacatatccc aaagctgaat 600
 tatggtagac aaaactcttc cacttttagt gcatcaactt cttatttggtg taataagaaa 660
 attgggaaaa cgatcttcaa tatgcttacc aagctgtgat tccaaatatt acgtaaatac 720
 acttgcaaag gaggatgttt ttagtagcaa tttgtactga tggtagggg ccaagagata 780
 tatcttagag ggagggctga gggtttgaag tccaactcct aagccagtgc cagaagagcc 840

aaggacaggt acggctgtca tcacttagac ctccacctgt ggagccacac cctagggttg 900
 gccaatctac tcccaggagc agggagggca ggagccaggg ctgggcataa aagtcagggc 960
 agagccatct attgcttaca tttgcttctg acacaactgt gttcactagc aacctcaaac 1020
 agacaccatg gtgcatctga ctcttaagc tttaaaaaac ctccacatc tccccctgaa 1080
 cctgaaacat aaaatgaatg caattgttgt tgtaacttg tttattgcag cttataatgg 1140
 ttacaaataa agcaatagca tcacaaattt cacaaataaa gcttacttgt acagctcgtc 1200
 catgccgaga gtgatccggg cggcggtcac gaactccagc aggaccatgt gatcgcgctt 1260
 ctggttgggg tctttgctca gggcggtactg ggtgctcagg tagtggttgt cgggcagcag 1320

cacggggccg tcgccgatgg ggggtgttctg ctggtagtgg tcggcgagct gcacgctgcc 1380
 gtcctcgatg ttgtggcgga tcttgaagtt caccttgatg cgttcttctt gcttgtcggc 1440
 catgatatag acgttgtggc tgtttagtgt gtactccagc ttgtgccccg ggatgttgcc 1500
 gtcctccttg aagtcgatgc ccttcagctc gatgcggttc accagggtgt cgccctcgaa 1560
 cttcacctcg gcgcgggtct tgtagtggcc gtcgtccttg aagaagatgg tgcgctcctg 1620
 gacgtagcct tcgggcatgg cggacttgaa gaagtcgtgc tgcttcatgt ggtcggggta 1680
 gcggctgaag cactgcacgc cgtaggtcag ggtggtcacg aggttgggcc agggcacggg 1740
 cagcttgccg gtggtgcaga tgaacttcag ggtcagcttg ccgtaggtgg catcgccctc 1800

gccctcgccg gacacgctga acttgtggcc gtttacgtcg ccgtccagct cgaccaggat 1860
 gggcaccacc ccggtgaaca gtcctcggcc ctgtctcacc atggtggcgg gcgcggcccg 1920
 atctgacggt tcaactaaac agctctgctt atatagagct cggggagcag aagcgcgca 1980
 acagaagcga gaagcgaact gattgggttag ttcaaataag gcacagggtc atttcaggtc 2040
 cttggggcac cctggaaaca tctgatggtt ctctagaaac tgctgagggc gggaccgcat 2100
 ctggggacca tctgttcttg gccctgagcc ggggcaggaa ctgcttacca cagatactct 2160
 gtttggcccc tattctgtcg ttccaactgt tcttggccct gagccggggc aggaactgct 2220
 taccacagat atcctgtttg gcccatattc tctgtttct ctgttccgc ggcgagatcg 2280

agaccatcct ggtaacaca gtgaaacccc gtctctacta aaaaaataca aaaaattagc 2340
 cgggcttggg gcgggtgcc tgtagtccca gctactatgg aggtgaggc gggagaatgg 2400
 cgtgaacgcg gggggcggag cttgcagtga gcagagatca ggggccactg cactccagcc 2460
 tgggcgacag agagagactc tgtctcaaaa aaaagaaaaa aaaaatttag tagactagct 2520
 aaaaaaatcc agagatagtt attgatgcat atgtaaaagt cttcaatat ttacaagtac 2580
 aatgaaaaaa aaataacctt gaattaagt tagaactcat tgacaatgtt tcaaaggatg 2640
 tgagggataa actaaaattt gggcagtaca tctgttctct gtgtacttgg aacagaggga 2700
 gaaaatctgg gctgaaata ttgttatagg agttagcaca tgaaggtgac aactaaatta 2760

tttggagtag atggagtcac cagcacatgt gaatagtttt agaagaaat gacccaagat 2820
 agaactttgg agagccccc aatttaaata aaatcagtat aagagaagag gaagaaacca 2880
 aatggtatag tagtctaat tgtttcttag tgacaaaaga ataacctgaa tattagatta 2940
 gctgcctata tgetctctga atcaatttca ttcaacatgc aacagtcttg gaacctatca 3000
 gggaccacag tcagccaggc aagcacatct gcccaagcca aggttggagg catgcagctg 3060
 tgggggtctg tgaaaacact tgaggagca gataactggg ccaacctga ctcagtgttt 3120
 ctggaggcca acaggactgc tgagtcattc tgtgggggtg gaggtgggac aagggaagg 3180

ggtgaatggt actgctgatt acaacctctg gtgctgcctc cccctcctgt ttatctgaga 3240

 gaggcctcac tggagctaga gacaagaagg taaaaaacgg ctgacaaaag aagtcctggt 3300
 atcctctatg atgggagaag gaaactagct aaagggaaga ataaattaga gaaaaactgg 3360
 aatgactgaa tcggaacaag gcaaaggcta taaaaaaat tagcagtatc ctcttggggg 3420
 ccccttcccc acactatctc aatgcaata tctgtctgaa acggtccctg gctaaactcc 3480
 acccatgggt tggccagcct tgccttgaca aggcaactt gaccaatagt cttagagtat 3540
 ccagtgaggc caggggccgg cggctggcta gggatgaaga ataaaaggaa gcacccttca 3600
 gcagtccac acactcgctt ctggaacgtc tgaggttatc aataagctcc tagtcagac 3660
 gccatggtcc atcttacacc ggtggagaag tctgccgtta ctgccctgtg gggcaagggtg 3720

 aacgtggatg aagttggtgg tgaggccctg ggcaggttgg tatcaaggtt acaagacagg 3780
 ttttaaggaga ccaatagaaa ctgggcatgt ggagacagag aagactcttg ggtttctgat 3840
 aggcactgac tctctctgcc tatttgctta ttttccacc cttaggctgc tgggtgtcta 3900
 cccttggacc cagaggttct ttgagtcctt tggggatctg tccactcctg atgctgttat 3960
 gggcaacctt aaggtgaagg ctcatggcaa gaaagtgtc ggtgccttta gtgatggcct 4020
 ggctcacctg gacaacctca agggcacctt tgccacactg agtgagctgc actgtgacaa 4080
 gctgcacgtg gatcctgaga acttcagggt gagtctatgg gacgcttgat gttttctttc 4140
 ccttctttt ctatggttaa gtatcatgtca taggaagggg ataagtaaca gggtagagt 4200

 tagaatggga aacagacgaa tgattgcac agtgtggaag tctcaggatc gttttagttt 4260
 cttttatttg ctgttataa caattgttt cttttgtta attcttgctt tcttttttt 4320
 tcttctccg aatttttact attatactta atgccttaac atttgtata acaaaaggaa 4380
 atatctctga gatacattaa gtaacttaaa aaaaaacttt acacagtctg cctagtacat 4440
 tactatttgg aatataatgt tgcttatttg catattcata atcgctgaca ggaaccccta 4500
 tgatggagt tggccactcc ctctctgcgc gctcgctgc tctactgaggc cgggcgacca 4560
 aaggtcgccc gacgcccggt ctttccccgg gcggcctcag tgagcgagcg agcgcgcagc 4620
 tggcgttaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg cagcctgaat 4680

 ggcaatggc gattccgttg caatggctgg cgtaaatatt gttctggata ttaccagcaa 4740
 ggccgatagt ttgattctt ctactcaggc aagtgatgtt attactaatc aaagaagtat 4800
 tgcgacaacg gttaatttgc gtgatggaca gactctttta ctcggtggcc tctactgatta 4860
 taaaaacact tctcaggatt ctggcgtaac gttcctgtct aaaatccctt taatcggcct 4920
 cctgttttag tcccgtctg attctaacga ggaaagcacg ttatagctgc tcgtcaaagc 4980
 aaccatagta cgcgccctgt agcggcgcat taagcgcggc ggggtgtgtg gttacgcgca 5040

gcgtagccgc tacacttgcc agcgccctag cgcccgtcc tttcgcttc tccccttct 5100
 ttctcgccac gttagccggc tttccccgtc aagctctaaa tcgggggctc ctttagggt 5160

tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttcac 5220
 glagtgggac atcgccctga tagacgggtt ttcccccctt gacgttgag tccacgttct 5280
 ttaatagtgg actcttgctc caaactggaa caacactcaa ccctatctcg gtctattctt 5340
 ttgatttata agggattttg ccgatttcgg cctattgggt aaaaaatgag ctgatttaac 5400
 aaaaatttaa cggaatttt aacaaaatat taacgtttac aatttaaata ttgcttata 5460
 caatcttctt gtttttggg cttttctgat tatcaaccgg ggtacatatg attgacatgc 5520
 tagttttacg attaccgttc atcgattctc ttgtttgctc cagactctca ggcaatgacc 5580
 tgatagcctt ttagagacc tctcaaaaat agctaccctc tccggcatga atttatcagc 5640

tagaacggtt gaatatcata ttgatggatga ttgactgtc tccggccttt ctcaccggtt 5700
 tgaatcttta cctacacatt actcaggcat tgcatttaaa atatatgagg gttctaaaaa 5760
 tttttatcct tgcgttgaaa taaaggcttc tcccgaataa gtattacagg gtcataatgt 5820
 ttttggtaca accgatttag ctttatgtc tgaggcttta ttgcttaatt ttgctaattc 5880
 ttgacctgac ctgtatgatt tattggatgt tggaatcgcc tgatgcggtt ttttctcctt 5940
 acgcatctgt gcggtatttc acaccgata tgggtgactc tcagtacaat ctgctctgat 6000
 gccgcatagt taagccagcc ccgacaccgc ccaacaccgc ctgacgcgcc ctgacgggct 6060
 tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt 6120

cagaggtttt caccgtcatc accgaaacgc gcgagacgaa agggcctcgt gatacgccta 6180
 tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcaggtgg cacttttcgg 6240
 ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg 6300
 ctcatgagac aataacctg ataatgtct caataatatt gaaaaaggaa gagtatgagt 6360
 attcaacatt tccgtgtcgc cttatttccc ttttttgcgg cttttgcct tctgtttt 6420
 gctcaccag aaacgtggt gaaagtataa gatgctgaag atcagttggg tgcacgagt 6480
 ggttacatcg aactgcatc caacagcggg aagatccttg agagttttcg ccccgaagaa 6540
 cgttttccaa tgatgagcac ttttaaagt ctgctatgtg gcgcggtatt atccgtatt 6600

gacgccgggc aagagcaact cggtagccgc atacactatt ctcagaatga cttaggttag 6660
 tactcaccag tcacagaaaa gcactttacg gatggcatga cagtaagaga attatgcagt 6720
 gctgccataa ccatgagatg taacactgac gccaaactac ttctgacaac gatcggagga 6780
 ccgaaggagc taaccgttt tttgcacaac atgggggac atgtaactcg cttgatcgt 6840
 tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta 6900

gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg 6960
caacaattaa tagactggat ggagcgcat aaagttgcag gaccatttct gcgctcggcc 7020
cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt 7080

atcattgcag cactggggcc agatggtaag ccctcccgtg tcgtagtat ctacacgacg 7140
gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg 7200
attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa 7260
cttcattttt aatttaaaag gatctagggtg aagatccttt ttgataatct catgaccaa 7320
atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga 7380
tcttcttgag atcctttttt tctgcgcgtg atctgctgct tgcaaacaaa aaaaccaccg 7440
ctaccagcgg tggttttgtt gccggatcaa gagctacca ctccttttcc gaaggtaact 7500
ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta gttagccac 7560

cacttcaaga actctgtagc accgcctaca tacctcgtc tgctaactct gttaccagtg 7620
gtgctgccg gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg 7680
gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga 7740
acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc 7800
gaaggagaga aggcggacag gtatccggtg agcggcaggg tcggaacagg agagcgcacg 7860
agggagcttc cagggggaaa cgcttggtat ctttatagtc ctgtcgggtt tcgccacctc 7920
tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc 7980
agcaacgcgg cttttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt 8040

cctgcgttat cccctgattc tgttgataac cgtattaccg cctttgagtg agctgatacc 8100
gtcgcgccga gccgaacgac cgagcgcagc gattcagtga gcgaggaagc ggaagagcgc 8160
ccaatacgca aaccgcctct ccccgcgctg tggccgattc attaatg 8207

<210> 36

<211> 8009

<212> DNA

<213> Artificial Sequence

<220>

<223> template 1322

<400> 36

cagctgcgcg ctgcctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
tttggtcgcc cggcctcagt gagegagcga gcgcgcagag agggagtggc caactccatc 120

actaggggtt cctacgcgta gatctagtta gatgtcccca gttaacctcc tatttgacac 180

cactgattac ccattgata gtcacacttt gggttgtaag tgacttttta tttatttgta 240

tttttgactg cattaagagg tctctagttt tttatctctt gtttccaaa acctaataag 300

taactaatgc acagagcaca ttgatttgta tttattctat ttttagacat aatttattag 360

catgcatgag caaattaaga aaaacaacaa caaatgaatg catatatatg tatatgtatg 420

tgtgtatata tacacacata tatatatata ttttttcttt tcttaccaga aggttttaat 480

ccaaataagg agaagatatg cttagaaccg aggtagagtt ttcattccatt ctgtcctgta 540

agtattttgc atattctgga gacgcaggaa gagatccatc tacatatccc aaagctgaat 600

tatggtagac aaaactcttc cacttttagt gcatcaactt cttatttttg taataagaaa 660

attgggaaaa cgatcttcaa tatgcttacc aagctgtgat tccaaatatt acgtaaatac 720

acttgcaaaag gaggatgttt ttagtagcaa ttgtactga tggatgggg ccaagagata 780

tatcttagag ggagggctga gggtttgaa tccaactcct aagccagtgc cagaagagcc 840

aaggacaggt gctttaaaaa acctcccaca tctccccctg aacctgaaac ataaaatgaa 900

tgcaattgtt gttgttaact tgtttattgc agcttataat ggttacaaat aaagcaatag 960

catcacaaat ttcacaaata aagcttactt gtacagctcg tccatgccga gagtgatccc 1020

ggcggcggtc acgaactcca gcaggacat gtgatcgcg tctctgttgg ggtctttgct 1080

cagggcggtc tgggtgtcga ggtagtggtt gtcgggcagc agcacggggc cgtcgccgat 1140

gggggtgttc tgctggtagt ggtcggcgag ctgcacgctg ccgtcctcga tgttgtggcg 1200

gatcttgaag ttcacctga tgccgttctt ctgcttctcg gccatgatat agacgttgtg 1260

gctgtttag ttgtactcca gcttctgccc caggatgttg ccgtcctcct tgaagtcgat 1320

gcccttcagc tcgatcggtt tcaccagggt gtgcacctcg aacttcacct cggcgcggtt 1380

ctttagttg ccgtcgtcct tgaagaagat ggtgcgtcc tggacgtagc cttcgggcat 1440

ggcggacttg aagaagtctg gctgcttcat gtggtcgggg tagcggctga agcactgcac 1500

gccgtaggtc aggtgtgtca cgagggtggg ccagggcacg ggcagcttgc cgggtgtgca 1560

gatgaacttc aggttcagct tgccgtaggt ggcatgccc tcgccctcgc cggacacgct 1620

gaacttgttg ccgtttacgt cgccgtccag ctgcaccagg atgggcacca ccccggtgaa 1680

cagctcctcg ccttctga ccatgttggt ggcgcgcccg cgatctgacg gttcactaaa 1740

cgagctctgc ttatatagag ctcggggagc agaagcgcg gaacagaagc gagaagcgaa 1800

ctgattggtt agttcaaata aggcacaggg tcatttcagg tccttggggc accctggaaa 1860

catctgatgg ttctctagaa actgctgagg gcgggaccgc atctggggac catctgttct 1920

tggccctgag ccggggcagg aactgcttac cacagatc ctgtttggcc catattctgc 1980

tgttccaact gticttggcc ctgagccggg gcaggaactg cttaccacag atatcctgtt 2040
 tggcccatat tctcctgttt ctctgttccc gcggcgagat cgagaccatc ctggctaaca 2100

 cagtgaacc ccgtctctac taaaaaata caaaaaatta gccgggcttg gtggcgggtg 2160
 cctgtagtcc cagctactat ggaggctgag gcgggagaat ggctgaacg cggggggcgg 2220
 agcttgacgt gagcagagat cagggggccac tgcactccag cctgggcgac agagagagac 2280
 tctgtctcaa aaaaaagaaa aaaaaattt agtagactag ctaaaaaaat ccagagatag 2340
 ttattgatgc atatgtaaaa gtcttccaat atttacaagt acaatgaaaa aaaaataacc 2400
 ttgaattaag tgtagaactc attgacaatg tttcaaagga tgtgagggat aaactaaaat 2460
 ttgggcagta catgctgttc ctgtgtactt ggaacagagg gagaaaatct gggctggaaa 2520
 tattgttata ggagttagca catgaagggtg acaactaaat tatttggagt agatggagtc 2580

 accagcacat gtgaatagtt ttagaatgaa atgaccaag atagaacttt ggagagcccc 2640
 caaatttaaa taaatcagt ataagagaag aggaagaaac caaatggtat actagtctaa 2700
 attgtttctt agtgacaaaa gaataacctg aatattagat tagctgccta tatgtctctt 2760
 gaatcaattt cattcaacat gcaacagttc tggaacctat cagggaccac agtcagccag 2820
 gcaagcacat ctgccaagc caagggtgga ggcatgcagc tgtgggggtc tgtgaaaaca 2880
 cttgaggag cagataactg ggccaacct gactcagtgc ttctggagc caacaggact 2940
 gctgagtcac cctgtggggg tggagggtggg acaagggaag ggggtgaatg gtactgtga 3000
 ttacaacctc tgggtgtgcc tccccctct gtttatctga gagaggcctc actggagcta 3060

 gagacaagaa ggtaaaaaac ggctgacaaa agaagtcctg gtatcctcta tgatgggaga 3120
 aggaaactag ctaaagggaa gaataaatta gaaaaaact ggaatgactg aatcggaaca 3180
 aggcaaagc tataaaaaa attagcagta tctcttggg ggcccttcc ccacactatc 3240
 tcaatgcaaa tatctgtctg aaacggtecc tggctaaact ccacccatgg gttggccagc 3300
 cttgccttga caaggcaaac ttgaccaata gtcttagagt atccagttag gccaggggcc 3360
 ggcggtggc tagggatgaa gaataaaagg aagcaccctt cagcagttcc acacactcgc 3420
 ttctggaacg tctgaggtta tcaataagct cctagtcag acgcatggt ccatcttaca 3480
 ccggtggaga agtctgccgt tactgccctg tggggcaagg tgaacgtgga tgaagtgggt 3540

 ggtgaggccc tgggcaggtt ggtatcaagg ttacaagaca ggtttaagga gaccaataga 3600
 aactgggcat gtggagacag agaagactct tgggtttctg ataggcactg actctctctg 3660
 cctattggtc tattttccca cccttaggct gctggtggtc tacccttga cccagagggt 3720
 ctttgagtc tttggggatc tgtccactcc tgatgtgtt atgggcaacc ctaagtgaa 3780
 ggctcatggc aagaaagtgc tcggtgcctt tagtgatggc ctggctcacc tggacaacct 3840

caagggcacc ttigccacac tgagttagct gcactgtgac aagctgcacg tggatcctga 3900
gaacttcagg gtgagtctat gggacgcttg atgttttctt tccccttctt ttctatggtt 3960
aagttcatgt cataggaagg ggataagtaa cagggtacag tttagaatgg gaaacagacg 4020

aatgattgca tcagtgtgga agtctcagga tcgttttagt tttttttatt tgctgttcat 4080
aacaattggt tttttttgtt taattcttgc tttttttttt tttcttctcc gcaattttta 4140
ctattatact taatgcctta acattgtgta taacaaaagg aaatatctct gagatacatt 4200
aagtaactta aaaaaaaact ttacacagtc tgcctagtag attactatgt ggaatatatg 4260
tgtgcttatt tgcatattca taatcgtcga cagggaacccc tagtgatgga gttggccact 4320
ccctctctgc gcgtctgctc gctcactgag gccgggagac caaaggtcgc ccgacgcccc 4380
ggctttgccc gggcgccctc agtgagcgag cgagcgcgca gctggcgtaa tagcgaagag 4440
gccccaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg gcgattccgt 4500

tgcaatggct ggcggtaata ttgttctgga tattaccagc aaggccgata gtttgagtgc 4560
ttctactcag gcaagtgatg ttattactaa tcaaagaagt attgcgaca cggttaattt 4620
gcgtgatgga cagactcttt tactcggtgg cctcactgat tataaaaaca cttctcagga 4680
ttctggcgta ccgttccgtg ctaaaatccc tttaatcggc ctctgttta gctcccgtc 4740
tgattctaac gaggaagca cgttatacgt gctcgtcaaa gcaaccatag tacgcgccct 4800
gtagcggcgc attaagcgcg gcgggtgtgg tggttacgag cagcgtgacc gctacacttg 4860
ccagcgccct agcgcccgct ctttccgtt tcttcccttc ctttctcgcc acgttcgccc 4920
gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac 4980

ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccacgcctc 5040
gatagacggt ttttcgccc ttgacgttgg agtccagtt ctttaatatg ggactcttgt 5100
tccaaactgg aacaacactc aacctatct cggctctatt ttttgattta taagggattt 5160
tgccgatttc ggctatttgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt 5220
ttaacaaaat attaacgttt acaatttaaa tatttgctta tacaatcttc ctgtttttgg 5280
ggcttttctg attatacaac ggggtacata tgattgacat gctagtttta cgattaccgt 5340
tcacgattc tctgtttgc tccagactct caggcaatga cctgatagcc tttgtagaga 5400
cctctcaaaa atagctaccc tctccggcat gaatttatca gctagaacgg ttgaatatca 5460

tattgatggt gatttgactg tctccggcct tttcaccgc tttgaatctt tacctacaca 5520
ttactcaggc attgcattta aaatatatga gggttctaaa aatttttatc cttgcgttga 5580
aataaaggct tctcccgcaa aagtattaca gggtcataat gtttttggt caaccgattt 5640
agctttatgc tctgaggctt tattgcttaa ttttgcta tctttgcctt gcctgtatga 5700

tttattggat gttggaatcg cctgatgcgg tattttctcc ttacgcatct gtgcggtatt 5760
 tcacaccgca tatggtgcac tctcagtaca atctgctctg atgccgcata gttaagccag 5820
 ccccgacacc cgccaacacc cgctgacgcg ccttgacggg cttgtctgct cccggcatcc 5880
 gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagagggtt ttcacgtca 5940

tcaccgaaac gcgcgagacg aaaggcctc gtgatacgcc ttttttata ggttaatgtc 6000
 atgataataa tggtttctta gacgtcaggt ggcacttttc ggggaaatgt gcgcggaacc 6060
 cctatttggtt tatttttcta aatacattca aatatgtatc cgctcatgag acaataaccc 6120
 tgataaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca tttccgtgtc 6180
 gcccttattc cttttttgc ggcattttgc cttcctgttt ttgctcacc agaacgctg 6240
 gtgaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat cgaactggat 6300
 ctcaacagcg gtaagatcct tgagagtttt cggcccgaag aacgttttcc aatgatgagc 6360
 acttttaag ttctgctatg tggcgcggta ttatccgta ttgacgccgg gcaagagcaa 6420

ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtactcacc agtcacagaa 6480
 aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgcat aaccatgagt 6540
 gataaactg cgccaactt acttctgaca acgatcggag gaccgaagga gctaaccgct 6600
 tttttgcaca acatggggga tcatgtaact cgccttgatc gttgggaacc ggagctgaat 6660
 gaagccatac caaacgacga gcgtgacacc acgatgcctg tagcaatggc aacaacgttg 6720
 cgcaactat taactggcga actacttact ctgcttccc ggcaacaatt aatagactgg 6780
 atggaggcgg ataaagtgc aggaccactt ctgcgctcgg cccttcggc tggttggttt 6840
 attgctgata aatctggagc cggtagcgt gggtctcgcg gtatcattgc agcactgggg 6900

ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca ggcaactatg 6960
 gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca ttggtaactg 7020
 tcagaccaag ttactcata tatactttag attgatttaa aacttcattt ttaatttaa 7080
 aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt 7140
 tcgttccact gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt 7200
 tttctgcgcg taatctgctg cttgcaaaca aaaaaccac cgctaccagc ggtggtttgt 7260
 ttgccggatc aagagctacc aactctttt ccgaaggtaa ctggcttcag cagagcgag 7320
 ataccaata ctgtccttct agttagccg tagttaggcc accacttcaa gaactctgta 7380

gcaccgccta catacctcgc tctgctaate ctgttaccag tggctgctgc cagtggcgat 7440
 aagtcgtgtc ttaccgggtt ggactcaaga cgatagtac cggataaggc gcagcggtcg 7500
 ggctgaacgg ggggttcgtg cacacagccc agcttgagc gaacgaccta caccgaactg 7560

agatacctac agcgtgagct atgagaaagc gccacgcttc ccgaagggag aaaggcggac 7620
 aggtatccgg taagcggcag ggtcggaca ggagagcgca cgaggagct tccaggggga 7680
 aacgcctggt atctttatag tctgtcggg ttccgccacc tctgacttga gcgtcgattt 7740
 ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta 7800
 cggttcctgg ccttttgctg gccttttgct cacatgttct ttcctgcgtt atcccctgat 7860

tctgtggata accgtattac cgcctttgag tgagctgata ccgctcgccg cagccgaacg 7920
 accgagcgca gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg caaacgcct 7980
 ctccccgcgc gttggccgat tcattaatg 8009

<210> 37

<211> 1606

<212> DNA

<213> Artificial Sequence

<220>

<223> gene sequence of HBB hemoglobin subunit beta [Homo sapiens (human)] Gene ID: 3043 on NCBI (updated on 21-Apr-2018), located at 11p15.4 of the human genome. The sequence starts at the beginning of

Exon 1. The sequence is identified on the NCBI database as: NC_000011.10:c5227071-5225466 Homo sapiens chromosome 11, GRCh38.p12 Primary Assembly

<400> 37

acatttgctt ctgacacaac tgtgttcact agcaacctca aacagacacc atggtgcatc 60
 tgactcctga ggagaagtct gccgttactg cctgtgagg caagtgaaac gtggatgaag 120
 ttggtggtga ggccctgggc aggttggtat caaggttaca agacaggttt aaggagacca 180
 atagaaactg ggcatgtgga gacagagaag actcttgggt ttctgatagg cactgactct 240
 ctctgcctat tggctatatt tccaccctt aggtctctgg tggctaccc ttggaccag 300

aggttctttg agtcctttgg ggatctgtcc actcctgatg ctgttatggg caaccetaag 360
 gtgaaggctc atggcaagaa agtgctcggg gccttttagt atggcctggc tcacctggac 420
 aacctcaagg gcacctttgc cacactgagt gagctgcact gtgacaagct gcacgtggat 480
 cctgagaact tcagggtgag tctatgggac gcttgatgtt ttctttcccc ttcttttcta 540
 tggttaagtt catgtcatag gaaggggata agtaacaggg tacagtttag aatgggaaac 600

agacgaatga ttgcatcagt gtggaagtct caggatcggt ttagtttctt ttatttgctg 660
 ttcataacaa ttgttttctt ttgtttaatt ctgctttct tttttttct tctccgcaat 720
 ttttactatt atacttaatg ccttaacatt gtgtataaca aaaggaaata tctctgagat 780

acattaagta acttaaaaaa aaactttaca cagtctgcct agtacattac tatttggaat 840
 atatgtgtgc ttatttgcat attcataatc tccctacttt attttctttt atttttaatt 900
 gatacataat cattatacat atttatgggt taaagtgtaa tgttttaata tgtgtacaca 960
 tattgaccaa atcagggttaa ttttgcatth gtaattttta aaaatgcttt ctcttttta 1020
 tatacttttt tgtttatctt atttctaata ctttccttaa tctctttctt tcagggcaat 1080
 aatgatacaa tgiatcatgc ctctttgcac cattctaaag aataacagtg ataatttctg 1140
 ggtaaggca atagcaatat ctctgcataa aaatatttct gcatataaat tgtaactgat 1200
 gtaagagggt tcatattgct aatagcagct acaatccagc taccattctg cttttattht 1260

atggttggga taaggctgga ttattctgag tccaagctag gcccttttgc taatcatggt 1320
 catactctt atcttctcc cacagctctt gggcaacgtg ctggtctgtg tgctggccca 1380
 tcactttggc aaagaattca ccccaccagt gcaggctgcc tatcagaaag tgggtgctgg 1440
 tgtggcta at gccctggccc acaagtatca ctaagctgc tttcttgctg tccaatttct 1500
 attaaagggt ctttgttcc ctaagtccaa ctactaaact gggggatatt atgaagggcc 1560
 ttgagcatct ggattctgcc taataaaaaa catttattht cattgc 1606

<210> 38

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 38

gtcgagaagt ctgcagtcac tgctctatgg gggaaa 36

<210> 39

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> HBB-1250
 <400> 39
 aggcctttttg ttccccaga 20

<210> 40
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> HBB-1250
 <400> 40
 agccttcacc ttagggttgc 20

<210> 41
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 >
 <223> SCL-386
 <400> 41
 gggttggcca atctactccc 20

<210> 42
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> SCL-386
 <400> 42
 cctctgggtc caagggtaga 20

<210> 43
 <211> 19
 <212> DNA

<213> Artificial Sequence	
<220>	
<223> ddPCR	
<400> 43	
cataaaagtc agggcagag	19
<210> 44	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> ddPCR	
<400> 44	
gtctccttaa acctgtcttg	20
<210> 45	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> LINC01206	
<400> 45	
caaaaagcaa aatttgggga ta	22
<210> 46	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> LINC01206	
<400> 46	
cttttagccc agtgccagac	20
<210> 47	

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> MIR7974

<400> 47

atcagcccct ctttctggat

20

<210> 48

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> MIR7974

<400> 48

agtgcagtgg tgccatcata

20

<210> 49

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> HBD

<400> 49

cagatcccca aaggactcaa

20

<210> 50

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> HBD

<400> 50

gcggtgggga gatatgtaga	20
<210> 51	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> TULP4	
<400> 51	
cacgccagga tgtaagctct	20
<210> 52	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> TULP4	
<400> 52	
tctgaggcaa aagtgaaga	20
<210> 53	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
>	
<223> DENND3	
<400> 53	
gggggtttct atccctcact	20
<210> 54	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	

<223> DENND3	
<400> 54	
caagagggtc aggttgagga	20
<210> 55	
<211> 53	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> HBB-Miseq (with adapter)	
<400> 55	
tcgtcggcag cgtcagatgt gtataagaga caggggttgg ccaatctact ccc	53
<210> 56	
<211> 54	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> HBB-Miseq (with adapter)	
<400> 56	
gtctcgtggg ctcgagatg tgtataagag acagcctctg ggtccaaggg taga	54
<210> 57	
<211> 55	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> HBD-Miseq (with adapter)	
<400> 57	
tcgtcggcag cgtcagatgt gtataagaga cagcacaac taatgaaacc ctgct	55
<210> 58	
<211> 55	
<212> DNA	

<213> Artificial Sequence

<220>

<223> HBD-Miseq (with adapter)

<400> 58

gtctcgtggg ctcgagatg tgtataagag acagtctaca catgcccagt ttcca 55

<210>

59

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> GTC HDR FAM

<400> 59

ctcctgtcga gaagtctgc 19

<210> 60

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> GAA HDR FAM

<400> 60

ctcccgaaga gaagtctgc 19

<210> 61

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> GTG HDR FAM

<400> 61

ctcctgtgga gaagtctgc 19

<210> 62
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> GAG WT HEX
 <400> 62
 tgactcctgt cgagaagt 18

<210> 63
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> REF HEX
 <400> 63
 gttcactagc aacctcaaac agacacc 27

<210> 64
 <211> 59
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> E6V GAG > GTC
 <400> 64
 tcagggcaga gccatctatt gcttacattt gcttctgaca caactgtgtt cactagcaa 59

<210> 65
 <211> 59
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> E6V GAG > GTC
 <400> 65

cctcaaacag acaccatggt gcatctgact cctgtcgaga agtctgccgt tactgcct 59

<210> 66

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > GTC

<400> 66

gtggggcaag gtgaacgtgg atgaagttgg tggtagggcc ctgggcaggt 50

<210> 67

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > CCC GAA

<400> 67

tcagggcaga gccatctatt gcttacattt gcttctgaca caactgtgtt cactagcaa 59

<210> 68

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > CCC GAA

<400> 68

cctcaaacag acaccatggt gcatctgact cccgaagaga agtctgccgt tactgcct 59

<210> 69

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > CCC GAA

<400> 69

gtggggcaag gtgaacgtgg atgaagttgg tggtagggcc ctgggcaggt 50

<210> 70

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > GTG

<400> 70

tcagggcaga gccatctatt gcttacattt gcttctgaca caactgtgtt cactagcaa 59

<210> 71

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > GTG

<400> 71

cctcaaacag acaccatggt gcacttgact cctgtggaga agtctgccgt tactgcct 59

<210> 72

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> E6V GAG > GTG

<400> 72

gtggggcaag gtgaacgtgg atgaagttgg tggtagggcc ctgggcaggt 50

<210> 73

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 73

caacctcaaa cagacaccat ggtgcatctg actcctgagg agaagtctgc cgttactgcc 60

ctgtggggca aggtg

75

<210> 74

<211> 72

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 74

caacctcaaa cagacaccat ggtgcacctg actcctgaga agtctgccgt tactgccctg 60

tggggcaagg tg

72

<210> 75

<211> 70

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 75

caacctcaaa cagacaccat ggtgcacctg actcctgaag tctgccgtta ctgccctgtg 60

gggcaaggtg

70

<210> 76

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 76

caacctcaaa cagacaccat ggtgcacctg actcctgccg ttactgccct gtggggcaag 60
gtg 63

<210> 77

<211> 69

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 77

caacctcaaa cagacaccat ggtgcacctg actcctgagt ctgccgttac tgcctgtgg 60
ggcaagggtg 69

<210> 78

<211> 72

<

212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 78

caacctcaaa cagacaccat ggtgcacctg actcctgaga agtctgccgt tactgccctg 60
tggggcaagg tg 72

<210> 79

<211> 74

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 79

caacctcaaa cagacaccat ggtgcacctg actcctgaga gaagtctgcc gttactgccc 60
tgtggggcaa ggtg 74

<210> 80

<211> 72

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 80

caacctcaaa cagacaccat ggtgcacctg actcctgaga agtctgccgt tactgccctg 60
tggggcaagg tg 72

<210> 81

<211> 74

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 81

caacctcaaa cagacaccat ggtgcacctg actcctgaga gaagtctgcc gttactgccc 60
tgtggggcaa ggtg 74