(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0017098
(43) 공개일자 2024년02월06일

(51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) *A61P 7/00* (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/113 (2013.01)
A61K 48/005 (2013.01)
(21) 출원번호 10-2024-7002710(분할)
(22) 출원일자(국제) 2018년02월05일
심사청구일자 없음
(62) 원출원 특허 10-2019-7026089
원출원일자(국제) 2018년02월05일
심사청구일자 2021년02월04일
(85) 번역문제출일자 2024년01월23일
(86) 국제출원번호 PCT/IB2018/050712
(87) 국제공개번호 WO 2018/142364
국제공개일자 2018년08월09일
(30) 우선권주장
62/455,464 2017년02월06일 미국(US)

(71) 출원인
노파르티스 아게
스위스 체하-4056 바젤 리히트스트라쎄 35
인텔리아 테라퓨틱스, 인크.
미국, 매사추세츠 02139, 캠브릿지, 이리 스트리트 40
(72) 발명자
미카닌, 크레이그 스티븐
미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 250 노바티스 인스티튜츠 포 바이오메디칼 리서치, 인크. 내
슈메트, 크리스티안
미국 92121 캘리포니아주 샌디에고 존 제이 홉킨스 드라이브 10675 디비에이 더 지노믹스 인스티튜트 오브 더 노바티스 리서치 파운데이션 (지엔에프) 노바티스 인스티튜트 포 펑셔날 지노믹스, 인크. 내
(뒷면에 계속)
(74) 대리인
양영준, 이상영

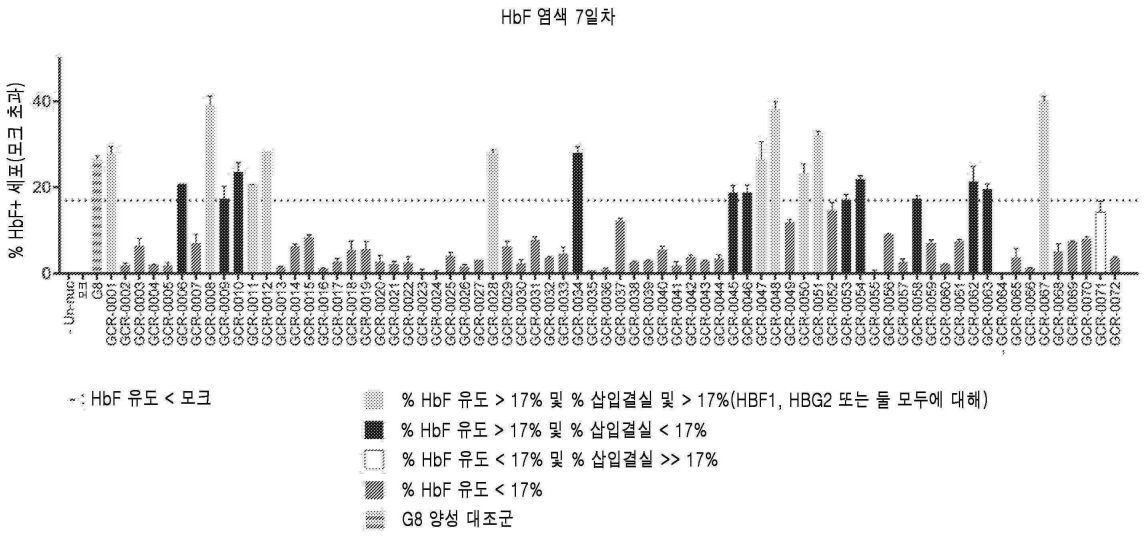
전체 청구항 수 : 총 92 항

(54) 발명의 명칭 **혈색소병증의 치료를 위한 조성물 및 방법**

(57) 요 약

본 발명은 혈색소병증의 치료를 위한 게놈 편집 시스템, 시약 및 방법에 관한 것이다.

대 표 도



(52) CPC특허분류
A61P 7/00 (2018.01)
C12N 9/22 (2013.01)
C12N 2310/20 (2017.05)
(72) 발명자
스니드, 제니퍼
미국 92121 캘리포니아주 샌디에고 존 제이 홉킨스 드라이브 10675 디비에이 더 지노믹스 인스티튜트 오브 더 노바티스 리서치 파운데이션 (지엔에프) 노바티스 인스티튜트 포 펑셔날 지노믹스, 인크. 내
스티븐슨, 수잔 씨.
미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 250 노바티스 인스티튜츠 포 바이오메디칼 리서치, 인크. 내
양, 이
미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 250 노바티스 인스티튜츠 포 바이오메디칼 리서치, 인크. 내

명 세 서

청구범위

청구항 1

tracr 및 crRNA를 포함하는 gRNA 분자로서, 여기서, crRNA는

a) Chr11:5,250,094 내지 5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

b) Chr11:5,255,022 내지 5,255,164, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

c) 비결실(nondeletional) HFPH 영역(예를 들어 인간 비결실 HPFH 영역)의 표적 서열에 상보적인 표적화 도메인;

d) Chr11:5,249,833 내지 Chr11:5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

e) Chr11:5,254,738 내지 Chr11:5,255,164, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

f) Chr11:5,249,833 내지 5,249,927, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

g) Chr11:5,254,738 내지 5,254,851, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

h) Chr11:5,250,139 내지 5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인; 또는

i) 이들의 조합

을 포함하는, gRNA 분자.

청구항 2

제1항에 있어서, 표적화 도메인이 SEQ ID NO: 1 내지 SEQ ID NO: 72 또는 이들의 단편 중 임의의 하나를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 3

제2항 있어서, 표적화 도메인이 SEQ ID NO: 67, SEQ ID NO: 1, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 28, SEQ ID NO: 34, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 58, SEQ ID NO: 62, SEQ ID NO: 63 또는 이들의 단편 중 임의의 하나를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 4

제2항에 있어서, 표적화 도메인이

a) SEQ ID NO: 67, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 28, SEQ ID NO: 34, SEQ ID NO: 48, SEQ ID NO: 51 또는 이들의 단편; 또는

b) SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 54 또는 이들의 단편

중 임의의 하나를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 5

제2항 내지 제4항 중 어느 한 항에 있어서, 표적화 도메인이 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 6

제5항에 있어서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산이 나열된 표적화 도메인 서열의 3' 말단에 배치된 17, 18, 19 또는 20개 연속 핵산인, gRNA 분자.

청구항 7

제5항에 있어서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산이 나열된 표적화 도메인 서열의 5' 말단에 배치된 17, 18, 19 또는 20개 연속 핵산인, gRNA 분자.

청구항 8

제5항에 있어서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산이 나열된 표적화 도메인 서열의 5' 또는 3' 핵산 중 어느 것도 포함하지 않는, gRNA 분자.

청구항 9

제2항 내지 제8항 중 어느 한 항에 있어서, 표적화 도메인이 나열된 표적화 도메인 서열로 구성되는, gRNA 분자.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 표적화 도메인이 SEQ ID NO: 67을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, gRNA 분자가 이중 가이드 RNA 분자인, gRNA 분자.

청구항 12

제1항 내지 제10항 중 어느 한 항에 있어서, gRNA 분자가 단일 가이드 RNA 분자인, gRNA 분자.

청구항 13

제12항에 있어서,

(a) SEQ ID NO: 195;

(b) SEQ ID NO: 231; 또는

(c) 상기 (a) 또는 (b) 중 어느 하나를 포함하고, 3' 말단에서 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 추가로 포함하며; 임의의 (a) 내지 (c)의 서열이 표적화 도메인에 대해 3', 선택적으로 바로(immediately) 3'에 배치되는, gRNA 분자.

청구항 14

제1항에 있어서, 서열:

(a) SEQ ID NO: 174;

(b) SEQ ID NO: 175; 또는

(c) SEQ ID NO: 176

을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 15

제1항에 있어서,

(a) SEQ ID NO: 177을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

(b) SEQ ID NO: 177을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

(c) SEQ ID NO: 178을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

(d) SEQ ID NO: 178을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서,

a) gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 삽입결실(indel)이 gRNA 분자의 표적화 도메인에 상보적인 표적 서열에서 또는 그 부근에서 형성되고/되거나;

b) gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, HBG1 프로모터 영역에서 gRNA 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 gRNA 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열을 포함하는, 예를 들어 실질적으로 모든 서열을 포함하는 결실이 생성되는, gRNA 분자.

청구항 17

제16항에 있어서, 삽입결실이 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않고, 선택적으로 삽입결실이 비결실 HPFH 또는 전사 인자 결합 부위의 뉴클레오타이드를 포함하지 않는, gRNA 분자.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내로 도입될 때, 집단의 세포 중 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 40%, 예를 들어 적어도 약 50%, 예를 들어 적어도 약 55%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 75%에서 gRNA 분자의 표적화 도메인에 상보적인 표적 서열에서 또는 그 부근에서 삽입결실이 형성되는, gRNA 분자.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포(예를 들어 세포 집단) 내에 도입될 때,

(a) 태아 헤모글로빈의 발현이 상기 세포 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손에서 증가되며, 선택적으로 태아 헤모글로빈의 상기 발현이, gRNA 분자가 도입되지 않은 세포 집단, 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손의 집단에서의 태아 헤모글로빈의 발현 수준에 비해, 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 25%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 35%, 예를 들어 적어도 약 40%만큼 증가되며;

(b) 상기 세포 또는 세포 집단, 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손이 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하며;

(c) 예를 들어 차세대 시퀀싱(next generation sequencing) 및/또는 뉴클레오타이드 삽입 검정법(nucleotide

insertional assay)에 의해 검출 가능한 바와 같이 어떠한 표적-외(off-target) 삽입결실도 상기 세포에서 형성되지 않으며, 예를 들어 어떠한 표적-외 삽입결실도 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 형성되지 않고/않거나;

(d) 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 어떠한 표적-외 삽입결실도 세포 집단의 세포 중 약 5% 초과, 예를 들어 약 1% 초과, 예를 들어 약 0.1% 초과, 예를 들어 약 0.01% 초과에서 검출되지 않는, gRNA 분자.

청구항 20

제16항 내지 제19항 중 어느 한 항에 있어서, 세포가 포유류, 영장류 또는 인간 세포이며(또는 세포 집단이 이를 포함하며), 예를 들어 인간 세포이며, 선택적으로 상기 세포가 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아를 앓고 있는 환자로부터 수득되는, gRNA 분자.

청구항 21

제20항에 있어서, 세포가 HSPC, 선택적으로 CD34+ HSPC, 선택적으로 CD34+CD90+ HSPC인(또는 세포 집단이 이를 포함하는), gRNA 분자.

청구항 22

제16항 내지 제21항 중 어느 한 항에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가(autologous) 또는 동종이계(allogeneic)인, gRNA 분자.

청구항 23

조성물로서,

1) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함) 및 Cas9 분자;

2) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함) 및 Cas9 분자를 인코딩하는 핵산;

3) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함)를 인코딩하는 핵산 및 Cas9 분자;

4) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함)를 인코딩하는 핵산 및 Cas9 분자를 인코딩하는 핵산; 또는

5) 임의의 상기 1) 내지 4), 및 주형 핵산; 또는

6) 임의의 상기 1) 내지 4), 및 주형 핵산을 인코딩하는 서열을 포함하는 핵산

을 포함하는, 조성물.

청구항 24

제1항 내지 제22항 중 어느 한 항의 제1 gRNA 분자를 포함하며, Cas9 분자를 추가로 포함하는 조성물로서, 선택적으로 여기서, Cas9 분자는 활성 또는 불활성 *s. 피오게네스*(*s. pyogenes*) Cas9이며, 선택적으로 Cas9 분자는 SEQ ID NO: 205, 또는 SEQ ID NO: 205에 적어도 95%, 96%, 97%, 98% 또는 99% 서열 상동성(sequence homology)을 갖는 서열을 포함하는, 조성물.

청구항 25

제23항 또는 제24항에 있어서, Cas9 분자가

(a) SEQ ID NO: 233;

(b) SEQ ID NO: 234;

(c) SEQ ID NO: 235;

(d) SEQ ID NO: 236;

(e) SEQ ID NO: 237;

(f) SEQ ID NO: 238;

(g) SEQ ID NO: 239;

(h) SEQ ID NO: 240;

(i) SEQ ID NO: 241;

(j) SEQ ID NO: 242;

(k) SEQ ID NO: 243; 또는

(l) SEQ ID NO: 244

를 포함하는, 예를 들어 이로 구성되는, 조성물.

청구항 26

제23항 내지 제25항 중 어느 한 항에 있어서, 제1 gRNA 분자 및 Cas9 분자가 리보핵 단백질 복합체(RNP; ribonuclear protein complex)에 존재하는, 조성물.

청구항 27

제23항 내지 제26항 중 어느 한 항에 있어서, 전기천공에 적합한 배지 내에서 제제화되는, 조성물.

청구항 28

제23항 내지 제27항 중 어느 한 항에 있어서, 각각의 상기 gRNA 분자가 본원에 기재된 Cas9 분자와 함께 RNP에 존재하고, 여기서, 각각의 상기 RNP가 약 10 uM 미만, 예를 들어 약 3 uM 미만, 예를 들어 약 1 uM 미만, 예를 들어 약 0.5 uM 미만, 예를 들어 약 0.3 uM 미만, 예를 들어 약 0.1 uM 미만의 농도로 존재하고, 선택적으로 상기 RNP의 농도가 약 2 uM이거나 약 1 uM이며, 선택적으로, 조성물이 세포, 예를 들어 HSPC의 집단을 추가로 포함하는, 조성물.

청구항 29

제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자를 인코딩하는 핵산 서열.

청구항 30

제29항의 핵산을 포함하는 벡터로서, 선택적으로 상기 벡터가 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노-연관 바이러스(AAV) 벡터, 단순 포진 바이러스(HSV) 벡터, 플라스미드, 미니서클(minicircle), 나노플라스미드 및 RNA 벡터로 구성된 군으로부터 선택되는, 벡터.

청구항 31

세포(예를 들어 세포 집단) 내의 표적 서열에서 또는 그 부근에서 상기 세포를 변경시키는(예를 들어 핵산의 구조(예를 들어 서열)를 변경시키는) 방법으로서, 상기 세포(예를 들어 세포 집단)를

1) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자 및 Cas9 분자;

2) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자 및 Cas9 분자를 인코딩하는 핵산;

3) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자를 인코딩하는 핵산 및 Cas9 분자;

4) 제1항 내지 제22항 중 어느 한 항의 하나 이상의 gRNA 분자를 인코딩하는 핵산 및 Cas9 분자를 인코딩하는 핵산;

5) 임의의 상기 1) 내지 4), 및 주형 핵산;

6) 임의의 상기 1) 내지 4), 및 주형 핵산을 인코딩하는 서열을 포함하는 핵산;

7) 제23항 내지 제28항 중 어느 한 항의 조성물; 또는

8) 제30항의 벡터

와 접촉시키는(예를 들어 이것 내로 도입시키는) 단계를 포함하는, 방법.

청구항 32

제31항에 있어서, 세포가 동물 세포, 예를 들어 포유류, 영장류 또는 인간 세포이며, 예를 들어 인간 세포이며; 선택적으로 상기 세포가 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아를 앓고 있는 환자로부터 수득되는, 방법.

청구항 33

제31항 또는 제32항에 있어서, 세포가 HSPC, 선택적으로 CD34+ HSPC, 선택적으로 CD34+CD90+ HSPC인, 방법.

청구항 34

제31항 내지 제33항 중 어느 한 항에 있어서, 세포가 CD34+ 세포에 대해 농화되었던 세포 집단을 포함하는 조성물에 배치되는, 방법.

청구항 35

제31항 내지 제34항 중 어느 한 항에 있어서, 세포(예를 들어 세포 집단)가 골수, 말초 혈액(예를 들어 동원된 말초 혈액), 또는 제대혈로부터 단리되었던 것인, 방법.

청구항 36

제31항 내지 제35항 중 어느 한 항에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가 또는 동종이계인, 방법.

청구항 37

제31항 내지 제36항 중 어느 한 항에 있어서,

a) 변경이 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 초래하고/하거나;

b) 변경이, HBG1 프로모터 영역에서 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 초래하고, 선택적으로 결실이 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는, 방법.

청구항 38

제31항 내지 제37항 중 어느 한 항에 있어서,

(a) 방법이 세포 집단을 초래하며, 여기서, 집단 중 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 40%, 예를 들어 적어도 약 50%, 예를 들어 적어도 약 55%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 75%가 변경되었으며, 예를 들어 삽입결실을 포함하고, 선택적으로 삽입결실이 표 2 내지 표 7에 열거된 삽입결실로부터 선택되고, 선택적으로 집단의 세포가 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않으며;

(b) 변경이 적혈구계 계통의 분화된 세포(예를 들어 적혈구)로 분화할 수 있는 세포(예를 들어 세포 집단)를 초래하고, 상기 분화된 세포가 예를 들어 비변경된 세포(예를 들어 세포 집단)에 비해 증가된 수준의 태아 헤모글로빈을 나타내며;

(c) 변경이 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수

있는 세포 집단을 초래하고, 분화된 세포의 상기 집단이 예를 들어 비변경된 세포의 집단에 비해 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가지고/가지거나;

(d) 변경이 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있는 세포(예를 들어 세포 집단)를 초래하고, 상기 분화된 세포가 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, 방법.

청구항 39

제31항 내지 제38항 중 어느 한 항의 방법에 의해 변경된 세포, 또는 제31항 내지 제38항 중 어느 한 항의 방법에 의해 수득 가능한 세포.

청구항 40

표 7-2에 기재된 삽입결실을 포함하는 세포로서, 선택적으로 세포는 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않는, 세포.

청구항 41

제1항 내지 제22항 중 어느 한 항의 제1 gRNA 분자, 또는 제23항 내지 제28항 중 어느 한 항의 조성물, 제29항의 핵산, 또는 제30항의 벡터를 포함하는 세포.

청구항 42

제39항 내지 제41항 중 어느 한 항에 있어서, 세포가 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있고, 여기서, 상기 분화된 세포가 예를 들어 gRNA 분자를 포함하도록 변형되지 않은 동일한 유형의 세포에 비해, 증가된 수준의 태아 헤모글로빈을 나타내며, 선택적으로 분화된 세포(예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구)가, 예를 들어 gRNA 분자를 포함하도록 변형되지 않은 동일한 유형의 분화된 세포에 비해, 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, 세포.

청구항 43

제39항 내지 제42항 중 어느 한 항에 있어서, 줄기세포 확장제(expander)와 접촉되었던 세포.

청구항 44

제43항에 있어서, 줄기세포 확장제가

a) (1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민;

b) 메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트;

c) 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀;

d) (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올; 또는

e) 이들의 조합(예를 들어 (1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)

인, 세포.

청구항 45

세포, 예를 들어 제39항 내지 제44항 중 어느 한 항의 세포로서,

a) 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근

에서 삽입결실을 포함하고/하거나;

b) HBG1 프로모터 영역에서 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 포함하고, 선택적으로 결실이 5,250,092 내지 5,249,833, -가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는, 세포.

청구항 46

제39항 내지 제45항 중 어느 한 항에 있어서, 세포가 동물 세포, 예를 들어 포유류, 영장류 또는 인간 세포이며, 예를 들어 인간 세포이며; 선택적으로 상기 세포가 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아를 앓고 있는 환자로부터 수득되는, 세포.

청구항 47

제39항 내지 제46항 중 어느 한 항에 있어서, 세포가 HSPC, 선택적으로 CD34+ HSPC, 선택적으로 CD34+CD90+ HSPC인, 세포.

청구항 48

제39항 내지 제47항 중 어느 한 항에 있어서, 세포(예를 들어 세포 집단)가 골수, 말초 혈액(예를 들어 동원된 말초 혈액), 또는 제대혈로부터 단리되었던 것인, 세포.

청구항 49

제39항 내지 제48항 중 어느 한 항에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가 또는 동종이계인, 세포.

청구항 50

제39항 내지 제49항 중 어느 한 항의 세포를 포함하는 세포 집단으로서, 선택적으로 집단의 세포 중 적어도 약 50%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 80%, 예를 들어 적어도 약 90%(예를 들어 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%)가 제39항 내지 제49항 중 어느 한 항에 따른 세포인, 세포 집단.

청구항 51

제50항에 있어서, 세포 집단이 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 분화된 세포의 상기 집단이 예를 들어 동일한 유형의 비변형된 세포 집단에 비해, 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 17%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가지며; 선택적으로 분화된 세포의 집단의 F 세포가 1개 세포당 평균적으로 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, 세포 집단.

청구항 52

제50항 또는 제51항에 있어서,

1) 세포가 투여되는 환자의 체중 1 kg 당 적어도 1e6개 CD34+ 세포;

2) 세포가 투여되는 환자의 체중 1 kg 당 적어도 2e6개 CD34+ 세포;

3) 세포가 투여되는 환자의 체중 1 kg 당 적어도 3e6개 CD34+ 세포;

4) 세포가 투여되는 환자의 체중 1 kg 당 적어도 4e6개 CD34+ 세포; 또는

5) 세포가 투여되는 환자의 체중 1 kg 당 2e6 내지 10e6개 CD34+ 세포

를 포함하는, 세포 집단.

청구항 53

제50항 내지 제52항 중 어느 한 항에 있어서, 집단의 세포 중 적어도 약 40%, 예를 들어 적어도 약 50%(예를 들어 적어도 약 60%, 적어도 약 70%, 적어도 약 80% 또는 적어도 약 90%)가 CD34+ 세포이며, 선택적으로 집단의 세포 중 적어도 약 10%, 예를 들어 적어도 약 15%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%가 CD34+CD90+ 세포인, 세포 집단.

청구항 54

제50항 내지 제53항 중 어느 한 항에 있어서, 세포 집단이 제대혈, 말초 혈액(예를 들어 동원된 말초 혈액), 또는 골수로부터 유래되거나, 예를 들어 골수로부터 유래되는, 세포 집단.

청구항 55

제50항 내지 제54항 중 어느 한 항에 있어서, 세포 집단이 포유류 세포, 예를 들어 인간 세포를 포함하며, 예를 들어 이로 구성되며, 선택적으로 세포 집단이 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아를 앓고 있는 환자로부터 수득되는, 세포 집단.

청구항 56

제50항 내지 제55항 중 어느 한 항에 있어서, 세포 집단이 (i) 이러한 세포 집단이 투여되는 환자에 비해 자가이거나, (ii) 이러한 세포 집단이 투여되는 환자에 비해 동종이계인, 세포 집단.

청구항 57

예를 들어 제50항 내지 제56항 중 어느 한 항에 있어서, 표 7-2에 기재된 바와 같은 삽입결실 패턴을 포함하는 세포(예를 들어 CD34+ 세포) 집단으로서, 선택적으로 표 7-2에 기재된 삽입결실 패턴의 삽입결실은 집단의 세포 중 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90% 또는 적어도 95%에서 검출 가능한 것인, 세포(예를 들어 CD34+ 세포) 집단.

청구항 58

제39항 내지 제57항 중 어느 한 항의 세포 또는 세포 집단을 포함하며, 선택적으로 약제학적으로 허용 가능한 배지, 예를 들어 동결보존에 적합한 약제학적으로 허용 가능한 배지를 포함하는, 조성물.

청구항 59

혈색소병증을 치료하는 방법으로서, 제39항 내지 제57항 중 어느 한 항의 세포 또는 세포 집단, 또는 제58항의 조성물을 환자에게 투여하는 단계를 포함하는, 방법.

청구항 60

포유류에서 태아 헤모글로빈 발현을 증가시키는 방법으로서, 제39항 내지 제57항 중 어느 한 항의 세포 또는 세포 집단, 또는 제58항의 조성물을 환자에게 투여하는 단계를 포함하는, 방법.

청구항 61

제60항에 있어서, 혈색소병증이 베타-탈라쎄미아 또는 겸상 적혈구 질병인, 방법.

청구항 62

세포(예를 들어 세포 집단)를 제조하는 방법으로서,

(a) 세포(예를 들어 세포 집단)(예를 들어 HSPC(예를 들어 HSPC의 집단))를 제공하는 단계;

(b) 상기 세포(예를 들어 상기 세포 집단)를 줄기세포 확장제를 포함하는 세포 배양 배지에서 생체외에서 배양하는 단계; 및

(c) 상기 세포 내로 제1항 내지 제22항 중 어느 한 항의 제1 gRNA 분자, 제1항 내지 제22항 중 어느 한 항의 제

1 gRNA 분자를 인코딩하는 핵산 분자, 제23항 내지 제28항 중 어느 한 항의 조성물, 제29항의 핵산, 또는 제30항의 벡터를 도입하는 단계

를 포함하는, 방법.

청구항 63

제62항에 있어서, 단계 (c)의 상기 도입 후, 상기 세포(예를 들어 세포 집단)가 분화된 세포(예를 들어 분화된 세포의 집단), 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구계 계통의 세포 집단), 예를 들어 적혈구(예를 들어 적혈구 집단)로 분화할 수 있고, 상기 분화된 세포(예를 들어 분화된 세포의 집단)가 예를 들어 단계 (c)를 거치지 않은 동일한 세포에 비해, 증가된 태아 헤모글로빈을 생산하는, 방법.

청구항 64

제62항 또는 제63항에 있어서, 줄기세포 확장제가

a) (1r,4r)-N1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민;

b) 메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트;

c) 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀;

d) (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올; 또는

e) 이들의 조합(예를 들어 (1r,4r)-N1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)

인, 방법.

청구항 65

제62항 내지 제64항 중 어느 한 항에 있어서,

세포 배양 배지가 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)를 포함하며, 선택적으로 세포 배양 배지가 인간 인터루킨-6(IL-6)을 추가로 포함하며; 선택적으로 세포 배양 배지가 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L), 인간 줄기세포 인자(SCF) 및 존재한다면 인간 IL-6을 각각 약 10 ng/mL 내지 약 1000 ng/mL 범위의 농도, 선택적으로 각각 약 50 ng/mL의 농도, 예를 들어 50 ng/mL의 농도로 포함하는, 방법.

청구항 66

제62항 내지 제65항 중 어느 한 항에 있어서, 세포 배양 배지가 줄기세포 확장제를 약 1 nM 내지 약 1 mM 범위의 농도, 선택적으로 약 1 uM 내지 약 100 nM 범위의 농도, 선택적으로 약 500 nM 내지 약 750 nM 범위의 농도, 선택적으로 약 500 nM의 농도, 예를 들어 500 nM의 농도, 또는 약 750 nM의 농도, 예를 들어 750 nM의 농도로 포함하는, 방법.

청구항 67

제62항 내지 제66항 중 어느 한 항에 있어서, 단계 (b)의 배양이 단계 (c)의 도입 전에 배양 기간을 포함하며, 선택적으로 단계 (c)의 도입 전의 배양 기간이 적어도 12시간, 예를 들어 약 1일 내지 약 12일의 기간 동안이며, 예를 들어 약 1일 내지 약 6일의 기간 동안이며, 예를 들어 약 1일 내지 약 3일의 기간 동안이며, 예를 들어 약 1일 내지 약 2일의 기간이고, 예를 들어 약 2일의 기간 동안인, 방법.

청구항 68

제62항 내지 제67항 중 어느 한 항에 있어서, 단계 (b)의 배양이 단계 (c)의 도입 후에 배양 기간을 포함하며, 선택적으로 단계 (c)의 도입 후의 배양 기간이 적어도 12시간, 예를 들어 약 1일 내지 약 12일의 기간 동안이며, 예를 들어 약 1일 내지 약 6일의 기간 동안이며, 예를 들어 약 2일 내지 약 4일의 기간 동안이며, 예를 들어 약 2일의 기간 동안이거나 약 3일의 기간 동안이거나, 약 4일의 기간 동안인, 방법.

청구항 69

제62항 내지 제68항 중 어느 한 항에 있어서, 세포 집단이 생체외에서 적어도 3배, 예를 들어 적어도 4배, 예를 들어 적어도 5배, 예를 들어 적어도 10배 확장되는, 방법.

청구항 70

제62항 내지 제69항 중 어느 한 항에 있어서, 단계 (c)의 도입이 전기천공을 포함하는, 방법.

청구항 71

제62항 내지 제70항 중 어느 한 항에 있어서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 인간 세포(예를 들어 인간 세포의 집단)인, 방법.

청구항 72

제71항에 있어서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 골수, 말초 혈액(예를 들어 동원된 말초 혈액) 또는 제대혈로부터 단리된 것인, 방법.

청구항 73

제72항에 있어서,

(i) 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 골수로부터 단리되며, 예를 들어 혈색소병증을 앓고 있는 환자의 골수로부터 단리되며, 선택적으로 혈색소병증이 겸상 적혈구 질병 또는 탈라쎄미아이고, 선택적으로 탈라쎄미아가 베타 탈라쎄미아이거나;

(ii) 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 말초 혈액으로부터 단리되며, 예를 들어 혈색소병증을 앓고 있는 환자의 말초 혈액으로부터 단리되며, 선택적으로 혈색소병증이 겸상 적혈구 질병 또는 탈라쎄미아이며, 선택적으로 탈라쎄미아가 베타 탈라쎄미아이고; 선택적으로 말초 혈액이 동원된 말초 혈액이며, 선택적으로 동원된 말초 혈액이 Plerixafor, G-CSF 또는 이들의 조합을 사용하여 동원된 것인, 방법.

청구항 74

제62항 내지 제73항 중 어느 한 항에 있어서, 단계 (a)에서 제공되는 세포 집단이 CD34+ 세포에 대해 농화된 것인, 방법.

청구항 75

제62항 내지 제74항 중 어느 한 항에 있어서, 단계 (c)의 도입에 후속하여, 세포(예를 들어 세포 집단)가 동결보존되는, 방법.

청구항 76

제62항 내지 제75항 중 어느 한 항에 있어서, 단계 (c)의 도입에 후속하여, 세포(예를 들어 세포 집단)가

a) 제1 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하고/하거나;

b) HBG1 프로모터 영역에서 제1 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 제1 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 포함하고, 선택적으로 삽입결실, 예를 들어 결실이 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는, 방법.

청구항 77

제62항 내지 제76항 중 어느 한 항에 있어서,

(a) 단계 (c)의 도입 후, 세포 집단의 세포 중 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%,

적어도 약 80%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%가 제1 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하고, 선택적으로 삽입결실이 표 2 내지 표 7에 열거된 삽입결실로부터 선택되고, 선택적으로 집단의 세포 중 어떠한 세포도 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않으며;

(b) 단계 (c)의 도입 후, 세포(예를 들어 세포 집단)가 적혈구계 계통의 분화된 세포(예를 들어 적혈구)로 분화할 수 있고, 상기 분화된 세포가 예를 들어 비변경된 세포(예를 들어 세포 집단)에 비해 증가된 수준의 태아 헤모글로빈을 나타내며;

(c) 단계 (c)의 도입 후, 세포 집단이 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 분화된 세포의 상기 집단이 예를 들어 비변경된 세포의 집단에 비해, 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가지며;

(d) 단계 (c)의 도입 후, 세포(예를 들어 세포 집단)가 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있고, 상기 분화된 세포(예를 들어 분화된 세포의 집단)가 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하며;

(e) 단계 (c)의 도입 후, 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실도 상기 세포에서 형성되지 않으며, 예를 들어 어떠한 표적-외 삽입결실도 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 형성되지 않고/않거나;

(f) 단계 (c)의 도입 후, 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 어떠한 표적-외 삽입결실도 세포 집단의 세포 중 약 5% 초과, 예를 들어 약 1% 초과, 예를 들어 약 0.1% 초과, 예를 들어 약 0.01% 초과에서 검출되지 않는, 방법.

청구항 78

제62항 내지 제77항 중 어느 한 항의 방법에 의해 수득 가능한 세포(예를 들어 세포 집단).

청구항 79

세포, 예를 들어 변경된 세포, 예를 들어 제78항의 세포로서, 여기서:

(a) 세포 집단의 세포 중 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%가 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하며, 선택적으로 삽입결실은 표 2 내지 표 7에 열거된 삽입결실로부터 선택되고, 선택적으로 집단의 어떠한 세포도 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않으며;

(b) 세포(예를 들어 세포 집단)가 적혈구계 계통의 분화된 세포(예를 들어 적혈구)로 분화할 수 있고, 상기 분화된 세포가 예를 들어 비변경된 세포(예를 들어 세포 집단)에 비해, 증가된 수준의 태아 헤모글로빈을 나타내며;

(c) 세포 집단이 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 분화된 세포의 상기 집단이 예를 들어 비변경된 세포의 집단에 비해 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가지며;

(d) 세포(예를 들어 세포 집단)가 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있고, 상기 분화된 세포(예를 들어 분화된 세포의 집단)가 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하며;

(e) 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실도 상기 세포에서 형성되지 않으며, 예를 들어 어떠한 표적-외 삽입결실도 HBG1 및/또는 HBG2 프로

모터 영역의 외부에서 형성되지 않으며;

(f) 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 어떠한 표적-외 삽입결실도 세포 집단의 세포 중 약 5% 초과, 예를 들어 약 1% 초과, 예를 들어 약 0.1% 초과, 예를 들어 약 0.01% 초과에서 검출되지 않고/않거나;

(g) 상기 세포 또는 이의 자손이 선택적으로 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 검출함으로써 검출되는 바와 같이, 세포 또는 이의 자손이 이식되는 환자에서 이식 후 16주 초과, 20주 초과 또는 24주 초과 시에 검출 가능하고, 선택적으로 삽입결실은 표 2 내지 표 7에 열거된 삽입결실로부터 선택되는, 세포.

청구항 80

제78항 또는 제79항에 있어서, 세포가 동물 세포, 예를 들어 포유류, 영장류 또는 인간 세포이며, 예를 들어 인간 세포이며; 선택적으로 상기 세포가 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아를 앓고 있는 환자로부터 수득되는, 세포.

청구항 81

제78항 내지 제80항 중 어느 한 항에 있어서, 세포가 HSPC, 선택적으로 CD34+ HSPC, 선택적으로 CD34+CD90+ HSPC인, 세포.

청구항 82

제78항 내지 제81항 중 어느 한 항에 있어서, 세포(예를 들어 세포 집단)가 골수, 말초 혈액(예를 들어 동원된 말초 혈액), 또는 제대혈로부터 단리되었던 것인, 세포.

청구항 83

제78항 내지 제82항 중 어느 한 항에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가 또는 동종이계인, 세포.

청구항 84

혈색소병증을 치료하는 방법으로서, 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 세포 또는 세포 집단을 포함하는 조성물을 인간 환자에게 투여하는 단계를 포함하는, 방법.

청구항 85

인간 환자에서 태아 헤모글로빈 발현을 증가시키는 방법으로서, 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 세포 또는 세포 집단을 포함하는 조성물을 상기 인간 환자에게 투여하는 단계를 포함하는, 방법.

청구항 86

제84항에 있어서, 혈색소병증이 베타-탈라쎄미아 또는 겸상 적혈구 질병인, 방법.

청구항 87

제84항 내지 제86항 중 어느 한 항에 있어서, 인간 환자에게 인간 환자의 체중 1 kg 당 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 적어도 약 2e6개 세포, 예를 들어 인간 환자의 체중 1 kg 당 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 적어도 약 2e6개 CD34+ 세포를 포함하는 조성물이 투여되는, 방법.

청구항 88

제84항 내지 제87항 중 어느 한 항에 있어서, 세포 또는 세포 집단, 또는 이의 자손이 선택적으로 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 검출함으로써 검출된 바와 같이 투여 후 16주 초과, 20주 초과 또는 24주 초과 시에 인간 환자에서 검출 가

능하며, 선택적으로 삽입결실이 표 2 내지 표 7에 열거된 삽입결실로부터 선택되고; 선택적으로 투여 후 16주 초과, 20주 초과 또는 24주 초과 시에 참조 세포(예를 들어 CD34+ 세포) 집단에서 검출된 삽입결실의 수준이 투여 직전에 세포 집단에서 검출된 삽입결실의 수준에 비해, 50% 이하, 40% 이하, 30% 이하, 20% 이하, 10% 이하, 5% 이하 또는 1% 이하만큼 감소되는, 방법.

청구항 89

약제로서 사용하기 위한, 제1항 내지 제22항 중 어느 한 항의 gRNA 분자, 제23항 내지 28항 중 어느 한 항 또는 제58항의 조성물, 제29항의 핵산, 제30항의 벡터, 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 세포 또는 세포 집단.

청구항 90

약제의 제조에 사용하기 위한, 제1항 내지 제22항 중 어느 한 항의 gRNA 분자, 제23항 내지 28항 중 어느 한 항 또는 제58항의 조성물, 제29항의 핵산, 제30항의 벡터, 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 세포 또는 세포 집단.

청구항 91

질병의 치료에 사용하기 위한, 제1항 내지 제22항 중 어느 한 항의 gRNA 분자, 제23항 내지 28항 중 어느 한 항 또는 제58항 중의 조성물, 제29항의 핵산, 제30항의 벡터, 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 세포 또는 세포 집단.

청구항 92

질병의 치료에 사용하기 위한, 제1항 내지 제22항 중 어느 한 항의 gRNA 분자, 제23항 내지 28항 중 어느 한 항 또는 제58항의 조성물, 제29항의 핵산, 제30항의 벡터, 제39항 내지 제57항 또는 제78항 내지 제83항 중 어느 한 항의 세포 또는 세포 집단으로서, 여기서, 질병이 혈색소병증이며, 선택적으로 혈색소병증이 겸상 적혈구 질병 또는 탈라쎄미아(예를 들어 베타-탈라쎄미아)인, gRNA 분자, 조성물, 핵산, 벡터, 세포 또는 세포 집단.

발명의 설명

기 술 분 야

[0001] 관련 출원에 대한 교차 참조

[0002] 본 출원은 2017년 2월 6일에 출원된 미국 가특허 출원 62/455,464에 대해 우선권을 주장하며, 이의 내용은 그 전체가 참조로 본 명세서에 포함된다.

[0003] 서열 목록

[0004] 본 출원은 ASCII 포맷으로 전자적으로 제출된 서열 목록을 포함하고, 상기 서열 목록은 그 전체가 참조로 본 명세서에 포함된다. 2018년 1월 30일에 생성된 상기 ASCII 카피는 명칭이 PAT057603-WO-PCT_SL.txt이고 크기가 258,837 바이트이다.

배 경 기 술

[0005] RISPR(Clustered Regularly Interspaced Short Palindromic Repeats, 주기적 간격으로 분포하고 짧은 회문 구조의 반복서열)은 바이러스 공격에 대하여 방어하기 위한 적응 면역 시스템으로서 박테리아에서 진화하였다. 바이러스에 노출 시, 바이러스 DNA의 짧은 절편이 박테리아 게놈의 CRISPR 유전자좌 내로 통합된다. RNA는 바이러스 서열을 포함하는 CRISPR 유전자좌의 일부분으로부터 전사된다. 바이러스 게놈에 상보적인 서열을 함유하는 해당 RNA는 바이러스 게놈 내의 서열로의 Cas9 단백질의 표적화를 매개한다. Cas9 단백질은 바이러스 표적을 절단하여 침묵시킨다.

[0006] 최근, CRISPR/Cas 시스템은 진핵 세포에서 게놈 편집을 위해 적응되었다. 부위-특이적 단일 가닥 절단부(SSB; single strand break) 또는 이중 가닥 절단부(DSB; double strand break)의 도입은, 예를 들어 비상동 말단-접합(NHEJ; non-homologous end-joining) 또는 상동성-지시 수선(HDR; homology-directed repair)을 통한 표적 서열 변경을 가능하게 한다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0007] 이론으로 결부시키고자 하는 것은 아니지만, 본 발명은 부분적으로는, 예를 들어 본원에 기재된 바와 같은 CRISPR 시스템, 예를 들어 Cas9 CRISPR 시스템이 예를 들어 변형된 세포의 자손, 예를 들어 적혈구 자손에서 태아 헤모글로빈(HbF) 발현을 증가시키고/시키거나 베타 글로빈(예를 들어 질병-유발 돌연변이를 갖는 베타 글로빈 유전자)의 발현을 저하시키기 위해 본원에 기재된 바와 같이 예를 들어 비결실(nondeltional) HPFH 영역에서 세포(예를 들어 조혈 줄기 및 전구 세포(HSPC))를 변형시키는 데 사용될 수 있고, 변형된 세포(예를 들어 변형된 HSPC)가 혈색소병증, 예를 들어 겸상 적혈구 질병 및 베타 탈라쎄미아(beta thalassemia)를 치료하는 데 사용될 수 있다는 발견을 기초로 한다. 일 양태에서, 어떠한 공지된 HPFH 돌연변이 또는 결실도 맵핑(map)되지 않은 게놈의 영역을 표적화하는 예를 들어 본원에 기재된 바와 같은 유전자 편집 시스템, 예를 들어 CRISPR 시스템을 세포(예를 들어 HSPC)에 도입하는 것은, 유기체 내로 효율적으로 생착하고, 생착된 유기체 내에서 장기간 지속되고, 증가된 태아 헤모글로빈 발현을 갖는 적혈구를 포함하여 이로 분화할 수 있는 변형된 HSPC(예를 들어 본원에 기재된 바와 같이 예를 들어 하나 이상의 삽입결실(indel)을 포함하는 HSPC)를 생성한다는 것이 본원에서 놀랍게도 제시되었다. 또한, 이들 변형된 HSPC는 예를 들어 줄기세포 확장제(expander)(예를 들어 본원에 기재된 바와 같음)의 존재 하에, 줄기세포능(stemness)을 유지하면서도 이들 HSPC의 확장 및 증식을 유발하는 조건 하에 생체외에서 배양될 수 있다. 예를 들어 본원에 기재된 바와 같은 유전자 편집 시스템, 예를 들어 CRISPR 시스템이 겸상 적혈구 질병 환자로부터 유래된 HPSC 내로 도입될 때, 변형된 세포 및 이들의 자손(예를 들어 적혈구계(erythroid) 자손)은 놀랍게도 비변형된 세포 집단에 비해, 태아 헤모글로빈의 상향조절을 보여줄 뿐만 아니라 겸상 베타-글로빈의 유의한 저하, 및 겸상 적혈구 수의 유의한 저하와 정상 적혈구 수의 증가를 보여준다.

[0008] 따라서, 일 양태에서, 본 발명은 본원에 기재된 바와 같은 하나 이상의, 예를 들어 하나의 gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 Cas CRISPR 시스템, 예를 들어 Cas9 CRISPR 시스템, 예를 들어 *S. 피오게네스(S. pyogenes)* Cas9 CRISPR 시스템)을 제공한다. 본원에 기재된 임의의 gRNA 분자는 이러한 시스템, 및 본원에 기재된 방법 및 세포에 사용될 수 있다.

[0009] 일 양태에서, 본 발명은 tracr 및 crRNA를 포함하는 gRNA 분자를 제공하며, 여기서, crRNA는:

[0010] a) 비결실 HFPH 영역(예를 들어 인간 비결실 HPFH 영역)의 표적 서열에 상보적인 표적화 도메인;

[0011] b) Chr11:5,249,833 내지 Chr11:5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[0012] c) Chr11:5,254,738 내지 Chr11:5,255,164, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[0013] d) Chr11:5,250,094 내지 5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[0014] e) Chr11:5,255,022 내지 5,255,164, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[0015] f) Chr11:5,249,833 내지 5,249,927, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[0016] g) Chr11:5,254,738 내지 5,254,851, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[0017] h) Chr11:5,250,139 내지 5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인; 또는

[0018] i) 이들의 조합

[0019] 을 포함한다.

[0020] 실시형태에서, 표적화 도메인은 SEQ ID NO: 1 내지 SEQ ID NO: 72 중 임의의 하나를 포함하며, 예를 들어 이로 구성된다. 실시형태에서, 표적화 도메인은 SEQ ID NO: 1, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 28, SEQ ID NO: 34, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 58, SEQ ID NO: 62, SEQ ID NO: 63 또는 SEQ ID NO: 67 중 임의의 하나를 포함하며, 예를 들어 이로 구성된다. 실시형태에서, 표적화 도메인은 a) SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 28, SEQ ID NO: 34, SEQ ID NO: 48, SEQ ID NO: 51 또는 SEQ ID NO: 67; 또는 b) SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12 또는 SEQ ID NO: 54 중 임의의 하나를 포함하며, 예를 들어 이로 구성된다. 실시형태에서, gRNA 분자는 SEQ ID NO: 8을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다. 실시형태에서, gRNA 분자는 SEQ ID NO: 67을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다. 실시형태에서, gRNA 분자는 임의의 상기 서열의 단편을 포함하는(예를 들어 이로 구성되는) 표적화 도메인을 포함한다.

[0021] 임의의 상기 언급된 양태 및 실시형태에서, gRNA 분자는 본원에 기재된 영역 및/또는 특성을 추가로 가질 수 있다. 실시형태에서, gRNA 분자는 임의의 상기 언급된 표적화 도메인의 단편을 포함한다. 실시형태에서, 표적화 도메인은 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산을 포함하며, 예를 들어 이로 구성된다. 실시형태에서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산은 나열된 표적화 도메인 서열의 3' 말단에 배치된 17, 18, 19 또는 20개 연속 핵산이다. 다른 실시형태에서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산은 나열된 표적화 도메인 서열의 5' 말단에 배치된 17, 18, 19 또는 20개 연속 핵산이다. 다른 실시형태에서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산은 나열된 표적화 도메인 서열의 5' 또는 3' 핵산 중 어느 것도 포함하지 않는다. 실시형태에서, 표적화 도메인은 나열된 표적화 도메인 서열로 구성된다.

[0022] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, crRNA의 일부 및 tracr의 일부는 혼성화하여, SEQ ID NO: 182 또는 183을 포함하는 플래그폴(flagpole)을 형성한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 플래그폴은 이러한 플래그폴의 crRNA 부분에 대해 3'에 위치한 제1 플래그폴 연장(extension)을 추가로 포함하고, 여기서, 상기 제1 플래그폴 연장은 SEQ ID NO: 184를 포함한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 플래그폴은 이러한 플래그폴의 crRNA 부분에 대해 3'에 위치한 제2 플래그폴 연장 및 존재한다면 제1 플래그폴 연장을 추가로 포함하고, 여기서, 상기 제2 플래그폴 연장은 SEQ ID NO: 185를 포함한다.

[0023] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, tracr는 SEQ ID NO: 224 또는 SEQ ID NO: 225를 포함한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, tracr는 SEQ ID NO: 232를 포함하며, 선택적으로 3' 말단에 부가적인 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 추가로 포함한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, crRNA는 5’ 으로부터 3'까지, [표적화 도메인]-: a) SEQ ID NO: 182; b) SEQ ID NO: 183; c) SEQ ID NO: 199; d) SEQ ID NO: 200; e) SEQ ID NO: 201; f) SEQ ID NO: 202; 또는 g) SEQ ID NO: 226을 포함한다.

[0024] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, tracr는 5’ 으로부터 3'까지: a) SEQ ID NO: 187; b) SEQ ID NO: 188; c) SEQ ID NO: 203; d) SEQ ID NO: 204; e) SEQ ID NO: 224; f) SEQ ID NO: 225; g) SEQ ID NO: 232; h) SEQ ID NO: 227; i) SEQ ID NO: 228; j) SEQ ID NO: 229; k) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 j); l) 3' 말단에, 적어도 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 k); 또는 m) 5' 말단(end)에(예를 들어 5' 말단(terminus)에), 적어도 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 l)을 포함한다.

[0025] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 표적화 도메인 및 tracr는 개별 핵산 분자 상에 배치된다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 표적화 도메인 및 tracr는 개별 핵산 분자 상에 배치되고, 표적화 도메인을 포함하는 핵산 분자는 선택적으로 이러한 표적화 도메인에 대해 바로(immediately) 3'에 배치된 SEQ ID NO: 201을 포함하고, tracr를 포함하는 핵산 분자는 SEQ ID NO: 224를 포함하며, 예를 들어 이로 구성된다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 플래그폴의

crRNA 부분은 SEQ ID NO: 201 또는 SEQ ID NO: 202를 포함한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, tracr는 SEQ ID NO: 187 또는 188, 및 선택적으로 제1 플래그폴 연장이 존재한다면 SEQ ID NO: 187 또는 188에 대해 5'에 배치된 제1 tracr 연장을 포함하고, 상기 제1 tracr 연장은 SEQ ID NO: 189를 포함한다.

[0026] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 표적화 도메인 및 tracr는 예를 들어 단일 핵산 분자 상에 배치되며, 여기서, tracr는 표적화 도메인에 대해 3'에 배치된다. 일 양태에서, gRNA 분자는 표적화 도메인에 대해 3' 및 tracr에 대해 5'에 배치된 루프를 포함한다. 실시형태에서, 루프는 SEQ ID NO: 186을 포함한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, gRNA 분자는 5' 으로부터 3'까지, [표적화 도메인]-: (a) SEQ ID NO: 195; (b) SEQ ID NO: 196; (c) SEQ ID NO: 197; (d) SEQ ID NO: 198; (e) SEQ ID NO: 231; 또는 (f) 3' 말단에 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 (a) 내지 (e)를 포함한다.

[0027] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 표적화 도메인 및 tracr는 단일 핵산 분자 상에 배치되고, 여기서, 상기 핵산 분자는 상기 표적화 도메인, 및 선택적으로 상기 표적화 도메인에 대해 바로 3'에 배치된 SEQ ID NO: 231을 포함하며, 예를 들어 이로 구성된다.

[0028] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, gRNA 분자를 포함하여 핵산 분자 중 1개 또는 선택적으로 1개 초과는:

[0029] a) 상기 핵산 분자 또는 분자들의 3' 말단에 하나 이상의, 예를 들어 3개의, 포스포로티오에이트 변형;

[0030] b) 상기 핵산 분자 또는 분자들의 5' 말단에 하나 이상의, 예를 들어 3개의, 포스포로티오에이트 변형;

[0031] c) 상기 핵산 분자 또는 분자들의 3' 말단에 하나 이상의, 예를 들어 3개의, 2'-O-메틸 변형;

[0032] d) 상기 핵산 분자 또는 분자들의 5' 말단에 하나 이상의, 예를 들어 3개의, 2'-O-메틸 변형;

[0033] e) 상기 핵산 분자 또는 분자들의 말단 방향으로 4 번째, 말단 방향으로 3 번째 및 말단 방향으로 2 번째 3' 잔기 각각에서 2' O-메틸 변형;

[0034] f) 상기 핵산 분자 또는 분자들의 말단 방향으로 4 번째, 말단 방향으로 3 번째 및 말단 방향으로 2 번째 5' 잔기 각각에서 2' O-메틸 변형; 또는

[0035] f) 이들의 임의의 조합

[0036] 을 포함한다.

[0037] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0038] (a) SEQ ID NO: 74;

[0039] (b) SEQ ID NO: 75; 또는

[0040] (c) SEQ ID NO: 76.

[0041] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0042] (a) SEQ ID NO: 77을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0043] (b) SEQ ID NO: 77을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0044] (c) SEQ ID NO: 78을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0045] (d) SEQ ID NO: 78을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0046] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0047] (a) SEQ ID NO: 79;

[0048] (b) SEQ ID NO: 80; 또는

[0049] (c) SEQ ID NO: 81.

[0050] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0051] (a) SEQ ID NO: 82를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0052] (b) SEQ ID NO: 82를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0053] (c) SEQ ID NO: 83을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0054] (d) SEQ ID NO: 83을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0055] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0056] (a) SEQ ID NO: 84;

[0057] (b) SEQ ID NO: 85; 또는

[0058] (c) SEQ ID NO: 86.

[0059] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0060] (a) SEQ ID NO: 87을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0061] (b) SEQ ID NO: 87을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0062] (c) SEQ ID NO: 88을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0063] (d) SEQ ID NO: 88을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0064] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0065] (a) SEQ ID NO: 89;

[0066] (b) SEQ ID NO: 90; 또는

[0067] (c) SEQ ID NO: 91.

[0068] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0069] (a) SEQ ID NO: 92를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0070] (b) SEQ ID NO: 92를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0071] (c) SEQ ID NO: 93을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0072] (d) SEQ ID NO: 93을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0073] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0074] (a) SEQ ID NO: 94;

[0075] (b) SEQ ID NO: 95; 또는

[0076] (c) SEQ ID NO: 96.

[0077] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0078] (a) SEQ ID NO: 97을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0079] (b) SEQ ID NO: 97을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0080] (c) SEQ ID NO: 98을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0081] (d) SEQ ID NO: 98을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0082] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0083] (a) SEQ ID NO: 99;

[0084] (b) SEQ ID NO: 100; 또는

[0085] (c) SEQ ID NO: 101.

[0086] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0087] (a) SEQ ID NO: 102를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0088] (b) SEQ ID NO: 102를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0089] (c) SEQ ID NO: 103을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0090] (d) SEQ ID NO: 103을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0091] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0092] (a) SEQ ID NO: 104;

[0093] (b) SEQ ID NO: 105; 또는

[0094] (c) SEQ ID NO: 106.

[0095] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0096] (a) SEQ ID NO: 107을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0097] (b) SEQ ID NO: 107을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0098] (c) SEQ ID NO: 108을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0099] (d) SEQ ID NO: 108을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0100] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0101] (a) SEQ ID NO: 109;

[0102] (b) SEQ ID NO: 110; 또는

[0103] (c) SEQ ID NO: 111.

[0104] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0105] (a) SEQ ID NO: 112를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0106] (b) SEQ ID NO: 112를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0107] (c) SEQ ID NO: 113을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0108] (d) SEQ ID NO: 113을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0109] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0110] (a) SEQ ID NO: 114;

[0111] (b) SEQ ID NO: 115; 또는

[0112] (c) SEQ ID NO: 116.

[0113] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0114] (a) SEQ ID NO: 117을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0115] (b) SEQ ID NO: 117을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0116] (c) SEQ ID NO: 118을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0117] (d) SEQ ID NO: 118을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0118] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0119] (a) SEQ ID NO: 119;

[0120] (b) SEQ ID NO: 120; 또는

[0121] (c) SEQ ID NO: 121.

[0122] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0123] (a) SEQ ID NO: 122를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0124] (b) SEQ ID NO: 122를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0125] (c) SEQ ID NO: 123을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0126] (d) SEQ ID NO: 123을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0127] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0128] (a) SEQ ID NO: 124;

[0129] (b) SEQ ID NO: 125; 또는

[0130] (c) SEQ ID NO: 126.

[0131] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0132] (a) SEQ ID NO: 127을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0133] (b) SEQ ID NO: 127을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0134] (c) SEQ ID NO: 128을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0135] (d) SEQ ID NO: 128을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0136] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0137] (a) SEQ ID NO: 129;

[0138] (b) SEQ ID NO: 130; 또는

[0139] (c) SEQ ID NO: 131.

[0140] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0141] (a) SEQ ID NO: 132를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0142] (b) SEQ ID NO: 132를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0143] (c) SEQ ID NO: 133을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0144] (d) SEQ ID NO: 133을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0145] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0146] (a) SEQ ID NO: 134;

[0147] (b) SEQ ID NO: 135; 또는

[0148] (c) SEQ ID NO: 136.

[0149] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0150] (a) SEQ ID NO: 137을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0151] (b) SEQ ID NO: 137을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0152] (c) SEQ ID NO: 138을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0153] (d) SEQ ID NO: 138을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0154] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0155] (a) SEQ ID NO: 139;

[0156] (b) SEQ ID NO: 140; 또는

[0157] (c) SEQ ID NO: 141.

[0158] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0159] (a) SEQ ID NO: 142를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0160] (b) SEQ ID NO: 142를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0161] (c) SEQ ID NO: 143을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0162] (d) SEQ ID NO: 143을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0163] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0164] (a) SEQ ID NO: 144;

[0165] (b) SEQ ID NO: 145; 또는

[0166] (c) SEQ ID NO: 146.

[0167] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0168] (a) SEQ ID NO: 147을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0169] (b) SEQ ID NO: 147을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0170] (c) SEQ ID NO: 148을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0171] (d) SEQ ID NO: 148을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0172] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0173] (a) SEQ ID NO: 149;

[0174] (b) SEQ ID NO: 150; 또는

[0175] (c) SEQ ID NO: 151.

[0176] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0177] (a) SEQ ID NO: 152를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0178] (b) SEQ ID NO: 152를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0179] (c) SEQ ID NO: 153을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0180] (d) SEQ ID NO: 153을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0181] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0182] (a) SEQ ID NO: 154;

[0183] (b) SEQ ID NO: 155; 또는

[0184] (c) SEQ ID NO: 156.

[0185] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0186] (a) SEQ ID NO: 157을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0187] (b) SEQ ID NO: 157을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0188] (c) SEQ ID NO: 158을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0189] (d) SEQ ID NO: 158을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0190] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0191] (a) SEQ ID NO: 159;

[0192] (b) SEQ ID NO: 160; 또는

[0193] (c) SEQ ID NO: 161.

[0194] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0195] (a) SEQ ID NO: 162를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0196] (b) SEQ ID NO: 162를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0197] (c) SEQ ID NO: 163을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0198] (d) SEQ ID NO: 163을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0199] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0200] (a) SEQ ID NO: 164;

[0201] (b) SEQ ID NO: 165; 또는

[0202] (c) SEQ ID NO: 166.

[0203] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0204] (a) SEQ ID NO: 167을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0205] (b) SEQ ID NO: 167을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0206] (c) SEQ ID NO: 168을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0207] (d) SEQ ID NO: 168을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0208] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0209] (a) SEQ ID NO: 169;

[0210] (b) SEQ ID NO: 170; 또는

[0211] (c) SEQ ID NO: 171.

[0212] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0213] (a) SEQ ID NO: 172를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0214] (b) SEQ ID NO: 172를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0215] (c) SEQ ID NO: 173을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0216] (d) SEQ ID NO: 173을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0217] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0218] (a) SEQ ID NO: 174;

[0219] (b) SEQ ID NO: 175; 또는

[0220] (c) SEQ ID NO: 176.

[0221] 일 양태에서, 본 발명은 하기 서열을 포함하는, 예를 들어 이로 구성되는 gRNA 분자를 제공한다:

[0222] (a) SEQ ID NO: 177을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[0223] (b) SEQ ID NO: 177을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[0224] (c) SEQ ID NO: 178을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[0225] (d) SEQ ID NO: 178을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr.

[0226] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은 gRNA 분자를 제공하며, 여기서:

[0227] a) gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 삽입결실이 gRNA 분자의 표적화 도메인에 상보적인 표적 서열에서 또는 그 부근에서 형성되고/되거나;

[0228] b) gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, HBG1 프로모터 영역 내의 gRNA 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역 내의 gRNA 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열을 포함하여, 예를 들어 실질적으로 모든 서열을 포함하여 결실이 생성된다. 실시형태에서, 삽입결실은 비결실 HPFH 또는 전사 인자 결합 부위의 뉴클레오타이드를 포함하지 않는다.

[0229] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은 gRNA 분자를 제공하며, 여기서 gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내에 도입될 때, 삽입결실은 집단의 세포 중 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 40%, 예를 들어 적어도 약 50%, 예를 들어 적어도 약 55%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 75%에서 gRNA 분자의 표적화 도메인에 상보적인 표적 서열에서 또는 그 부근에서 형성된다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 삽입결실은 HBG1 프로모터 영역의 적어도 하나의 뉴클레오타이드 또는 HBG2 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함한다. 실시형태에서, 집단의 세포 중 적어도 약 15%는 HBG1 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실 및 HBG2 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실을 포함한다. 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, HBG1 프로모터 영역의 적어도 하나의 뉴클레오

타이드를 포함하는 삽입결실을 포함하는 집단의 세포의 퍼센트는 HBG2 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실을 포함하는 집단의 세포의 퍼센트와 적어도 약 5%, 예를 들어 적어도 약 10%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%만큼 상이하다. 실시형태에서, 삽입결실은 차세대 시퀀싱(NGS)에 의해 측정된 바와 같다.

[0230] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은 gRNA 분자를 제공하며, 여기서 gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 태아 헤모글로빈의 발현은 상기 세포 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손에서 증가된다. 실시형태에서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내에 도입될 때, 상기 집단 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손의 집단 내의 F 세포의 퍼센트는, gRNA 분자가 도입되지 않은 세포 집단 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손의 집단 내의 F 세포의 퍼센트에 비해, 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 25%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 35%, 예를 들어 적어도 약 40%만큼 증가된다. 실시형태에서, 상기 세포 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손은 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산한다.

[0231] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은 gRNA 분자를 제공하며, 여기서 gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 어떠한 표적-외(off-target) 삽입결실도 상기 세포에서 형성되지 않으며, 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 예를 들어 어떠한 표적-외 삽입결실도 HBG1 및/또는 HBG2 프로모터 영역의 외부(예를 들어 유전자 내, 예를 들어 유전자의 코딩 영역 내)에서 형성되지 않는다.

[0232] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은 gRNA 분자를 제공하며, 여기서 gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내에 도입될 때, 어떠한 표적-외 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부(예를 들어 유전자 내, 예를 들어 유전자의 코딩 영역 내)에서 어떠한 표적-외 삽입결실도 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 세포 집단의 세포 중 약 5% 초과, 예를 들어 약 1% 초과, 예를 들어 약 0.1% 초과, 예를 들어 약 0.01% 초과에서 검출되지 않는다.

[0233] 임의의 상기 언급된 양태 및 실시형태를 포함하여 일 양태에서, 세포는 포유류, 영장류 또는 인간 세포이며(또는 세포 집단은 이를 포함하며), 예를 들어 인간 세포이며, 예를 들어 세포는 HSPC이며(또는 세포 집단은 이를 포함하며), 예를 들어 HSPC는 CD34+이며, 예를 들어 HSPC는 CD34+CD90+이다. 실시형태에서, 세포는 상기 세포가 투여되는 환자에 관하여 자가(autologous)이다. 다른 실시형태에서, 세포는 상기 세포가 투여되는 환자에 관하여 동종이계(allogeneic)이다.

[0234] 일 양태에서, 본원에 기재된 gRNA 분자, 게놈 편집 시스템(예를 들어 CRISPR 시스템) 및/또는 방법은 하기 특성 중 하나 이상을 포함하거나 초래하는, 예를 들어 본원에 기재된 바와 같은 세포에 관한 것이다:

[0235] (a) 본원에 기재된 세포 집단의 세포 중 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%는 본원에 기재된 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하며, 선택적으로 여기서 삽입결실은 표 2 내지 표 7에 열거된 삽입결실로부터 선택되며, 선택적으로 집단의 어떠한 세포도 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않는다;

[0236] (b) 본원에 기재된 세포(예를 들어 세포 집단)은 적혈구계 계통의 분화된 세포(예를 들어 적혈구)로 분화할 수 있고, 여기서, 상기 분화된 세포는 예를 들어 비변경된 세포(예를 들어 세포 집단)에 비해 증가된 수준의 태아 헤모글로빈을 나타낸다;

[0237] (c) 본원에 기재된 세포 집단은 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 여기서, 분화된 세포의 상기 집단은 예를 들어 비변경된 세포의 집단에 비해 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가진다;

[0238] (d) 본원에 기재된 세포(예를 들어 세포 집단)는 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈

구)로 분화할 수 있고, 여기서, 상기 분화된 세포(예를 들어 분화된 세포의 집단)는 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산한다;

[0239] (e) 어떠한 표적-외 삽입결실도 본원에 기재된 세포에서 형성되지 않으며, 예를 들어 어떠한 표적-외 삽입결실도 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 HBG1 및/또는 HBG2 프로모터 영역의 외부(예를 들어 유전자 내, 예를 들어 유전자의 코딩 영역 내)에서 형성되지 않는다;

[0240] (f) 어떠한 표적-외 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부(예를 들어 유전자 내, 예를 들어 유전자의 코딩 영역 내)에서 어떠한 표적-외 삽입결실도 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 본원에 기재된 세포 집단의 세포 중 약 5% 초과, 예를 들어 약 1% 초과, 예를 들어 약 0.1% 초과, 예를 들어 약 0.01% 초과에서 검출되지 않는다;

[0241] (g) 본원에 기재된 세포 또는 이의 자손은 선택적으로 청구항 제1항 내지 제22항 중 어느 한 항의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 검출함으로써 검출되는 바와 같이, 상기 세포 또는 자손이 이식된 환자에서 이식 후 16주 초과, 20주 초과 또는 24주 초과 시에, 검출 가능하며, 예를 들어 골수에서 검출 가능하거나 말초 혈액에서 검출 가능하며, 선택적으로 여기서, 삽입결실은 표 2 내지 표 7에 열거된 삽입결실로부터 선택되며, 선택적으로 삽입결실은 대형(large) 결실 삽입결실이다;

[0242] (h) 본원에 기재된 세포 집단은 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 여기서, 분화된 세포의 상기 집단은 예를 들어 비변경된 세포의 집단에 비해, 감소된 퍼센트의 겸상 적혈구(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30%, 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80% 또는 적어도 약 90% 더 낮은 퍼센트의 겸상 적혈구)를 포함하고/하거나;

[0243] (i) 본원에 기재된 세포 또는 세포 집단은 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 여기서, 분화된 세포의 상기 집단은 예를 들어 비변경된 세포의 집단에 비해, 감소된 수준(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30%, 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80% 또는 적어도 약 90% 더 낮은 수준)의 겸상 헤모글로빈(HbS)을 생산하는 세포를 포함한다.

[0244] 일 양태에서, 본 발명은:

[0245] 1) 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 양태 및 실시형태의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함), 및 예를 들어 본원에 기재된 Cas9 분자;

[0246] 2) 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 양태 및 실시형태의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함), 및 예를 들어 본원에 기재된 Cas9 분자를 인코딩하는 핵산;

[0247] 3) 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 양태 및 실시형태의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함)를 인코딩하는 핵산, 및 예를 들어 본원에 기재된 Cas9 분자;

[0248] 4) 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 양태 및 실시형태의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함)를 인코딩하는 핵산, 및 예를 들어 본원에 기재된 Cas9 분자를 인코딩하는 핵산; 또는

[0249] 5) 임의의 상기 1) 내지 4), 및 주형 핵산; 또는

[0250] 6) 임의의 상기 1) 내지 4), 및 주형 핵산을 인코딩하는 서열을 포함하는 핵산

[0251] 을 포함하는 조성물을 제공한다.

[0252] 일 양태에서, 본 발명은 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 양태 및 실시형태의 제1 gRNA 분자를 포함하며, 예를 들어 본원에 기재된 Cas9 분자를 추가로 포함하는 조성물을 제공하며, 예를 들어 여기서 Cas9 분자는 활성 또는 불활성 *S. 피오게네스* Cas9이며, 예를 들어 Cas9 분자는 SEQ ID NO: 205를 포함한다. 양태에서, Cas9 분자는: (a) SEQ ID NO: 233; (b) SEQ ID NO: 234; (c) SEQ ID NO: 235; (d) SEQ ID NO: 236; (e) SEQ ID NO: 237; (f) SEQ ID NO: 238; (g) SEQ ID NO: 239; (h) SEQ ID NO: 240; (i) SEQ ID NO: 241; (j) SEQ ID NO: 242; (k) SEQ ID NO: 243 또는 (l) SEQ ID NO: 244를 포함하며, 예를 들어 이로 구성된다.

[0253] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 제1 gRNA 분자 및 Cas9 분자는 리보핵

단백질 복합체(RNP; ribonuclear protein complex)에 존재한다.

[0254] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은 제2 gRNA 분자; 제2 gRNA 분자 및 제3 gRNA 분자; 또는 제2 gRNA 분자, 선택적으로 제3 gRNA 분자 및 선택적으로 제4 gRNA 분자를 추가로 포함하는 조성물을 제공하며, 여기서, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자는 본원에 기재된 gRNA 분자이며, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 gRNA 분자이고, 조성물의 각각의 gRNA 분자는 상이한 표적 서열에 상보적이다. 실시형태에서, 제1 gRNA 분자, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자 중 2개 이상은 동일한 유전자 또는 영역 내의 표적 서열에 상보적이다. 실시형태에서, 제1 gRNA 분자, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자는 6000개 이하의 뉴클레오타이드, 5000개 이하의 뉴클레오타이드, 500개 이하의 뉴클레오타이드, 400개 이하의 뉴클레오타이드, 300개 이하의 뉴클레오타이드, 200개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 90개 이하의 뉴클레오타이드, 80개 이하의 뉴클레오타이드, 70개 이하의 뉴클레오타이드, 60개 이하의 뉴클레오타이드, 50개 이하의 뉴클레오타이드, 40개 이하의 뉴클레오타이드, 30개 이하의 뉴클레오타이드, 20개 이하의 뉴클레오타이드 또는 10개 이하의 뉴클레오타이드 이격된(apart) 표적 서열에 상보적이다. 실시형태에서, 제1 gRNA 분자, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자 중 2개 이상은, HBG1 프로모터 영역의 표적 서열에 상보적인 표적화 도메인을 포함하는 적어도 하나의 gRNA 분자 및 HBG2 프로모터 영역의 표적 서열에 상보적인 표적화 도메인을 포함하는 적어도 하나의 gRNA 분자를 포함한다. 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 조성물은 제1 gRNA 분자 및 제2 gRNA 분자를 포함하며(예를 들어 이로 구성되며), 여기서, 제1 gRNA 분자 및 제2 gRNA 분자는: (a) 독립적으로 선택되고 예를 들어 본원에 기재된 비결실 HPFH 영역을 표적화하며, 상이한 표적 서열에 상보적이거나; (b) 표 1의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나; (c) 표 2의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나; (d) 표 3a의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나; (e) 표 3b의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나; 또는 (f) 임의의 상기 언급된 양태 및 실시형태의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이다.

[0255] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 조성물은 제1 gRNA 분자 및 제2 gRNA 분자를 포함하며, 여기서:

[0256] a) 제1 gRNA 분자는:

[0257] i) Chr11:5,249,833 내지 Chr11:5,250,237(hg38);

[0258] ii) Chr11:5,250,094 내지 5,250,237(hg38);

[0259] iii) Chr11:5,249,833 내지 5,249,927(hg38); 또는

[0260] iv) Chr11:5,250,139 내지 5,250,237(hg38)

[0261] 내의 적어도 1개 뉴클레오타이드를 포함하는(예를 들어 20개 연속 뉴클레오타이드를 포함하는) 표적 서열에 상보적이며;

[0262] b) 제2 gRNA 분자는:

[0263] i) Chr11:5,254,738 내지 Chr11:5,255,164(hg38);

[0264] ii) Chr11:5,255,022 내지 5,255,164(hg38); 또는

[0265] iii) Chr11:5,254,738 내지 5,254,851(hg38)

[0266] 내의 적어도 1개 뉴클레오타이드를 포함하는(예를 들어 20개 연속 뉴클레오타이드를 포함하는) 표적 서열에 상보적이다.

[0267] 일 양태에서, 조성물의 gRNA 분자 구성성분에 관하여, 이러한 조성물은 제1 gRNA 분자 및 제2 gRNA 분자로 구성된다.

[0268] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 각각의 상기 gRNA 분자는 예를 들어 본원에 기재된 Cas9 분자와 함께 리보핵 단백질 복합체(RNP) 내에 존재한다.

[0269] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 조성물은 주형 핵산을 포함하며, 여기서, 주형 핵산은 제1 gRNA 분자의 표적 서열에서 또는 그 부근에서 뉴클레오타이드에 상응하는 뉴클레오타

이드를 포함한다. 실시형태에서, 주형 핵산은: (a) 인간 베타 글로빈, 예를 들어 돌연변이 G16D, E22A 및 T87Q 중 하나 이상을 포함하는 인간 베타 글로빈, 또는 이의 단편; 또는 (b) 인간 감마 글로빈, 또는 이의 단편을 인코딩하는 핵산을 포함한다.

[0270] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 조성물은 전기천공에 적합한 배지에서 제제화된다.

[0271] 임의의 상기 언급된 조성물 양태 및 실시형태를 포함하여 일 양태에서, 상기 조성물의 각각의 상기 gRNA 분자는 본원에 기재된 Cas9 분자와 함께 RNP에 존재하고, 여기서, 각각의 상기 RNP는 약 10 uM 미만, 예를 들어 약 3 uM 미만, 예를 들어 약 1 uM 미만, 예를 들어 약 0.5 uM 미만, 예를 들어 약 0.3 uM 미만, 예를 들어 약 0.1 uM 미만의 농도로 존재한다. 실시형태에서, RNP는 약 1 uM의 농도로 존재한다. 실시형태에서, RNP는 약 2 uM의 농도로 존재한다. 실시형태에서, 상기 농도는 예를 들어 본원에 기재된 바와 같은 세포를 포함하는 조성물 내의 RNP의 농도이며, 선택적으로 여기서, 세포 및 RNP를 포함하는 조성물은 전기천공에 적합하다.

[0272] 일 양태에서, 본 발명은 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 하나 이상의 gRNA 분자를 인코딩하는 핵산 서열을 제공한다. 실시형태에서, 핵산은 하나 이상의 gRNA 분자를 인코딩하는 서열에 작동적으로 연결된 프로모터를 포함하며, 예를 들어 프로모터는 RNA 중합효소 II 또는 RNA 중합효소 III에 의해 인지되는 프로모터이거나, 예를 들어 프로모터는 U6 프로모터 또는 HI 프로모터이다.

[0273] 임의의 상기 언급된 핵산 양태 및 실시형태를 포함하여 일 양태에서, 핵산은 추가로, Cas9 분자, 예를 들어 임의의 SEQ ID NO: 205, SEQ ID NO: 233, SEQ ID NO: 234, SEQ ID NO: 235, SEQ ID NO: 236, SEQ ID NO: 237, SEQ ID NO: 238, SEQ ID NO: 239, SEQ ID NO: 240, SEQ ID NO: 241, SEQ ID NO: 242, SEQ ID NO: 243 또는 SEQ ID NO: 244를 포함하는, 예를 들어 이로 구성되는 Cas9 분자를 인코딩한다. 실시형태에서, 상기 핵산은 Cas9 분자를 인코딩하는 서열에 작동적으로 연결된 프로모터, 예를 들어 EF-1 프로모터, CMV IE 유전자 프로모터, EF-1α 프로모터, 유비퀴틴 C 프로모터 또는 포스포글리세레이트 키나제(PGK) 프로모터를 포함한다.

[0274] 일 양태에서, 본 발명은 임의의 상기 언급된 핵산 양태 및 실시형태의 핵산을 포함하는 벡터를 제공한다. 실시형태에서, 벡터는 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노-연관 바이러스(AAV) 벡터, 단순 포진 바이러스(HSV) 벡터, 플라스미드, 미니서클(minicircle), 나노플라스미드 및 RNA 벡터로 구성된 군으로부터 선택된다.

[0275] 일 양태에서, 본 발명은 세포(예를 들어 세포 집단) 내의 표적 서열에서 또는 그 부근에서 상기 세포를 변경시키는(예를 들어 핵산의 구조(예를 들어 서열)를 변경시키는) 방법을 제공하며, 상기 방법은 상기 세포(예를 들어 세포 집단)를:

[0276] 1) 본원에 기재된(예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의) 하나 이상의 gRNA 분자 및 예를 들어 본원에 기재된 Cas9 분자;

[0277] 2) 본원에 기재된(예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의) 하나 이상의 gRNA 분자 및 예를 들어 본원에 기재된 Cas9 분자를 인코딩하는 핵산;

[0278] 3) 본원에 기재된(예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의) 하나 이상의 gRNA 분자를 인코딩하는 핵산 및 예를 들어 본원에 기재된 Cas9 분자;

[0279] 4) 본원에 기재된(예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의) 하나 이상의 gRNA 분자를 인코딩하는 핵산 및 예를 들어 본원에 기재된 Cas9 분자를 인코딩하는 핵산;

[0280] 5) 임의의 상기 1) 내지 4), 및 주형 핵산;

[0281] 6) 임의의 상기 1) 내지 4), 및 주형 핵산을 인코딩하는 서열을 포함하는 핵산;

[0282] 7) 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물; 또는

[0283] 8) 본원에 기재된 벡터, 예를 들어 임의의 상기 언급된 벡터 양태 및 실시형태의 벡터

[0284] 와 접촉시키는(예를 들어 이것 내로 도입시키는) 단계를 포함한다.

[0285] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 일 양태에서, gRNA 분자 또는 gRNA 분자를 인코딩하는 핵산, 및 Cas9 분자 또는 이러한 Cas9 분자를 인코딩하는 핵산은 단일 조성물에서 제제화된다. 또 다른 양태에서, gRNA 분자 또는 gRNA 분자를 인코딩하는 핵산, 및 Cas9 분자 또는 이러한 Cas9 분자를 인코딩하는 핵산은 1

개 초과의 조성물에서 제제화된다. 일 양태에서, 1개 초과의 조성물은 동시에 또는 순차적으로 전달된다.

[0286] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 본원에 기재된 방법의 일 양태에서, 세포는 동물 세포이며, 예를 들어 세포는 포유류, 영장류 또는 인간 세포이며, 예를 들어 세포는 조혈 줄기 또는 전구 세포(HSPC)(예를 들어 HSPC의 집단)이며, 예를 들어 세포는 CD34+ 세포이고, 예를 들어 세포는 CD34+CD90+ 세포이다. 본원에 기재된 방법의 실시형태에서, 세포는 CD34+ 세포에 대해 농화된 세포 집단을 포함하는 조성물에 배치된다. 본원에 기재된 방법의 실시형태에서, 세포(예를 들어 세포 집단)는 골수, 동원된 말초 혈액, 또는 제대혈로부터 단리되었다. 본원에 기재된 방법의 실시형태에서, 세포는 상기 세포가 투여되는 환자에 관하여 자가 또는 동종이계, 예를 들어 자가이다.

[0287] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 본원에 기재된 방법의 일 양태에서, a) 변경은 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 초래하거나; 또는 b) 변경은 HBG1 프로모터 영역에서 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 초래한다. 방법의 양태에서, 삽입결실은 약 40개 미만 뉴클레오타이드, 예를 들어 30개 미만 뉴클레오타이드, 예를 들어 20개 미만 뉴클레오타이드, 예를 들어 10개 미만 뉴클레오타이드의 삽입 또는 결실이며, 예를 들어 단일 뉴클레오타이드 결실이다.

[0288] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 본원에 기재된 방법의 일 양태에서, 이러한 방법은 세포 집단을 초래하며, 여기서, 집단 중 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 40%, 예를 들어 적어도 약 50%, 예를 들어 적어도 약 55%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 75%가 변경되었으며, 예를 들어 삽입결실을 포함한다.

[0289] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 본원에 기재된 방법의 일 양태에서, 변경은 적혈구계 계통의 분화된 세포(예를 들어 적혈구)로 분화할 수 있는 세포(예를 들어 세포 집단)를 초래하고, 여기서, 상기 분화된 세포는 예를 들어 비변경된 세포(예를 들어 세포 집단)에 비해 증가된 수준의 태아 헤모글로빈을 나타낸다.

[0290] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 본원에 기재된 방법의 일 양태에서, 변경은 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있는 세포 집단을 초래하고, 여기서, 분화된 세포의 상기 집단은 예를 들어 비변경된 세포의 집단에 비해, 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가진다.

[0291] 임의의 상기 언급된 방법 양태 및 실시형태를 포함하여 본원에 기재된 방법의 일 양태에서, 변경은 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있는 세포를 초래하고, 여기서, 상기 분화된 세포는 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산한다.

[0292] 일 양태에서, 본 발명은 본원에 기재된 방법, 예를 들어 임의의 상기 언급된 방법 양태 및 실시형태의 방법에 의해 변경된 세포를 제공한다.

[0293] 일 양태에서, 본 발명은 본원에 기재된 방법, 예를 들어 임의의 상기 언급된 방법 양태 및 실시형태의 방법에 의해 수득 가능한 세포를 제공한다.

[0294] 일 양태에서, 본 발명은 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 또는 실시형태의 제1 gRNA 분자, 또는 본원에 기재된, 예를 들어 임의의 상기 언급된 조성물 양태 또는 실시형태의 조성물, 본원에 기재된, 예를 들어 임의의 상기 언급된 핵산 양태 또는 실시형태의 핵산, 또는 본원에 기재된, 예를 들어 임의의 상기 언급된 벡터 양태 또는 실시형태의 벡터를 포함하는 세포를 제공한다.

[0295] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 세포는 추가로, 예를 들어 본원에 기재된 Cas9 분자, 예를 들어 SEQ ID NO: 205, SEQ ID NO: 233, SEQ ID NO: 234, SEQ ID NO: 235, SEQ ID NO: 236, SEQ ID NO: 237, SEQ ID NO: 238, SEQ ID NO: 239, SEQ ID NO: 240, SEQ ID NO: 241,

SEQ ID NO: 242, SEQ ID NO: 243 또는 SEQ ID NO: 244 중 임의의 하나를 포함하는 Cas9 분자를 포함한다.

[0296] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 세포는 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 또는 실시형태의 제2 gRNA 분자, 또는 상기 gRNA 분자를 인코딩하는 핵산을 포함하거나, 포함하였거나, 포함할 것이며, 여기서, 제1 gRNA 분자 및 제2 gRNA 분자는 동일하지 않은(nonidentical) 표적화 도메인을 포함한다.

[0297] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 태아 헤모글로빈의 발현은 gRNA 분자를 포함하도록 변형되지 않은 동일한 세포 유형의 세포 또는 이의 자손에 비해, 상기 세포 또는 이의 자손(예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손)에서 증가된다.

[0298] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 세포는 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있고, 여기서, 상기 분화된 세포는 예를 들어 gRNA 분자를 포함하도록 변형되지 않은 동일한 유형의 세포에 비해, 증가된 수준의 태아 헤모글로빈을 나타낸다.

[0299] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 분화된 세포(예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구)는 예를 들어 gRNA 분자를 포함하도록 변형되지 않은 동일한 유형의 분화된 세포에 비해, 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산한다.

[0300] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 세포는 줄기세포 확장제, 예를 들어: a) (1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민; b) 메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트; c) 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀; d) (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올; 또는 e) 이들의 조합(예를 들어 (1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)으로부터 선택되는 줄기세포 확장제와 접촉되었거나, 예를 들어 생체외에서 접촉되었다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이다.

[0301] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 세포는: a) 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 또는 실시형태의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실; 또는 b) HBG1 프로모터 영역에서 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 또는 실시형태의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 본원에 기재된, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 또는 실시형태의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 포함한다. 일 양태에서, 삽입결실은 약 40개 미만 뉴클레오타이드, 예를 들어 30개 미만 뉴클레오타이드, 예를 들어 20개 미만 뉴클레오타이드, 예를 들어 10개 미만 뉴클레오타이드의 삽입 또는 결실이며, 예를 들어 삽입결실은 단일 뉴클레오타이드 결실이다.

[0302] 임의의 상기 언급된 세포 양태 및 실시형태를 포함하여 본원에 기재된 세포의 일 양태에서, 세포는 동물 세포이고, 예를 들어 세포는 포유류, 영장류 또는 인간 세포이다. 일 양태에서, 세포는 조혈 줄기 또는 전구 세포(HSPC)(예를 들어 HSPC의 집단)이며, 예를 들어 세포는 CD34+ 세포이며, 예를 들어 세포는 CD34+CD90+ 세포이다. 실시형태에서, 세포(예를 들어 세포 집단)는 골수, 동원된 말초 혈액, 또는 제대혈로부터 단리되었다. 실시형태에서, 세포는 상기 세포가 투여되는 환자에 관하여 자가이다. 실시형태에서, 세포는 상기 세포가 투여되는 환자에 관하여 동종이계이다.

[0303] 일 양태에서, 본 발명은 본원에 기재된 세포 집단, 예를 들어 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포를 포함하는 세포 집단을 제공한다. 양태에서, 본 발명은 세포 집단을 제공하며, 여기서, 집단의 세포 중 적어도 약 50%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 80%, 예를 들어 적어도 약 90%(예를 들어 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약

98% 또는 적어도 약 99%)는 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포이다. 양태에서, 세포 집단(예를 들어 세포 집단의 세포)은 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 여기서, 분화된 세포의 상기 집단은 예를 들어 동일한 유형의 비변형된 세포의 집단에 비해, 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 17%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 가진다. 양태에서, 분화된 세포의 집단의 F 세포는 1개 세포 당 평균 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산한다.

[0304] 임의의 상기 언급된 세포 집단 양태 및 실시형태를 포함하여 일 양태에서, 본 발명은: 1) 세포가 투여되는 환자의 체중 1 kg 당 적어도 1e6개 CD34+ 세포; 2) 세포가 투여되는 환자의 체중 1 kg 당 적어도 2e6개 CD34+ 세포; 3) 세포가 투여되는 환자의 체중 1 kg 당 적어도 3e6개 CD34+ 세포; 4) 세포가 투여되는 환자의 체중 1 kg 당 적어도 4e6개 CD34+ 세포; 또는 5) 세포가 투여되는 환자의 체중 1 kg 당 2e6 내지 10e6개 CD34+ 세포를 포함하는 세포 집단을 제공한다. 실시형태에서, 집단의 세포 중 적어도 약 40%, 예를 들어 적어도 약 50%(예를 들어 적어도 약 60%, 적어도 약 70%, 적어도 약 80% 또는 적어도 약 90%)는 CD34+ 세포이다. 실시형태에서, 집단의 세포 중 적어도 약 5%, 예를 들어 적어도 약 10%, 예를 들어 적어도 약 15%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%는 CD34+CD90+ 세포이다. 실시형태에서, 세포 집단은 제대혈, 말초 혈액(예를 들어 동원된 말초 혈액), 또는 골수로부터 유래되며, 예를 들어 골수로부터 유래된다. 실시형태에서, 세포 집단은 포유류 세포, 예를 들어 인간 세포를 포함하며, 예를 들어 이로 구성된다. 실시형태에서, 세포 집단은 이것이 투여되는 환자에 비해 자가이다. 다른 실시형태에서, 세포 집단은 이것이 투여되는 환자에 비해 동종이계이다.

[0305] 일 양태에서, 본 발명은 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포, 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 포함하는 조성물을 제공한다. 일 양태에서, 조성물은 약제학적으로 허용 가능한 배지, 예를 들어 동결보존에 적합한 약제학적으로 허용 가능한 배지를 포함한다.

[0306] 일 양태에서, 본 발명은 혈색소병증을 치료하는 방법을 제공하며, 상기 방법은 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포, 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단, 또는 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물을 환자에게 투여하는 단계를 포함한다.

[0307] 일 양태에서, 본 발명은 포유류에서 태아 헤모글로빈 발현을 증가시키는 방법을 제공하며, 상기 방법은 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포, 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단, 또는 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물을 환자에게 투여하는 단계를 포함한다. 양태에서, 혈색소병증은 베타-탈라쎄미아이다. 양태에서, 혈색소병증은 겸상 적혈구 질병이다.

[0308] 일 양태에서, 본 발명은 세포(예를 들어 세포 집단)를 제조하는 방법을 제공하며, 상기 방법은:

[0309] (a) 세포(예를 들어 세포 집단)(예를 들어 HSPC(예를 들어 HSPC의 집단))를 제공하는 단계;

[0310] (b) 상기 세포(예를 들어 상기 세포 집단)를, 줄기세포 확장제를 포함하는 세포 배양 배지에서 생체외에서 배양하는 단계; 및

[0311] (c) 상기 세포 내에, 예를 들어 본원에 기재된 제1 gRNA 분자, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 제1 gRNA 분자; 제1 gRNA 분자를 인코딩하는 핵산 분자; 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물; 또는 본원에 기재된 벡터, 예를 들어 임의의 상기 언급된 양태 및 실시형태의 벡터를 도입하는 단계를 포함한다. 방법의 양태에서, 단계 (c)의 상기 도입 후, 상기 세포(예를 들어 세포 집단)는 분화된 세포(예를 들어 분화된 세포의 집단), 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구계 계통의 세포 집단), 예를 들어 적혈구(예를 들어 적혈구 집단)로 분화할 수 있고, 여기서, 상기 분화된 세포(예를 들어 분화된 세포의 집단)는 예를 들어 단계 (c)를 거치지 않은 동일한 세포에 비해, 증가된 태아 헤모글로빈을 생산한다. 방법의 양태에서, 줄기세포 확장제는: a) (1r,4r)-N1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민; b) 메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트; c) 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀; d) (S)-2-(6-(2-(1H-인돌-3-일)에

틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올; 또는 e) 이들의 조합(예를 들어 (1r,4r)-N1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)이다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이다. 양태에서, 세포 배양 배지는 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)를 포함한다. 양태에서, 세포 배양 배지는 인간 인터루킨-6(IL-6)를 추가로 포함한다. 양태에서, 세포 배양 배지는 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)를 각각 약 10 ng/mL 내지 약 1000 ng/mL 범위의 농도, 예를 들어 각각 약 50 ng/mL의 농도, 예를 들어 각각 50 ng/mL의 농도로 포함한다. 양태에서, 세포 배양 배지는 인간 인터루킨-6(IL-6)를 약 10 ng/mL 내지 약 1000 ng/mL 범위의 농도, 예를 들어 약 50 ng/mL의 농도, 예를 들어 50 ng/mL의 농도로 포함한다. 양태에서, 세포 배양 배지는 줄기세포 확장제를 약 1 nM 내지 약 1 mM 범위의 농도, 예를 들어 약 1 uM 내지 약 100 nM 범위의 농도, 예를 들어 약 500 nM 내지 약 750 nM 범위의 농도로 포함한다. 양태에서, 세포 배양 배지는 줄기세포 확장제를 약 500 nM의 농도, 예를 들어 500 nM의 농도로 포함한다. 양태에서, 세포 배양 배지는 줄기세포 확장제를 약 750 nM의 농도, 예를 들어 750 nM의 농도로 포함한다.

[0312] 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (b)의 배양은 단계 (c)의 도입 전에 배양 기간을 포함하며, 예를 들어 단계 (c)의 도입 전에 배양 기간은 적어도 12시간이며, 예를 들어 약 1일 내지 약 12일의 기간 동안이며, 예를 들어 약 1일 내지 약 6일의 기간 동안이며, 예를 들어 약 1일 내지 약 3일의 기간 동안이며, 예를 들어 약 1일 내지 약 2일의 기간 동안이고, 예를 들어 약 2일의 기간 동안이다. 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (b)의 배양은 단계 (c)의 도입 후에 배양 기간을 포함하며, 예를 들어 단계 (c)의 도입 후에 배양 기간은 적어도 12시간이며, 예를 들어 약 1일 내지 약 12일의 기간 동안이며, 예를 들어 약 1일 내지 약 6일의 기간 동안이며, 예를 들어 약 2일 내지 약 4일의 기간 동안이며, 예를 들어 약 2일의 기간 동안이거나 약 3일의 기간 동안이거나 약 4일의 기간 동안이다. 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 세포 집단은 예를 들어 단계 (b)에 따라 배양되지 않은 세포에 비해, 적어도 4배, 예를 들어 적어도 5배, 예를 들어 적어도 10배 확장된다.

[0313] 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (c)의 도입은 전기천공을 포함한다. 양태에서, 전기천공은 1 내지 5 펄스, 예를 들어 1 펄스를 포함하고, 여기서, 각각의 펄스는 700 볼트 내지 2000 볼트 범위의 펄스 전압에 존재하고, 10 ms 내지 100 ms 범위의 펄스 기간을 가진다. 양태에서, 전기천공은 1 펄스를 포함하며, 예를 들어 이로 구성된다. 양태에서, 펄스(또는 1 초과의 펄스) 전압은 1500 내지 1900 볼트의 범위이며, 예를 들어 1700 볼트이다. 양태에서, 1 펄스 또는 1 초과의 펄스의 펄스 기간은 10 ms 내지 40 ms의 범위이며, 예를 들어 20 ms이다.

[0314] 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)는 인간 세포(예를 들어 인간 세포의 집단)이다. 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)는 골수, 말초 혈액(예를 들어 동원된 말초 혈액) 또는 제대혈로부터 단리된다. 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)는 골수로부터 단리되며, 예를 들어 혈색소병증을 앓고 있는 환자의 골수로부터 단리된다.

[0315] 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (a)에서 제공되는 세포 집단은 CD34+ 세포에 대해 농화된다.

[0316] 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (c)의 도입에 후속하여, 세포(예를 들어 세포 집단)는 동결보존된다.

[0317] 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (c)의 도입에 후속하여, 세포(예를 들어 세포 집단)는: a) 제1 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실; 또는 b) HBG1 프로모터 영역에서 제1 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 제1 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들

어 실질적으로 모든 서열을 포함하는 결실을 포함한다.

[0318] 방법의 임의의 상기 언급된 양태 및 실시형태를 포함하여 세포(예를 들어 세포 집단)를 제조하는 방법의 양태에서, 단계 (c)의 도입 후, 세포 집단의 세포 중 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%는 제1 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함한다.

[0319] 일 양태에서, 본 발명은 본원에 기재된, 예를 들어 임의의 상기 언급된 세포를 제조하는 방법 양태 및 실시형태에 기재된 세포(예를 들어 세포 집단)를 제조하는 방법에 의해 수득 가능한 세포(예를 들어 세포 집단)를 제공한다.

[0320] 일 양태에서, 본 발명은 인간 환자에서 혈색소병증을 치료하는 방법을 제공하며, 상기 방법은 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포; 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 포함하는 조성물을 인간 환자에게 투여하는 단계를 포함한다. 양태에서, 혈색소병증은 베타-탈라쎄미아이다. 양태에서, 혈색소병증은 겸상 적혈구 질병이다.

[0321] 일 양태에서, 본 발명은 인간 환자에서 태아 헤모글로빈 발현을 증가시키는 방법을 제공하며, 상기 방법은 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포; 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 포함하는 조성물을 상기 인간 환자에게 투여하는 단계를 포함한다. 양태에서, 인간 환자는 베타-탈라쎄미아를 갖는다. 양태에서, 인간 환자는 겸상 적혈구 질병을 갖는다.

[0322] 혈색소병증을 치료하는 방법 또는 태아 헤모글로빈 발현을 증가시키는 방법의 양태에서, 인간 환자에게 이러한 인간 환자의 체중 1 kg 당 적어도 약 1e6개 세포(예를 들어 본원에 기재된 바와 같은 세포), 예를 들어 이러한 인간 환자의 체중 1 kg 당 적어도 약 1e6개 CD34+ 세포(예를 들어 본원에 기재된 바와 같은 세포)를 포함하는 조성물이 투여된다. 혈색소병증을 치료하는 방법 또는 태아 헤모글로빈 발현을 증가시키는 방법의 양태에서, 인간 환자에게 이러한 인간 환자의 체중 1 kg 당 적어도 약 2e6개 세포(예를 들어 본원에 기재된 바와 같은 세포), 예를 들어 이러한 인간 환자의 체중 1 kg 당 적어도 약 2e6개 CD34+ 세포(예를 들어 본원에 기재된 바와 같은 세포)를 포함하는 조성물이 투여된다. 혈색소병증을 치료하는 방법 또는 태아 헤모글로빈 발현을 증가시키는 방법의 양태에서, 인간 환자에게 이러한 인간 환자의 체중 1 kg 당 약 2e6개 세포(예를 들어 본원에 기재된 바와 같은 세포), 예를 들어 이러한 인간 환자의 체중 1 kg 당 약 2e6개 CD34+ 세포(예를 들어 본원에 기재된 바와 같은 세포)를 포함하는 조성물이 투여된다. 혈색소병증을 치료하는 방법 또는 태아 헤모글로빈 발현을 증가시키는 방법의 양태에서, 인간 환자에게 이러한 인간 환자의 체중 1 kg 당 적어도 약 3e6개 세포(예를 들어 본원에 기재된 바와 같은 세포), 예를 들어 이러한 인간 환자의 체중 1 kg 당 적어도 약 3e6개 CD34+ 세포(예를 들어 본원에 기재된 바와 같은 세포)를 포함하는 조성물이 투여된다. 혈색소병증을 치료하는 방법 또는 태아 헤모글로빈 발현을 증가시키는 방법의 양태에서, 인간 환자에게 이러한 인간 환자의 체중 1 kg 당 약 3e6개 세포(예를 들어 본원에 기재된 바와 같은 세포), 예를 들어 이러한 인간 환자의 체중 1 kg 당 약 3e6개 CD34+ 세포(예를 들어 본원에 기재된 바와 같은 세포)를 포함하는 조성물이 투여된다. 혈색소병증을 치료하는 방법 또는 태아 헤모글로빈 발현을 증가시키는 방법의 양태에서, 인간 환자에게 이러한 인간 환자의 체중 1 kg 당 약 2e6 내지 약 10e6개 세포(예를 들어 본원에 기재된 바와 같은 세포), 예를 들어 이러한 인간 환자의 체중 1 kg 당 약 2e6 내지 약 10e6개 CD34+ 세포(예를 들어 본원에 기재된 바와 같은 세포)를 포함하는 조성물이 투여된다.

[0323] 일 양태에서, 본 발명은: 약제로서 사용하기 위한, 본원에 기재된 gRNA 분자, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 gRNA 분자; 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물; 본원에 기재된 핵산, 예를 들어 임의의 상기 언급된 핵산 양태 및 실시형태의 핵산; 본원에 기재된 벡터, 예를 들어 임의의 상기 언급된 벡터 양태 및 실시형태의 벡터; 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포; 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 제공한다.

[0324] 일 양태에서, 본 발명은: 약제의 제조에 사용하기 위한, 본원에 기재된 gRNA 분자, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 gRNA 분자; 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물; 본원에 기재된 핵산, 예를 들어 임의의 상기 언급된 핵산 양태 및 실시형태의 핵산; 본원에 기재된 벡터, 예를 들어 임의의 상기 언급된 벡터 양태 및 실시형태의 벡터; 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포; 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 제공한다.

[0325] 일 양태에서, 본 발명은: 질병의 치료에 사용하기 위한, 본원에 기재된 gRNA 분자, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 gRNA 분자; 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물; 본원에 기재된 핵산, 예를 들어 임의의 상기 언급된 핵산 양태 및 실시형태의 핵산; 본원에 기재된 벡터, 예를 들어 임의의 상기 언급된 벡터 양태 및 실시형태의 벡터; 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포; 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 제공한다.

[0326] 일 양태에서, 본 발명은: 질병의 치료에 사용하기 위한, 본원에 기재된 gRNA 분자, 예를 들어 임의의 상기 언급된 gRNA 분자 양태 및 실시형태의 gRNA 분자; 본원에 기재된 조성물, 예를 들어 임의의 상기 언급된 조성물 양태 및 실시형태의 조성물; 본원에 기재된 핵산, 예를 들어 임의의 상기 언급된 핵산 양태 및 실시형태의 핵산; 본원에 기재된 벡터, 예를 들어 임의의 상기 언급된 벡터 양태 및 실시형태의 벡터; 본원에 기재된 세포, 예를 들어 임의의 상기 언급된 세포 양태 및 실시형태의 세포; 또는 본원에 기재된 세포 집단, 예를 들어 임의의 상기 언급된 세포 집단 양태 및 실시형태의 세포 집단을 제공하며, 여기서, 질병은 혈색소병증, 예를 들어 베타-탈라세미아 또는 겸상 적혈구 질병이다.

도면의 간단한 설명

[0327] **도 1:** 편집 후 7일째 HbF 유도이다. 시험된 각각의 gRNA 표적 서열에 대해, 모크(mock) 형질감염을 기초로 백그라운드 수준에 대해 보정된(corrected) 유도된 HbF 발현을 갖는 세포의 퍼센트가 표준 편차를 표시하는 오차 막대와 함께 평균으로서 제시된다. BCL11A의 엑손 2에 대한 gRNA G8은 양성 대조군으로서 역할을 한다. 17%에서 점선은 분석을 위해 선택된 역치 수준을 표시한다. 범례에서 표시된 바와 같이 다양한 회색 음영법은 HBG1 또는 HBG2 표적 유전자좌에서의 편집율에 대한 HbF 유도율에 관한 것이다.

도 2: HBG1 표적 유전자좌에서의 편집 효율이다. 시험된 각각의 gRNA에 대해, NGS에 의해 검출된 삽입결실의 퍼센트는 표준 편차를 표시하는 오차 막대와 함께 평균으로서 제시된다. BCL11A의 엑손 2에 대한 gRNA G8은 양성 대조군으로서 역할을 한다. 어떠한 NGS 데이터도 수득되지 않은 2개의 가이드는 화살표에 의해 표시된다.

도 3: HBG2 표적 유전자좌에서의 편집 효율이다. 시험된 각각의 gRNA에 대해, NGS에 의해 검출된 삽입결실의 퍼센트는 표준 편차를 표시하는 오차 막대와 함께 평균으로서 제시된다. BCL11A의 엑손 2에 대한 gRNA G8은 양성 대조군으로서 역할을 한다. 어떠한 NGS 데이터도 수득되지 않은 16개의 가이드는 화살표에 의해 표시된다.

도 4: HBG1 프로모터 영역에서 고성능 gRNA 표적 서열(예를 들어 7일차에 17% 초과의 HbF 상향조절), 공지된 비결실 HPFH 다형성 및 전사 인자 결합 부위의 위치의 개요이다. 도면은 SEQ ID NOS 293~312를 각각 출현(appearance)의 순서로 개시한다.

도 5: HBG2 프로모터 영역에서 고성능 gRNA 표적 서열(예를 들어 7일차에 17% 초과의 HbF 상향조절), 공지된 비결실 HPFH 다형성 및 전사 인자 결합 부위의 위치의 개요이다. 도면은 SEQ ID NOS 313~332를 각각 출현의 순서로 개시한다.

도 6: NGS 및 유세포분석에 의해 평가된 바와 같이, 상이한 Cas9 변이체에 의한 CD34+ HSPC에서 표적화된 B2M 유전자좌에서의 편집 효율이다. NLS는 SV40 NLS이며; His6(SEQ ID NO: 247) 또는 His8(SEQ ID NO: 248)은 각각 6(SEQ ID NO: 247) 또는 8(SEQ ID NO: 248) 히스티딘 잔기를 지칭하며; TEV는 담배 식각 바이러스 절단 부위이고; Cas9는 야생형 S. 피오게네스 Cas9이고 - 돌연변이 또는 변이체는 표시된 바와 같다.

도 7: 표시된 표적화 도메인의 sgRNA를 함유하는 RNP를 이용한 HSPC의 전기천공에 뒤이은 적혈구계 분화 후 7일(검정색 막대), 14일(연회색 막대) 또는 21일(진회색 막대)째에 유세포분석에 의한 HbF 양성 세포의 검출 및 정량화이다. 각각의 시점에서 sgRNA로 처리되지 않은 대조군 배양물에 대한 퍼센트 HbF+ 세포는 차감되었다(subtracted). 평균 + 표준 편차가 표시된다(n=2 기술적 복제물).

도 8: 표시된 표적화 도메인의 sgRNA를 함유하는 RNP를 이용한 HSPC의 전기천공에 뒤이은 적혈구계 분화 후 7일(열린 검정색 막대), 14일(열린 연회색 막대) 또는 21일(열린 진회색 막대)째에 유세포분석에 의한 HbF 양성 세포의 검출 및 정량화이다. 각각의 시점에서 sgRNA로 처리되지 않은 대조군 배양물에 대한 퍼센트 HbF+ 세포는 차감되었다. 2명의 독립적인 세포 공여자의 평균(막대)이 각각의 공여자에 대한 값과 함께 제시된다(원형은 제1 공여자이고, 삼각형은 제2 공여자임).

도 9. 표시된 반응으로부터의 PCR 생성물: 실시예에서 기재된 바와 같이, 표시된 표적화 도메인의 sgRNA를 함유하는 RNP를 이용한 전기천공 후의 세포 또는 sgRNA로 처리되지 않은 대조군 세포로부터의 P1, P2 또는 P3의 시

각화이다. 예상된 생성물은 하기와 같다. P1: 야생형/소형(small) 삽입결실 대립유전자 또는 4.9 kb 역위(inversion) 대립유전자에 대해 7.7 kb, 4.9 kb 결실된 대립유전자에 대해 2.8 kb. P2: 야생형/소형 삽입결실 대립유전자에 대해 3.8 kb, 4.9 kb 결실 또는 4.9 kb 역위 대립유전자에 대해 생성물 없음. P3: 4.9 kb 역위 대립유전자에 대해 1.8 kb, 야생형/소형 삽입결실 대립유전자 또는 4.9 kb 결실 대립유전자에 대해 생성물 없음. L은 DNA 참조 사다리이다. *은 이 밴드로부터의 DNA가 단리되었고 차세대 시퀀싱을 거쳤음을 표시한다.

도 10: 4.9 kb 결실을 정량화하기 위해 디지털 방울 PCR(digital droplet PCR) 검정법에 대한 프라이머 및 프로브 결합 부위의 게놈 위치를 표시하는 도식도이다. 프라이머(5.2 kb Fwd 및 5.2 kb Rev) 및 프로브(FAM 프로브)는 HBG2의 다운스트림 및 HBG1의 업스트림에서 유전자간(intergenic) 부위에 결합한다. 프로브는 HBG2의 업스트림에 제2 결합 부위를 갖지만, 해당 영역은 프라이머에 의해 결합되지 않는다. HBG1 및 HBG2 프로모터 내의 표적화 도메인이 위치하는 영역이 표시된다. 2개의 표적화 도메인 영역을 개입하는 서열이 결실된다면, HBG1과 HBG2 사이의 프라이머/프로브 결합 부위는 상실될 것이다.

도 11: GCR-0067 표적화 도메인의 sgRNA를 함유하는 RNP를 이용한 전기천공 후 세포 시료에 대한 HSPC 하위집단에 대한 소팅 도식도이다. 표시된 세포 표면 마커로 표적화된 면역염색 후 세포 형광의 도트 플롯이 제시된다. 하기 집단이 제시된 바와 같이 소팅되었다: P5는 CMP(CD34+CD45RA-CD38+)이며, P9는 MPP(CD34+CD45RA-CD38-CD90-CD49f-)이며, P10은 ST-HSC(CD34+CD45RA-CD38-CD90-CD49f+)이고, P11은 LT-HSC(CD34+CD45RA-CD38-CD90+CD49f+)이다. 총 CD34+ 세포가 또한 소팅되었다(제시되지 않음).

도 12: GCR-0067 표적화 도메인의 sgRNA를 함유하는 RNP를 이용한 전기천공 후 소팅된 HSPC 하위집단의 편집 퍼센트이다. HBG1 삽입결실 및 HBG2 삽입결실은 HBG1 또는 HBG2 프로모터 영역 표적화 도메인 각각에서 또는 그 부근에서 PCR 앰플리콘의 차세대 시퀀싱에 의해 식별된 소형 삽입 및 결실의 퍼센트를 표시한다(이전에 기재된 4.9 kb 결실 또는 역위를 갖는 대립유전자는 증폭되지 않음을 주목함). HBG1-HBG2 결실은 실시예에 기재된 디지털 방울 PCR 검정법에 의해 결정된 바와 같이 4.9 kb 결실을 갖는 대립유전자의 퍼센트를 표시한다. 총 편집은, HBG1-HBG2 결실이 없는 퍼센트에 더해진 HBG1-HBG2 결실의 퍼센트에 HBG2 삽입결실 퍼센트를 곱하여 계산된 근사치이다.

도 13: 도 13a는 GCR-0067의 표적화 도메인을 포함하는 sgRNA의 도입 후 언급된 세포 유형에서 HBG1 유전자좌에서 관찰된 모든 삽입결실의 합계를 보여준다. 삽입결실은 각각의 막대의 상부에서 가장 빈번하게 관찰된 삽입결실과 함께 배열된다. HBG1과 HBG2 유전자좌 사이에 대형 4.9 kb 결실을 포함하는 세포의 분획은 이러한 검정법에서 정량화되지 않는다. 가장 우세한 삽입결실의 각각의 막대 내의 수는, 야생형 게놈 서열로부터 뉴클레오타이드 차이의 수를 표시한다(-는 결실을 가리키고; +는 삽입을 가리킴). 도 13b는 GCR-0067의 표적화 도메인을 포함하는 sgRNA의 도입 후 언급된 세포 유형에서 HBG2 유전자좌에서 관찰된 모든 삽입결실의 합계를 보여준다. 삽입결실은 각각의 막대의 상부에서 가장 빈번하게 관찰된 삽입결실과 함께 배열된다. HBG1과 HBG2 유전자좌 사이에 대형 4.9 kb 결실을 포함하는 세포의 분획은 이러한 검정법에서 정량화되지 않는다. 가장 우세한 삽입결실의 각각의 막대 내의 수는, 야생형 게놈 서열로부터 뉴클레오타이드 차이의 수를 표시한다(-는 결실을 가리키고; +는 삽입을 가리킴). CMP는 CD34+CD45RA-CD38+ 세포이며; MPP는 CD34+CD45RA-CD38-CD90-CD49f- 세포이며; ST-HSC는 CD34+CD45RA-CD38-CD90-CD49f+ 세포이고; LT-HSC는 CD34+CD45RA-CD38-CD90+CD49f+ 세포이다.

도 14: 표시된 표적화 도메인의 sgRNA를 함유하는 RNP, 또는 sgRNA로 처리되지 않은 대조군 배양물을 이용한 전기천공 후 표시된 하위유형, CFU-GEMM(진회색 막대), CFU-G/M/GM(중회색 막대) 또는 BFU-E/CFU-E(연회색 막대)에 상응하는 콜로니의 퍼센트이다. 평균 +/- 표준 편차가 표시된다(n=2명의 독립적인 공여자).

도 15: GCR-0067 표적화 도메인의 sgRNA를 함유하는 RNP, 또는 sgRNA로 처리되지 않은 대조군 배양물을 이용한 전기천공 후 화합물 4를 포함하는 배지에서 표시된 바와 같은 총 핵화된(nucleated) 세포(TNC), CD34 양성 세포(CD34+), 및 CD34와 CD90 이중 양성 세포(CD34+CD90+)의 배수 증식이다. 평균 +/- 표준 편차가 표시된다(n=2명의 독립적인 공여자). 3명의 독립적인 세포 공여자의 평균(막대)이 각각의 공여자에 대한 값과 함께 제시된다(정사각형은 공여자 A이며, 삼각형은 공여자 B이고, 원형은 공여자 C임). 편집된 배양물과 대조군 배양물 사이의 차이는 독립 표본 t-검정(unpaired t-test)(GraphPad Prism)에 의해 유의하지 않았다(ns).

도 16: 유세포분석에 의한 세포 집단의 대표적인 게이팅(gating)이다. 표시된 세포 표면 마커, 또는 이소형(isotype) 대조군(이소형)으로 표적화된 면역염색 후 세포 형광의 도트 플롯이 제시된다. 볼드체 박스로 표시된 게이트는 표시된 집단의 퍼센트를 정량화하는 데 사용되었고, 이소형 대조군 표지된 세포를 배제하도록 설정되었다. DAPI 음성으로 예비-게이팅되고 세포성 포워드 스캐터(scatter) 및 사이드 스캐터 게이트 내에 있는 살아있는 세포만 제시되고, Cas9 단독으로 전기천공된 공여자 C로부터 유래된다.

도 17: GCR-0067 표적화 도메인의 sgRNA를 함유하는 RNP(검정색 막대) 또는 sgRNA로 처리되지 않은 대조군 배양물(회색 막대)을 이용한 전기천공 후, 유세포분석에 의해 평가된 바와 같이, 표시된 세포성 세포 표면 표현형을 갖는 세포의 퍼센트이다. 세포를 화합물 4를 포함하는 배지 내에서 RNP의 전기천공 후 2일째에 확장시키고, 도 16에 기재된 바와 같이 유세포분석에 의해 평가하였다. 평균 + 표준 편차가 표시된다(n=3명의 독립적인 공여자). 독립 표본 t-검정(GraphPad Prism)에 의해, 주어진 집단에 대해 편집된 세포와 비편집된 세포 사이에 유의한 차이가 존재하지 않았다.

도 18a. Cas9를 이용하여 편집되고 sgRNA가 없는 겸상 적혈구 질병 환자(SCD1) 모크로부터 유래된 모크 편집 HSC에서 글로빈 하위단위의 CE-MS 정량화이다. 세포가 적혈구계 계통으로 분화된 후, 세포는 정상 수준의 a-글로빈, 겸상 동형접합성(homozygosity)으로 인한 제로 수준의 정상적인 b-글로빈, 높은 수준의 겸상 b-글로빈 하위단위 및 낮은 수준의 태아 g-글로빈을 보여주었다.

도 18b. 겸상 적혈구 환자(SCD1)로부터 유래된 게놈-편집된 HSPC에서 글로빈 하위단위의 CE-MS 정량화이다. HSC를 편집한 후, 시료 환자로부터 유래된 적혈구계 세포는 태아 g-글로빈의 40% 상향 조절, 및 겸상 b-글로빈 하위단위의 병발하는(concurrent) 50% 하향조절을 실증하였다.

도 19. 유전자 편집된 세포의 생착을 연구하기 위한 도식적 프로토콜이다. Cas9, 및 CR001128의 표적화 도메인을 포함하는 sgRNA(sg1128)를 이용한 HSC 유전자-편집의 이식이다. $5x10^5$개 인간 CD34+ 세포를 gRNA로 모크 편집하거나 sg1128로 유전자 편집하였으며, 뒤이어 2 Gy 방사선 조사된 NOD.Cg-*Prkdc*^{scid}*I12rg*^{tm1Wjl}/SzJ(NSG) 수여체(recipient) 내로 주사하였다. 4, 8, 12, 16주째에 마우스에서 채혈하고, 이식-후 16주째에 골수 세포를 수합하였다.

도 20. 이식-후 16주째에 도 19에 제시된 실험 프로토콜을 사용한 골수 재구성이다.

도 21. 다수의 시점에서 도 19에 제시된 실험 프로토콜을 사용한, 말초 혈액 및 골수 내 골수계, B 및 T 림프계 세포의 재구성이다. N=5/그룹이며, 데이터는 4회의 독립적인 실험에서 최소 내지 최대를 보여준다.

도 22. BCL11A 유전자의 적혈구계-특이적 인핸서 영역으로부터의 gRNA(sg-G1128; sg1128로도 지칭됨)와 비교하여, 감마 글로빈 프로모터 영역으로부터의 sgRNA(sg-G0008, sg-G0051, sg-G0010, sg-G0048, sg-G0067)를 이용하여 편집된 HSC의 줄기세포 기능을 평가하기 위한 이식 연구의 개략도이다. 전기천공 후 24시간째에 세포를 배양물 내에 놔두었다. 전기천공 전과 후 둘 모두에서 배양 조건은 StemSpan SFEM+IL6, SCF, TPO, Flt3L; 750 nM 화합물 4이었다.

도 23: NSG 마우스에서 20주에 걸친 인간 생착 및 계통 분석이다. **도 23a**) 18주에 걸친 말초 혈액 키메라 현상(chimerism)이며; **도 23b**) 18주째에 말초 혈액에서의 계통 분포이다. 골수 분석: **도 23c)** 9주째에 인간 세포의 골수 분석이며; **도 23d**) 9주째에 골수 내에서 인간 CD45+ 생착 및 인간 세포의 계통 분포이며; **도 23e)** 생착 후 20주째에 골수 내에서 인간 CD45+ 생착이고; **도 23f)** 20주째에 골수 내에서 생착된 세포의 계통 분포이다.

도 24a. BCL11A 유전자의 적혈구계-특이적 인핸서 영역으로부터의 sgRNA(sg1128)와 비교하여, 감마 글로빈 프로모트(promote) 영역과 상동성인 sgRNA(sg-G0008, sg-G0051, sg-G0010, sg-G0048, sg-G0067)를 이용하여 편집된 HSC의 생착 효율이다. 이식 후 8주째에 NSG 마우스에서 인간 세포 생착을 보여준다. N=10/그룹이며, 3회의 독립적인 실험이다. 그래프는 풀링된(pooled) 데이터를 보여준다.

도 24b. BCL11A 유전자의 적혈구계-특이적 인핸서 영역으로부터의 sgRNA(sg1128)와 비교하여, 감마 글로빈 프로모트 영역과 상동성인 sgRNA(sg-G0008, sg-G0051, sg-G0010, sg-G0048, sg-G0067)를 이용하여 편집된 HSC의 생착 효율이다. 이식 후 20주째에 NSG 마우스에서 인간 세포 생착을 보여준다. N=10/그룹이며, 3회의 독립적인 실험이다. 그래프는 풀링된 데이터를 보여준다.

도 25. 유전자-편집된 CD34+ 세포 이식된 NSG 마우스의 다계통 재구성이다. N=10/그룹이며, 데이터는 1개의 대표적인 실험으로부터의 것이다.

도 26. 이식-전과 이식-후에 높은 편집 효율이 유지되었다. "이식-전": 편집 시 그러나 이식 전에 NGS에 의해 측정된 바와 같은 인간 CD34+ 세포에서 개별 sgRNA의 편집 효율 및 삽입결실 패턴이다. "이식-후 8주": 마우스에서 골수 이식 후 8주째에 NGS에 의해 측정된 바와 같은 인간 CD34+ 세포에서 편집 효율 및 삽입결실 패턴이다. "이식-후 20주": 마우스에서 골수 이식 후 20주째에 NGS에 의해 측정된 바와 같은 인간 CD34+ 세포에서 편집 효율 및 삽입결실 패턴이다. 전기천공을 1개 그룹 당 3벌 중복으로 수행하였으며, 데이터는 1개의 대표

적인 실험으로부터의 것이다. 이 도면과 관련하여 사용된 바와 같이, "삽입결실"은 200개 미만 nt의 모든 삽입 결실의 합계를 지칭하고; "대형(large) 결실"은 각각의 gRNA에 대한 예측된 HBG1 결합 부위와 HBG2 결합 부위 사이에서 서열의 결실을 지칭한다.

도 27. RNP를 이용한 편집-후 CD34+ 세포의 NGS 분석이다. 특이적인 영역을 표적화하는 sgRNA는 x-축 상에 표시된다. 도 27a: RNP의 전기천공-후 2일째에 CD34+ 세포의 NGS 분석이다. 도 27b: 이식-후 9주째에(27b) 편집된 골수 CD34+ 세포를 이식한 NSG 마우스로부터의 전체 골수의 NGS 분석이다. 도 27c: 이식-후 20주째에 편집된 골수 CD34+ 세포를 이식한 NSG 마우스로부터의 전체 골수의 NGS 분석이다. 삽입 삽입결실은 검정색으로 표시되며, 결실 삽입결실(gRNA sg-G51, sg-G48 및 sg-G67 각각에 대한 결합 HBG1과 HBG2 표적 서열 사이의 영역의 절제(excision)를 포함하는 대형 결실은 배제함)은 회색으로 표시된다. 총 % 편집은 막대의 높이로 표시된다. N=10이며, 데이터는 1회의 독립적인 실험으로부터의 평균±SEM으로서 제시된다.

도 28. 유전자-편집된, 장기간 생착된 인간 HSC는 적혈구계 세포 분화 시 증가된 수준의 HbF를 생성할 수 있었다. 8-주 또는 20-주 이식된 NSG 마우스의 골수로부터 $5x10^4$개 인간 CD34+ 세포를 소팅하였다. 소팅된 세포를 적혈구계 분화 배지 내로 14 내지 21일 동안 접종하였다. 배양물 중 성숙한 적혈구를 HbF 발현에 대해 검정하였고, HbF+ 세포의 수를 유세포분석에 의해 계수하였다. 모크 대조군은 gRNA 없는 Cas9로 편집되고 유전자-편집된 대조군과 동일한 방식으로 NSG 마우스 내로 이식된 CD34+ 세포를 나타낸다. N=10/그룹이며, 3회의 독립적인 실험이다.

도 29. HBG1 및/또는 HBG2 가이드 RNA에 대한 표적-외 활성을 Cas9-과발현 HEK-293 세포에서 dsDNA 올리고-삽입 방법을 사용하여 평가하였다. 검출된 표적-적중 부위(열린 원형) 및 잠재적인 표적-외 부위(닫힌 원형)가 표시되며; y-축은 검출 빈도를 표시한다. 모든 gRNA를, CRxxxxxx 식별자에 의해 표시된 표적화 도메인과 함께 dgRNA 포맷에서 시험하였다.

도 30. HBG1 및/또는 HBG2 가이드 RNA에 대한 표적-외 활성을 Cas9-과발현 HEK-293 세포에서 dsDNA 올리고-삽입 방법을 사용하여 평가하였다. 검출된 표적-적중 부위(열린 원형) 및 잠재적인 표적-외 부위(닫힌 원형)가 표시되며; y-축은 검출 빈도를 표시한다. gRNA를, CRxxxxx 식별자에 의해 표시된 표적화 도메인과 함께 dgRNA 포맷에서, 또는 Gxxxxxx 식별자에 의해 표시된 표적화 도메인과 함께 sgRNA 포맷에서 시험하였다.

도 31a. 유전자-편집 시, 건강한 개체의 동원된 말초 혈액으로부터 유래된 세포의 CD34+ 세포 수치이다. 세포를 0일차에 해동시키고, 3일차에 전기천공 전에 3일 동안 배양하였다. 전기천공 후 10일에 걸쳐 ISHAGE에 의한 CD34+ 세포 수의 계수이다. 2회의 독립적인 실험을 2벌 중복으로 수행하였다. 총 N=5이었다. 그래프는 1회의 실험으로부터의 데이터를 평균±SEM과 함께 보여준다.

도 31b. 유전자-편집 시, 건강한 개체의 동원된 말초 혈액으로부터 유래된 세포의 CD34+ 세포 확장이다. 세포를 0일차에 해동시키고, 3일차에 전기천공 전에 3일 동안 배양하였다. 편집 시 3일차, 7일차 및 10일차에 총 단핵 세포의 확장이다. 2회의 독립적인 실험을 2벌 중복으로 수행하였다. 총 N=5이었다. 그래프는 1회의 실험으로부터의 데이터를 평균±SEM과 함께 보여준다.

도 31c. 유전자-편집 시, 건강한 개체의 동원된 말초 혈액으로부터 유래된 세포의 CD34+ 세포 생존력(viability)이다. 세포를 0일차에 해동시키고, 3일차에 전기천공 전에 3일 동안 배양하였다. 편집-후 동일한 시점에서 단핵 세포의 생존력이다. 2회의 독립적인 실험을 2벌 중복으로 수행하였다. 총 N=5이었다. 그래프는 1회의 실험으로부터의 데이터를 평균±SEM과 함께 보여준다.

도 32a. 건강한 개체의 말초 혈액으로부터 동원된 CD34+ 세포에서 sg1128 및 sg0067의 편집 효율이다. sg1128(ESH의 BCL11A +58 영역을 표적화함) 및 sg0067(HbG-1 및 HbG-2 유전자 클러스터를 표적화함)을 사용한 편집 시 NGS에 의해 포착된 삽입결실의 퍼센트가 제시된다. 이 그래프는 또한, 단지 소형 삽입결실로서 sg1128에 대한 총 편집 효율이 이러한 sgRNA에 의해 발생되었음을 표시한다. 5회 초과의 독립적인 실험을 2벌 중복 또는 3벌 중복으로 수행하였으며, n=2~3/실험이다.

도 32b. 건강한 개체의 말초 혈액으로부터 동원된 CD34+ 세포에서 sg1128 및 sg0067의 편집 효율이다. sg0067(HbG-1 및 HbG-2 유전자 클러스터를 표적화함)을 사용한 편집 시 NGS에 의해 포착된 삽입결실의 퍼센트이다. sg0067의 총 편집 효율 및 편집 패턴이 제시된다. sg0067의 편집 패턴은 5 kb 대형 결실(검정색 막대로 표시됨) 및 더 소형의 삽입결실(회색 막대로 표시됨)로 구성된다. 5회 초과의 독립적인 실험을 2벌 중복 또는 3벌 중복으로 수행하였으며, n=2~3/실험이다.

도 33. BCL11A의 CRISPR 넉다운 또는 HBG1/2 영역에서 삽입결실/결실 형성 시 환자 시료로부터의 적혈구계 세포에서 g-글로빈 전사체의 배수 변화이다. 건강한 공여자의 동원된 말초 혈액으로부터 유래된 CD34+ 세포를 이전의 절차에 기재된 바와 같이 *시험관내에서* CRISPR에 의해 편집하고 적혈구계 계통으로 분화시켰다. 적혈구계 분화의 11일차에, 세포를 배양물로부터 수합하고, qPCR을 받게 하여 g-글로빈 및 b-글로빈 전사체를 측정하고, GAPDH에 대해 정규화하였다(normalizing). 실험을 독립적으로 2회, 각각 2벌 중복으로 수행하였으며, n=2~3/실험이다. 데이터는 하나의 연구로부터의 풀링된 공여자의 평균±SEM을 보여준다.

도 34. 유전자 편집 시 건강한 개체로부터의 HbF+ 세포의 계수이다. 건강한 공여자의 동원된 말초 혈액으로부터 유래된 CD34+ 세포를 이전의 절차에 기재된 바와 같이 *시험관내에서* CRISPR에 의해 편집하고 적혈구계 계통으로 분화시켰다. 적혈구계 분화의 10일차에, 세포를 항-HbF-FITC 항체로 염색하여, HbF+ 세포를 유세포분석에 의해 계수하였다. 실험을 독립적으로 2회, 각각 2벌 중복으로 수행하였으며, n=2~3/실험이다. 데이터는 하나의 연구로부터의 평균±SEM을 보여준다.

도 35a. 유전자-편집 시 겸상 적혈구 질병 개체의 말초 혈액으로부터 유래된 CD34+ 세포의 확장 및 생존력이다. 세포를 전기천공 전 6 내지 10일 동안 배양하였다. D0는 전기천공일을 지칭한다. 전기천공 후 10일에 걸쳐 ISHAGE에 의한 CD34+ 세포의 절대 수치가 제시된다. N=4이며, 데이터는 평균±SEM을 보여준다. 4회의 독립적인 실험을 2벌 중복으로 수행하였다. 막대는 각각의 시점에서 좌측으로부터 우측까지 모크, sg1128, sg0067이다.

도 35b. 유전자-편집 시 겸상 적혈구 질병 개체의 말초 혈액으로부터 유래된 CD34+ 세포의 확장 및 생존력이다. 세포를 전기천공 전 6 내지 10일 동안 배양하였다. D0는 전기천공일을 지칭한다. 전기천공 후 10일에 걸쳐 ISHAGE에 의한 CD34+ 세포의 퍼센트가 제시된다. N=4이며, 데이터는 평균±SEM을 보여준다. 4회의 독립적인 실험을 2벌 중복으로 수행하였다. 막대는 각각의 시점에서 좌측으로부터 우측까지 모크, sg1128, sg0067이다.

도 35c. 유전자-편집 시 겸상 적혈구 질병 개체의 말초 혈액으로부터 유래된 CD34+ 세포의 확장 및 생존력이다. 세포를 전기천공 전 6 내지 10일 동안 배양하였다. D0는 전기천공일을 지칭한다. 편집 시 3일차, 7일차 및 10일차에 총 단핵 세포의 확장이 제시된다. N=4이며, 데이터는 평균±SEM을 보여준다. 4회의 독립적인 실험을 2벌 중복으로 수행하였다. 막대는 각각의 시점에서 좌측으로부터 우측까지 모크, sg1128, sg0067이다.

도 35d. 유전자-편집 시 겸상 적혈구 질병 개체의 말초 혈액으로부터 유래된 CD34+ 세포의 확장 및 생존력이다. 세포를 전기천공 전 6 내지 10일 동안 배양하였다. D0는 전기천공일을 지칭한다. 편집-후 3일차, 7일차 및 10일차에 단핵 세포의 생존력이 제시된다. N=4이며, 데이터는 평균±SEM을 보여준다. 4회의 독립적인 실험을 2벌 중복으로 수행하였다. 막대는 각각의 시점에서 좌측으로부터 우측까지 모크, sg1128, sg0067이다.

도 36. 겸상 적혈구 질병 환자 시료로부터의 CD34+ 세포에서 sg1128 및 sg0067의 편집 효율이다. sg0067의 편집 패턴은 대형 결실(회색) 및 소형 삽입결실의 합계(검정색)에 의해 표시되었다. 데이터는 4명의 상이한 겸상 적혈구 질병 환자(SCD1 내지 4)로부터의 CD34+ 세포를 이용하여 2벌 중복으로 수행된 4회의 독립적인 편집 실험의 평균±SEM을 보여준다.

도 37. 콜로니-형성 단위 검정법에 의해 측정된 바와 같이 HSPC의 *시험관내* 다계통 분화 능력(capacity)이다. BFU-E는 블라스트-형성 단위(blast-forming unit), 적혈구계이며; CFU-GM은 콜로니-형성 단위, 과립구, 단핵구이고; CFU-GEMM은 콜로니-형성 단위, 과립구, 적혈구계, 단핵구, 거핵구이다. 그래프는 3명의 상이한 겸상 적혈구 질병 환자(SCD1 내지 3)로부터의 CD34+ 세포로부터 3회의 독립적인 실험을 보여준다. 실험을 3벌 중복으로 수행하였다. 데이터는 평균±SEM을 나타낸다.

도 38. BCL11A의 CRISPR 넉다운 또는 g-글로빈 유전자 클러스터에서 삽입결실/결실 형성 시 환자 시료로부터의 적혈구계 세포에서 g-글로빈 전사체의 배수 변화이다. 3명의 겸상 적혈구 질병 환자(SCD1 내지 3)로부터 유래된 CD34+ 세포를 실시예에 기재된 바와 같이 *시험관내에서* CRISPR에 의해 편집하고 적혈구계 계통으로 분화시켰다. 적혈구계 분화 11일차에, 세포를 배양물로부터 수합하고, qPCR을 받게 하여 g-글로빈 및 b-글로빈 전사체를 측정하고, GAPDH에 대해 정규화하였다. 실험을 2벌 중복으로 수행하였다. 데이터는 풀링된 공여자의 평균±SEM을 보여준다.

도 39. 유전자 편집 시 겸상 적혈구 질병 환자 시료에서 HbF+ 세포의 계수이다. 3명의 겸상 적혈구 질병 환자(SCD1 내지 3)로부터 유래된 CD34+ 세포를 실시예에 기재된 바와 같이 *시험관내에서* CRISPR에 의해 편집하고 적혈구계 계통으로 분화시켰다. 적혈구계 분화 11일차, 14일차 및 21일차에, 세포를 항-HbF-FITC 항체로 염색시켜 HbF+ 세포를 유세포분석에 의해 계수하였다. 실험을 2벌 중복으로 수행하였다. 데이터는 평균±SEM을 보여준다. SCD1 내지 3으로부터 결과는 11일차에 제시되며; SCD-1 및 SCD-2로부터의 결과는 14일차에 제시되고; SCD-2 및

SCD-3으로부터의 결과는 21일차에 제시된다.

도 40. 유전자-편집된 겸상 적혈구 질병 환자 시료에서 유세포분석에 의한 태아 헤모글로빈 발현의 측정이다. 3명의 겸상 적혈구 질병 환자(SCD1 내지 3)로부터 유래된 CD34+ 세포를 실시예에 기재된 바와 같이 *시험관내에서* CRISPR에 의해 편집하고 적혈구계 계통으로 분화시켰다. 적혈구계 분화 11일차, 14일차 및 21일차에, 세포를 항-HbF-FITC 항체로 염색시켜 각각의 세포의 HbF 발현 강도를 유세포분석에 의해 측정하였다. 실험을 2벌 중복으로 수행하였다. 데이터는 평균±SEM을 보여준다. MFI는 평균 형광 강도이다. SCD1 내지 3으로부터 결과는 11일차에 제시되며; SCD-1 및 SCD-2로부터의 결과는 14일차에 제시되고; SCD-2 및 SCD-3으로부터의 결과는 21일차에 제시된다.

도 41. 환자 시료의 CRISPR 편집 시 겸상 적혈구 대 정상 세포의 수의 계수이다. 3명의 겸상 적혈구 질병 환자(SCD1 내지 3)로부터 유래된 CD34+ 세포를 이전의 절차에 기재된 바와 같이 *시험관내에서* CRISPR에 의해 편집하고 적혈구계 계통으로 분화시켰다. 적혈구계 분화 21일차에, 세포를 4일 동안 % 저산소증 챔버를 거치게 하고, 고정시키고, 뒤이어 단일 세포 이미징 유세포분석을 받게 하였다. 도 41a(좌측 패널)는 단일 세포 이미징에 의해 계수된 바와 같이 유전자-편집된 환자 시료에서 겸상 적혈구의 수의 변화를 보여준다. 도 41b(우측 패널)는 단일 세포 이미징에 의해 계수된 바와 같이 환자 시료에서 유전자 편집 후 정상 세포의 수의 변화를 보여준다. 3회의 독립적인 실험을 수행하였으며, 각각의 실험을 2벌 중복으로 수행하였다. $4x10^4$개 단일 세포 이미지를 각각의 환자로부터 계수하였다. 그래프는 풀링된 데이터로부터 평균±SEM을 보여준다.

발명을 실시하기 위한 구체적인 내용

[0328] 정의

[0329] 용어 "CRISPR 시스템", "Cas 시스템" 또는 "CRISPR/Cas 시스템"은 RNA-가이드 뉴클레아제 또는 다른 효과기 분자 및 gRNA 분자를 함께 포함하여, RNA-가이드된 뉴클레아제 또는 다른 효과기 분자에 의해 표적 서열에서 핵산의 변형을 안내 및 달성하기에 필요하고 충분한 분자 세트를 지칭한다. 일 실시형태에서, CRISPR 시스템은 gRNA 및 Cas 단백질, 예를 들어 Cas9 단백질을 포함한다. Cas9 또는 변형된 Cas9 분자를 포함하는 이러한 시스템은 본원에서 "Cas9 시스템" 또는 "CRISPR/Cas9 시스템"으로 지칭된다. 일례에서, gRNA 분자 및 Cas 분자는 복합체화되어 리보핵 단백질(RNP) 복합체를 형성할 수 있다.

[0330] 용어 "가이드 RNA", "가이드 RNA 분자", "gRNA 분자" 또는 "gRNA"는 상호 교환적으로 사용되고, RNA-가이드된 뉴클레아제 또는 다른 효과기 분자(전형적으로gRNA 분자와 복합체로)의 표적 서열로의 특이적 안내를 촉진하는 핵산 분자 세트를 지칭한다. 일부 실시형태에서, 상기 안내는 gRNA의 일부분이 DNA에 혼성화되는 것(예를 들어 gRNA 표적화 도메인을 통해)을 통해, 및 gRNA 분자의 일부분이 RNA-가이드된 뉴클레아제 또는 다른 효과기 분자에 결합하는 것(예를 들어 적어도 gRNA tracr를 통해)에 의해 달성된다. 실시형태에서, gRNA 분자는 본원에서 "단일 가이드 RNA" 또는 "sgRNA" 등으로 지칭되는 단일 인접 폴리뉴클레오타이드 분자로 구성된다. 다른 실시형태에서, gRNA 분자는 본원에서 "이중 가이드 RNA" 또는 "dgRNA" 등으로 지칭되는, 통상적으로 혼성화를 통해, 스스로 관련될 수 있는, 복수의, 통상적으로 2개의 폴리뉴클레오타이드 분자로 구성된다. gRNA 분자는 하기에 더욱 상세하게 기재되지만, 일반적으로 표적화 도메인 및 tracr를 포함한다. 실시형태에서, 표적화 도메인 및 tracr는 단일 폴리뉴클레오타이드 상에 배치된다. 다른 실시형태에서, 표적화 도메인 및 tracr는 별개의 폴리뉴클레오타이드 상에 배치된다.

[0331] 용어가 gRNA와 관련되어 사용되는 바와 같이는 용어 "표적화 도메인"은, 표적 서열, 예를 들어 세포의 핵산 내, 예를 들어 유전자 내, 표적 서열을 인지하는, 예를 들어 표적 서열에 상보적인 gRNA 분자의 부분이다.

[0332] 용어가 gRNA와 관련되어 사용되는 바와 같이는 용어 "crRNA"는 표적화 도메인 및, 플래그폴 영역을 형성하기 위해 tracr과 상호 작용하는 영역을 포함하는 gRNA 분자의 일부분이다.

[0333] 용어 "표적 서열"은 gRNA 표적화 도메인에 상보적인, 예를 들어 완전히 상보적인 핵산 서열을 지칭한다. 실시형태에서, 표적 서열은 게놈 DNA 상에 배치된다. 일 실시형태에서, 표적 서열은 뉴클레아제 또는 다른 효과기 활성을 갖는 단백질에 의해 인지되는 프로토스페이서 인접 모티프(PAM) 서열, 예를 들어 Cas9에 의해 인지되는 PAM 서열에 인접해 있다(DNA의 동일 가닥 상에서 또는 상보적 가닥 상에서). 실시형태에서, 표적 서열은 글로빈 유전자의 발현에 영향을 미치는, 예를 들어 베타 글로빈 또는 태아 헤모글로빈(HbF)의 발현에 영향을 미치는 유전자 또는 유전자좌 내의 표적 서열이다. 실시형태에서, 표적 서열은 비결실 HPFH 영역 내의 표적 서열이며, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역 내에 존재한다.

[0334] gRNA분자와 관련되어 본원에 사용된 용어 "플래그폴"은 crRNA 및 tracr가 서로 결합하거나 또는 서로 혼성화되는 gRNA의 부분을 지칭한다.

[0335] gRNA분자와 관련되어 본원에 사용된 용어 "tracr"은 뉴클레아제 또는 다른 효과기 분자에 결합하는 gRNA의 부분을 지칭한다. 실시형태에서, tracr는 Cas9에 특이적으로 결합하는 핵산 서열을 포함한다. 실시형태에서, tracr는 플래그폴의 일부를 형성하는 핵산 서열을 포함한다.

[0336] 용어 "Cas9" 또는 "Cas9 분자"는 DNA 절단을 담당하는 박테리아 II형 CRISPR/Cas 시스템 유래의 효소를 지칭한다. Cas9는 또한 야생형 단백질뿐만 아니라 이의 기능적 및 비기능적 돌연변이를 포함한다. 실시형태에서, Cas9는 *S.피오게네스*의 Cas9이다.

[0337] 핵산과 관련되어 사용된 용어 "상보적"은 염기들, A의 T 또는 U와의, 및 G의 C와의 쌍을 지칭한다. 용어 상보적은 완전히 상보적인, 즉, 전체 참조 서열에 걸쳐 A 대 T 또는 U 쌍 및 G 대 C 쌍을 형성하는 핵산 분자뿐만 아니라 적어도 80%, 85%, 90%, 95%, 99% 상보적인 분자도 지칭한다.

[0338] 상동성-지시 수선 또는 상동 재조합과 관련되어 사용된 "주형 핵산"은 절단 부위에서의 유전자 수선(삽입)을 위해 CRISPR 시스템 공여자 서열에 의해 변형 부위에 삽입되는 핵산을 지칭한다.

[0339] 용어가 본원에 사용되는 바와 같이는 "삽입결실"은 gRNA 분자를 포함하는 조성물, 예를 들어 CRISPR 시스템에 노출된 후 발생한, 참조 핵산에 비해 하나 이상의 뉴클레오타이드의 삽입, 하나 이상의 뉴클레오타이드의 결실, 또는 뉴클레오타이드의 삽입 및 결실의 조합을 포함하는 핵산을 지칭한다. 삽입결실은 gRNA 분자를 포함하는 조성물에 노출된 후 핵산을 시퀀싱하는 것, 예를 들어 NGS에 의해 결정될 수 있다. 삽입결실 부위에 대해, 참조 서열의 약 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1뉴클레오타이드 내에 적어도 하나의 삽입 또는 결실을 포함하는 경우, 또는 상기 참조 부위의 일부 또는 전부와 중첩되는(예를 들어 gRNA 분자, 예를 들어 본원에 기재된 gRNA 분자의 표적화 도메인에 상보적인 부위의 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1 뉴클레오타이드와 중첩되는 또는 내에 적어도 하나의 삽입 또는 결실을 포함하는) 경우, 삽입결실은 참조 부위(예를 들어 gRNA 분자의 표적화 도메인에 상보적인 부위) "에 또는 그 부근에" 존재한다고 한다. 실시형태에서, 삽입결실은 예를 들어 약 1 kb 초과, 약 2 kb 초과, 약 3 kb 초과, 약 4 kb 초과, 약 5 kb 초과, 약 6 kb 초과 또는 약 10 kb 초과의 핵산을 포함하는 대형 결실이다. 실시형태에서, 대형 결실의 5' 말단, 3' 말단, 또는 5' 말단과 3' 말단 둘 모두는 본원에 기재된 gRNA 분자의 표적 서열에서 또는 그 부근에 배치된다. 실시형태에서, 대형 결실은, HBG1 프로모터 영역 내에 배치된 예를 들어 본원에 기재된 gRNA 분자의 표적 서열과 HBG2 프로모터 영역 내에 배치된 예를 들어 본원에 기재된 gRNA 분자의 표적 서열 사이에 배치된 약 4.9 kb의 DNA를 포함한다.

[0340] 용어가 본원에 사용되는 바와 같이는 "삽입결실 패턴(indel pattern)"은 gRNA 분자를 포함하는 조성물에의 노출 후 발생한 삽입결실 세트를 지칭한다. 실시형태에서, 삽입결실 패턴은 출현 빈도에 의해 상위 3개의 삽입결실들로 구성된다. 실시형태에서, 삽입결실 패턴은 출현 빈도에 의해 상위 5개의 삽입결실들로 구성된이다. 실시형태에서, 삽입결실 패턴은 모든 시퀀싱 리드(read)에 비해 약 5% 초과의 빈도로 존재하는 삽입결실로 구성된다. 실시형태에서, 삽입결실 패턴은 삽입결실 시퀀싱 리드(즉, 변형되지 않은 참조 핵산 서열로 구성되지 않은 리드)의 총 개수에 비해 약 10% 초과의 빈도로 존재하는 삽입결실로 구성된다. 실시형태에서, 삽입결실 패턴은 상위 5개의 가장 빈번하게 관찰된 삽입결실 중 임의의 3개를 포함한다. 삽입결실 패턴은 예를 들어 gRNA 분자에 노출된 세포 집단의 세포를 시퀀싱함에 의해 결정될 수 있다.

[0341] 용어가 본원에 사용되는 바와 같이는 "표적-외 삽입결실"은 gRNA 분자의 표적화 도메인의 표적 서열 외의 부위의 또는 그 부근의 삽입결실을 지칭한다. 이러한 부위는 예를 들어 gRNA의 표적화 도메인의 서열에 비해 1,2,3,4,5 또는 그 이상의 미스매치 뉴클레오타이드를 포함할 수 있다. 예시적인 실시형태에서, 그러한 부위는 인 실리코(in silico) 예측된 표적-외 부위의 표적화된 시퀀싱을 이용하여, 또는 당분야에 공지된 삽입 방법에 의해 검출된다. 본원에 기재된 gRNA에 관하여, 표적-외 삽입결실의 예는 HBG1 및/또는 HBG2 프로모터 영역의 외부의 서열에서 형성된 삽입결실이다. 예시적인 실시형태에서, 표적-외 삽입결실은 유전자의 서열에서, 예를 들어 유전자의 코딩 서열 내에서 형성된다.

[0342] 용어 단수형("a" 및 "an")은 관사의 문법적 목적어의 하나 또는 하나 초과(즉, 적어도 하나)를 지칭한다. 예로서, "요소"는 하나의 요소 또는 하나를 초과하는 요소를 의미한다.

[0343] 양, 시간의 지속 등과 같은 측정 가능한 값을 지칭할 때 용어 "약"은 이러한 변동이 개시된 방법을 수행하기에 적절한 동안, 특정 값으로부터 ± 20% 또는 일부 경우 ± 10%, 또는 일부 경우 ± 5%, 또는 일부 경우 ± 1%, 또는 일부 경우 ± 0.1%의 변동을 포괄하는 의미이다.

[0344] 용어 "항원" 또는 "Ag"는 면역 반응을 유발하는 분자를 지칭한다. 이 면역 반응은 항체 생산, 또는 특정 면역학적으로-적격한 세포의 활성화, 또는 둘 모두를 수반할 수 있다. 당업자는 사실상 모든 단백질 또는 펩타이드를 포함하는 임의의 거대 분자가 항원으로서 역할할 수 있음을 이해할 것이다. 추가적으로, 항원은 재조합 또는 게놈 DNA로부터 유래될 수 있다. 당업자는 면역 반응을 유도하는 단백질을 인코딩하는 뉴클레오타이드 서열 또는 부분 뉴클레오타이드 서열을 포함하는 임의의 DNA가 따라서, 그 용어가 본원에 사용되는 바와 같이, "항원"을 인코딩하는 것을 이해할 것이다. 추가적으로, 당업자는 항원이 유전자의 전장 뉴클레오타이드 서열에 의해서만 인코딩될 필요는 없다는 것을 이해할 것이다. 본 발명이 하나를 초과하는 유전자의 부분 뉴클레오타이드 서열의 사용을 포함하나 이에 제한되지 않는다는 것 및 이들 뉴클레오타이드 서열이 원하는 면역 반응을 유도하는 폴리펩타이드를 인코딩하기 위해 다양한 조합으로 배열된다는 것은 쉽게 명백하다. 더욱이, 당업자는 항원이 "유전자"에 의해 인코딩될 필요가 전혀 없음을 이해할 것이다. 항원은 합성되거나 생물학적 시료에서 유래될 수 있고, 또는 폴리펩타이드 외의 거대 분자일 수 있다는 것은 쉽게 명백하다. 이러한 생물학적 시료은 조직 시료, 세포 또는 다른 생물학적 성분을 갖는 유체를 포함할 수 있으나, 이들로 한정되는 것은 아니다.

[0345] 용어 "자가"는 물질이 이후 재-도입될 동일한 개체에서 유래된 임의의 물질을 지칭한다.

[0346] 용어 "동종이계"는 물질이 도입되는 개체와 동일한 종의 상이한 동물에서 유래된 임의의 물질을 지칭한다. 하나 이상의 유전자좌의 유전자가 동일하지 않은 경우 둘 이상의 개체는 서로 동종이계라고 한다. 일부 양태에서, 동일한 종의 개체로부터의 동종이계 물질은 항원적으로 상호 작용하기에 충분히 유전적으로 같지 않을 수 있다.

[0347] 용어 "이종(xenogenic)"은 상이한 종의 동물에서 유래된 이식을 지칭한다.

[0348] 용어가 본원에 사용되는 바와 같이 "에서 유래된"은 제1 과 제2 분자 사이의 관계를 표시한다. 이는 일반적으로 제1 분자와 제2 분자 사이의 구조적 유사성을 지칭하며, 제2 분자에서 유래된 제1 분자에 대해 공정 또는 공급원 제한을 함축 또는 포함하지 않는다.

[0349] 용어 "인코딩"은 정의된 서열의 뉴클레오타이드(예를 들어 rRNA, tRNA 및 mRNA) 또는 정의된 서열의 아미노산 서열 및 그로부터 발생하는 생물학적 특성을 갖는 생물학적 과정에서의 다른 중합체 및 거대 분자의 합성을 위한 주형으로서 역할하는, 유전자, cDNA 또는 mRNA와 같은 폴리뉴클레오타이드 내 특정 뉴클레오타이드 서열의 내재적 특성을 지칭한다. 따라서 유전자, cDNA 또는 RNA는 그 유전자에 해당하는 mRNA의 전사 및 번역이 세포 또는 다른 생물학적 시스템에서 단백질을 생산하는 경우 단백질을 인코딩한다. 그 뉴클레오타이드 서열이 mRNA 서열과 동일하며 통상적으로 서열 목록에 제공되는 코딩 가닥, 및 유전자 또는 cDNA의 전사를 위한 주형으로 사용되는 비-코딩 가닥 둘다 단백질 또는 그 유전자 또는 cDNA의 다른 산물을 인코딩하는 것으로 지칭될 수 있다.

[0350] 달리 특정되지 않는 한, "아미노산 서열을 인코딩하는 뉴클레오타이드 서열"은 서로의 축퇴성 형태이고 동일한 아미노산 서열을 인코딩하는 모든 뉴클레오타이드 서열을 포함한다. 단백질 또는 RNA를 인코딩하는 구(phrase) 핵산 서열은 또한 일부 형태에서 단백질을 인코딩하는 뉴클레오타이드 서열이 인트론을 함유할 수 있는 정도까지 인트론을 포함할 수 있다.

[0351] 용어 "유효량" 또는 "치료적 유효량"은 본원에서 상호교환적으로 사용되며, 특정 생물학적 결과를 달성하기에 효과적인 본원에 기재된 화합물, 제형, 물질 또는 조성물의 양을 지칭한다.

[0352] 용어 "내인성(endogenous)"은 유기체, 세포, 조직 또는 시스템 내부에서 유래된 또는 생산된 임의의 물질을 지칭한다.

[0353] 용어 "외인성(exogenous)"은 유기체, 세포, 조직 또는 시스템 외부로부터 도입된 또는 생산된 임의의 물질을 지칭한다.

[0354] 용어 "발현"은 프로모터에 의해 추진된 특정 뉴클레오타이드 서열의 전사 및/또는 번역을 지칭한다.

[0355] 용어 "전달 벡터"는 분리된 핵산을 포함하고 세포의 내부로 분리된 핵산을 전달하기 위해 사용될 수 있는 물질 조성을 지칭한다. 선형 폴리뉴클레오타이드, 이온성 또는 양친성 화합물 관련된 폴리뉴클레오타이드, 플라스미드 및 바이러스를 포함하는 다수의 벡터가 당분야에 공지되어 있지만, 이들로 한정되는 것은 아니다. 따라서, 용어 "전달 벡터"는 자율적으로 복제하는 플라스미드 또는 바이러스를 포함한다. 용어는 예를 들어 폴리리신 화합물, 리포좀 등과 같은 핵산의 세포로의 전달을 촉진하는 비플라스미드 및 비바이러스 화합물을 추가로 포함하는 것으로 또한 해석되어야 한다. 바이러스 전달 벡터의 예로는 아데노바이러스 벡터, 아데노-관련 바이러스 벡터, 레트로바이러스 벡터, 렌티바이러스 벡터 등이 있으나, 이들로 한정되는 것은 아니다.

[0356] 용어 "발현 벡터"는 발현될 뉴클레오타이드 서열에 작동 가능하게 연결된 발현 조절 서열을 포함하는 재조합 폴

리뉴클레오타이드를 포함하는 벡터를 지칭한다. 발현 벡터는 발현을 위한 충분한 시스-작용 요소를 포함하며; 발현을 위한 다른 요소는 숙주 세포에 의해 또는 시험관내 발현 시스템 내에서 공급될 수 있다. 발현 벡터는 재조합 폴리뉴클레오타이드를 혼입시키는 코스미드, 플라스미드(예를 들어 네이키드(naked) 또는 리포좀 내 함유된) 및 바이러스(예를 들어 렌티바이러스, 레트로바이러스, 아데노바이러스 및 아데노-관련 바이러스)를 포함하는 당분야에 공지된 모든 것을 포함한다.

[0357] 용어 "상동의" 또는 "동일성"은 2개의 중합체 분자 사이, 예를 들어 2개의 핵산 분자, 예를 들어 2개의 DNA 분자 또는 2개의 RNA 분자 사이, 또는 2개의 폴리펩타이드 분자 사이의 하위단위 서열 동일성을 지칭한다. 2개의 분자 둘 모두에서의 하위단위 위치가 동일한 단량체 하위단위에 의해 점유될 때; 예를 들어 2개의 DNA 분자의 각각에서의 위치가 아데닌에 의해 점유되는 경우, 그들은 그 위치에서 상동 또는 동일하다. 2개의 서열 사이의 상동성은 매치되거나 상동의 위치의 개수의 직접 함수이며; 예를 들어 2개의 서열에서 위치의 절반(예를 들어 10개 하위단위 길이의 중합체에서 5개 위치)이 상동이면, 2개의 서열은 50% 상동이고; 위치의 90%(예를 들어 10개 중 9개)가 매치되거나 상동이면, 2개의 서열은 90% 상동이다.

[0358] 용어 "단리된"은 자연 상태로부터 변경되거나 제거되는 것을 의미한다. 예를 들어 살아있는 동물 내 자연적으로 존재하는 핵산 또는 펩타이드는 "단리된"이 아니지만, 그 자연적인 상태의 공존 물질로부터 부분적으로 또는 완전히 분리된 동일한 핵산 또는 펩타이드는 "단리된"이다. 단리된 핵산 또는 단백질은 실질적으로 정제된 형태로 존재할 수 있거나, 예를 들어 숙주 세포와 같은, 본래가 아닌(non-native) 환경에서 존재할 수 있다.

[0359] 용어 "작동 가능하게 연결된" 또는 "전사 제어"는 이종 핵산 서열을 발현을 야기하는, 조절 서열과 이종 핵산 서열 사이의 기능적 연결을 지칭한다. 예를 들어 제1 핵산 서열이 제2 핵산 서열과 기능적 관계에 배치될 때 제1 핵산 서열은 제2 핵산 서열과 작동 가능하게 연결된다. 예를 들어 프로모터가 코딩 서열의 전사 또는 발현에 영향을 주는 경우 프로모터는 코딩 서열에 작동 가능하게 연결된다. 작동 가능하게 연결된 DNA 서열은 서로와 인접할 수 있으며, 예를 들어 2개의 단백질 코딩 영역을 연결하기 위해 필요한 경우, 동일한 리딩 프레임 내에 있다.

[0360] 용어 면역원성 조성물의 "비경구" 투여는 예를 들어 피하(s.c.), 정맥 내(i.v.), 근육 내(i.m.) 또는 흉골 내 주사, 종양 내, 또는 주입 기술을 포함한다.

[0361] 용어 "핵산" 또는 "폴리뉴클레오타이드"는 단일 또는 이중 가닥 형태의 데 옥시리보핵산(DNA) 또는 리보핵산(RNA) 및 이의 중합체를 지칭한다. 명확히 제한되지 않는 한, 용어는 참조 핵산과 유사한 결합 특성을 갖고 자연적 발생 뉴클레오타이드와 유사한 방식으로 대사되는 천연 뉴클레오타이드의 공지된 유사체를 함유하는 핵산을 포괄한다. 달리 표시하지 않는 한, 특정 핵산 서열은 또한 명시적으로 표시된 서열뿐만 아니라 보존적으로 변형된 그의 변이체(예를 들어 축퇴성 코돈 치환), 대립유전자, 이종상동체(ortholog), SNP 및 상보적 서열을 암시적으로 포괄한다. 구체적으로, 축퇴성 코돈 치환은 하나 이상의 선택된(또는 모든) 코돈의 세 번째 위치가 혼합 염기 및/또는 데옥시이노신 잔기로 치환된 서열을 생성함에 의해 달성할 수 있다(문헌[Batzer et al., Nucleic Acid Res.19 : 5081 (1991)]; 문헌[Ohtsuka et al., J.Biol.Chem.260 : 2605-2608(1985)] 및 문헌[Rossolini et al., Mol. Cell. Probes 8 : 91-98(1994)]).

[0362] 용어 "펩타이드", "폴리펩타이드" 및 "단백질"은 상호교환적으로 사용되며, 펩타이드 결합에 의해 공유 결합된 아미노산 잔기로 구성된 화합물을 지칭한다. 단백질 또는 펩타이드는 적어도 2개의 아미노산을 함유해야 하며, 단백질 또는 펩타이드 서열을 포함할 수 있는 아미노산의 최대 개수에 대해 제한은 없다. 폴리펩타이드는 펩타이드 결합에 의해 서로 연결된 2개 이상의 아미노산을 포함하는 임의의 펩타이드 또는 단백질을 포함한다. 본원에 사용된 바와 같이, 용어는, 예를 들어 펩타이드, 올리고 펩타이드 및 올리고머로 당분야에서 통상적으로 지칭되는, 단쇄 및, 많은 유형이 있는, 단백질로 당분야에서 일반적으로 지칭되는 장쇄 둘 모두 지칭한다. "폴리펩타이드"에는 예를 들어 생물학적 활성 단편, 실질적으로 상동 폴리펩타이드, 올리고펩타이드, 동종이합체, 이종이합체, 폴리펩타이드의 변이체, 변형된 폴리펩타이드, 유도체, 유사체, 융합 단백질 등이 포함된다. 폴리펩타이드에는 천연 펩타이드, 재조합 펩타이드, 또는 이의 조합이 포함된다.

[0363] 용어 "프로모터"는 폴리뉴클레오타이드 서열의 특이적인 전사를 개시하기 위해 필요한, 세포의 합성 기구 또는 도입된 합성 기구에 의해 인지되는 DNA 서열을 지칭한다.

[0364] 용어 "프로모터/조절 서열"은 프로모터/조절 서열에 작동 가능하게 연결된 유전자 산물의 발현을 위해 필요한 핵산 서열을 지칭한다. 일부 경우, 이 서열은 코어 프로모터 서열일 수 있고, 다른 경우에, 이 서열은 또한 인핸서 서열 및 유전자 산물의 발현을 위해 필요한 다른 조절 요소를 포함할 수 있다. 프로모터/조절 서열은, 예

를 들어 조직 특이적 방법으로 유전자 산물을 발현하는 것일 수 있다.

[0365] 용어 "구성적" 프로모터는 유전자 산물을 인코딩 또는 특정하는 폴리뉴클레오타이드와 작동 가능하게 연결된 경우, 세포의 대부분 또는 모든 생리적 조건 하에서 유전자 산물이 세포에서 생산되도록 야기하는 뉴클레오타이드 서열을 지칭한다.

[0366] 용어 "유도성(inducible)" 프로모터는 유전자 산물을 인코딩 또는 특정하는 폴리뉴클레오타이드와 작동 가능하게 연결된 경우, 실질적으로 프로모터에 대응하는 유도제가 세포 내에 존재할 때에만 유전자 산물이 세포에서 생산되도록 야기하는 뉴클레오타이드 서열을 지칭한다.

[0367] 용어 "조직 특이적" 프로모터는 유전자에 의해 인코딩 또는 특정된 폴리뉴클레오타이드와 작동 가능하게 연결된 경우, 실질적으로 세포가 프로모터에 대응하는 조직 유형의 세포인 경우에만 세포에서 유전자 산물이 생산되도록 야기하는 뉴클레오타이드 서열을 지칭한다.

[0368] 메신저 RNA(mRNA)와 관련되어 본원에 사용된, 5'캡(RNA 캡, RNA 7-메틸구아노신 캡 또는 RNA m7G 캡으로 또한 지칭됨)은 "프론트" 또는 전사 시작 직후 진핵 세포 메신저 RNA의 5' 말단에 부가된 변형된 구아닌 뉴클레오타이드이다. 5'캡은 첫 번째 전사되는 뉴클레오타이드에 연결된 말단 군으로 구성된다. 그 존재는 리보솜에 의한 인지 및 RNase로부터의 보호에 대단히 중요하다. 캡 부가는 전사에 연결되며, 각자가 서로에게 영향을 미치도록, 동시-전사적으로(co-transcriptionally) 발생한다. 전사 시작 직후, 합성되는 mRNA의 5' 말단은 RNA 중합효소와 관련된 캡 합성 복합체에 의해 결합된다. 이 효소 복합체는 mRNA 캡핑에 필요한 화학 반응을 촉매한다. 합성은 다단계 생화학 반응으로 진행한다. 캡핑 모이어티는 안정성 또는 번역 효율과 같은 mRNA의 기능성을 조절하도록 변형될 수 있다.

[0369] 본원에 사용된 "시험관내 전사된 RNA"는 시험관내에서 합성된 RNA, 바람직하게는 mRNA를 지칭한다. 일반적으로, 시험관내 전사된 RNA는 시험관내 전사 벡터로부터 생성된다. 시험관내 전사 벡터는 시험관내 전사된 RNA를 생성하기 위해 사용되는 주형을 포함한다.

[0370] 본원에 사용된 "폴리(A)"는 폴리아데닐화에 의해 mRNA에 붙은 일련의 아데노신이다. 일시적 발현을 위한 작제물의 바람직한 실시형태에서, 폴리A는 50과 5000 사이(SEQ ID NO: 190), 바람직하게는 64 초과, 더욱 바람직하게는 100 초과, 가장 바람직하게는 300 또는 400 초과이다. 폴리(A) 서열은, 배치, 안정성 또는 번역 효율과 같은 mRNA 기능성을 조절하기 위해 화학적으로 또는 효소적으로 변형될 수 있다.

[0371] 본원에 사용된 "폴리아데닐화"는 폴리아데닐일 모이어티 또는 이의 변형된 변이체의 메신저 RNA 분자로의 공유 결합을 지칭한다. 진핵 생물에서 대부분의 메신저 RNA(mRNA) 분자는 3' 말단에서 폴리아데닐화된다. 3'폴리(A) 꼬리는 효소인 폴리아데닐레이트 중합효소의 작용을 통해 pre-mRNA에 부가된 아데닌 뉴클레오타이드(종종 수백 개)의 긴 서열이다. 고등 진핵 생물에서 폴리(A) 꼬리는 특정 서열인 폴리아데닐화 신호를 함유하는 전사체에 부가된다. 폴리(A) 꼬리 및 그에 결합된 단백질은 엑소뉴클레아제에 의한 분해로부터 mRNA를 보호하는데 도움을 준다. 폴리아데닐화는 또한 전사 종결, 핵으로부터의 mRNA의 외수송, 및 번역을 위해 중요하다. 폴리아데닐화는 DNA가 RNA로 전사 된 직후에 핵에서 발생하지만, 추가적으로, 이후에 세포질에서도 발생할 수 있다. 전사가 종결된 후, mRNA 사슬은 RNA 중합효소와 관련된 엔도뉴클레아제 복합체의 작용을 통해 절단된다. 절단 부위는 통상적으로 절단 부위 부근의 염기 서열 AAUAAA의 존재를 특징으로 한다. mRNA가 절단된 후, 아데노신 잔기가 절단 부위에서 유리 3' 말단에 부가된다.

[0372] 본원에 사용된, "일시적(transient)"은 수 시간, 수 일 또는 수 주의 기간 동안 통합되지 않은 전이 유전자의 발현을 지칭하며, 발현 기간은 게놈에 통합되는 경우 또는 숙주 세포에서 안정한 플라스미드 복제 단물 내에 함유되는 경우의 유전자 발현을 위한 기간보다 적다.

[0373] 본원에 사용된 용어 "치료하다", "치료" 및 "치료하는"은, 하나 이상의 치료법(예를 들어 gRNA 분자, CRISPR 시스템, 또는 본 발명의 변형된 세포와 같은 하나 이상의 치료제)의 투여로부터 발생하는, 장애, 예를 들어 혈색소병증의 진행, 중증도 및/또는 지속기간의 감소 또는 개선, 또는 장애, 예를 들어 혈색소병증의 하나 이상의 증상(바람직하게는, 하나 이상의 식별 가능한 증상)의 개선을 지칭한다. 특정 실시형태에서, 용어 "치료하다", "치료" 및 "치료하는"은 환자에 의해 식별되지 않는, 혈색소병증 장애의 적어도 하나의 측정 가능한 신체적 매개변수의 개선을 지칭한다. 다른 실시형태에서, 용어 "치료하다", "치료" 및 "치료하는"은 신체적으로, 예를 들어 식별 가능한 증상의 안정화에 의해, 생리적으로, 예를 들어 신체적 매개변수의 안정화에 의해, 또는 둘 모두에 의해 장애의 진행을 저해하는 것을 지칭한다. 다른 실시형태에서, 용어 "치료하다", "치료" 및 "치료하는"은 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 베타-탈라세미아의 증상의 감소 또는 안정화를 지칭한다.

[0374] 용어 "신호 전달 경로"는 세포의 한 부분에서 세포의 또 다른 부분으로의 신호 전달에 중요한 역할을 하는 다양한 신호 전달 분자 사이의 생화학적 관계를 지칭한다. 문구 "세포 표면 수용체"는 신호를 받고 세포막 건너 신호를 전달할 수 있는 분자 및 분자의 복합체를 포함한다.

[0375] 용어 "대상체"는 면역 반응이 유도될 수있는 살아있는 유기체(예를 들어 포유류, 인간)를 포함하는 것을 의미한다.

[0376] 용어 "실질적으로 정제된" 세포는 본질적으로 다른 세포 유형이 없는 세포를 지칭한다. 실질적으로 정제된 세포는 또한 자연적으로 발생하는 상태에서 정상적으로 관련되어 있는 다른 세포 유형으로부터 분리된 세포를 지칭한다. 일부 경우, 실질적으로 정제된 세포 집단은 균질한 세포 집단을 지칭한다. 다른 경우, 이 용어는 단순히 자연 상태에서 자연적으로 관련되어 있는 세포로부터 분리된 세포를 지칭한다. 일부 양태에서, 세포는 시험관내 배양된다. 다른 양태에서, 세포는 시험관내 배양되지 않는다.

[0377] 본원에 사용된 용어 "치료적"은 치료를 의미한다. 치료적 효과는 질병 상태의 감소, 억제, 완화 또는 박멸에 의해 수득된다.

[0378] 본원에 사용된 용어 "예방"은 질병 또는 질병 상태의 예방 또는 보호적 처치를 의미한다.

[0379] 용어 "형질감염된" 또는 "형질전환된" 또는 "형질도입된"은 외인성 핵산 및/또는 단백질이 숙주 세포 내로 전달되거나 도입되는 과정을 지칭한다. "형질감염된" 또는 "형질전환된" 또는 "형질도입된" 세포는 외인성 핵산 및/또는 단백질로 형질감염, 형질전환 또는 형질도입된 것이다. 세포는 일차 대상체 세포 및 그의 자손을 포함한다.

[0380] 용어 "특이적으로 결합한다"는 시료 내에 존재하는 결합 파트너(예를 들어 단백질 또는 핵산)를 인지하고 결합하지만, 시료 내의 다른 분자를 실질적으로 인지 또는 결합하지 않는 분자를 지칭한다.

[0381] 용어 "생물학적 동등성의"는 참조 화합물의 참조 용량 또는 참조량에 의해 생산된 효과와 동등한 효과를 생산하기 위해 필요한, 참조 화합물 외의 제제의 양을 지칭한다.

[0382] 본원에 사용된 "불응성"은 치료에 반응하지 않는 질병, 예를 들어 혈색소병증을 지칭한다. 실시형태에서, 불응성 혈색소병증은 치료 시작 전 또는 치료 초기에 내성을 나타낼 수 있다. 다른 실시형태에서, 불응성 혈색소병증은 치료 동안 내성이 생길 수 있다. 불응성 혈색소병증은 내성 헤모글로빈 병증으로도 불린다.

[0383] 본원에 사용된 "재발된"은 개선 기간 후, 예를 들어 치료법, 예를 들어 혈색소병증 치료법의 이전 처치 후 질병(예를 들어 혈색소병증) 또는 혈색소병증과 같은 질병의 징후 및 증상의 복귀를 지칭한다.

[0384] 범위: 본 개시 전반에 걸쳐, 본 발명의 다양한 양태가 범위 형식으로 제시될 수 있다. 범위 형식의 기재는 단지 편의 및 간결을 위한 것이며, 본 발명의 범위에 대한 융통성없는 제한으로 해석되어서는 안된다는 것이 이해되어야 한다. 따라서, 범위의 기재는 모든 가능한 하위 범위뿐만 아니라 그 범위 내의 개별 수치를 구체적으로 개시한 것으로 고려되어야 한다. 예를 들어 1 내지 6과 같은 범위의 기재는 1 내지 3, 1 내지 4, 1 내지 5, 2 내지 4, 2 내지 6, 3 내지 6 등과 같은 하위 범위뿐만 아니라 그 범위 내의 개별 숫자, 예를 들어 1, 2, 2.7, 3, 4, 5, 5.3 및 6을 구체적으로 개시한 것으로 고려되어야 한다. 또 다른 예로서, 95 내지 99% 동일성과 같은 범위에는 95%, 96%, 97%, 98% 또는 99% 동일성을 갖는 어떤 것이 포함되고, 96 내지 99%, 96 내지 98%, 96 내지 97%, 97 내지 99%, 97 내지 98% 및 98 내지 99% 동일성과 같은 하위 범위가 포함된다. 이것은 범위의 폭에 관계없이 적용된다.

[0385] 용어 "BCL11a"는 B-세포 림프종/백혈병 11A, RNA 중합효소 II 코어 프로모터 근위 영역 서열-특이적 DNA 결합 단백질, 및 모든 인트론 및 엑손과 함께 상기 단백질을 인코딩하는 유전자를 지칭한다. 이 유전자는 C2H2 형 징크 핑거 단백질을 인코딩한다. BCL11A는 태아 헤모글로빈 생산 억제에 중요한 역할을 하는 것으로 밝혀졌다. BCL11a는 B-세포 CLL/림프종 11A(징크 핑거 단백질), CTIP1, EVI9, 동종지향성 바이러스 통합 부위 9 단백질 동족체, COUP-TF-상호 작용 단백질 1, 징크 핑거 단백질 856, KIAA1809, BCL-11A, ZNF856, EVI-9 및 B-세포 CLL/림프종 11A로 또한 알려져 있다. 용어는 BLC11a의 모든 이소폼(isoform) 및 스플라이스 변이체를 포괄한다. BCL11a를 인코딩하는 인간 유전자는 염색체 위치 2p16.1(Ensembl에 의한)로 맵핑된다. 인간 및 쥐의 아미노산 및 핵산 서열은 GenBank, UniProt 및 Swiss-Prot.와 같은 공개 데이터베이스에서 찾을 수 있으며, 인간 BCL11a의 게놈 서열은 GenBank에서 NC_000002.12에서 찾을 수 있다. BCL11a 유전자는 모든 인트론 및 엑손을 포함하여 이 게놈 위치를 지칭한다. 다수의 알려진 BCL11a의 이소형이 있다.

[0386] 인간 BCL11a의 이소폼 1을 인코딩하는 mRNA의 서열은 NM_022893에서 찾을 수 있다.

[0387] 인간 BCL11a의 이소폼 1의 펩타이드 서열은

```
        10         20         30         40         50
MSRRKQGKPQ HLSKREFSPE PLEAILTDDE PDHGPLGAPE GDHDLLTCGQ
        60         70         80         90        100
CQMNFPLGDI LIFIEHKRKQ CNGSLCLEKA VDKPPSPSPI EMKKASNPVE
       110        120        130        140        150
VGIQVTPEDD DCLSTSSRGI CPKQEHIADK LLHWRGLSSP RSAHGALIPT
```

[0388]

```
       160        170        180        190        200
PGMSAEYAPQ GICKDEPSSY TCTTCKQPFT SAWFLLQHAQ NTHGLRIYLE
       210        220        230        240        250
SEHGSPLTPR VGIPSGLGAE CPSQPPLHGI HIADNNPFNL LRIPGSVSRE
       260        270        280        290        300
ASGLAEGRFP PTPPLFSPPP RHHLDPHRIE RLGAEEMALA THHPSAFDRV
       310        320        330        340        350
LRLNPMAMEP PAMDFSRRLR ELAGNTSSPP LSPGRPSPMQ RLLQPFQPGS
       360        370        380        390        400
KPPFLATPPL PPLQSAPPPS QPPVKSKSCE FCGKTFKFQS NLVVHRRSHT
       410        420        430        440        450
GEKPYKCNLC DHACTQASKL KRHMKTHMHK SSPMTVKSDD GLSTASSPEP
       460        470        480        490        500
GTSDLVGSAS SALKSVVAKF KSENDPNLIP ENGDEEEEED DEEEEEEEEE
       510        520        530        540        550
EEEELTESER VDYGFGLSLE AARHHENSSR GAVVGVGDES RALPDVMQGM
       560        570        580        590        600
VLSSMQHFSE AFHQVLGEKH KRGHLAEAEG HRDTCDEDSV AGESDRIDDG
       610        620        630        640        650
TVNGRGCSPG ESASGGLSKK LLLGSPSSLS PFSKRIKLEK EFDLPPAAMP
       660        670        680        690        700
NTENVYSQWL AGYAASRQLK DPFLSFGDSR QSPFASSSEH SSENGSLRFS
       710        720        730        740        750
TPPGELDGGI SGRSGTGSGG STPHISGPGP GRPSSKEGRR SDTCEYCGKV
       760        770        780        790        800
FKNCSNLTVH RRSHTGERPY KCELCNYACA QSSKLTRHMK THGQVGKDVY
       810        820        830
KCEICKMPFS VYSTLEKHMK KWHSDRVLNN DIKTE
```

[0389]

[0390] 이다.

[0391] SEQ ID NO: 245(식별자 Q9H165-1; 및 NM_022893.3; 및 수탁 ADL14508.1).

[0392] 다른 BCL11a 단백질 이소폼의 서열은 하기에서 제공된다:

[0393] 이소폼 2: Q9H165-2

[0394] 이소폼 3: Q9H165-3

[0395] 이소폼 4: Q9H165-4

[0396] 이소폼 5: Q9H165-5

[0397] 이소폼 6: Q9H165-6.

[0398] 본원에 사용된, 인간 BCL11a 단백질은 또한 그의 전장에 걸쳐 BCL11a 이소형 1 내지 6과 적어도 약 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%의 서열 동일성을 갖는 단백질을 포괄하며, 이러한 단백질은 여전히 BCL11a의 기능 중 적어도 하나를 갖는다.

[0399] 본원에 사용된 용어 "글로빈 유전자좌"는 배아(ε), 태아(G(γ) 및 A(γ)), 성인 글로빈 유전자(γ 및 β), 유

전자좌 제어 영역 및 DNase I 과민성 부위를 포함하는 인간 염색체 11의 영역을 지칭한다.

[0400] 핵산과 관련되어 사용된 용어 "상보적"은 염기들, A의 T 또는 U와의, 및 G의 C와의 쌍을 지칭한다. 용어 상보적은 완전히 상보적인, 즉, 전체 참조 서열에 걸쳐 A 대 T 또는 U 쌍 및 G 대 C 쌍을 형성하는 핵산 분자뿐만 아니라 적어도 80%, 85%, 90%, 95%, 99% 상보적인 분자도 지칭한다.

[0401] 용어 "비결실 HPFH"는 하나 이상의 뉴클레오타이드의 삽입 또는 결실을 포함하지 않으며 태아 헤모글로빈의 유전성 지속증(hereditary persistence)을 초래하고 성인 적혈구에서 증가된 태아 헤모글로빈을 특징으로 하는 돌연변이를 지칭한다. 예시적인 실시형태에서, 비결실 HPFH는 문헌[Nathan and Oski's Hematology and Oncology of Infancy and Childhood, 8th Ed., 2015, Orkin SH, Fisher DE, Look T, Lux SE, Ginsburg D, Nathan DG, Eds., Elsevier Saunders]에 기재된 돌연변이, 예를 들어 표 21-5에 기재된 비결실(nondeltional) HPFH 돌연변이이며, 이의 전체 내용은 참조로 본 명세서에 포함된다. 용어 "비결실 HPFH 영역"은 비결실 HPFH를 포함하거나 부근에 존재하는 게놈 부위를 지칭한다. 예시적인 실시형태에서, 비결실 HPFH 영역은 HBG1 프로모터 영역(Chr11:5,249,833 내지Chr11:5,250,237, hg38; - 가닥)의 핵산 서열, HBG2 프로모터 영역(Chr11:5,254,738 내지 Chr11:5,255,164, hg38; - 가닥)의 핵산 서열 또는 이들의 조합이다. 예시적인 실시형태에서, 비결실 HPFH 영역은 문헌[Nathan and Oski's Hematology and Oncology of Infancy and Childhood, 8th Ed., 2015, Orkin SH, Fisher DE, Look T, Lux SE, Ginsburg D, Nathan DG, Eds., Elsevier Saunders]에 기재된(예를 들어 그 안의 표 21-5에 기재된) 비결실 HPFH 중 하나 이상을 포함한다. 예시적인 실시형태에서, 비결실 HPFH 영역은 chr11:5,250,094~5,250,237, - 가닥, hg38에서 핵산 서열; 또는 chr11:5,255,022~5,255,164, - 가닥, hg38에서 핵산 서열; 또는 chr11:5,249,833~5,249,927, - 가닥에서 핵산 서열, hg38; 또는 chr11:5,254,738~5,254,851, - 가닥, hg38에서 핵산 서열; 또는 chr11:5,250,139~5,250,237, - 가닥, hg38에서 핵산 서열; 또는 이들의 조합이다.

[0402] 용어가 본원에 사용되는 바와 같이 "BCL11a 인핸서"는 BCL11a의 발현 또는 기능에 영향을 미치는, 예를 들어 증강시키는 핵산 서열을 지칭한다. 예를 들어 문헌[Bauer et al., Science, vol. 342, 2013, pp. 253-257]을 참조한다. BCL11a 인핸서는 예를 들어 어떤 세포 유형, 예를 들어 적혈구계 계통의 세포에서만 작동될 수 있다. BCL11a 인핸서의 일례는 BCL11a 유전자 유전자의 엑손 2와 엑손 3 사이의 핵산 서열이다(예를 들어 hg38에 기록된 +55: Chr2 : 60497676~60498941, +58: Chr2: 60494251~60495546; +62 : Chr2: 60490409~60491734 위치에 있는 또는 위치에 해당하는 핵산). 실시형태에서, BCL11a 인핸서는 BCL11a 유전자의 엑손 2와 엑손 3 사이의 핵산 서열의 +62 영역이다. 실시형태에서, BCL11a 인핸서는 BCL11a 유전자의 엑손 2와 엑손 3 사이의 핵산 서열의 +58 영역이다. 실시형태에서, BCL11a 인핸서는 BCL11a 유전자의 엑손 2와 엑손 3 사이의 핵산 서열의 +55 영역이다.

[0403] 용어 "조혈 줄기 및 전구 세포" 또는 "HSPC"는 상호교환적으로 사용되고, 조혈 줄기세포("HSC")와 조혈 전구 세포("HPC") 둘 모두를 포함하는 세포 집단을 지칭한다. 이러한 세포는, 예를 들어 CD34+ 인 것을 특징으로 한다. 예시적인 실시형태에서, HSPC는 골수로부터 단리된다. 다른 예시적 실시형태에서, HSPC는 말초 혈액으로부터 단리된다. 다른 예시적 실시형태에서, HSPC는 제대혈로부터 단리된다. 일 실시형태에서, HSPC는 CD34+/CD38-/CD90+/CD45RA-를 특징으로 한다. 실시형태에서, HSPC는 CD34+/CD90+/CD49f+ 세포를 특징으로 한다. 실시형태에서, HSPC는 CD34+ 세포를 특징으로 한다. 실시형태에서, HSPC는 CD34+/CD90+ 세포를 특징으로 한다. 실시형태에서, HSPC는 CD34+/CD90+/CD45RA- 세포를 특징으로 한다.

[0404] 본원에 사용된 "줄기 세포 확장제"는 세포, 예를 들어 HSPC, HSC 및/또는 HPC가 상기 제제가 없는 동일한 세포 유형에 비해 더 빠른 속도로 증식, 예를 들어 수가 증가하도록 야기하는 화합물을 지칭한다. 하나의 예시적인 양태에서, 줄기 세포 확장제는 아릴 탄화수소 수용체 경로의 저해제이다. 줄기 세포 확장제의 추가적인 예는 하기에 제공된다. 실시형태에서, 증식, 예를 들어 수의 증가는 생체외에서 달성된다.

[0405] 용어 "생착" 또는 "생착하다"는 수용자 예를 들어 포유류 또는 인간 대상체의 신체 내로의 세포 또는 조직, 예를 들어 HSPC의 집단의 혼입을 지칭한다. 일례에서, 생착은 수용자 내에서 생착된 세포의 성장, 확장 및/또는 분화를 포함한다. 실시예에서, HSPC의 생착은 수용자 신체 내에서의 상기 HSPC의 적혈구계 세포로의 분화 및 성장을 포함한다.

[0406] 본원에 사용된 용어 "조혈 전구 세포"(HPC)"는 자가 재생을 위한 제한된 능력 및, 조혈 계층 내 배치에 따라(문헌[Doulatov et al., Cell Stem Cell 2012]), 다-계통 분화(예를 들어 골수성, 림프계), 단일 계통 분화(예를 들어 골수성 또는 림프계) 또는 세포 유형 한정된 분화(예를 들어 적혈구계 전구 세포)를 위한 잠재력을 갖는

원시 조혈 세포를 지칭한다.

[0407] 본원에 사용된 바와 같이 "조혈 줄기세포"(HSC; hematopoietic stem cell)는 자가-재생하고, 과립구(예를 들어 전골수구(promyelocyte), 호중구, 호산구, 호염기구), 적혈구(예를 들어 망상적혈구(reticulocyte), 적혈구), 혈소판(thrombocyte)(예를 들어 거핵아구(megakaryoblast), 혈소판 생성 거핵구, 혈소판(platelet)) 및 단핵구(예를 들어 단핵구, 대식세포)를 포함하는 보다 성숙한 혈액 세포로 분화하는 능력을 갖는 미성숙 혈액 세포를 지칭한다. HSC는 명세서 전체에서 줄기세포와 상호교환적으로 기재된다. 이러한 세포는 CD34+ 세포를 포함할 수 있거나 포함하지 않을 수 있는 것으로 당업계에 공지되어 있다. CD34+ 세포는 CD34 세포 표면 마커를 발현하는 미성숙 세포이다. CD34+ 세포는 상기 정의된 줄기세포 특성을 갖는 준세포(subcellular) 집단을 포함하는 것으로 여겨진다. HSC는 원시 전구 세포(예를 들어 다능성 전구 세포) 및/또는 특이적인 조혈 계통에 전념하는 전구 세포(예를 들어 림프계 전구 세포)를 유발할 수 있는 다능성 세포인 것으로 당업계에 잘 공지되어 있다. 특이적인 조혈 계통에 전념하는 줄기세포는 T 세포 계통, B 세포 계통, 수지상 세포 계통, 랑게르한스 세포 계통 및/또는 림프계 조직-특이적 대식세포 세포 계통일 수 있다. 또한, HSC는 장기간 HSC(LT-HSC) 및 단기간 HSC(ST-HSC)를 지칭하기도 한다. ST-HSC는 LT-HSC보다 더 활성이고 더 증식적이다. 그러나, LT-HSC는 무제한 자가 재생을 갖는 반면(즉, 이들은 성인기 전체에서 생존함), ST-HSC는 제한된 자가 재생을 갖는다(즉, 이들은 단지 제한된 기간 동안 생존함). 임의의 이들 HSC는 임의의 본원에 기재된 방법에 사용될 수 있다. 선택적으로, ST-HSC는 이들이 고도로 증식적이고 따라서 HSC 및 이들의 자손의 수를 신속하게 증가시키기 때문에 유용하다. 조혈 줄기세포는 선택적으로, 혈액 생성물로부터 수득된다. 혈액 생성물은 조혈 기원의 세포를 함유하는 신체 또는 이러한 신체의 기관으로부터 수득된 생성물을 포함한다. 이러한 공급원은 비-분획화된(un-fractionated) 골수, 탯줄, 말초 혈액(예를 들어 동원된 말초 혈액, 예를 들어 G-CSF 또는 Plerixafor®(AMD3100), 또는 G-CSF와 Plerixafor®(AMD3100)의 조합과 같은 동원제(mobilization agent)를 이용하여 동원된 말초 혈액), 간, 흉선, 림프 및 비장을 포함한다. 상기 언급된 미정제(crude) 또는 비-분획화된 혈액 생성물은 모두, 조혈 줄기세포 특징을 갖는 세포에 대해 당업자에게 공지된 방식으로 농화될 수 있다. 일 실시형태에서, HSC는 CD34+/CD38-/CD90+/CD45RA-를 특징으로 한다. 실시형태에서, HSC는 CD34+/CD90+/CD49f+ 세포를 특징으로 한다. 실시형태에서, HSC는 CD34+ 세포를 특징으로 한다. 실시형태에서, HSC는 CD34+/CD90+ 세포를 특징으로 한다. 실시형태에서, HSC는 CD34+/CD90+/CD45RA- 세포를 특징으로 한다.

[0408] 세포의 맥락에서 "확장" 또는 "확장하다"는 세포의 초기 세포 집단으로부터, 동일하거나 동일하지 않을 수 있는, 특징적인 세포 유형 또는 세포 유형들의 수의 증가를 지칭한다. 확장을 위해 사용된 초기 세포는 확장으로부터 생성된 세포와 동일하지 않을 수 있다.

[0409] "세포 집단"은 생물학적 공급원, 예를 들어 혈액 산물 또는 조직으로부터 단리되고, 하나를 초과하는 세포에서 유래된 진핵 포유류, 바람직하게는 인간 세포를 지칭한다.

[0410] 세포 집단의 맥락에 사용될 때 "농화된"은 하나 이상의 마커(예를 들어 CD34+)의 존재에 기반하여 선택된 세포 집단을 지칭한다.

[0411] 용어 "CD34+ 세포"는 표면에 CD34 마커를 발현하는 세포를 지칭한다. CD34+ 세포는 예를 들어 유세포 분석 및 형광 표지된 항-CD34 항체를 이용하여 검출되고 계수될 수 있다.

[0412] "CD34+ 세포로 농화된"은 세포 집단이 CD34 마커의 존재에 기반하여 선택되었다는 것을 의미한다. 따라서, 선택 방법 후의 세포 집단 내 CD34+ 세포의 백분율은 CD34 마커에 기반한 선택 단계 전의 초기 세포 집단 내 CD34+ 세포의 백분율보다 높다. 예를 들어 CD34+ 세포는 CD34+ 세포로 농화된 세포 집단 내 세포의 적어도 50%, 60%, 70%, 80% 또는 적어도 90%에 해당할 수 있다.

[0413] 용어 "F 세포" 및 "F-세포"는 태아 헤모글로빈을 함유 및/또는 생산(예를 들어 발현)하는 일반적으로 적혈구(예를 들어 적혈구)인 세포를 지칭한다. 예를 들어 F 세포는 검출 가능한 수준의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 예를 들어 F-세포는 적어도 5 피코그램의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 또 다른 실시예에서, F-세포는 적어도 6 피코그램의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 또 다른 실시예에서, F-세포는 적어도 7 피코그램의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 또 다른 실시예에서, F-세포는 적어도 8 피코그램의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 또 다른 실시예에서, F-세포는 적어도 9 피코그램의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 또 다른 실시예에서, F-세포는 적어도 10 피코그램의 태아 헤모글로빈을 함유하거나 생산하는 세포이다. 태아 헤모글로빈의 수준은 본원에 기재된 검정 또는 당분야에 공지된 다른 방법, 예를 들어 항-태반 헤모글로빈 검출 시약을 이용하는 유세포 분석, 고성능 액체 크로마토그라피, 질량 분광분석, 또는 효소-관련 면역흡착 측정을 사용하여 측정될 수 있다.

[0414] 달리 명시되지 않는 한, 모든 게놈 또는 염색체 좌표는 hg38에 따른다.

[0415] 상세한 설명

[0416] 본원에 기재된 gRNA 분자, 조성물 및 방법은 CRISPR/Cas9 시스템을 사용하는 진핵 세포에서의 게놈 편집에 관한 것이다. 특히, 본원에 기재된 gRNA 분자, 조성물 및 방법은 글로빈 수준의 조절에 관한 것이며, 예를 들어 글로빈 유전자 및 단백질의 발현 및 생산 조절에 유용하다. gRNA 분자, 조성물 및 방법은 혈색소병증의 치료에 유용할 수 있다.

[0417] I. gRNA 분자

[0418] gRNA 분자는 아래에서 더 충분히 기재된 바와 같이, 다수의 도메인을 가질 수 있으나, gRNA 분자는 전형적으로 적어도 하나의 crRNA 도메인(표적화 도메인을 포함함) 및 tracr를 포함한다. CRISPR 시스템의 성분으로서 사용되는, 본 발명의 gRNA 분자는 표적 부위에서 또는 그 부근에서 DNA를 변형(예를 들어 서열을 변형)하는데 유용하다. 이러한 변형은 예를 들어 표적 부위를 포함하는 유전자의 기능적 산물의 감소된 또는 제거된 발현을 야기하는 결실 및/또는 삽입을 포함한다. 이들 용도, 및 추가적 용도는 아래에서 더 충분히 기재된다.

[0419] 실시형태에서, 단분자, 또는 sgRNA는 바람직하게는 5'으로부터 3'까지:(표적 서열에 상보적인 표적화 도메인 및 플래그폴의 부분을 형성하는 영역(즉, crRNA 플래그폴 영역)을 함유하는) crRNA; 루프; 및 (crRNA 플래그폴 영역에 상보적인 도메인, 및 뉴클레아제 또는 다른 효과기 분자, 예를 들어 Cas 분자, 예를 들어 Cas9 분자에 추가적으로 결합하는 도메인을 함유하는) tracr를 포함하고, 다음의 형식을 취할 수 있다(5' 으로부터 3'까지):

[0420] [표적화 도메인]-[crRNA 플래그폴 영역]-[선택적인 제1 플래그폴 연장]-[루프]-[선택적인 제1 tracr 연장]-[tracr 플래그폴 영역]-[tracr 뉴클레아제 결합 도메인].

[0421] 실시형태에서, tracr 뉴클레아제 결합 도메인은 Cas 단백질, 예를 들어 Cas9 단백질에 결합한다.

[0422] 실시형태에서, 이분자, 또는 dgRNA는 2개의 폴리뉴클레오타이드를 포함한다; 제1, 바람직하게는 5’ 으로부터 3’ 까지:(표적 서열에 상보적인 표적화 도메인 및 플래그폴의 부분을 형성하는 영역을 함유하는) crRNA; 및 제2, 바람직하게는 5’ 으로부터 3’ 까지:(crRNA 플래그폴 영역에 상보적인 도메인, 및 뉴클레아제 또는 다른 효과기 분자, 예를 들어 Cas 분자, 예를 들어 Cas9 분자에 추가적으로 결합하는 도메인을 함유하는) tracr, 그리고 다음의 형식을 취할 수 있다(5’ 으로부터 3’ 까지):

[0423] 폴리뉴클레오타이드1(crRNA): [표적화 도메인]-[crRNA 플래그폴 영역]-[선택적인 제1 플래그폴 연장]-[선택적인 제2 플래그폴 연장]

[0424] 폴리뉴클레오타이드2(tracr): [선택적인 제1 tracr 연장]-[tracr 플래그폴 영역] - [tracr 뉴클레아제 결합 도메인]

[0425] 실시형태에서, tracr 뉴클레아제 결합 도메인은 Cas 단백질, 예를 들어 Cas9 단백질에 결합한다.

[0426] 일부 양태에서, 표적화 도메인은 본원에 기재된 표적화 도메인 서열, 예를 들어 표 1에 기재된 표적화 도메인, 또는 표 1에 기재된 표적화 도메인 서열의 17, 18, 19 또는 20(바람직하게는 20)개의 연속 뉴클레오타이드를 포함하거나 이로 구성된 표적화 도메인을 포함하거나 또는 이로 구성된다.

[0427] 일부 양태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 5’ 으로부터 3’ 까지: GUUUUAGAGCUA(SEQ ID NO: 182)를 포함한다.

[0428] 일부 양태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 5’ 으로부터 3’ 까지: GUUUAAGAGCUA(SEQ ID NO: 183)를 포함한다.

[0429] 일부 양태에서, 루프는 5’ 으로부터 3’ 까지: GAAA(SEQ ID NO: 186)를 포함한다.

[0430] 일부 양태에서, tracr는 5’ 으로부터 3’ 까지: UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO: 187) 를 포함하며, 바람직하게는 SEQ ID NO 182를 포함하는 gRNA 분자에 사용된다.

[0431] 일부 양태에서, tracr는 5’ 으로부터 3’ 까지: UAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO: 188) 를 포함하며, 바람직하게는 SEQ ID NO 183을 포함하는 gRNA 분자에 사용된다.

[0432] 일부 양태에서, gRNA는 또한 3' 말단에 추가적인 U 핵산을 포함할 수 있다. 예를 들어 gRNA는 3' 말단에 추가적인 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 U 핵산(SEQ ID NO: 249)을 포함할 수 있다. 실시형태에서, gRNA는 3' 말단에 추가적인 4개의 U 핵산을 포함한다. dgRNA의 경우, dgRNA의 하나 이상의 폴리뉴클레오타이드(예를 들어 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드)는 3' 말단에 추가적인 U 핵산을 포함할 수 있다. 예를 들어 dgRNA의 경우, dgRNA의 하나 이상의 폴리뉴클레오타이드(예를 들어 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드)는 추가적인 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 U 핵산(SEQ ID NO: 249)을 3' 말단에 포함할 수 있다. 실시형태에서, dgRNA의 경우, dgRNA의 하나 이상의 폴리뉴클레오타이드(예를 들어 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드)는 3' 말단에 추가적인 4개의 U 핵산을 포함한다. dgRNA의 실시형태에서, 오직 tracr를 포함하는 폴리뉴클레오타이드만 추가적인 U 핵산(들), 예를 들어 4개의 U 핵산을 포함한다. dgRNA의 실시형태에서, 오직 표적화 도메인을 포함하는 폴리뉴클레오타이드만 추가적인 U 핵산(들)을 포함한다. dgRNA의 실시형태에서, 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드 둘 모두 추가적인 U 핵산, 예를 들어 4개의 U 핵산을 포함한다.

[0433] 일부 양태에서, gRNA는 또한 3' 말단에 추가적인 A 핵산을 포함할 수 있다. 예를 들어 gRNA는 3' 말단에 추가적인 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 A 핵산(SEQ ID NO: 250)을 포함할 수 있다. 실시형태에서, gRNA는 3' 말단에 추가적인 4개의 A 핵산을 포함한다. dgRNA의 경우, dgRNA의 하나 이상의 폴리뉴클레오타이드(예를 들어 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드)는 3' 말단에 추가적인 A 핵산을 포함할 수 있다. 예를 들어 dgRNA의 경우, dgRNA의 하나 이상의 폴리뉴클레오타이드(예를 들어 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드)는 추가적인 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 A 핵산(SEQ ID NO: 250)을 3' 말단에 포함할 수 있다. 실시형태에서, dgRNA의 경우, dgRNA의 하나 이상의 폴리뉴클레오타이드(예를 들어 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드)는 3' 말단에 추가적인 4개의 A 핵산을 포함한다. dgRNA의 실시형태에서, 오직 tracr를 포함하는 폴리뉴클레오타이드만 추가적인 A 핵산(들), 예를 들어 4개의 A 핵산을 포함한다. dgRNA의 실시형태에서, 오직 표적화 도메인을 포함하는 폴리뉴클레오타이드만 추가적인 A 핵산(들)을 포함한다. dgRNA의 실시형태에서, 표적화 도메인을 포함하는 폴리뉴클레오타이드 및 tracr를 포함하는 폴리뉴클레오타이드 둘 모두 추가적인 U 핵산, 예를 들어 4개의 A 핵산을 포함한다.

[0434] 실시형태에서, gRNA 분자의 하나 이상의 폴리뉴클레오타이드는 5' 말단에 캡을 포함할 수 있다.

[0435] 실시형태에서, 단분자, 또는 sgRNA는, 바람직하게는 5’ 으로부터 3’ 까지: (표적 서열에 상보적인 표적화 도메인을 함유하는) crRNA; crRNA 플래그폴 영역; 제1 플래그폴 연장; 루프; (제1 플래그폴 연장의 적어도 일부분에 상보적인 도메인을 함유하는) 제1 tracr 연장; 및 (crRNA 플래그폴 영역에 상보적인 도메인, 및 Cas9 분자에 추가적으로 결합하는 도메인을 함유하는) tracr를 포함한다. 일부 양태에서, 표적화 도메인은, 본원에 기재된 표적화 도메인 서열, 예를 들어 표 1에 기재된 표적화 도메인, 또는 표 1에 기재된 표적화 도메인 서열의 17, 18, 19 또는 20(바람직하게는 20)개의 연속 뉴클레오타이드, 예를 들어 표 1에 기재된 표적화 도메인 서열의 3' 17, 18, 19 또는 20(바람직하게는 20)개의 연속 뉴클레오타이드를 포함하거나 이로 구성된 표적화 도메인을 포함한다.

[0436] 제1 플래그폴 연장 및/또는 제1 tracr 연장을 포함하는 양태에서, 플래그폴, 루프 및 tracr 서열은 상기 기재된 바와 같을 수 있다. 일반적으로 임의의 제1 플래그폴 연장 및 제1 tracr 연장은 그들이 상보적이라면 이용될 수 있다. 실시형태에서, 제1 플래그폴 연장 및 제1 tracr 연장은 3, 4, 5, 6, 7, 8, 9, 10개 이상의 상보적 뉴클레오타이드로 구성된다.

[0437] 일부 양태에서, 제1 플래그폴 연장은 5’ 으로부터 3’ 까지: UGCUG(SEQ ID NO: 184)를 포함한다. 일부 양태에서, 제1 플래그폴 연장은 SEQ ID NO: 184로 구성된다.

[0438] 일부 양태에서, 제1 tracr 연장은 5’ 으로부터 3’ 까지: CAGCA(SEQ ID NO: 189)를 포함한다. 일부 양태에서, 제1 tracr 연장은 SEQ ID NO: 189로 구성된다.

[0439] 실시형태에서, dgRNA는 2개의 핵산 분자를 포함한다. 일부 양태에서, dgRNA는 바람직하게는 5’ 으로부터 3’ 까지: 표적 서열에 상보적인 표적화 도메인; crRNA 플래그폴 영역; 선택적으로 제1 플래그폴 연장; 및 선택적으로, 제2 플래그폴 연장을 함유하는 제1 핵산; 및 바람직하게는 5’ 으로부터 3’ 까지: 선택적으로 제1 tracr 연장; 및 (crRNA 플래그폴 영역에 상보적인 도메인, 및 Cas, 예를 들어 Cas9, 분자에 추가로 결합하는 도메인을 함유하는) tracr를 포함하는 제2 핵산(본원에서 tracr로 지칭될 수 있고, Cas 분자, 예를 들어 Cas9 분

자에 결합하는 적어도 하나의 도메인을 포함함)을 포함한다. 제2 핵산은 추가적으로 3' 말단(예를 들어 tracr에 대해 3')에 추가적인 U 핵산을 포함할 수 있다. 예를 들어 tracr는 3' 말단(예를 들어 tracr에 대해 3')에 추가적인 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 U 핵산(SEQ ID NO: 249)을 포함할 수 있다. 제2 핵산은 추가적으로 또는 교대로 3' 말단(예를 들어 tracr에 대해 3')에 추가적인 A 핵산을 포함할 수 있다. 예를 들어 tracr는 3' 말단(예를 들어 tracr에 대해 3')에 추가적인 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 A 핵산(SEQ ID NO: 250)을 포함할 수 있다. 일부 양태에서, 표적화 도메인은 본원에 기재된 표적화 도메인 서열, 예를 들어 표 1에 기재된 표적화 도메인, 또는 표 1에 기재된 표적화 도메인 서열의 17, 18, 19 또는 20(바람직하게는 20)개의 연속 뉴클레오타이드를 포함하거나 이로 구성된 표적화 도메인을 포함한다.

[0440] dgRNA를 수반하는 양태에서, crRNA 플래그폴 영역, 선택적인 제1 플래그폴 연장, 선택적인 제1 tracr 연장 및 tracr 서열은 상기 기재된 바와 같을 수 있다.

[0441] 일부 양태에서, 선택적인 제2 플래그폴 연장은 5’ 으로부터 3’ 까지: UUUUG(SEQ ID NO: 185)를 포함한다.

[0442] 실시형태에서, gRNA 분자의 3' 1, 2, 3, 4 또는 5 뉴클레오타이드, 5' 1, 2, 3, 4 또는 5 뉴클레오타이드, 또는 3' 및 5' 둘 모두의 1, 2, 3, 4 또는 5 뉴클레오타이드(및 dgRNA 분자의 경우, 표적화 도메인을 포함하는 폴리뉴클레오타이드 및/또는 tracr를 포함하는 폴리뉴클레오타이드)는 아래 섹션 XIII에서 더 충분히 기재된 바와 같이, 변형된 핵산이다.

[0443] 도메인이 아래에 간략히 논의된다:

[0444] 1) 표적화 도메인:

[0445] 표적화 도메인의 선택에 대한 안내는 예를 들어 문헌[Fu Y el al. NAT BIOTECHNOL 2014(doi: 10.1038/nbt.2808)] 및 문헌[Sternberg SH el al. Nature 2014(doi: 10.1038/nature13011)]에서 찾을 수 있다.

[0446] 표적화 도메인은 표적 핵산 상의 표적 서열에 상보적인, 예를 들어 적어도 80, 85, 90, 95 또는 99% 상보적인, 예를 들어 완전히 상보적인 뉴클레오타이드 서열을 포함한다. 표적화 도메인은 RNA 분자의 부분이며, 따라서 염기 우라실(U)을 포함할 것인 반면, gRNA 분자를 인코딩하는 임의의 DNA는 염기 티민(T)을 포함할 것이다. 이론에 얽매이기를 원하지 않지만, 표적 서열과 표적화 도메인의 상보성이 표적 핵산과 gRNA 분자/Cas9 분자 복합체의 상호 작용의 특이성에 기여한다고 여겨진다. 표적화 도메인과 표적 서열 쌍에서, 표적화 도메인에서의 우라실 염기는 표적 서열에서의 아데닌 염기와 쌍을 이룰 것으로 이해된다.

[0447] 일 실시형태에서, 표적화 도메인은 5 내지 50, 예를 들어 10 내지 40, 예를 들어 10 내지 30, 예를 들어 15 내지 30, 예를 들어 15 내지 25개 뉴클레오타이드 길이이다. 일 실시형태에서, 표적화 도메인은 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25개 뉴클레오타이드 길이이다. 일 실시형태에서, 표적화 도메인은 16개 뉴클레오타이드 길이이다. 일 실시형태에서, 표적화 도메인은 17개 뉴클레오타이드 길이이다. 일 실시형태에서, 표적화 도메인은 18개 뉴클레오타이드 길이이다. 일 실시형태에서, 표적화 도메인은 19개 뉴클레오타이드 길이이다. 일 실시형태에서, 표적화 도메인은 20개 뉴클레오타이드 길이이다. 실시형태에서, 상기 언급된 16, 17, 18, 19 또는 20개 뉴클레오타이드는 표 1에 기재된 표적화 도메인으로부터 5'-16, 17, 18, 19 또는 20개 뉴클레오타이드를 포함한다. 실시형태에서, 상기 언급된 16, 17, 18, 19 또는 20개 뉴클레오타이드는 표 1에 기재된 표적화 도메인으로부터 3'- 16, 17, 18, 19 또는 20개 뉴클레오타이드를 포함한다.

[0448] 임의의 이론으로 결부시키고자 하는 것은 아니지만, 표적화 도메인의 3' 말단에 배치된 표적화 도메인의 8, 9, 10, 11 또는 12개의 핵산은 표적 서열을 표적화하는데 중요하며, 따라서 표적화 도메인의 “코어” 영역으로 지칭될 수 있다고 여겨진다. 실시형태에서, 코어 도메인은 표적 서열과 완전히 상보적이다.

[0449] 표적화 도메인이 상보적인 표적 핵산의 가닥은 본원에서 표적 서열로 지칭된다. 일부 양태에서, 표적 서열은 염색체 상에 배치되고, 예를 들어 유전자 내의 표적이다. 일부 양태에서, 표적 서열은 유전자의 엑손 내에 배치된다. 일부 양태에서, 표적 서열은 유전자의 인트론 내에 배치된다. 일부 양태에서, 표적 서열은 조절 요소의 결합 부위, 예를 들어 관심 유전자의 프로모터 또는 전사 인자 결합 부위를 포함하거나 인접한다(예를 들어 10, 20, 30, 40, 50, 100, 200, 300, 400, 500 또는 1000개 핵산 내). 도메인의 뉴클레오타이드의 일부 또는 전부는 변형, 예를 들어 본원 섹션 XIII에서 확인되는 변형을 가질 수 있다.

[0450] 2) crRNA 플래그폴 영역:

[0451] 플래그폴은 crRNA 및 tracr 둘 모두로부터의 부분을 함유한다. crRNA 플래그폴 영역은 tracr의 일부분과 상보적

이고, 실시형태에서, 적어도 일부 생리적 조건, 예를 들어 정상적인 생리적 조건 하에서 이중 영역을 형성하기에 충분한 tracr의 일부분에 대한 상보성을 갖는다. 실시형태에서, crRNA 플래그폴 영역은 5 내지 30뉴클레오타이드 길이이다. 실시형태에서, crRNA 플래그폴 영역은 5 내지 25 뉴클레오타이드 길이이다. crRNA 플래그폴 영역은 박테리아 CRISPR 배열로부터의 반복 서열의 자연적 발생 부분과 상동성을 공유할 수 있거나, 이로부터 유래될 수 있다. 실시형태에서, 본원에 개시된 crRNA 플래그폴 영역, 예를 들어 S. 피오게네스 또는 S. 써모필루스, crRNA 플래그폴 영역과 적어도 50%의 상동성을 갖는다.

[0452] 실시형태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 SEQ ID NO: 182를 포함한다. 실시형태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 SEQ ID NO: 182와 적어도 50%, 60%, 70%, 80%, 85%, 90%, 95% 또는 99%의 상동성을 갖는 서열을 포함한다. 실시형태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 SEQ ID NO: 182의 적어도 5, 6, 7, 8, 9, 10 또는 11개 뉴클레오타이드를 포함한다. 실시형태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 SEQ ID NO: 183을 포함한다. 실시형태에서, 플래그폴은 SEQ ID NO: 183과 적어도 50%, 60%, 70%, 80%, 85%, 90%, 95% 또는 99%의 상동성을 갖는 서열을 포함한다. 실시형태에서, 플래그폴, 예를 들어 crRNA 플래그폴 영역은 SEQ ID NO: 183의 적어도 5, 6, 7, 8, 9, 10 또는 11개 뉴클레오타이드를 포함한다.

[0453] 도메인의 뉴클레오타이드의 일부 또는 전부는 변형, 예를 들어 본원 섹션 XIII에 기재된 변형을 가질 수 있다.

[0454] 3) 제1 플래그폴 연장

[0455] 제1 tracr 연장을 포함하는 tracr가 사용되는 경우, crRNA는 제1 플래그폴 연장을 포함할 수 있다. 일반적으로 임의의 제1 플래그폴 연장 및 제1 tracr 연장은 그들이 상보적이라면 이용될 수 있다. 실시형태에서, 제1 플래그폴 연장 및 제1 tracr 연장은 3, 4, 5, 6, 7, 8, 9, 10개 이상의 상보적 뉴클레오타이드로 구성된다.

[0456] 제1 플래그폴 연장은 제1 tracr 연장의 뉴클레오타이드와 상보적인, 예를 들어 80%, 85%, 90%, 95% 또는 99%, 예를 들어 완전히 상보적인 뉴클레오타이드를 포함할 수 있다. 일부 양태에서, 제1 tracr 연장의 상보적인 뉴클레오타이드와 혼성화되는 제1 플래그폴 연장 뉴클레오타이드는 연속적이다. 일부 양태에서, 제1 tracr 연장의 상보적인 뉴클레오타이드와 혼성화되는 제1 플래그폴 연장 뉴클레오타이드는 불연속적이고, 예를 들어 제1 tracr 연장의 뉴클레오타이드와 염기쌍 형성하지 않는 뉴클레오타이드에 의해 분리된 2개 이상의 혼성화 영역을 포함한다. 일부 양태에서, 제1 플래그폴 연장은 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개 이상의 뉴클레오타이드를 포함한다. 일부 양태에서, 제1 플래그폴 연장은 5’으로부터 3’까지: UGCUG(SEQ ID NO: 184)를 포함한다. 일부 양태에서, 제1 플래그폴 연장은 SEQ ID NO: 184로 구성된다. 일부 양태에서, 제1 플래그폴 연장은 SEQ ID NO: 184와 적어도 80%, 85%, 90%, 95% 또는 99% 상동성을 갖는 핵산을 포함한다.

[0457] 제1 tracr 연장의 뉴클레오타이드의 일부 또는 전부는 변형, 예를 들어 본원 섹션 XIII에서 확인되는 변형을 가질 수 있다.

[0458] 3) 루프

[0459] 루프는 sgRNA의 crRNA 플래그폴 영역(또는 존재하는 경우, 선택적으로 제1 플래그폴 연장)을 tracr(또는 존재하는 경우, 선택적으로 제1 tracr 연장)와 연결하는 역할을 한다. 루프는 crRNA 플래그폴 영역과 tracr를 공유 또는 비공유 결합으로 연결할 수 있다. 실시형태에서, 연결은 공유결합이다. 실시형태에서, 루프는 crRNA 플래그폴 영역과 tracr를 공유결합으로 결합시킨다. 실시형태에서, 루프는 제1 플래그폴 연장과 제1 tracr 연장을 공유결합으로 결합시킨다. 실시형태에서, 루프는 crRNA 플래그폴 영역과 crRNA 플래그폴 영역에 혼성화되는 tracr의 도메인 사이에 삽입된 공유 결합이거나 이를 포함한다. 전형적으로, 루프는 하나 이상의, 예를 들어 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 뉴클레오타이드를 포함한다.

[0460] dgRNA 분자에서 2개의 분자는 crRNA의 적어도 일부분(예를 들어 crRNA 플래그폴 영역)과 tracr의 적어도 일부분(예를 들어 crRNA 플래그폴 영역에 상보적인 tracr의 도메인) 사이의 혼성화에 의해 관련될 수 있다.

[0461] 매우 다양한 루프가 sgRNA에서의 사용에 적합하다. 루프는 공유 결합으로 구성될 수 있거나, 1개 또는 수개의 뉴클레오타이드, 예를 들어 1, 2, 3, 4 또는 5개 뉴클레오타이드 길이만큼 짧을 수 있다. 실시형태에서, 루프는 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 또는 25개 이상의 뉴클레오타이드 길이이다. 실시형태에서, 루프는 2 내지 50, 2 내지 40, 2 내지 30, 2 내지 20, 2 내지 10 또는 2 내지 5개 뉴클레오타이드 길이이다. 실시형태에서, 루프는 자연적 발생 서열과 상동성을 공유하거나 또는 이로부터 유래된다. 실시형태에서, 루프는 본원에 개시된 루프와 적어도 50%의 상동성을 갖는다. 실시형태에서, 루프는 SEQ ID NO: 186을 포함한다.

[0462] 도메인의 뉴클레오타이드의 일부 또는 전부는 변형, 예를 들어 본원 섹션 XIII에 기재된 변형을 가질 수 있다.

[0463] 4) 제2 플래그폴 연장

[0464] 실시형태에서, dgRNA는 본원에서 제2 플래그폴 연장으로 지칭되는 추가적인 서열(crRNA 플래그폴 영역에 대해 3'), 또는 존재하는 경우, 제1 플래그폴 연장을 포함할 수 있다. 실시형태에서, 제2 플래그폴 연장은 2 내지 10, 2 내지 9, 2 내지 8, 2 내지 7, 2 내지 6, 2 내지 5 또는 2 내지 4개 뉴클레오타이드 길이이다. 실시형태에서, 제2 플래그폴 연장은 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 이상 뉴클레오타이드 길이이다. 실시형태에서, 제2 플래그폴 연장은 SEQ ID NO: 185를 포함한다.

[0465] 5) Tracr:

[0466] tracr는 뉴클레아제(예를 들어 Cas9) 결합에 필요한 핵산 서열이다. 임의의 이론으로 결부시키고자 하는 것은 아니지만, 각각의 Cas9 종은 특정 tracr 서열과 관련되어 있다고 여겨진다. Tracr 서열은 sgRNA 및 dgRNA 시스템 둘 모두에 이용된다. 실시형태에서, tracr는 *S. 피오게네스* tracr로부터의 서열, 또는 이로부터 유래된 서열을 포함한다. 일부 양태에서, tracr는 crRNA의 플래그폴 부분에 혼성화되는 일부분을 가지며, 예를 들어 적어도 일부 생리적 조건 하에서 이중 영역을 형성하기에 충분한 crRNA 플래그폴 영역에 대한 상보성을 갖는다(때때로 본원에서 tracr 플래그폴 영역 또는 crRNA 플래그폴 영역에 상보적인 tracr 도메인으로 지칭됨). 실시형태에서, crRNA 플래그폴 영역과 혼성화되는 tracr 도메인은 crRNA 플래그폴 영역의 상보적인 뉴클레오타이드와 혼성화되는 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개 뉴클레오타이드를 포함한다. 일부 양태에서, crRNA 플래그폴 영역의 상보적인 뉴클레오타이드와 혼성화되는 tracr 뉴클레오타이드는 연속적이다. 일부 양태에서, crRNA 플래그폴 영역의 상보적인 뉴클레오타이드와 혼성화되는 tracr 뉴클레오타이드는 불연속적이고, 예를 들어 crRNA 플래그폴 영역의 뉴클레오타이드와 염기쌍 형성하지 않는 뉴클레오타이드에 의해 분리된 2개 이상의 혼성화 영역을 포함한다. 일부 양태에서, crRNA 플래그폴 영역에 혼성화되는 tracr의 부분은 5’ 으로부터 3’ 까지 UAGCAAGUUAAAA(SEQ ID NO: 191)를 포함한다. 일부 양태에서, crRNA 플래그폴 영역에 혼성화되는 tracr 부분은 5’ 으로부터 3’ 까지 UAGCAAGUUUAAA(SEQ ID NO: 192)를 포함한다. 실시형태에서, crRNA 플래그폴 영역과 혼성화되는 서열은 tracr 상에서 뉴클레아제, 예를 들어 Cas 분자, 예를 들어 Cas9 분자와 추가적으로 결합하는 tracr의 서열의 5'에 배치된다.

[0467] tracr는 뉴클레아제, 예를 들어 Cas 분자, 예를 들어 Cas9 분자에 추가적으로 결합하는 도메인을 추가로 포함한다. 임의의 이론으로 결부시키고자 하는 것은 아니지만, 상이한 종으로부터의 Cas9는 상이한 tracr 서열에 결합한다고 여겨진다. 일부 양태에서, tracr는 *S. 피오게네스* Cas9 분자에 결합하는 서열을 포함한다. 일부 양태에서, tracr는 본원에 개시된 Cas9 분자에 결합하는 서열을 포함한다. 일부 양태에서, Cas9 분자에 추가적으로 결합하는 도메인은 5’ 으로부터 3’ 까지: UAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO: 193)를 포함한다. 일부 양태에서, Cas9 분자에 추가적으로 결합하는 도메인은 5’ 으로부터 3’ 까지: UAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU(SEQ ID NO: 194)를 포함한다.

[0468] 일부 실시형태에서, tracr는 SEQ ID NO: 187을 포함한다. 일부 실시형태에서, tracr는 SEQ ID NO: 188을 포함한다.

[0469] tracr의 뉴클레오타이드의 일부 또는 전부는 변형, 예를 들어 본원 섹션 XIII에서 확인되는 변형을 가질 수 있다. 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 gRNA의 5' 말단, 3' 말단 또는 5' 및 3' 말단 둘 모두에 역위된 무염기성(abasic) 잔기를 포함한다. 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 폴리뉴클레오타이드의 5' 말단에 있는 잔기들 사이의 하나 이상의 포스포로티오에이트 결합, 예를 들어 처음 2개의 5' 잔기들 사이, 처음 3개의 5' 잔기들 각각의 사이, 처음 4개의 5' 잔기들 각각의 사이, 또는 처음 5개의 5' 잔기들 각각의 사이의 포스포로티오에이트 결합을 포함한다. 실시형태에서, gRNA 또는 gRNA 성분은 폴리뉴클레오타이드의 3' 말단에 있는 잔기들 사이의 하나 이상의 포스포로티오에이트 결합, 예를 들어 처음 2개의 3' 잔기들 사이, 처음 3개의 3' 잔기들 각각의 사이, 처음 4개의 3' 잔기들 각각의 사이, 또는 처음 5개의 3' 잔기들 각각의 사이의 포스포로티오에이트 결합을 대안적으로 또는 추가로 포함할 수 있다. 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 처음 4개의 5' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 5' 말단(들)에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 및 처음 4개의 3' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 3' 말단(들)에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨)을 포함한다. 실시형태에서, 상기 기재된 포스포로티오에이트 변형 중 임의의 것

은 폴리뉴클레오타이드의 5' 말단, 3' 말단 또는 5' 및 3' 말단 둘 모두에서의 역위된 무염기성 잔기와 조합된다. 이러한 실시형태에서, 역위된 무염기성 뉴클레오타이드는 포스페이트 결합 또는 포스포로티오에이트 결합에 의해 5' 및/또는 3'뉴클레오타이드에 연결될 수 있다. 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 2'O-메틸 변형을 포함하는 하나 이상의 뉴클레오타이드를 포함한다. 실시형태에서, 처음 1, 2, 3 이상의 5' 잔기들 각각은 2'O-메틸 변형을 포함한다. 실시형태에서, 처음 1, 2, 3 이상의 3' 잔기들 각각은 2'O-메틸 변형을 포함한다. 실시형태에서, 말단 방향으로 4 번째, 말단 방향으로 3 번째, 및 말단 방향으로 2 번째 3' 잔기들은 2'O-메틸 변형을 포함한다. 실시형태에서, 처음 1, 2, 3 이상의 5' 잔기들 각각은 2'O-메틸 변형을 포함하고, 처음 1, 2, 3 이상의 3' 잔기들 각각은 2'O-메틸 변형을 포함한다. 실시형태에서, 처음 3개의 5' 잔기들 각각은 2'O-메틸 변형을 포함하고, 처음 3개의 3' 잔기들 각각은 2'O-메틸 변형을 포함한다. 실시형태에서, 처음 3개의 5' 잔기들 각각은 2'O-메틸 변형을 포함하고, 말단 방향으로 4 번째, 말단 방향으로 3 번째, 및 말단 방향으로 2 번째 3' 잔기들은 2'O-메틸 변형을 포함한다. 실시형태에서, 예를 들어 상기 기재된 2'O-메틸 변형 중 임의의 것은, 예를 들어 상기 기재된 하나 이상의 포스포로티오에이트 변형 및/또는 예를 들어 상기 기재된 하나 이상의 역위된 무염기성의 변형과 조합될 수 있다. 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 처음 4개의 5' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 4개의 3' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 3개의 5' 잔기들 각각에서 2'O-메틸 변형, 및 처음 3개의 3' 잔기들 각각에서 2'O-메틸 변형을 포함하며, 예를 들어 이로 구성된다. 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 처음 4개의 5' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 4개의 3' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 3개의 5' 잔기들 각각에서 2'O-메틸 변형, 및 말단 방향으로 4 번째, 말단 방향으로 3 번째, 및 말단 방향으로 2 번째 3' 잔기 각각에서 2'O-메틸 변형을 포함하며, 예를 들어 이로 구성된다.

[0470] 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 처음 4개의 5' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 4개의 3' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 3개의 5' 잔기들 각각에서 2'O-메틸 변형, 처음 3개의 3' 잔기들 각각에서 2'O-메틸 변형 및 5' 및 3' 말단 각각에 추가적인 역위된 무염기성 잔기를 포함하며, 예를 들어 이로 구성된다.

[0471] 실시형태에서, gRNA(예를 들어 sgRNA 또는 dgRNA의 tracr 및/또는 crRNA), 예를 들어 상기 기재된 gRNA 또는 gRNA 성분 중 임의의 것은 처음 4개의 5' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 4개의 3' 잔기들 각각의 사이의 포스포로티오에이트 결합(예를 들어 폴리뉴클레오타이드(들)의 5' 말단에 3개의 포스포로티오에이트 결합을 포함하고, 예를 들어 이로 구성됨), 처음 3개의 5' 잔기들 각각에서 2'O-메틸 변형, 및 말단 방향으로 4 번째, 말단 방향으로 3 번째, 및 말단 방향으로 2 번째 3' 잔기 각각에서 2'O-메틸 변형, 및 5' 및 3' 말단 각각에 추가적인 역위된 무염기성 잔기를 포함하며, 예를 들어 이로 구성된다.

[0472] 실시형태에서, gRNA는 dgRNA이고:

[0473] crRNA:

[0474] mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAGCUAU*mG*mC*mU(SEQ ID NO: 251)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타내고;

[0475] tracr:

[0476] AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU(SEQ ID NO: 224)(선택적으로

5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 포함하며, 예를 들어 이로 구성된다.

[0477] 실시형태에서, gRNA는 dgRNA이고:

[0478] crRNA:

[0479] mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAGCUAU*mG*mC*mU(SEQ ID NO: 251)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타내고;

[0480] tracr:

[0481] mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU*mU*mU*mU(SEQ ID NO: 246)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타낸다.

[0482] 실시형태에서, gRNA는 dgRNA이고:

[0483] crRNA:

[0484] mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGUU*mU*mU*mG(SEQ ID NO: 252)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타내고;

[0485] tracr:

[0486] AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU(SEQ ID NO: 224)(선택적으로 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 포함하며, 예를 들어 이로 구성된다.

[0487] 실시형태에서, gRNA는 dgRNA이고:

[0488] crRNA:

[0489] mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGUU*mU*mU*mG(SEQ ID NO: 252)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타내고;

[0490] tracr:

[0491] mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU*mU*mU*mU(SEQ ID NO: 246)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)을 나타낸다.

[0492] 실시형태에서, gRNA는 dgRNA이고:

[0493] crRNA:

[0494] NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGUUUUG(SEQ ID NO: 253)를 포함하고, 예를 들어 이로 구성되며, 여기서 N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타내고;

[0495] tracr:

[0496] mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU*mU*mU*mU(SEQ ID NO: 246)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)을 나타낸다.

[0497] 실시형태에서, gRNA는 sgRNA이고:

[0498] NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU(S

EQ ID NO: 254)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타낸다.

[0499] 실시형태에서, gRNA는 sgRNA이고:

[0500] mN*mN*mN*NNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU(SEQ ID NO: 255)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타낸다.

[0501] 실시형태에서, gRNA는 sgRNA이고:

[0502] mN*mN*mN*NNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U(SEQ ID NO: 256)를 포함하고, 예를 들어 이로 구성되며, 여기서 m은 2'O-메틸 변형을 갖는 염기를 나타내고, *은 포스포로티오에이트 결합을 나타내고, N은 예를 들어 본원에 기재된 표적화 도메인의 잔기(선택적으로, 5' 및/또는 3' 말단에 역위된 무염기성 잔기를 가짐)를 나타낸다.

[0503] 6) 제1 Tracr 연장

[0504] gRNA가 제1 플래그폴 연장을 포함하는 경우, tracr는 제1 tracr 연장을 포함할 수 있다. 제1 tracr 연장은 제1 플래그폴 연장의 뉴클레오타이드와 상보적인, 예를 들어 80%, 85%, 90%, 95% 또는 99%, 예를 들어 완전히 상보적인 뉴클레오타이드를 포함할 수 있다. 일부 양태에서, 제1 플래그폴 연장의 상보적인 뉴클레오타이드와 혼성화되는 제1 tracr 연장 뉴클레오타이드는 연속적이다. 일부 양태에서, 제1 플래그폴 연장의 상보적인 뉴클레오타이드와 혼성화되는 제1 tracr 연장 뉴클레오타이드는 불연속적이고, 예를 들어 제1 플래그폴 연장의 뉴클레오타이드와 염기쌍 형성하지 않는 뉴클레오타이드에 의해 분리된 2개 이상의 혼성화 영역을 포함한다. 일부 양태에서, 제1 tracr 연장은 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개 이상의 뉴클레오타이드를 포함한다. 일부 양태에서, 제1 tracr 연장은 SEQ ID NO: 189를 포함한다. 일부 양태에서, 제1 tracr 연장은 SEQ ID NO: 189와 적어도 80%, 85%, 90%, 95% 또는 99% 상동인 핵산을 포함한다.

[0505] 제1 tracr 연장의 뉴클레오타이드의 일부 또는 전부는 변형, 예를 들어 본원 섹션 XIII에서 확인되는 변형을 가질 수 있다.

[0506] 일부 실시형태에서, sgRNA는 5’ 으로부터 3’ 까지 표적화 도메인에 대해 3'에 배치되는:

a)
GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA
AGUGGCACCGAGUCGGUGC (SEQ ID NO: 195);

b)

GUUUAAGAGCUAGAAAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA
AGUGGCACCGAGUCGGUGC (SEQ ID NO: 196);

c)

GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
AACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 197);

d)

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUC
[0507] AACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 198);

[0508] e) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개의 우라실(U) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6, 또는 7개의 우라실(U) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 d) 중 임의의 것;

[0509] f) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개의 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6, 또는 7개의 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 d) 중 임의의 것; 또는

[0510] g) 5' 말단(예를 들어 5' 말단, 예를 들어 표적화 도메인에 대해 5')에 적어도 1, 2, 3, 4, 5, 6 또는 7개의 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6, 또는 7개의 아데닌(A) 뉴클레오타이드를 추가로 포함하

는, 임의의 상기 a) 내지 f) 중 임의의 것을 포함할 수 있다. 실시형태에서, 임의의 상기 a) 내지 g) 중 임의의 것은 표적화 도메인에 대해 3'에 직접 배치된다.

[0511] 실시형태에서, 본 발명의 sgRNA는 5’으로부터 3’까지: [표적화 도메인]-

[0512] GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 231).

[0513] 를 포함하며, 예를 들어 이로 구성된다.

[0514] 실시형태에서, 본 발명의 sgRNA는 5’으로부터 3’까지: [표적화 도메인]-

[0515] GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUC AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 227).

[0516] 를 포함하며, 예를 들어 이로 구성된다.

[0517] 일부 실시형태에서, dgRNA는:

[0518] 5’으로부터 3’까지, 바람직하게는 표적화 도메인에 대해 3'에 직접 배치된:

a) GUUUUAGAGCUA (SEQ ID NO: 182);

b) GUUUAAGAGCUA (SEQ ID NO: 183);

c) GUUUUAGAGCUAUGCUG (SEQ ID NO: 199);

d) GUUUAAGAGCUAUGCUG (SEQ ID NO: 200);

e) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 201);

f) GUUUAAGAGCUAUGCUGUUUUG (SEQ ID NO: 202); 또는

[0519] g) GUUUUAGAGCUAUGCU (SEQ ID NO: 226):

[0520] 를 포함하는 crRNA, 및 5’으로부터 3’까지:

a)
UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGC (SEQ ID NO: 187);

b)
UAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGC (SEQ ID NO: 188);

c)
CAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGA GUCGGUGC (SEQ ID NO: 203);

d)
CAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGA GUCGGUGC (SEQ ID NO: 204);

e)
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 224);

f)
AACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 225);

g)
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC

[0521] GAGUCGGUGC (SEQ ID NO: 232)

h)
GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUC
AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 227);

i)
AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG
UCGGUGCUUU (SEQ ID NO: 228);

j)
GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACU
[0522] UGAAAAAGUGGCACCGAGUCGGUGCUUU (SEQ ID NO: 229);

[0523] k) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개의 우라실(U) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6, 또는 7개의 우라실(U) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 j) 중 임의의 것;

[0524] l) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개의 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6, 또는 7개의 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 j) 중 임의의 것; 또는

[0525] m) 5' 말단(예를 들어 5'말단)에 적어도 1, 2, 3, 4, 5, 6 또는 7개의 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6, 또는 7개의 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 l) 중 임의의 것;을 포함하는 tracr를 포함할 수 있다.

[0526] 실시형태에서, 상기 k)의 서열은, 예를 들어 U6 프로모터가 전사에 사용되는 경우 3'서열 UUUUUU를 포함한다. 실시형태에서, 상기 k)의 서열은, 예를 들어 HI 프로모터가 전사에 사용되는 경우 3'서열 UUUU를 포함한다. 실시형태에서, 상기 k)의 서열은, 예를 들어 사용된 pol-III 프로모터의 종결 신호에 따라, 가변적인 개수의 3'U를 포함한다. 실시형태에서, 상기 k)의 서열은 T7 프로모터가 사용되는 경우 DNA 주형으로부터 유래된 가변적인 3'서열을 포함한다. 실시형태에서, 상기 k)의 서열은, 예를 들어 시험관내 전사가 RNA 분자를 생성하기 위해 사용되는 경우, DNA 주형으로부터 유래된 가변적인 3'서열을 포함한다. 실시형태에서, 상기 k)의 서열은 예를 들어 pol-II 프로모터가 전사를 추진하는데 사용되는 경우, DNA 주형으로부터 유래된 가변적인 3'서열을 포함한다.

[0527] 실시형태에서, crRNA는 표적화 도메인 및, 표적화 도메인에 대해 3'에 배치된(예를 들어 표적화 도메인에 대해 3'에 직접 배치된) SEQ ID NO: 201을 포함하는, 예를 들어 이로 구성되는 서열을 포함하고, 예를 들어 이로 구성되고, tracr는 AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU(SEQ ID NO: 224)를 포함하며, 예를 들어 이로 구성된다.

[0528] 실시형태에서, crRNA는 표적화 도메인 및, 표적화 도메인에 대해 3'에 배치된(예를 들어 표적화 도메인의 3'에 직접 배치된) SEQ ID NO: 202를 포함하는, 예를 들어 이로 구성되는 서열을 포함하고, 예를 들어 이로 구성되고, tracr는 AACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU(SEQ ID NO: 225)를 포함하며, 예를 들어 이로 구성된다.

[0529] 실시형태에서, crRNA는 표적화 도메인 및, 표적화 도메인에 대해 3'에 배치된(예를 들어 표적화 도메인에 대해 3'에 직접 배치된) GUUUUAGAGCUAUGCU(SEQ ID NO: 226)을 포함하는, 예를 들어 이로 구성되는 서열을 포함하고, tracr는 GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU(SEQ ID NO: 227)를 포함하며, 예를 들어 이로 구성된다.

[0530] 실시형태에서, crRNA는 표적화 도메인 및, 표적화 도메인에 대해 3'에 배치된(예를 들어 표적화 도메인에 대해 3'에 직접 배치된) GUUUUAGAGCUAUGCU(SEQ ID NO: 226)를 포함하는, 예를 들어 이로 구성되는 서열을 포함하고, tracr는 AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU(SEQ ID NO: 228)를 포함하며, 예를 들어 이로 구성된다.

[0531] 실시형태에서, crRNA는 표적화 도메인 및, 표적화 도메인에 대해 3'에 배치된(예를 들어 표적화 도메인에 대해 3'에 직접 배치된) GUUUUAGAGCUAUGCUGUUUUG(SEQ ID NO: 201)를 포함하는, 예를 들어 이로 구성되는 서열을 포함하고, tracr는 GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU(SEQ ID NO: 229)를 포함하며, 예를 들어 이로 구성된다.

[0532] **II. 비결실 HPFH 영역에 관한 gRNA 표적화 도메인**

[0533] 본 발명의 다양한 양태에서, 예를 들어 글로빈 유전자, 예를 들어 태아 헤모글로빈 유전자 또는 헤모글로빈 베타 유전자의 발현을 변경시키는 데 사용하기 위한, 본 발명의 gRNA 분자에 대한 비결실 HPFH 영역에 관한 표적화 도메인이 하기 표에 제공된다.

[0534] [표 1]

[0535] 비결실 HPFH 영역에 관한 gRNA 표적화 도메인. SEQ ID NO:는 gRNA 표적화 도메인 서열을 지칭한다.

[0536]

표적화 도메인 ID	표적 프로모터 영역	gRNA 표적화 도메인 서열	게놈 위치 (hg38) 1	가닥	게놈 위치 (hg38) 2 (존재한다면)	가닥	SEQ ID NO:
HBG1 프로모터 영역 내에서만 표적 서열을 갖는 gRNA 표적화 도메인							
GCR-0001	HBG1	AGUCCUGGUAUCCUCUAUGA	chr11:5250169-5250189	-			1
GCR-0002	HBG1	AAUUAGCAGUAUCCUCUUGG	chr11:5250063-5250083	-			2
GCR-0003	HBG1	AGAAUAAAUUAGAGAAAAAC	chr11:5250123-5250143	-			3
GCR-0004	HBG1	AAAAAUUAGCAGUAUCCUCU	chr11:5250066-5250086	-			4
GCR-0005	HBG1	AAAAUUAGCAGUAUCCUCUU	chr11:5250065-5250085	-			5
GCR-0006	HBG1	AAAAACUGGAAUGACUGAAU	chr11:5250109-5250129	-			6
GCR-0007	HBG1	CUCCCAUCAUAGAGGAUACC	chr11:5250163-5250183	+			7
GCR-0008	HBG1	GGAGAAGGAAACUAGCUAAA	chr11:5250147-5250167	-			8
GCR-0009	HBG1	GUUUCCUUCUCCCAUCAUAG	chr11:5250155-5250175	+			9
GCR-0010	HBG1	GGGAGAAGGAAACUAGCUAA	chr11:5250148-5250168	-			10
GCR-0011	HBG1	CACUGGAGCUAGAGACAAGA	chr11:5250213-5250233	-			11
GCR-0012	HBG1	AGAGACAAGAAGGUAAAAAA	chr11:5250203-5250223	-			12
GCR-0013	HBG1	AAAUUAGCAGUAUCCUCUUG	chr11:5250064-5250084	-			13
GCR-0014	HBG1	GUCCUGGUAUCCUCUAUGAU	chr11:5250168-5250188	-			14
GCR-0015	HBG1	GUAUCCUCUAUGAUGGGAGA	chr11:5250162-5250182	-			15
HBG2 프로모터 영역 내에서만 표적 서열을 갖는 gRNA 표적화 도메인							
GCR-0017	HBG2	AUUAAGCAGCAGUAUCCUCU	chr11:5254990-5255010	-			17
GCR-0022	HBG2	AGAAUAAAUUAGAGAAAAAU	chr11:5255051-5255071	-			22
GCR-0029	HBG2	AGAAGUCCUGGUAUCUUCUA	chr11:5255100-5255120	-			29
GCR-	HBG2	UUAAGCAGCA	chr11:5254989-	-			32

[0537]

[0538]

0032		GUAUCCUCUU	5255009				
GCR-0034	HBG2	AAAAAUUGGAAUGACUGAAU	chr11:5255037-5255057	-			34
GCR-0046	HBG2	GGGAGAAGAAAACUAGCUAA	chr11:5255076-5255096	-			46
GCR-0051	HBG2	GGAGAAGAAAACUAGCUAAA	chr11:5255075-5255095	-			51
GCR-0052	HBG2	CUCCCACCAUAGAAGAUACC	chr11:5255091-5255111	+			52
GCR-0054	HBG2	AGUCCUGGUAUCUUCUAUGG	chr11:5255097-5255117	-			54
GCR-0058	HBG2	GUCCUGGUAUCUUCUAUGGU	chr11:5255096-5255116	-			58
GCR-0060	HBG2	UAAGCAGCAGUAUCCUCUUG	chr11:5254988-5255008	-			60
GCR-0069	HBG2	AAGCAGCAGUAUCCUCUUGG	chr11:5254987-5255007	-			69
HBG1 및 HBG2 프로모터 영역 내에 표적화 도메인을 갖는 gRNA							
GCR-0016	HBG1/HBG2	CCUAGCCAGCCGCCGGCCCC	chr11:5249895-5249915	+	chr11:5254819-5254839	+	16
GCR-0018	HBG1/HBG2	UAUCCAGUGAGGCCAGGGGC	chr11:5249910-5249930	-	chr11:5254834-5254854	-	18
GCR-0019	HBG1/HBG2	CAUUGAGAUAGUGUGGGGAA	chr11:5250036-5250056	+	chr11:5254960-5254980	+	19
GCR-0020	HBG1/HBG2	CCAGUGAGGCCAGGGGCCGG	chr11:5249907-5249927	-	chr11:5254831-5254851	-	20
GCR-0021	HBG1/HBG2	GUGGGGAAGGGGCCCCCAAG	chr11:5250048-5250068	+	chr11:5254972-5254992	+	21
GCR-0023	HBG1/HBG2	CCAGGGGCCGGCGGCUGGCU	chr11:5249898-5249918	-	chr11:5254822-5254842	-	23
GCR-0024	HBG1/HBG2	UGAGGCCAGGGGCCGGCGGC	chr11:5249903-5249923	-	chr11:5254827-5254847	-	24
GCR-0025	HBG1/HBG2	CAGUUCCACACACUCGCUUC	chr11:5249846-5249866	-	chr11:5254770-5254790	-	25
GCR-0026	HBG1/HBG2	CCGCCGGCCCCUGGCCUCAC	chr11:5249904-5249924	+	chr11:5254828-5254848	+	26

GCR-0027	HBG1/HBG2	GUUUGCCUUGUCAAGGCUAU	chr11:5249949-5249969	+	chr11:5254873-5254893	+	27
GCR-0028	HBG1/HBG2	GGCUAGGGAUGAAGAAUAAA	chr11:5249882-5249902	-	chr11:5254806-5254826	-	28
GCR-0030	HBG1/HBG2	CAGGGGCCGGCGGCUGGCUA	chr11:5249897-5249917	-	chr11:5254821-5254841	-	30
GCR-0031	HBG1/HBG2	ACUGGAUACUCUAAGACUAU	chr11:5249922-5249942	+	chr11:5254846-5254866	+	31
GCR-0033	HBG1/HBG2	CCCUGGCUAAACUCCACCCA	chr11:5249995-5250015	-	chr11:5254919-5254939	-	33
GCR-0035	HBG1/HBG2	UUAGAGUAUCCAGUGAGGCC	chr11:5249916-5249936	-	chr11:5254840-5254860	-	35
GCR-0036	HBG1/HBG2	CCCAUGGGUGGAGUUUAGCC	chr11:5249991-5250011	+	chr11:5254915-5254935	+	36
GCR-0037	HBG1/HBG2	AGGCAAGGCUGGCCAACCCA	chr11:5249975-5249995	+	chr11:5254899-5254919	+	37
GCR-0038	HBG1/HBG2	UAGAGUAUCCAGUGAGGCCA	chr11:5249915-5249935	-	chr11:5254839-5254859	-	38
GCR-0039	HBG1/HBG2	UAUCUGUCUGAAACGGUCCC	chr11:5250012-5250032	-	chr11:5254936-5254956	-	39
GCR-0040	HBG1/HBG2	AUUGAGAUAGUGUGGGGAAG	chr11:5250037-5250057	+	chr11:5254961-5254981	+	40
GCR-0041	HBG1/HBG2	CUUCAUCCCUAGCCAGCCGC	chr11:5249888-5249908	+	chr11:5254812-5254832	+	41
GCR-0042	HBG1/HBG2	GCUAUUGGUCAAGGCAAGGC	chr11:5249964-5249984	+	chr11:5254888-5254908	+	42
GCR-0043	HBG1/HBG2	AUGCAAAUAUCUGUCUGAAA	chr11:5250019-5250039	-	chr11:5254943-5254963	-	43
GCR-0044	HBG1/HBG2	GCAUUGAGAUAGUGUGGGGA	chr11:5250035-5250055	+	chr11:5254959-5254979	+	44
GCR-0045	HBG1/HBG2	UGGUCAAGUUUGCCUUGUCA	chr11:5249942-5249962	+	chr11:5254866-5254886	+	45
GCR-0047	HBG1/HBG2	GGCAAGGCUGGCCAACCCAU	chr11:5249976-5249996	+	chr11:5254900-5254920	+	47
GCR-0048	HBG1/HBG2	ACGGCUGACAAAAGAAGUCC	chr11:5250184-5250204	-	chr11:5255112-5255132	-	48

[0539]

GCR-0049	HBG1/HBG2	CGAGUGUGUGGAACUGCUGA	chr11:5249850-5249870	+	chr11:5254774-5254794	+	49
GCR-0050	HBG1/HBG2	CCUGGCUAAACUCCACCCAU	chr11:5249994-5250014	-	chr11:5254918-5254938	-	50
GCR-0053	HBG1/HBG2	CUUGUCAAGGCUAUUGGUCA	chr11:5249955-5249975	+	chr11:5254879-5254899	+	53
GCR-0055	HBG1/HBG2	AUAUUUGCAUUGAGAUAGUG	chr11:5250029-5250049	+	chr11:5254953-5254973	+	55
GCR-0056	HBG1/HBG2	GCUAAACUCCACCCAUGGGU	chr11:5249990-5250010	-	chr11:5254914-5254934	-	56
GCR-0057	HBG1/HBG2	ACGUUCCAGAAGCGAGUGUG	chr11:5249838-5249858	+	chr11:5254762-5254782	+	57
GCR-0059	HBG1/HBG2	UAUUUGCAUUGAGAUAGUGU	chr11:5250030-5250050	+	chr11:5254954-5254974	+	59
GCR-0061	HBG1/HBG2	GGAAUGACUGAAUCGGAACA	chr11:5250102-5250122	-	chr11:5255030-5255050	-	61
GCR-0062	HBG1/HBG2	CUUGACCAAUAGCCUUGACA	chr11:5249957-5249977	-	chr11:5254881-5254901	-	62
GCR-0063	HBG1/HBG2	CAAGGCUAUUGGUCAAGGCA	chr11:5249960-5249980	+	chr11:5254884-5254904	+	63
GCR-0064	HBG1/HBG2	AAGGCUGGCCAACCCAUGGG	chr11:5249979-5249999	+	chr11:5254903-5254923	+	64
GCR-0065	HBG1/HBG2	ACUCGCUUCUGGAACGUCUG	chr11:5249835-5249855	-	chr11:5254759-5254779	-	65
GCR-0066	HBG1/HBG2	AUUUGCAUUGAGAUAGUGUG	chr11:5250031-5250051	+	chr11:5254955-5254975	+	66
GCR-0067	HBG1/HBG2	ACUGAAUCGGAACAAGGCAA	chr11:5250096-5250116	-	chr11:5255024-5255044	-	67
GCR-0068	HBG1/HBG2	CCAUGGGUGGAGUUUAGCCA	chr11:5249992-5250012	+	chr11:5254916-5254936	+	68
GCR-0070	HBG1/HBG2	AGAGUAUCCAGUGAGGCCAG	chr11:5249914-5249934	-	chr11:5254838-5254858	-	70
GCR-0071	HBG1/HBG2	GAGUGUGUGGAACUGCUGAA	chr11:5249851-5249871	+	chr11:5254775-5254795	+	71
GCR-0072	HBG1/HBG2	UAGUCUUAGAGUAUCCAGUG	chr11:5249921-5249941	-	chr11:5254845-5254865	-	72

[0540]

[0541] 하기 표 2는, gRNA 분자에 포함되었을 때, 실시예에 기재된 방법에 따라 7일째에 태아 헤모글로빈에서(예를 들어 변형된 HSPC로부터 분화된 적혈구계 세포에서) 적어도 17% 증가를 초래하는 표적화 도메인을 보여준다. 임의의 이들 표적화 도메인을 포함하는 gRNA 분자는 본원에서 수합해서 Tier 2 gRNA 분자로 지칭된다.

[0542] [표 2]

[0543] Tier 2 gRNA 분자에 대한 표적화 도메인

표적화 도메인 ID
GCR-0001
GCR-0006
GCR-0008
GCR-0009
GCR-0010
GCR-0011
GCR-0012
GCR-0028
GCR-0034
GCR-0045
GCR-0046
GCR-0047
GCR-0048
GCR-0050
GCR-0051
GCR-0053
GCR-0054
GCR-0058
GCR-0062
GCR-0063
GCR-0067

[0544]

[0545] 하기 표 3a 및 표 3b는, gRNA 분자에 포함되었을 때, 실시예에 기재된 방법에 따라 7일째에 태아 헤모글로빈에서(예를 들어 변형된 HSPC로부터 분화된 적혈구계 세포에서) 최고 증가를 초래하는 표적화 도메인을 보여준다. 임의의 이들 표적화 도메인을 포함하는 gRNA 분자는 본원에서 수합해서 Tier 1(예를 들어 Tirr 1a 또는 Tier 1b) gRNA 분자로 지칭된다.

[0546] [표 3a]

[0547] Tier 1a gRNA 분자에 대한 표적화 도메인

표적화 도메인 ID
GCR-0006
GCR-0008
GCR-0028
GCR-0034
GCR-0048
GCR-0067

[0548]

[0549] [표 3b]

[0550] Tier 1b gRNA 분자에 대한 표적화 도메인

표적화 도메인 ID
GCR-0001
GCR-0008
GCR-0009
GCR-0010
GCR-0012
GCR-0054

[0551]

[0552] **III. gRNA 설계 방법**

[0553] 표적 서열을 선택, 설계 및 검증하기 위한 방법을 포함하여, gRNA를 설계하기 위한 방법이 본원에 기재된다. 예시적인 표적화 도메인이 또한 본원에 제공된다. 본원에서 논의된 표적화 도메인은 본원에 기재된 gRNA에 포함될 수 있다.

[0554] 표적-외 분석뿐만 아니라 표적 서열의 선택 및 검증 방법은 예를 들어 문헌[Mali el al., 2013 SCIENCE 339(6121): 823-826]; 문헌[Hsu et al, 2013 NAT BIOTECHNOL, 31(9): 827-32]; 문헌[Fu et al, 2014 NAT BIOTECHNOL, doi: 10.1038/nbt.2808. PubMed PM ID: 24463574]; 문헌[Heigwer et al, 2014 NAT METHODS 11(2): 122-3. doi: 10.1038/nmeth.2812. PubMed PMID: 24481216]; 문헌[Bae el al, 2014 BIOINFORMATICS PubMed PMID: 24463181]; 문헌[Xiao A el al, 2014 BIOINFORMATICS PubMed PMID: 24389662]에 기재되어 있다.

[0555] 예를 들어 소프트웨어 툴은 사용자의 표적 서열 내에서 gRNA의 선택을 최적화하기 위하여, 예를 들어 게놈 전체에 걸쳐 총 표적-외 활성을 최소화하기 위하여 사용될 수 있다. 표적-외 활성은 절단 외의 것일 수 있다. 예를 들어 S. 피오게네스 Cas9를 사용하는, 각각의 가능한 gRNA 선택을 위하여, 툴은 일정 개수(예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)까지의 미스매치된 염기쌍을 함유하는 게놈 전체에 걸친 모든 표적-외 서열(예를 들어 NAG 또는 NGG PAM 이전)을 확인할 수 있다. 각 표적-외 서열에서의 절단 효율은, 예를 들어 실험적으로 유도된 가중치 구조를 사용하여 예측될 수 있다. 그 후 각각의 가능한 gRNA는 총 예상 표적-외 절단에 따라 순위가 매겨진다; 최상위 순위의 gRNA는 가장 많은 표적-적중(on-target) 절단 및 가장 적은 표적-외(off-target) 절단을 가질 가능성이 있는 것에 해당한다. 다른 기능, 예를 들어 CRISPR 구축을 위한 자동화된 시약 설계, 적중(on-target) Surveyor 검정을 위한 프라이머 설계, 및 차세대 시퀀싱을 통한 표적-외 절단의 대용량 고효율(high-throughput) 검출 및 정량화를 위한 프라이머 설계가 툴에 또한 포함될 수 있다. 후보 gRNA 분자는 당분야에 공지된 방법 또는 본원에 기재된 바에 의해 평가될 수 있다.

[0556] 잠재적인 gRNA 분자의 초기 목록을 생성하기 위해 소프트웨어 알고리즘이 사용될 수 있음에도 불구하고, 절단 효율 및 특이성이 반드시 예측된 값을 반영하지 않을 것이며, gRNA 분자는 전형적으로, 예를 들어 절단 효율, 삽입결실 형성, 절단 특이성 및 원하는 표현형의 변화를 결정하기 위해, 특정 세포주, 예를 들어 일차 인간 세포주, 예를 들어 인간 HSPC, 예를 들어 인간 CD34+ 세포에서의 스크리닝을 필요로 한다. 이들 특성은 본원에 기재된 방법에 의해 검정될 수 있다.

[0557] 본 발명의 양태에서, GCR-0001(SEQ ID NO: 1, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0001 #1:

AGUCCUGGUAUCCUCUAUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 74)

sgRNA GCR-0001 #2:

mA*mG*mU*CCUGGUAUCCUCUAUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU (SEQ
ID NO: 75)

sgRNA GCR-0001 #3:

mA*mG*mU*CCUGGUAUCCUCUAUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U (SEQ
ID NO: 76)

dgRNA GCR-0001 #1:

[0558] crRNA: AGUCCUGGUAUCCUCUAUGAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 77)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0001 #2:

crRNA: mA*mG*mU*CCUGGUAUCCUCUAUGAGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 78)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0001 #3:

crRNA: mA*mG*mU*CCUGGUAUCCUCUAUGAGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 78)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0559] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0560] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0561] 본 발명의 양태에서, GCR-0006(SEQ ID NO: 6, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0006 #1:

AAAAACUGGAAUGACUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 79)

sgRNA GCR-0006 #2:

mA*mA*mA*AACUGGAAUGACUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
[0562] (SEQ ID NO: 80)

sgRNA GCR-0006 #3:

mA*mA*mA*AACUGGAAUGACUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 81)

dgRNA GCR-0006 #1:

crRNA: AAAAACUGGAAUGACUGAAUGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 82)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0006 #2:

crRNA:
mA*mA*mA*AACUGGAAUGACUGAAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 83)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0006 #3:

crRNA:
mA*mA*mA*AACUGGAAUGACUGAAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 83)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0563] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0564] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0565] 본 발명의 양태에서, GCR-0008(SEQ ID NO: 8, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0008 #1:

GGAGAAGGAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 84)

sgRNA GCR-0008 #2:

mG*mG*mA*GAAGGAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 85)

sgRNA GCR-0008 #3:

mG*mG*mA*GAAGGAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 86)

dgRNA GCR-0008 #1:

crRNA: GGAGAAGGAAACUAGCUAAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 87)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0008 #2:

crRNA:
mG*mG*mA*GAAGGAAACUAGCUAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 88)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0008 #3:

crRNA:
mG*mG*mA*GAAGGAAACUAGCUAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ

[0566] ID NO: 88)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC

[0567] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0568] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0569] 본 발명의 양태에서, GCR-0009(SEQ ID NO: 9, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0009 #1:

GUUUCCUUCUCCCAUCAUAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 89)

sgRNA GCR-0009 #2:

mG*mU*mU*UCCUUCUCCCAUCAUAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU (SEQ
ID NO: 90)

sgRNA GCR-0009 #3:

mG*mU*mU*UCCUUCUCCCAUCAUAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U (SEQ
ID NO: 91)

dgRNA GCR-0009 #1:

crRNA: GUUUCCUUCUCCCAUCAUAGGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 92)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

[0570] dgRNA GCR-0009 #2:

crRNA: mG*mU*mU*UCCUUCUCCCAUCAUAGGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 93)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0009 #3:

crRNA: mG*mU*mU*UCCUUCUCCCAUCAUAGGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 93)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0571] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0572] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0573] 본 발명의 양태에서, GCR-0010(SEQ ID NO: 10, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0010 #1:

GGGAGAAGGAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 94)

sgRNA GCR-0010 #2:

mG*mG*mG*AGAAGGAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 95)

sgRNA GCR-0010 #3:

mG*mG*mG*AGAAGGAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
[0574] (SEQ ID NO: 96)

dgRNA GCR-0010 #1:

crRNA: GGGAGAAGGAAACUAGCUAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 97)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0010 #2:

crRNA:
mG*mG*mG*AGAAGGAAACUAGCUAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 98)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0010 #3:

crRNA:
mG*mG*mG*AGAAGGAAACUAGCUAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 98)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0575] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0576] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0577] 본 발명의 양태에서, GCR-0011(SEQ ID NO: 11, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0011 #1:

CACUGGAGCUAGAGACAAGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
[0578] GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 99)

sgRNA GCR-0011 #2:

mC*mA*mC*UGGAGCUAGAGACAAGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 100)

sgRNA GCR-0011 #3:

mC*mA*mC*UGGAGCUAGAGACAAGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 101)

dgRNA GCR-0011 #1:

crRNA: CACUGGAGCUAGAGACAAGAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
102)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0011 #2:

crRNA:
mC*mA*mC*UGGAGCUAGAGACAAGAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 103)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUC
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0011 #3:

crRNA:
mC*mA*mC*UGGAGCUAGAGACAAGAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 103)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0579] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0580] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0581] 본 발명의 양태에서, GCR-0012(SEQ ID NO: 12, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0012 #1:

AGAGACAAGAAGGUAAAAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 104)

sgRNA GCR-0012 #2:

mA*mG*mA*GACAAGAAGGUAAAAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 105)

sgRNA GCR-0012 #3:

mA*mG*mA*GACAAGAAGGUAAAAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 106)

dgRNA GCR-0012 #1:

crRNA: AGAGACAAGAAGGUAAAAAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
107)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0012 #2:

crRNA:
mA*mG*mA*GACAAGAAGGUAAAAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 108)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
[0582] GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0012 #3:

crRNA:
mA*mG*mA*GACAAGAAGGUAAAAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 108)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0583] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0584] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0585] 본 발명의 양태에서, GCR-0028(SEQ ID NO: 28, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0028 #1:

GGCUAGGGAUGAAGAAUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 109)

sgRNA GCR-0028 #2:

mG*mG*mC*UAGGGAUGAAGAAUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 110)

sgRNA GCR-0028 #3:

mG*mG*mC*UAGGGAUGAAGAAUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 111)

dgRNA GCR-0028 #1:

crRNA: GGCUAGGGAUGAAGAAUAAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
[0586] 112)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0028 #2:

crRNA:
mG*mG*mC*UAGGGAUGAAGAAUAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 113)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0028 #3:

crRNA:
mG*mG*mC*UAGGGAUGAAGAAUAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 113)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0587] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0588] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0589] 본 발명의 양태에서, GCR-0034(SEQ ID NO: 34, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0034 #1:

AAAAAUUGGAAUGACUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 114)

[0590] sgRNA GCR-0034 #2:

mA*mA*mA*AAUUGGAAUGACUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 115)

sgRNA GCR-0034 #3:

mA*mA*mA*AAUUGGAAUGACUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 116)

dgRNA GCR-0034 #1:

crRNA: AAAAAUUGGAAUGACUGAAUGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 117)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0034 #2:

crRNA:
mA*mA*mA*AAUUGGAAUGACUGAAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 118)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0034 #3:

crRNA:
mA*mA*mA*AAUUGGAAUGACUGAAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 118)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0591]

[0592] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0593] 본 발명의 양태에서, GCR-0045(SEQ ID NO: 45, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0045 #1:

UGGUCAAGUUUGCCUUGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 119)

sgRNA GCR-0045 #2:

mU*mG*mG*UCAAGUUUGCCUUGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 120)

sgRNA GCR-0045 #3:

mU*mG*mG*UCAAGUUUGCCUUGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 121)

dgRNA GCR-0045 #1:

crRNA: UGGUCAAGUUUGCCUUGUCAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
122)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0045 #2:

crRNA:
mU*mG*mG*UCAAGUUUGCCUUGUCAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 123)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

[0594] dgRNA GCR-0045 #3:

crRNA:
mU*mG*mG*UCAAGUUUGCCUUGUCAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 123)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0595] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0596] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0597] 본 발명의 양태에서, GCR-0046(SEQ ID NO: 46, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0046 #1:

GGGAGAAGAAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 124)

sgRNA GCR-0046 #2:

mG*mG*mG*AGAAGAAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 125)

sgRNA GCR-0046 #3:

mG*mG*mG*AGAAGAAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 126)

dgRNA GCR-0046 #1:

crRNA: GGGAGAAGAAAACUAGCUAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:

[0598] 127)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0046 #2:

crRNA:
mG*mG*mG*AGAAGAAAACUAGCUAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 128)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0046 #3:

crRNA:
mG*mG*mG*AGAAGAAAACUAGCUAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 128)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC

[0599] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0600] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0601] 본 발명의 양태에서, GCR-0047(SEQ ID NO: 47, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0047 #1:

GGCAAGGCUGGCCAACCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 129)

[0602] sgRNA GCR-0047 #2:

mG*mG*mC*AAGGCUGGCCAACCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU (SEQ ID NO: 130)

sgRNA GCR-0047 #3:

mG*mG*mC*AAGGCUGGCCAACCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U (SEQ ID NO: 131)

dgRNA GCR-0047 #1:

crRNA: GGCAAGGCUGGCCAACCCAUGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 132)

tracr: AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0047 #2:

crRNA: mG*mG*mC*AAGGCUGGCCAACCCAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 133)

tracr: mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0047 #3:

crRNA: mG*mG*mC*AAGGCUGGCCAACCCAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 133)

tracr: AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0603]

[0604] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0605] 본 발명의 양태에서, GCR-0048(SEQ ID NO: 48, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0048 #1:

ACGGCUGACAAAAGAAGUCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 134)

sgRNA GCR-0048 #2:

mA*mC*mG*GCUGACAAAAGAAGUCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU (SEQ
ID NO: 135)

sgRNA GCR-0048 #3:

mA*mC*mG*GCUGACAAAAGAAGUCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U (SEQ
ID NO: 136)

dgRNA GCR-0048 #1:

crRNA: ACGGCUGACAAAAGAAGUCCGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
137)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0048 #2:

crRNA: mA*mC*mG*GCUGACAAAAGAAGUCCGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 138)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0048 #3:

crRNA: mA*mC*mG*GCUGACAAAAGAAGUCCGUUUUAGAGCUAUGCUGUU*mU*mU*mG
[0606] (SEQ ID NO: 138)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0607] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0608] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0609] 본 발명의 양태에서, GCR-0050(SEQ ID NO: 50, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0050 #1:

CCUGGCUAAACUCCACCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 139)

sgRNA GCR-0050 #2:

mC*mC*mU*GGCUAAACUCCACCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU (SEQ
ID NO: 140)

sgRNA GCR-0050 #3:

mC*mC*mU*GGCUAAACUCCACCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U (SEQ
ID NO: 141)

dgRNA GCR-0050 #1:

crRNA: CCUGGCUAAACUCCACCCAUGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 142)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

[0610] dgRNA GCR-0050 #2:

crRNA: mC*mC*mU*GGCUAAACUCCACCCAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 143)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0050 #3:

crRNA: mC*mC*mU*GGCUAAACUCCACCCAUGUUUUAGAGCUAUGCUGUU*mU*mU*mG
(SEQ ID NO: 143)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0611] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0612] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0613] 본 발명의 양태에서, GCR-0051(SEQ ID NO: 51, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0051 #1:

GGAGAAGAAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 144)

sgRNA GCR-0051 #2:

mG*mG*mA*GAAGAAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 145)

sgRNA GCR-0051 #3:

mG*mG*mA*GAAGAAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
[0614] (SEQ ID NO: 146)

dgRNA GCR-0051 #1:

crRNA: GGAGAAGAAAACUAGCUAAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 147)

tracr: AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0051 #2:

crRNA: mG*mG*mA*GAAGAAAACUAGCUAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 148)

tracr: mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0051 #3:

crRNA: mG*mG*mA*GAAGAAAACUAGCUAAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 148)

tracr: AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0615]

[0616] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0617] 본 발명의 양태에서, GCR-0053(SEQ ID NO: 53, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0053 #1:

CUUGUCAAGGCUAUUGGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 149)

sgRNA GCR-0053 #2:

mC*mU*mU*GUCAAGGCUAUUGGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 150)

sgRNA GCR-0053 #3:

mC*mU*mU*GUCAAGGCUAUUGGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 151)

dgRNA GCR-0053 #1:

crRNA: CUUGUCAAGGCUAUUGGUCAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
152)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0053 #2:

crRNA:
mC*mU*mU*GUCAAGGCUAUUGGUCAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 153)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0053 #3:

crRNA:
mC*mU*mU*GUCAAGGCUAUUGGUCAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 153)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0618]

[0619] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0620] 본 발명의 양태에서, GCR-0054(SEQ ID NO: 54, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0054 #1:

AGUCCUGGUAUCUUCUAUGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 154)

sgRNA GCR-0054 #2:

mA*mG*mU*CCUGGUAUCUUCUAUGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 155)

sgRNA GCR-0054 #3:

mA*mG*mU*CCUGGUAUCUUCUAUGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 156)

dgRNA GCR-0054 #1:

crRNA: AGUCCUGGUAUCUUCUAUGGGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
157)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0054 #2:

crRNA:
mA*mG*mU*CCUGGUAUCUUCUAUGGGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
[0621] ID NO: 158)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0054 #3:

crRNA:
mA*mG*mU*CCUGGUAUCUUCUAUGGGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 158)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0622] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0623] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0624] 본 발명의 양태에서, GCR-0058(SEQ ID NO: 58, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0058 #1:

GUCCUGGUAUCUUCUAUGGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 159)

sgRNA GCR-0058 #2:

mG*mU*mC*CUGGUAUCUUCUAUGGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 160)

sgRNA GCR-0058 #3:

mG*mU*mC*CUGGUAUCUUCUAUGGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 161)

[0625] dgRNA GCR-0058 #1:

crRNA: GUCCUGGUAUCUUCUAUGGUGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 162)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0058 #2:

crRNA:
mG*mU*mC*CUGGUAUCUUCUAUGGUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 163)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0058 #3:

crRNA:
mG*mU*mC*CUGGUAUCUUCUAUGGUGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 163)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0626] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0627] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0628] 본 발명의 양태에서, GCR-0062(SEQ ID NO: 62, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0062 #1:

CUUGACCAAUAGCCUUGACAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
[0629] GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 164)

sgRNA GCR-0062 #2:

mC*mU*mU*GACCAAUAGCCUUGACAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU (SEQ ID NO: 165)

sgRNA GCR-0062 #3:

mC*mU*mU*GACCAAUAGCCUUGACAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U (SEQ ID NO: 166)

dgRNA GCR-0062 #1:

crRNA: CUUGACCAAUAGCCUUGACAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 167)

tracr: AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0062 #2:

crRNA: mC*mU*mU*GACCAAUAGCCUUGACAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 168)

tracr: mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0062 #3:

crRNA: mC*mU*mU*GACCAAUAGCCUUGACAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ ID NO: 168)

tracr: AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUUU (SEQ ID NO: 224).

[0630]

[0631] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0632] 본 발명의 양태에서, GCR-0063(SEQ ID NO: 63, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0063 #1:

CAAGGCUAUUGGUCAAGGCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 169)

sgRNA GCR-0063 #2:

mC*mA*mA*GGCUAUUGGUCAAGGCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 170)

sgRNA GCR-0063 #3:

mC*mA*mA*GGCUAUUGGUCAAGGCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 171)

dgRNA GCR-0063 #1:

crRNA: CAAGGCUAUUGGUCAAGGCAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
172)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0063 #2:

crRNA:
mC*mA*mA*GGCUAUUGGUCAAGGCAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 173)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

[0633] dgRNA GCR-0063 #3:

crRNA:
mC*mA*mA*GGCUAUUGGUCAAGGCAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 173)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0634] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0635] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0636] 본 발명의 양태에서, GCR-0067(SEQ ID NO: 67, 하기에서 밑줄이 그어진 비변형된 서열)의 표적화 도메인을 포함하는 gRNA, 예를 들어 하기 기재된 gRNA 분자 중 하나는 예를 들어 본원에 기재된 1개 초과의 gRNA 분자를 수반하는 양태를 포함하여 본 발명의 CRISPR 시스템, 방법, 세포 및 다른 양태와 실시형태에서, 유용하다:

sgRNA GCR-0067 #1:

ACUGAAUCGGAACAAGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 174)

sgRNA GCR-0067 #2:

mA*mC*mU*GAAUCGGAACAAGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
(SEQ ID NO: 175)

sgRNA GCR-0067 #3:

mA*mC*mU*GAAUCGGAACAAGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUA
AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
(SEQ ID NO: 176)

dgRNA GCR-0067 #1:

crRNA: ACUGAAUCGGAACAAGGCAAGUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:
[0637] 177)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGCUUUUUUU (SEQ ID NO: 224)

dgRNA GCR-0067 #2:

crRNA:
mA*mC*mU*GAAUCGGAACAAGGCAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 178)

tracr:
mA*mA*mC*AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGAGUCGGUGCUUUU*mU*mU*mU (SEQ ID NO: 73)

dgRNA GCR-0067 #3:

crRNA:
mA*mC*mU*GAAUCGGAACAAGGCAAGUUUUAGAGCUAUGCUGUU*mU*mU*mG (SEQ
ID NO: 178)

tracr:
AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
[0638] GAGUCGGUGCUUUUUUU (SEQ ID NO: 224).

[0639] 상기 기재된 각각의 gRNA 분자에서, "*"는 인접한 뉴클레오타이드 사이의 포스포로티오에이트 결합을 나타내고, "mN"(여기서 N은 A, G, C 또는 U임)은 2'-OMe 변형된 뉴클레오타이드를 나타낸다. 실시형태에서, 임의의 본원에 기재된, 예를 들어 상기 기재된 gRNA 분자는 예를 들어 본원에 기재된 바와 같은 Cas9 분자와 복합체화하여, 리보핵 단백질 복합체(RNP)를 형성한다. 이러한 RNP는 특히, 본 발명의, 예를 들어 본원에 기재된 방법, 세포 및 다른 양태와 실시형태에서, 유용하다.

[0640] **IV. Cas 분자**

[0641] **Cas9 분자**

[0642] 바람직한 실시형태에서, Cas 분자는 Cas9 분자이다. 다양한 종의 Cas9 분자가 본원에 기재된 방법 및 조성물에 사용될 수 있다. S. 피오게네스 Cas9 분자가 본원에 개시의 많은 것의 대상체이지만, 본원에 나열된 다른 종의 Cas9 단백질의, 이로부터 유래된, 또는 이에 기반한 Cas9 분자가 사용될 수도 있다. 즉, 다른 Cas9 분자, 예를 들어 S. 써모필루스(S. thermophilus), 스타필로코커스 아우레우스(Staphylococcus aureus) 및/또는 네이세리아 메닌기티디스(Neisseria meningitides) Cas9 분자가 본원에 기재된 시스템, 방법 및 조성물에 사용될 수 있다. 추가적인 Cas9 종에는: 에시도보락스 아베나에(Acidovorax avenae), 악티노바실러스 플레우로뉴모니아에(Actinobacillus pleuropneumoniae), 악티노바실러스 숙시노게네스(Actinobacillus succinogenes), 악티노바실러스 수이스(Actinobacillus suis), 악티노마이세스 종(Actinomyces sp.) 사이클리필루스 데니트리피칸스(cycliphilus denitrificans), 아미노모나스 파우시보란스(Aminomonas paucivorans), 바실러스 세레우스

(Bacillus cereus), 바실러스 스미티(Bacillus smithii), 바실러스 투린기엔시스(Bacillus thuringiensis), 박테로이데스 종(Bacteroides sp.), 블라스토피렐룰라 마리나(Blastopirellula marina), 브라디리조비움 종(Bradyrhizobium sp.), 브레비바실러스 라템스포루스(Brevibacillus latemsporus), 캄필로박터 콜리(Campylobacter coli), 캄필로박터 제주니(Campylobacter jejuni), 캄필로박터 라드(Campylobacter lad), 칸디다투스 푸니세이스피릴룸(Candidatus Puniceispirillum), 클로스트리디움 셀룰롤리티쿰(Clostridium cellulolyticum), 클로스트리디움 페르프린겐스(Clostridium perfringens), 코리네박테리움 악콜렌스(Corynebacterium accolens), 코리네박테리움 디프테리아(Corynebacterium diphtheria), 코리네박테리움 마트루초티(Corynebacterium matruchotii), 디노로세오박터 슬리바에(Dinoroseobacter sliibae), 유박테리움 돌리춤(Eubacterium dolichum), 감마 프로테오박테리움(Gamma proteobacterium), 글루코나세토박터 디아조트로피쿠스(Gluconacetobacter diazotrophicus), 해모필루스 파라인플루엔자에(Haemophilus parainfluenzae), 해모필루스 스푸토룸(Haemophilus sputorum), 헬리코박터 카나덴시스(Helicobacter canadensis), 헬리코박터 시나에디(Helicobacter cinaedi), 헬리코박터 무스텔라에(Helicobacter mustelae), 일리오박터 폴리트로푸스(Ilyobacter polytropus), 킨겔라 킨가에(Kingella kingae), 락토바실러스 크리스파투스(Lactobacillus crispatus), 리스테리아 이바노비(Listeria ivanovii), 리스테리아 모노사이토게네스(Listeria monocytogenes), 리스테리아세아에 박테리움(Listeriaceae bacterium), 메틸로시스티스 종(Methylocystis sp.), 메틸로시누스 트리초스포리움(Methylosinus trichosporium), 모빌룬쿠스 물리에리스(Mobiluncus mulieris), 네이세리아 바실리포르미스(Neisseria bacilliformis), 네이세리아 시네레아(Neisseria cinerea), 네이세리아 플라베센스(Neisseria flavescens), 네이세리아 락타미카, 네이세리아 종(Neisseria sp.), 네이세리아 와드스워르티(Neisseria wadsworthii), 니트로소모나스 종(Nitrosomonas sp.), 파르비바쿨룸 라바멘티보란스(Parvibaculum lavamentivorans), 파스테우렐라 물토시다(Pasteurella multocida), 파스콜락토박테리움 숙시나투텐스(Phascolarctobacterium succinatutens), 랄스토니아 시지기(Ralstonia syzygii), 로도슈도모나스 팔루스트리스(Rhodopseudomonas palustris), 로도불룸 종(Rhodovulum sp.), 시몬시엘라 무엘레리(Simonsiella muelleri), 스핀고모나스 종(Sphingomonas sp.), 스포로락토바실러스 비네아에(Sporolactobacillus vineae), 스타필로코커스 루그두넨시스(Staphylococcus lugdunensis), 스트렙토코커스 종(Streptococcus sp), 서브돌리그라눌룸 종(Subdoligranulum sp.), 티스트렐라 모빌리스(Tistrella mobilis), 트레포네마 종(Treponema sp.) 또는 베르미네프로박터 에이세니아에(Verminephrobacter eiseniae)가 포함된다.

[0643] 그 용어가 본원에 사용되는 바와 같이, Cas9 분자는 gRNA 분자(예를 들어 tracr 도메인의 서열)와 상호 작용할 수 있으며, gRNA 분자와 협력하여, 표적 서열 및 PAM 서열을 포함하는 부위로 배치(예를 들어 표적화 또는 유도화) 할 수 있는 분자를 지칭한다.

[0644] 실시형태에서, Cas9 분자는 표적 핵산 분자를 절단할 수 있으며, 본원에서 활성 Cas9 분자로 지칭될 수 있다. 실시형태에서, 활성 Cas9 분자는 다음 활성 중 하나 이상을 포함한다: 니카제 활성, 즉 핵산 분자의 단일 가닥, 예를 들어 비상보성 가닥 또는 상보성 가닥을 절단하는 능력; 이중 가닥 뉴클레아제 활성, 즉 이중 가닥 핵산의 두 가닥을 모두 절단하고 이중 가닥 절단부을 생성하는 능력, 실시형태에서, 이는 2개의 니카제 활성의 존재임; 엔도뉴클레아제 활성; 엑소 뉴클레아제 활성; 및 헬리카제 활성, 즉 이중 가닥 핵산의 나선형 구조를 푸는 능력.

[0645] 실시형태에서, 효소적으로 활성인 Cas9분자는 DNA 가닥을 둘 모두 절단하며 이중 가닥 절단부을 야기한다. 실시형태에서, Cas9 분자는 오직 하나의 가닥, 예를 들어 gRNA가 혼성화하는 가닥, 또는 gRNA와 혼성화하는 가닥에 상보적인 가닥을 절단한다. 실시형태에서, 활성 Cas9 분자는 HNH-유사 도메인과 관련된 절단 활성을 포함한다. 실시형태에서, 활성 Cas9 분자는 N-말단의 RuvC-유사 도메인과 관련된 절단 활성을 포함한다. 실시형태에서, 활성 Cas9 분자는 HNH-유사 도메인과 관련된 절단 활성 및 N-말단의 RuvC-유사 도메인과 관련된 절단 활성을 포함한다. 실시형태에서, 활성 Cas9 분자는 활성의, 또는 절단적격의, HNH-유사 도메인 및 비활성의, 또는 절단 부적격의, N-말단의 RuvC-유사 도메인을 포함한다. 실시형태에서, 활성 Cas9 분자는 비활성의, 또는 절단 부적격의, HNH-유사 도메인 및 활성의, 또는 절단적격의, N-말단의 RuvC-유사 도메인을 포함한다.

[0646] 실시형태에서, 표적 핵산과 상호작용하고 표적 핵산을 절단하는 활성 Cas9 분자의 능력은 PAM 서열 의존적이다. PAM 서열은 표적 핵산 내 서열이다. 실시형태에서, 표적 핵산의 절단은 PAM 서열의 상류에서 발생한다. 상이한 박테리아 종으로부터의 활성 Cas9 분자는 상이한 서열 모티프(예를 들어 PAM 서열)를 인지할 수 있다. 실시형태에서, S.피오게네스의 활성 Cas9 분자는 서열 모티프 NGG를 인지하고, 그 서열로부터 상류의 표적 핵산 서열 1 내지 10, 예를 들어 3 내지 5 염기쌍의 절단을 지시한다. 예를 들어 문헌[Mali et al., Science 2013; 339(6121): 823-826]을 참조한다. 실시형태에서, S. 써모필루스의 활성 Cas9 분자는 서열 모티프 NGGNG 및

NNAGAAW(W= A 또는 T)를 인지하고, 이들 서열로부터 상류의 코어 표적 핵산 서열 1 내지 10, 예를 들어 3 내지 5 염기쌍의 절단을 지시한다. 예를 들어 문헌[Horvath et al., Science 2010; 327(5962):167-170], 및 문헌[Deveau et al., J Bacteriol 2008; 190(4): 1390-1400]을 참조한다. 실시형태에서, S. 뮤탄스의 활성 Cas9 분자는 서열 모티프 NGG 또는 NAAR(R- A 또는 G)을 인지하고, 이 서열로부터 상류의 코어 표적 핵산 서열 1 내지 10, 예를 들어 3 내지 5 염기쌍의 절단을 지시한다. 예를 들어 문헌[Deveau et al., J Bacteriol 2008; 190(4): 1390-1400]을 참조한다.

[0647] 실시형태에서, S. 아우레우스의 활성 Cas9 분자는 서열 모티프 NNGRR(R= A 또는 G)을 인지하고, 그 서열로부터 상류의 표적 핵산 서열 1 내지 10, 예를 들어 3 내지 5 염기쌍의 절단을 지시한다. 예를 들어 문헌[Ran F. et al., NATURE, vol. 520, 2015, pp. 186-191]을 참조한다. 실시형태에서, N. 메닌기티디스의 활성 Cas9 분자는 서열 모티프 NNNNGATT를 인지하고, 그 서열로부터 상류의 표적 핵산 서열 1 내지 10, 예를 들어 3 내지 5 염기쌍의 절단을 지시한다. 예를 들어 문헌[Hou et al., PNAS Early Edition 2013, 1-6]을 참조한다. PAM 서열을 인지하는 Cas9 분자의 능력은, 예를 들어 문헌[Jinek et al., Science 2012, 337:816]에 기재된 형질전환 분석을 이용하여 결정될 수 있다.

[0648] 일부 Cas9 분자는 gRNA 분자와 상호작용하는 능력을 갖고, gRNA 분자와 함께 코어 표적 도메인으로 유도(예를 들어 표적화 또는 배치)되지만, 표적 핵산을 절단할 수 없거나 또는 효율적인 속도로 절단할 수 없다. 절단 활성을 갖지 않거나 또는 실질적인 절단 활성을 갖지 않는 Cas9 분자는 본원에서 비활성 Cas9(효소적으로 비활성인 Cas9), 비작동 Cas9, 또는 dCas9분자로 지칭된다. 예를 들어 비활성Cas9 분자는 절단 활성이 없거나 또는 실질적으로 적은, 예를 들어 본원에 기재된 분석에 의해 측정하여, 참조 Cas9분자의 20%, 10%, 5%, 1% 또는 0.1% 미만의 절단 활성을 가질 수 있다.

[0649] 예시적인 자연적 발생 Cas9 분자는 문헌[Chylinski et al., RNA Biology 2013; 10:5, 727-737]에 기재되어 있다. 이러한 Cas9 분자는 클러스터 1 박테리아과, 클러스터 2 박테리아과, 클러스터 3 박테리아과, 클러스터 4 박테리아과, 클러스터 5 박테리아과, 클러스터 6 박테리아과, 클러스터 7 박테리아과, 클러스터 8 박테리아과, 클러스터 9 박테리아과, 클러스터 10 박테리아과, 클러스터 11 박테리아과, 클러스터 12 박테리아과, 클러스터 13 박테리아과, 클러스터 14 박테리아과, 클러스터 15 박테리아과, 클러스터 16 박테리아과, 클러스터 17 박테리아과, 클러스터 18 박테리아과, 클러스터 19 박테리아과, 클러스터 20 박테리아과, 클러스터 21 박테리아과, 클러스터 22 박테리아과, 클러스터 23 박테리아과, 클러스터 24 박테리아과, 클러스터 25 박테리아과, 클러스터 26 박테리아과, 클러스터 27 박테리아과, 클러스터 28 박테리아과, 클러스터 29 박테리아과, 클러스터 30 박테리아과, 클러스터 31 박테리아과, 클러스터 32 박테리아과, 클러스터 33 박테리아과, 클러스터 34 박테리아과, 클러스터 35 박테리아과, 클러스터 36 박테리아과, 클러스터 37 박테리아과, 클러스터 38 박테리아과, 클러스터 39 박테리아과, 클러스터 40 박테리아과, 클러스터 41 박테리아과, 클러스터 42 박테리아과, 클러스터 43 박테리아과, 클러스터 44 박테리아과, 클러스터 45 박테리아과, 클러스터 46 박테리아과, 클러스터 47 박테리아과, 클러스터 48 박테리아과, 클러스터 49 박테리아과, 클러스터 50 박테리아과, 클러스터 51 박테리아과, 클러스터 52 박테리아과, 클러스터 53 박테리아과, 클러스터 54 박테리아과, 클러스터 55 박테리아과, 클러스터 56 박테리아과, 클러스터 57 박테리아과, 클러스터 58 박테리아과, 클러스터 59 박테리아과, 클러스터 60 박테리아과, 클러스터 61 박테리아과, 클러스터 62 박테리아과, 클러스터 63 박테리아과, 클러스터 64 박테리아과, 클러스터 65 박테리아과, 클러스터 66 박테리아과, 클러스터 67 박테리아과, 클러스터 68 박테리아과, 클러스터 69 박테리아과, 클러스터 70 박테리아과, 클러스터 71 박테리아과, 클러스터 72 박테리아과, 클러스터 73 박테리아과, 클러스터 74 박테리아과, 클러스터 75 박테리아과, 클러스터 76 박테리아과, 클러스터 77 박테리아과, 또는 클러스터 78 박테리아과의 Cas9 분자를 포함한다.

[0650] 예시적인 자연적 발생 Cas9 분자는 클러스터 1 박테리아과의 Cas9 분자를 포함한다. 예로는 S. 피오게네스(예를 들어 균주 SF370, MGAS10270, MGAS10750, MGAS2096, MGAS315, MGAS5005, MGAS6180, MGAS9429, NZ131 및 SSI-1), S. 써모필루스(예를 들어 균주 LMD-9), S. 슈도포르시누스(S. pseudoporcinus)(예를 들어 균주 SPIN 20026), S. 뮤탄스(예를 들어 균주 UA 159, NN2025), S. 마카카에(S. macacae)(예를 들어 균주 NCTC11558), S. 갈롤리리쿠스(S. gallolylicus)(예를 들어 균주 UCN34, ATCC BAA-2069), S. 에쿠이네스(S. equines)(예를 들어 균주 ATCC 9812, MGCS 124), S. 디스달락티아에(S. dysdalactiae)(예를 들어 균주 GGS 124), S. 보비스(S. bovis)(예를 들어 균주 ATCC 700338), S. 안기노수스(S. anginosus)(예를 들어 균주 F0211), S. 아갈락티아에(S. agalactiae)(예를 들어 균주 NEM316, A909), 리스테리아 모노사이토게네스(예를 들어 균주 F6854), 리스테리아 이노쿠아(Listeria innocua)(L. 이노쿠아, 예를 들어 균주 Clip11262), 엔테로코커스 이탈리쿠스(Enterococcus italicus)(예를 들어 균주 DSM 15952) 또는 엔테로코커스 파에시움(Enterococcus faecium)(예를

들어 균주 1,231,408)의 Cas9 분자가 포함된다. 추가적인 예시적인 Cas9 분자는 네이세리아 메닌기티디스의 Cas9 분자(문헌[Hou et al. PNAS Early Edition 2013, 1-6]) 및 S. 아우레우스 Cas9 분자이다.

[0651] 실시형태에서, Cas9 분자, 예를 들어 활성 Cas9 분자 또는 비활성 Cas9 분자는, 본원에 기재된 임의의 Cas9 분자 서열 또는 자연적 발생 Cas9 분자 서열, 예를 들어 본원에 나열되거나 또는 문헌[Chylinski et al., RNA Biology 2013, 10:5, 'I2'I-T,1; Hou et al. PNAS Early Edition 2013, 1-6]에 기재된 종으로부터의 Cas9 분자와 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 상동성을 갖거나; 이와 비교할 때 1%, 2%, 5%, 10%, 15%, 20%, 30% 또는 40% 이하의 아미노산 잔기에서 상이하거나; 이와 적어도 1, 2, 5, 10 또는 20개의 아미노산만큼이지만, 100, 80, 70, 60, 50, 40 또는 30개 이하의 아미노산만큼 상이하거나; 또는 이와 동일한 아미노산 서열을 포함한다.

[0652] 실시형태에서, Cas9 분자는, *S. 피오게네스* Cas9 분자와 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 상동성을 갖거나; 이와 비교할 때 1%, 2%, 5%, 10%, 15%, 20%, 30% 또는 40% 이하의 아미노산 잔기에서 상이하거나; 이와 적어도 1, 2, 5, 10 또는 20개의 아미노산만큼이지만, 100, 80, 70, 60, 50, 40 또는 30개 이하의 아미노산만큼 상이하거나; 또는 이와 동일한 아미노산 서열을 포함한다.

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

[0653] Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn

```
    210                 215                 220
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510
```

[0654]

```
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
```

[0655]

```
                805                 810                 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
        995                 1000                1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys
    1010                1015                1020
Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser
1025                1030                1035                1040
Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu
                1045                1050                1055
Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile
            1060                1065                1070
Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser
        1075                1080                1085
Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly
    1090                1095                1100
```

[0656]

```
Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
1105                1110                1115                1120
Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser
                1125                1130                1135
Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly
            1140                1145                1150
Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile
        1155                1160                1165
Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala
    1170                1175                1180
Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys
1185                1190                1195                1200
Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser
                1205                1210                1215
Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr
            1220                1225                1230
Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
        1235                1240                1245
Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His
    1250                1255                1260
Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val
1265                1270                1275                1280
Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys
                1285                1290                1295
His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu
            1300                1305                1310
Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp
        1315                1320                1325
Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp
    1330                1335                1340
Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1345                1350                1355                1360
Asp Leu Ser Gln Leu Gly Gly Asp
                1365
```

[0657]

[0658] (SEQ ID NO: 205)

[0659] 실시형태에서, Cas9 분자는 비하전 또는 비극성 아미노산, 예를 들어 알라닌을 상기 위치에 도입하는 양전하 아미노산(예를 들어 리신, 아르기닌 또는 히스티딘)에 대한 하나 이상의 돌연변이를 포함하는 SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, 돌연변이는 Cas9의 nt-그루브 내의 하나 이상의 양전하 아미노산에 대한 것이다.

[0660] 실시형태에서, Cas9 분자는 SEQ ID NO: 205의 위치 855에서의 돌연변이, 예를 들어SEQ ID NO: 205의 위치 855 위치에 비하전 아미노산, 예를 들어 알라닌으로의 돌연변이를 포함하는 SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205에 비해 오직 SEQ ID NO: 205의 위치 855에서만, 예를 들어 비하전 아미노산, 예를 들어 알라닌으로의 돌연변이를 갖는다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205의 위치 810에서의 돌연변이, 위치 1003에서의 돌연변이 및/또는 위치 1060에서의 돌연변이, 예를 들어 SEQ ID NO: 205의 위치 810, 위치 1003 및/또는 위치 1060에서 알라닌으로의 돌연변이를 포함하는, SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205에 비해 오직 SEQ ID NO: 205의 위치 810, 위치 1003 및 위치 1060에서만 돌연변이를 가지며, 예를 들어 각 돌연변이는 비하전 아미노산, 예를 들어 알라닌으로이다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205의 위치 848에서의 돌연변이, 위치 1003에서의 돌연변이 및/또는 위치 1060에서의 돌연변이, 예를 들어 SEQ ID NO: 205의 위치 848, 위치 1003 및/또는 위치 1060에서 알라닌으로의 돌연변이를 포함하는, SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205에 비해 오직 SEQ ID NO: 205의 위치 848, 위치 1003 및 위치 1060에서만 돌연변이를 가지며, 예를 들어 각 돌연변이는 비하전 아미노산, 예를 들어 알라닌으로이다. 실시형태에서, Cas9 분자는 Science DOI: 10.1126/science.aad5227에서 2015년 12월 1일 온라인 이용 가능한 문헌[Slaymaker et al.,

Science Express]에 기재된 Cas9 분자이다.

[0661] 실시형태에서, Cas9 분자는 하나 이상의 돌연변이를 포함하는 SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치80 에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 80에 류신을 포함한다(즉, C80L 돌연변이를 갖는SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치574 에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 574에 글루탐산을 포함한다(즉, C574E 돌연변이를 갖는SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치80 에 돌연변이 및 위치 574에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 80에 류신, 및 SEQ ID NO: 205의 위치 574에 글루탐산을 포함한다(즉, C80L 돌연변이 및 C574E 돌연변이를 갖는SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 임의의 이론으로 결부시키고자 하는 것은 아니지만, 이러한 돌연변이는 Cas9 분자의 용해 특성을 개선시키는 것으로 여겨진다.

[0662] 실시형태에서, Cas9 분자는 하나 이상의 돌연변이를 포함하는 SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치147 에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 147에 티로신을 포함한다(즉, D147Y 돌연변이를 갖는 SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치411 에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 411에 트레오닌을 포함한다(즉, P411T 돌연변이를 갖는SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치147에 돌연변이 및 위치411 에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 147에 티로신, 및 SEQ ID NO: 205의 위치 411에 트레오닌을 포함한다(즉, D147Y 돌연변이 및 P411T 돌연변이를 갖는SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 임의의 이론으로 결부시키고자 하는 것은 아니지만, 이러한 돌연변이는 예를 들어 효모에서 Cas9 분자의 표적화 효율을 개선시키는 것으로 여겨진다.

[0663] 실시형태에서, Cas9 분자는 하나 이상의 돌연변이를 포함하는 SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 변이체는 SEQ ID NO: 205의 위치1135 에 돌연변이를 포함하며, 예를 들어 SEQ ID NO: 205의 위치 1135에 글루탐산을 포함한다(즉, D1135E 돌연변이를 갖는SEQ ID NO: 205를 포함하며, 예를 들어 이로 구성된다). 임의의 이론으로 결부시키고자 하는 것은 아니지만, 이러한 돌연변이는 NGG PAM 서열 대 NAG PAM 서열에 대한 Cas9 분자의 선택성을 개선시키는 것으로 여겨진다.

[0664] 실시형태에서, Cas9 분자는 비하전 또는 비극성 아미노산, 예를 들어 알라닌을 일정 위치에 도입하는 하나 이상의 돌연변이를 포함하는 SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205의 위치 497에서의 돌연변이, 위치 661에서의 돌연변이, 위치 695에서의 돌연변이 및/또는 위치 926에서의 돌연변이, 예를 들어 SEQ ID NO: 205의 위치 497, 위치 661, 위치 695 및/또는 위치 926에서 알라닌으로의 돌연변이를 포함하는, SEQ ID NO: 205의 *S. 피오게네스* Cas9 변이체이다. 실시형태에서, Cas9 분자는 SEQ ID NO: 205에 비해 오직 SEQ ID NO: 205의 위치 497, 위치 661, 위치 695 및 위치 926에서만 돌연변이를 가지며, 예를 들어 각 돌연변이는 비하전 아미노산, 예를 들어 알라닌으로이다. 임의의 이론으로 결부시키고자 하는 것은 아니지만, 이러한 돌연변이는 표적-외 부위에서의 Cas9 분자에 의한 절단을 감소하는 것으로 여겨진다.

[0665] 본원에 기재된 Cas9 분자에 대한 돌연변이는 조합될 수 있고, 본원에 기재된 임의의 융합 또는 다른 변형 및 본원에 기재된 검정에서 시험된 Cas9 분자 중 임의의 것과 조합될 수 있다는 것이 이해될 것이다.

[0666] 다양한 유형의 Cas 분자가 본원에 개시된 발명을 실시하기 위해 사용될 수 있다. 일부 실시형태에서, II형 Cas 시스템의 Cas 분자가 사용된다. 다른 실시형태에서, 다른 Cas 시스템의 Cas 분자가 사용된다. 예를 들어 I형 또는 III형 Cas 분자가 사용될 수 있다. 예시적인 Cas 분자(및 Cas 시스템)는 예를 들어 문헌[Haft et al, PLoS COMPUTATIONAL BIOLOGY 2005, 1(6): e60] 및 문헌[Makarova et al, NATURE REVIEW MICROBIOLOGY 201 1, 9:467-477]에 기재되어 있으며, 두 참조 문헌 모두의 내용은 그 전문이 참조로 본원에 포함된다.

[0667] 실시형태에서, Cas9 분자는 다음 활성 중 하나 이상을 포함한다: 니카제 활성; 이중 가닥 손상 활성(예를 들어 엔도뉴클레아제 및/또는 엑소 뉴클레아제 활성); 헬리카제 활성; 또는 gRNA 분자와 함께 표적 핵산에 배치하는 능력.

[0668] **변경된 Cas9 분자**

[0669] 자연적 발생 Cas9 분자는 니카제 활성, 뉴클레아제 활성(예를 들어 엔도뉴클레아제 및/또는 엑소뉴클레아제 활성); 헬리카제 활성; gRNA 분자와 기능적으로 관련되는 능력; 및 핵산 상의 부위를 표적화하는(또는 부위에 배

치하는) 능력(예를 들어 PAM 인지 및 특이성)을 포함하는 다수의 특성을 갖는다. 실시형태에서, Cas9 분자는 이들 특성의 모두 또는 서브세트를 포함할 수 있다. 전형적인 실시형태에서, Cas9 분자는 gRNA 분자와 상호작용하고, gRNA 분자와 협력하여 핵산 내 부위에 배치하는 능력을 갖는다. 다른 활성, 예를 들어 PAM 특이성, 절단 활성 또는 헬리카제 활성은 Cas9 분자에서 더욱 광범위하게 다를 수 있다.

[0670] 원하는 특성을 갖는 Cas9 분자는 다수의 방법으로, 예를 들어 원하는 특성을 갖는 변경된 Cas9 분자를 제공하기 위하여 모 Cas9 분자, 예를 들어 자연적 발생 Cas9 분자의 변경에 의해 제조될 수 있다. 예를 들어 모 Cas9 분자에 비해 하나 이상의 돌연변이 또는 차이가 도입될 수 있다. 이러한 돌연변이 및 차이는 치환(예를 들어 보존적 치환 또는 비필수 아미노산의 치환); 삽입; 또는 결실을 포함한다. 실시형태에서, Cas9 분자는 하나 이상의 돌연변이 또는 차이, 예를 들어 참조 Cas9 분자에 비해 적어도 1, 2, 3, 4, 5, 10, 15, 20, 30, 40 또는 50개의 돌연변이지만 200, 100 또는 80개 미만의 돌연변이를 포함할 수 있다.

[0671] 실시형태에서, 돌연변이 또는 돌연변이들은 Cas9 활성, 예를 들어 본원에 기재된 Cas9 활성에 대해 실질적인 효과를 갖지 않는다. 실시형태에서, 돌연변이 또는 돌연변이들은 Cas9 활성, 예를 들어 본원에 기재된 Cas9 활성에 대해 실질적인 효과를 갖는다. 실시형태에서, 예시적인 활성은 PAM 특이성, 절단 활성 및 헬리카제 활성 중 하나 이상을 포함한다. 돌연변이(들)는 예를 들어 하나 이상의 RuvC-유사 도메인, 예를 들어 N-말단의 RuvC-유사 도메인; HNH-유사 도메인; RuvC-유사 도메인 및 HNH-유사 도메인 외부의 영역에 존재할 수 있다. 일부 실시형태에서, 돌연변이(들)는 N-말단의 RuvC-유사 도메인에 존재한다. 일부 실시형태에서, 돌연변이(들)는 HNH-유사 도메인에 존재한다. 일부 실시형태에서, 돌연변이들은 N-말단의 RuvC-유사 도메인 및 HNH-유사 도메인 둘 모두에 존재한다.

[0672] 특정 서열, 예를 들어 치환이 하나 이상의 활성, 예를 들어 표적화 활성, 절단 활성 등에 영향을 미칠 수 있는지 여부는, 예를 들어 돌연변이가 보존적인지 여부를 평가함에 의해 또는 섹션 III에 기재된 방법에 의해 평가 또는 예측될 수 있다. 실시형태에서, Cas9 분자의 맥락에 사용된 "비필수" 아미노산 잔기는 Cas9 활성(예를 들어 절단 활성)을 파괴하지 않고 또는 더욱 바람직하게는 실질적으로 변경하지 않고, Cas9 분자, 예를 들어 자연적 발생 Cas9 분자, 예를 들어 활성 Cas9 분자의 야생형 서열로부터 변경될 수 있는 잔기인 반면, "필수" 아미노산 잔기를 변화시키는 것은 활성(예를 들어 절단 활성)의 실질적인 손실을 야기한다.

[0673] **변경된 PAM 인지를 갖는 또는 PAM 인지를 갖지 않는 Cas9 분자**

[0674] 자연적 발생 Cas9 분자는 특정 PAM 서열, 예를 들어 S. 피오게네스, S.써모필루스, S.뮤탄스, S.아우레우스 및 N. 메닌기티디스에 대해 상기 기재된 PAM 인지 서열을 인지할 수 있다.

[0675] 실시형태에서, Cas9 분자는 자연적 발생 Cas9 분자와 동일한 PAM 특이성을 갖는다. 다른 실시형태에서, Cas9 분자는 자연적 발생 Cas9 분자와 관련되지 않은 PAM 특이성, 또는 가장 가까운 서열 상동성을 갖는 자연적 발생 Cas9 분자와 관련되지 않은 PAM 특이성을 갖는다. 예를 들어 자연적 발생 Cas9 분자는, 예를 들어 PAM 인지를 변경하도록, 예를 들어 표적-외 부위를 감소시키고/시키거나 특이성을 개선하기 위해 Cas9 분자가 인지하는 PMA 서열을 변경하도록; 또는 PAM 인지 필요를 제거하도록, 변경될 수 있다. 실시형태에서, Cas9 분자는, 예를 들어 표적-외 부위를 감소시키고 특이성을 증가시키기 위해, 예를 들어 PAM 인지 서열의 길이를 증가시키고/시키거나 고수준의 동일성으로 Cas9 특이성을 개선하도록 변경될 수 있다. 실시형태에서, PAM 인지 서열의 길이는 적어도 4, 5, 6, 7, 8, 9, 10 또는 15개의 아미노산 길이이다. 상이한 PAM 서열을 인지하고/하거나 감소된 표적-외 활성을 갖는 Cas9 분자는 유도 진화(directed evolution)를 이용하여 생성될 수 있다. Cas9 분자의 유도 진화를 위해 사용될 수 있는 예시적인 방법 및 시스템은, 예를 들어 문헌[Esvelt el al, Nature 2011, 472(7344): 499-503]에 기재되어 있다. 후보 Cas9 분자는, 예를 들어 본원에 기재된 방법에 의해 평가될 수 있다.

[0676] **절단하지 않는 및 변형된 절단 Cas9 분자**

[0677] 실시형태에서, Cas9 분자는 자연적 발생 Cas9 분자와 상이한, 예를 들어 가장 가까운 상동성을 갖는 자연적 발생 Cas9 분자와 상이한 절단 특성을 포함한다. 예를 들어 Cas9 분자는 자연적 발생 Cas9 분자, 예를 들어 S. 피오게네스의 Cas9 분자와 다음과 같이 상이할 수 있다: 예를 들어 자연적 발생 Cas9 분자(예를 들어 S. 피오게네스의 Cas9 분자)에 비해 예를 들어 감소 또는 증가된, 이중 가닥 절단부의 절단(엔도뉴클레아제 및/또는 엑소뉴클레아제 활성)을 조절하는 능력; 예를 들어 자연적 발생 Cas9 분자(예를 들어 S. 피오게네스의 Cas9 분자)에 비해, 예를 들어 감소 또는 증가된, 핵산의 단일 가닥, 예를 들어 핵산 분자의 비상보적 가닥 또는 핵산 분자의 상보적 가닥의 절단을 조절하는 능력(니카제 활성); 또는 핵산 분자, 예를 들어 이중 가닥 또는 단일 가닥 핵산 분자를 절단하는 능력이 제거될 수 있다.

[0678] **변형된 절단 활성 Cas9 분자**

[0679] 실시형태에서, 활성 Cas9 분자는 다음의 활성 중 하나 이상을 포함한다: N-말단의 RuvC-유사 도메인과 관련된 절단 활성; HNH-유사 도메인과 관련된 절단 활성; HNH 도메인과 관련된 절단 활성 및 N-말단의 RuvC-유사 도메인과 관련된 절단 활성.

[0680] 실시형태에서, Cas9 분자는 Cas9 니카제고, 예를 들어 오직 DNA의 단일 가닥만 절단한다. 실시형태에서, Cas9 니카제는SEQ ID NO: 205의 위치 10에서의 돌연변이 및/또는 위치 840에서의 돌연변이를 포함하고, 예를 들어 SEQ ID NO: 205에 대해 D10A 및/또는 H840A 돌연변이를 포함한다.

[0681] **절단하지 않는 불활성 Cas9 분자**

[0682] 실시형태에서, 변경된 Cas9 분자는 핵산 분자(이중 가닥 또는 단일 가닥 핵산 분자 어느 것도)를 절단하지 않거나, 또는 예를 들어 본원에 기재된 검정에 의해 측정하여, 상당히 더 적은 효율로, 예를 들어 참조 Cas9 분자의 절단 활성의 20%, 10%, 5%, 1% 또는 0.1% 미만으로 핵산 분자를 절단하는 불활성 Cas9 분자이다. 참조Cas9 분자는 자연적으로 발생한 변형되지 않은 Cas9 분자, 예를 들어 S. 피오게네스, S.써모필루스, S.아우레우스 또는 N. 메닌기티디스의 Cas9 분자와 같은 자연적 발생 Cas9 분자일 수 있다. 실시형태에서, 참조Cas9 분자는 가장 가까운 서열 동일성 또는 상동성을 갖는 자연적 발생 Cas9 분자이다. 실시형태에서, 불활성 Cas9 분자는 실질적인 N-말단의 RuvC-유사 도메인과 관련된 절단 활성 및 HNH-유사 도메인과 관련된 절단 활성이 없다.

[0683] 실시형태에서, Cas9분자는 dCas9이다. 문헌[Tsai et al.(2014), Nat. Biotech. 32:569-577].

[0684] 촉매적 불활성 Cas9 분자는 전사 억제제와 융합될 수 있다. 불활성 Cas9 융합 단백질은 gRNA와 복합체 형성하고, gRNA의 표적화 도메인에 의해 특정된 DNA 서열로 배치하지만, 활성 Cas9와 달리, 이는 표적 DNA를 절단하지 않을 것이다. 효과기 도메인, 예를 들어 전사 억제 도메인의 불활성 Cas9로의 융합은 gRNA에 의해 특정된 임의의 DNA 부위로의 효과기 모집을 가능하게 한다. 유전자의 프로모터 영역으로의 Cas9 융합 단백질의 부위 특이적 표적화는 프로모터 영역에 중합효소가 결합하는 것, 예를 들어 전사 활성화를 증가 또는 저해하는 핵산에 전사 인자(예를 들어 전사 활성화제) 및/또는 전사 인핸서와 융합된 Cas9가 결합하는 것을 차단하거나 영향을 미칠 수 있다. 대안적으로, 유전자의 전사 억제제 또는 프로모터 영역으로의 Cas9-융합의 부위 특이적 표적화는 전사 활성화를 감소시키기 위해 사용될 수 있다.

[0685] 불활성 Cas9 분자에 융합될 수 있는 전사 억제제 또는 전사 억제제 도메인에는 크루펠(kruppel) 관련 박스(KRAB 또는 SKD), Mad mSIN3 상호작용 도메인(SID) 또는 ERF 억제제 도메인(ERD)이 포함될 수 있다.

[0686] 또 다른 실시형태에서, 불활성 Cas9 분자는 염색질을 변형하는 단백질과 융합될 수 있다. 예를 들어 불활성 Cas9 분자는 이질염색질 단백질 1(HP1), 히스톤 리신 메틸트랜스퍼라제(예를 들어 SUV39H1, SUV39H2, G9A, ESET/SETDB1, Pr-SET7/8, SUV4-20H1, RIZ1), 히스톤 리신 데메틸레이트(예를 들어 LSD1/BHC1 10, SpLsd1/Sw, 1/Saf1 10, Su(var)3-3, JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, Rph1, JARID1A/RBP2, JAR1D1B/PLU-1, JAR1D1C/SMCX, JARID1D/SMCY, Lid, Jhn2, Jmj2), 히스톤 리신 데아세틸라제(예를 들어 HDAC1, HDAC2, HDAC3, HDAC8, Rpd3, Hos1, Cir6, HDAC4, HDAC5, HDAC7, HDAC9, Hda1, Cir3, SIRT1, SIRT2, Sir2, Hst1, Hst2, Hst3, Hst4, HDAC11) 및 DNA 메틸라제(DNMT1, DNMT2a/DMNT3b, MET1)에 융합될 수 있다. 불활성 Cas9-염색질 변형 분자 융합 단백질은 표적 유전자의 발현을 감소시키기 위해 염색질 상태를 변경하는 데 사용될 수 있다.

[0687] 이종 서열(예를 들어 전사 억제제 도메인)은 불활성 Cas9 단백질의 N- 또는 C-말단에 융합될 수 있다. 대안적인 실시형태에서, 이종 서열(예를 들어 전사 억제제 도메인)은 불활성 Cas9 단백질의 내부 부분(즉, N-말단 또는 C-말단 외의 일부분)에 융합될 수 있다.

[0688] Cas9 분자/gRNA 분자 복합체의 표적 핵산에 결합하고 표적 핵산을 절단하는 능력은, 예를 들어 섹션 III에서 본원에 기재된 방법에 의해 평가될 수 있다. 활성 Cas9 또는 불활성 Cas9와 같은 Cas9 분자의 활성은 단독으로 또는 gRNA 분자와의 복합체에서, 유전자 발현 검정 및 염색질-기반 검정, 예를 들어 염색질 면역침전(ChiP) 및 염색질 생체내 검정(CiA)을 포함하는 당분야에 잘 알려진 방법에 의해 또한 평가될 수 있다.

[0689] **다른 Cas9 분자 융합**

[0690] 실시형태에서, Cas9 분자, 예를 들어 *S. 피오게네스*의 Cas9는 추가적인 활성을 부여하는 하나 이상의 아미노산 서열을 추가로 포함할 수 있다.

[0691] 일부 양태에서, Cas9 분자는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 NLS와 같은 하나 이상의 핵 국재화

서열(NLS)을 포함할 수 있다. 일부 실시형태에서, Cas9 분자는 아미노 말단에 또는 그 부근에 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 NLS, 카르복시 말단에 또는 그 부근에 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 NLS, 또는 이들의 조합(예를 들어 아미노 말단에 하나 이상의 NLS 및 카르복시 말단에 하나 이상의 NLS)를 포함한다. 하나를 초과하는 NLS가 존재할 때, 각각은 다른 것들과 독립적으로 선택될 수 있어, 단일 NLS는 하나를 초과하는 카피로 및/또는 하나 이상의 카피로 존재하는 하나 이상의 다른 NLS와 조합하여 존재할 수 있다. 일부 실시형태에서, NLS는 NLS의 가장 가까운 아미노산이 N 말단 또는 C 말단으로부터 폴리펩타이드 사슬을 따라 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 이상의 아미노산 이내일 때, N- 또는 C-말단 부근이라고 여겨진다. 전형적으로, NLS는 단백질 표면상에 노출된 하나 이상의 양전하 리신 또는 아르기닌의 짧은 서열로 구성되지만, 다른 유형의 NLS가 알려져 있다. NLS의 비제한적인 예로는 아미노산 서열 PKKKRKV(SEQ ID NO: 206)를 갖는 SV40 바이러스 대형 T-항원의 NLS; 서열 KRPAATKKAGQAKKKK(SEQ ID NO: 207)의 뉴클레오플라스민으로부터의 NLS(예를 들어 뉴클레오플라스민 2부분으로 된 NLS; 아미노산 서열 PAAKRVKLD(SEQ ID NO: 208) 또는 RQRRNELKRSP(SEQ ID NO: 209)를 갖는 c-myc NLS; 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY(SEQ ID NO: 210)를 갖는 hRNPA1 M9 NLS, 임포틴-알파로부터의 IBB 도메인의 서열 RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV(SEQ ID NO: 211); 근종 T 단백질의 서열 VSRKRPRP(SEQ ID NO: 212) 및 PPKKARED(SEQ ID NO: 213); 인간 p53의 서열 PQPKKKPL(SEQ ID NO: 214); 마우스 c-ab1 IV의 서열 SALIKKKKKMAP(SEQ ID NO: 215); 인플루엔자 바이러스 NS1의 서열 DRLRR(SEQ ID NO: 216) 및 PKQKKRK(SEQ ID NO: 217); 간염 바이러스 델타 항원의 서열 RKLKKKIKKL(SEQ ID NO: 218); 마우스 Mx1 단백질의 서열 REKKKFLKRR(SEQ ID NO: 219); 인간 폴리(ADP-리보스) 중합효소의 서열 KRKGDEVDGVDEVAKKKSKK(SEQ ID NO: 220); 및 스테로이드 호르몬 수용체(인간) 글루코코르티코이드의 서열 RKCLQAGMNLEARKTKK(SEQ ID NO: 221)를 포함하거나 이로부터 유래된 NLS 서열이 포함된다. 다른 적합한 NLS 서열은 당분야에 공지되어있다(예를 들어 문헌 [Sorokin, Biochemistry(Moscow)(2007) 72:13, 1439-1457; Lange J Biol Chem.(2007) 282:8, 5101-5]).

[0692] 실시형태에서, Cas9 분자, 예를 들어 *S. 피오게네스* Cas9 분자는, 예를 들어 Cas9 분자에 대해 N 말단 배치되는, SV40의 NLS 서열을 포함한다. 실시형태에서, Cas9 분자, 예를 들어 *S. 피오게네스* Cas9 분자는 Cas9 분자에 대해 N 말단 배치되는 SV40의 NLS 서열 및 Cas9 분자에 대해 C 말단 배치되는 SV40의 NLS 서열을 포함한다. 실시형태에서, Cas9 분자, 예를 들어 *S. 피오게네스* Cas9 분자는 Cas9 분자에 대해 N 말단 배치되는 SV40의 NLS 서열 및 Cas9 분자에 대해 C 말단 배치되는 뉴클레오플라스민의 NLS 서열을 포함한다. 임의의 상기 실시형태에서, 분자는 예를 들어 N 말단 또는 C 말단에 태그, 예를 들어 His 태그, 예를 들어 His(6) 태그(SEQ ID NO: 247) 또는 His(8) 태그(SEQ ID NO : 248)를 추가로 포함할 수 있다.

[0693] 일부 양태에서, Cas9 분자는 Cas9 분자가 특이적으로 인지되도록 하는 하나 이상의 아미노산 서열, 예를 들어 태그를 포함할 수 있다. 일 실시형태에서, 태그는 히스티딘 태그, 예를 들어 적어도 3, 4, 5, 6, 7, 8, 9, 10개 이상의 히스티딘 아미노산을 포함하는 히스티딘 태그이다. 실시형태에서, 히스티딘 태그는 His6 태그(6개 히스티딘)(SEQ ID NO: 247)이다. 다른 실시형태에서, 히스티딘 태그는 His8 태그(8개 히스티딘)(SEQ ID NO: 248)이다. 실시형태에서, 히스티딘 태그는 링커에 의해 Cas9 분자의 하나 이상의 다른 부분으로부터 분리될 수 있다. 실시형태에서, 링커는 GGS이다. 그러한 융합의 예는 Cas9 분자 iProt106520이다.

[0694] 일부 양태에서, Cas9 분자는 프로테아제에 의해 인지되는 하나 이상의 아미노산 서열을 포함(예를 들어 프로테아제 절단 부위를 포함)할 수 있다. 실시형태에서, 절단 부위는 담배 식각 바이러스(TEV) 절단 부위, 예를 들어 서열 ENLYFQG(SEQ ID NO: 230)를 포함한다. 일부 양태에서, 프로테아제 절단 부위, 예를 들어 TEV 절단 부위는 태그, 예를 들어 His6(SEQ ID NO: 247) 또는 His8 태그(SEQ ID NO: 248)와 같은 His 태그와 Cas9 분자의 나머지 사이에 배치된다. 임의의 이론으로 결부시키고자 하는 것은 아니지만, 이러한 도입은, 예를 들어 Cas9 분자의 정제를 위한 태그의 사용, 및 태그가 Cas9 분자 기능을 방해하지 않도록 후속적 절단을 감안할 것으로 여겨진다.

[0695] 실시형태에서, Cas9 분자(예를 들어 본원에 기재된 Cas9 분자)는 N-말단 NLS 및 C-말단 NLS를 포함(예를 들어 N-말단에서 C-말단으로 NLS-Cas9-NLS를 포함)하며, 예를 들어 각각의 NLS는 SV40 NLS(PKKKRKV(SEQ ID NO: 206))이다. 실시형태에서, Cas9 분자(예를 들어 본원에 기재된 Cas9 분자)는 N-말단 NLS, C-말단 NLS 및 C-말단 His6 태그(SEQ ID NO: 247)를 포함(예를 들어 N-말단에서 C-말단으로 NLS-Cas9-NLS-His 태그를 포함)하며, 예를 들어 각 NLS는 SV40 NLS(PKKKRKV(SEQ ID NO: 206))이다. 실시형태에서, Cas9 분자(예를 들어 본원에 기재된 Cas9 분자)는 N-말단 His 태그(예를 들어 His6 태그(SEQ ID NO: 247)), N-말단 NLS 및 C-말단 NLS를 포함(예를 들어 N-말단에서 C-말단으로 His 태그-NLS-Cas9-NLS를 포함)하며, 예를 들어 각 NLS는 SV40 NLS(PKKKRKV(SEQ ID NO: 206))이다. 실시형태에서, Cas9 분자(예를 들어 본원에 기재된 Cas9 분자)는 N-말단 NLS 및 C-말단 His

태그(예를 들어 His6 태그(SEQ ID NO: 247))를 포함(예를 들어 N-말단에서 C-말단으로 His 태그-Cas9-NLS를 포함)하며, NLS는 SV40 NLS(PKKKRKV(SEQ ID NO: 206))이다. 실시형태에서, Cas9 분자(예를 들어 본원에 기재된 Cas9 분자)는 N-말단 NLS 및 C-말단 His 태그(예를 들어 His6 태그(SEQ ID NO: 247))를 포함(예를 들어 N-말단에서 C-말단으로 NLS-Cas9-His 태그를 포함)하며, 예를 들어 NLS는 SV40 NLS(PKKKRKV(SEQ ID NO: 206))이다. 실시형태에서, Cas9 분자(예를 들어 본원에 기재된 Cas9 분자)는 N-말단 His 태그(예를 들어 His8 태그(SEQ ID NO: 248)), N-말단 절단 도메인(예를 들어 담배 식각 바이러스(TEV) 절단 도메인(예를 들어 서열 ENLYFQG(SEQ ID NO: 230)를 포함)), N-말단 NLS(예를 들어 SV40 NLS; SEQ ID NO: 206), 및 C-말단 NLS(예를 들어 SV40 NLS; SEQ ID NO: 206)를 포함(예를 들어 N-말단에서 C-말단으로 His 태그-TEV-NLS-Cas9-NLS를 포함)한다. 임의의 상기 실시형태에서, Cas9는 SEQ ID NO: 205의 서열을 갖는다. 대안적으로, 임의의 상기 실시형태에서, Cas9는, 예를 들어 본원에 기재된, SEQ ID NO: 205의 Cas9 변이체의 서열을 갖는다. 임의의 상기 실시형태에서, Cas9 분자는 His 태그와 분자의 또 다른 부분 사이에 링커, 예를 들어 GGS 링커를 포함한다. 상기 기재된 예시적인 Cas9 분자의 아미노산 서열이 아래에 제공된다. "iProt" 식별자는 도 6에서의 것과 매치된다.

[0696] iProt105026(단백질의 조제물에 따라, iProt106154, iProt106331, iProt106545, 및 PID426303으로도 지칭됨)(SEQ ID NO: 233):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL
FDSGETAEAT RLKRTARRRY TRRKNRICYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK
HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG
DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE
KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA
AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF
DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH

[0697] QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET

ITPWNFEEVV DKGASAQSFI ERMTNFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE
GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI ECFDSVEISG VEDRFNASLG
TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK
RRRYTGWGRL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMQLIHDDSL TFKEDIQKAQ
VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK
GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEKLY LYYLQNGRDM YVDQELDINR
LSDYDVDHIV PQSFLKDDSI DNKVLTRSDK NRGKSDNVPS EEVVKKMKNY
WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRQITKHVA QILDSRMNTK
YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP
KLESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKRPL
IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR
KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL
EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY
EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI
REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL
SQLGGDSRAD PKKKRKVHHH HHH

iProt106518 (SEQ ID NO: 234):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL
FDSGETAEAT RLKRTARRRY TRRKNRILYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK
HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG
DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE
KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA
AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF
DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH
QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET
ITPWNFEEVV DKGASAQSFI ERMTNFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE
GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI EEFDSVEISG VEDRFNASLG
TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK
RRRYTGWGRL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMQLIHDDSL TFKEDIQKAQ
VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK
GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEKLY LYYLQNGRDM YVDQELDINR
LSDYDVDHIV PQSFLKDDSI DNKVLTRSDK NRGKSDNVPS EEVVKKMKNY
WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRQITKHVA QILDSRMNTK
YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP
KLESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKRPL
[0698] IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR

KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL
EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY
EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI
REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL
SQLGGDSRAD PKKKRKVHHH HHH

iProt106519 (SEQ ID NO: 235):

MGSSHHHHHH HHENLYFQGS MDKKYSIGLD IGTNSVGWAV ITDEYKVPSK
KFKVLGNTDR HSIKKNLIGA LLFDSGETAE ATRLKRTARR RYTRRKNRIC YLQEIFSNEM
AKVDDSFFHR LEESFLVEED KKHERHPIFG NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD
LRLIYLALAH MIKFRGHFLI EGDLNPDNSD VDKLFIQLVQ TYNQLFEENP INASGVDAKA
ILSARLSKSR RLENLIAQLP GEKKNGLFGN LIALSLGLTP NFKSNFDLAE DAKLQLSKDT
YDDDLDNLLA QIGDQYADLF LAAKNLSDAI LLSDILRVNT EITKAPLSAS MIKRYDEHHQ
DLTLLKALVR QQLPEKYKEI FFDQSKNGYA GYIDGGASQE EFYKFIKPIL EKMDGTEELL
VKLNREDLLR KQRTFDNGSI PHQIHLGELH AILRRQEDFY PFLKDNREKI EKILTFRIPY
YVGPLARGNS RFAWMTRKSE ETITPWNFEE VVDKGASAQS FIERMTNFDK NLPNEKVLPK
HSLLYEYFTV YNELTKVKYV TEGMRKPAFL SGEQKKAIVD LLFKTNRKVT
VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI IKDKDFLDNE ENEDILEDIV
LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRYTGWG RLSRKLINGI RDKQSGKTIL
DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL HEHIANLAGS PAIKKGILQT
VKVVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER MKRIEEGIKE LGSQILKEHP
VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDH IVPQSFLKDD SIDNKVLTRS
DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL TKAERGGLSE
LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS KLVSDFRKDF
QFYKVREINN YHHAHDAYLN AVVGTALIKK YPKLESEFVY GDYKVYDVRK
MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
ATVRKVLSMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA
YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDLIIKLPK
YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFTLTNLGA
PAAFKYFDTT IDRKRYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGDGG GSPKKKRKV

iProt106520 (SEQ ID NO: 236):

MAHHHHHHGG SPKKKRKVDK KYSIGLDIGT NSVGWAVITD EYKVPSKKFK
VLGNTDRHSI KKNLIGALLF DSGETAEATR LKRTARRRYT RRKNRICYLQ EIFSNEMAKV
[0699] DDSFFHRLEE SFLVEEDKKH ERHPIFGNIV DEVAYHEKYP TIYHLRKKLV DSTDKADLRL

IYLALAHMIK FRGHFLIEGD LNPDNSDVDK LFIQLVQTYN QLFEENPINA SGVDAKAILS
ARLSKSRRLE NLIAQLPGEK KNGLFGNLIA LSLGLTPNFK SNFDLAEDAK LQLSKDTYDD
DLDNLLAQIG DQYADLFLAA KNLSDAILLS DILRVNTEIT KAPLSASMIK RYDEHHQDLT
LLKALVRQQL PEKYKEIFFD QSKNGYAGYI DGGASQEEFY KFIKPILEKM DGTEELLVKL
NREDLLRKQR TFDNGSIPHQ IHLGELHAIL RRQEDFYPFL KDNREKIEKI LTFRIPYYVG
PLARGNSRFA WMTRKSEETI TPWNFEEVVD KGASAQSFIE RMTNFDKNLP NEKVLPKHSL
LYEYFTVYNE LTKVKYVTEG MRKPAFLSGE QKKAIVDLLF KTNRKVTVKQ LKEDYFKKIE
CFDSVEISGV EDRFNASLGT YHDLLKIIKD KDFLDNEENE DILEDIVLTL TLFEDREMIE
ERLKTYAHLF DDKVMKQLKR RRYTGWGRLS RKLINGIRDK QSGKTILDFL KSDGFANRNF
MQLIHDDSLT FKEDIQKAQV SGQGDSLHEH IANLAGSPAI KKGILQTVKV VDELVKVMGR
HKPENIVIEM ARENQTTQKG QKNSRERMKR IEEGIKELGS QILKEHPVEN TQLQNEKLYL
YYLQNGRDMY VDQELDINRL SDYDVDHIVP QSFLKDDSID NKVLTRSDKN RGKSDNVPSE
EVVKKMKNYW RQLLNAKLIT QRKFDNLTKA ERGGLSELDK AGFIKRQLVE
TRQITKHVAQ ILDSRMNTKY DENDKLIREV KVITLKSKLV SDFRKDFQFY KVREINNYHH
AHDAYLNAVV GTALIKKYPK LESEFVYGDY KVYDVRKMIA KSEQEIGKAT
AKYFFYSNIM NFFKTEITLA NGEIRKRPLI ETNGETGEIV WDKGRDFATV RKVLSMPQVN
IVKKTEVQTG GFSKESILPK RNSDKLIARK KDWDPKKYGG FDSPTVAYSV LVVAKVEKGK
SKKLKSVKEL LGITIMERSS FEKNPIDFLE AKGYKEVKKD LIIKLPKYSL FELENGRKRM
LASAGELQKG NELALPSKYV NFLYLASHYE KLKGSPEDNE QKQLFVEQHK HYLDEIIEQI
SEFSKRVILA DANLDKVLSA YNKHRDKPIR EQAENIIHLF TLTNLGAPAA FKYFDTTIDR
KRYTSTKEVL DATLIHQSIT GLYETRIDLS QLGGDSRADP KKKRKV

iProt106521 (SEQ ID NO: 237):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL
FDSGETAEAT RLKRTARRRY TRRKNRICYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK
HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG
DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE
KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA
AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF
DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH
QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET
ITPWNFEEVV DKGASAQSFI ERMTNFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE
GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI ECFDSVEISG VEDRFNASLG
TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK
RRRYTGWGRL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMQLIHDDSL TFKEDIQKAQ
VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK
[0700] GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEKLY LYYLQNGRDM YVDQELDINR

LSDYDVDHIV PQSFLKDDSI DNKVLTRSDK NRGKSDNVPS EEVVKKMKNY
WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRQITKHVA QILDSRMNTK
YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP
KLESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKRPL
IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR
KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL
EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY
EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI
REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL
SQLGGDSRAD HHHHHH

iProt106522 (SEQ ID NO: 238):

MAHHHHHHGG SDKKYSIGLD IGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA
LLFDSGETAE ATRLKRTARR RYTRRKNRIC YLQEIFSNEM AKVDDSFFHR LEESFLVEED
KKHERHPIFG NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI
EGDLNPDNSD VDKLFIQLVQ TYNQLFEENP INASGVDAKA ILSARLSKSR RLENLIAQLP
GEKKNGLFGN LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLDNLLA QIGDQYADLF
LAAKNLSDAI LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLLKALVR QQLPEKYKEI
FFDQSKNGYA GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI
PHQIHLGELH AILRRQEDFY PFLKDNREKI EKILTFRIPY YVGPLARGNS RFAWMTRKSE
ETITPWNFEE VVDKGASAQS FIERMTNFDK NLPNEKVLPK HSLLYEYFTV YNELTKVKYV
TEGMRKPAFL SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS
LGTYHDLLKI IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ
LKRRRYTGWG RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK
AQVSGQGDSL HEHIANLAGS PAIKKGILQT VKVVDELVKV MGRHKPENIV IEMARENQTT
QKGQKNSRER MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI
NRLSDYDVDH IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK
NYWRQLLNAK LITQRKFDNL TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN
TKYDENDKLI REVKVITLKS KLVSDFRKDF QFYKVREINN YHHAHDAYLN AVVGTALIKK
YPKLESEFVY GDYKVYDVRK MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR
PLIETNGETG EIVWDKGRDF ATVRKVLSMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI
ARKKDWDPKK YGGFDSPTVA YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID
FLEAKGYKEV KKDLIIKLPK YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS
HYEKLKGSPE DNEQKQLFVE QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK
PIREQAENII HLFTLTNLGA PAAFKYFDTT IDRKRYTSTK EVLDATLIHQ SITGLYETRI
DLSQLGGDSR ADPKKKRKV

[0701]

iProt106658 (SEQ ID NO: 239):

MGSSHHHHHH HHENLYFQGS MDKKYSIGLD IGTNSVGWAV ITDEYKVPSK
KFKVLGNTDR HSIKKNLIGA LLFDSGETAE ATRLKRTARR RYTRRKNRIC YLQEIFSNEM
AKVDDSFFHR LEESFLVEED KKHERHPIFG NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD
LRLIYLALAH MIKFRGHFLI EGDLNPDNSD VDKLFIQLVQ TYNQLFEENP INASGVDAKA
ILSARLSKSR RLENLIAQLP GEKKNGLFGN LIALSLGLTP NFKSNFDLAE DAKLQLSKDT
YDDDLDNLLA QIGDQYADLF LAAKNLSDAI LLSDILRVNT EITKAPLSAS MIKRYDEHHQ
DLTLLKALVR QQLPEKYKEI FFDQSKNGYA GYIDGGASQE EFYKFIKPIL EKMDGTEELL
VKLNREDLLR KQRTFDNGSI PHQIHLGELH AILRRQEDFY PFLKDNREKI EKILTFRIPY
YVGPLARGNS RFAWMTRKSE ETITPWNFEE VVDKGASAQS FIERMTNFDK NLPNEKVLPK
HSLLYEYFTV YNELTKVKYV TEGMRKPAFL SGEQKKAIVD LLFKTNRKVT
VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI IKDKDFLDNE ENEDILEDIV
LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRYTGWG RLSRKLINGI RDKQSGKTIL
DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL HEHIANLAGS PAIKKGILQT
VKVVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER MKRIEEGIKE LGSQILKEHP
VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDH IVPQSFLKDD SIDNKVLTRS
DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL TKAERGGLSE
LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS KLVSDFRKDF
QFYKVREINN YHHAHDAYLN AVVGTALIKK YPKLESEFVY GDYKVYDVRK
MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
ATVRKVLSMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA
YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDLIIKLPK
YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFTLTNLGA
PAAFKYFDTT IDRKRYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGDGG GSPKKKRKV

iProt106745 (SEQ ID NO: 240):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL
FDSGETAEAT RLKRTARRRY TRRKNRICYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK
HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG
DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE
KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA
AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF
DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH
QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET
[0702] ITPWNFEEVV DKGASAQSFI ERMTNFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE

GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI ECFDSVEISG VEDRFNASLG
TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK
RRRYTGWGRL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMQLIHDDSL TFKEDIQKAQ
VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK
GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEKLY LYYLQNGRDM YVDQELDINR
LSDYDVDHIV PQSFLKDDSI DNAVLTRSDK NRGKSDNVPS EEVVKKMKNY
WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRQITKHVA QILDSRMNTK
YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP
KLESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKRPL
IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR
KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL
EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY
EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI
REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL
SQLGGDSRAD PKKKRKVHHH HHH

iProt106746 (SEQ ID NO: 241):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL
FDSGETAEAT RLKRTARRRY TRRKNRICYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK
HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG
DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE
KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA
AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF
DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH
QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET
ITPWNFEEVV DKGASAQSFI ERMTNFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE
GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI ECFDSVEISG VEDRFNASLG
TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK
RRRYTGWGRL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMQLIHDDSL TFKEDIQKAQ
VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK
GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEALY LYYLQNGRDM YVDQELDINR
LSDYDVDHIV PQSFLKDDSI DNKVLTRSDK NRGKSDNVPS EEVVKKMKNY
WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRQITKHVA QILDSRMNTK
YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP
ALESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKAPL
IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR
[0703] KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL

EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL SQLGGDSRAD PKKKRKVHHH HHH

iProt106747 (SEQ ID NO: 242):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL FDSGETAEAT RLKRTARRRY TRRKNRICYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET ITPWNFEEVV DKGASAQSFI ERMTNFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI ECFDSVEISG VEDRFNASLG TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK RRRYTGWGRL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMQLIHDDSL TFKEDIQKAQ VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEKLY LYYLQNGRDM YVDQELDINR LSDYDVDHIV PQSFLADDSI DNKVLTRSDK NRGKSDNVPS EEVVKKMKNY WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRQITKHVA QILDSRMNTK YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP ALESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKAPL IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL SQLGGDSRAD PKKKRKVHHH HHH

iProt106884 (SEQ ID NO: 243):

MAPKKKRKVD KKYSIGLDIG TNSVGWAVIT DEYKVPSKKF KVLGNTDRHS IKKNLIGALL FDSGETAEAT RLKRTARRRY TRRKNRICYL QEIFSNEMAK VDDSFFHRLE ESFLVEEDKK HERHPIFGNI VDEVAYHEKY PTIYHLRKKL VDSTDKADLR LIYLALAHMI KFRGHFLIEG DLNPDNSDVD KLFIQLVQTY NQLFEENPIN ASGVDAKAIL SARLSKSRRL ENLIAQLPGE

[0704]

KKNGLFGNLI ALSLGLTPNF KSNFDLAEDA KLQLSKDTYD DDLDNLLAQI GDQYADLFLA
AKNLSDAILL SDILRVNTEI TKAPLSASMI KRYDEHHQDL TLLKALVRQQ LPEKYKEIFF
DQSKNGYAGY IDGGASQEEF YKFIKPILEK MDGTEELLVK LNREDLLRKQ RTFDNGSIPH
QIHLGELHAI LRRQEDFYPF LKDNREKIEK ILTFRIPYYV GPLARGNSRF AWMTRKSEET
ITPWNFEEVV DKGASAQSFI ERMTAFDKNL PNEKVLPKHS LLYEYFTVYN ELTKVKYVTE
GMRKPAFLSG EQKKAIVDLL FKTNRKVTVK QLKEDYFKKI ECFDSVEISG VEDRFNASLG
TYHDLLKIIK DKDFLDNEEN EDILEDIVLT LTLFEDREMI EERLKTYAHL FDDKVMKQLK
RRRYTGWGAL SRKLINGIRD KQSGKTILDF LKSDGFANRN FMALIHDDSL TFKEDIQKAQ
VSGQGDSLHE HIANLAGSPA IKKGILQTVK VVDELVKVMG RHKPENIVIE MARENQTTQK
GQKNSRERMK RIEEGIKELG SQILKEHPVE NTQLQNEKLY LYYLQNGRDM YVDQELDINR
LSDYDVDHIV PQSFLKDDSI DNKVLTRSDK NRGKSDNVPS EEVVKKMKNY
WRQLLNAKLI TQRKFDNLTK AERGGLSELD KAGFIKRQLV ETRAITKHVA QILDSRMNTK
YDENDKLIRE VKVITLKSKL VSDFRKDFQF YKVREINNYH HAHDAYLNAV VGTALIKKYP
KLESEFVYGD YKVYDVRKMI AKSEQEIGKA TAKYFFYSNI MNFFKTEITL ANGEIRKRPL
IETNGETGEI VWDKGRDFAT VRKVLSMPQV NIVKKTEVQT GGFSKESILP KRNSDKLIAR
KKDWDPKKYG GFDSPTVAYS VLVVAKVEKG KSKKLKSVKE LLGITIMERS SFEKNPIDFL
EAKGYKEVKK DLIIKLPKYS LFELENGRKR MLASAGELQK GNELALPSKY VNFLYLASHY
EKLKGSPEDN EQKQLFVEQH KHYLDEIIEQ ISEFSKRVIL ADANLDKVLS AYNKHRDKPI
REQAENIIHL FTLTNLGAPA AFKYFDTTID RKRYTSTKEV LDATLIHQSI TGLYETRIDL
SQLGGDSRAD PKKKRKVHHH HHH

iProt 20109496 (SEQ ID NO: 244)

MAPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD
SGETAEATRLKRTARRRYTRRKNRILYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHP
IFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ
EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFL
KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT
NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIEEFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVL
TLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD
ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQ
NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK

[0705]

HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN
AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLA
NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNS
DKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP
IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH
YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE
QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDS
RADHHHHHH

[0706]

[0707] **Cas9 분자를 인코딩하는 핵산**

[0708] Cas9 분자, 예를 들어 활성 Cas9 분자 또는 불활성 Cas9 분자를 인코딩하는 핵산이 본원에 제공된다.

[0709] Cas9 분자를 인코딩하는 예시적인 핵산은 문헌[Cong et al, SCIENCE 2013, 399(6121):819-823]; 문헌[Wang et al, CELL 2013, 153(4):910-918]; 문헌[Mali et al., SCIENCE 2013, 399(6121):823-826]; 문헌[Jinek et al, SCIENCE 2012, 337(6096):816-821]에 기재되어 있다.

[0710] 실시형태에서, Cas9 분자를 인코딩하는 핵산은 합성 핵산 서열일 수 있다. 예를 들어 합성 핵산 분자는, 예를 들어 섹션 XIII에 기재된 바와 같이, 화학적으로 변형될 수 있다. 실시형태에서, Cas9 mRNA는 다음의 특성 중 하나 이상, 예를 들어 모두를 갖는다: 캡핑되고, 폴리아데닐화되고, 5-메틸시티딘 및/또는 슈도우리딘으로 치환

된다.

[0711] 추가적으로 또는 대안적으로, 합성 핵산 서열은 코돈 최적화될 수 있으며, 예를 들어 적어도 하나의 흔하지 않은(non-common) 코돈 또는 덜 흔한 코돈이 흔한 코돈으로 대체되었다. 예를 들어 합성 핵산은 최적화된, 예를 들어 본원에 기재된 것과 같은 포유류 발현 시스템에서의 발현에 최적화된, 메신저 mRNA의 합성을 지시할 수 있다.

[0712] S. 피오게네스의 Cas9 분자를 인코딩하는 예시적인 코돈 최적화된 핵산 서열이 하기에 제공된다.

atggataaaa agtacagcat cgggctggac atcggtacaa actcagtggg gtgggccgtg 60

attacggacg agtacaaggt accctccaaa aaatttaaag tgctgggtaa cacggacaga 120

cactctataa agaaaaatct tattggagcc ttgctgttcg actcaggcga gacagccgaa 180

gccacaaggt tgaagcggac cgccaggagg cggtatacca ggagaaagaa ccgcatatgc 240

tacctgcaag aaatcttcag taacgagatg gcaaaggttg acgatagctt tttccatcgc 300

ctggaagaat cctttcttgt tgaggaagac aagaagcacg aacggcaccc catctttggc 360

[0713] aatattgtcg acgaagtggc atatcacgaa aagtacccga ctatctacca cctcaggaag 420

aagctggtgg actctaccga taaggcggac ctcagactta tttatttggc actcgcccac 480

atgattaaat ttagaggaca tttcttgatc gagggcgacc tgaacccgga caacagtgac 540

gtcgataagc tgttcatcca acttgtgcag acctacaatc aactgttcga agaaaaccct 600

ataaatgctt caggagtcga cgctaaagca atcctgtccg cgcgcctctc aaaatctaga 660

agacttgaga atctgattgc tcagttgccc ggggaaaaga aaaatggatt gtttggcaac 720

ctgatcgccc tcagtctcgg actgacccca aatttcaaaa gtaacttcga cctggccgaa 780

gacgctaagc tccagctgtc caaggacaca tacgatgacg acctcgacaa tctgctggcc 840

cagattgggg atcagtacgc cgatctcttt ttggcagcaa agaacctgtc cgacgccatc 900

ctgttgagcg atatcttgag agtgaacacc gaaattacta aagcacccct tagcgcatct 960

atgatcaagc ggtacgacga gcatcatcag gatctgaccc tgctgaaggc tcttgtgagg 1020

caacagctcc ccgaaaaata caaggaaatc ttctttgacc agagcaaaaa cggctacgct 1080

ggctatatag atggtggggc cagtcaggag gaattctata aattcatcaa gcccattctc 1140

gagaaaatgg acggcacaga ggagttgctg gtcaaactta acagggagga cctgctgcgg 1200

aagcagcgga cctttgacaa cgggtctatc ccccaccaga ttcatctggg cgaactgcac 1260

gcaatcctga ggaggcagga ggatttttat cctttcttа aagataaccg cgagaaaata 1320

gaaaagattc ttacattcag gatcccgtac tacgtgggac ctctcgcccg gggcaattca 1380

cggtttgcct ggatgacaag gaagtcagag gagactatta caccttggaa cttcgaagaa 1440

gtggtggaca agggtgcatc tgcccagtct ttcatcgagc ggatgacaaa ttttgacaag 1500

aacctcccta atgagaaggt gctgcccaaa cattctctgc tctacgagta ctttaccgtc 1560

tacaatgaac tgactaaagt caagtacgtc accgagggaa tgaggaagcc ggcattcctt 1620

agtggagaac agaagaaggc gattgtagac ctgttgttca agaccaacag gaaggtgact 1680

gtgaagcaac ttaaagaaga ctactttaag aagatcgaat gttttgacag tgtggaaatt 1740

tcaggggttg aagaccgctt caatgcgtca ttggggactt accatgatct tctcaagatc 1800

ataaaggaca aagacttcct ggacaacgaa gaaaatgagg atattctcga agacatcgtc 1860

ctcaccctga ccctgttcga agacagggaa atgatagaag agcgcttgaa aacctatgcc 1920

[0714] cacctcttcg acgataaagt tatgaagcag ctgaagcgca ggagatacac aggatgggga 1980

agattgtcaa ggaagctgat caatggaatt agggataaac agagtggcaa gaccatactg 2040

gatttcctca aatctgatgg cttcgccaat aggaacttca tgcaactgat tcacgatgac 2100

tctcttacct tcaaggagga cattcaaaag gctcaggtga gcgggcaggg agactcccttt 2160

catgaacaca tcgcgaattt ggcaggttcc cccgctatta aaaagggcat ccttcaaact 2220

gtcaaggtgg tggatgaatt ggtcaaggta atgggcagac ataagccaga aaatattgtg 2280

atcgagatgg cccgcgaaaa ccagaccaca cagaagggcc agaaaaatag tagagagcgg 2340

atgaagagga tcgaggaggg catcaaagag ctgggatctc agattctcaa agaacacccc 2400

gtagaaaaca cacagctgca gaacgaaaaa ttgtacttgt actatctgca gaacggcaga 2460

gacatgtacg tcgaccaaga acttgatatt aatagactgt ccgactatga cgtagaccat 2520

atcgtgcccc agtccttcct gaaggacgac tccattgata acaaagtctt gacaagaagc 2580

gacaagaaca ggggtaaaag tgataatgtg cctagcgagg aggtggtgaa aaaaatgaag 2640

aactactggc gacagctgct taatgcaaag ctcattacac aacggaagtt cgataatctg 2700

acgaaagcag agagaggtgg cttgtctgag ttggacaagg cagggtttat taagcggcag 2760

ctggtggaaa ctaggcagat cacaaagcac gtggcgcaga ttttggacag ccggatgaac 2820

acaaaatacg acgaaaatga taaactgata cgagaggtca aagttatcac gctgaaaagc 2880

aagctggtgt ccgattttcg gaaagacttc cagttctaca aagttcgcga gattaataac 2940

taccatcatg ctcacgatgc gtacctgaac gctgttgtcg ggaccgcctt gataaagaag 3000

tacccaaagc tggaatccga gttcgtatac ggggattaca aagtgtacga tgtgaggaaa 3060

atgatagcca agtccgagca ggagattgga aaggccacag ctaagtactt cttttattct 3120

aacatcatga attttttaa gacggaaatt accctggcca acggagagat cagaaagcgg 3180

ccccttatag agacaaatgg tgaaacaggt gaaatcgtct gggataaggg cagggatttc 3240

gctactgtga ggaaggtgct gagtatgcca caggtaaata tcgtgaaaaa aaccgaagta 3300

cagaccggag gattttccaa ggaaagcatt ttgcctaaaa gaaactcaga caagctcatc 3360

gcccgcaaga aagattggga ccctaagaaa tacgggggat ttgactcacc caccgtagcc 3420

tattctgtgc tggtggtagc taaggtggaa aaaggaaagt ctaagaagct gaagtccgtg 3480

[0715] aaggaactct tgggaatcac tatcatggaa agatcatcct ttgaaaagaa ccctatcgat 3540

ttcctggagg ctaagggtta caaggaggtc aagaaagacc tcatcattaa actgccaaaa 3600

tactctctct tcgagctgga aaatggcagg aagagaatgt tggccagcgc cggagagctg 3660

caaaagggaa acgagcttgc tctgccctcc aaatatgtta attttctcta tctcgcttcc 3720

cactatgaaa agctgaaagg gtctcccgaa gataacgagc agaagcagct gttcgtcgaa 3780

cagcacaagc actatctgga tgaaataatc gaacaaataa gcgagttcag caaaagggtt 3840

atcctggcgg atgctaattt ggacaaagta ctgtctgctt ataacaagca ccgggataag 3900

cctattaggg aacaagccga gaatataatt cacctcttta cactcacgaa tctcggagcc 3960

cccgccgcct tcaaatactt tgatacgact atcgaccgga aacggtatac cagtaccaaa 4020

gaggtcctcg atgccaccct catccaccag tcaattactg gcctgtacga aacacggatc 4080

gacctctctc aactgggcgg cgactag 4107

[0716] (SEQ ID NO: 222)

[0717] SEQ ID NO: 244를 포함하는 Cas9 분자를 인코딩하는 예시적인 코돈 최적화된 핵산 서열이 하기에 제공된다:

ATGGCTCCGAAGAAAAAGCGTAAAGTGGATAAAAAATACAGCATTGGTCTGGACATTGG
CACGAACTCAGTGGGTTGGGCGGTCATCACGGATGAATATAAGGTCCCGTCAAAAAAGT
TCAAAGTGCTGGGCAACACCGATCGCCATTCGATTAAAAAGAATCTGATCGGCGCGCTG
CTGTTTGATAGCGGTGAAACCGCGGAAGCAACGCGTCTGAAACGTACCGCACGTCGCCG
TTACACGCGCCGTAAAAATCGTATTCTGTATCTGCAGGAAATCTTTAGCAACGAAATGGC
GAAAGTTGATGACTCATTTTTCCACCGCCTGGAAGAATCGTTTCTGGTCGAAGAAGACAA
AAAGCATGAACGTCACCCGATTTTCGGTAATATCGTTGATGAAGTCGCGTACCATGAAAA
ATATCCGACGATTTACCATCTGCGTAAAAAACTGGTGGATTCAACCGACAAAGCCGATCT
GCGCCTGATTTACCTGGCACTGGCTCATATGATCAAATTTCGTGGCCACTTCCTGATTGA
AGGTGACCTGAACCCGGATAACTCTGACGTTGATAAGCTGTTCATCCAGCTGGTCCAAAC
CTATAATCAGCTGTTCGAAGAAAACCCGATCAATGCAAGTGGCGTTGATGCGAAGGCCA
TTCTGTCCGCTCGCCTGAGTAAATCCCGCCGTCTGGAAAACCTGATTGCACAACTGCCGG
GCGAAAAGAAAAACGGCCTGTTTGGTAATCTGATCGCTCTGTCACTGGGTCTGACGCCGA
ACTTTAAATCGAATTTCGACCTGGCAGAAGATGCTAAGCTGCAGCTGAGCAAAGATACC
TACGATGACGATCTGGACAACCTGCTGGCGCAAATTGGTGACCAGTATGCCGACCTGTTT
CTGGCGGCCAAAAATCTGTCAGATGCCATTCTGCTGTCGGACATCCTGCGCGTGAACACC
GAAATCACGAAAGCGCCGCTGTCAGCCTCGATGATTAAACGCTACGATGAACATCACCA
GGACCTGACCCTGCTGAAAGCACTGGTTCGTCAGCAACTGCCGGAAAAGTACAAGGAAA
[0718] TTTTCTTTGACCAATCTAAGAACGGCTATGCAGGTTACATCGATGGCGGTGCTAGTCAGG

AAGAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAAATGGATGGCACGGAAGAACTG
CTGGTGAAACTGAATCGTGAAGATCTGCTGCGTAAACAACGCACCTTTGACAACGGCAG
CATTCCGCATCAGATCCACCTGGGTGAACTGCATGCGATTCTGCGCCGTCAGGAAGATTT
TTATCCGTTCCTGAAAGACAACCGTGAAAAAATTGAAAAGATCCTGACGTTTCGCATCCC
GTATTACGTTGGCCCGCTGGCGCGTGGTAATAGCCGCTTCGCCTGGATGACCCGCAAATC
TGAAGAAACCATTACGCCGTGGAACTTTGAAGAAGTGGTTGATAAAGGTGCAAGCGCTC
AGTCTTTTATCGAACGTATGACCAATTTCGATAAAAACCTGCCGAATGAAAAGGTCCTGC
CGAAACATAGCCTGCTGTATGAATACTTTACCGTGTACAACGAACTGACGAAAGTGAAG
TATGTTACCGAAGGCATGCGCAAACCGGCGTTTCTGTCTGGTGAACAGAAAAAAGCCAT
TGTGGATCTGCTGTTCAAGACCAATCGTAAAGTTACGGTCAAACAGCTGAAGGAAGATT
ACTTCAAAAAGATCGAAGAATTCGACAGCGTGGAAATTTCTGGCGTTGAAGATCGTTTCA
ACGCCAGTCTGGGTACCTATCATGACCTGCTGAAGATCATCAAGGACAAGGATTTTCTGG
ATAACGAAGAAAATGAAGACATTCTGGAAGATATCGTGCTGACCCTGACGCTGTTCGAA
GATCGTGAAATGATTGAAGAACGCCTGAAAACGTACGCACACCTGTTTGACGATAAAGT
TATGAAGCAGCTGAAACGCCGTCGCTATACCGGCTGGGGTCGTCTGTCTCGCAAACTGAT
TAATGGCATCCGCGATAAGCAAAGTGGTAAAACGATTCTGGATTTCCTGAAATCCGACG
GCTTTGCCAACCGTAATTTCATGCAGCTGATCCATGACGATAGTCTGACCTTTAAGGAAG
ACATTCAGAAAGCACAAGTGTCAGGCCAGGGTGATTCGCTGCATGAACACATTGCGAAC
CTGGCCGGCTCCCCGGCTATTAAAAAGGGTATCCTGCAGACCGTCAAAGTCGTGGATGA
ACTGGTGAAGGTTATGGGCCGTCACAAACCGGAAAACATTGTGATCGAAATGGCGCGCG
AAAATCAGACCACGCAAAAGGGTCAGAAAAACTCACGTGAACGCATGAAGCGCATTGA
AGAAGGCATCAAAGAACTGGGTTCGCAGATTCTGAAAGAACATCCGGTTGAAAACACCC
AGCTGCAAAATGAAAAACTGTACCTGTATTACCTGCAAAATGGCCGTGACATGTATGTCG
ATCAGGAACTGGACATCAACCGCCTGAGCGACTATGATGTCGACCACATTGTGCCGCAG
AGCTTTCTGAAGGACGATTCTATCGATAATAAAGTGCTGACCCGTTCTGATAAGAACCGC
GGTAAAAGCGACAATGTTCCGTCTGAAGAAGTTGTCAAAAAGATGAAGAACTACTGGCG
TCAACTGCTGAATGCGAAGCTGATTACGCAGCGTAAATTCGATAACCTGACCAAGGCGG
AACGCGGCGGTCTGAGTGAACTGGATAAGGCCGGCTTTATCAAACGTCAACTGGTGGAA
ACCCGCCAGATTACGAAACATGTTGCCCAGATCCTGGATTCCCGCATGAACACGAAATAT
GACGAAAATGATAAGCTGATTCGTGAAGTCAAAGTGATCACCCTGAAGAGTAAGCTGGT
GTCCGATTTCCGTAAGGACTTTCAGTTCTACAAAGTTCGCGAAATTAACAATTACCATCA
CGCACACGATGCTTATCTGAATGCAGTGGTTGGCACCGCTCTGATCAAAAAGTATCCGAA
ACTGGAAAGCGAATTTGTGTATGGTGATTACAAAGTCTATGACGTGCGCAAGATGATTGC
GAAAAGTGAACAGGAAATCGGCAAGGCGACCGCCAAGTACTTTTTCTATTCCAACATCA
TGAACTTTTTCAAGACCGAAATCACGCTGGCAAATGGCGAAATTCGTAAACGCCCGCTG
ATCGAAACCAACGGCGAAACGGGTGAAATTGTGTGGGATAAAGGTCGTGACTTCGCGAC
CGTTCGCAAAGTCCTGTCAATGCCGCAAGTGAATATCGTTAAAAAGACCGAAGTTCAGA

[0719]

CGGGCGGTTTTAGTAAAGAATCCATCCTGCCGAAGCGTAACTCGGATAAACTGATTGCGC
GCAAAAAGGATTGGGACCCGAAAAAGTACGGCGGTTTTGATAGTCCGACCGTTGCATAT
TCCGTCCTGGTCGTGGCTAAAGTTGAAAAAGGCAAGAGTAAAAAGCTGAAGTCCGTCAA
AGAACTGCTGGGTATTACCATCATGGAACGTAGCTCTTTTGAAAAGAACCCGATTGACTT
CCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGATCTGATTATCAAGCTGCCGAAAT
ATTCGCTGTTCGAACTGGAAAACGGTCGTAAACGCATGCTGGCAAGCGCTGGCGAACTG
CAGAAGGGTAATGAACTGGCACTGCCGTCTAAATATGTGAACTTTCTGTACCTGGCTAGC
CATTATGAAAAACTGAAGGGTTCTCCGGAAGATAACGAACAGAAGCAACTGTTCGTTGA
ACAACATAAACACTACCTGGATGAAATCATCGAACAGATCTCAGAATTCTCGAAACGCG
TCATTCTGGCGGATGCCAATCTGGACAAAGTGCTGAGCGCGTATAACAAGCATCGTGAT
AAACCGATTCGCGAACAGGCCGAAAATATTATCCACCTGTTTACCCTGACGAACCTGGGC
GCACCGGCAGCTTTTAAATACTTCGATACCACGATCGACCGTAAGCGCTATACCAGCACG
AAAGAAGTTCTGGATGCTACCCTGATTCATCAGTCAATCACCGGTCTGTATGAAACGCGT
ATTGACCTGAGCCAACTGGGCGGTGATAGCCGTGCCGACCATCACCATCACCATCACTAA

[0720] TAG (SEQ ID NO: 223)

[0721] 상기 Cas9 서열이 C-말단에서 펩타이드 또는 폴리펩타이드와 융합되는 경우(예를 들어 C-말단에서 전사 억제제와 융합된 불활성 Cas9), 정지 코돈이 제거될 것으로 이해된다.

[0722] Cas9 분자, 예를 들어 본원에 기재된 Cas9 분자의 생산을 위한 핵산, 벡터 및 세포가 본원에 또한 제공된다. 폴리펩타이드 분자의 재조합 생산은 당업자에게 공지된 기술을 사용하여 달성될 수 있다. 예를 들어 본원에 기재된 바와 같은 폴리펩타이드 분자, 예컨대 Cas9 분자의 재조합 생산을 위한 분자 및 방법은 본원에 기재된다. 이와 관련하여 사용된 바와 같이, "재조합" 분자 및 생산은 재조합 수단에 의해 제조되거나, 발현되거나, 생성되거나 단리된 모든 폴리펩타이드(예를 들어 Cas9 분자, 예를 들어 본원에 기재된 바와 같음), 예컨대 관심 분자를 인코딩하는 핵산에 대해 유전자 이식(transgenic) 또는 염색체 이식(transchromosomal)인 동물(예를 들어 마우스)로부터 단리된 폴리펩타이드, 이로부터 제조된 하이브리도마, 이러한 분자를 발현하도록 형질변환된 숙주 세포로부터, 예를 들어 트랜스펙토마(transfectoma)로부터 단리된 분자, 재조합, 조합적 라이브러리로부터 단리된 분자, 및 이러한 분자(또는 이의 일부)를 인코딩하는 유전자의 모두 또는 일부를 다른 DNA 서열에 스플라이싱하는 단계를 수반하는 임의의 다른 수단에 의해 제조되거나, 발현되거나, 생성되거나 단리된 분자를 포함한다. 재조합 생산은 숙주 세포, 예를 들어 본원에 기재된 분자, 예를 들어 Cas9 분자, 예를 들어 본원에 기재된 Cas9 분자를 인코딩하는 핵산을 포함하는 숙주 세포로부터일 수 있다.

[0723] 예를 들어 본원에 기재된 바와 같은 분자(예를 들어 Cas9 분자 및/또는 gRNA 분자)를 인코딩하는 핵산 분자가 본원에 제공된다. 구체적으로, SEQ ID NO: 233 내지 SEQ ID NO: 244 중 임의의 하나를 인코딩하거나, 임의의 SEQ ID NO: 233 내지 SEQ ID NO: 244의 단편을 인코딩하거나, 임의의 SEQ ID NO: 233 내지 SEQ ID NO: 244와 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99% 서열 상동성을 포함하는 폴리펩타이드를 인코딩하는 서열을 포함하는 핵산 분자가 제공된다.

[0724] 임의의 상기 기재된 핵산 분자를 포함하는 예를 들어 본원에 기재된 바와 같은 벡터가 본원에 제공된다. 실시형태에서, 상기 핵산 분자는 프로모터, 예를 들어 이러한 벡터가 도입되는 숙주 세포에서 작동 가능한 프로모터에 작동적으로 연결된다.

[0725] 본원에 기재된 하나 이상의 핵산 분자 및/또는 벡터를 포함하는 숙주 세포가 본원에 제공된다. 실시형태에서, 숙주 세포는 원핵 숙주 세포이다. 실시형태에서, 숙주 세포는 진핵 숙주 세포이다. 실시형태에서, 숙주 세포는 효모 또는 e. 콜라이 세포이다. 실시형태에서, 숙주 세포는 포유류 세포, 예를 들어 인간 세포이다. 이러한 숙주 세포는 본원에 기재된 재조합 분자, 예를 들어 본원에 기재된 바와 같은 예를 들어 Cas9 또는 gRNA 분자의 생산에 사용될 수 있다.

[0726] **VI. 후보 분자의 기능적 분석**

[0727] 후보 Cas9 분자, 후보 gRNA 분자, 후보 Cas9 분자/gRNA 분자 복합체는 당분야에 공지된 방법에 의해 또는 본원에 기재된 바와 같이 평가될 수 있다. 예를 들어 Cas9 분자의 엔도뉴클레아제 활성을 평가하기 위한 예시적인 방법은 예를 들어 문헌[Jinek el al., SCIENCE 2012; 337(6096):816-821]에 기재되어 있다.

[0728] **VII. 주형 핵산(핵산 도입용)**

[0729] 본원에 사용된 용어 "주형 핵산" 또는 "공여자 주형"은 본 발명의 CRISPR 시스템에 의해 변형된, 예를 들어 절단된 표적 서열에서 또는 그 부근에서 삽입되는 핵산을 지칭한다. 실시형태에서, 표적 서열의 또는 그 부근의 핵산 서열은 전형적으로 절단 부위(들)에 또는 그 부근에, 주형 핵산의 서열의 일부 또는 전부를 갖도록 변형된다. 실시형태에서, 주형 핵산은 단일 가닥이다. 대안적인 실시형태에서, 주형 핵산은 이중 가닥이다. 실시형태에서, 주형 핵산은 DNA, 예를 들어 이중 가닥 DNA이다. 대안적인 실시형태에서, 주형 핵산은 단일 가닥 DNA이다.

[0730] 실시형태에서, 주형 핵산은 글로빈 단백질, 예를 들어 베타 글로빈을 인코딩하는 서열을 포함하며, 예를 들어 베타 글로빈 유전자를 포함한다. 실시형태에서, 핵산에 의해 인코딩되는 베타 글로빈은 하나 이상의 돌연변이, 예를 들어 항-겸상적혈구화 돌연변이를 포함한다. 실시형태에서, 핵산에 의해 인코딩되는 베타 글로빈은 돌연변이 T87Q를 포함한다. 실시형태에서, 핵산에 의해 인코딩되는 베타 글로빈은 돌연변이 G16D를 포함한다. 실시형태에서, 핵산에 의해 인코딩되는 베타 글로빈은 돌연변이 E22A를 포함한다. 실시형태에서, 베타 글로빈 유전자는 돌연변이 G16D, E22A 및 T87Q를 포함한다. 실시형태에서, 주형 핵산은 하나 이상의 조절 요소, 예를 들어 프로모터(예를 들어 인간 베타 글로빈 프로모터), 3' 인핸서, 및/또는 글로빈 유전자좌 제어 영역(예를 들어 하나 이상의 DNAseI 과민성 부위(예를 들어 인간 글로빈 유전자좌의 HS2, HS3 및/또는 HS4))의 적어도 일부분을 추가로 포함한다.

[0731] 다른 실시형태에서, 주형 핵산은 감마 글로빈을 인코딩하는 서열을 포함하며, 예를 들어 감마 글로빈 유전자를

포함한다. 실시형태에서, 주형 핵산은 감마 글로빈 단백질의 하나를 초과하는 카피를 인코딩하는 서열을 포함하며, 예를 들어 2개 이상, 예를 들어 2개의 감마 글로빈 유전자 서열을 포함한다. 실시형태에서, 주형 핵산은 하나 이상의 조절 요소, 예를 들어 프로모터 및/또는 인핸서를 추가로 포함한다.

[0732] 실시형태에서, 주형 핵산은 상동성 지시된 수선 사건에 참여함으로써 표적 위치의 구조를 변경한다. 실시형태에서, 주형 핵산은 표적 위치의 서열을 변경한다. 실시형태에서, 주형 핵산은 변형된 또는 자연적 발생이 아닌 염기의 표적 핵산으로의 도입을 야기한다.

[0733] 본원에 기재된 유전자 또는 경로의 돌연변이는 본원에서 논의된 접근법 중 하나를 이용하여 교정될 수 있다. 실시형태에서, 본원에 기재된 유전자 또는 경로의 돌연변이는 주형 핵산을 이용하는 상동성 지시된 수선(HDR)에 의해 교정된다. 실시형태에서, 본원에 기재된 유전자 또는 경로의 돌연변이는 주형 핵산을 이용하는 상동 재조합(HR)에 의해 교정된다. 실시형태에서, 본원에 기재된 유전자 또는 경로의 돌연변이는 주형 핵산을 이용하는 비상동 말단 접합(NHEJ) 수선에 의해 교정된다. 다른 실시형태에서, 관심 분자를 인코딩하는 핵산은 본 발명의 CRISPR 시스템에 의해 변형되는 부위에 또는 그 부근에 삽입될 수 있다. 실시형태에서, 주형 핵산은 예를 들어 본원에 기재된, 관심 분자를 인코딩하는 핵산 서열에 작동 가능하게 연결된 조절 요소, 예를 들어 하나 이상의 프로모터 및/또는 인핸서를 포함한다.

[0734] HDR 또는 HR 수선 및 주형 핵산

[0735] 본원에 기재된 바와 같이, 뉴클레아제-유도된 상동성 지시된 수선(HDR) 또는 상동 재조합(HR)은 표적 서열을 변경하고 게놈 내의 돌연변이를 교정(예를 들어 수선 또는 편집)하는 데 사용될 수 있다. 이론에 얽매이기를 원하지 않지만, 표적 서열의 변경은 공여자 주형 또는 주형 핵산에 기반한 수선에 의해 발생하는 것으로 여겨진다. 예를 들어 공여자 주형 또는 주형 핵산은 표적 서열의 변경을 대비한다. 플라스미드 공여자 또는 선형 이중 가닥 주형이 상동 재조합을 위한 주형으로서 사용될 수 있음이 고려된다. 단일 가닥 공여자 주형이 표적 서열과 공여자 주형 사이에서의 상동성 지시 수선(예를 들어 단일 가닥 어닐링)이라는 대안적 방법에 의해 표적 서열을 변경하기 위한 주형으로서 사용될 수 있음이 추가로 고려된다. 표적 서열의 공여자 주형-영향 변경은 Cas9 분자에 의한 절단에 의존할 수 있다. Cas9에 의한 절단은 이중 가닥 절단부, 하나의 단일 가닥 단절, 또는 2개의 단일 가닥 단절을 포함할 수 있다.

[0736] 실시형태에서, 돌연변이는 하나의 이중 가닥 절단부 또는 두 개의 단일 가닥 단절에 의해 교정될 수 있다. 실시형태에서, 돌연변이는(1)하나의 이중 가닥 절단부, (2)두 개의 단일 가닥 단절, (3)단절이 표적 서열의 각 편 상에서 발생하는 두 개의 이중 가닥 절단부, (4)이중 가닥 절단부 및 두 개의 단일 가닥 단절이 표적 서열의 각 편 상에서 발생하는 하나의 이중 가닥 절단부 및 두 개의 단일 가닥 단절, (5)한 쌍의 단일 가닥 단절이 표적 서열의 각 편 상에서 발생하는 네 개의 단일 가닥 단절, 또는(6)하나의 단일 가닥 단절을 형성하는 CRISPR/Cas9 시스템 및 주형을 제공함으로써 교정될 수 있다.

[0737] 이중 가닥 절단부 매개 교정

[0738] 실시형태에서, 이중 가닥 절단은 HNH-유사 도메인과 관련된 절단 활성 및 RuvC-유사 도메인, 예를 들어 N-말단 RuvC-유사 도메인과 관련된 절단 활성을 갖는 Cas9분자, 예를 들어 야생형 Cas9에 의해 영향을 받는다. 이러한 실시형태는 오직 단일 gRNA만을 필요로 한다.

[0739] 단일 가닥 단절 매개 교정

[0740] 다른 실시형태에서, 2개의 단일 가닥 단절 또는 닉(nick)은 니카제 활성, 예를 들어 HNH-유사 도메인과 관련된 절단 활성 또는 N-말단 RuvC-유사 도메인과 관련된 절단 활성을 갖는 Cas9 분자에 의해 영향을 받는다. 이러한 실시형태는 2개의 gRNA를 필요로 하며, 각각의 단일 가닥 단절의 배치를 위한 것이다. 실시형태에서, 니카제 활성을 갖는 Cas9 분자는 gRNA가 혼성화되는 가닥을 절단 하지만, gRNA가 혼성화되는 가닥에 상보적인 가닥은 절단하지 않는다. 실시형태에서, 니카제 활성을 갖는 Cas9 분자는 gRNA가 혼성화되는 가닥을 절단하지 않지만, 오히려 gRNA가 혼성화되는 가닥에 상보적인 가닥을 절단한다.

[0741] 실시형태에서, 니카제는 HNH 활성을 갖는다, 예를 들어 불활성화된 RuvC 활성을 갖는 Cas9분자, 예를 들어 D10A 돌연변이와 같은 D10에서의 돌연변이를 갖는 Cas9 분자. D10A는 RuvC를 불활성화시키며; 따라서, Cas9 니카제는 (오직) HNH 활성을 가지며, gRNA가 혼성화되는 가닥(예를 들어 NGG PAM을 갖지 않는, 상보적 가닥)을 절단할 것이다. 다른 실시형태에서, H840, 예를 들어 H840A 돌연변이를 갖는 Cas9 분자는 니카제로서 사용될 수 있다. H840A는 HNH를 불활성화시키며; 따라서, Cas9 니카제는(오직) RuvC 활성을 가지며 비 상보적 가닥(예를 들어

NGG PAM을 갖고 그 서열이 gRNA와 동일한 가닥)을 절단한다.

[0742] 니카제 및 2개의 gRNA가 2개의 단일 가닥 닉을 위치시키는 데 사용되는 실시형태에서, 하나의 닉은 표적 핵산의 + 가닥 상에 있고 하나의 닉은 - 가닥 상에 있다. PAM은 바깥 쪽을 향하고 있다. gRNA가 약 0 내지 50부터, 0 내지 100 또는 0 내지 200 뉴클레오타이드에 의해 분리되도록 gRNA는 선택될 수 있다. 실시형태에서, 2개의 gRNA의 표적화 도메인에 상보적인 표적 서열 간에 중첩이 없다. 실시형태에서, gRNA는 중첩되지 않으며 50, 100 또는 200 뉴클레오타이드만큼에 의해 분리된다. 실시형태에서, 2개의 gRNA의 사용은 예를 들어 표적-외 결합을 감소시킴으로써 특이성을 증가시킬 수 있다(Ran el al., CELL 2013).

[0743] 실시형태에서, 단일 닉은 HDR을 유도하기 위해 사용될 수 있다. 단일 닉은 주어진 절단 부위에서 HDR, HR 또는 NHEJ의 비율을 증가시키는 데 사용될 수 있음이 본원에서 고려된다.

[0744] 표적 위치에 관련하여 이중 가닥 절단부 또는 단일 가닥 단절의 배치

[0745] 가닥 중 하나에서 이중 가닥 절단부 또는 단일 가닥 단절은 그러한 교정이 발생하도록 표적 위치에 충분히 가까워야 한다. 실시형태에서, 거리는 50, 100, 200, 300, 350 또는 400 뉴클레오타이드 이하이다. 이론에 얽매이기를 원하지 않지만, 단절은 말단 절제 동안의 엑소뉴클레아제-매개된 제거의 대상체가 되는 영역 내에 있도록 표적 위치에 충분히 가까워야 한다고 여겨진다. 표적 위치와 단절 사이의 거리가 너무 큰 경우, 돌연변이는 말단 절제에 포함되지 못할 수 있으며, 따라서, 공여자 서열이 말단 절제 영역 내의 서열을 교정하는 데에만 사용될 수 있기 때문에, 돌연변이는 교정되지 않을 수 있다.

[0746] gRNA(단분자(또는 키메라) 또는 모듈식 gRNA) 및 Cas9 뉴클레아제가 HDR- 또는 HR-매개 교정을 유도할 목적으로 이중 가닥 절단부을 유도하는 실시형태에서, 절단 부위는 표적 위치로부터 0 내지 200 bp 사이(예를 들어 0 내지 175, 0 내지 150, 0 내지 125, 0 내지 100, 0 내지 75, 0 내지 50, 0 내지 25, 25 내지 200, 25 내지 175, 25 내지 150, 25 내지 125, 25 내지 100, 25 내지 75, 25 내지 50, 50 내지 200, 50 내지 175, 50 내지 150, 50 내지 125, 50 내지 100, 50 내지 75, 75 내지 200, 75 내지 175, 75 내지 150, 75 내지 125, 75 내지 100 bp) 떨어져 있다. 실시형태에서, 절단 부위는 표적 위치로부터 0 내지 100 bp 사이(예를 들어 0 내지 75, 0 내지 50, 0 내지 25, 25 내지 100, 25 내지 75, 25 내지 50, 50 내지 100, 50 내지 75 또는 75 내지 100 bp) 떨어져 있다.

[0747] Cas9 니카제와 복합체를 형성하는 2개의 gRNA(독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA)가 HDR-매개 교정을 유도할 목적으로 2개의 단일 가닥 단절을 유도하는 실시형태에서, 더 가까운 닉은 표적 위치로부터 0 내지 200 bp 사이(예를 들어 0 내지 175, 0 내지 150, 0 내지 125, 0 내지 100, 0 내지 75, 0 내지 50, 0 내지 25, 25 내지 200, 25 내지 175, 25 내지 150, 25 내지 125, 25 내지 100, 25 내지 75, 25 내지 50, 50 내지 200, 50 내지 175, 50 내지 150, 50 내지 125, 50 내지 100, 50 내지 75, 75 내지 200, 75 내지 175, 75 내지 150, 75 내지 125, 75 내지 100 bp) 떨어져 있고, 2개의 닉은 이상적으로 서로의 25 내지 55 bp 내(예를 들어 25 내지 50, 25 내지 45, 25 내지 40, 25 내지 35, 25 내지 30, 30 내지 55, 30 내지 50, 30 내지 45, 30 내지 40, 30 내지 35, 35 내지 55, 35 내지 50, 35 내지 45, 35 내지 40, 40 내지 55, 40 내지 50, 40 내지 45bp)에 있을 것이며, 서로로부터 100 bp 이하 떨어져 있을(예를 들어 서로로부터 90, 80, 70, 60, 50, 40, 30, 20, 10 또는 5 bp 이하로 떨어져 있을) 것이다. 실시형태에서, 절단 부위는 표적 위치로부터 0 내지 100 bp 사이(예를 들어 0 내지 75, 0 내지 50, 0 내지 25, 25 내지 100, 25 내지 75, 25 내지 50, 50 내지 100, 50 내지 75 또는 75 내지 100 bp) 떨어져 있다.

[0748] 일 실시형태에서, 2개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 표적 위치의 양 편 상에 이중 가닥 절단부을 위치시키도록 배열된다. 대안적인 실시형태에서, 3개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 이중 가닥 절단부(즉, Cas9 뉴클레아제와 복합체 형성하는 하나의 gRNA) 및 2개의 단일 가닥 단절 또는 쌍을 이루는 단일 가닥 단절(즉, Cas9 니카제와 복합체 형성하는 2개의 gRNA)을 표적 위치의 각 편 상에 위치시키도록 배열된다(예를 들어 제1 gRNA는 표적 위치의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 표적 위치의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 또 다른 실시형태에서, 4개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 표적 위치의 각 편 상에 2 쌍의 단일 가닥 단절(즉, Cas9 니카제와 복합체 형성하는 2개의 gRNA의 2 쌍)을 생성하도록 배열된다(예를 들어 제1 gRNA는 표적 위치의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 표적 위치의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 이중 가닥 절단부(들) 또는 쌍에서 2개의 단일 가닥 닉 중 더 가까운 것은 이상적으로 표적 위치의 0 내지 500 bp 이내(예를 들어 표적 위치로부터 450, 400, 350, 300, 250, 200, 150, 100, 50 또는 25 bp이하)에 있을 것이다. 니카제가 사용될 때, 쌍에서 2개의 닉은 서로의 25 내지 55 bp

이내(예를 들어 25 내지 50, 25 내지 45, 25 내지 40, 25 내지 35, 25 내지 30, 50 내지 55, 45 내지 55, 40 내지 55, 35 내지 55, 30 내지 55, 30 내지 50, 35 내지 50, 40 내지 50, 45 내지 50, 35 내지 45, 또는 40 내지 45 bp 사이)에 있으며, 서로로부터 100 bp 이하(예를 들어 90, 80, 70, 60, 50, 40, 30, 20 또는 10 bp 이하) 떨어져 있다.

[0749] 일 실시형태에서, 2개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 표적 위치의 양 편 상에 이중 가닥 절단부을 위치시키도록 배열된다. 대안적인 실시형태에서, 3개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 이중 가닥 절단부(즉, Cas9 뉴클레아제와 복합체 형성하는 하나의 gRNA) 및 2개의 단일 가닥 단절 또는 쌍을 이루는 단일 가닥 단절(즉, Cas9 니카제와 복합체 형성하는 2개의 gRNA)을 2개의 표적 서열 상에 위치시키도록 배열된다(예를 들어 제1 gRNA는 표적 서열의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 삽입 부위의 표적 서열의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 또 다른 실시형태에서, 4개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 삽입 부위의 각 편 상에 2 쌍의 단일 가닥 단절(즉, Cas9 니카제와 복합체 형성하는 2개의 gRNA의 2 쌍)을 생성하도록 배열된다(예를 들어 제1 gRNA는 본원에 기재된 표적 서열의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 본원에 기재된 표적 서열의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 이중 가닥 절단부(들) 또는 쌍에서 2개의 단일 가닥 닉 중 더 가까운 것은 이상적으로 표적 위치의 0 내지 500 bp 이내(예를 들어 표적 위치로부터 450, 400, 350, 300, 250, 200, 150, 100, 50 또는 25 bp이하)에 있을 것이다. 니카제가 사용될 때, 쌍에서 2개의 닉은 서로의 25 내지 55 bp 이내(예를 들어 25 내지 50, 25 내지 45, 25 내지 40, 25 내지 35, 25 내지 30, 50 내지 55, 45 내지 55, 40 내지 55, 35 내지 55, 30 내지 55, 30 내지 50, 35 내지 50, 40 내지 50, 45 내지 50, 35 내지 45, 또는 40 내지 45 bp 사이)에 있으며, 서로로부터 100 bp 이하(예를 들어 90, 80, 70, 60, 50, 40, 30, 20 또는 10 bp 이하) 떨어져 있다.

[0750] 상동성 아암(arm)의 길이

[0751] 상동성 아암은 예를 들어 절제된 단일 가닥 돌출부가 공여자 주형 내에서 상보적 영역을 찾을 수 있도록 하기 위해서, 적어도 말단 절제가 발생할 수 있는 영역까지 연장되어야 한다. 전체 길이는 플라스미드 크기 또는 바이러스 패키징 한계와 같은 매개변수에 의해 제한될 수 있다. 실시형태에서, 상동성 아암은 반복 요소, 예를 들어 ALU 반복, LINE 반복 내로 연장되지 않는다. 주형은 동일하거나 상이한 길이의 2개의 상동성 아암을 가질 수 있다.

[0752] 예시적인 상동성 아암 길이는 적어도 25, 50, 100, 250, 500, 750 또는 1000 뉴클레오타이드를 포함한다.

[0753] 본원에 사용된 표적 위치는 Cas9 분자-의존적 과정에 의해 변형되는 표적 핵산(예를 들어 염색체)상의 부위를 지칭한다. 예를 들어 표적 위치는 표적 핵산의 변형된 Cas9 분자 절단 및 표적 위치의 주형 핵산 지시된 변형, 예를 들어 교정일 수 있다. 실시형태에서, 표적 위치는 하나 이상의 뉴클레오타이드가 첨가되는 표적 핵산 상의 2개의 뉴클레오타이드, 예를 들어 인접한 뉴클레오타이드 사이의 부위일 수 있다. 표적 위치는 주형 핵산에 의해 변경(예를 들어 교정) 되는 하나 이상의 뉴클레오타이드를 포함할 수 있다. 실시형태에서, 표적 위치는 표적 서열(예를 들어 gRNA가 결합하는 서열) 내에 있다. 실시형태에서, 표적 위치는 표적 서열(예를 들어 gRNA가 결합하는 서열)의 업스트림 또는 다운스트림이다.

[0754] 전형적으로, 주형 서열은 표적 서열과 단절 매개된 또는 촉매된 재조합을 겪는다. 실시형태에서, 주형 핵산은 Cas9 매개된 절단 사건에 의해 절단되는 표적 서열상의 부위에 상응하는 서열을 포함한다. 실시형태에서, 주형 핵산은 제1 Cas9 매개 사건에서 절단되는 표적 서열 상의 제1 부위 및 제2 Cas9 매개 사건에서 절단되는 표적 서열 상의 제2 부위 둘 모두에 상응하는 서열을 포함한다.

[0755] 실시형태에서, 주형 핵산은 번역된 서열의 코딩 서열에서의 변경을 야기하는 서열, 예를 들어 돌연변이 대립유전자를 야생형 대립유전자로 형질전환하는, 야생형 대립유전자를 돌연변이 대립유전자로 형질전환하는, 및/또는 정지 코돈을 도입하는 것과 같은, 단백질 산물에서 하나의 아미노산의 또 다른 것으로의 치환, 아미노산 잔기의 삽입, 아미노산 잔기의 결실, 또는 넌센스 돌연변이를 야기하는 것을 포함할 수 있다.

[0756] 다른 실시형태에서, 주형 핵산은 비-코딩 서열에서의 변경, 예를 들어 엑손에서의 또는 5'또는 3' 비-번역 또는 비-전사 영역에서의 변경을 야기하는 서열을 포함할 수 있다. 이러한 변경은 제어 요소, 예를 들어 프로모터, 인핸서에서의 변경, 및 시스-작용 또는 트랜스-작용 제어 요소에서의 변경을 포함한다.

[0757] 주형 핵산은 통합될 때:

[0758] 양성 제어 요소의 활성 감소;

[0759] 양성 제어 요소의 활성 증가;

[0760] 음성 제어 요소의 활성 감소;

[0761] 음성 제어 요소의 활성 증가;

[0762] 유전자의 발현 감소;

[0763] 유전자의 발현 증가;

[0764] 장애 또는 질병에 대한 내성 증가;

[0765] 바이러스 진입에 대한 내성 증가;

[0766] 돌연변이 교정 또는 원하지 않는 아미노산 잔기 변경;

[0767] 유전자 산물의 생물학적 특성 부여, 증가, 파괴 또는 감소, 예를 들어 효소의 효소 활성 증가, 또는 유전자 산물의 또 다른 분자와 상호작용하는 능력 증가; 를 야기하는 서열을 포함할 수 있다.

[0768] 주형 핵산은:

[0769] 표적 서열의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 이상 뉴클레오타이드의 서열에서의 변화를 야기하는 서열을 포함할 수 있다.

[0770] 실시형태에서, 주형 핵산은 20+/- 10, 30+/- 10, 40+/- 10, 50+/- 10, 60+/- 10, 70+/- 10, 80+/- 10, 90+/- 10, 100+/- 10, 1 10+/- 10, 120+/- 10, 130+/- 10, 140+/- 10, 150+/- 10, 160+/- 10, 170+/- 10, 1 80+/- 10, 190+/- 10, 200+/- 10, 210+/-10, 220+/- 10, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 2000, 2000 내지 3000 또는 3000을 초과하는 뉴클레오타이드 길이이다.

[0771] 주형 핵산은 다음의 성분을 포함한다:

[0772] [5' 상동성 아암]-[삽입 서열]-[3' 상동성 아암].

[0773] 상동성 아암은 염색체 내로의 재조합을 대비하며, 이는 원하지 않는 요소, 예를 들어 돌연변이 또는 특징을 대체 서열로 대체할 수 있다. 실시형태에서, 상동성 아암은 가장 먼 절단 부위의 옆에 배치된다.

[0774] 실시형태에서, 5' 상동성 아암의 3' 말단은 대체 서열의 5' 말단의 바로 옆 위치이다. 실시형태에서, 5' 상동성 아암은 대체 서열의 5' 말단으로부터 5'로 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 180, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 또는 2000 뉴클레오타이드 연장될 수 있다.

[0775] 실시형태에서, 3' 상동성 아암의 5' 말단은 대체 서열의 3' 말단의 바로 옆 위치이다. 실시형태에서, 3' 상동성 아암은 대체 서열의 3' 말단으로부터 3'로 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 180, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 또는 2000 뉴클레오타이드 연장될 수 있다.

[0776] 본원에서는 일정 서열 반복 요소, 예를 들어 Alu 반복, LINE 요소를 포함하는 것을 피하기 위하여 하나 또는 두 개의 상동성 아암이 단축될 수 있음이 고려된다. 예를 들어 5' 상동성 아암은 서열 반복 요소를 피하기 위하여 단축될 수 있다. 다른 실시형태에서, 3' 상동성 아암은 서열 반복 요소를 피하기 위하여 단축될 수 있다. 일부 실시형태에서, 5' 및 3' 상동성 아암 둘 모두는 일정 서열 반복 요소를 포함하는 것을 피하기 위하여 단축될 수 있다.

[0777] 본원에서는 돌연변이를 교정하기 위한 주형 핵산은 단일 가닥 올리고 뉴클레오타이드(ssODN)로서의 사용을 위해 설계될 수 있음이 고려된다. ssODN을 사용할 때, 5' 및 3' 상동성 아암은 약 200 염기쌍(bp) 길이까지의, 예를 들어 적어도 25, 50, 75, 100, 125, 150, 175 또는 200 bp 길이의 범위일 수 있다. 올리고뉴클레오타이드 합성에서의 개선이 계속됨에 따라 ssODN을 위해 더 긴 상동성 아암 또한 고려된다.

[0778] 유전자 표적화를 위한 NHEJ 접근법

[0779] 본원에 기재된 바와 같이, 뉴클레아제-유도된 비-상동 말단-접합(NHEJ)은 유전자 특이적 넉아웃을 표적화하는 데 사용될 수 있다. 뉴클레아제-유도된 NHEJ는 관심 유전자에서 서열을 제거(예를 들어 결실)하는 데 또한 사용될 수 있다.

[0780] 이론에 얽매이기를 원하지 않지만, 실시형태에서, 본원에 기재된 방법과 관련된 게놈 변경은 뉴클레아제-유도된 NHEJ 및 NHEJ 수선 경로의 오류-빈발 성질에 의존한다. NHEJ는 2개의 말단을 함께 접합함으로써 DNA의 이중 가닥 절단부을 수선한다; 그러나, 일반적으로, 원래의 서열은 2개의 맞는 말단이, 이중 가닥 절단부에 의해 형성되었던 것과 정확히 같이, 완벽하게 결찰되는 경우에만 복구된다. 이중 가닥 절단부의 DNA 말단은 빈번하게, 말단이 재접합하기 전에, 하나 또는 양 가닥에서, 뉴클레오타이드의 첨가 또는 제거를 야기하는 효소적 가공의 대상체가 된다. 이는 NHEJ 수선 부위의 DNA 서열에 삽입 및/또는 결실(삽입결실) 돌연변이의 존재를 야기한다. 이들 돌연변이의 3분의 2는 리딩 프레임을 변경하여, 비-기능성 단백질을 생산할 수 있다. 추가적으로, 리딩 프레임을 유지하지만, 상당량의 서열을 삽입 또는 결실하는 돌연변이는 단백질의 기능성을 파괴할 수 있다. 중대한 기능성 도메인에서의 돌연변이는 단백질의 중대하지 않은 영역에서의 돌연변이보다 덜 견딜 수 있을 가능성이 있기 때문에, 이는 유전자좌 의존적이다.

[0781] NHEJ에 의해 생성된 삽입결실 돌연변이는 사실상 예측할 수 없다; 그러나 주어진 단절 부위에서 일정 삽입결실 서열이 유리하여 집단 내에서 과도하게 나타난다. 결실 길이는 폭넓게 달라질 수 있으며; 가장 통상적으로 1 내지 50 bp 범위이지만, 쉽게 100 내지 200 bp 초과에 도달할 수 있다. 삽입은 더 짧은 경향이 있고 종종 단절 부위 바로 옆을 둘러싸는 서열의 짧은 중복을 포함한다. 그러나, 대형 삽입을 수득하는 것이 가능하며, 이들 경우에, 삽입된 서열은 종종 세포 내에 존재하는 게놈의 다른 영역 또는 플라스미드 DNA임이 밝혀졌다.

[0782] NHEJ는 돌연변이 유발 과정이기 때문에, 특정 최종 서열의 생성이 필요하지 않는 한 작은 서열 모티프를 결실하는 데 또한 사용될 수 있다. 이중 가닥 절단부이 짧은 표적 서열 부근으로 표적화되는 경우, NHEJ 수선에 의해 야기된 결실 돌연변이는 종종 원하지 않는 뉴클레오타이드를 포괄하고, 따라서 제거한다. 더 큰 DNA 부분을 결실하기 위해, 서열의 각 편 상에 하나씩 2개의 이중 가닥 절단부을 도입하는 것은, 말단들 사이에 중간 서열 전체가 제거되는 NHEJ를 야기할 수 있다. 이들 접근법은 둘 모두 특정 DNA 서열을 결실하는 데 사용될 수 있다; 그러나, NHEJ의 오류-빈발 성질은 여전히 수선 부위에 삽입결실 돌연변이를 생산할 수 있다.

[0783] 이중 가닥 절단하는 Cas9 분자 및 단일 가닥, 또는 니카제, Cas9 분자 둘 모두는 본원에 기재된 방법 및 조성물에서 NHEJ-매개 삽입결실을 생성하기 위하여 사용될 수 있다. 유전자, 예를 들어 코딩 영역, 예를 들어 관심 유전자의 초기 코딩 영역으로 표적화된 NHEJ-매개 삽입결실은 관심 유전자를 넉아웃(즉, 발현 제거)하기 위하여 사용될 수 있다. 예를 들어 관심 유전자의 초기 코딩 영역은, 전사 개시 부위를 바로 뒤따르는, 코딩 서열의 제 1 엑손 내의, 또는 전사 개시 부위의 500 bp 내(예를 들어 500, 450, 400, 350, 300, 250, 200, 150, 100 또는 50 bp 미만)의 서열을 포함한다.

[0784] 표적 위치에 관련하여 이중 가닥 또는 단일 가닥 단절의 배치

[0785] gRNA 및 Cas9 뉴클레아제가 NHEJ-매개 삽입결실을 유도할 목적으로 이중 가닥 절단부을 생성하는 실시형태에서, gRNA, 예를 들어 단분자(또는 키메라) 또는 모듈식 gRNA 분자는 하나의 이중 가닥 절단부을 표적 위치의 뉴클레오타이드에 근접하게 위치시키도록 배열된다. 실시형태에서, 절단 부위는 표적 위치로부터 0 내지 500 bp 사이(예를 들어 표적 위치로부터 500, 400, 300, 200, 100, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1 bp 미만) 떨어져있다.

[0786] Cas9 니카제와 복합체 형성하는 2개의 gRNA가 NHEJ-매개 삽입결실을 유도할 목적으로 2개의 단일 가닥 단절을 유도하는 실시형태에서, 2개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는, 표적 위치의 뉴클레오타이드의 NHEJ 수선을 대비하기 위해 2개의 단일 가닥 단절을 위치시키도록 배열된다. 실시형태에서, gRNA들은 본질적으로 이중 가닥 절단부을 모방 하여, 상이한 가닥 상의 동일한 위치 또는 서로의 수 뉴클레오타이드 내에 절단을 위치시키도록 배열된다. 실시형태에서, 더 가까운 닉은 표적 위치로부터 0 내지 30 bp 사이(예를 들어 표적 위치로부터 30, 25, 20, 1, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1 bp 미만) 떨어져있고, 두 개의 닉은 서로의 25 내지 55 bp 내(예를 들어 25 내지 50, 25 내지 45, 25 내지 40, 25 내지 35, 25 내지 30, 50 내지 55, 45 내지 55, 40 내지 55, 35 내지 55, 30 내지 55, 30 내지 50, 35 내지 50, 40 내지 50, 45 내지 50, 35 내지 45, 또는 40 내지 45 bp 사이)에 있으며, 서로로부터 100 bp 이하(예를 들어 90, 80, 70, 60, 50, 40, 30, 20 또는 10 bp 이하) 떨어져 있다. 실시형태에서, gRNA들은 표적 위치의 뉴클레오타이드의 각 편 상에 단일 가닥 단절을 배치하도록 배열된다.

[0787] 이중 가닥 절단하는 Cas9 분자 및 단일 가닥, 또는 니카제, Cas9 분자 둘 모두는 본원에 기재된 방법 및 조성물에서 표적 위치의 양 편에 단절을 생성하기 위하여 사용될 수 있다. 2개의 절단 사이의 핵산 서열을 제거하기 위하여 이중 가닥 또는 쌍을 이루는 단일 가닥 단절은 표적 위치의 양 편 상에 생성될 수 있다(예를 들어 2개의 단절 사이의 영역이 결실됨). 일 실시형태에서, 2개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는

모듈식 gRNA는 표적 위치의 양 편 상에 이중 가닥 절단부을 위치시키도록 배열된다(예를 들어 제1 gRNA는 본원에 기재된 유전자 또는 경로의 돌연변이의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 본원에 기재된 유전자 또는 경로의 돌연변이의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 대안적인 실시형태에서, 3개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 이중 가닥 절단부(즉, Cas9 뉴클레아제와 복합체 형성하는 하나의 gRNA) 및 2개의 단일 가닥 단절 또는 쌍을 이루는 단일 가닥 단절(즉, Cas9 니카제와 복합체 형성하는 2개의 gRNA)을 표적 위치의 각 편 상에 위치시키도록 배열된다(예를 들어 제1 gRNA는 본원에 기재된 유전자 또는 경로의 돌연변이의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 본원에 기재된 유전자 또는 경로의 돌연변이의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 또 다른 실시형태에서, 4개의 gRNA, 예를 들어 독립적으로, 단분자(또는 키메라) 또는 모듈식 gRNA는 표적 위치의 각 편 상에 2 쌍의 단일 가닥 단절(즉, Cas9 니카제와 복합체 형성하는 2개의 gRNA의 2 쌍)을 생성하도록 배열된다(예를 들어 제1 gRNA는 본원에 기재된 유전자 또는 경로의 돌연변이의 업스트림(즉, 5')을 표적화하는 데 사용되며 제2 gRNA는 본원에 기재된 유전자 또는 경로의 돌연변이의 다운스트림(즉, 3')을 표적화하는 데 사용된다). 이중 가닥 절단부(들) 또는 쌍에서 2개의 단일 가닥 닉 중 더 가까운 것은 이상적으로 표적 위치의 0 내지 500 bp 이내(예를 들어 표적 위치로부터 450, 400, 350, 300, 250, 200, 150, 100, 50 또는 25 bp이하)에 있을 것이다. 니카제가 사용될 때, 쌍에서 2개의 닉은 서로의 25 내지 55 bp 이내(예를 들어 25 내지 50, 25 내지 45, 25 내지 40, 25 내지 35, 25 내지 30, 50 내지 55, 45 내지 55, 40 내지 55, 35 내지 55, 30 내지 55, 30 내지 50, 35 내지 50, 40 내지 50, 45 내지 50, 35 내지 45, 또는 40 내지 45 bp 사이)에 있으며, 서로로부터 100 bp 이하(예를 들어 90, 80, 70, 60, 50, 40, 30, 20 또는 10 bp 이하) 떨어져 있다.

[0788] 다른 실시형태에서, 주형 핵산의 삽입은 미세상동 말단 접합(MMEJ)에 의해 매개될 수 있다. 예를 들어 2015년 12월 17일 온라인으로 발행되고, 그 내용의 전문이 참조로 포함되는 문헌[Saksuma et al., "MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems." *Nature Protocols* **11**, 118-133(2016) doi:10.1038/nprot.2015.140]을 참조한다.

[0789] **VIII. 하나를 초과하는 gRNA 분자를 포함하는 시스템**

[0790] 이론에 얽매이기를 의도하지 않지만, 놀랍게도 연속적 핵산 상에 근접하게 위치된 2개의 표적 서열의 표적화(예를 들어 각각의 표적 서열에서 또는 그 부근에서 단일- 또는 이중-가닥 단절을 각각 유도하는 2개의 gRNA 분자/Cas9 분자 복합체에 의하여)는 2개의 표적 서열 사이에 위치된 핵산 서열(또는 핵산 서열의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%)의 절제(예를 들어 결실)를 유도한다는 것이 본원에 나타났다. 일부 양태에서, 본 개시는 연속적 핵산, 예를 들어 염색체, 예를 들어 그의 인트론, 엑손 및 조절 요소를 포함하는 유전자 또는 유전자좌 상에 근접한 표적 서열들을 표적화하는 표적화 도메인들을 포함하는 2개 이상의 gRNA 분자의 사용을 대비한다. 사용은 예를 들어 하나 이상의 Cas9 분자(또는 2개 이상의 gRNA 분자 및/또는 하나 이상의 Cas9 분자를 인코딩하는 핵산)와 함께 2개 이상의 gRNA 분자를 세포 내로 도입하는 것에 의할 수 있다.

[0791] 일부 양태에서, 2개 이상의 gRNA 분자의 표적 서열은 연속 핵산 상에서 적어도 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10,000, 11,000, 12,000, 13,000, 14,000 또는 15,000개 뉴클레오타이드 이격되어 위치하나, 연속 핵산 상에서 25,000개 이하 뉴클레오타이드 이격되어 위치한다. 실시형태에서, 표적 서열은 약 4000 내지 약 6000개 뉴클레오타이드 이격되어 위치한다. 일 실시형태에서, 표적 서열은 약 4000개 뉴클레오타이드 이격되어 위치한다. 일 실시형태에서, 표적 서열은 약 5000개 뉴클레오타이드 이격되어 위치한다. 일 실시형태에서, 표적 서열은 약 6000개 뉴클레오타이드 이격되어 위치한다.

[0792] 일부 양태에서, 복수의 gRNA 분자는 각각 동일한 유전자 또는 유전자의 유전자좌 내의 서열을 표적화한다. 예시적인 일 양태에서, 2개 이상의 gRNA 분자의 표적 서열은 HBG1 프로모터 영역에 위치한다. 예시적인 일 양태에서, 2개 이상의 gRNA 분자의 표적 서열은 HBG2 프로모터 영역에 위치한다. 또 다른 양태에서, 복수의 gRNA 분자는 각각 2개 이상의 상이한 유전자 또는 유전자의 유전자좌 내의 서열을 표적화한다. 예시적인 일 양태에서, 하나 이상의 gRNA 분자의 표적 서열은 HBG1 프로모터 영역에 위치하고, 하나 이상의 다른 gRNA 분자의 표적 서열은 HBG2 프로모터 영역에 위치한다.

[0793] 일부 양태에서, 본 발명은 복수의, 예를 들어 2개 이상의, 예를 들어 2개의 본 발명의 gRNA 분자를 포함하는 조성물 및 세포를 제공하며, 여기서, 복수의 gRNA 분자는 15,000개 미만, 14,000개 미만, 13,000개 미만, 12,000개 미만, 11,000개 미만, 10,000개 미만, 9,000개 미만, 8,000개 미만, 7,000개 미만, 6,000개 미만, 5,000개 미만, 4,000개 미만, 3,000개 미만, 2,000개 미만, 1,000개 미만, 900개 미만, 800개 미만, 700개 미만, 600개

미만, 500개 미만, 400개 미만, 300개 미만, 200개 미만, 100개 미만, 90개 미만, 80개 미만, 70개 미만, 60개 미만, 50개 미만, 40개 미만 또는 30개 미만의 뉴클레오타이드 이격된 서열을 표적화한다. 일 실시형태에서, 표적 서열은 듀플렉스 핵산의 동일한 가닥 상에 존재한다. 일 실시형태에서, 표적 서열은 듀플렉스 핵산의 상이한 가닥 상에 존재한다.

[0794] 일 실시형태에서, 본 발명은 듀플렉스 핵산의 동일한 또는 상이한 가닥 상에서 25,000개 미만, 20,000개 미만, 15,000개 미만, 14,000개 미만, 13,000개 미만, 12,000개 미만, 11,000개 미만, 10,000개 미만, 9,000개 미만, 8,000개 미만, 7,000개 미만, 6,000개 미만, 5,000개 미만, 4,000개 미만, 3,000개 미만, 2,000개 미만, 1,000개 미만, 900개 미만, 800개 미만, 700개 미만, 600개 미만, 500개 미만, 400개 미만, 300개 미만, 200개 미만, 100개 미만, 90개 미만, 80개 미만, 70개 미만, 60개 미만, 50개 미만, 40개 미만 또는 30개 미만의 뉴클레오타이드가 이격되어 배치된 2개의 gRNA 결합 부위 사이에 배치된 핵산을 절제하는(예를 들어 결실시키는) 방법을 제공한다. 일 실시형태에서, 이러한 방법은 각각의 gRNA 결합 부위와 연관된 PAM 부위 사이에 배치된 뉴클레오타이드 중 50% 초과, 60% 초과, 70% 초과, 80% 초과, 85% 초과, 86% 초과, 87% 초과, 88% 초과, 89% 초과, 90% 초과, 91% 초과, 92% 초과, 93% 초과, 94% 초과, 95% 초과, 96% 초과, 97% 초과, 98% 초과, 99% 초과 또는 100%의 결실을 제공한다. 실시형태에서, 이러한 결실은 각각의 gRNA 결합 부위와 연관된 하나 이상의 PAM 부위 내에 하나 이상의 뉴클레오타이드를 추가로 포함한다. 실시형태에서, 이러한 결실은 또한, 각각의 gRNA 결합 부위와 연관된 하나 이상의 PAM 부위 사이의 영역의 외부에 하나 이상의 뉴클레오타이드를 포함한다.

[0795] 일 양태에서, 2개 이상의 gRNA 분자는 유전자 제어 요소, 예를 들어 프로모터 결합 부위, 인핸서 영역 또는 리프레서(repressor) 영역의 측면에 존재하는 표적 서열을 표적화하는 표적화 도메인을 포함하며, 따라서, 개입 서열(또는 이러한 개입 서열의 일부)의 절제는 관심 유전자의 상향조절 또는 하향조절을 유발한다. 다른 실시형태에서, 2개 이상의 gRNA 분자는 유전자의 측면에 존재하는 서열을 표적화하는 표적화 도메인을 포함하며, 따라서, 개입 서열(또는 이의 일부)의 절제는 관심 유전자의 결실을 유발한다.

[0796] 일 실시형태에서, 2개 이상의 gRNA 분자는 각각 표 1, 예를 들어 표 2, 또는 예를 들어 표 3a 또는 표 3b의 표적화 도메인을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다. 실시형태에서, 2개 이상의 gRNA 분자는 각각, 본원에 기재된 방법에 의해 생체외에서 gRNA에 의해 편집된 HSPC로부터 (예를 들어 편집 후 7일차에) 분화된 적혈구 집단에서 F 세포의 수의 적어도 15% 상향조절을 초래하는 상기 gRNA 분자의 표적화 도메인을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다. 양태에서, 2개 이상의 gRNA 분자는 동일한 유전자 또는 영역, 예를 들어 HBG1 프로모터 영역 또는 HBG2 프로모터 영역 내의 서열과 상보적인 표적화 도메인을 포함한다. 양태에서, 2개 이상의 gRNA 분자는 상이한 유전자 또는 영역의 서열, 예를 들어 HBG1 프로모터 영역에서 하나 또는 HBG2 프로모터 영역에서 하나와 상보적인 표적화 도메인을 포함한다.

[0797] 일 양태에서, 2개 이상의 gRNA 분자는 유전자 제어 요소, 예를 들어 프로모터 결합 부위, 인핸서 영역 또는 리프레서 영역의 측면에 존재하는 표적 서열을 표적화하는 표적화 도메인을 포함하며, 따라서, 개입 서열(또는 이러한 개입 서열의 일부)의 절제는 관심 유전자의 상향조절 또는 하향조절을 유발한다. 또 다른 실시형태에서, 2개 이상의 gRNA 분자는 유전자의 측면에 존재하는 서열을 표적화하는 표적화 도메인을 포함하며, 따라서, 개입 서열(또는 이의 일부)의 절제는 관심 유전자의 결실을 유발한다. 예를 들어, 2개 이상의 gRNA 분자는 HBG1 유전자의 측면에 존재하는 표적 서열을 표적화하는 표적화 도메인을 포함하며(예를 들어 하나의 gRNA 분자는 HBG1 프로모터 영역 내의 표적 서열을 표적화하고, 제2 gRNA 분자는 HBG2 프로모터 영역 내의 표적 서열을 표적화함), 따라서, HBG1 유전자가 절제된다.

[0798] 일 실시형태에서, 2개 이상의 gRNA 분자는 표 1로부터 선택되는 표적화 도메인을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다.

[0799] 양태에서, 2개 이상의 gRNA 분자는 상기 표 2에 열거된 표적화 도메인을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다. 양태에서, 2개 이상의 gRNA 분자는 상기 표 3a에 열거된 gRNA의 표적화 도메인 서열을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다. 양태에서, 2개 이상의 gRNA 분자는 상기 표 3b에 열거된 표적화 도메인 서열을 포함하는, 예를 들어 이로 구성되는 표적화 도메인을 포함한다.

[0800] 비결실 HPFH 영역, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역 내의 서열을 표적화하는 표적화 도메인을 포함하는 gRNA 분자는 부가적으로, 예를 들어 BCL11a 인핸서 영역(예를 들어 +55, +58 또는 +62 BCL11a 인핸서 영역) 내의 표적 서열에 상보적인 표적화 도메인을 포함하는 gRNA 분자, 및/또는 결실 HPFH 유전자좌의 표적 서열에 상보적인 표적화 도메인을 포함하는 gRNA 분자와 함께 사용될 수 있다. 이러한 결실 HPFH 유전자좌는 당업계에 공지되어 있으며, 예를 들어 문헌[Sankaran VG et al. NEJM (2011) 365:807-814(그 전체가 참조로 본 명

세서에 포함됨)]에 기재된 것들이다.

[0801] **IX. gRNA의 특성**

[0802] 본원에서 추가로 놀랍게도, 단일 sgRNA 분자가 1개 초과의 유전자좌(예를 들어 높은 서열 상동성을 갖는 유전자좌)에서 표적 서열을 가질 수 있고, 이러한 유전자좌가 동일한 염색체 상에, 예를 들어 약 15,000개 미만의 뉴클레오타이드, 약 14,000개 미만의 뉴클레오타이드, 약 13,000개 미만의 뉴클레오타이드, 약 12,000개 미만의 뉴클레오타이드, 약 11,000개 미만의 뉴클레오타이드, 약 10,000개 미만의 뉴클레오타이드, 약 9,000개 미만의 뉴클레오타이드, 약 8,000개 미만의 뉴클레오타이드, 약 7,000개 미만의 뉴클레오타이드, 약 6,000개 미만의 뉴클레오타이드, 약 5,000개 미만의 뉴클레오타이드, 약 4,000개 미만의 뉴클레오타이드 또는 약 3,000개 미만의 뉴클레오타이드(예를 들어 약 4,000 내지 약 6,000개 뉴클레오타이드 이격되어) 내에 존재할 때, 이러한 gRNA 분자는 개입 서열(또는 이의 일부)의 절제를 초래하여, 이로써 유익한 효과, 예를 들어 변형된 HSPC로부터 분화된 적혈구계 세포(본원에 기재된 바와 같음)에서 태아 헤모글로빈의 상향조절을 초래할 수 있는 것으로 제시되었다. 따라서, 일 양태에서, 본 발명은 2개의 유전자좌에서 표적 서열을 갖는 gRNA 분자를, 예를 들어 동일한 염색체 상의, (예를 들어 표 1에 기재된 바와 같은) 예를 들어 HBG1 프로모터 영역 및 HBG2 프로모터 영역에 표적 서열을 갖는 유전자좌에 제공한다. 이론으로 결부시키고자 하는 것은 아니지만, 이러한 gRNA가 1개 초과의 위치에서(예를 들어 2개 영역 각각 내의 표적 서열에서) 게놈의 절단을 초래할 수 있고, 후속적인 수선(repair)이 개입 핵산 서열의 결실을 초래할 수 있는 것으로 여겨진다. 다시, 이론으로 결부시키고자 하는 것은 아니지만, 상기 개입 서열의 결실은 하나 이상의 단백질의 발현 또는 기능에 요망되는 효과를 가질 수 있다.

[0803] 추가로 놀랍게도, 하나의 유전자 또는 영역 내의 서열에서만 상보적인, 예를 들어 (HBG2 프로모터 영역이 아니라) HBG1 프로모터 영역 내의 표적 서열에 상보적이거나, 또는 (HBG1 프로모터 영역이 아니라) HBG2 프로모터 영역 내의 표적 서열에 상보적인 표적화 도메인을 포함하는 gRNA 분자는 (본원에 기재된 바와 같은) 변형된 HSPC로부터 분화된 적혈구계 세포에서 태아 헤모글로빈의 유의한 상향조절을 초래할 수 있는 것으로 제시되었다. 따라서 양태에서, 본 발명은, 단일 비결실 HPFH 영역, 예를 들어 (예를 들어 표 1에 기재된 바와 같이) HBG1 또는 HBG2 프로모터 영역 내의 표적 서열에 상보적이지만 임의의 다른 유전자 또는 영역에서 완전히 (예를 들어 100%) 상보적인 표적 서열 매치는 갖지 않는 표적화 도메인을 포함하는 gRNA 분자를 제공한다.

[0804] 본원에서 추가로 놀랍게도, gRNA 분자, 및 상기 gRNA 분자를 포함하는 CRISPR 시스템은 동일한 세포 유형, 전달 방법 및 crRNA/tracr 구성성분을 사용한 다수의 실험에 걸쳐 유사하거나 동일한 삽입결실 패턴을 생산하는 것으로 제시되었다. 이론으로 결부시키고자 하는 것은 아니지만, 일부 삽입결실 패턴은 다른 삽입결실 패턴보다 더 유리할 수 있는 것으로 여겨진다. 예를 들어, "프레임시프트 돌연변이"를 초래하는 삽입 및/또는 결실(예를 들어 1- 또는 2-염기쌍 삽입 또는 결실, 또는 n/3이 정수(whole number)가 아닌(n이 삽입 또는 결실에서 뉴클레오타이드의 수인) 임의의 삽입 또는 결실)을 주로 포함하는 삽입결실은 기능성 단백질의 발현을 감소시키거나 없애는 데 있어서 유익할 수 있는 것으로 여겨진다. 마찬가지로, "대형 결실"(예를 들어 본원에 기재된 바와 같이 예를 들어 gRNA에 대한 제1 결합 부위와 제2 결합 부위 사이에 배치된 서열을 포함하는, 10, 11, 12, 13, 14, 15, 20, 25 또는 30개 초과의 뉴클레오타이드, 예를 들어 1 kb 초과, 2 kb 초과, 3 kb 초과, 5 kb 초과 또는 10 kb 초과의 결실)을 주로 포함하는 삽입결실은 또한, 예를 들어 프로모터 결합 부위와 같은 중대한 조절 서열을 제거하는 데 있어서, 또는 기능성 단백질의 발현에 유사하게 효과를 가질 수 있는 유전자좌의 구조 또는 기능을 변경시키는 데 있어서 유익할 수 있다. 주어진 gRNA/CRISPR 시스템에 의해 유도된 삽입결실 패턴이 놀랍게도 주어진 세포 유형, gRNA 및 CRISPR 시스템에 대해 일관적으로 재생되는 것으로 확인되긴 하였지만, 본원에 기재된 바와 같이, 어떠한 단일 삽입결실 구조도 gRNA/CRISPR 시스템의 도입 시 주어진 세포에서 필연적으로 생산되지 않을 것이다.

[0805] 따라서, 본 발명은 유익한 삽입결실 패턴 또는 구조를 생성하는, 예를 들어 대형 결실로 주로 구성된 삽입결실 패턴 또는 구조를 갖는 gRNA 분자를 제공한다. 이러한 gRNA 분자는 본원에 기재된 바와 같이 시험 세포(예를 들어 HEK293 세포)에서 또는 관심 세포, 예를 들어 HSPC 세포에서 후보 gRNA 분자에 의해 생성된 삽입결실 패턴 또는 구조를 NGS에 의해 평가함으로서 선택될 수 있다. 실시예에 제시된 바와 같이, 요망되는 세포 집단 내에 도입 시, gRNA의 표적 서열에 또는 부근에 대형 결실을 갖는 세포의 유의한 분획을 포함하는 세포 집단을 초래하는 gRNA 분자가 발견되었다. 일부 경우, 대형 결실 삽입결실의 형성률은 75%, 80%, 85%, 90% 이상만큼 높다. 따라서, 본 발명은 본원에 기재된 gRNA 분자의 표적 부위에 또는 부근에 예를 들어 본원에 기재된 바와 같은 대형 결실을 갖는 세포를 적어도 약 40%(예를 들어 적어도 약 45%, 적어도 약 50%, 적어도 약 55%, 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95% 또는 적어도 약 99%) 포함하는 세포 집단을 제공한다. 본 발명은 또한, 본원에 기재된 gRNA 분자의 표적

부위에 또는 부근에 예를 들어 본원에 기재된 바와 같은 대형 결실을 갖는 세포를 적어도 약 50%(예를 들어 적어도 약 55%, 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95% 또는 적어도 약 99%) 포함하는 세포 집단을 제공한다.

[0806] 따라서, 본 발명은 본 발명의 치료 방법에 사용하기 위한 gRNA 분자를 선택하는 방법을 제공하며, 상기 방법은: 1) 복수의 gRNA 분자를 관심 표적에 제공하는 단계, 2) 상기 gRNA 분자의 사용에 의해 생성된 삽입결실 패턴 또는 구조를 평가하는 단계, 3) 프레임시프트 돌연변이, 대형 결실 또는 이들의 조합으로 주로 구성된 삽입결실 패턴 또는 구조를 형성하는 gRNA 분자를 선택하는 단계, 및 4) 상기 선택된 gRNA를 본 발명의 방법에 사용하는 단계를 포함한다.

[0807] 따라서, 본 발명은 본 발명의 치료 방법에 사용하기 위한 gRNA 분자를 선택하는 방법을 제공하며, 상기 방법은: 1) 복수의 gRNA 분자를, 예를 들어 1개 초과의 위치에서 표적 서열을 갖는 관심 표적을 제공하는 단계, 2) 상기 gRNA 분자의 사용에 의해 생성된 삽입결실 패턴 또는 구조를 평가하는 단계, 3) 예를 들어 상기 gRNA 분자에 노출된 세포 집단의 세포 중 적어도 약 25% 이상에서, 2개의 표적 서열 사이에 위치한 핵산 서열을 포함하는 서열의 절제를 형성하는 gRNA 분자를 선택하는 단계, 및 4) 상기 선택된 gRNA를 본 발명의 방법에 사용하는 단계를 포함한다.

[0808] 본 발명은 추가로, 세포를 변경시키는 방법, 및 변경된 세포를 제공하며, 여기서, 특정 삽입결실 패턴은 해당 세포 유형에서 주어진 gRNA/CRISPR 시스템을 이용하여 일관적으로 생산된다. 본원에 기재된 gRNA/CRISPR 시스템을 이용하여 관찰되는 가장 빈번하게 발생하는 상위 5개 삽입결실은 실시예의 방법을 사용하여 결정될 수 있고, 예를 들어 실시예에 개시된다. 실시예에 제시된 바와 같이, 세포 집단이 발생되며, 여기서, 유의한 분획의 세포(예를 들어 세포 집단)는 상위 5개 삽입결실 중 하나를 포함하며, 여기서, 상위 5개 삽입결실 중 하나는 집단의 세포 중 30% 초과, 40% 초과, 50% 초과, 60% 이상에 존재한다. 따라서, 본 발명은 주어진 gRNA/CRISPR 시스템을 이용하여 관찰된 상위 5개 삽입결실 중 임의의 하나의 삽입결실을 포함하는 세포, 예를 들어 (본원에 기재된 바와 같은) HSPC를 제공한다. 추가로, 본 발명은 예를 들어 NGS에 의해 평가 시, 주어진 gRNA/CRISPR 시스템에 대해 본원에 기재된 상위 5개 삽입결실 중 하나를 포함하는 세포를 높은 퍼센트로 포함하는 세포, 예를 들어 (본원에 기재된 바와 같은) HSPC의 집단을 제공한다. 삽입결실 패턴 분석과 연관되어 사용 시, "높은 퍼센트"는 주어진 gRNA/CRISPR 시스템에 대해 본원에 기재된 상위 5개 삽입결실 중 하나를 포함하는 집단의 세포 중 적어도 약 50%(예를 들어 적어도 약 55%, 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95% 또는 적어도 약 99%)를 지칭한다. 다른 실시형태에서, 세포 집단은 주어진 gRNA/CRISPR 시스템에 대해 본원에 기재된 상위 5개 삽입결실 중 하나를 갖는 세포를 적어도 약 25%(예를 들어 약 25% 내지 약 60%, 예를 들어 약 25% 내지 약 50%, 예를 들어 약 25% 내지 약 40%, 예를 들어 약 25% 내지 약 35%) 포함한다.

[0809] 또한, 소정의 gRNA 분자가 표적 세포 유형의 게놈 내의 표적-외 서열(예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부의 표적-외 서열)에서 삽입결실을 생성하지 않거나, 표적 부위에서의 삽입결실 생성 빈도에 비해 표적-외 부위(HBG1 및/또는 HBG2 프로모터 영역의 외부의 표적-외 서열)에서 매우 낮은 빈도의(예를 들어 집단 내에서 5% 미만의 세포에서) 삽입결실을 생산하는 것으로 발견되었다. 따라서, 본 발명은, 표적 세포 유형에서 표적-외 삽입결실 형성을 나타내지 않거나, 5% 미만의 표적-외 삽입결실 형성 빈도, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 임의의 표적-외 부위에서 5% 미만의 빈도의 삽입결실을 생산하는 gRNA 분자 및 CRISPR 시스템을 제공한다. 실시형태에서, 본 발명은 표적 세포 유형에서 임의의 표적-외 삽입결실 형성을 나타내지 않는 gRNA 분자 및 CRISPR 시스템을 제공한다. 따라서, 본 발명은 추가로, 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 세포, 예를 들어 HSPC 집단을 제공하며, 이는 본원에 기재된 gRNA 분자의 표적 부위에 또는 부근에 삽입결실(예를 들어 프레임시프트 삽입결실, 또는 예를 들어 본원에 기재된 바와 같은 주어진 gRNA/CRISPR 시스템에 의해 생산된 상위 5개 삽입결실 중 임의의 하나)을 포함하지만, gRNA 분자의 임의의 표적-외 부위에서 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 임의의 표적-외 부위에서 삽입결실을 포함하지 않는다. 다른 실시형태에서, 본 발명은 추가로, 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 HSPC의 집단을 제공하며, 이는 본원에 기재된 gRNA 분자의 표적 부위에 또는 부근에 삽입결실(예를 들어 프레임시프트 삽입결실, 또는 예를 들어 본원에 기재된 바와 같은 주어진 gRNA/CRISPR 시스템에 의해 생산된 상위 5개 삽입결실 중 임의의 하나)을 갖는 세포를 적어도 20%, 예를 들어 적어도 30%, 예를 들어 적어도 40%, 예를 들어 적어도 50%, 예를 들어 적어도 60%, 예를 들어 적어도 70%, 예를 들어 적어도 75% 포함하지만, gRNA 분자의 임의의 표적-외 부위에서 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 임의의 표적-외 부위에서 삽입결실을 포함하는 세포를 5% 미만, 예를 들어 4% 미만, 3% 미만, 2% 미만 또는 1% 미만으로

포함한다. 다른 실시형태에서, 본 발명은 추가로, 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 HSPC의 집단을 제공하며, 이는 HBG1 및/또는 HBG2 프로모터 영역 내에(예를 들어 gRNA의 표적 서열에 적어도 90% 상동성인 서열에서 또는 그 부근에서) 삽입결실을 갖는 세포를 적어도 20%, 예를 들어 적어도 30%, 예를 들어 적어도 40%, 예를 들어 적어도 50%, 예를 들어 적어도 60%, 예를 들어 적어도 70%, 예를 들어 적어도 75%, 예를 들어 적어도 80%, 예를 들어 적어도 90%, 예를 들어 적어도 95% 포함하지만, HBG1 및/또는 HBG2 프로모터 영역의 외부에서 임의의 표적-외 부위에서 또는 부근에서 삽입결실을 포함하는 세포를 5% 미만, 예를 들어 4% 미만, 3% 미만, 2% 미만 또는 1% 미만 포함한다. 실시형태에서, 표적-외 삽입결실은 유전자의 서열 내에서, 예를 들어 유전자의 코딩 서열 내에서 형성된다. 실시형태에서, 어떠한 표적-외 삽입결실도, 예를 들어 본원에 기재된 바와 같은 관심 세포에서 유전자의 서열 내에서, 예를 들어 유전자의 코딩 서열 내에서 형성되지 않는다.

[0810] **X. 전달/작제물**

[0811] 성분, 예를 들어 Cas9 분자 또는 gRNA 분자, 또는 둘 모두는 다양한 형태로 전달, 제형화 또는 투여될 수 있다. 비제한적 예로서, gRNA 분자 및 Cas9 분자는(하나 이상의 조성물에) 제형화되고, 게놈 편집 사건이 필요한 세포로 직접 전달되거나 투여될 수 있다. 대안적으로, 하나 이상의 성분, 예를 들어 Cas9 분자 또는 gRNA 분자 또는 둘 모두를 인코딩하는 핵산은(하나 이상의 조성물에) 제형화되거나, 전달되거나 또는 투여될 수 있다. 일 양태에서, gRNA 분자는 gRNA 분자를 인코딩하는 DNA로서 제공되고, Cas9 분자는 Cas9 분자를 인코딩하는 DNA로서 제공된다. 일 실시형태에서, gRNA 분자 및 Cas9 분자는 별개의 핵산 분자 상에서 인코딩된다. 일 실시형태에서, gRNA 분자 및 Cas9 분자는 동일한 핵산 분자 상에서 인코딩된다. 일 양태에서, gRNA 분자는 RNA로서 제공되고, Cas9 분자는 Cas9 분자를 인코딩하는 DNA로서 제공된다. 일 실시형태에서, gRNA 분자는 예를 들어 본원에 기재된 하나 이상의 변형과 함께 제공된다. 일 양태에서, gRNA 분자는 RNA로서 제공되고, Cas9 분자는 Cas9 분자를 인코딩하는 mRNA로서 제공된다. 일 양태에서, gRNA 분자는 RNA로서 제공되고, Cas9 분자는 단백질로서 제공된다. 일 실시형태에서, gRNA 및 Cas9 분자는 리보핵 단백질 복합체(RNP)로서 제공된다. 일 양태에서, gRNA 분자는 gRNA 분자를 인코딩하는 DNA로서 제공되고, Cas9 분자는 단백질로서 제공된다.

[0812] 전달, 예를 들어 RNP의 전달(예를 들어 본원에 기재된 HSPC 세포로의)은 예를 들어 전기천공(예를 들어 당분야에 공지된) 또는 핵산 및/또는 폴리펩타이드 분자에 대해 세포막을 투과성으로 만드는 다른 방법에 의해 달성될 수 있다. 실시형태에서, CRISPR 시스템, 예를 들어 본원에 기재된 RNP는 4D-뉴클레오펙터(Lonza), 예를 들어 4D-뉴클레오펙터(Lonza) 상의 프로그램 CM-137을 사용한 전기천공에 의해 전달된다. 실시형태에서, CRISPR 시스템, 예를 들어 본원에 기재된 RNP는 약 800 볼트 내지 약 2000 볼트의 전압, 예를 들어 약 1000 볼트 내지 약 1800 볼트, 예를 들어 약 1200 볼트 내지 약 1800 볼트, 예를 들어 약 1400 볼트 내지 약 1800 볼트, 예를 들어 약 1600 볼트 내지 약 1800 볼트, 예를 들어 약 1700 볼트, 예를 들어 1700 볼트의 전압을 사용하는 전기천공에 의해 전달된다. 실시형태에서, 펄스 폭/길이 는 약 10 ms 내지 약 50 ms, 예를 들어 약 10 ms 내지 약 40 ms, 예를 들어 약 10 ms 내지 약 30 ms, 예를 들어 약 15 ms 내지 약 25 ms, 예를 들어 약 20 ms, 예를 들어 20 ms이다. 실시형태에서, 1, 2, 3, 4, 5 이상, 예를 들어 2, 예를 들어 1 펄스가 사용된다. 실시형태에서, CRISPR 시스템, 예를 들어 본원에 기재된 RNP는 약 1700 볼트(예를 들어 1700 볼트)의 전압, 약 20 ms(예를 들어 20 ms)의 펄스 폭을 사용하는 전기천공에 의해, 단일(1) 펄스를 사용하여 전달된다. 실시형태에서, 전기천공은 네온 전기천공기를 사용하여 달성된다. 막을 투과성으로 만들기 위한 추가적인 기술은 당분야에 공지되어 있으며, 예를 들어 세포 압착(예를 들어 그 전체가 참조로 본 명세서에 포함된 WO2015/023982 및 WO2013/059343에 기재됨), 나노 니들(nanoneedle)(예를 들어 그 전체가 참조로 본 명세서에 포함된 문헌[Chiappini et al., Nat. Mat., 14; 532-39], 또는 US2014/0295558에 기재됨) 및 나노 스트로(nanostraw)(예를 들어 그 전체가 참조로 본 명세서에 포함된 문헌[Xie, ACS Nano, 7(5); 4351-58]에 기재됨)를 포함한다.

[0813] 성분이 DNA에 인코딩되어 전달될 때, DNA는 전형적으로 발현을 초래하기 위하여 예를 들어 프로모터를 포함하는 제어 영역을 포함할 것이다. Cas9 분자 서열에 유용한 프로모터로는 CMV, EF-1알파, MSCV, PGK, CAG 제어 프로모터가 포함된다. gRNA에 유용한 프로모터로는 H1, EF-1a 및 U6 프로모터가 포함된다. 성분의 발현을 조정하기 위하여 유사하거나 다른 강도를 갖는 프로모터가 선택될 수 있다. Cas9 분자를 인코딩하는 서열은 핵 국재화 신호(NLS), 예를 들어 SV40 NLS를 포함할 수 있다. 실시형태에서, Cas9 분자 또는 gRNA 분자를 위한 프로모터는 독립적으로 유도성, 조직 특이적 또는 세포 특이적일 수 있다.

[0814] Cas9 분자 및 또는 gRNA 분자의 DNA-기반 전달

[0815] Cas9 분자 및/또는 gRNA 분자를 인코딩하는 DNA는 당분야에 공지된 방법에 의해 또는 본원에 기재된 바와 같이 대상체에게 투여되거나 세포 내로 전달될 수 있다. 예를 들어 Cas9-인코딩 및/또는 gRNA-인코딩 DNA는 예를 들

어 벡터(예를 들어 바이러스 또는 비-바이러스 벡터), 비-벡터 기반 방법(예를 들어 네이키드 DNA 또는 DNA 복합체를 이용하여) 또는 이의 조합에 의해 전달될 수 있다.

[0816] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 벡터(예를 들어 바이러스 벡터/바이러스, 플라스미드, 미니 서클 또는 나노 플라스미드)에 의해 전달된다.

[0817] 벡터는 Cas9 분자 및/또는 gRNA 분자를 인코딩하는 서열을 포함할 수 있다. 벡터는 또한, 예를 들어 Cas9 분자 서열에 융합된 신호 펩타이드(예를 들어 핵 국재화, 핵소체 국재화, 미토콘드리아 국재화를 위한)를 인코딩하는 서열을 포함할 수 있다. 예를 들어 벡터는 Cas9 분자를 인코딩하는 서열에 융합된 하나 이상의 핵 국재화 서열(예를 들어 SV40로부터의)을 포함할 수 있다.

[0818] 하나 이상의 조절/제어 요소, 예를 들어 프로모터, 인핸서, 인트론, 폴리아데닐화 신호, 코작(Kozak) 공통 서열, 내부 리보솜 진입 부위(IRES), 2A 서열 및 스플라이스 수용체 또는 공여자는 벡터 내에 포함될 수 있다. 일부 실시형태에서, 프로모터는 RNA 중합효소 II에 의해 인지된다(예를 들어 CMV 프로모터). 다른 실시형태에서, 프로모터는 RNA 중합효소 III에 의해 인지된다(예를 들어 U6 프로모터). 일부 실시형태에서, 프로모터는 조절되는 프로모터(예를 들어 유도성 프로모터)이다. 다른 실시형태에서, 프로모터는 구성적 프로모터이다. 일부 실시형태에서, 프로모터는 조직 특이적 프로모터이다. 일부 실시형태에서, 프로모터는 바이러스 프로모터이다. 다른 실시형태에서, 프로모터는 비-바이러스 프로모터이다.

[0819] 일부 실시형태에서, 벡터 또는 전달 비히클은 미니 서클이다. 일부 실시형태에서, 벡터 또는 전달 비히클은 나노 플라스미드이다.

[0820] 일부 실시형태에서, 벡터 또는 전달 비히클은(예를 들어 재조합 바이러스의 생성을 위한) 바이러스 벡터이다. 일부 실시형태에서, 바이러스는 DNA 바이러스(예를 들어 dsDNA 또는 ssDNA 바이러스)이다. 다른 실시형태에서, 바이러스는 RNA 바이러스(예를 들어 ssRNA 바이러스)이다.

[0821] 예시적인 바이러스 벡터/바이러스에는 예를 들어 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-관련 바이러스(AAV), 우두 바이러스, 폭스바이러스 및 단순 헤르페스 바이러스가 포함된다. 바이러스 벡터 기술은 당 분야에 잘 공지되어 있으며, 예를 들어 문헌[Sambrook et al., 2012, MOLECULAR CLONING: A LABORATORY MANUAL, volumes 1-4, Cold Spring Harbor Press, NY)] 및 다른 바이러스학 및 분자 생물학 매뉴얼에 기재되어 있다.

[0822] 일부 실시형태에서, 바이러스는 분열성 세포를 감염시킨다. 다른 실시형태에서, 바이러스는 비-분열성 세포를 감염시킨다. 일부 실시형태에서, 바이러스는 분열성 및 비-분열성 세포를 둘 모두 감염시킨다. 일부 실시형태에서, 바이러스는 숙주 게놈으로 통합될 수 있다. 일부 실시형태에서, 바이러스는 예를 들어 인간에서 감소된 면역을 갖도록 조작된다. 일부 실시형태에서, 바이러스는 복제-적격이다. 다른 실시형태에서, 바이러스는 예를 들어 비리온 복제 및/또는 패키징의 추가적인 라운드에 필요한 유전자에 대한 하나 이상의 코딩 영역이 다른 유전자로 대체되거나 결실된, 복제-결함이다. 일부 실시형태에서, 바이러스는 Cas9 분자 및/또는 gRNA 분자의 일시적 발현을 야기한다. 다른 실시형태에서, 바이러스는 Cas9 분자 및/또는 gRNA 분자의, 오래 지속되는, 예를 들어 적어도 1주, 2주, 1개월, 2개월, 3개월, 6개월, 9개월, 1년, 2년 또는 영구적 발현을 야기한다. 바이러스의 패키징 용량은, 예를 들어 적어도 약 4 kb 내지 적어도 약 30 kb, 예를 들어 적어도 5 kb, 10 kb, 15 kb, 20 kb, 25 kb, 30 kb, 35 kb, 40 kb, 45 kb 또는 50 kb로 다를 수 있다.

[0823] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 재조합 레트로바이러스에 의해 전달된다. 일부 실시형태에서, 레트로바이러스(예를 들어 몰로니(Moloney) 마우스 백혈병 바이러스)는, 예를 들어 숙주 게놈으로의 통합을 가능하게 하는, 역전사 효소를 포함한다. 일부 실시형태에서, 레트로바이러스는 복제-적격이다. 다른 실시형태에서, 레트로바이러스는 예를 들어 비리온 복제 및 패키징의 추가적인 라운드에 필요한 유전자에 대한 하나 이상의 코딩 영역이 다른 유전자로 대체되거나 결실된, 복제-결함이다.

[0824] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 재조합 렌티바이러스에 의해 전달된다. 예를 들어 렌티바이러스는, 예를 들어 바이러스 복제를 위해 요구되는 하나 이상의 유전자를 포함하지 않는, 복제-결함이다.

[0825] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 재조합 아데노바이러스에 의해 전달된다. 일부 실시형태에서, 아데노바이러스는 인간에서 감소된 면역을 갖도록 조작된다.

[0826] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 재조합 AAV에 의해 전달된다. 일부 실시형태에서, AAV는 그 게놈을 숙주 세포, 예를 들어 본원에 기재된 표적 세포의 게놈 내로 통합할 수 있다. 일부 실시형태에서,

AAV는 자가-상보적인 아데노-관련 바이러스(scAAV), 예를 들어 함께 어닐링하여 이중 가닥 DNA를 형성하는 양 가닥을 패키징하는 scAAV이다. 개시된 방법에 사용될 수 있는 AAV 혈청형에는, 예를 들어 AAV1, AAV2, 변형 AAV2(예를 들어 Y444F, Y500F, Y730F 및/또는 S662V에서의 변형), AAV3, 변형 AAV3(예를 들어 Y705F, Y73 1F 및/또는 T492V에서의 변형), AAV4, AAV5, AAV6, 변형 AAV6(예를 들어 S663V 및/또는 T492V에서의 변형), AAV8. AAV8.2, AAV9, AAVrh10이 포함되고, AAV2/8, AAV2/5 및 AAV2/6과 같은 위형 AAV 또한 개시된 방법에 사용될 수 있다.

[0827] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 혼종 바이러스, 예를 들어 본원에 기재된 하나 이상의 바이러스의 혼종에 의해 전달된다.

[0828] 패키징 세포는 숙주 또는 표적 세포를 감염시킬 수 있는 바이러스 입자를 형성하기 위해 사용된다. 이러한 세포에는, 아데노바이러스를 패키징할 수 있는 293 세포, 및, 레트로바이러스를 패키징할 수 있는 ψ2 세포 또는 PA317 세포가 포함된다. 유전자 치료에 사용되는 바이러스 벡터는 통상적으로 핵산 벡터를 바이러스 입자로 패키징하는 생산체 세포주에 의해 생성된다. 벡터는 전형적으로 패키징 및 숙주 또는 표적 세포로의 후속적 통합(적용 가능한 경우)에 요구되는 최소한의 바이러스 서열을 함유하며, 다른 바이러스 서열은 발현될 단백질을 인코딩하는 발현 카세트에 의해 대체된다. 예를 들어 유전자 치료에 사용되는 AAV 벡터는 전형적으로 숙주 또는 표적 세포에서의 패키징 및 유전자 발현에 요구되는 AAV 게놈으로부터의 역위 말단 반복(ITR) 서열만을 보유한다. 누락된 바이러스 기능은 패키징 세포주에 의해 트란스(trans)로 공급된다. 이후, 바이러스 DNA는 다른 AAV 유전자, 즉 rep 및 cap을 인코딩하지만 ITR 서열이 결여된 헬퍼 플라스미드를 함유하는 세포주에서 패키징된다. 세포주는 또한 헬퍼로서 아데노바이러스로 감염된다. 헬퍼 바이러스는 AAV 벡터의 복제 및 헬퍼 플라스미드로부터의 AAV 유전자의 발현을 촉진한다. 헬퍼 플라스미드는 ITR 서열의 결여로 인해 상당한 양에서 패키징되지 않는다. 아데노바이러스에 의한 오염은, 예를 들어 아데노바이러스가 AAV보다 더 민감한 열 처리에 의해 감소될 수 있다.

[0829] 실시형태에서, 바이러스 벡터는 세포 유형 및/또는 조직 유형 인지 능력을 갖는다. 예를 들어 바이러스 벡터는 상이한/대안적 바이러스 외피 당단백질로 위형화될 수 있고; 세포 유형-특이적 수용체로 조작될 수 있고(예를 들어 펩타이드 리간드, 단일 사슬 항체, 성장 인자와 같은 표적화 리간드를 혼입시키기 위한 바이러스 외피 당단백질의 유전적 변형); 및/또는 바이러스 당단백질을 인지하는 하나의 말단 및 표적 세포 표면의 모이어티를 인지하는 다른 말단으로 이중 특이성을 갖는 분자 가교(예를 들어 리간드-수용체, 단일클론 항체, 아비딘-비오틴 및 화학적 접합)를 갖도록 조작될 수 있다.

[0830] 실시형태에서, 바이러스 벡터는 세포 유형 특이적 발현을 달성한다. 예를 들어 조직 특이적 프로모터는 표적 세포에서만 전이유전자(Cas9 및 gRNA)의 발현을 한정하도록 작제될 수 있다. 벡터의 특이성은 또한 전이유전자 발현의 마이크로RNA-의존적 제어에 의해 매개될 수 있다. 실시형태에서, 바이러스 벡터는 바이러스 벡터와 표적 세포막의 융합 효율을 증가시켰다. 예를 들어 세포 내로의 바이러스 흡수를 증가시키기 위해 융합-적격 헤마글루틴(HA)과 같은 융합 단백질이 혼입될 수 있다. 실시형태에서, 바이러스 벡터는 핵 국재화 능력을 갖는다. 예를 들어(세포 분열 동안의) 세포벽의 분해를 요구하며 따라서 비-분열성 세포를 감염시키지 않을 바이러스 는 바이러스의 매트릭스 단백질 내에 핵 국재화 펩타이드를 혼입시켜 비-증식성 세포의 형질도입이 가능하도록 변경될 수 있다.

[0831] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 비-벡터 기반 방법(예를 들어 네이키드 DNA 또는 DNA 복합체를 사용하는)에 의해 전달된다. 예를 들어, DNA는 이를테면 유기적으로 변형된 실리카 또는 실리케이트(Ormosil), 전기천공, 유전자 총, 초음파 천공, 자기 주입, 지질-매개 형질 감염, 덴드리머, 무기 나노 입자, 포스페이트 칼슘 또는 이의 조합에 의해 전달될 수 있다.

[0832] 일부 실시형태에서, Cas9- 및/또는 gRNA-인코딩 DNA는 벡터 및 비-벡터 기반 방법의 조합에 의해 전달된다. 예를 들어 비로좀(virosome)은 불활성화된 바이러스(예를 들어 HIV 또는 인플루엔자 바이러스)와 조합된 리포좀을 포함하며, 바이러스 또는 리포좀 방법 단독보다, 예를 들어 호흡 상피 세포에서, 더 효율적인 유전자 전달을 야기할 수 있다.

[0833] 실시형태에서, 전달 비히클은 비-바이러스 벡터이다. 실시형태에서, 비-바이러스 벡터는 무기 나노 입자(예를 들어 나노 입자의 표면으로 페이로드(payload)에 부착된)이다. 예시적인 무기 나노 입자에는, 예를 들어 자성 나노 입자(예를 들어 Fe $_{l}$vlnO_2), 또는 실리카가 포함된다. 나노 입자의 외부 표면은 페이로드의 부착(예를 들어 접합 또는 포착)을 가능하게 하는 양전하 중합체(예를 들어 폴리에틸렌이민, 폴리라이신, 폴리세린)와 접합될 수 있다. 실시형태에서, 비-바이러스 벡터는 유기 나노 입자(예를 들어 나노 입자 내부에 페이로드의 포착)

이다. 예시적인 유기 나노 입자에는 예를 들어 폴리에틸렌 글리콜(PEG)로 코팅된 중성 헬퍼 지질과 함께 양이온성 지질을 함유하는SNALP 리포좀 및 지질 코팅으로 코팅된 프로타민과 핵산 복합체가 포함된다.

[0834] CRISPR 시스템, 또는 CRISPR 시스템 또는 이의 성분을 인코딩하는 핵산(예를 들어 벡터)의 전달을 위한 예시적인 지질 및/또는 중합체에는, 예를 들어 그 각각의 내용의 전문이 참조로 본원에 포함된 WO2011/076807, WO2014/136086, WO2005/060697, WO2014/140211, WO2012/031046, WO2013/103467, WO2013/006825, WO2012/006378, WO2015/095340, 및 WO2015/095346에 기재된 것들이 포함된다. 실시형태에서, 비히클은 나노 입자 및 리포좀의 표적 세포 갱신을 증가시키는 표적화 변형, 예를 들어 세포 특이적 항원, 단일클론 항체, 단일 사슬 항체, 압타머, 중합체, 당 및 세포 침투 펩타이드를 갖는다. 실시형태에서, 비히클은 융합 생성 및 엔도좀-불안정화 펩타이드/중합체를 사용한다. 실시형태에서, 비히클은 산-촉발된 입체 형태 변화를 겪는다(예를 들어 화물의 엔도좀 탈출을 가속화하기 위해). 실시형태에서, 예를 들어 세포 구획에서의 방출을 위해 자극-절단 가능한 중합체가 사용된다. 예를 들어 환원성 세포 환경에서 절단되는 디설파이드 기반 양이온성 중합체가 사용될 수 있다.

[0835] 실시형태에서, 전달 비히클은 생물학적 비-바이러스 전달 비히클이다. 실시형태에서, 비히클은 약화된 박테리아(예를 들어 침입성이지만 발병을 방지하기 위해 약화되도록 자연적으로 또는 인공적으로 조작되고 전이유전자를 발현하는(예를 들어 리스테리아 모노사이토게네스(Listeria monocytogenes), 일정 살모넬라 균주, 비피도박테리움 론굼(Bifidobacterium longum) 및 변형된 에스케리치아 콜라이(에스케리키아 콜라이)), 특정 조직을 표적화하기 위한 영양 및 조직-특이적 향성(tropism)을 갖는 박테리아, 표적 조직 특이성을 변경하기 위해 변형된 표면 단백질을 갖는 박테리아)이다. 실시형태에서, 비히클은 유전적으로 변형된 박테리오파지(예를 들어 큰 패키징 용량, 더 적은 면역원성을 갖고, 포유류 플라스미드 유지 서열을 함유하며, 혼입된 표적화 리간드를 갖는, 조작된 파지)이다. 실시형태에서, 비히클은 포유류 바이러스 유사 입자이다. 예를 들어 변형된 바이러스 입자가 생성될 수 있다(예를 들어 "비어있는" 입자의 정제 후 원하는 화물과 바이러스의 생체외 조립에 의해). 비히클은 또한 표적 조직 특이성을 변경하기 위해 표적화 리간드를 혼입하도록 조작될 수 있다. 실시형태에서, 비히클은 생물학적 리포좀이다. 예를 들어 생물학적 리포좀은 인간 세포에서 유래된 인지질 기반 입자(예를 들어 구형 구조로 분해된 적혈구인, 대상체에서 유래된 적혈구 고스트(예를 들어 조직 표적화는 다양한 조직 또는 세포-특이적 리간드의 부착에 의해 달성될 수 있다), 또는 분비성 엑소좀 - 대상체(즉, 환자) 유래된 식균 작용 기원의 막-결합 나노베지클(30 내지 100 nm)(예를 들어 다양한 세포 유형으로부터 생산될 수 있으며 따라서 표적화 리간드의 필요 없이 세포들에 의해 잡힐 수 있다) 이다.

[0836] 실시형태에서, Cas 시스템의 성분, 예를 들어 본원에 기재된 Cas9 분자 성분 및/또는 gRNA 분자 성분 외의 하나 이상의 핵산 분자(예를 들어 DNA 분자)가 전달된다. 실시형태에서, 핵산 분자는 Cas 시스템의 하나 이상의 성분이 전달되는 것과 동시에 전달된다. 실시형태에서, 핵산 분자는 Cas9 시스템의 하나 이상의 성분이 전달되기 전 또는 후(예를 들어 약 30분, 1시간, 2시간, 3시간, 6시간, 9시간, 12시간, 1일, 2일, 3일, 1주, 2주 또는 4주 미만)에 전달된다. 실시형태에서, 핵산 분자는 Cas9 시스템의 하나 이상의 성분, 예를 들어 Cas9 분자 성분 및/또는 gRNA 분자 성분이 전달되는 것과 상이한 수단에 의해 전달된다. 핵산 분자는 본원에 기재된 전달 방법 중 임의의 것에 의해 전달될 수 있다. 예를 들어 핵산 분자는 바이러스 벡터, 예를 들어 통합-결핍 렌티바이러스에 의해 전달될 수 있고, Cas9 분자 성분 및/또는 gRNA 분자 성분은 전기천공에 의해 전달될 수 있어, 예를 들어 핵산(예를 들어 DNA)에 의해 야기되는 독성이 감소될 수 있다. 실시형태에서, 핵산 분자는 치료 단백질, 예를 들어 본원에 기재된 단백질을 인코딩한다. 실시형태에서, 핵산 분자는 RNA 분자, 예를 들어 본원에 기재된 RNA 분자를 인코딩한다.

[0837] Cas9 분자를 인코딩하는 RNA의 전달

[0838] Cas9 분자(예를 들어 활성 Cas9 분자, 불활성 Cas9 분자 또는 불활성 Cas9 융합 단백질) 및/또는 gRNA 분자를 인코딩하는 RNA는 당분야에 공지된 방법에 의해 또는 본원에 기재된 바와 같이, 세포, 예를 들어 본원에 기재된 표적 세포로 전달될 수 있다. 예를 들어 Cas9- 인코딩 및/또는 gRNA-인코딩 RNA는, 예를 들어 미세주사, 전기천공, 지질-매개 형질 감염, 펩타이드-매개 전달, 또는 이의 조합에 의해 전달될 수 있다.

[0839] Cas9 분자의 단백질로서 전달

[0840] Cas9 분자(예를 들어 활성 Cas9 분자, 불활성 Cas9 분자 또는 불활성 Cas9 융합 단백질)는 당분야에 공지된 방법에 의해 또는 본원에 기재된 바와 같이, 세포로 전달될 수 있다. 예를 들어 Cas9 단백질 분자는, 예를 들어 미세주사, 전기천공, 지질-매개 형질 감염, 펩타이드-매개 전달, 세포 압착 또는 마모(예를 들어 나노 니들에 의해), 또는 이의 조합에 의해 전달될 수 있다. 전달은 gRNA를 인코딩하는 DNA, 또는 gRNA, 예를 들어 리보핵

단백질 복합체(RNP)로 gRNA 및 Cas9 단백질을 미리 복합체 형성하는 것과 동반할 수 있다.

[0841] 양태에서, 예를 들어 본원에 기재된 Cas9 분자는 단백질로서 전달되고, gRNA 분자는 하나 이상의 RNA로서(예를 들어 본원에 기재된 dgRNA 또는 sgRNA로서) 전달된다. 실시형태에서, Cas9 단백질은 세포로의 전달 전에, 예를 들어 본원에 기재된 바와 같이, 리보핵 단백질 복합체("RNP")로서 gRNA 분자와 복합체 형성된다. 실시형태에서, RNP는 당분야에 공지된 임의의 방법, 예를 들어 전기천공에 의해 예를 들어 본원에 기재된, 세포 내로 전달될 수 있다. 본원에 기재된 바와 같이, 그리고 임의의 이론으로 결부시키고자 하는 것은 아니지만, 세포로 전달되는 RNP의 농도가 감소되는 때에도, 예를 들어 본원에 기재된 표적 세포에서 표적 서열에서의 높은 편집 %(예를 들어 85% 초과, 90% 초과, 95% 초과, 98% 초과, 또는 99% 초과)을 야기하는 gRNA 분자 및 Cas9분자를 사용하는 것이 바람직할 수 있다. 또한, 임의의 이론으로 결부시키고자 하는 것은 아니지만, 표적 세포에서(낮은 RNP 농도에서를 포함하여) 표적 서열에서의 높은 편집 %을 생산하는 gRNA분자를 포함하는 감소된 또는 낮은 농도의 RNP를 전달하는 것은, 표적-외 편집 사건의 빈도 및 수를 감소시킬 수 있기 때문에 유익할 수 있다. 일 양태에서, 낮은 또는 감소된 농도의 RNP가 사용되는 경우, dgRNA 분자를 갖는 RNP를 생성하기 위하여 다음의 예시적인 절차가 사용될 수 있다:

[0842] 1. Cas9 분자 및 tracr를 고농도 용액(예를 들어 세포로 전달되는 최종 RNP 농도보다 높은 농도)으로 제공하고, 두 성분이 평형을 이룰 수 있게 한다;

[0843] 2. crRNA 분자를 제공하고, 성분들이 평형을 이룰 수 있게 한다(이로써 RNP의 고농도 용액을 형성한다);

[0844] 3. RNP 용액을 원하는 농도로 희석한다;

[0845] 4. 상기 원하는 농도의 상기 RNP를 예를 들어 전기천공에 의해 표적 세포로 전달한다.

[0846] 상기 절차는 상기 단계 2를 생략하고, 단계 1에서 Cas9 분자 및 sgRNA 분자를 용액 내에서 고농도로 제공하고, 구성성분을 평형화시킴으로써 sgRNA 분자와 함께 사용되기 위해 변형될 수 있다. 실시형태에서, Cas9 분자 및 각각의 gRNA 구성성분은 용액 내에서 1:2 비(ratio)(Cas9:gRNA), 예를 들어 1:2 몰비의 Cas9:gRNA 분자로 제공된다. dsRNA 분자가 사용되는 경우, 상기 비, 예를 들어 몰비는 1:2:2(Cas9:tracr:crRNA)이다. 실시형태에서, RNP는 20 uM 이상의 농도, 예를 들어 약 20 uM 내지 약 50 uM의 농도로 형성된다. 실시형태에서, RNP는 10 uM 이상의 농도, 예를 들어 약 10 uM 내지 약 30 uM의 농도로 형성된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 10 uM 이하의 최종 농도(예를 들어 약 0.01 uM 내지 약 10 uM의 농도)까지 희석된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 3 uM 이하의 최종 농도(예를 들어 약 0.01 uM 내지 약 3 uM의 농도)까지 희석된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 1 uM 이하의 최종 농도(예를 들어 약 0.01 uM 내지 약 1 uM의 농도)까지 희석된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 0.3 uM 이하의 최종 농도(예를 들어 약 0.01 uM 내지 약 0.3 uM의 농도)까지 희석된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 3 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 2 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 1 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 0.3 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 0.1 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 0.05 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 0.03 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 (예를 들어 본원에 기재된) 표적 세포를 전달하기 위해 상기 표적 세포를 포함하는 용액 내에서 약 0.01 uM의 최종 농도로 제공된다. 실시형태에서, RNP는 전기천공에 적합한 배지 내에서 제제화된다. 실시형태에서, RNP는 예를 들어 본원에 기재된 전기천공 조건을 사용하여 전기천공에 의해 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 HSPC 세포에 전달된다.

[0847] 양태에서, 유전자 편집 시스템(예를 들어 CRISPR 시스템)의 구성성분 및/또는 유전자 편집 시스템(예를 들어 CRISPR 시스템)의 하나 이상의 구성성분을 인코딩하는 핵산은 세포를 기계적으로 섭동시킴으로써, 예를 들어 상기 세포를 이러한 세포를 구속하는 포어(pore) 또는 채널을 통해 통과시킴으로써 세포 내로 도입된다. 이러한

섭동(purturbation)은 예를 들어 본원에 기재된 바와 같은 유전자 편집 시스템(예를 들어 CRISPR 시스템)의 구성성분 및/또는 유전자 편집 시스템(예를 들어 CRISPR 시스템)의 하나 이상의 구성성분을 인코딩하는 핵산을 포함하는 용액 내에서 달성될 수 있다. 실시형태에서, 이러한 섭동은 예를 들어 본원, 예를 들어 실시예 및 PCT 특허 출원 PCT/US17/54110(그 전체가 참조로 본 명세서에 포함됨)에 기재된 바와 같이 TRIAMF 시스템을 사용하여 달성된다.

[0848] 성분들의 두 가지 방식 또는 차등적 전달

[0849] Cas 시스템의 성분, 예를 들어 Cas9 분자 성분 및 gRNA 분자 성분의 별도의 전달, 더욱 구체적으로, 상이한 방식에 의한 성분들의 전달은, 예를 들어 조직 특이성 및 안전성을 개선함으로써 성능을 향상시킬 수 있다.

[0850] 실시형태에서, Cas9 분자 및 gRNA 분자는 상이한 방식(또는 때때로 본원에서 차등적 방식으로 지칭됨)에 의해 전달된다. 본원에 사용되는 상이한 또는 차등적 방식은 대상체 성분 분자, 예를 들어 Cas9 분자, gRNA 분자 또는 주형 핵산에 대해 상이한 약력학적 또는 약동학적 특성을 부여하는 전달 방식을 지칭한다. 예를 들어 전달 방식은 상이한 조직 분포, 상이한 반감기, 또는 예를 들어 선택된 구획, 조직 또는 기관에서의 상이한 시간 분포를 야기할 수 있다.

[0851] 일부 전달 방식, 예를 들어 자율적 복제 또는 세포 핵산 내로의 삽입에 의한, 세포 또는 세포의 자손에서 지속되는 핵산 벡터에 의한 전달은 성분의 더 지속적인 발현 및 존재를 야기한다.

[0852] **XI. 치료 방법**

[0853] 본원에 기재된 Cas9 시스템, 예를 들어 하나 이상의 gRNA 분자 및 하나 이상의 Cas9 분자는 포유류, 예를 들어 인간에서 질병의 치료에 유용하다. 용어 "치료하다", "치료된", "치료하는", 및 "치료"는 질병의 증상, 합병증 또는 생화학적 징후의 개시를 방지 또는 지연하며, 증상을 경감 또는 질병, 질환, 또는 장애의 추가적인 발달을 정지 또는 저해하기 위해 cas9 시스템, 예를 들어 하나 이상의 gRNA 분자 및 하나 이상의 cas9 분자의 세포로의 투여를 포함한다. 치료는 또한 질병의 증상, 합병증 또는 생화학적 징후의 개시를 방지 또는 지연하며, 증상을 경감 또는 질병, 질환, 또는 장애의 추가적인 발달을 정지 또는 저해하기 위해, 본 발명의 gRNA 분자(또는 하나를 초과하는 gRNA 분자)의 도입 또는 본원에 기재된 CRISPR 시스템의 도입, 또는 본원에 기재된 상기 세포를 제조하는 방법 중 임의의 것에 의해 변형된 하나 이상의(예를 들어 집단의) 세포, 예를 들어 HSPC의 투여를 포함할 수 있다. 치료는 예방적(질병의 발병을 방지 또는 지연시키는 것, 또는 이의 임상적 또는 준임상적 증상의 표현을 방지하는 것) 또는 질병의 표현 후 증상의 치료적 억제 또는 경감일 수 있다. 치료는 본원에 기재된 치료적 기준에 의해 측정될 수 있다. 따라서, 본 발명의 "치료"의 방법은 또한 질병 또는 질환의 하나 이상의 증상을 치료, 중증도를 감소, 또는 개선하기 위하여, 이러한 치료의 부재 시에 기대되는 것 이상으로 대상체의 건강 또는 생존을 연장시키기 위하여, 상기 세포 내로의 cas9 시스템(예를 들어 하나 이상의 gRNA 분자 및 하나 이상의 Cas9 분자)의 도입에 의해 변경된 세포를 대상체에게 투여하는 것을 포함한다. 예를 들어 "치료"는 환자에서 질병 증상의 적어도 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 이상 만큼의 경감을 포함한다.

[0854] 본원, 예를 들어 표 1에 기재된 표적화 도메인을 포함하는 gRNA 분자를 포함하는 Cas9 시스템, 및 방법과 세포(예를 들어 본원에 기재된 바와 같음)는 혈색소병증의 치료에 유용하다.

[0855] **혈색소병증**

[0856] 혈색소병증은 적혈구(RBC)의 생산 감소 및/또는 파괴(용혈) 증가가 있는 유전적 기원의 다수의 빈혈을 포괄한다. 이들은 또한 산소 농도를 유지하는 능력 손상을 수반하는 비정상적 헤모글로빈 생산을 야기하는 유전적 결함을 포함한다. 일부 이러한 장애는 정상적인 β-글로빈을 충분한 양으로 생산하지 못하는 것을 수반하지만 다른 것은 정상적인 β-글로빈을 완전히 생산하지 못하는 것을 수반한다. β-글로빈 단백질과 관련된 이들 장애는 일반적으로 β-혈색소병증으로 지칭된다. 예를 들어 β-탈라세미아는 HbA 부족 또는 부재를 초래하는, β-글로빈 유전자의 발현에서의 부분적 또는 완전한 결함에 기인한다. 겸상 적혈구 빈혈은 비정상적(겸상) 헤모글로빈(HbS)의 생산을 초래하는, β-글로빈 구조 유전자에서의 점 돌연변이에 기인한다. HbS는 특히 탈산분해된 조건에서 중합되기 쉽다. HbS 적혈구는 정상적 적혈구보다 더 손상되기 쉽고 더 쉽게 용혈을 겪어, 결국 빈혈을 초래한다.

[0857] 실시형태에서, 알파 글로빈 또는 베타 글로빈에서의 유전적 결함은 본원에 기재된 Cas9 분자 및 gRNA 분자, 예를 들어 CRISPR 시스템을 사용하는 예를 들어 상동 재조합에 의해 교정된다.

[0858] 실시형태에서, 알파 글로빈 또는 베타 글로빈의 야생형(예를 들어 돌연변이되지 않은) 카피를 인코딩하는 유전자는 세포의 게놈 내로, 예를 들어 세이프 하버(safe harbor) 부위에, 예를 들어 AAVS1 세이프 하버 부위에, 본원에 기재된 CRISPR 시스템 및 방법을 이용하는 상동 재조합에 의해 삽입된다.

[0859] 실시형태에서, 혈색소병증-관련 유전자는 본원에 기재된 Cas9 분자 및 gRNA 분자를 사용하여 표적화된다. 예시적인 표적에는, 예를 들어 감마-글로빈 유전자의 제어와 관련된 유전자가 포함된다. 실시형태에서, 표적은 비결실 HPFH 영역이다.

[0860] 태아 헤모글로빈(또한 헤모글로빈 F 또는 HbF 또는 $\alpha_2\gamma_2$)은 2개의 성체 알파-글로빈 폴리펩타이드 및 2개의 태아 베타-유사 감마-글로빈 폴리펩타이드의 사합체이다. HbF는 자궁에서의 최종 7개월의 발달 동안의 인간 태아에서 및 대략적으로 6개월령까지의 신생아에서 주요 산소 수송 단백질이다. 기능적으로, 태아 헤모글로빈은, 성체 형태보다 더 높은 친화성으로 산소를 결합할 수 있어, 발달하고 있는 태아에게 모체 혈류로부터의 산소에 대한 더 좋은 접근을 제공한다는 점에서 성체 헤모글로빈과 가장 많이 상이하다.

[0861] 신생아에서, 태아 헤모글로빈은 생후 약 6개월까지 성체 헤모글로빈에 의해 거의 완전히 대체된다. 성체에서, 태아 헤모글로빈 생산은 약리학적으로 재활성화될 수 있으며, 혈색소병증과 같은 질병의 치료에 유용하다. 예를 들어 혈색소병증을 갖는 일정 환자에서, 더 높은 수준의 감마-글로빈 발현은 결함있거나 손상된 베타-글로빈 유전자 생산을 부분적으로 보상할 수 있어, 이들 질병에서 임상적 중증도를 개선할 수 있다. 증가된 HbF 수준 또는 F-세포(HbF 함유 적혈구) 수는 혈색소병증, 예를 들어 베타-탈라세미아 메이저 및 겸상 적혈구 빈혈의 질병 중증도를 개선할 수 있다.

[0862] 놀랍게 발견된 바와 같이, 증가된 HbF 수준 또는 F-세포 수치는 세포, 예를 들어 HSPC 및/또는 HSPC(예를 들어 본원에 기재된 하나 이상의 gRNA 분자에 의해 변형된 HSPC)로부터 분화된 세포 내의 하나 이상의 비결실 HPFH 영역에서 연관된 삽입결실 형성일 수 있다. 일 실시형태에서, 세포는 조혈 줄기세포 또는 전구 세포이다.

[0863] 겸상 적혈구 질병

[0864] 겸상 적혈구 질병은 헤모글로빈에 영향을 미치는 장애 군이다. 이 장애를 갖는 인간들은 적혈구를 낫 또는 초승달 모양으로 일그러뜨릴 수 있는, 비정형 헤모글로빈 분자(헤모글로빈 S)를 갖는다. 이 장애의 특징은 적은 수의 적혈구(빈혈), 반복 감염, 및 주기적 통증 사건을 포함한다.

[0865] HBB 유전자에서의 돌연변이가 겸상 적혈구 질병을 야기한다. HBB 유전자는 베타-글로빈을 만들기 위한 지시를 제공한다. 다양한 형태의 베타-글로빈은 HBB 유전자에서의 상이한 돌연변이에 기인한다. 하나의 특정 HBB 유전자 돌연변이는 헤모글로빈 S(HbS)로 알려진 베타-글로빈의 비정상적 형태를 생산한다. HBB 유전자에서의 다른 돌연변이는 헤모글로빈 C(HbC) 및 헤모글로빈 E(HbE)와 같은 추가적인 베타-글로빈의 비정상적 형태를 초래한다. HBB 유전자 돌연변이는 또한 매우 낮은 수준의 베타-글로빈, 즉 베타 탈라세미아를 야기할 수 있다.

[0866] 겸상 적혈구 질병을 갖는 인간들에서, 헤모글로빈 내의 베타-글로빈 하위단위 중 적어도 하나는 헤모글로빈 S로 대체된다. 겸상 적혈구 질병의 통상적인 형태인, 겸상 적혈구 빈혈에서, 헤모글로빈 S는 헤모글로빈 내의 베타-글로빈 하위단위 둘 모두를 대체한다. 다른 유형의 겸상 적혈구 질병에서, 헤모글로빈 내의 오직 하나의 베타-글로빈 하위단위이 헤모글로빈 S로 대체된다. 다른 베타-글로빈 하위단위은 헤모글로빈 C와 같은 상이한 비정상적 변이체로 대체된다. 예를 들어 겸상 헤모글로빈 C(HbSC) 질병을 갖는 인간들은 베타-글로빈 대신 헤모글로빈 S 및 헤모글로빈 C를 갖는 헤모글로빈 분자를 갖는다. 헤모글로빈 S 및 베타 탈라세미아를 생산하는 돌연변이들이 함께 발생하는 경우, 개체는 헤모글로빈 S-베타 탈라세미아(HbSBetaThal) 질병을 갖는다.

[0867] 베타 탈라세미아

[0868] 베타 탈라세미아는 헤모글로빈 생산을 감소시키는 혈액 장애이다. 베타 탈라세미아를 갖는 인간들에서, 낮은 수준의 헤모글로빈이 신체의 많은 부분에서 산소의 부족을 초래한다. 병에 걸린 개체는 또한 적혈구의 부족(빈혈)을 가져, 창백한 피부, 허약, 피로 및 더 심각한 합병증을 야기할 수 있다. 베타 탈라세미아를 갖는 인간들은 비정상적 혈전을 발생시킬 위험이 증가한다.

[0869] 베타 탈라세미아는 증상의 중증도에 따라 두 가지 유형으로 분류된다: 탈라세미아 메이저(쿨리(Cooley's) 빈혈로도 알려진) 및 탈라세미아 인터메디아(intermedia). 두 가지 유형 중에서, 탈라세미아 메이저가 더 중증이다.

[0870] HBB 유전자에서의 돌연변이는 베타 탈라세미아를 야기한다. HBB 유전자는 베타-글로빈을 만들기 위한 지시를 제공한다. HBB 유전자에서의 일부 돌연변이는 임의의 베타-글로빈의 생산을 방해한다. 베타-글로빈의 부재는 베타

-제로(B^0) 탈라세미아로 지칭된다. 다른 HBB 유전자 돌연변이는 일부 베타-글로빈이 생산되지만 감소된 양으로 생산되도록 한다, 즉 베타-플러스(B^+) 탈라세미아. 두 가지 유형을 갖는 인간들은 탈라세미아 메이저 및 탈라셈아 인터메디아로 진단되어 왔다.

[0871] 실시형태에서, 제1 유전자 또는 유전자좌를 표적화하는 Cas9 분자/gRNA 분자 복합체는 제2 유전자, 예를 들어 제2 유전자에서의 돌연변이를 특징으로 하는 장애를 치료하는 데 사용된다. 예로서, 예를 들어 편집 또는 페이로드 전달에 의한 제1 유전자의 표적화는 제2 유전자, 예를 들어 돌연변이체 제2 유전자의 영향을 보상하거나, 이로부터의 추가 손상을 저해할 수 있다. 실시형태에서, 대상체에 의해 옮겨진 제1 유전자의 대립유전자(들)는 장애의 원인이 아니다. 예를 들어, 본원에 제시된 바와 같이, 비결실 HPFH 영역, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역에서 삽입결실 형성을 유도하는 gRNA 분자는 변형된 HSPC로부터 분화된 적혈구계 세포(본원에 기재된 바와 같음)에서 태아 헤모글로빈의 상향조절을 초래할 수 있고, 이론으로 결부시키고자 하는 것은 아니지만, 이러한 태아 헤모글로빈 상향조절은 겸상 돌연변이를 갖는 HBB 유전자에 대해 보상하고, 보정한다.

[0872] 일 양태에서, 본 발명은 증가된 태아 헤모글로빈(HbF)을 필요로 하는 포유류, 예를 들어 인간의 치료에 관한 것이다.

[0873] 일 양태에서, 본 발명은 혈색소병증으로 진단되거나 또는 혈색소병증을 발병할 위험이 있는 포유류, 예를 들어 인간의 치료에 관한 것이다.

[0874] 일 양태에서, 혈색소병증은 β-혈색소병증이다. 일 양태에서, 혈색소병증은 겸상 적혈구 질병이다. 일 양태에서, 혈색소병증은 베타 탈라세미아이다.

[0875] **혈색소병증의 치료 방법**

[0876] 또 다른 양태에서, 본 발명은 치료 방법을 제공한다. 양태에서, 본 발명의 gRNA 분자, CRISPR 시스템 및/또는 세포는 이를 필요로 하는 환자를 치료하기 위해 사용된다. 양태에서, 환자는 포유류, 예를 들어 인간이다. 양태에서 환자는 혈색소병증을 갖는다. 실시형태에서, 환자는 겸상 적혈구 질병을 갖는다. 실시형태에서, 환자는 베타 탈라세미아를 갖는다.

[0877] 일 양태에서, 치료 방법은 포유류, 예를 들어 인간에게 하나 이상의 gRNA 분자, 예를 들어 표 1에 기재된 표적화 도메인을 포함하는 하나 이상의 gRNA 분자, 및 본원에 기재된 하나 이상의 cas9 분자를 투여하는 것을 포함한다.

[0878] 일 양태에서, 치료 방법은 포유류에게 세포 집단을 투여하는 것을 포함하며, 세포 집단은, 하나 이상의 gRNA 분자, 예를 들어 표 1에 기재된 표적화 도메인을 포함하는 하나 이상의 gRNA 분자, 및 본원에 기재된 하나 이상의 cas9 분자, 예를 들어 본원에 기재된 CRISPR 시스템이 투여된 포유류, 예를 들어 인간으로부터의 세포 집단이다. 일 실시형태에서, 세포로의 하나 이상의 gRNA 분자 또는 CRISPR 시스템의 투여는 생체내에서 달성된다. 일 실시형태에서, 세포로의 하나 이상의 gRNA 분자 또는 CRISPR 시스템의 투여는 생체외에서 달성된다.

[0879] 일 양태에서, 치료 방법은 포유류, 예를 들어 인간에게, 하나 이상의 gRNA 분자, 예를 들어 표 1에 기재된 표적화 도메인을 포함하는 하나 이상의 gRNA 분자, 및 본원에 기재된 하나 이상의 cas9 분자를 포함하거나 한때 포함하였던 세포, 또는 상기 세포의 자손을 포함하는 세포 집단을 유효량 투여하는 것을 포함한다. 일 실시형태에서, 세포는 포유류에 대하여 동종이계이다. 일 실시형태에서, 세포는 포유류에 대하여 자가 유래이다. 일 실시형태에서, 세포는 포유류로부터 수확되고, 생체외에서 조작되고, 포유류로 복귀된다.

[0880] 양태에서, 하나 이상의 gRNA 분자, 예를 들어 표 1에 기재된 표적화 도메인을 포함하는 하나 이상의 gRNA 분자, 및 본원에 기재된 하나 이상의 cas9 분자를 포함하거나 한때 포함하였던 세포, 또는 상기 세포의 자손은 줄기 세포 또는 전구 세포를 포함한다. 일 양태에서, 줄기 세포는 조혈 줄기 세포이다. 일 양태에서, 전구 세포는 조혈 전구 세포이다. 일 양태에서, 세포는 조혈 줄기 세포 및 조혈 전구 세포 둘 모두를 포함하며, 예를 들어 HSPC이다. 일 양태에서, 세포는 CD34+ 세포를 포함하며, 예를 들어 이로 구성된다. 일 양태에서, 세포는 실질적으로 CD34- 세포가 없다. 일 양태에서, 세포는 CD34+/CD90+ 줄기 세포를 포함하며, 예를 들어 이로 구성된다. 일 양태에서, 세포는 CD34+/CD90- 세포를 포함하며, 예를 들어 이로 구성된다. 양태에서, 세포는 상기 기재되거나 본원에 기재된 세포 유형 중 하나 이상을 포함하는 집단이다.

[0881] 일 실시형태에서, 본 개시내용은 하기 방법을 필요로 하는 포유류, 예를 들어 인간에서 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 베타-탈라쎄미아를 치료하는 방법, 또는 태아 헤모글로빈 발현을 증가시키는 방법을 제

공하며, 상기 방법은:

[0882] a) 포유류로부터 HSPC의 집단(예를 들어 CD34+ 세포)을 제공하는, 예를 들어 수합하거나 단리하는 단계;

[0883] b) 선택적으로 적어도 하나의 줄기세포 확장제를 포함하는 유효량의 조성물의 존재 하에 생체외에서, 예를 들어 세포 배양 배지에서 상기 세포를 제공하여, 이로써 HSPC(예를 들어 CD34+ 세포)의 상기 집단을 비처리된 집단보다 큰 정도까지 확장시키는 단계;

[0884] c) HSPC(예를 들어 CD34+ 세포)의 집단을 유효량의: 본원에 기재된 표적화 도메인, 예를 들어 표 1에 기재된 표적화 도메인을 포함하는 적어도 하나의 gRNA 분자, 또는 상기 gRNA 분자를 인코딩하는 핵산, 및 예를 들어 본원에 기재된 적어도 하나의 Cas9 분자, 또는 상기 cas9 분자를 인코딩하는 핵산, 예를 들어 본원에 기재된 바와 같은 하나 이상의 RNP를 포함하는 조성물과, 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계;

[0885] d) 집단의 세포 중 적어도 일부(예를 들어 집단의 HSPC, 예를 들어 CD34+ 세포 중 적어도 일부)에서 적어도 하나의 변형을 유발하여, 이로써 예를 들어 상기 HSPC가 적혈구계 계통의 세포, 예를 들어 적혈구로 분화될 때, 태아 헤모글로빈 발현이 예를 들어 단계 c)에 따라 접촉되지 않은 세포에 비해 증가되는 단계; 및

[0886] f) 상기 변형된 HSPC(예를 들어 CD34+ 세포)를 포함하는 세포 집단을 포유류에게 복귀시키는 단계

[0887] 를 포함한다.

[0888] 일 양태에서, HSPC는 이들 HSPC가 복귀되는 포유류에게 동종이계이다. 일 양태에서, HSPC는 이들 HSPC가 복귀되는 포유류에게 자가이다. 양태에서, HSPC는 골수로부터 단리된다. 양태에서, HSPC는 말초 혈액, 예를 들어 동원된 말초 혈액으로부터 단리된다. 양태에서, 동원된 말초 혈액은 G-CSF가 투여된 대상체로부터 단리된다. 양태에서, 동원된 말초 혈액은 G-CSF 이외의 동원제, 예를 들어, Plerixafor®(AMD3100)가 투여된 대상체로부터 단리된다. 다른 양태에서, 동원된 말초 혈액은 G-CSF와 Plerixafor®(AMD3100)의 조합이 투여된 대상체로부터 단리된다. 양태에서, HSPC는 제대혈로부터 단리된다. 실시형태에서, 세포는 혈색소병증 환자, 예를 들어 겸상 적혈구 질병 환자 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아 환자로부터 유래된다.

[0889] 방법의 추가의 실시형태에서, 방법은, 예를 들어 상기 공급원으로부터 HSPC(예를 들어 CD34+ 세포)의 집단을 제공한 후에, HSPC(예를 들어 CD34+ 세포)에 대하여 세포 집단을 농화하는 단계를 추가로 포함한다. 방법의 실시형태에서, 상기 농화 후, 세포, 예를 들어 HSPC의 집단은 실질적으로 CD34- 세포가 없다.

[0890] 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 70%의 살아있는 세포를 포함한다. 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 75%의 살아있는 세포를 포함한다. 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 80%의 살아있는 세포를 포함한다. 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 85%의 살아있는 세포를 포함한다. 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 90%의 살아있는 세포를 포함한다. 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 95%의 살아있는 세포를 포함한다. 실시형태에서, 포유류로 복귀되는 세포 집단은 적어도 99%의 살아있는 세포를 포함한다. 생존능은 예를 들어 당분야에 공지된 바와 같이 세포 생존능 마커에 대해 세포 집단의 대표적인 부분을 염색함으로써 결정될 수 있다.

[0891] 또 다른 실시형태에서, 본 개시내용은 하기 방법을 필요로 하는 포유류, 예를 들어 인간에서 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 베타-탈라쎄미아를 치료하는 방법, 또는 태아 헤모글로빈 발현을 증가시키는 방법을 제공하며, 상기 방법은:

[0892] a) 포유류, 예를 들어 포유류의 골수로부터 HSPC의 집단(예를 들어 CD34+ 세포)을 제공하는, 예를 들어 수합하거나 단리하는 단계;

[0893] b) 단계 a)의 세포 집단으로부터 CD34+ 세포를 단리하는 단계;

[0894] c) 상기 CD34+ 세포를 생체외에서 제공하고, 약 0.5 내지 약 0.75 마이크로몰 농도의 적어도 하나의 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 유효량의 조성물의 존재 하에 상기 세포를 예를 들어 세포 배양 배지에서 배양하여, 이로써 CD34+ 세포의 상기 집단을 비처리된 집단보다 큰 정도까지 확장시키는 단계;

[0895] d) CD34+ 세포의 집단의 세포를 유효량의: 예를 들어 본원에 기재된 바와 같은 Cas9 분자 및 예를 들어 본원에 기재된 바와 같은 gRNA 분자를 포함하는 조성물 내로 도입하는 단계로서, 예를 들어 선택적으로 Cas9 분자 및 gRNA 분자가 예를 들어 본원에 기재된 바와 같은 RNP 형태로 존재하고, 선택적으로 상기 도입이 예를 들어 본원

에 기재된 바와 같이 상기 세포 내로의 상기 RNP의 전기천공에 의한 것인 단계;

[0896] e) 집단의 세포 중 적어도 일부(예를 들어 집단의 HSPC, 예를 들어 CD34+ 세포 중 적어도 일부)에서 적어도 하나의 유전자 변형을 유발하여, 이로써 예를 들어 본원에 기재된 바와 같은 삽입결실이 단계 d)에서 도입된 gRNA의 표적화 도메인에 상보적인 게놈 부위에서 또는 그 부근에서 생성되는 단계;

[0897] f) 선택적으로, 부가적으로 상기 도입 후, 약 0.5 내지 약 0.75 마이크로몰 농도의 적어도 하나의 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 유효량의 조성물의 존재 하에 상기 세포를 예를 들어 세포 배양 배지에서 배양하여, 세포를 적어도 2배, 예를 들어, 적어도 4배, 예를 들어, 적어도 5배 확장시키는 단계;

[0898] g) 상기 세포를 동결보존시키는 단계; 및

[0899] h) 세포를 포유류에게 복귀시키는 단계

[0900] 를 포함하고, 여기서,

[0901] 포유류에게 복귀되는 세포는 1) 적혈구계 계통의 세포, 예를 들어 적혈구로 분화하는 능력을 유지하며; 2) 적혈루로 분화된 경우, 예를 들어 단계 e)의 gRNA에 의해 비변형된 세포에 비해 증가된 수준의 태아 헤모글로빈을 생산하는, 예를 들어 1개 세포 당 적어도 6 피코그램의 태아 헤모글로빈을 생산하는 세포를 포함한다.

[0902] 일 양태에서, HSPC는 이들 HSPC가 복귀되는 포유류에게 동종이계이다. 일 양태에서, HSPC는 이들 HSPC가 복귀되는 포유류에게 자가이다. 양태에서, HSPC는 골수로부터 단리된다. 양태에서, HSPC는 말초 혈액, 예를 들어 동원된 말초 혈액으로부터 단리된다. 양태에서, 동원된 말초 혈액은 G-CSF가 투여된 대상체로부터 단리된다. 양태에서, 동원된 말초 혈액은 G-CSF 이외의 동원제, 예를 들어 Plerixafor®(AMD3100)가 투여된 대상체로부터 단리된다. 다른 양태에서, 동원된 말초 혈액은 G-CSF와 Plerixafor®(AMD3100)의 조합이 투여된 대상체로부터 단리된다. 양태에서, HSPC는 제대혈로부터 단리된다. 실시형태에서, 세포는 혈색소병증 환자, 예를 들어 겸상 적혈구 질병 환자 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아 환자로부터 유래된다.

[0903] 상기 방법의 실시형태에서, 열거된 단계 b)는 실질적으로 CD34- 세포가 없는 세포 집단을 야기한다.

[0904] 방법의 추가의 실시형태에서, 방법은, 예를 들어 상기 공급원으로부터 HSPC(예를 들어 CD34+ 세포)의 집단을 제공한 후에, HSPC(예를 들어 CD34+ 세포)에 대하여 세포 집단을 농화하는 단계를 추가로 포함한다.

[0905] 이들 방법의 추가의 실시형태에서, 분화하는 능력을 갖고, 증가된 태아 헤모글로빈 발현을 갖는 변형된 HSPC(예를 들어 CD34+ 줄기 세포)의 집단은 포유류 내로 재도입되기 전에 동결보존 및 저장된다. 실시형태에서, 적혈구계 계통의 세포, 예를 들어 적혈구로 분화하는 능력을 갖고/갖거나 적혈구계 계통의 세포, 예를 들어 적혈구로 분화될 때, 증가된 수준의 태아 헤모글로빈을 생산하는 동결보존된 HSPC의 집단은 해동되어 포유류로 재도입된다. 이들 방법의 추가의 실시형태에서, 방법은 포유류의 내인성 조혈 전구 또는 줄기 세포를 제거 또는 감소시키기 위하여 화학 요법 및/또는 방사선 요법을 포함한다. 이들 방법의 추가의 실시형태에서, 방법은 포유류의 내인성 조혈 전구 또는 줄기 세포를 제거 또는 감소시키기 위하여 화학 요법 및/또는 방사선 요법의 것을 포함하지 않는다. 이들 방법의 추가의 실시형태에서, 방법은 포유류의 내인성 조혈 전구 또는 줄기 세포를 부분적으로 감소시키기 위하여(예를 들어 부분적 림프구 감소) 화학 요법 및/또는 방사선 요법을 포함한다. 실시형태에서, 환자는 포유류로의 변형된 HSPC의 재도입 전에, 완전히 림프구 감소시키는 투여량의 부술판(busulfan)으로 처리된다. 실시형태에서, 환자는 포유류로의 변형된 HSPC의 재도입 전에, 부분적으로 림프구 감소시키는 투여량의 부술판으로 처리된다.

[0906] 실시형태에서, 세포는 예를 들어 본원에 기재된 바와 같은(예를 들어 표 1에 열거된 표적화 도메인을 포함하는) 비결실 HPFH 영역에 대한 gRNA와 복합체화된, 예를 들어 본원에 기재된 바와 같은 Cas9 분자를 포함하는 RNP와 접촉된다.

[0907] 실시형태에서, 줄기세포 확장제는 화합물 1이다. 실시형태에서, 줄기세포 확장제는 화합물 2이다. 실시형태에서, 줄기세포 확장제는 화합물 3이다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이고, 2 내지 0.1 마이크로몰, 예를 들어 1 내지 0.25 마이크로몰, 예를 들어 0.75 내지 0.5 마이크로몰의 농도로 존재한다. 실시형태에서, 줄기세포 확장제는 WO2010/059401에 기재된 분자(예를 들어 WO2010/059401의 실시예 1에 기재된

분자)이다.

[0908] 실시형태에서, 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 HSPC는 이러한 세포를 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계 전에, 약 1시간 내지 약 15일의 기간, 예를 들어 약 12시간 내지 약 12일의 기간, 예를 들어 약 12시간 내지 4일의 기간, 예를 들어 약 1일 내지 약 4일의 기간, 예를 들어 약 1일 내지 약 2일의 기간, 예를 들어 약 1일의 기간 또는 약 2일의 기간 동안 생체외에서 배양된다. 실시형태에서, 상기 접촉 단계 전에 상기 배양은 예를 들어 본원에 기재된 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.25 uM 내지 약 1 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.75 내지 0.5 마이크로몰 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 조성물(예를 들어 세포 배양 배지)에서이다. 실시형태에서, 세포는 이러한 세포를 예를 들어 본원에 기재된 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.25 uM 내지 약 1 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.75 내지 약 0.5 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 예를 들어 세포 배양 배지에서 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계 후, 약 1일 이하, 예를 들어 약 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1시간(들) 이하의 기간 동안 생체외에서 배양된다. 다른 실시형태에서, 세포는 이러한 세포를 예를 들어 본원에 기재된 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.25 uM 내지 약 1 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.75 내지 약 0.5 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 예를 들어 세포 배양 배지에서 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계 후, 약 1시간 내지 약 15일의 기간, 예를 들어 약 12시간 내지 약 10일의 기간, 예를 들어 약 1일 내지 약 10일의 기간, 예를 들어 약 1일 내지 약 5일의 기간, 예를 들어 약 1일 내지 약 4일의 기간, 예를 들어 약 2일 내지 약 4일의 기간, 예를 들어 약 2일, 약 3일 또는 약 4일의 기간 동안 생체외에서 배양된다. 실시형태에서, 세포는 약 1시간 내지 약 20일의 기간, 예를 들어 약 6 내지 12일의 기간, 예를 들어 약 6, 약 7, 약 8, 약 9, 약 10, 약 11 또는 약 12일의 기간 동안 생체외에서 배양된다(예를 들어 상기 접촉 단계 전에 배양되고/되거나 상기 접촉 단계 후에 배양된다).

[0909] 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 약 $1x10^6$개 세포(예를 들어 적어도 약 $1x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 약 $2x10^6$개 세포(예를 들어 적어도 약 $2x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 약 $3x10^6$개 세포(예를 들어 적어도 약 $3x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 약 $4x10^6$개 세포(예를 들어 적어도 약 $4x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 약 $5x10^6$개 세포(예를 들어 적어도 약 $5x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 약 $6x10^6$개 세포(예를 들어 적어도 약 $6x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 $1x10^6$개 세포(예를 들어 적어도 $1x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 $2x10^6$개 세포(예를 들어 적어도 $2x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 $3x10^6$개 세포(예를 들어 적어도 $3x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 $4x10^6$개 세포(예를 들어 적어도 $4x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 $5x10^6$개 세포(예를 들어 적어도 $5x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 적어도 $6x10^6$개 세포(예를 들어 적어도 $6x10^6$개

CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 약 $1x10^6$개 세포(예를 들어 약 $1x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 약 $2x10^6$개 세포(예를 들어 약 $2x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 약 $3x10^6$개 세포(예를 들어 약 $3x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 약 $4x10^6$개 세포(예를 들어 약 $4x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 약 $5x10^6$개 세포(예를 들어 약 $5x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 1 kg 당 약 $6x10^6$개 세포(예를 들어 약 $6x10^6$개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 환자 체중 1 kg 당 약 2 x 10^6개 세포(예를 들어 약 2 x 10^6개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 환자 체중 1 kg 당 적어도 2 x 10^6개 세포(예를 들어 약 2 x 10^6개 CD34+ 세포)를 포함한다. 실시형태에서, 포유류에게 복귀되는 변형된 HSPC를 포함하는 세포 집단은 환자 체중 1 kg 당 2 x 10^6개 세포(예를 들어 약 2 x 10^6개 CD34+ 세포) 내지 10 x 10^6개 세포(예를 들어 약 2 x 10^6개 CD34+ 세포)를 포함한다. 실시형태에서, 변형된 세포를 포함하는 세포는 환자 내에 주입된다. 실시형태에서, 변형된 HSPC를 포함하는 세포가 환자에게 주입되기 전에, 환자는 림프구결실 요법(lymphodepleting therapy)으로 치료받거나, 예를 들어 부술판으로 치료받거나, 예를 들어 완전 림프구결실 부술판 투약 방식(regimen)으로 치료받거나, 예를 들어 감소된 강도의 부술판 림프구결실 투약 방식으로 치료받는다.

[0910] 실시형태에서, 상기 기재된 임의의 방법은 예를 들어 포유류에게 세포를 재도입한 후 적어도 15일, 예를 들어, 적어도 20, 적어도 30, 적어도 40 적어도 50 또는 적어도 60일째에 측정된 바와 같이, 방법에 사용된 gRNA 분자의 표적화 도메인에 상보적인 게놈 부위에서 또는 그 분근에서 삽입결실을 포함하는 이의 순환성 CD34+ 세포 중 적어도 80%를 갖는 환자를 초래한다. 이론으로 결부시키고자 하는 것은 아니지만, 놀랍게도, 예를 들어 본원에 기재된 바와 같은 유전자 편집 시스템, 예를 들어 CRISPR 시스템, 예를 들어 HBG1 및/또는 HBG2 영역을 표적화하는 gRNA 분자를 포함하는 CRISPR 시스템이 HSPC 내로 도입되고 이들 세포가 유기체 내로 이식될 때 관찰된 삽입결실 및 삽입결실 패턴(대형 결실을 포함함)에서, 소정의 gRNA는 편집된 세포 집단 및 이의 자손, 유기체에서 검출 가능한 채로 남아 있고 8주, 12주, 16주 또는 20주 초과 동안 지속되는 삽입결실 및 삽입결실 패턴(대형 삽입결실을 포함함)을 포함하는 세포를 생산하는 것으로 발견되었다. 이론으로 결부시키고자 하는 것은 아니지만, 유기체(예를 들어 환자) 내로의 도입 후 예를 들어 16주보다 길게 또는 20주보다 길게 검출 가능한 세포 집단 내에서 지속되는 삽입결실 패턴 또는 특정 삽입결실(대형 결실을 포함함)을 포함하는 세포 집단은, 유기체 또는 환자 내로 도입 시 하나 이상의 삽입결실(대형 결실을 포함함)을 상실하는 세포(또는 이들의 자손)보다 예를 들어 본원에 기재된 질병 또는 질환(예를 들어 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아)의 장기간 완화를 생산하는 데 유익할 수 있을 것이다. 실시형태에서, 지속적 삽입결실 또는 삽입결실 패턴은 상향 조절된 태아 헤모글로빈(예를 들어 상기 세포의 적혈구계 자손에서)과 연관이 있다. 따라서, 실시형태에서, 본 개시내용은 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 HSPC의 집단을 제공하며, 이러한 집단은 유기체, 예를 들어 환자 내로의 도입 후 8주 초과, 12주 초과, 16주 초과 또는 20주 초과 동안 혈액 및/또는 골수)에서 지속되는(예를 들어 세포 집단 또는 이의 자손에서 예를 들어 검출 가능한 채로 남아 있는) 하나 이상의 삽입결실(대형 결실을 포함함)을 포함한다.

[0911] 실시형태에서, 상기 기재된 임의의 방법은 예를 들어 포유류 내로 세포의 재도입 후 적어도 15일, 예를 들어, 적어도 20, 적어도 30, 적어도 40 적어도 50 또는 적어도 60일째에 측정된 바와 같이 이러한 방법에 사용되는 gRNA 분자의 표적화 도메인에 상보적인 게놈 부위에서 또는 그 부근에서 삽입결실을 포함하는 환자의 골수 CD34+ 세포 중 적어도 20%를 갖는 환자를 초래한다.

[0912] 실시형태에서, 포유류 내에 재도입되는 HSPC는 생체내에서 적혈구계 계통의 세포, 예를 들어 적혈구로 분화할 수 있고, 상기 분화된 세포는 증가된 태아 헤모글로빈 수준을 나타내며, 예를 들어 1개 세포 당 적어도 6 피코그램의 태아 헤모글로빈, 예를 들어 1개 세포 당 적어도 7 피코그램의 태아 헤모글로빈, 1개 세포 당 적어도 8 피코그램의 태아 헤모글로빈, 1개 세포 당 적어도 9 피코그램의 태아 헤모글로빈, 1개 세포 당 적어도 10 피코그램의 태아 헤모글로빈, 예를 들어 1개 세포 당 약 9 내지 약 10 피코그램의 태아 헤모글로빈을 포함하며, 따라서 예를 들어 혈색소병증이 포유류에서 치료된다.

[0913] 세포가 증가된 태아 헤모글로빈을 갖는 것을 특징으로 하는 경우, 해당 세포의 자손, 예를 들어 분화된 자손이 증가된 태아 헤모글로빈을 나타내는 실시형태를 포함하는 것으로 이해될 것이다. 예를 들어 본원에 기재된 방법에서, 변경된 또는 변형된 CD34+ 세포(또는 세포 집단)는 증가된 태아 헤모글로빈을 발현할 수 없으나, 적혈구계 계통의 세포, 예를 들어 적혈구로 분화 시, 세포는 증가된 태아 헤모글로빈, 예를 들어 유사한 조건 하에 비변형된 또는 비변경된 세포에 비해 증가된 태아 헤모글로빈을 발현한다.

[0914] **XII. 세포의 배양 방법 및 제조 방법**

[0915] 본 개시내용은 본원에 기재된 gRNA 분자를 이용하여 변형된 또는 변형되는 세포, 예를 들어 HSPC, 예를 들어 조혈 줄기세포, 예를 들어 CD34+ 세포를 배양하는 방법을 제공한다.

[0916] **DNA 수선 경로 저해제**

[0917] 이론으로 결부시키고자 하는 것은 아니지만, 특정 표적 서열에서 주어진 gRNA 분자에 의해 생산된 삽입결실 패턴은 세포 내에서 각각의 활성 DNA 수선 기전(예를 들어 비-상동성 말단 접합, 미세상동성-매개 말단 접합 등)의 생산물인 것으로 여겨진다. 이론으로 결부시키고자 하는 것은 아니지만, 편집되는 세포를, 요망되는 삽입결실을 생산하지 않는 DNA 수선 경로의 저해제와 접촉시킴으로써, 특히 선호할 만한 삽입결실이 선택되거나 농화될 수 있는 것으로 여겨진다. 따라서, 본 발명의 gRNA 분자, CRISPR 시스템, 방법 및 다른 양태는 이러한 저해제와 조합하여 수행될 수 있다. 이러한 저해제의 예는 예를 들어 WO2014/130955에 기재된 것들을 포함하며, 이의 내용은 그 전체가 참조로 본 명세서에 포함된다. 실시형태에서, 저해제는 DNAPKc 저해제, 예를 들어 NU7441이다.

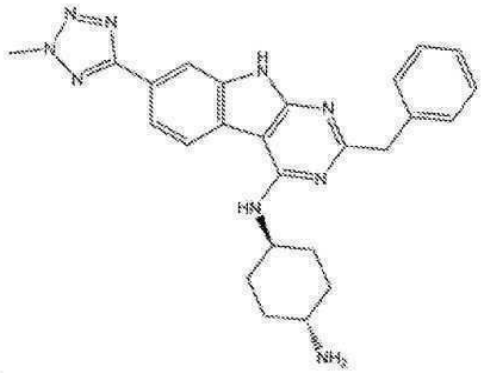
[0918] **줄기세포 확장제**

[0919] 일 양태에서, 본 발명은 변형된 또는 변형되는 세포, 예를 들어 HSPC, 예를 들어 CD34+ 세포를, 작용제로 처리되지 않은 세포에 비해 증가된 확장율, 증가된 확장 수준 또는 증가된 생착을 초래하는 하나 이상의 작용제와 함께 배양하는 것에 관한 것이다. 이러한 작용제는 본원에서 줄기세포 확장제로 지칭된다.

[0920] 일 양태에서, 작용제, 예를 들어 줄기세포 확장제로 처리되지 않은 세포에 비해 증가된 확장율 또는 증가된 확장 수준을 초래하는 하나 이상의 작용제는 아릴 탄화수소 수용체(AHR) 경로의 저해제인 작용제를 포함한다. 양태에서, 줄기세포 확장제는 WO2013/110198에 개시된 화합물 또는 WO2010/059401에 개시된 화합물이고, 이들의 내용은 그 전체가 참조로 포함된다.

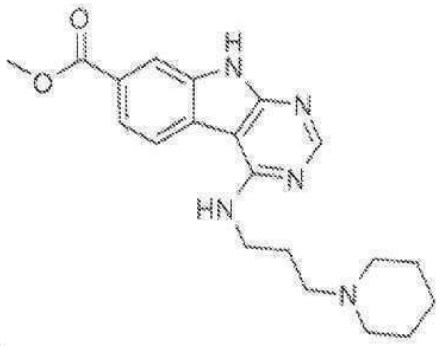
[0921] 일 양태에서, 작용제로 처리되지 않은 세포에 비해 증가된 확장율 또는 증가된 확장 수준을 초래하는 하나 이상의 작용제는 예를 들어 WO2013/110198에 개시된 바와 같은 피리미도[4,5-b]인돌 유도체이며, 이의 내용은 그 전체가 참조로 본 명세서에 포함된다. 일 실시형태에서, 이러한 작용제는 화합물 1((1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민)이다:

[0922] **화합물 1**

[0923] 

[0924] 또 다른 양태에서, 이러한 작용제는 화합물 2(메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트)이다:

[0925] **화합물 2:**

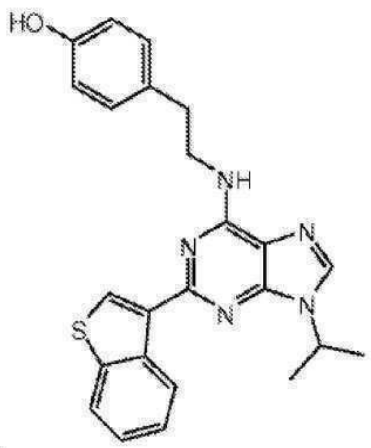


[0926]

[0927] 또 다른 양태에서, 작용제로 처리되지 않은 세포에 비해 증가된 확장율 또는 증가된 확장 수준을 초래하는 하나 이상의 작용제는 예를 들어 WO2010/059401에 개시된 작용제이며, 이의 내용은 그 전체가 참조로 본 명세서에 포함된다.

[0928] 일 실시형태에서, 줄기세포 확장제는 화합물 3: 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀이며, 즉, 하기 구조를 갖는, WO2010/059401의 실시예 1의 화합물이다:

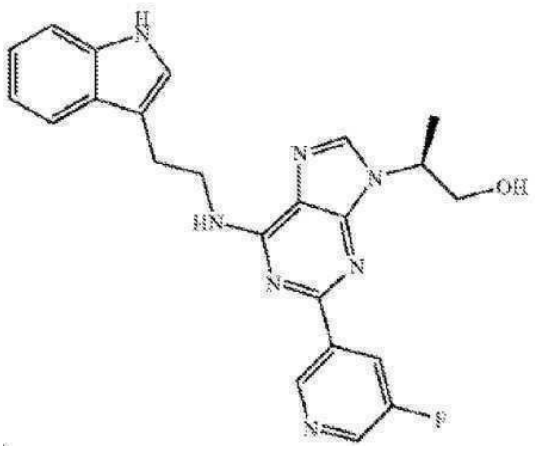
[0929] **화합물 3:**



[0930]

[0931] 또 다른 양태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올((S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이며, 즉, 하기 구조를 갖는, WO2010/059401에 따른 화합물 157S이다:

[0932] **(S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올:**



[0933]

[0934] 실시형태에서, HSPC의 집단은 CRISPR 시스템(예를 들어 본 발명의 gRNA 분자 및/또는 Cas9 분자)을 상기 HSPC 내에 도입하기 전에, 줄기세포 확장제, 예를 들어 화합물 1, 화합물 2, 화합물 3, (S)-2-(6-(2-(1H-인돌-3-일)

에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올 또는 이들의 조합(예를 들어 화합물 1과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)과 접촉된다. 실시형태에서, HSPC의 집단은 CRISPR 시스템(예를 들어 본 발명의 gRNA 분자 및/또는 Cas9 분자)을 상기 HSPC 내에 도입한 후에, 줄기세포 확장제, 예를 들어 화합물 1, 화합물 2, 화합물 3, (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올 또는 이들의 조합(예를 들어 화합물 1과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)과 접촉된다. 실시형태에서, HSPC의 집단은 CRISPR 시스템(예를 들어 본 발명의 gRNA 분자 및/또는 Cas9 분자)을 상기 HSPC 내에 도입하기 전과 후 모두에서, 줄기세포 확장제, 예를 들어 화합물 1, 화합물 2, 화합물 3, (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올 또는 이들의 조합(예를 들어 화합물 1과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)과 접촉된다.

[0935] 실시형태에서, 줄기세포 확장제는 동일한 배지이지만 줄기세포 확장제의 부재 하의 HSPC에 비해, HSPC의 확장 수준을 증가시키기에 효과적인 양으로 존재한다. 실시형태에서, 줄기세포 확장제는 약 0.01 내지 약 10 uM, 예를 들어 약 0.1 uM 내지 약 1 uM 범위의 농도로 존재한다. 실시형태에서, 줄기세포 확장제는 세포 배양 배지 내에서 약 1 uM, 약 950 nM, 약 900 nM, 약 850 nM, 약 800 nM, 약 750 nM, 약 700nM, 약 650 nM, 약 600 nM, 약 550 nM, 약 500 nM, 약 450 nM, 약 400 nM, 약 350 nM, 약 300 nM, 약 250 nM, 약 200 nM, 약 150 nM, 약 100 nM, 약 50 nM, 약 25 nM, 또는 약 10 nM의 농도로 존재한다. 실시형태에서, 줄기세포 확장제는 약 500 nM 내지 약 750 nM 범위의 농도로 존재한다.

[0936] 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이며, 이러한 확장제는 세포 배양 배지 내에서 약 0.01 내지 약 10 마이크로몰(uM; micromolar) 범위의 농도로 존재한다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이며, 이러한 확장제는 세포 배양 배지 내에서 약 0.1 내지 약 1 마이크로몰(uM) 범위의 농도로 존재한다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이며, 이러한 확장제는 세포 배양 배지 내에서 약 0.75 마이크로몰(uM)의 농도로 존재한다. 실시형태에서, 줄기세포 확장제는 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올이며, 이러한 확장제는 세포 배양 배지 내에서 약 0.5 마이크로몰(uM)의 농도로 존재한다. 임의의 상기 실시형태에서, 세포 배양 배지는 부가적으로 화합물 1을 포함한다.

[0937] 실시형태에서, 줄기세포 확장제는 화합물 1과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 혼합물이다.

[0938] 실시형태에서, 본 발명의 세포는 CD34+ 세포의 2 내지 10,000배 확장, 예를 들어 CD34+ 세포의 2 내지 1000배 확장, 예를 들어 CD34+ 세포의 2 내지 100배 확장, 예를 들어 CD34+ 세포의 20 내지 200배 확장을 유발하기에 충분한 시간 및 충분한 시간으로 하나 이상의 줄기세포 확장제 분자와 접촉된다. 본원에 기재된 바와 같이, 하나 이상의 줄기세포 확장제와의 접촉은, 세포가 예를 들어 본원에 기재된 바와 같은 CRISPR 시스템과 접촉되기 전에, 세포가 예를 들어 본원에 기재된 바와 같은 CRISPR 시스템과 접촉된 후에, 또는 이들의 조합 시 존재할 수 있다. 일 실시형태에서, 세포는 CD34+ 세포, 예를 들어 상기 세포 내로 도입된 CRISPR/Cas9 시스템의 gRNA의 표적화 도메인에 대해 상보성을 갖는 표적 부위에서 또는 그 부근에서 삽입결실을 포함하는 CD34+ 세포의 적어도 2배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올과 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포, 예를 들어 상기 세포 내로 도입된 CRISPR/Cas9 시스템의 gRNA의 표적화 도메인에 대해 상보성을 갖는 표적 부위에서 또는 그 부근에서 삽입결실을 포함하는 CD34+ 세포의 적어도 4배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올과 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포, 예를 들어 상기 세포 내로 도입된 CRISPR/Cas9 시스템의 gRNA의 표적화 도메인에 대해 상보성을 갖는 표적 부위에서 또는 그 부근에서 삽입결실을 포함하는 CD34+ 세포의 적어도 5배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올과 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포의 적어도 10배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나

이상의 줄기세포 확장제 분자와 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포의 적어도 20배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자와 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포의 적어도 30배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자와 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포의 적어도 40배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자와 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포의 적어도 50배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자와 접촉된다. 일 실시형태에서, 세포는 CD34+ 세포의 적어도 60배 확장을 유발하기에 충분한 시간 동안 및 충분한 양으로 하나 이상의 줄기세포 확장제 분자와 접촉된다. 실시형태에서, 세포는 약 1 내지 60일, 예를 들어 약 1 내지 50일, 예를 들어 약 1 내지 40일, 예를 들어 약 1 내지 30일, 예를 들어, 1 내지 20일, 예를 들어 약 1 내지 10일, 예를 들어 약 7일, 예를 들어 약 1 내지 5일, 예를 들어 약 2 내지 5일, 예를 들어 약 2 내지 4일, 예를 들어 약 2일 또는 예를 들어 약 4일의 기간 동안 하나 이상의 줄기세포 확장제 분자와 접촉된다.

[0939] 실시형태에서, 예를 들어 본원에 기재된 바와 같은 세포, 예를 들어 HSPC는 이러한 세포를 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계 전에, 약 1시간 내지 약 10일의 기간, 예를 들어 약 12시간 내지 약 5일의 기간, 예를 들어 약 12시간 내지 4일의 기간, 예를 들어 약 1일 내지 약 4일의 기간, 예를 들어 약 1일 내지 약 2일의 기간, 예를 들어 약 1일의 기간 또는 약 2일의 기간 동안 생체외에서 배양된다. 실시형태에서, 상기 접촉 단계 전의 상기 배양은 예를 들어 본원에 기재된 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.25 uM 내지 약 1 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.75 내지 0.5 마이크로몰 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 조성물(예를 들어 세포 배양 배지)에서이다. 실시형태에서, 세포는 이러한 세포를 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계 후, 예를 들어 본원에 기재된 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.25 uM 내지 약 1 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.75 내지 0.5 마이크로몰 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 예를 들어 세포 배양 배지에서 약 1일 이하, 예를 들어 약 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1시간(들) 이하의 기간 동안 생체외에서 배양된다. 다른 실시형태에서, 세포는 이러한 세포를 예를 들어 본원에 기재된 CRISPR 시스템과 접촉시키는 단계 후, 예를 들어 본원에 기재된 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.25 uM 내지 약 1 uM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 약 0.75 내지 0.5 마이크로몰 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 예를 들어 세포 배양 배지에서 약 1시간 내지 약 14일의 기간, 예를 들어 약 12시간 내지 약 10일의 기간, 예를 들어 약 1일 내지 약 10일의 기간, 예를 들어 약 1일 내지 약 5일의 기간, 예를 들어 약 1일 내지 약 4일의 기간, 예를 들어 약 2일 내지 약 4일의 기간, 예를 들어 약 2일, 약 3일 또는 약 4일의 기간 동안 생체외에서 배양된다.

[0940] 실시형태에서, 세포 배양 배지는 화학적으로 정의된 배지이다. 실시형태에서, 세포 배양 배지는 부가적으로, 예를 들어 StemSpan SFEM(Stemcell Technologies; Cat no. 09650)을 함유할 수 있다. 실시형태에서, 세포 배양 배지는 대안적으로 또는 부가적으로 예를 들어 HSC Brew, GMP(Miltenyi)를 함유할 수 있다. 실시형태에서, 세포 배양 배지는 혈청-무함유이다. 실시형태에서, 배지에 트롬보포이에틴(Tpo), 인간 Flt3 리간드(Flt-3L), 인간 줄기세포 인자(SCF), 인간 인터루킨-6, L-글루타민, 및/또는 페니실린/스트렙토마이신이 보충될 수 있다. 실시형태에서, 배지에 트롬보포이에틴(Tpo), 인간 Flt3 리간드(Flt-3L), 인간 줄기세포 인자(SCF), 인간 인터루킨-6 및 L-글루타민이 보충된다. 다른 실시형태에서, 배지에 트롬보포이에틴(Tpo), 인간 Flt3 리간드(Flt-3L), 인간 줄기세포 인자(SCF) 및 인간 인터루킨-6가 보충된다. 다른 실시형태에서, 배지에 트롬보포이에틴(Tpo), 인간 Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)가 보충되지만, 인간 인터루킨-6은 보충되지 않는다. 다른 실시형태에서, 배지에 인간 Flt3 리간드(Flt-3L), 인간 줄기세포 인자(SCF)가 보충되지만, 인간 트롬보포이에틴(Tpo) 또는 인간 인터루킨-6은 보충되지 않는다. 트롬보포이에틴(Tpo), 인간 Flt3 리간드(Flt-3L), 인간 줄기세포 인자(SCF), 인간 인터루킨-6, 및/또는 L-글루타민이 배지에 존재할 때, 이들은 각각 약 1 ng/mL 내지 약 1000 ng/mL 범위의 농도, 예를 들어 약 10 ng/mL 내지 약 500 ng/mL 범위의 농도, 예를 들어 약 10 ng/mL 내지 약 100 ng/mL 범위의 농도, 예를 들어 약 25 ng/mL 내지 약 75 ng/mL 범위의 농도, 예를 들어 약 50 ng/mL의

농도로 존재한다. 실시형태에서, 각각의 보충된 구성성분은 동일한 농도로 존재한다. 다른 실시형태에서, 각각의 보충된 구성성분은 상이한 농도로 존재한다. 일 실시형태에서, 배지는 StemSpan SFEM(Stemcell Technologies; Cat no. 09650), 50 ng/mL의 트롬보포이에틴(Tpo), 50 ng/mL의 인간 Flt3 리간드(Flt-3L), 50 ng/mL의 인간 줄기세포 인자(SCF) 및 50 ng/mL의 인간 인터루킨-6(IL-6)을 포함한다. 일 실시형태에서, 배지는 StemSpan SFEM(Stemcell Technologies; Cat no. 09650), 50 ng/mL의 트롬보포이에틴(Tpo), 50 ng/mL의 인간 Flt3 리간드(Flt-3L) 및 50 ng/mL의 인간 줄기세포 인자(SCF)를 포함하고, IL-6은 포함하지 않는다. 실시형태에서, 배지는 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 0.75 μM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 추가로 포함한다. 실시형태에서, 배지는 줄기세포 확장제, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올, 예를 들어 0.5 μM 농도의 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 추가로 포함한다. 실시형태에서, 배지는 1% L-글루타민 및 2% 페니실린/스트렙토마이신을 추가로 포함한다. 실시형태에서, 세포 배양 배지는 혈청-무함유이다.

[0941] **XII. 병용 요법**

[0942] 본 개시는 하나 이상의 다른 치료 방식 및/또는 제제 제제와 조합하여, 본원에 기재된 gRNA 분자, 또는 본원에 기재된 gRNA 분자로 변형된 세포(예를 들어 조혈 줄기 세포, 예를 들어 CD34+ 세포)의 사용을 고려한다. 따라서, gRNA 분자 또는 본원에 기재된 gRNA 분자로 변형된 세포의 사용에 추가하여, 혈색소병증을 치료하기 위한 하나 이상의 "표준" 요법을 대상체에게 또한 투여할 수 있다.

[0943] 혈색소병증을 치료하기 위한 하나 이상의 추가적인 요법에는, 예를 들어 추가적인 줄기 세포 이식, 예를 들어 조혈 줄기 세포 이식이 포함될 수 있다. 줄기 세포 이식은 동종이계 또는 자가 유래일 수 있다.

[0944] 혈색소병증을 치료하기 위한 하나 이상의 추가적인 요법에는, 예를 들어 혈액 수혈 및/또는 철 킬레이트화(예를 들어 제거) 요법이 포함될 수 있다. 공지된 철 킬레이트화제에는 예를 들어 데페록사민 및 데페라시록스가 포함된다.

[0945] 혈색소병증을 치료하기 위한 하나 이상의 추가적인 요법에는, 예를 들어 엽산 보충제 또는 하이드록시우레아(예를 들어 5-하이드록시우레아)가 포함될 수 있다. 혈색소병증을 치료하기 위한 하나 이상의 추가적인 요법은 하이드록시우레아일 수 있다. 실시형태에서, 하이드록시우레아는 예를 들어 1일 10 내지 35 mg/kg, 예를 들어 1일 10 내지 20 mg/kg의 투여량으로 투여될 수 있다. 실시형태에서, 하이드록시우레아는 1일 10 mg/kg의 투여량으로 투여된다. 실시형태에서, 하이드록시우레아는 1일 10 mg/kg의 투여량으로 투여된다. 실시형태에서, 하이드록시우레아는 1일 20 mg/kg의 투여량으로 투여된다. 실시형태에서, 하이드록시우레아는 예를 들어 본원에 기재된, 본 발명의 세포(또는 세포 집단), 예를 들어 CD34+ 세포(또는 세포 집단) 전에 및/또는 후에 투여된다.

[0946] 하나 이상의 추가적인 치료제에는, 예를 들어 항-p-셀렉틴 항체, 예를 들어 SelG1(Selexys)가 포함될 수 있다. P-셀렉틴 항체는 예를 들어 그 각각의 내용의 전문이 본원에 포함된, PCT 공개 WO1993/021956, PCT 공개 WO1995/034324, PCT 공개 WO2005/100402, PCT 공개 WO2008/069999, 미국 특허 출원 공개 US2011/0293617, 미국 특허 제5800815호, 미국 특허 제6667036호, 미국 특허 제8945565호, 미국 특허 제8377440호 및 미국 특허 제9068001호에 기재되어 있다.

[0947] 하나 이상의 추가적인 제제에는 예를 들어 태아 헤모글로빈을 상향 조절하는 저분자가 포함될 수 있다. 이러한 분자의 예로는 TN1(예를 들어 본원에 참조로 포함된 문헌[Nam, T. et al., ChemMedChem 2011, 6, 777-780, DOI: 10.1002/cmdc.201000505]에 기재됨)이 포함된다.

[0948] 하나 이상의 추가적인 요법에는 또한 당분야에 공지된 방사선조사 또는 다른 골수 절제 요법이 포함될 수 있다. 이러한 요법의 예는 부술판이다. 이러한 추가적인 요법은 대상체로의 본 발명의 세포의 도입 전에 수행될 수 있다. 본원에 기재된 치료 방법(예를 들어 본원에 기재된 방법에 의해 변형된(예를 들어 HbF 생산을 증가시키기 위해 본원에 기재된 CRISPR 시스템으로 변형된) 세포(예를 들어 HSPC)의 투여를 포함하는 치료 방법) 실시형태에서, 방법은 골수 절제 단계를 포함하지 않는다. 실시형태에서, 방법은 부분적인 골수 절제 단계를 포함한다.

[0949] 본원에 기재된 요법(예를 들어 HSPC, 예를 들어 본원에 기재된 CRISPR 시스템을 사용하여 변형된 HSPC의 집단를 투여하는 것을 포함)은 또한 추가적인 치료제와 조합될 수 있다. 실시형태에서, 추가적인 치료제는 HDAC 저해제, 예를 들어 파노비노스탯(panobinostat)이다. 실시형태에서, 추가적인 치료제는 PCT 공개 WO2014/150256호에 기재된 화합물, 예를 들어 WO2014/150256의 표 1에 기재된 화합물, 예를 들어 GBT440이다. HDAC 저해제의

다른 예로는, 예를 들어 수베로일아닐리드 하이드록삼산(SAHA)이 포함된다. 하나 이상의 추가적인 제제에는 예를 들어 DNA 메틸화 저해제가 포함될 수 있다. 이러한 제제는 감소된 BCL11a 활성을 갖는 세포에서 HbF 유도를 증가시키는 것으로 나타났다(예를 들어 본원에 참조로 포함된 문헌[Jian Xu et al, Science 334, 993(2011); DOI: 0.1126/science.1211053] 참조). 다른 HDAC 저해제에는 당분야에 공지된 임의의 HDAC 저해제, 예를 들어 트리코스타틴 A, HC 독소, DACI-2, FK228, DACI-14, 데푸디신, DACI-16, 투바신, NK57, MAZ1536, NK125, 스크립타이드, 피록사미드, MS-275, ITF-2357, MCG-D0103, CRA-024781, CI-994, 및 LBH589(예를 들어 그 전문이 본원에 참고로 포함된 문헌[Bradner JE, et al., *PNAS*, 2010(vol. 107:28), 12617-12622] 참조).

[0950] 본원에 기재된 gRNA 분자, 또는 본원에 기재된 gRNA 분자로 변형된 세포(예를 들어 조혈 줄기 세포, 예를 들어 CD34+ 세포) 및 병용-치료제 또는 병용-요법은 동일한 제형으로 또는 별도로 투여될 수 있다. 별도의 투여의 경우, 본원에 기재된 gRNA 분자 또는 본원에 기재된 gRNA 분자로 변형된 세포는 병용-치료 또는 병용-요법 전, 후 또는 이와 동시에 투여될 수 있다. 하나의 제제는 다른 제제의 투여를 수 분 내지 수 주 범위의 간격 만큼 선행하거나 뒤따를 수 있다. 2 이상의 상이한 종류의 치료제가 대상체에게 별도로 개별적으로 적용되는 실시형태에서, 이들 상이한 종류의 제제가 표적 조직 또는 세포 상에 유리하게 조합된 효과를 여전히 발휘할 수 있도록, 일반적으로, 각 전달의 시간 사이에 상당한 시간의 기간이 만료되지 않았음을 보장할 것이다.

[0951] **XIII. 변형된 뉴클레오시드, 뉴클레오타이드, 및 핵산**

[0952] 변형된 뉴클레오시드 및 변형된 뉴클레오타이드는 핵산, 예를 들어 특히 gRNA에 존재할 수 있으나, 또한 다른 형태의 RNA, 예를 들어 mRNA, RNAi 또는 siRNA에도 존재할 수 있다. 본원에 기재된 "뉴클레오시드"는 5-탄당 분자(펜토스 또는 리보스) 또는 이의 유도체, 및 유기 염기, 퓨린 또는 피리미딘, 또는 이의 유도체를 함유하는 화합물로 정의된다. 본원에 기재된 "뉴클레오타이드"는 포스페이트기를 추가로 포함하는 뉴클레오시드로 정의된다.

[0953] 변형된 뉴클레오시드 및 뉴클레오타이드는:

[0954] (i) 포스포디에스테르 백본 연결 내의 하나 또는 둘 다의 비-연결 포스페이트 산소 및/또는 하나 이상의 연결 포스페이트 산소의 변경, 예를 들어 대체;

[0955] (ii) 리보스 당의 구성 요소, 예를 들어 리보스 당 상의 2' 하이드록실의 변경, 예를 들어 대체;

[0956] (iii) 포스페이트 모이어티의 "데포스포(dephospho)" 링커로의 대규모 대체;

[0957] (iv) 비정준 핵염기로의 변형 또는 치환을 포함하는, 자연적 발생 핵염기의 변형 또는 대체;

[0958] (v) 리보스-포스페이트 백본의 대체 또는 변형;

[0959] (vi) 올리고뉴클레오타이드의 3' 말단 또는 5' 말단의 변형, 예를 들어 말단 포스페이트기의 제거, 변형 또는 대체, 또는 모이어티, 캡 또는 링커의 접합; 및

[0960] (vii) 당의 변형 또는 대체; 중 하나 이상을 포함할 수 있다.

[0961] 위에서 나열된 변형들은 조합되어 2, 3, 4 이상의 변형을 가질 수 있는 변형된 뉴클레오시드 및 뉴클레오타이드를 제공할 수 있다. 예를 들어 변형된 뉴클레오시드 또는 뉴클레오타이드는 변형된 당 및 변형된 핵염기를 가질 수 있다. 실시형태에서, gRNA의 모든 염기가 변형되고, 예를 들어 모든 염기가 변형된 포스페이트기를 가지며, 예를 들어 모두가 포스포로티오에이트기이다. 실시형태에서, 단 분자 또는 모듈식 gRNA 분자의 포스페이트기의 전부 또는 실질적으로 전부가 포스포로티오에이트기로 대체된다. 실시형태에서, gRNA의 5개 3'-말단 염기 중 하나 이상 및/또는 gRNA의 5개 5'-말단 염기 중 하나 이상은 포스포로티오에이트기로 변형된다.

[0962] 실시형태에서, 변형된 뉴클레오타이드, 예를 들어 본원에 기재된 변형을 갖는 뉴클레오타이드는 핵산, 예를 들어 "변형된 핵산"으로 혼입될 수 있다. 일부 실시형태에서, 변형된 핵산은 1개, 2개, 3개 이상의 변형된 뉴클레오타이드를 포함한다. 일부 실시형태에서, 변형된 핵산 내의 위치의 적어도 5%(예를 들어 적어도 약 5%, 적어도 약 10%, 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30%, 적어도 약 35%, 적어도 약 40%, 적어도 약 45%, 적어도 약 50%, 적어도 약 55%, 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 또는 약 100%)는 변형된 뉴클레오타이드이다.

[0963] 변형되지 않은 핵산은 예를 들어 세포 뉴클레아제에 의한 분해되기 쉬울 수 있다. 예를 들어 뉴클레아제는 핵산 포스포디에스테르 결합을 가수분해할 수 있다. 따라서, 일 양태에서, 본원에 기재된 변형된 핵산은 예를 들어 뉴클레아제에 대한 안정성을 도입하기 위해 하나 이상의 변형된 뉴클레오시드 또는 뉴클레오타이드를 함유 할

수 있다.

[0964] 일부 실시형태에서, 본원에 기재된 변형된 뉴클레오시드, 변형된 뉴클레오타이드 및 변형된 핵산은 생체내 및 생체외 둘 다에서 세포 집단으로 도입될 때 감소된 선천 면역 반응을 나타낼 수 있다. 용어 "선천 면역 반응"은 일반적으로 바이러스 또는 박테리아 기원의 단일 가닥 핵산을 포함하는 외인성 핵산에 대한 세포 반응을 포함하며, 사이토카인 발현 및 방출, 특히 인터페론의 유도 및 세포 사멸을 수반한다. 일부 실시형태에서, 본원에 기재된 변형된 뉴클레오시드, 변형된 뉴클레오타이드 및 변형된 핵산은 주홈(major groove) 상호 작용 파트너와 핵산의 결합을 파괴할 수 있다. 일부 실시형태에서, 본원에 기재된 변형된 뉴클레오시드, 변형된 뉴클레오타이드 및 변형된 핵산은 생체내 및 생체외 둘 다에서 세포 집단으로 도입될 때 감소된 선천 면역 반응을 나타낼 수 있고, 또한, 주홈 상호 작용 파트너와 핵산의 결합을 파괴 할 수 있다.

[0965] 화학적 기의 정의

[0966] 본원에 사용되는 "알킬"은 직쇄 또는 분지형인 포화 탄화수소기를 지칭하는 것으로 여겨진다. 알킬기의 예로는 메틸(Me), 에틸(Et), 프로필(예를 들어 n-프로필 및 이소프로필), 부틸(예를 들어 n-부틸, 이소부틸, t-부틸), 펜틸(예를 들어 n-펜틸, 이소펜틸, 네오펜틸) 등이 포함된다. 알킬기는 1개 내지 약 20개, 2개 내지 약 20개, 1개 내지 약 12개, 1개 내지 약 8개, 1개 내지 약 6개, 1개 내지 약 4개, 또는 1개 내지 약 3개의 탄소 원자를 함유할 수 있다.

[0967] 본원에 사용되는 "아릴"은 모노사이클릭 또는 폴리사이클릭(예를 들어 2개, 3개 또는 4개의 융합된 고리를 갖는) 방향족 탄화수소, 예를 들어 페닐, 나프틸, 안트라세닐, 페난트레닐, 인다닐, 인데닐 등을 지칭한다.

[0968] 실시형태에서, 아릴기는 6개 내지 약 20개의 탄소 원자를 갖는다.

[0969] 본원에 사용되는 "알케닐"은 적어도 1개의 이중 결합을 함유하는 지방족 기를 지칭한다. 본원에 사용되는 "알키닐"은 2 내지 12개의 탄소 원자를 포함하고, 하나 이상의 삼중 결합을 갖는 것을 특징으로 하는 직선 또는 분지형 탄화수소 사슬을 지칭한다. 알키닐 기의 예로는 에티닐, 프로파르길, 및 3-헥시닐을 포함하지만, 이에 제한되는 것은 아니다.

[0970] 본원에 사용되는 "아릴알킬" 또는 "아랄킬"은 알킬 수소 원자가 아릴기로 대체된 알킬 모이어티를 지칭한다. 아랄킬은 하나를 초과하는 수소 원자가 아릴기로 대체된 기를 포함한다. "아릴알킬" 또는 "아랄킬"의 예로는 벤질, 2-페닐에틸, 3-페닐프로필, 9-플루오레닐, 벤즈하이드릴, 및 트리틸기가 포함된다.

[0971] 본원에 사용되는 "사이클로알킬"은 3개 내지 12개의 탄소를 갖는 사이클릭, 바이사이클릭, 트리사이클릭, 또는 폴리사이클릭 비방향족 탄화수소기를 지칭한다. 사이클로알킬 모이어티의 예로는 사이클로프로필, 사이클로펜틸, 및 사이클로헥실이 포함되지만, 이에 제한되는 것은 아니다.

[0972] 본원에 사용되는 "헤테로사이클릴"은 헤테로사이클릭 고리계의 1 가 라디칼을 지칭한다. 대표적인 헤테로사이클릴에는 비제한적으로, 테트라하이드로푸라닐, 테트라하이드로티에닐, 피롤리디닐, 피롤리도닐, 피페리디닐, 피롤리닐, 피페라지닐, 디옥사닐, 디옥솔라닐, 디제피닐, 옥사제피닐, 티제피닐, 및 모르폴리닐이 포함된다.

[0973] 본원에 사용되는 "헤테로아릴"은 헤테로방향족 고리계의 1가 라디칼을 지칭한다. 헤테로아릴 모이어티의 예로는 이미다졸릴, 옥사졸릴, 티아졸릴, 트리아졸릴, 피롤릴, 푸라닐, 인돌릴, 티오페닐 피라졸릴, 피리디닐, 피라지닐, 피리다지닐, 피리미디닐, 인돌리지닐, 퓨리닐, 나프티리디닐, 퀴놀릴, 및 프테리디닐이 포함되지만, 이에 제한되는 것은 아니다.

[0974] 포스페이트 백본 변형

[0975] 포스페이트기

[0976] 일부 실시형태에서, 변형된 뉴클레오타이드의 포스페이트기는 하나 이상의 산소를 상이한 치환기로 대체함으로써 변형될 수 있다. 추가적으로, 변형된 뉴클레오타이드, 예를 들어 변형된 핵산에 존재하는 변형된 뉴클레오타이드는, 변형되지 않은 포스페이트 모이어티의 본원에 기재된 변형된 포스페이트로의 대규모 대체를 포함할 수 있다. 일부 실시형태에서, 포스페이트 백본의 변형은 비하전 링커 또는 하전된 링커가 비대칭 전하 분포를 갖도록 야기하는 변경을 포함할 수 있다.

[0977] 변형된 포스페이트기의 예로는 포스포로티오에이트, 포스포로셀레네이트, 보라노 포스페이트, 보라노 포스페이트 에스테르, 수소 포스포네이트, 포스포로아미데이트, 알킬 또는 아릴 포스포네이트 및 포스포트리에스테르가 포함된다. 일부 실시형태에서, 포스페이트 백본 모이어티에서 비-가교 포스페이트 산소 원자 중 하나는 다음의

기 중 임의의 것으로 대체될 수 있다: 황(S), 셀레늄(Se), BR_3(여기서, R은 예를 들어 수소, 알킬, 또는 아릴일 수 있음), C(예를 들어 알킬기, 아릴기 등), H, NR_2(여기서, R은 예를 들어 수소, 알킬, 또는 아릴일 수 있음), 또는 OR(여기서, R은 예를 들어 알킬 또는 아릴일 수 있음). 변형되지 않은 포스페이트기에서 인 원자는 아키랄이다. 그러나, 상기 원자 또는 원자 그룹 중 하나로 비-가교 산소 중 하나를 대체하는 것은 인 원자를 키랄로 만들 수 있으며; 다시 말하면 이 방식으로 변형된 포스페이트기에서 인 원자는 입체발생 중심이다. 입체발생 인 원자는 "R" 배열(본원에서 Rp) 또는 "S" 배열(본원에서 Sp)를 보유할 수 있다.

[0978] 포스포로디티오에이트는 비-가교 산소 둘 다 황으로 대체된다. 포스포로디티오에이트에서 인 중심은 올리고리보뉴클레오타이드 부분입체이성질체의 형성을 불가능하게 하는 아키랄이다. 일부 실시형태에서, 비-가교 산소 하나 또는 둘 다의 변형은 또한 비-가교 산소를 S, Se, B, C, H, N, 및 OR(R은 예를 들어 알킬 또는 아릴일 수 있음)로부터 독립적으로 선택되는 기로 대체하는 것을 포함할 수 있다.

[0979] 포스페이트 링커는 또한 가교 산소(즉, 뉴클레오시드에 포스페이트를 연결하는 산소)를 질소(가교 포스포로아미데이트), 황(가교 포스포로티오에이트) 및 탄소(가교 메틸렌포스포네이트)로 대체함으로써 변형될 수 있다. 대체는 연결 산소 중 임의의 것 또는 연결 산소의 둘 다에서 발생할 수 있다.

[0980] 포스페이트기의 대체

[0981] 포스페이트기는 비-인 함유 커넥터(non-phosphrous containing connector)에 의해 대체될 수 있다. 실시형태에서, 하전 포스페이트기는 중성 모이어티로 대체될 수 있다.

[0982] 포스페이트기를 대체할 수 있는 모이어티의 예로는, 비제한적으로, 예를 들어 메틸 포스포네이트, 하이드록실아미노, 실록산, 카르보네이트, 카르복시메틸, 카르바메이트, 아미드, 티오에테르, 에틸렌 옥사이드 링커, 설포네이트, 설폰아미드, 티오포름아세탈, 포름아세탈, 옥심, 메틸렌이미노, 메틸렌메틸이미노, 메틸렌히드라조, 메틸렌디메틸히드라조 및 메틸렌옥시메틸이미노가 포함된다.

[0983] 리보포스페이트 백본의 대체

[0984] 포스페이트 링커 및 리보스 당이 뉴클레아제 내성 뉴클레오시드 또는 뉴클레오타이드 대용물로 대체된, 핵산을 모방할 수 있는 스캐폴드가 또한 작제될 수 있다. 일부 실시형태에서, 핵염기는 대용물 백본에 의해 묶일 수 있다. 예로는, 비제한적으로, 모르폴리노, 사이클로부틸, 피롤리딘 및 펩티드 핵산(PNA) 뉴클레오시드 대용물을 포함할 수 있다.

[0985] 당 변형

[0986] 변형된 뉴클레오시드 및 변형된 뉴클레오타이드는 당기에 하나 이상의 변형을 포함할 수 있다. 예를 들어 2' 하이드록실기(OH)는 다수의 상이한 "옥시" 또는 "데옥시" 치환기로 대체되거나 변형될 수 있다. 일부 실시형태에서, 하이드록실은 더 이상 탈양성자화되지 않아 2'-알콕사이드 이온을 형성할 수 있으므로, 2' 하이드록실기에의 변형은 핵산의 안정성을 향상할 수 있다. 2'-알콕사이드는 링커 인 원자 상에서 분자내 친핵성 공격에 의한 분해를 촉매할 수 있다.

[0987] "옥시"-2' 하이드록실기 변형의 예는 알콕시 또는 아릴옥시(OR, 여기서 "R"은, 예를 들어 알킬, 사이클로알킬, 아릴, 아랄킬, 헤테로아릴 또는 당일 수 있음); 폴리에틸렌글리콜(PEG), $O(CH_2CH_2O)_nCH2CH_2OR$ 여기서, R은 예를 들어 H 또는 선택적으로 치환된 알킬일 수 있고, n은 0 내지 20(예를 들어 0 내지 4, 0 내지 8, 0 내지 10, 0 내지 16, 1 내지 4, 1 내지 8, 1 내지 10, 1 내지 16, 1 내지 20, 2 내지 4, 2 내지 8, 2 내지 10, 2 내지 16, 2 내지 20, 4 내지 8, 4 내지 10, 4 내지 16, 및 4 내지 20)의 정수일 수 있음)을 포함할 수 있다. 일부 실시형태에서, "옥시"-2' 하이드록실기 변형은 "잠금(locked)" 핵산(LAN)을 포함할 수 있으며, 2' 하이드록실은, 예를 들어 $Ci-_6$ 알킬렌 또는 Cj-6 헤테로알킬렌 가교에 의해, 동일한 리보스 당의 4' 탄소에 연결될 수 있으며, 여기서 예시적인 가교는 메틸렌, 프로필렌, 에테르, 또는 아미노 가교; O-아미노(여기서, 아미노는 예를 들어 NH_2; 알킬아미노, 디알킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 또는 디헤테로아릴아미노, 에틸렌디아민, 또는 폴리아미노일 수 있음) 및 아미노알콕시, $O(CH_2)_n$-아미노(여기서, 아미노는, 예를 들어 NH_2; 알킬아미노, 디알킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 또는 디헤테로아릴아미노, 에틸렌디아민, 또는 폴리아미노일 수 있음)을 포함할 수 있다. 일부 실시형태에서, "옥시"-2' 하이드록실기 변형은 메톡시에틸기(MOE), ($OCH_2CH_2OCH_3$, 예를 들어 PEG 유도체)를 포함할 수 있다.

[0988] "데옥시" 변형은 수소(즉, 데옥시리보스 당, 예를 들어 부분적으로 ds RNA의 돌출된 부분에서); 할로(예를 들어 브로모, 클로로, 플루오로, 또는 요오도); 아미노(여기서, 아미노는, 예를 들어 NH_2; 알킬아미노, 디아킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 디헤테로아릴아미노, 또는 아미노산일 수 있음); $NH(CH_2CH_2NH)_nCH_2CH_2$-아미노(여기서, 아미노는, 예를 들어 본원에 기재된 바와 같을 수 있음), -NHC(O)R(여기서, R은, 예를 들어 알킬, 사이클로알킬, 아릴, 아랄킬, 헤테로아릴 또는 당일 수 있음), 시아노; 메르캅토; 알킬-티오-알킬; 티오알콕시; 및 알킬, 사이클로알킬, 아릴, 알케닐 및 알키닐을 포함할 수 있으며, 이는 예를 들어 본원에 기재된 아미노로 선택적으로 치환될 수 있다.

[0989] 당기는 또한 리보스 내의 상응하는 탄소와 반대의 입체화학 배열을 보유하는 하나 이상의 탄소를 함유할 수 있다. 따라서, 변형된 핵산은, 당으로서 예를 들어 아라비노스를 함유하는 뉴클레오타이드를 포함할 수 있다. 뉴클레오타이드 "단량체"는 당의 Γ 위치에서 알파 연결, 예를 들어 알파-뉴클레오시드를 가질 수 있다. 변형된 핵산은 또한, C- 에서 핵염기가 결여된, "무염기성" 당을 포함할 수 있다. 이들 무염기성 당은 또한 구성 요소 당 원자 중 하나 이상에서 추가로 변형될 수 있다. 변형된 핵산은 또한 L 형태인 하나 이상의 당, 예를 들어 L-뉴클레오시드를 포함할 수 있다.

[0990] 일반적으로, RNA는 당기 리보스를 포함하며, 이는 산소를 갖는 5-원 고리이다. 예시적인 변형된 뉴클레오시드 및 변형된 뉴클레오타이드는 비제한적으로, 리보스 중 산소의(예를 들어 황(S), 셀레늄(Se), 또는 예를 들어 메틸렌 또는 에틸렌과 같은 알킬렌으로의) 대체; 이중 결합의 첨가(예를 들어 리보스를 사이클로펜테닐 또는 사이클로헥세닐로 대체); 리보스의 고리 축소(예를 들어 사이클로부탄 또는 옥세탄의 4-원 고리 형성); 리보스의 고리 확장(예를 들어 안하이드로헥시톨, 알트리톨, 만니톨, 사이클로헥사닐, 사이클로헥세닐, 및 포스포르아미데이트 백본을 또한 갖는 모르폴리노와 같은, 예를 들어 추가적인 탄소 또는 헤테로원자를 갖는 6원- 또는 7-원 고리 형성)을 포함할 수 있다. 일부 실시형태에서, 변형된 뉴클레오타이드는 다중사이클릭 형태(예를 들어 트리사이클로; 및 "비잠금(unlocked)" 형태, 예를 들어 글리콜 핵산(GNA)(예를 들어 R-GNA 또는 S-GNA, 여기서 리보스는 포스포디에스테르 결합에 부착된 글리콜 단위로 대체됨), 트레오스 핵산(TNA, 여기서 리보스는 α-L-트레오푸라노실-(3'→2')로 대체됨)을 포함할 수 있다.

[0991] 핵염기 상의 변형

[0992] 변형된 핵산으로 혼입될 수 있는, 본원에 기재된 변형된 뉴클레오시드 및 변형된 뉴클레오타이드는, 변형된 핵염기를 포함할 수 있다. 핵염기의 예에는 아데닌 (A), 구아닌 (G), 시토신 (C), 및 우라실(U)이 포함되지만, 이에 제한되는 것은 아니다. 이들 핵염기는 변형되거나 완전히 대체되어 변형된 핵산으로 혼입될 수 있는 변형된 뉴클레오시드 및 변형된 뉴클레오타이드를 제공할 수 있다. 뉴클레오타이드의 핵염기는 퓨린, 피리미딘, 퓨린 또는 피리미딘 유사체로부터 독립적으로 선택될 수 있다. 일부 실시형태에서, 핵염기는, 예를 들어 자연적-발생 및 합성 염기 유도체를 포함할 수 있다.

[0993] 우라실

[0994] 일부 실시형태에서, 변형된 핵염기는 변형된 우라실이다. 변형된 우라실을 갖는 예시적인 핵염기 및 뉴클레오시드는 비제한적으로, 슈도우리딘(ψ), 피리딘-4-온 리보뉴클레오시드, 5-아자-우리딘, 6-아자-우리딘, 2-티오-5-아자-우리딘, 2-티오-우리딘(s2U), 4-티오-우리딘(s4U), 4-티오-슈도우리딘, 2-티오-슈도우리딘, 5-하이드록시-우리딘(ho^5U), 5-아미노알릴-우리딘, 5-할로-우리딘(예를 들어 5-요오도-우리딘 또는 5-브로모-우리딘), 3-메틸-우리딘(m^3U), 5-메톡시-우리딘(mo^5U), 우리딘 5-옥시아세트산(cmo^5U), 우리딘 5-옥시아세트산 메틸 에스테르(mcmo^U), 5-카르복시메틸-우리딘(cm^5U), 1-카르복시메틸-슈도우리딘, 5-카르복시하이드록시메틸-우리딘(chm^5U), 5-카르복시하이드록시메틸-우리딘 메틸 에스테르($mchm^5U$), 5-메톡시카르보닐메틸-우리딘(mcm^5U), 5-메톡시카르보닐메틸-2-티오-우리딘(mcm^5s2U), 5-아미노메틸-2-티오-우리딘(nm^5s2U), 5-메틸아미노메틸-우리딘(mnm^5U), 5-메틸아미노메틸-2-티오-우리딘(mnm^5s2U), 5-메틸아미노메틸-2-셀레노-우리딘(mnm^5se^2U), 5-카르바모일메틸-우리딘(ncm^5U), 5-카르복시메틸아미노메틸-우리딘($cmnm^5U$), 5-카르복시메틸아미노메틸-2-티오-우리딘(cmnm ＼s2U), 5-프로피닐-우리딘, 1-프로피닐-슈도우리딘, 5-타우리노메틸-우리딘(xcm^5U), 1-타우리노메틸-슈도우리딘, 5-타우리노메틸-2-티오-우리딘(Trn^5s2U), 1-타우리노메틸-4-티오-슈도우리딘, 5-메틸-우리딘(m^5U, 즉 핵염기 데옥시티민을 갖는), 1-메틸-슈도우리딘(ι τ ι'ψ), 5-메틸-2-티오-우리딘(m^5s2U), 1-메틸-4-티오-슈도

우리딘(m' s \|/), 4-티오-1-메틸-슈도우리딘, 3-메틸-슈도우리딘(m'V), 2-티오-1-메틸-슈도우리딘, 1-메틸-1-데아자-슈도우리딘, 2-티오-1-메틸-1-데아자-슈도우리딘, 디하이드로우리딘 (D), 디하이드로슈도우리딘, 5,6-디하이드로우리딘, 5-메틸-디하이드로우리딘(m^5D), 2-티오-디하이드로우리딘, 2-티오-디하이드로슈도우리딘, 2-메톡시-우리딘, 2-메톡시-4-티오-우리딘, 4-메톡시-슈도우리딘, 4-메톡시-2-티오-슈도우리딘, N1-메틸-슈도우리딘, 3-(3-아미노-3-카르복시프로필)우리딘(acp^3U), 1-메틸-3-(3-아미노-3-카르복시프로필 슈도우리딘 5-(이소펜테닐아미노메틸)우리딘(inm^5U), 5-(이소펜테닐아미노메티])-2-티오-우리딘(inm^5s2U), a-티오-우리딘, 2'-O-메틸-우리딘(Um), 5,2'-O-디메틸-우리딘(m^5Um), 2'-O-메틸-슈도우리딘(ψ π ι), 2-티오-2'-O-메틸-우리딘(s2Um), 5-메톡시카르보닐메틸-2'-O-메틸-우리딘(mcm^5Um), 5-카르바모일메틸-2'-O-메틸-우리딘(ncm^5Um), 5-카르복시메틸아미노메틸-2'-O-메틸-우리딘($cmnm^5Um$), 3,2'-O-디메틸-우리딘(m^3Um), 5-(이소펜테닐아미노메틸)-2'-O-메틸-우리딘(inm^5Um), 1-티오-우리딘, 디옥시티미딘, 2'-F-아라-우리딘, 2'-F-우리딘, 2'-OH-아라-우리딘, 5-(2-카르보메톡시비닐) 우리딘, 5-[3-(1-E-프로페닐아미노)우리딘, 피라졸로[3,4-d]피리미딘, 잔틴, 및 하이포잔틴을 포함한다.

[0995] 시토신

[0996] 실시형태에서, 변형된 핵염기는 변형된 시토신이다. 변형된 시토신을 갖는 예시적인 핵염기 및 뉴클레오시드는 비제한적으로, 5-아자-시티딘, 6-아자-시티딘, 슈도이소시티딘, 3-메틸-시티딘(m^3C), N4-아세틸-시티딘(act), 5-포르밀-시티딘(f^5C), N4-메틸-시티딘(m^4C), 5-메틸-시티딘(m^5C), 5-할로-시티딘(예를 들어 5-요오도-시티딘), 5-하이드록시메틸-시티딘(hm^5C), 1-메틸-슈도이소시티딘, 피롤로-시티딘, 피롤로-슈도이소시티딘, 2-티오-시티딘(s2C), 2-티오-5-메틸-시티딘, 4-티오-슈도이소시티딘, 4-티오-1-메틸-슈도이소시티딘, 4-티오-1-메틸-1-데아자-슈도이소시티딘, 1-메틸-1-데아자-슈도이소시티딘, 제불라린, 5-아자-제불라린, 5-메틸-제불라린, 5-아자-2-티오-제불라린, 2-티오-제불라린, 2-메톡시-시티딘, 2-메톡시-5-메틸-시티딘, 4-메톡시-슈도이소시티딘, 4-메톡시-1-메틸-슈도이소시티딘, 리시딘(k^2C), a-티오-시티딘, 2'-O-메틸-시티딘(Cm), 5,2'-O-디메틸-시티딘(m^5Cm), N4-아세틸-2'-O-메틸-시티딘(ac^4Cm), N4,2'-O-디메틸-시티딘(m^4Cm), 5-포르밀-2'-O-메틸-시티딘(f^5Cm), N4,N4,2'-O-트리메틸-시티딘(m^4_2Cm), 1-티오-시티딘, 2'-F-아라-시티딘, 2'-F-시티딘, 및 2'-OH-아라-시티딘을 포함한다.

[0997] 아데닌

[0998] 일부 실시형태에서, 변형된 핵염기는 변형된 아데닌이다. 변형된 아데닌을 갖는 예시적인 핵염기 및 뉴클레오시드는, 비제한적으로, 2-아미노-퓨린, 2,6-디아미노퓨린, 2-아미노-6-할로-퓨린(예를 들어 2-아미노-6-클로로-퓨린), 6-할로-퓨린(예를 들어 6-클로로-퓨린), 2-아미노-6-메틸-퓨린, 8-아지도-아데노신, 7-데아자-아데닌, 7-데아자-8-아자-아데닌, 7-데아자-2-아미노-퓨린, 7-데아자-8-아자-2-아미노-퓨린, 7-데아자-2,6-디아미노퓨린, 7-데아자-8-아자-2,6-디아미노퓨린, 1-메틸-아데노신(m^1A), 2-메틸-아데닌(m A), N6-메틸-아데노신(m^6A), 2-메틸티오-N6-메틸-아데노신($ms2m^6A$), N6-이소펜테닐-아데노신(i^6A), 2-메틸티오-N6-이소펜테닐-아데노신(ms^2i^6A), N6-(cis-하이드록시이소펜테닐)아데노신(io^6A), 2-메틸티오-N6-(cis-하이드록시이소펜테닐)아데노신($ms2io^6A$), N6-글리시닐카르바모일-아데노신(g^6A), N6-트레오닐카르바모일-아데노신(t^6A), N6-메틸-N6-트레오닐카르바모일-아데노신(m^6t^6A), 2-메틸티오-N6-트레오닐카르바모일-아데노신(ms^2g^6A), N6,N6-디메틸-아데노신(m^6_2A), N6-하이드록시노르발릴카르바모일-아데노신(hn^6A), 2-메틸티오-N6-하이드록시노르발릴카르바모일-아데노신($ms2hn^6A$), N6-아세틸-아데노신(ac^6A), 7-메틸-아데닌, 2-메틸티오-아데닌, 2-메톡시-아데닌, a-티오-아데노신, 2'-O-메틸-아데노신(Am), N6,2'-O-디메틸-아데노신(m^6Am), N6-메틸-2'-데옥시아데노신, N6,N6,2'-O-트리메틸-아데노신(m^6_2Am), 1,2'-O-디메틸-아데노신(m^1Am), 2'-O-리보실아데노신(포스페이트)(Ar(p)), 2-아미노-N6-메틸-퓨린, 1-티오-아데노신, 8-아지도-아데노신, 2'-F-아라-아데노신, 2'-F-아데노신, 2'-OH-아라-아데노신, 및 N6-(19-아미노-펜타옥사논아데실)-아데노신을 포함한다.

[0999] 구아닌

[1000] 일부 실시형태에서, 변형된 핵염기는 변형된 구아닌이다. 변형된 구아닌을 갖는 예시적인 핵염기 및 뉴클레오시드는 비제한적으로, 이노신 (i), 1-메틸-이노신(m '1), 와이오신(imG), 메틸와이오신(mimG), 4-데메틸-와이오신(imG-14), 이소와이오신(imG2), 와이부토신(yW), 퍼록시와이부토신(o^2yW), 하이드록시와이부토신(OHyW), 저변형 하이드록시와이부토신(OHyW*), 7-데아자-구아노신, 퀘우오신(Q), 에폭시퀘우오신(oQ), 갈락토실-퀘우오신(galQ), 만노실-퀘우오신(manQ), 7-시아노-7-데아자-구아노신(preQ0), 7-아미노메틸-7-데아자-구아노신(preQi), 아르캐오신(G+), 7-데아자-8-아자-구아노신, 6-티오-구아노신, 6-티오-7-데아자-구아노신, 6-티오-7-데아자-8-아자-구아노신, 7-메틸-구아노신(m^7G), 6-티오-7-메틸-구아노신, 7-메틸-이노신, 6-메톡시-구아노신, 1-메틸-구아노신(m'G), N2-메틸-구아노신(m^2G), N2,N2-디메틸-구아노신(m^2 $_2$G), N2,7-디메틸-구아노신(m^2,7G), N2, N2,7-디메틸-구아노신(m^2,2,7G), 8-옥소-구아노신, 7-메틸-8-옥소-구아노신, 1-메트 티오-구아노신, N2-메틸-6-티오-구아노신, N2,N2-디메틸-6-티오-구아노신, a-티오-구아노신, 2'-O-메틸-구아노신(Gm), N2-메틸-2'-O-메틸-구아노신($m_{3/4}$m), N2,N2-디메틸-2'-O-메틸-구아노신(m^2 $_2$Gm), 1-메틸-2'-O-메틸-구아노신(m'Gm), N2,7-디메틸-2'-O-메틸-구아노신(m^2,7Gm), 2'-O-메틸-이노신(Im), 1,2'-O-디메틸-이노신(m'Im), O^6-페닐-2'-디옥시이노신, 2'-O-리보실구아노신(포스페이트)(Gr(p)), 1-티오-구아노신, O^6-메틸] -구아노신, O^6-메티]-2'-데옥시구아노신, 2'-F-아라-구아노신, 및 2'-F-구아노신을 포함한다.

[1001] 변형된 gRNA

[1002] 일부 실시형태에서, 변형된 핵산은 변형된 gRNA일 수 있다. 일부 실시형태에서, gRNA는 3' 말단에서 변형될 수 있다. 이 실시형태에서, gRNA는 3' 말단 U 리보스에서 변형될 수 있다. 예를 들어 U 리보스의 2개의 말단 하이드록실기는 알데히드기로 산화되고 리보스 고리의 개구가 수반되어 변형된 뉴클레오시드를 제공할 수 있으며, 여기서, U는 변형되지 않은 또는 변형된 우리딘일 수 있다.

[1003] 또 다른 실시형태에서, 3' 말단 U는 2'3' 사이클릭 포스페이트로 변형될 수 있으며, 여기서, U는 변형되지 않은 또는 변형된 우리딘일 수 있다. 일부 실시형태에서, gRNA 분자는, 예를 들어 본원에 기재된 변형된 뉴클레오타이드 중 하나 이상을 혼입함으로써, 분해에 대하여 안정화될 수 있는 3' 뉴클레오타이드를 함유할 수 있다. 이 실시형태에서, 예를 들어 우리딘은 변형된 우리딘, 예를 들어 5-(2-아미노)프로필 우리딘, 및 5-브로모 우리딘, 또는 본원에 기재된 변형된 우리딘 중 임의의 것으로 대체될 수 있으며; 아데노신 및 구아노신은 변형된 아데노신 및 구아노신, 예를 들어 8-위치에서의 변형을 갖는 것, 예를 들어 8-브로모 구아노신, 또는 본원에 기재된 변형된 아데노신 또는 구아노신 중 임의의 것으로 대체될 수 있다. 일부 실시형태에서, 데아자 뉴클레오타이드, 예를 들어 7-데아자-아데노신은 gRNA로 혼입될 수 있다. 일부 실시형태에서, O- 및 N-알킬화 뉴클레오타이드, 예를 들어 N6-메틸 안 데노신은 gRNA로 혼입될 수 있다. 일부 실시형태에서, 예를 들어 2' OH-기가 H, -OR, -R (여기서, R은 예를 들어 메틸, 알킬, 사이클로알킬, 아릴, 아랄킬, 헤테로아릴 또는 당일 수 있음), 할로, -SH, -SR(여기서, R은 예를 들어 알킬, 사이클로알킬, 아릴, 아랄킬, 헤테로아릴 또는 당일 수 있음), 아미노(여기서, 아미노는 예를 들어 NH_2; 알킬아미노, 디알킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 디헤테로아릴아미노, 또는 아미노산일 수 있음); 또는 시아노(-CN)로부터 선택되는 기로 대체되는, 당-변형 리보뉴클레오타이드가 혼입될 수 있다. 일부 실시형태에서, 포스페이트 백본은 본원에 기재된 바와 같이, 예를 들어 포스포티오에이트기로 변형될 수 있다. 일부 실시형태에서, gRNA의 돌출부 영역의 뉴클레오타이드는 각각 독립적으로 변형된 2'-당 변형된, 예를 들어 2-F 2'-O-메틸, 티미딘(T), 2'-O-메톡시에틸-5-메틸우리딘(Teo), 2'-O-메톡시에틸아데노신(Aeo), 2'-O-메톡시에틸-5-메틸시티딘(m5Ceo), 및 이의 임의의 조합을 포함하지만 이에 제한되지 않는, 변형된 또는 변형되지 않은 뉴클레오타이드일 수 있다.

[1004] 실시형태에서, RNA 분자, 예를 들어 gRNA 분자의 단일 가닥 돌출부 내의 뉴클레오타이드의 하나 이상 또는 모두는 데옥시뉴클레오타이드이다.

[1005] **약제학적 조성물**

[1006] 본 발명의 약제학적 조성물은 본원에 기재된 gRNA 분자, 예를 들어 본원에 기재된 복수의 gRNA 분자, 또는 본원에 기재된 하나 이상의 gRNA 분자로 변형된 하나 이상의 세포를 포함하는 세포(예를 들어 세포 집단, 예를 들어 조혈 줄기 세포, 예를 들어 CD34+ 세포의 집단)를 약제학적으로 또는 생리적으로 허용되는 담체, 희석제 또는

부형제와 조합하여 포함할 수 있다. 이러한 조성물은 중성 완충 식염수, 포스페이트 완충 식염수 등의 완충제; 글루코스, 만노스, 수 크로스 또는 덱스트란, 만니톨과 같은 탄수화물, 만니톨; 단백질; 폴리펩티드 또는 글리신과 같은 아미노산; 항산화제; EDTA 또는 글루타티온과 같은 킬레이트화제; 보조제(예를 들어 알루미늄 하이드록사이드); 및 보존제를 포함할 수 있다. 본 발명의 조성물은 일 양태에서 정맥내 투여용으로 제형화된다.

[1007] 본 발명의 약제학적 조성물은 치료될(또는 예방될) 질병에 적절한 방식으로 투여될 수 있다. 적절한 투여량은 임상 시험에 의해 결정될 수 있지만, 투여량 및 빈도는 환자의 질환, 및 환자 질병의 유형 및 중증도와 같은 인자에 의해 결정될 것이다.

[1008] 일 실시형태에서, 약제학적 조성물은 예를 들어 내독소, 미코플라스마, 마우스 항체, 혼합 인간 혈청, 소 혈청 알부민, 소 혈청, 배양 배지 성분, 원하지 않는 CRISPR 시스템 성분, 박테리아 및 균류로 구성된 군으로부터 선택되는, 오염 물질이 실질적으로 없고, 예를 들어 검출 가능한 수준의 오염 물질이 없다. 일 실시형태에서, 박테리아는 알칼리게네스 파에칼리스(Alcaligenes faecalis), 칸디다 알비칸스(Candida albicans), 에스케리치아 콜라이(에스케리키아 콜라이), 해모필루스 인플루엔자(Haemophilus influenza), 네이세리아 메닌기티데스(Neisseria meningitides), 슈도모나스 아에루기노사(Pseudomonas aeruginosa), 스타필로코커스 아우레우스(Staphylococcus aureus), 스트렙토코커스 뉴모니아(Streptococcus pneumonia), 및 스트렙토코커스 피오게네스 그룹 A(Streptococcus pyogenes group A)로 구성된 군으로부터 선택된 적어도 하나이다.

[1009] 대상체 조성물의 투여는 에어로졸 흡입, 주사, 섭취, 수혈, 삽입 또는 이식을 포함하는 임의의 편리한 방식으로 수행될 수 있다. 본원에 기재된 조성물은 환자에게 경동맥으로, 피하, 피내, 종양내, 결절내, 골수내, 근육내, 정맥내(i.v.) 주사에 의해 또는 복강내 투여될 수 있다. 일 양태에서, 본 발명의 조성물은 피내 또는 피하 주사에 의해 환자에게 투여된다. 일 양태에서, 본 발명의 세포 조성물은 i.v. 주사에 의해 투여된다.

[1010] 환자에게 투여되는 상기 치료의 투여량은 치료되는 질환의 정확한 본질 및 치료의 수용자에 따라 다를 것이다. 인간 투여를 위한 투여량의 크기 조정은 당분야에서 허용되는 관행에 따라 수행될 수 있다.

[1011] **세포**

[1012] 본 발명은 또한 본 발명의 gRNA 분자 또는 상기 gRNA 분자를 인코딩하는 핵산을 포함하는 세포에 관한 것이다.

[1013] 양태에서, 세포는 본원에 기재된 과정에 의해 제조된 세포이다.

[1014] 실시형태에서, 세포는 조혈 줄기 세포(예를 들어 조혈 줄기 및 전구 세포; HSPC), 예를 들어 CD34+ 줄기 세포이다. 실시형태에서, 세포는 CD34+/CD90+ 줄기 세포이다. 실시형태에서, 세포는 CD34+/CD90- 줄기 세포이다. 실시형태에서, 세포는 인간 조혈 줄기 세포이다. 실시형태에서, 세포는 자가 유래이다. 실시형태에서, 세포는 동종이계이다.

[1015] 실시형태에서, 세포는 골수, 예를 들어 자가 유래 골수로부터 유래된다. 실시형태에서, 세포는 말초 혈액, 예를 들어 동원된 말초 혈액, 예를 들어 자가 동원된 말초 혈액으로부터 유래된다. 동원된 말초 혈액을 이용하는 실시형태에서, 세포는 동원제를 투여한 환자로부터 단리된다. 실시형태에서, 동원제는 G-CSF이다. 실시형태에서, 동원제는 Plerixafor®(AMD3100)이다. 실시형태에서, 동원제는 G-CSF와 Plerixafor®(AMD3100))의 조합이다. 실시형태에서, 세포는 제대혈, 예를 들어 동종이계 제대혈로부터 유래된다. 실시형태에서, 세포는 혈색소병증 환자, 예를 들어 겸상 적혈구 질병 환자 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아 환자로부터 유래된다.

[1016] 실시형태에서, 세포는 포유류이다. 실시형태에서, 세포는 인간이다. 실시형태에서, 세포는 혈색소병증 환자, 예를 들어 겸상 적혈구 질병 환자 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아 환자로부터 유래된다.

[1017] 일 양태에서, 본 발명은 예를 들어 본원에 기재된 바와 같은 CRISPR 시스템의 부분으로서 상기 세포 내로 도입된, 예를 들어 본원에 기재된 바와 같은 gRNA 분자 또는 gRNA 분자들에 대해 상보성을 갖는 핵산 서열에서 또는 그 부근에서(예를 들어 이의 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 뉴클레오타이드 이내에서) 변형 또는 변경, 예를 들어 삽입결실을 포함하는 세포를 제공한다. 실시형태에서, 세포는 CD34+ 세포이다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포는 다양한 계통의 세포로 분화하는 능력을 유지하며, 예를 들어 적혈구계 계통의 세포로 분화하는 능력을 유지시킨다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포는, 배양 중에, 예를 들어 줄기 세포 확장제, 예를 들어 화합물 4를 포함하는 배양 중에, 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 또는 적어도 10회 이상의 배가를 겪었거나 겪을 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포는 배양 중에, 예를 들어 줄기 세포 확장제 분자, 예를 들어 본원에 기재된, 예를 들어 화합물 4를

포함하는 배양 중에, 적어도 5, 예를 들어 약 5회 배가를 겪었거나 겪을 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포는, 변형되지 않거나 변경되지 않은 유사한 세포에 비해, 증가된 태아 헤모글로빈 수준(예를 들어 발현 수준 및/또는 단백질 수준)을 나타내는, 예를 들어 태아 헤모글로빈 단백질 수준에서 적어도 20%의 증가를 나타내는 세포, 예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구를 나타내거나 및/또는 이로 분화될 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포는, 변형되지 않거나 변경되지 않은 유사한 세포에 비해, 증가된 태아 헤모글로빈 수준(예를 들어 발현 수준 및/또는 단백질 수준)을 나타내며, 예를 들어 적어도 6 피코그램, 예를 들어 적어도 7 피코그램, 적어도 8 피코그램, 적어도 9 피코그램, 또는 적어도 10 피코그램의 태아 헤모글로빈을 생산하는 세포, 예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구를 나타내거나 및/또는 이로 분화될 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포는, 변형되지 않거나 변경되지 않은 유사한 세포에 비해, 증가된 태아 헤모글로빈 수준(예를 들어 발현 수준 및/또는 단백질 수준)을 나타내며, 예를 들어 약 6 내지 약 12, 약 6 내지 약 7, 약 7 내지 약 8, 약 8 내지 약 9, 약 9 내지 약 10, 약 10 내지 약 11 또는 약 11 내지 약 12 피코그램의 태아 헤모글로빈을 생산하는 세포, 예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구를 나타내거나 및/또는 이로 분화될 수 있다.

[1018] 일 양태에서, 본 발명은, 예를 들어 본원에 기재된, CRISPR 시스템의 부분으로서, 상기 세포 내로 도입된, 예를 들어 본원에 기재된, gRNA 분자 또는 gRNA 분자들에 대한 상보성을 갖는 핵산 서열에 또는 그 부근에(예를 들어 그의 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1 뉴클레오타이드 이내에) 변형 또는 변경, 예를 들어 삽입결실을 갖는 세포를 포함하는 세포 집단을 제공한다. 실시형태에서, 예를 들어 본 발명의 gRNA 및/또는 CRISPR 시스템의 도입 후 2일차에, 예를 들어 본원에 기재된 바와 같은 NGS에 의해 측정하여, 집단의 세포의 적어도 50%, 예를 들어 적어도 60%, 적어도 70%, 적어도 80% 또는 적어도 90%가 변형 또는 변경을 갖는다(예를 들어 적어도 하나의 변형 또는 변경을 갖는다). 실시형태에서, 예를 들어 본 발명의 gRNA 및/또는 CRISPR 시스템의 도입 후 2일차에, 예를 들어 본원에 기재된 바와 같은 NGS에 의해 측정하여, 집단의 세포의 적어도 90%, 예를 들어 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99%가 변형 또는 변경을 갖는다(예를 들어 적어도 하나의 변형 또는 변경을 갖는다). 실시형태에서, 세포 집단은 CD34+ 세포를 포함하며, 예를 들어 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95% 또는 적어도 약 98%의 CD34+ 세포를 포함한다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포를 포함하는 세포 집단은 다양한 계통의 세포를 생산, 예를 들어 이로 분화하는 능력을 유지하고, 예를 들어 적혈구계 계통의 세포를 생산, 예를 들어 이로 분화하는 능력을 유지시킨다. 실시형태에서, 세포 집단, 예를 들어 CD34+ 세포의 집단은 배양 중에, 예를 들어 줄기 세포 확장제, 예를 들어 화합물 4를 포함하는 배양 중에, 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 또는 적어도 10회 이상의 집단 배가를 겪었거나 겪을 수 있다. 실시형태에서, 변경된 또는 변형된 세포의 집단, 예를 들어 CD34+ 세포의 집단은 배양 중에, 예를 들어 줄기 세포 확장제 분자, 예를 들어 본원에 기재된, 예를 들어 화합물 4를 포함하는 배양 중에, 적어도 5, 예를 들어 약 5회 집단 배가를 겪었거나 겪을 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포를 포함하는 세포 집단은, 변형되지 않거나 변경되지 않은 유사한 세포에 비해, 증가된 태아 헤모글로빈 수준(예를 들어 발현 수준 및/또는 단백질 수준)을 나타내는, 예를 들어 태아 헤모글로빈 단백질 수준에서 적어도 20%의 증가를 나타내는 세포 집단, 예를 들어 적혈구계 계통의 세포 집단, 예를 들어 적혈구 집단을 나타내거나 및/또는 이로 분화될 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포를 포함하는 세포 집단은, 변형되지 않거나 변경되지 않은 유사한 세포에 비해, 증가된 태아 헤모글로빈 수준(예를 들어 발현 수준 및/또는 단백질 수준)을 나타내며, 예를 들어 세포 당 적어도 6 피코그램, 예를 들어 적어도 7 피코그램, 적어도 8 피코그램, 적어도 9 피코그램, 또는 적어도 10 피코그램의 태아 헤모글로빈을 생산하는 세포를 포함하는, 세포 집단, 예를 들어 적혈구계 계통의 세포 집단, 예를 들어 적혈구 집단을 나타내거나 및/또는 이로 분화될 수 있다. 실시형태에서, 변경된 또는 변형된 세포, 예를 들어 CD34+ 세포의 집단은, 변형되지 않거나 변경되지 않은 유사한 세포에 비해, 증가된 태아 헤모글로빈 수준(예를 들어 발현 수준 및/또는 단백질 수준)을 나타내며, 예를 들어 세포 당 약 6 내지 약 12, 약 6 내지 약 7, 약 7 내지 약 8, 약 8 내지 약 9, 약 9 내지 약 10, 약 10 내지 약 11 또는 약 11 내지 약 12 피코그램의 태아 헤모글로빈을 생산하는 세포를 포함하는, 세포 집단, 예를 들어 적혈구계 계통의 세포 집단, 예를 들어 적혈구 집단을 나타내거나 및/또는 이로 분화될 수 있다.

[1019] 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e3개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e4개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e5개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e7개 세포를 포

함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e8개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e9개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e10개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e11개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 적어도 약 1 e12개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e13개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 2 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 3 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 4 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 5 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 6 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 7 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 8 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 9 e6개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 2 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 3 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 4 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 5 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 6 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 7 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 8 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 9 e7개 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e8개 세포를 포함한다. 전술한 실시형태 중 임의의 것에서, 세포 집단은 적어도 약 50%(예를 들어 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95% 또는 적어도 약 99%)의 HSPC, 예를 들어 CD34+ 세포를 포함할 수 있다. 전술한 실시형태 중 임의의 것에서, 세포 집단은 약 60%의 HSPC, 예를 들어 CD34+ 세포를 포함할 수 있다. 실시형태에서, 예를 들어 본원에 기재된 세포 집단은 약 3 e7개 세포를 포함하고 약 2 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다.

[1020] 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1.5 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 2 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 3 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 4 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 5 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 6 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 7 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 8 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 9 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 2 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 3 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 4 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재

된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 5 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 6 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 7 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 8 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 9 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 1 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 2 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 3 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 4 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 적어도 약 5 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다.

[1021] 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 1 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 1.5 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 2 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 3 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 4 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 5 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 6 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 7 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 8 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 9 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 1 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 2 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 3 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 4 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 5 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 6 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 7 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 8 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 9 e7개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 1 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 2 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 3 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 4 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 5 e8개 HSPC, 예를 들어 CD34+ 세포를 포함한다.

[1022] 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 약 2 e6 내지 약 10 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다. 실시형태에서, 예를 들어 본원에 기재된, 세포 집단은 투여될 환자의 체중 킬로그램 당 2 e6 내지 10 e6개 HSPC, 예를 들어 CD34+ 세포를 포함한다.

[1023] 본 발명의 세포는 본 발명의 gRNA 분자 또는 상기 gRNA 분자를 인코딩하는 핵산, 및 본 발명의 Cas9 분자 또는 상기 Cas9 분자를 인코딩하는 핵산을 포함할 수 있다. 실시형태에서, 본 발명의 세포는 본 발명의 gRNA 분자 및

본 발명의 Cas9 분자를 포함하는 리보핵 단백질(RNP) 복합체를 포함할 수 있다.

[1024] 본 발명의 세포는 바람직하게는, 예를 들어 본원에 기재된 방법에 의해, 예를 들어 전기천공에 의해 또는 TRIAMF에 의해(특허 출원 PCT/US2017/54110에 기재된 바와 같으며, 그 전체가 참조로 본 명세서에 포함됨) 생체 외에서 본 발명의 gRNA 분자를 포함하도록 변형된다.

[1025] 본 발명의 세포는, 하나 이상의 유전자의 발현이 본 발명의 gRNA를 포함하는 CRISPR 시스템의 도입에 의해 변경된, 예를 들어 감소된 또는 저해된 세포를 포함한다. 예를 들어 본 발명의 세포는 비변형된 세포에 비해 감소된 수준의 베타 글로빈(예를 들어 겸상세포생성(sickling) 돌연변이를 포함하는 헤모글로빈 베타) 발현을 가질 수 있다. 또 다른 예로서, 본 발명의 세포는 비변형된 세포에 비해 증가된 수준의 태아 헤모글로빈 발현을 가질 수 있다. 대안적으로 또는 이에 더하여, 본 발명의 세포는 비변형된 세포로부터 분화된 세포에 비해 증가된 수준의 태아 헤모글로빈 발현을 갖는 또 다른 유형의 세포, 예를 들어 적혈구를 야기하거나, 예를 들어 이로 분화될 수 있다. 실시형태에서, 태아 헤모글로빈 수준에서의 증가는 적어도 약 20%, 적어도 약 30%, 적어도 약 40% 또는 적어도 약 50%이다. 대안적으로 또는 이에 더하여, 본 발명의 세포는 비변형된 세포로부터 분화된 세포에 비해 감소된 수준의 베타 글로빈(예를 들어 본원에서 겸상 베타 글로빈으로도 지칭되는, 겸상세포생성 돌연변이를 포함하는 헤모글로빈 베타) 발현을 갖는 또 다른 유형의 세포, 예를 들어 적혈구를 야기하거나, 예를 들어 이로 분화될 수 있다. 실시형태에서, 겸상 베타-글로빈의 수준에서의 저하는 적어도 약 20%, 적어도 약 30%, 적어도 약 40% 또는 적어도 약 50%이다.

[1026] 본 발명의 세포는, 하나 이상의 유전자의 발현이 본 발명의 gRNA를 포함하는 CRISPR 시스템의 도입에 의해 변경된, 예를 들어 감소된 또는 저해된 세포를 포함한다. 예를 들어 본 발명의 세포는 비변형된 세포에 비해, 감소된 수준의 헤모글로빈 베타, 예를 들어 돌연변이화된 또는 야생형 헤모글로빈 베타의 발현을 가질 수 있다. 또 다른 양태에서, 본 발명은, 본 발명의 gRNA를 포함하는 CRISPR 시스템이 도입된 세포로부터 유래된, 예를 들어 분화된 세포를 제공한다. 이러한 양태에서, 본 발명의 gRNA를 포함하는 CRISPR 시스템이 도입된 세포는 감소된 수준의 헤모글로빈 베타, 예를 들어 돌연변이화된 또는 야생형 헤모글로빈 베타를 나타낼 수는 없으나, 상기 세포로부터 유래된, 예를 들어 분화된 세포는 감소된 수준의 헤모글로빈 베타, 예를 들어 돌연변이화된 또는 야생형 헤모글로빈 베타를 나타낸다. 실시형태에서, 유래(derivation), 예를 들어 분화는 생체내에서(예를 들어 환자, 예를 들어 혈색소병증 환자, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타 탈라쎄미아 환자에서) 달성된다. 실시형태에서, 본 발명의 gRNA를 포함하는 CRISPR 시스템이 도입된 세포는 CD34+ 세포이고, 이로부터 유래된, 예를 들어 분화된 세포는 적혈구계 계통의 세포, 예를 들어 적혈구이다.

[1027] 본 발명의 세포는, 하나 이상의 유전자의 발현이 본 발명의 gRNA를 포함하는 CRISPR 시스템의 도입에 의해 변경된, 예를 들어 증가된 또는 촉진된 세포를 포함한다. 예를 들어 본 발명의 세포는 비변형된 세포에 비해 증가된 수준의 태아 헤모글로빈 발현을 가질 수 있다. 또 다른 양태에서, 본 발명은, 본 발명의 gRNA를 포함하는 CRISPR 시스템이 도입된 세포로부터 유래된, 예를 들어 분화된 세포를 제공한다. 이러한 양태에서, 본 발명의 gRNA를 포함하는 CRISPR 시스템이 도입된 세포는 증가된 수준의 태아 헤모글로빈을 나타낼 수는 없으나, 상기 세포로부터 유래된, 예를 들어 분화된 세포는 증가된 수준의 태아 헤모글로빈을 나타낸다. 실시형태에서, 유래, 예를 들어 분화는 생체내에서(예를 들어 환자, 예를 들어 혈색소병증 환자, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아 환자, 예를 들어 베타 탈라쎄미아 환자에서) 달성된다. 실시형태에서, 본 발명의 gRNA를 포함하는 CRISPR 시스템이 도입된 세포는 CD34+ 세포이고, 이로부터 유래된, 예를 들어 분화된 세포는 적혈구계 계통의 세포, 예를 들어 적혈구이다.

[1028] 또 다른 양태에서, 본 발명은 상기 세포 내로 도입된 gRNA 분자(예를 들어 gRNA 분자의 표적 서열) 또는 gRNA 분자들에 대해 상보성을 갖는 핵산 서열에서(예를 들어 핵산 서열 내에서) 또는 그 부근에서(예를 들어 핵산 서열의 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 뉴클레오타이드에서) 삽입결실을 포함하는 세포에 관한 것이다. 실시형태에서, 삽입결실은 프레임시프트 삽입결실이다. 실시형태에서, 세포는 대형 결실, 예를 들어 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb 이상의 결실을 포함한다. 실시형태에서, 대형 결실은 상기 세포 내로 도입된 gRNA 분자 또는 gRNA 분자들에 대한 2개의 결합 부위 사이에 배치된 핵산을 포함한다. 실시형태에서, 결실은 HBG1 프로모터 영역에 배치된 본원에 기재된 gRNA의 표적 서열과 HBG2 프로모터 영역에 배치된 본원에 기재된 gRNA의 표적 서열 사이에 배치된 약 4900 nt를 포함하며, 예를 들어 이로 구성된다. 실시형태에서, 삽입결실, 예를 들어 결실은 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는다.

[1029] 일 양태에서, 본 발명은 예를 들어 본원에 기재된 바와 같은 상기 세포 내로 도입된 예를 들어 본원에 기재된

바와 같은 gRNA 분자 또는 gRNA 분자들에 대해 상보성을 갖는 핵산 서열에서 또는 그 부근에서(예를 들어 핵산 서열의 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 뉴클레오타이드에서) 삽입결실을 포함하는 세포를 포함하는 세포 집단(예를 들어 본원에 기재된 바와 같음), 예를 들어 HSPC 집단에 관한 것이다. 실시형태에서, 삽입결실은 프레임시프트 삽입결실이다. 실시형태에서, 세포 집단은 대형 결실, 예를 들어 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb 이상의 결실을 포함하는 세포를 포함한다. 실시형태에서, 대형 결실은 상기 세포 내로 도입된 gRNA 분자 또는 gRNA 분자들에 대한 2개의 결합 부위 사이에 배치된 핵산을 포함한다. 실시형태에서, 결실은 HBG1 프로모터 영역에 배치된 본원에 기재된 gRNA의 표적 서열과 HBG2 프로모터 영역에 배치된 본원에 기재된 gRNA의 표적 서열 사이에 배치된 약 4900 nt를 포함하며, 예를 들어 이로 구성된다. 실시형태에서, 집단의 세포 중 1%, 0.5%, 0.1% 또는 0.001% 미만(예를 들어 집단의 세포 중 0%)은 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함한다. 실시형태에서, 집단의 세포 중 20% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 30% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 40% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 50% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 60% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 70% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 80% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 집단의 세포 중 90% 내지 100%는 상기 대형 결실, 삽입결실 또는 삽입결실들을 포함한다. 실시형태에서, 세포 집단은 예를 들어 대상체, 예를 들어 인간에서 다수의 세포 유형으로 분화하는 능력을 보유하며, 예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구로 분화하는 능력을 유지시킨다. 실시형태에서, 편집된 세포(예를 들어 본원에 기재된 바와 같은 예를 들어 HSPC 세포, 예를 들어 CD34+ 세포, 예를 들어 CD34+ 세포의 임의의 하위집단)는 예를 들어 세포 배양물에서 증식하는, 예를 들어 본원에 기재된 세포 배양 배지, 예를 들어 하나 이상의 줄기세포 확장제, 예를 들어 화합물 4를 포함하는 세포 배양 배지에서 예를 들어 1, 2, 3, 4, 5, 6, 7일 또는 그 이상 후(예를 들어 약 1 또는 2일 후), 적어도 5배, 적어도 10배, 적어도 15배, 적어도 20배, 적어도 25배, 적어도 30배, 적어도 35배, 적어도 40배, 적어도 45배, 적어도 50배 이상 증식하는 능력을 유지시킨다(및/또는 증식한다). 실시형태에서, 편집되고 분화된 세포(예를 들어 적혈구)는 증식하는, 예를 들어 실시예에 기재된 바와 같이 적혈구계 분화 배지(EDM)에서 약 7일 후 예를 들어 적어도 5배, 적어도 10배, 적어도 15배, 적어도 20배, 적어도 25배, 적어도 30배, 적어도 35배, 적어도 40배, 적어도 45배, 적어도 50배 이상 증식하고/하거나, 예를 들어 실시예에 기재된 바와 같이 예를 들어 적혈구계 분화 배지(EDM)에서 또는 대상체(예를 들어 포유류, 예를 들어 인간)에서 21일 후 적어도 30배, 적어도 35배, 적어도 40배, 적어도 45배, 적어도 50배, 적어도 55배, 적어도 60배, 적어도 65배, 적어도 70배, 적어도 75배, 적어도 80배, 적어도 85배, 적어도 90배, 적어도 95배, 적어도 100배, 적어도 110배, 적어도 120배, 적어도 130배, 적어도 140배, 적어도 150배, 적어도 200배, 적어도 300배, 적어도 400배, 적어도 500배, 적어도 600배, 적어도 700배, 적어도 800배, 적어도 900배, 적어도 1000배, 적어도 1100배, 적어도 1200배, 적어도 1300배, 적어도 1400배, 적어도 1500배 이상 증식하는 능력을 유지시킨다.

[1030] 일 실시형태에서, 본 발명은 세포, 예를 들어 CD34+ 세포의 집단을 제공하며, 이 중에서 집단의 세포 중 적어도 90%, 예를 들어 적어도 95%, 예를 들어 적어도 98%는 예를 들어 본원에 기재된 바와 같은 대형 결실 또는 하나 이상의 삽입결실을 포함한다. 이론으로 결부시키고자 하는 것은 아니지만, 세포 집단 내로의 본원에 기재된 바와 같은 gRNA 분자 또는 CRISPR 시스템의 도입은 상기 집단에서 삽입결실 및/또는 대형 결실의 패턴을 생산하고, 따라서 삽입결실 및/또는 대형 결실을 포함하는 각각의 집단의 세포는 동일한 삽입결실 및/또는 대형 결실을 나타내지 않을 수 있는 것으로 여겨진다. 실시형태에서, 삽입결실 및/또는 대형 결실은 본원에 기재된 gRNA 분자의 표적화 도메인에 상보적인 부위에서 또는 그 부근에서 하나 이상의 핵산을 포함하며; 여기서, 상기 세포는 적혈구계 계통의 세포, 예를 들어 적혈구로 분화하는 능력을 유지하고/하거나; 여기서, 세포 집단으로부터 분화된 상기 세포는 비변형된 세포의 유사한 집단으로부터 분화된 세포에 비해 증가된 수준의 태아 헤모글로빈을 가진다(예를 들어 이러한 집단은 더 높은 %의 F 세포를 가짐). 실시형태에서, 세포 집단은 생체외에서, 예를 들어 하나 이상의 줄기세포 확장제를 포함하는, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 배지에서 적어도 2배의 확장을 겪었다. 실시형태에서, 세포 집단은 생체외에서, 예를 들어 하나 이상의 줄기세포 확장제를 포함하는, 예를 들어 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올을 포함하는 배지에서 적어도 5배의 확장을 겪었다.

[1031] 실시형태에서, 삽입결실은 약 50개 미만 뉴클레오타이드, 예를 들어 약 45개 미만, 약 40개 미만, 약 35개

미만, 약 30개 미만 또는 약 25개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 25개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 20개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 15개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 10개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 9개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 9개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 7개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 6개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 5개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 4개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 3개 미만 뉴클레오타이드이다. 실시형태에서, 삽입결실은 약 2개 미만 뉴클레오타이드이다. 임의의 상기 언급된 실시형태에서, 삽입결실은 적어도 1개 뉴클레오타이드이다. 실시형태에서, 삽입결실은 1개 뉴클레오타이드이다. 실시형태에서, 대형 결실은 약 1 kb의 DNA를 포함한다. 실시형태에서, 대형 결실은 약 2 kb의 DNA를 포함한다. 실시형태에서, 대형 결실은 약 3 kb의 DNA를 포함한다. 실시형태에서, 대형 결실은 약 4 kb의 DNA를 포함한다. 실시형태에서, 대형 결실은 약 5 kb의 DNA를 포함한다. 실시형태에서, 대형 결실은 약 6 kb의 DNA를 포함한다. 실시형태에서, 대형 결실은 예를 들어 HBG1 프로모터 영역 내의 표적 서열과 HBG2 프로모터 영역 내의 표적 서열 사이에 배치된 약 4.9 kb의 DNA를 포함한다.

[1032] 실시형태에서, 세포 집단(예를 들어 본원에 기재된 바와 같음)은 본원에 기재된 gRNA 분자를 포함하는 CRISPR 시스템과 연관된 임의의 1, 2, 3, 4, 5 또는 6개의 가장 빈번하게 검출되는 삽입결실을 포함하는 삽입결실 및/또는 대형 결실의 패턴을 포함하며, 예를 들어 표 7-2에 기재된 1, 2, 3, 4, 5 또는 6개의 삽입결실 및 대형 결실을 포함한다(예를 들어 HBG1 표적 서열에서 또는 그 부근에서 검출된 1, 2, 3, 4, 5 또는 6개의 삽입결실 및 대형 결실을 포함하고/하거나 HBG2 표적 서열에서 또는 그 부근에서 검출된 1, 2, 3, 4, 5 또는 6개의 삽입결실 및 대형 결실을 포함함). 실시형태에서, 삽입결실 및/또는 대형 결실은 본원에 기재된 방법에 의해, 예를 들어 NGS 또는 qPCR에 의해 검출된다.

[1033] 일 양태에서, 세포 또는 세포 집단(예를 들어 본원에 기재된 바와 같음)은 예를 들어 본원에 기재된 방법에 의해 검출된 바와 같이 표적-외 부위에서 삽입결실 또는 대형 결실을 포함하지 않는다.

[1034] 실시형태에서, 본원에 기재된 세포 또는 세포 집단(예를 들어 겸상 적혈구 질병 환자로부터 유래됨)의 자손, 예를 들어 분화된 자손, 예를 들어 적혈구계(예를 들어 적혈구) 자손은 비변형된 세포보다 더 낮은 수준의 겸상 베타 글로빈 및/또는 더 높은 수준의 감마 글로빈을 생산한다. 실시형태에서, 본원에 기재된 세포 또는 세포 집단(예를 들어 겸상 적혈구 질병 환자로부터 유래됨)의 자손, 예를 들어 분화된 자손, 예를 들어 적혈구계(예를 들어 적혈구) 자손은 비변형된 세포보다 더 낮은 수준의 겸상 베타 글로빈 및 더 높은 수준의 감마 글로빈을 생산한다. 실시형태에서, 겸상 베타 글로빈은 비변형된 세포보다 적어도 약 20%, 적어도 약 30%, 적어도 약 40% 또는 적어도 약 50% 더 낮은 수준에서 생산된다. 실시형태에서, 감마 글로빈은 비변형된 세포보다 적어도 약 20%, 적어도 약 30%, 적어도 약 40%, 적어도 약 50%, 적어도 약 60% 또는 적어도 약 70% 더 높은 수준에서 생산된다.

[1035] 일 양태에서, 본 발명은 예를 들어 본원에 기재된 바와 같은 변형된 HSPC 또는 상기 HSPC로부터 분화된(예를 들어 생체외에서 또는 환자에서 분화된) 적혈구계 세포의 집단을 제공하며, 여기서, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90% 또는 적어도 95%의 세포는 F 세포이다. 실시형태에서, 세포 집단은 상기 세포 내로 도입된 예를 들어 본원에 기재된 바와 같은 gRNA 분자 또는 gRNA 분자들을 갖지 않은 유사한 세포 집단보다 더 높은 퍼센트의 F 세포를 함유한다(또는 이러한 F 세포를 함유하는 적혈구 집단으로 예를 들어 생체내에서 분화할 수 있음). 실시형태에서, 세포 집단은 상기 세포 내로 도입된 예를 들어 본원에 기재된 바와 같은 gRNA 분자 또는 gRNA 분자들을 갖지 않은 유사한 세포 집단에 비해 F 세포에서 적어도 20% 증가, 예를 들어 적어도 21% 증가, 적어도 22% 증가, 적어도 23% 증가, 적어도 24% 증가, 적어도 25% 증가, 적어도 26% 증가, 적어도 27% 증가, 적어도 28% 증가 또는 적어도 29% 증가를 가진다(또는 이러한 증가를 갖는 적혈구 집단으로 예를 들어 생체내에서 분화할 수 있음). 실시형태에서, 세포 집단은 상기 세포 내로 도입된 예를 들어 본원에 기재된 바와 같은 gRNA 분자 또는 gRNA 분자들을 갖지 않은 유사한 세포 집단에 비해 F 세포에서 적어도 30% 증가, 예를 들어 적어도 35% 증가, 적어도 40% 증가, 적어도 45% 증가, 적어도 50% 증가, 적어도 55% 증가, 적어도 60% 증가, 적어도 65% 증가, 적어도 70% 증가, 적어도 75% 증가, 적어도 80% 증가, 적어도 85% 증가, 적어도 90% 증가 또는 적어도 95% 증가를 가진다(또는 이러한 증가를 갖는 적혈구 집단으로 예를 들어 생체내에서 분화할 수 있음). 실시형태에서, 세포 집단은 상기 세포 내로 도입된 예를 들어 본원에 기재된 바와 같은 gRNA 분자 또는 gRNA 분자들을 갖지 않은 유사한 세포 집단에 비해 F 세포에서 10 내지 90%, 20% 내지 80%, 20% 내지 70%, 20% 내지 60%, 20% 내지 50%, 20% 내지 40%, 20% 내지 30%, 25% 내지 80%,

25% 내지 70%, 25% 내지 60%, 25% 내지 50%, 25% 내지 40%, 25% 내지 35%, 25% 내지 30%, 30% 내지 80%, 30% 내지 70%, 30% 내지 60%, 30% 내지 50%, 30% 내지 40% 또는 30% 내지 35% 증가를 가진다(또는 이러한 증가를 갖는 적혈구 집단으로 예를 들어 생체내에서 분화할 수 있음). 실시형태에서, 예를 들어 본원에 기재된 방법에 의해 생산된 바와 같은 세포 집단은, 이를 필요로 하는 환자에게 도입 시 상기 환자에서 예를 들어 본원에 기재된 바와 같은 혈색소병증, 예를 들어 겸상 적혈구 질병 및/또는 베타 탈라쎄미아를 치료하기에 충분한 수의 세포 및/또는 F 세포 %에서의 충분한 증가, 예를 들어 치료적 유효량을 포함한다. 실시형태에서, F 세포에서의 증가는 예를 들어 본원에 기재된 바와 같은 적혈구계 분화 검정법에서 측정된 바와 같다.

[1036] 본원에 기재된 임의의 실시형태 및 양태를 포함한 실시형태에서, 본 발명은 1개 세포 당 적어도 6 피코그램의 태아 헤모글로빈을 생산하는 F 세포를 포함하는, 예를 들어 본원에 기재된 임의의 gRNA, 방법 및/또는 CRISPR 시스템에 의해 변형된 바와 같은 세포, 예를 들어 세포 집단에 관한 것이다. 실시형태에서, F 세포는 1개 세포 당 적어도 7 피코그램의 태아 헤모글로빈을 생산한다. 실시형태에서, F 세포는 1개 세포 당 적어도 8 피코그램의 태아 헤모글로빈을 생산한다. 실시형태에서, F 세포는 1개 세포 당 적어도 9 피코그램의 태아 헤모글로빈을 생산한다. 실시형태에서, F 세포는 1개 세포 당 적어도 10 피코그램의 태아 헤모글로빈을 생산한다. 실시형태에서, F 세포는 1개 세포 당 평균적으로 6.0 내지 7.0 피코그램, 7.0 내지 8.0 피코그램, 8.0 내지 9.0 피코그램, 9.0 내지 10.0 피코그램, 10.0 내지 11.0 피코그램, 또는 11.0 내지 12.0 피코그램의 태아 헤모글로빈을 생산한다.

[1037] 실시형태에서, 예를 들어 본원에 기재된 바와 같은(예를 들어 삽입결실, 예를 들어 표 7-2에 기재된 대형 결실 또는 삽입결실을 포함함) 세포 또는 세포 집단(또는 이의 자손)은, 이것이 도입된 대상체의 세포에서 예를 들어 본원에 기재된 방법을 사용하여 삽입결실을 검출함으로써 검출 가능하며, 예를 들어 검출 가능하게 남아 있다. 실시형태에서, 세포 또는 세포 집단(또는 이의 자손)은, 상기 세포 또는 세포 집단이 대상체 내로 도입된 후 적어도 10주, 적어도 14주, 적어도 16주, 적어도 18주, 적어도 20주, 적어도 30주 적어도 40주, 적어도 50주 또는 그 이상 동안, 이것이 도입된 상기 대상체에서 검출 가능하다.

[1038] 실시형태에서, 하나 이상의 삽입결실(예를 들어 표 7-2에 기재된 대형 결실 또는 삽입결실)은, 본원에 기재된 세포 또는 세포 집단이 도입된 대상체의 세포(예를 들어 골수 및/또는 말초 혈액의 세포, 예를 들어 CD34+ 세포)에서 본원에 기재된 방법, 예를 들어 NGS에 의해 검출 가능하며, 예를 들어 검출 가능하게 남아 있다. 실시형태에서, 하나 이상의 삽입결실은, 본원에 기재된 세포 또는 세포 집단이 대상체 내로 도입된 후, 본원에 기재된 세포 또는 세포 집단이 도입된 대상체의 세포(예를 들어 골수 및/또는 말초 혈액의 세포, 예를 들어 CD34+ 세포)에서 적어도 10주, 적어도 14주, 적어도 16주, 적어도 18주, 적어도 20주, 적어도 30주 적어도 40주, 적어도 50주 이상 동안 검출 가능하다. 실시형태에서, 상기 하나 이상의 삽입결실의 검출 수준은 시간이 경과함에 따라 저하되지 않거나, 예를 들어 이식-전에 측정되거나 이식-후 2주째에 또는 이식-후 8주째에 측정된 검출 수준(예를 들어 하나 이상의 삽입결실을 포함하는 세포의 퍼센트)에 비해 이식-후 20주째에 측정된 경우, (예를 들어 이식-전 삽입결실 검출 수준에 비해, 또는 이식-후 2주째 또는 이식-후 8주째 검출 수준에 비해) 5% 미만, 10% 미만, 15% 미만, 20% 미만, 30% 미만, 40% 미만 또는 50% 미만만큼 저하된다.

[1039] 임의의 상기 실시형태를 포함하는 실시형태에서, 본 발명의 세포 및/또는 세포 집단은 Cas9 분자를 인코딩하는 핵산을 포함하지 않는 세포를 포함하며, 예를 들어 이로 구성된다.

[1040] **치료 방법**

[1041] 전달 시간

[1042] 실시형태에서, Cas 시스템의 성분, 예를 들어 본원에 기재된 Cas9 분자 성분 및/또는 gRNA 분자 성분 외의 하나 이상의 핵산 분자(예를 들어 DNA 분자)가 전달된다. 실시형태에서, 핵산 분자는 Cas 시스템의 하나 이상의 성분이 전달되는 것과 동시에 전달된다. 실시형태에서, 핵산 분자는 Cas 시스템의 하나 이상의 성분이 전달되기(예를 들어 약 30분, 1시간, 2시간, 3시간, 6시간, 9시간, 12시간, 1일, 2일, 3일, 1주, 2주 또는 4주 미만) 전 또는 후에 전달된다. 실시형태에서, 핵산 분자는 Cas 시스템의 하나 이상의 성분, 예를 들어 Cas9 분자 성분 및/또는 gRNA 분자 성분이 전달되는 것과 상이한 수단에 의해 전달된다. 핵산 분자는 본원에 기재된 전달 방법 중 임의의 것에 의해 전달될 수 있다. 예를 들어 핵산 분자는 바이러스 벡터, 예를 들어 통합-결함있는 렌티바이러스에 의해 전달될 수 있고, Cas9 분자 성분 및/또는 gRNA 분자 성분은 전기천공에 의해 전달될 수 있어, 예를 들어 핵산에 의해 야기되는 독성(예를 들어 DNA)이 감소될 수 있다. 실시형태에서, 핵산 분자는 치료 단백질, 예를 들어 본원에 기재된 단백질을 인코딩한다. 실시형태에서, 핵산 분자는 RNA 분자, 예를 들어 본원에 기재된 RNA 분자를 인코딩한다.

[1043] 성분들의 두 가지 방식 또는 차등적 전달

[1044] Cas 시스템의 성분, 예를 들어 Cas9 분자 성분 및 gRNA 분자 성분의 별도의 전달, 더욱 구체적으로, 상이한 방식에 의한 성분들의 전달은, 예를 들어 조직 특이성 및 안전성을 개선함으로써 성능을 향상시킬 수 있다. 실시형태에서, Cas9 분자 및 gRNA 분자는 상이한 방식(또는 때때로 본원에서 차등적 방식으로 지칭됨)에 의해 전달된다. 본원에 사용되는 상이한 또는 차등적 방식은 대상체 성분 분자, 예를 들어 Cas9 분자, gRNA 분자, 주형 핵산, 또는 페이로드에 대해 상이한 약력학적 또는 약동학적 특성을 부여하는 전달 방식을 지칭한다. 예를 들어 전달 방식은 상이한 조직 분포, 상이한 반감기, 또는 예를 들어 선택된 구획, 조직 또는 기관에서의 상이한 시간 분포를 야기할 수 있다.

[1045] 일부 전달 방식, 예를 들어 자율적 복제 또는 세포 핵산 내로의 삽입에 의한, 세포 또는 세포의 자손에서 지속되는 핵산 벡터에 의한 전달은 성분의 보다 지속적인 발현 및 존재를 야기한다. 예로는 바이러스, 예를 들어 아데노 관련 바이러스 또는 렌티바이러스, 전달이 포함된다.

[1046] 예로서, 성분, 예를 들어 Cas9 분자 및 gRNA 분자는 신체, 또는 특정 구획, 조직 또는 기관에서의 전달된 성분의 결과적인 반감기 또는 지속성 면에서 상이한 방식들에 의해 전달될 수 있다. 실시형태에서, gRNA 분자는 이러한 방식으로 전달될 수 있다. Cas9 분자 성분은 신체 또는 특정 구획 또는 조직 또는 기관에 대한 더 적은 지속성 또는 더 적은 노출을 야기하는 방식에 의해 전달될 수 있다.

[1047] 보다 일반적으로, 실시형태에서, 제1 전달 방식은 제1 성분을 전달하기 위해 사용되고 제2 전달 방식은 제2 성분을 전달하기 위해 사용된다. 제1 전달 방식은 제1 약력학적 또는 약동학적 특성을 부여한다. 제1 약력학적 특성은 예를 들어 신체, 구획, 조직 또는 기관에서, 성분 또는 성분을 인코딩하는 핵산의 분포, 지속성 또는 노출일 수 있다. 제2 전달 방식은 제2 약력학적 또는 약동학적 특성을 부여한다. 제2 약력학적 특성은 예를 들어 신체, 구획, 조직 또는 기관에서, 성분 또는 성분을 인코딩하는 핵산의 분포, 지속성 또는 노출일 수 있다.

[1048] 실시형태에서, 제1 약력학적 또는 약동학적 특성, 예를 들어 분포, 지속성 또는 노출은 제2 약력학적 또는 약동학적 특성보다 더 제한된다.

[1049] 실시형태에서, 제1 전달 방식은 약력학적 또는 약동학적 특성, 예를 들어 분포, 지속성 또는 노출을 최적화, 예를 들어 최소화하기 위해 선택된다.

[1050] 실시형태에서, 제2 전달 방식은 약력학적 또는 약동학적 특성, 예를 들어 분포, 지속성 또는 노출을 최적화, 예를 들어 최대화하기 위해 선택된다.

[1051] 실시형태에서, 제1 전달 방식은 상대적으로 지속적인 요소, 예를 들어 핵산, 예를 들어 플라스미드 또는 바이러스 벡터, 예를 들어 AAV 또는 렌티바이러스의 사용을 포함한다. 이러한 벡터는 상대적으로 지속적이기 때문에 이들로부터 전사된 산물은 상대적으로 지속적일 것이다.

[1052] 실시형태에서, 제2 전달 방식은 상대적으로 일시적인 요소, 예를 들어 RNA 또는 단백질을 포함한다.

[1053] 실시형태에서, 제1 성분은 gRNA를 포함하고, 전달 방식은 상대적으로 지속적이며, 예를 들어 gRNA는 플라스미드 또는 바이러스 벡터, 예를 들어 AAV 또는 렌티바이러스로부터 전사된다. 이들 유전자의 전사는 유전자가 단백질 산물을 인코딩하지 않고, gR A 는 별개로 작용할 수 없기 때문에 생리적 중요성이 거의 없을 것이다. 제2 성분, Cas9 분자는 전체 Cas9 분자/gRNA 분자 복합체가 반드시 짧은 시간 동안에만 존재하고 활성있도록 하기 위해, 일시적 방식으로, 예를 들어 mRNA 또는 단백질로서 전달된다.

[1054] 추가적으로, 성분들은 안전성 및 조직 특이성을 향상시키기 위해 서로 보완하는 상이한 분자 형태로, 또는 상이한 전달 벡터로 전달될 수 있다.

[1055] 차등 전달 방식의 사용은 성능, 안전성 및 표능을 향상할 수 있다. 예를 들어 최종적인 표적-외 변형의 가능성이 감소될 수 있다. 박테리아-유래된 Cas 효소로부터의 펩티드가 MHC 분자에 의해 세포 표면 상에 디스플레이되기 때문에, 덜 지속적인 방식에 의한 면역원성 성분, 예를 들어 Cas9 분자의 전달은 면역원성을 감소시킬 수 있다. 2-부분 전달 시스템은 이들 단점을 경감시킬 수 있다.

[1056] 차등 전달 방식은 상이하지만 겹쳐지는 표적 영역에 성분을 전달하기 위하여 사용될 수 있다. 활성 복합체의 형성은 표적 영역의 중첩 외부에서 최소화된다. 따라서, 실시형태에서, 제1 성분, 예를 들어 gRNA 분자는 제1 공간, 예를 들어 조직, 분포를 야기하는 제1 전달 방식에 의해 전달된다. 제2 성분, 예를 들어 Cas9 분자는 제2 공간, 예를 들어 조직, 분포를 야기하는 제2 전달 방식에 의해 전달된다. 실시형태에서, 제1 방식은 리포좀, 나

노 입자, 예를 들어 중합체 나노 입자, 및 핵산, 예를 들어 바이러스 벡터로부터 선택된 제1 요소를 포함한다. 제2 방식은 군으로부터 선택된 제2 요소를 포함한다. 실시형태에서, 제1 전달 방식은 제1 표적화 요소, 예를 들어 세포 특이적 수용체 또는 항체를 포함하고, 제2 전달 방식은 그 요소를 포함하지 않는다. 실시형태에서, 제2 전달 방식은 제2 표적화 요소, 예를 들어 제2 세포 특이적 수용체 또는 제2 항체를 포함한다.

[1057] Cas9 분자가 바이러스 전달 벡터, 리포좀 또는 중합체 나노 입자에서 전달되는 경우, 단일 조직만을 표적화하는 것이 바람직할 수 있을 때, 다수 조직으로 전달되고 치료 활성이 있을 가능성이 있다. 2-부분 전달 시스템은 이 문제를 해결하고 조직 특이성을 향상시킬 수 있다. gRNA 분자와 Cas9 분자가 구별되지만 겹치는 조직 향성을 갖는 분리된 전달 비히클 내에 패키징되는 경우, 완전히 기능적인 복합체는 두 벡터에 의해 표적화된 조직에서만 형성된다.

[1058] 후보 Cas 분자, 예를 들어 Cas9 분자, 후보 gRNA 분자, 후보 Cas9 분자/gRNA 분자 복합체, 및 후보 CRISPR 시스템은 당분야에 공지된 방법에 의해 또는 본원에 기재된 바와 같이 평가될 수 있다. 예를 들어 Cas9 분자의 엔도뉴클레아제 활성을 평가하기 위한 예시적인 방법은 예를 들어 문헌[Jinek el al., SCIENCE 2012; 337(6096):816-821]에 기재되어 있다.

[1059] 부가적인 양태는 하기 나열된 실시형태에서 기재된다.

[1060] 실시형태:

[1061] 1. tracr 및 crRNA를 포함하는 gRNA 분자로서, 여기서, crRNA는

[1062] a) 비결실 HFPH 영역(예를 들어 인간 비결실 HPFH 영역)의 표적 서열과 상보적인 표적화 도메인;

[1063] b) Chr11:5,249,833 내지 Chr11:5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[1064] c) Chr11:5,254,738 내지 Chr11:5,255,164, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[1065] d) Chr11:5,250,094 내지 5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[1066] e) Chr11:5,255,022 내지 5,255,164, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[1067] f) Chr11:5,249,833 내지 5,249,927, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[1068] g) Chr11:5,254,738 내지 5,254,851, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인;

[1069] h) Chr11:5,250,139 내지 5,250,237, - 가닥, hg38에서 게놈 핵산 서열 내의 표적 서열에 상보적인 표적화 도메인; 또는

[1070] i) 이들의 조합

[1071] 을 포함하는, gRNA 분자.

[1072] 2. 실시형태 1에 있어서, 표적화 도메인이 SEQ ID NO: 1 내지 SEQ ID NO: 72 또는 이들의 단편 중 임의의 하나를 포함하는, 예를 들어 이로 구성되는 gRNA 분자.

[1073] 3. 실시형태 2에 있어서, 표적화 도메인이 SEQ ID NO: 1, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 28, SEQ ID NO: 34, SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 58, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 67 또는 이들의 단편 중 임의의 하나를 포함하는, 예를 들어 이로 구성되는 gRNA 분자.

[1074] 4. 실시형태 2에 있어서, 표적화 도메인이

[1075] a) SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 28, SEQ ID NO: 34, SEQ ID NO: 48, SEQ ID NO: 51, SEQ ID NO: 67 또는 이들의 단편; 또는

[1076] b) SEQ ID NO: 1, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 54 또는 이들의 단편

[1077] 중 임의의 하나를 포함하는, 예를 들어 이로 구성되는 gRNA 분자.

[1078] 5. 실시형태 2 내지 4 중 어느 하나에 있어서, 표적화 도메인이 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산을 포함하는, 예를 들어 이로 구성되는 gRNA 분자.

[1079] 6. 실시형태 5에 있어서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산이 나열된 표적화 도메인 서열의 3' 말단에 배치된 17, 18, 19 또는 20개 연속 핵산인, gRNA 분자.

[1080] 7. 실시형태 5에 있어서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산이 나열된 표적화 도메인 서열의 5' 말단에 배치된 17, 18, 19 또는 20개 연속 핵산인, gRNA 분자.

[1081] 8. 실시형태 5에 있어서, 나열된 표적화 도메인 서열 중 임의의 하나의 17, 18, 19 또는 20개 연속 핵산이 나열된 표적화 도메인 서열의 5' 또는 3' 핵산 중 어느 것도 포함하지 않는, gRNA 분자.

[1082] 9. 실시형태 2 내지 8 중 어느 하나에 있어서, 표적화 도메인이 나열된 표적화 도메인 서열로 구성되는, gRNA 분자.

[1083] 10. 실시형태 1 내지 9 중 어느 하나에 있어서, crRNA의 일부 및 tracr의 일부가 혼성화하여, SEQ ID NO: 182 또는 183을 포함하는 플래그폴을 형성하는, gRNA 분자.

[1084] 11. 실시형태 10에 있어서, 플래그폴이 이러한 플래그폴의 crRNA 부분에 대해 3'에 위치한 제1 플래그폴 연장을 추가로 포함하고, 상기 제1 플래그폴 연장이 SEQ ID NO: 184를 포함하는, gRNA 분자.

[1085] 12. 실시형태 10 또는 11에 있어서, 플래그폴이 이러한 플래그폴의 crRNA 부분에 대해 3'에 위치한 제2 플래그폴 연장, 및 존재한다면 제1 플래그폴 연장을 추가로 포함하고, 상기 제2 플래그폴 연장이 SEQ ID NO: 185를 포함하는, gRNA 분자.

[1086] 13. 실시형태 1 내지 12 중 어느 하나에 있어서, tracr가 SEQ ID NO: 224 또는 SEQ ID NO: 225를 포함하는, gRNA 분자.

[1087] 14. 실시형태 1 내지 13 중 어느 하나에 있어서, tracr가, 3' 말단에 부가적인 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 선택적으로 추가로 포함하는, SEQ ID NO: 232를 포함하는, gRNA 분자.

[1088] 15. 실시형태 1 내지 14 중 어느 하나에 있어서, crRNA가 5' 으로부터 3'까지, [표적화 도메인]-:

[1089] a) SEQ ID NO: 182;

[1090] b) SEQ ID NO: 183;

[1091] c) SEQ ID NO: 199;

[1092] d) SEQ ID NO: 200;

[1093] e) SEQ ID NO: 201;

[1094] f) SEQ ID NO: 202; 또는

[1095] g) SEQ ID NO: 226

[1096] 을 포함하는, gRNA 분자.

[1097] 16. 실시형태 1 내지 9 또는 15 중 어느 하나에 있어서, tracr가 5' 으로부터 3'까지:

[1098] a) SEQ ID NO: 187;

[1099] b) SEQ ID NO: 188;

[1100] c) SEQ ID NO: 203;

[1101] d) SEQ ID NO: 204;

[1102] e) SEQ ID NO: 224;

[1103] f) SEQ ID NO: 225;

[1104] g) SEQ ID NO: 232;

[1105] h) SEQ ID NO: 227;

[1106] i) SEQ ID NO: 228;

[1107] j) SEQ ID NO: 229;

[1108] k) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 j);

[1109] l) 3' 말단에 적어도 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 k); 또는

[1110] m) 5' 말단에(예를 들어 5' 말단에), 적어도 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드, 예를 들어 1, 2, 3, 4, 5, 6 또는 7개 아데닌(A) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 a) 내지 l)

[1111] 를 포함하는, gRNA 분자.

[1112] 17. 실시형태 1 내지 9 중 어느 하나에 있어서, 표적화 도메인 및 tracr가 개별 핵산 분자 상에 배치되고, 표적화 도메인을 포함하는 핵산 분자가 선택적으로 표적화 도메인에 대해 바로 3'에 배치된 SEQ ID NO: 201을 포함하고, tracr를 포함하는 핵산 분자가 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1113] 18. 실시형태 13 또는 14에 있어서, 플래그폴의 crRNA 부분이 SEQ ID NO: 201 또는 SEQ ID NO: 202를 포함하는, gRNA 분자.

[1114] 19. 실시형태 1 내지 12 중 어느 하나에 있어서, tracr가 SEQ ID NO: 187 또는 188, 및 선택적으로 제1 플래그폴 연장이 존재한다면 SEQ ID NO: 187 또는 188에 대해 5'에 배치된 제1 tracr 연장을 포함하고, 상기 제1 tracr 연장은 SEQ ID NO: 189를 포함하는, gRNA 분자.

[1115] 20. 실시형태 1 내지 19 중 어느 하나에 있어서, 표적화 도메인 및 tracr가 개별 핵산 분자 상에 배치되는, gRNA 분자.

[1116] 21. 실시형태 1 내지 19 중 어느 하나에 있어서, 표적화 도메인 및 tracr가 단일 핵산 분자 상에 배치되고, tracr가 표적화 도메인에 대해 3'에 배치되는, gRNA 분자.

[1117] 22. 실시형태 21에 있어서, 표적화 도메인에 대해 3' 및 tracr에 대해 5'에 배치된 루프를 추가로 포함하는, gRNA 분자.

[1118] 23. 실시형태 22에 있어서, 루프가 SEQ ID NO: 186을 포함하는, gRNA 분자.

[1119] 24. 실시형태 1 내지 9 중 어느 하나에 있어서, 5' 으로부터 3'까지, [표적화 도메인]-:

[1120] (a) SEQ ID NO: 195;

[1121] (b) SEQ ID NO: 196;

[1122] (c) SEQ ID NO: 197;

[1123] (d) SEQ ID NO: 198;

[1124] (e) SEQ ID NO: 231; 또는

[1125] (f) 3' 말단에 1, 2, 3, 4, 5, 6 또는 7개 우라실(U) 뉴클레오타이드를 추가로 포함하는, 임의의 상기 (a) 내지 (e)

[1126] 를 포함하는, gRNA 분자.

[1127] 25. 실시형태 1 내지 9 또는 21 내지 24 중 어느 하나에 있어서, 표적화 도메인 및 tracr가 단일 핵산 분자 상에 배치되고, 상기 핵산 분자가 예를 들어 상기 표적화 도메인, 및 선택적으로 상기 표적화 도메인에 대해 바로 3'에 배치된 SEQ ID NO: 231을 포함하는, 예를 들어 이로 구성되는 gRNA 분자.

[1128] 26. 실시형태 1 내지 25 중 어느 하나에 있어서, gRNA 분자를 포함하는 1개 또는 선택적으로 1개 초과의 핵산

분자가

[1129] a) 상기 핵산 분자 또는 분자들의 3' 말단에 하나 이상의, 예를 들어 3개의, 포스포로티오에이트 변형;

[1130] b) 상기 핵산 분자 또는 분자들의 5' 말단에 하나 이상의, 예를 들어 3개의, 포스포로티오에이트 변형;

[1131] c) 상기 핵산 분자 또는 분자들의 3' 말단에 하나 이상의, 예를 들어 3개의, 2'-O-메틸 변형;

[1132] d) 상기 핵산 분자 또는 분자들의 5' 말단에 하나 이상의, 예를 들어 3개의, 2'-O-메틸 변형;

[1133] e) 상기 핵산 분자 또는 분자들의 말단 방향으로 4 번째, 말단 방향으로 3 번째 및 말단 방향으로 2 번째 3' 잔기 각각에서 2' O-메틸 변형;

[1134] f) 상기 핵산 분자 또는 분자들의 말단 방향으로 4 번째, 말단 방향으로 3 번째 및 말단 방향으로 2 번째 5' 잔기 각각에서 2' O-메틸 변형; 또는

[1135] f) 이들의 임의의 조합

[1136] 을 포함하는, gRNA 분자.

[1137] 27. 실시형태 1에 있어서, 서열:

[1138] (a) SEQ ID NO: 74;

[1139] (b) SEQ ID NO: 75; 또는

[1140] (c) SEQ ID NO: 76

[1141] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1142] 28. 실시형태 1에 있어서,

[1143] (a) SEQ ID NO: 77을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1144] (b) SEQ ID NO: 77을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1145] (c) SEQ ID NO: 78을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1146] (d) SEQ ID NO: 78을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1147] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1148] 29. 실시형태 1에 있어서, 서열:

[1149] (a) SEQ ID NO: 79;

[1150] (b) SEQ ID NO: 80; 또는

[1151] (c) SEQ ID NO: 81

[1152] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1153] 30. 실시형태 1에 있어서,

[1154] (a) SEQ ID NO: 82를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1155] (b) SEQ ID NO: 82를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1156] (c) SEQ ID NO: 83을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1157] (d) SEQ ID NO: 83을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이

로 구성되는 tracr

[1158] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1159] 31. 실시형태 1에 있어서, 서열:

[1160] (a) SEQ ID NO: 84;

[1161] (b) SEQ ID NO: 85; 또는

[1162] (c) SEQ ID NO: 86

[1163] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1164] 32. 실시형태 1에 있어서,

[1165] (a) SEQ ID NO: 87을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1166] (b) SEQ ID NO: 87을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1167] (c) SEQ ID NO: 88을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1168] (d) SEQ ID NO: 88을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1169] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1170] 33. 실시형태 1에 있어서, 서열:

[1171] (a) SEQ ID NO: 89;

[1172] (b) SEQ ID NO: 90; 또는

[1173] (c) SEQ ID NO: 91

[1174] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1175] 34. 실시형태 1에 있어서,

[1176] (a) SEQ ID NO: 92를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1177] (b) SEQ ID NO: 92를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1178] (c) SEQ ID NO: 93을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1179] (d) SEQ ID NO: 93을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1180] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1181] 35. 실시형태 1에 있어서, 서열:

[1182] (a) SEQ ID NO: 94;

[1183] (b) SEQ ID NO: 95; 또는

[1184] (c) SEQ ID NO: 96

[1185] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1186] 36. 실시형태 1에 있어서,

[1187] (a) SEQ ID NO: 97을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1188] (b) SEQ ID NO: 97을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1189] (c) SEQ ID NO: 98을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1190] (d) SEQ ID NO: 98을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1191] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1192] 37. 실시형태 1에 있어서, 서열:

[1193] (a) SEQ ID NO: 99;

[1194] (b) SEQ ID NO: 100; 또는

[1195] (c) SEQ ID NO: 101

[1196] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1197] 38. 실시형태 1에 있어서,

[1198] (a) SEQ ID NO: 102를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1199] (b) SEQ ID NO: 102를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1200] (c) SEQ ID NO: 103을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1201] (d) SEQ ID NO: 103을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1202] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1203] 39. 실시형태 1에 있어서, 서열:

[1204] (a) SEQ ID NO: 104;

[1205] (b) SEQ ID NO: 105; 또는

[1206] (c) SEQ ID NO: 106

[1207] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1208] 40. 실시형태 1에 있어서,

[1209] (a) SEQ ID NO: 107을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1210] (b) SEQ ID NO: 107을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1211] (c) SEQ ID NO: 108을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1212] (d) SEQ ID NO: 108을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1213] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1214] 41. 실시형태 1에 있어서, 서열:

[1215] (a) SEQ ID NO: 109;

[1216] (b) SEQ ID NO: 110; 또는

[1217] (c) SEQ ID NO: 111

[1218] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1219] 42. 실시형태 1에 있어서,

[1220] (a) SEQ ID NO: 112를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1221] (b) SEQ ID NO: 112를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1222] (c) SEQ ID NO: 113을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1223] (d) SEQ ID NO: 113을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1224] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1225] 43. 실시형태 1에 있어서, 서열:

[1226] (a) SEQ ID NO: 114;

[1227] (b) SEQ ID NO: 115; 또는

[1228] (c) SEQ ID NO: 116

[1229] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1230] 44. 실시형태 1에 있어서,

[1231] (a) SEQ ID NO: 117을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1232] (b) SEQ ID NO: 117을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1233] (c) SEQ ID NO: 118을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1234] (d) SEQ ID NO: 118을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1235] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1236] 45. 실시형태 1에 있어서, 서열:

[1237] (a) SEQ ID NO: 119;

[1238] (b) SEQ ID NO: 120; 또는

[1239] (c) SEQ ID NO: 121

[1240] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1241] 46. 실시형태 1에 있어서,

[1242] (a) SEQ ID NO: 122를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1243] (b) SEQ ID NO: 122를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1244] (c) SEQ ID NO: 123을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1245] (d) SEQ ID NO: 123을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1246] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1247] 47. 실시형태 1에 있어서, 서열:

[1248] (a) SEQ ID NO: 124;

[1249] (b) SEQ ID NO: 125; 또는

[1250] (c) SEQ ID NO: 126

[1251] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1252] 48. 실시형태 1에 있어서,

[1253] (a) SEQ ID NO: 127을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1254] (b) SEQ ID NO: 127을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1255] (c) SEQ ID NO: 128을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1256] (d) SEQ ID NO: 128을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1257] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1258] 49. 실시형태 1에 있어서, 서열:

[1259] (a) SEQ ID NO: 129;

[1260] (b) SEQ ID NO: 130; 또는

[1261] (c) SEQ ID NO: 131

[1262] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1263] 50. 실시형태 1에 있어서,

[1264] (a) SEQ ID NO: 132를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1265] (b) SEQ ID NO: 132를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1266] (c) SEQ ID NO: 133을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1267] (d) SEQ ID NO: 133을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1268] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1269] 51. 실시형태 1에 있어서, 서열:

[1270] (a) SEQ ID NO: 134;

[1271] (b) SEQ ID NO: 135; 또는

[1272] (c) SEQ ID NO: 136

[1273] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1274] 52. 실시형태 1에 있어서,

[1275] (a) SEQ ID NO: 137을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1276] (b) SEQ ID NO: 137을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1277] (c) SEQ ID NO: 138을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1278] (d) SEQ ID NO: 138을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1279] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1280] 53. 실시형태 1에 있어서, 서열:

[1281] (a) SEQ ID NO: 139;

[1282] (b) SEQ ID NO: 140; 또는

[1283] (c) SEQ ID NO: 141

[1284] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1285] 54. 실시형태 1에 있어서,

[1286] (a) SEQ ID NO: 142를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1287] (b) SEQ ID NO: 142를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1288] (c) SEQ ID NO: 143을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1289] (d) SEQ ID NO: 143을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1290] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1291] 55. 실시형태 1에 있어서, 서열:

[1292] (a) SEQ ID NO: 144;

[1293] (b) SEQ ID NO: 145; 또는

[1294] (c) SEQ ID NO: 146

[1295] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1296] 56. 실시형태 1에 있어서,

[1297] (a) SEQ ID NO: 147을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1298] (b) SEQ ID NO: 147을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1299] (c) SEQ ID NO: 148을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어

이로 구성되는 tracr; 또는

[1300] (d) SEQ ID NO: 148을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1301] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1302] 57. 실시형태 1에 있어서, 서열:

[1303] (a) SEQ ID NO: 149;

[1304] (b) SEQ ID NO: 150; 또는

[1305] (c) SEQ ID NO: 151

[1306] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1307] 58. 실시형태 1에 있어서,

[1308] (a) SEQ ID NO: 152를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1309] (b) SEQ ID NO: 152를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1310] (c) SEQ ID NO: 153을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1311] (d) SEQ ID NO: 153을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1312] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1313] 59. 실시형태 1에 있어서, 서열:

[1314] (a) SEQ ID NO: 154;

[1315] (b) SEQ ID NO: 155; 또는

[1316] (c) SEQ ID NO: 156

[1317] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1318] 60. 실시형태 1에 있어서,

[1319] (a) SEQ ID NO: 157을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1320] (b) SEQ ID NO: 157을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1321] (c) SEQ ID NO: 158을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1322] (d) SEQ ID NO: 158을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1323] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1324] 61. 실시형태 1에 있어서, 서열:

[1325] (a) SEQ ID NO: 159;

[1326] (b) SEQ ID NO: 160; 또는

[1327] (c) SEQ ID NO: 161

[1328] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1329] 62. 실시형태 1에 있어서,

[1330] (a) SEQ ID NO: 162를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1331] (b) SEQ ID NO: 162를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1332] (c) SEQ ID NO: 163을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1333] (d) SEQ ID NO: 163을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1334] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1335] 63. 실시형태 1에 있어서, 서열:

[1336] (a) SEQ ID NO: 164;

[1337] (b) SEQ ID NO: 165; 또는

[1338] (c) SEQ ID NO: 166

[1339] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1340] 64. 실시형태 1에 있어서,

[1341] (a) SEQ ID NO: 167을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1342] (b) SEQ ID NO: 167을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1343] (c) SEQ ID NO: 168을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1344] (d) SEQ ID NO: 168을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1345] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1346] 65. 실시형태 1에 있어서, 서열:

[1347] (a) SEQ ID NO: 169;

[1348] (b) SEQ ID NO: 170; 또는

[1349] (c) SEQ ID NO: 171

[1350] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1351] 66. 실시형태 1에 있어서,

[1352] (a) SEQ ID NO: 172를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1353] (b) SEQ ID NO: 172를 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1354] (c) SEQ ID NO: 173을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1355] (d) SEQ ID NO: 173을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이

로 구성되는 tracr

[1356] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1357] 67. 실시형태 1에 있어서, 서열:

[1358] (a) SEQ ID NO: 174;

[1359] (b) SEQ ID NO: 175; 또는

[1360] (c) SEQ ID NO: 176

[1361] 을 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1362] 68. 실시형태 1에 있어서,

[1363] (a) SEQ ID NO: 177을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr;

[1364] (b) SEQ ID NO: 177을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr;

[1365] (c) SEQ ID NO: 178을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 224를 포함하는, 예를 들어 이로 구성되는 tracr; 또는

[1366] (d) SEQ ID NO: 178을 포함하는, 예를 들어 이로 구성되는 crRNA, 및 SEQ ID NO: 73을 포함하는, 예를 들어 이로 구성되는 tracr

[1367] 를 포함하는, 예를 들어 이로 구성되는, gRNA 분자.

[1368] 69. 실시형태 1 내지 68 중 어느 하나에 있어서,

[1369] a) gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 삽입결실이 gRNA 분자의 표적화 도메인에 상보적인 표적 서열에서 또는 그 부근에서 형성되고/되거나;

[1370] b) gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, HBG1 프로모터 영역에서 gRNA 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 gRNA 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열을 포함하는, 예를 들어 실질적으로 모든 서열을 포함하는 결실이 생성되는, gRNA 분자.

[1371] 70. 실시형태 69에 있어서, 삽입결실이 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않고, 선택적으로 삽입결실이 비결실 HPFH 또는 전사 인자 결합 부위의 뉴클레오타이드를 포함하지 않는, gRNA 분자.

[1372] 71. 실시형태 1 내지 70 중 어느 한 항에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내로 도입될 때, 집단의 세포 중 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 40%, 예를 들어 적어도 약 50%, 예를 들어 적어도 약 55%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 75%, 예를 들어 적어도 약 80%, 예를 들어 적어도 약 85%, 예를 들어 적어도 약 90%, 예를 들어 적어도 약 95%에서 gRNA 분자의 표적화 도메인에 상보적인 표적 서열에서 또는 그 부근에서 삽입결실이 형성되는, gRNA 분자.

[1373] 72. 실시형태 69 내지 71 중 어느 하나에 있어서, 삽입결실이 HBG1 프로모터 영역의 적어도 하나의 뉴클레오타이드 또는 HBG2 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는, gRNA 분자.

[1374] 73. 실시형태 71 또는 72에 있어서, 집단의 세포 중 적어도 약 15%가, HBG1 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실 및 HBG2 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실을 포함하는, gRNA 분자.

[1375] 74. 실시형태 71 내지 73 중 어느 하나에 있어서, HBG1 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실을 포함하는 집단의 세포의 퍼센트가, HBG2 프로모터 영역의 적어도 하나의 뉴클레오타이드를 포함하는 삽입결실을 포함하는 집단의 세포의 퍼센트와 적어도 약 5%, 예를 들어 적어도 약 10%, 예를 들어 적어

도 약 20%, 예를 들어 적어도 약 30%만큼 상이한, gRNA 분자.

[1376] 75. 실시형태 69 내지 74 중 어느 하나에 있어서, 삽입결실이 차세대 시퀀싱(NGS)에 의해 측정된 바와 같은 것인, gRNA 분자.

[1377] 76. 실시형태 1 내지 75 중 어느 하나에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 태아 헤모글로빈의 발현이 상기 세포 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손에서 증가되는, gRNA 분자.

[1378] 77. 실시형태 76에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내에 도입될 때, 상기 집단 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손의 집단 내의 F 세포의 퍼센트가, gRNA 분자가 도입되지 않은 세포 집단 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손의 집단 내의 F 세포의 퍼센트에 비해, 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 25%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 35%, 예를 들어 적어도 약 40%만큼 증가되는, gRNA 분자.

[1379] 78. 실시형태 76에 있어서, 상기 세포 또는 이의 자손, 예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손이 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, gRNA 분자.

[1380] 79. 실시형태 1 내지 78 중 어느 하나에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 내로 도입될 때, 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실도 상기 세포에서 형성되지 않으며, 예를 들어 어떠한 표적-외 삽입결실도 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 형성되지 않는, gRNA 분자.

[1381] 80. 실시형태 1 내지 78 중 어느 하나에 있어서, gRNA 분자를 포함하는 CRISPR 시스템(예를 들어 본원에 기재된 바와 같은 RNP)이 세포 집단 내에 도입될 때, 예를 들어 차세대 시퀀싱 및/또는 뉴클레오타이드 삽입 검정법에 의해 검출 가능한 바와 같이 어떠한 표적-외 삽입결실, 예를 들어 HBG1 및/또는 HBG2 프로모터 영역의 외부에서 어떠한 표적-외 삽입결실도 세포 집단의 세포 중 약 5% 초과, 예를 들어 약 1% 초과, 예를 들어 약 0.1% 초과, 예를 들어 약 0.01% 초과에서 검출되지 않는, gRNA 분자.

[1382] 81. 실시형태 69 내지 80 중 어느 하나에 있어서, 세포(또는 세포 집단)가 포유류, 영장류 또는 인간 세포이며, 예를 들어 인간 세포인, gRNA 분자.

[1383] 82. 실시형태 81에 있어서, 세포가 HSPC인(또는 세포 집단이 이를 포함하는), gRNA 분자.

[1384] 83. 실시형태 82에 있어서, HSPC가 CD34+인, gRNA 분자.

[1385] 84. 실시형태 83에 있어서, HSPC가 CD34+CD90+인, gRNA 분자.

[1386] 85. 실시형태 69 내지 84 중 어느 하나에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가인, gRNA 분자.

[1387] 86. 실시형태 69 내지 84 중 어느 하나에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 동종이계인, gRNA 분자.

[1388] 87. 조성물로서,

[1389] 1) 실시형태 1 내지 86 중 어느 하나의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함) 및 Cas9 분자;

[1390] 2) 실시형태 1 내지 86 중 어느 하나의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함) 및 Cas9 분자를 인코딩하는 핵산;

[1391] 3) 실시형태 1 내지 86 중 어느 하나의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함)를 인코딩하는 핵산 및 Cas9 분자;

[1392] 4) 실시형태 1 내지 86 중 어느 하나의 하나 이상의 gRNA 분자(제1 gRNA 분자를 포함함)를 인코딩하는 핵산 및 Cas9 분자를 인코딩하는 핵산; 또는

[1393] 5) 임의의 상기 1) 내지 4), 및 주형 핵산; 또는

[1394] 6) 임의의 상기 1) 내지 4), 및 주형 핵산을 인코딩하는 서열을 포함하는 핵산

[1395] 을 포함하는, 조성물.

[1396] 88. 실시형태 1 내지 86 중 어느 하나의 제1 gRNA 분자를 포함하고, 선택적으로 Cas9 분자를 추가로 포함하는, 조성물.

[1397] 89. 실시형태 87 또는 88에 있어서, Cas9 분자가 활성 또는 불활성 *S. 피오게네스* Cas9인, 조성물.

[1398] 90. 실시형태 87 내지 89 중 어느 하나에 있어서, Cas9 분자가 SEQ ID NO: 205를 포함하는, 조성물.

[1399] 91. 실시형태 87 내지 89 중 어느 하나에 있어서, Cas9 분자가

[1400] (a) SEQ ID NO: 233;

[1401] (b) SEQ ID NO: 234;

[1402] (c) SEQ ID NO: 235;

[1403] (d) SEQ ID NO: 236;

[1404] (e) SEQ ID NO: 237;

[1405] (f) SEQ ID NO: 238;

[1406] (g) SEQ ID NO: 239;

[1407] (h) SEQ ID NO: 240;

[1408] (i) SEQ ID NO: 241;

[1409] (j) SEQ ID NO: 242;

[1410] (k) SEQ ID NO: 243; 또는

[1411] (l) SEQ ID NO: 244

[1412] 를 포함하는, 예를 들어 이로 구성되는, 조성물.

[1413] 92. 실시형태 88 내지 91 중 어느 하나에 있어서, 제1 gRNA 분자 및 Cas9 분자가 리보핵 단백질 복합체(RNP)에 존재하는, 조성물.

[1414] 93. 실시형태 87 내지 92 중 어느 하나에 있어서, 제2 gRNA 분자; 제2 gRNA 분자 및 제3 gRNA 분자; 또는 제2 gRNA 분자, 선택적으로, 제3 gRNA 분자, 및 선택적으로, 제4 gRNA 분자를 추가로 포함하고, 여기서, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자가 실시형태 1 내지 68 중 어느 하나의 gRNA 분자이고, 조성물의 각각의 gRNA 분자가 상이한 표적 서열에 상보적인, 조성물.

[1415] 94. 실시형태 93에 있어서, 제1 gRNA 분자, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자 중 2개 이상이 동일한 유전자 또는 영역 내의 표적 서열에 상보적인, 조성물.

[1416] 95. 실시형태 93 또는 94에 있어서, 제1 gRNA 분자, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자가 6000개 이하의 뉴클레오타이드, 5000개 이하의 뉴클레오타이드, 500개 이하, 400개 이하의 뉴클레오타이드, 300개 이하, 200개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 90개 이하의 뉴클레오타이드, 80개 이하의 뉴클레오타이드, 70개 이하의 뉴클레오타이드, 60개 이하의 뉴클레오타이드, 50개 이하의 뉴클레오타이드, 40개 이하의 뉴클레오타이드, 30개 이하의 뉴클레오타이드, 20개 이하의 뉴클레오타이드 또는 10개 이하의 뉴클레오타이드 이격되어 존재하는 표적 서열에 상보적인, 조성물.

[1417] 96. 실시형태 93에 있어서, 제1 gRNA 분자, 제2 gRNA 분자, 선택적인 제3 gRNA 분자 및 선택적인 제4 gRNA 분자 중 2개 이상이, HBG1 프로모터 영역의 표적 서열에 상보적인 표적화 도메인을 포함하는 적어도 하나의 gRNA 분자 및 HBG2 프로모터 영역의 표적 서열에 상보적인 표적화 도메인을 포함하는 적어도 하나의 gRNA 분자를 포함하는, 조성물.

[1418] 97. 실시형태 94 또는 95에 있어서, 제1 gRNA 분자 및 제2 gRNA 분자를 포함하고, 여기서, 제1 gRNA 분자 및 제2 gRNA 분자가

[1419] (a) 실시형태 1의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나;

[1420] (b) 실시형태 2의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나;

[1421] c) 실시형태 3의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나;

[1422] (d) 실시형태 4의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적이거나;

[1423] (e) 실시형태 27 내지 68 중 어느 하나의 gRNA 분자로부터 독립적으로 선택되고, 상이한 표적 서열에 상보적인 것인, 조성물.

[1424] 98. 실시형태 94 내지 96 중 어느 하나에 있어서, 제1 gRNA 분자 및 제2 gRNA 분자를 포함하고, 여기서,

[1425] a) 제1 gRNA 분자가:

[1426] i) Chr11:5,249,833 내지 Chr11:5,250,237(hg38);

[1427] ii) Chr11:5,250,094 내지 5,250,237(hg38);

[1428] iii) Chr11:5,249,833 내지 5,249,927(hg38); 또는

[1429] iv) Chr11:5,250,139 내지 5,250,237(hg38)

[1430] 내의 적어도 1개 뉴클레오타이드를 포함하는(예를 들어 20개 연속 뉴클레오타이드를 포함하는) 표적 서열에 상보적이고;

[1431] b) 제2 gRNA 분자가:

[1432] i) Chr11:5,254,738 내지 Chr11:5,255,164(hg38);

[1433] ii) Chr11:5,255,022 내지 5,255,164(hg38); 또는

[1434] iii) Chr11:5,254,738 내지 5,254,851(hg38)

[1435] 내의 적어도 1개 뉴클레오타이드를 포함하는(예를 들어 20개 연속 뉴클레오타이드를 포함하는) 표적 서열에 상보적인 것인, 조성물.

[1436] 99. 실시형태 87 내지 108 중 어느 하나에 있어서, 조성물의 gRNA 분자 구성성분에 관하여, 조성물이 제1 gRNA 분자 및 제2 gRNA 분자로 구성된 것인, 조성물.

[1437] 100. 실시형태 87 내지 109 중 어느 하나에 있어서, 각각의 상기 gRNA 분자가 본원에 기재된 Cas9 분자, 예를 들어 실시형태 90 또는 91의 Cas9 분자와의 리보핵 단백질 복합체(RNP)에 존재하는, 조성물.

[1438] 101. 실시형태 87 내지 100 중 어느 하나에 있어서, 주형 핵산을 포함하고, 여기서, 주형 핵산이 제1 gRNA 분자의 표적 서열에서 또는 그 부근에서 뉴클레오타이드에 상응하는 뉴클레오타이드를 포함하는, 조성물.

[1439] 102. 실시형태 101에 있어서, 주형 핵산이

[1440] (a) 인간 베타 글로빈, 예를 들어 돌연변이 G16D, E22A 및 T87Q 중 하나 이상을 포함하는 인간 베타 글로빈, 또는 이의 단편; 또는

[1441] (b) 인간 감마 글로빈, 또는 이의 단편

[1442] 을 인코딩하는 핵산을 포함하는, 조성물.

[1443] 103. 실시형태 87 내지 102 중 어느 하나에 있어서, 전기천공에 적합한 배지 내에서 제제화되는, 조성물.

[1444] 104. 실시형태 87 내지 103 중 어느 하나에 있어서, 각각의 상기 gRNA 분자가 본원에 기재된 Cas9 분자와의 RNP에 존재하고, 여기서, 각각의 상기 RNP가 약 10 uM 미만, 예를 들어 약 3 uM 미만, 예를 들어 약 1 uM 미만, 예를 들어 약 0.5 uM 미만, 예를 들어 약 0.3 uM 미만, 예를 들어 약 0.1 uM 미만의 농도로 존재하고, 선택적으로 상기 RNP의 농도가 약 2 uM이거나 약 1 uM이며, 선택적으로, 조성물이 세포, 예를 들어 HSPC 집단을 추가로 포함하는, 조성물.

[1445] 105. 실시형태 1 내지 68 중 어느 하나의 하나 이상의 gRNA 분자를 인코딩하는 핵산 서열.

[1446] 106. 실시형태 105에 있어서, 핵산이 하나 이상의 gRNA 분자를 인코딩하는 서열에 작동적으로 연결된 프로모터

를 포함하는, 핵산 서열.

[1447] 107. 실시형태 106에 있어서, 프로모터가 RNA 중합효소 II 또는 RNA 중합효소 III에 의해 인지되는 프로모터인, 핵산 서열.

[1448] 108. 실시형태 107에 있어서, 프로모터가 U6 프로모터 또는 HI 프로모터인, 핵산 서열.

[1449] 109. 실시형태 105 내지 108 중 어느 하나에 있어서, 핵산이 Cas9 분자를 추가로 인코딩하는, 핵산 서열.

[1450] 110. 실시형태 109에 있어서, Cas9 분자가 SEQ ID NO: 205, SEQ ID NO: 233, SEQ ID NO: 234, SEQ ID NO: 235, SEQ ID NO: 236, SEQ ID NO: 237, SEQ ID NO: 238, SEQ ID NO: 239, SEQ ID NO: 240, SEQ ID NO: 241, SEQ ID NO: 242, SEQ ID NO: 243 또는 SEQ ID NO: 244 중 임의의 것을 포함하는, 핵산 서열.

[1451] 111. 실시형태 109 또는 110에 있어서, 상기 핵산이 Cas9 분자를 인코딩하는 서열에 작동적으로 연결된 프로모터를 포함하는, 핵산 서열.

[1452] 112. 실시형태 111에 있어서, 프로모터가 EF-1 프로모터, CMV IE 유전자 프로모터, EF-1α 프로모터, 유비퀴틴 C 프로모터 또는 포스포글리세레이트 키나제(PGK) 프로모터인, 핵산 서열.

[1453] 113. 실시형태 105 내지 112 중 어느 하나의 핵산을 포함하는, 벡터.

[1454] 114. 실시형태 113에 있어서, 벡터가 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노-연관 바이러스(AAV) 벡터, 단순 포진 바이러스(HSV) 벡터, 플라스미드, 미니서클, 나노플라스미드 및 RNA 벡터로 구성된 군으로부터 선택되는, 벡터.

[1455] 115. 상기 세포 내의 표적 서열에서 또는 그 부근에서 세포(예를 들어 세포 집단)를 변경시키는(예를 들어 핵산의 구조(예를 들어 서열)를 변경시키는) 방법으로서, 상기 세포(예를 들어 세포 집단)를

[1456] 1) 실시형태 1 내지 68 중 어느 하나의 하나 이상의 gRNA 분자 및 Cas9 분자;

[1457] 2) 실시형태 1 내지 68 중 어느 하나의 하나 이상의 gRNA 분자 및 Cas9 분자를 인코딩하는 핵산;

[1458] 3) 실시형태 1 내지 68 중 어느 하나의 하나 이상의 gRNA 분자를 인코딩하는 핵산 및 Cas9 분자;

[1459] 4) 실시형태 1 내지 68 중 어느 하나의 하나 이상의 gRNA 분자를 인코딩하는 핵산 및 Cas9 분자를 인코딩하는 핵산;

[1460] 5) 임의의 상기 1) 내지 4), 및 주형 핵산;

[1461] 6) 임의의 상기 1) 내지 4), 및 주형 핵산을 인코딩하는 서열을 포함하는 핵산;

[1462] 7) 실시형태 87 내지 104 중 어느 하나의 조성물; 또는

[1463] 8) 실시형태 113 또는 114의 벡터

[1464] 와 접촉시키는(예를 들어 이것 내로 도입시키는) 단계를 포함하는, 방법.

[1465] 116. 실시형태 115에 있어서, gRNA 분자 또는 이러한 gRNA 분자를 인코딩하는 핵산, 및 Cas9 분자 또는 이러한 Cas9 분자를 인코딩하는 핵산이 단일 조성물에서 제제화되는, 방법.

[1466] 117. 실시형태 115에 있어서, gRNA 분자 또는 이러한 gRNA 분자를 인코딩하는 핵산, 및 Cas9 분자 또는 이러한 Cas9 분자를 인코딩하는 핵산이 1개 초과의 조성물에서 제제화되는, 방법.

[1467] 118. 실시형태 117에 있어서, 1개 초과의 조성물이 동시에 또는 순차적으로 전달되는, 방법.

[1468] 119. 실시형태 115 내지 118 중 어느 하나에 있어서, 세포가 동물 세포인, 방법.

[1469] 120. 실시형태 115 내지 118 중 어느 하나에 있어서, 세포가 포유류, 영장류 또는 인간 세포인, 방법.

[1470] 121. 실시형태 120에 있어서, 세포가 조혈 줄기 또는 전구 세포(HSPC)(예를 들어 HSPC의 집단)인, 방법.

[1471] 122. 실시형태 115 내지 121 중 어느 하나에 있어서, 세포가 CD34+ 세포인, 방법.

[1472] 123. 실시형태 115 내지 122 중 어느 하나에 있어서, 세포가 CD34+CD90+ 세포인, 방법.

[1473] 124. 실시형태 115 내지 123 중 어느 하나에 있어서, 세포가 CD34+ 세포에 대해 농화된 세포 집단을 포함하는

조성물에 배치되는, 방법.

[1474] 125. 실시형태 115 내지 124 중 어느 하나에 있어서, 세포(예를 들어 세포 집단)이 골수, 동원된 말초 혈액, 또는 제대혈로부터 단리된 것인, 방법.

[1475] 126. 실시형태 115 내지 125 중 어느 하나에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가 또는 동종이계이며, 선택적으로 환자가 혈색소병증 환자이고, 선택적으로 환자가 겸상 적혈구 질병 또는 탈라쎄미아, 선택적으로 베타 탈라쎄미아를 갖고 있는, 방법.

[1476] 127. 실시형태 115 내지 126 중 어느 하나에 있어서,

[1477] a) 변경이 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 초래하고/하거나;

[1478] b) 변경이, HBG1 프로모터 영역에서 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 하나 이상의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 초래하고, 선택적으로 결실이 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는, 방법.

[1479] 128. 실시형태 127에 있어서, 삽입결실이 약 40개 미만 뉴클레오타이드, 예를 들어 30개 미만 뉴클레오타이드, 예를 들어 20개 미만 뉴클레오타이드, 예를 들어 10개 미만 뉴클레오타이드의 삽입 또는 결실인, 방법.

[1480] 129. 실시형태 128에 있어서, 삽입결실이 단일 뉴클레오타이드 결실인, 방법.

[1481] 130. 실시형태 127 내지 129 중 어느 하나에 있어서, 방법이 세포 집단을 초래하며, 여기서, 집단 중 적어도 약 15%, 예를 들어 적어도 약 17%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%, 예를 들어 적어도 약 40%, 예를 들어 적어도 약 50%, 예를 들어 적어도 약 55%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 75%, 예를 들어 적어도 약 80%, 예를 들어 적어도 약 85%, 예를 들어 적어도 약 90%, 예를 들어 적어도 약 95%가 변경되었으며, 예를 들어 삽입결실, 선택적으로 표 2 내지 표 7에 열거된 삽입결실로부터 선택되는 삽입결실을 포함하고, 집단의 세포가 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실은 포함하지 않는, 방법.

[1482] 131. 실시형태 115 내지 130 중 어느 하나에 있어서, 변경이 적혈구계 계통의 분화된 세포(예를 들어 적혈구)로 분화할 수 있는 세포(예를 들어 세포 집단)를 초래하고, 상기 분화된 세포가 예를 들어 비변경된 세포(예를 들어 세포 집단)에 비해 증가된 수준의 태아 헤모글로빈을 나타내는, 방법.

[1483] 132. 실시형태 115 내지 131 중 어느 하나에 있어서, 변경이 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있는 세포 집단을 초래하고, 분화된 세포의 상기 집단이 예를 들어 비변경된 세포의 집단에 비해 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 갖는, 방법.

[1484] 133. 실시형태 115 내지 131 중 어느 하나에 있어서, 변경이 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있는 세포를 초래하고, 상기 분화된 세포가 1개 세포 당 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, 방법.

[1485] 134. 실시형태 115 내지 133 중 어느 하나의 방법에 의해 변경된 세포, 또는 실시형태 115 내지 133 중 어느 하나의 방법에 의해 수득 가능한 세포.

[1486] 135. 표 7-2에 기재된 삽입결실을 포함하는 세포로서, 선택적으로 세포는 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않는, 세포.

[1487] 136. 실시형태 1 내지 68 중 어느 하나의 제1 gRNA 분자, 또는 실시형태 87 내지 104 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 또는 실시형태 113 또는 114의 벡터를 포함하는 세포.

[1488] 137. 실시형태 136에 있어서, Cas9 분자를 포함하는, 세포.

[1489] 138. 실시형태 137에 있어서, Cas9 분자가 SEQ ID NO: 205, SEQ ID NO: 233, SEQ ID NO: 234, SEQ ID NO: 235,

SEQ ID NO: 236, SEQ ID NO: 237, SEQ ID NO: 238, SEQ ID NO: 239, SEQ ID NO: 240, SEQ ID NO: 241, SEQ ID NO: 242, SEQ ID NO: 243 또는 SEQ ID NO: 244 중 임의의 것을 포함하는, 세포.

[1490] 139. 실시형태 134 내지 138 중 어느 하나에 있어서, 세포가 실시형태 1 내지 68 중 어느 하나의 제2 gRNA 분자, 또는 실시형태 1 내지 68 중 어느 하나의 제2 gRNA 분자를 인코딩하는 핵산을 포함하거나, 포함하였거나, 포함할 것이며, 여기서, 제1 gRNA 분자 및 제2 gRNA 분자가 동일하지 않은 표적화 도메인을 포함하는, 세포.

[1491] 140. 실시형태 134 내지 139 중 어느 하나에 있어서, 태아 헤모글로빈의 발현이 gRNA 분자를 포함하도록 변형되지 않은 동일한 세포 유형의 세포 또는 이의 자손에 비해, 상기 세포 또는 이의 자손(예를 들어 이의 적혈구계 자손, 예를 들어 이의 적혈구 자손)에서 증가되는, 세포.

[1492] 141. 실시형태 134 내지 139 중 어느 하나에 있어서, 세포가 분화된 세포, 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구)로 분화할 수 있고, 여기서, 상기 분화된 세포가 예를 들어 gRNA 분자를 포함하도록 변형되지 않은 동일한 유형의 세포에 비해, 증가된 수준의 태아 헤모글로빈을 나타내는, 세포.

[1493] 142. 실시형태 140 또는 141에 있어서, 분화된 세포(예를 들어 적혈구계 계통의 세포, 예를 들어 적혈구)가, 예를 들어 gRNA 분자를 포함하도록 변형되지 않은 동일한 유형의 분화된 세포에 비해, 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, 세포.

[1494] 143. 실시형태 134 내지 142 중 어느 하나에 있어서, 줄기세포 확장제와 접촉되었던 세포.

[1495] 144. 실시형태 143에 있어서, 줄기세포 확장제가

[1496] a) (1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민;

[1497] b) 메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트;

[1498] c) 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀;

[1499] d) (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올; 또는

[1500] e) 이들의 조합(예를 들어 (1r,4r)-N^1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)

[1501] 인, 세포.

[1502] 145. 실시형태 144에 있어서, 줄기세포 확장제가 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올인, 세포.

[1503] 146. 세포, 예를 들어 실시형태 134 내지 145 중 어느 하나의 세포로서,

[1504] a) 실시형태 1 내지 68 중 어느 하나의 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하고/하거나;

[1505] b) HBG1 프로모터 영역에서 실시형태 1 내지 68 중 어느 하나의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 실시형태 1 내지 68 중 어느 하나의 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 포함하고, 선택적으로 결실이 5,250,092 내지 5,249,833, -가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는, 세포.

[1506] 147. 실시형태 146에 있어서, 삽입결실이 약 40개 미만 뉴클레오타이드, 예를 들어 30개 미만 뉴클레오타이드, 예를 들어 20개 미만 뉴클레오타이드, 예를 들어 10개 미만 뉴클레오타이드의 삽입 또는 결실인, 세포.

[1507] 148. 실시형태 146 또는 147에 있어서, 삽입결실이 단일 뉴클레오타이드 결실인, 세포.

[1508] 149. 실시형태 134 내지 148 중 어느 하나에 있어서, 세포가 동물 세포인, 세포.

[1509] 150. 실시형태 149에 있어서, 세포가 포유류, 영장류 또는 인간 세포인, 세포.

[1510] 151. 실시형태 134 내지 150 중 어느 하나에 있어서, 세포가 조혈 줄기 또는 전구 세포(HSPC)(예를 들어 HSPC의 집단)인, 세포.

[1511] 152. 실시형태 134 내지 151 중 어느 하나에 있어서, 세포가 CD34+ 세포인, 세포.

[1512] 153. 실시형태 152에 있어서, 세포가 CD34+CD90+ 세포인, 세포.

[1513] 154. 실시형태 134 내지 153 중 어느 하나에 있어서, 세포(예를 들어 세포 집단)가 골수, 동원된 말초 혈액, 또는 제대혈로부터 단리된 것인, 세포.

[1514] 155. 실시형태 134 내지 154 중 어느 하나에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 자가이며, 선택적으로 환자가 혈색소병증 환자이고, 선택적으로 환자가 겸상 적혈구 질병 또는 탈라쎄미아, 선택적으로 베타 탈라쎄미아를 갖고 있는, 세포.

[1515] 156. 실시형태 134 내지 154 중 어느 하나에 있어서, 세포는 상기 세포가 투여되는 환자에 관하여 동종이계인, 세포.

[1516] 157. 실시형태 134 내지 156 중 어느 하나의 세포를 포함하는, 세포 집단.

[1517] 158. 실시형태 157에 있어서, 집단의 세포 중 적어도 약 50%, 예를 들어 적어도 약 60%, 예를 들어 적어도 약 70%, 예를 들어 적어도 약 80%, 예를 들어 적어도 약 90%(예를 들어 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%)가 실시형태 134 내지 156 중 어느 하나에 따른 세포인, 세포 집단.

[1518] 159. 실시형태 157 또는 158에 있어서, 세포 집단이 분화된 세포의 집단, 예를 들어 적혈구계 계통의 세포 집단(예를 들어 적혈구 집단)으로 분화할 수 있고, 분화된 세포의 상기 집단이 예를 들어 동일한 유형의 비변형된 세포 집단에 비해, 증가된 퍼센트의 F 세포(예를 들어 적어도 약 15%, 적어도 약 17%, 적어도 약 20%, 적어도 약 25%, 적어도 약 30% 또는 적어도 약 40% 더 높은 퍼센트의 F 세포)를 갖는, 세포 집단.

[1519] 160. 실시형태 159에 있어서, 분화된 세포의 집단의 F 세포가 1개 세포 당 평균적으로 적어도 약 6 피코그램(예를 들어 적어도 약 7 피코그램, 적어도 약 8 피코그램, 적어도 약 9 피코그램, 적어도 약 10 피코그램, 또는 약 8 내지 약 9 피코그램, 또는 약 9 내지 약 10 피코그램)의 태아 헤모글로빈을 생산하는, 세포 집단.

[1520] 161. 실시형태 157 내지 160 중 어느 하나에 있어서,

[1521] 1) 세포가 투여되는 환자의 체중 1 kg 당 적어도 1e6개 CD34+ 세포;

[1522] 2) 세포가 투여되는 환자의 체중 1 kg 당 적어도 2e6개 CD34+ 세포;

[1523] 3) 세포가 투여되는 환자의 체중 1 kg 당 적어도 3e6개 CD34+ 세포;

[1524] 4) 세포가 투여되는 환자의 체중 1 kg 당 적어도 4e6개 CD34+ 세포; 또는

[1525] 5) 세포가 투여되는 환자의 체중 1 kg 당 2e6 내지 10e6개 CD34+ 세포

[1526] 를 포함하는, 세포 집단.

[1527] 162. 실시형태 157 내지 161 중 어느 하나에 있어서, 집단의 세포 중 적어도 약 40%, 예를 들어 적어도 약 50%(예를 들어 적어도 약 60%, 적어도 약 70%, 적어도 약 80% 또는 적어도 약 90%)가 CD34+ 세포인, 세포 집단.

[1528] 163. 실시형태 162에 있어서, 집단의 세포 중 적어도 약 10%, 예를 들어 적어도 약 15%, 예를 들어 적어도 약 20%, 예를 들어 적어도 약 30%가 CD34+CD90+ 세포인, 세포 집단.

[1529] 164. 실시형태 157 내지 163 중 어느 하나에 있어서, 세포 집단이 제대혈, 말초 혈액(예를 들어 동원된 말초 혈액), 또는 골수로부터 유래되거나, 예를 들어 골수로부터 유래되는, 세포 집단.

[1530] 165. 실시형태 157 내지 164 중 어느 하나에 있어서, 세포 집단이 포유류 세포, 예를 들어 인간 세포를 포함하며, 예를 들어 이로 구성되며, 선택적으로 세포 집단이 혈색소병증, 예를 들어 겸상 적혈구 질병 또는 탈라쎄미아, 예를 들어 베타-탈라쎄미아를 앓고 있는 환자로부터 수득되는, 세포 집단.

[1531] 166. 실시형태 157 내지 165 중 어느 하나에 있어서, 세포 집단이 (i) 이러한 세포 집단이 투여되는 환자에 비해 자가이거나, (ii) 이러한 세포 집단이 투여되는 환자에 비해 동종이계인, 세포 집단.

[1532] 167. 예를 들어 실시형태 157 내지 165 중 어느 하나에 있어서, 표 7-2에 기재된 바와 같은 삽입결실 패턴을 포

함하고, 선택적으로 표 7-2에 기재된 삽입결실 패턴의 삽입결실이 집단의 세포 중 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90% 또는 적어도 95%에서 검출 가능한 것인, 세포(예를 들어 CD34+ 세포) 집단.

[1533] 168. 실시형태 134 내지 156 중 어느 하나의 세포, 또는 실시형태 157 내지 167 중 어느 하나의 세포 집단을 포함하는 조성물.

[1534] 169. 실시형태 168에 있어서, 약제학적으로 허용 가능한 배지, 예를 들어 동결보존에 적합한 약제학적으로 허용 가능한 배지를 포함하는, 조성물.

[1535] 170. 혈색소병증을 치료하는 방법으로서, 실시형태 134 내지 156 중 어느 하나의 세포, 실시형태 157 내지 167 중 어느 하나의 세포 집단, 또는 실시형태 168 또는 169의 조성물을 환자에게 투여하는 단계를 포함하는, 방법.

[1536] 171. 포유류에서 태아 헤모글로빈 발현을 증가시키는 방법으로서, 실시형태 134 내지 156 중 어느 하나의 세포, 실시형태 157 내지 167 중 어느 하나의 세포 집단, 또는 실시형태 168 또는 169의 조성물을 환자에게 투여하는 단계를 포함하는, 방법.

[1537] 172. 실시형태 170에 있어서, 혈색소병증이 베타-탈라쎄미아 또는 겸상 적혈구 질병인, 방법.

[1538] 173. 세포(예를 들어 세포 집단)를 제조하는 방법으로서,

[1539] (a) 세포(예를 들어 세포 집단)(예를 들어 HSPC(예를 들어 HSPC의 집단))를 제공하는 단계;

[1540] (b) 상기 세포(예를 들어 상기 세포 집단)를 줄기세포 확장제를 포함하는 세포 배양 배지에서 생체외에서 배양하는 단계; 및

[1541] (c) 상기 세포 내로 실시형태 1 내지 86 중 어느 하나의 제1 gRNA 분자, 실시형태 1 내지 86 중 어느 하나의 제1 gRNA 분자를 인코딩하는 핵산 분자, 실시형태 87 내지 104 또는 168 내지 169 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 또는 실시형태 113 또는 114의 벡터를 도입하는 단계

[1542] 를 포함하는, 방법.

[1543] 174. 실시형태 173에 있어서, 단계 (c)의 상기 도입 후, 상기 세포(예를 들어 세포 집단)가 분화된 세포(예를 들어 분화된 세포의 집단), 예를 들어 적혈구계 계통의 세포(예를 들어 적혈구계 계통의 세포 집단), 예를 들어 적혈구(예를 들어 적혈구 집단)로 분화할 수 있고, 상기 분화된 세포(예를 들어 분화된 세포의 집단)가 예를 들어 단계 (c)를 거치지 않은 동일한 세포에 비해, 증가된 태아 헤모글로빈을 생산하는, 방법.

[1544] 175. 실시형태 173 또는 174에 있어서, 줄기세포 확장제가

[1545] a) (1r,4r)-N1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민;

[1546] b) 메틸 4-(3-피페리딘-1-일프로필아미노)-9H-피리미도[4,5-b]인돌-7-카르복실레이트;

[1547] c) 4-(2-(2-(벤조[b]티오펜-3-일)-9-이소프로필-9H-퓨린-6-일아미노)에틸)페놀;

[1548] d) (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올; 또는

[1549] e) 이들의 조합(예를 들어 (1r,4r)-N1-(2-벤질-7-(2-메틸-2H-테트라졸-5-일)-9H-피리미도[4,5-b]인돌-4-일)사이클로헥산-1,4-디아민과 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올의 조합)

[1550] 인, 방법.

[1551] 176. 실시형태 175에 있어서, 줄기세포 확장제가 (S)-2-(6-(2-(1H-인돌-3-일)에틸아미노)-2-(5-플루오로피리딘-3-일)-9H-퓨린-9-일)프로판-1-올인, 방법.

[1552] 177. 실시형태 173 내지 176 중 어느 하나에 있어서, 세포 배양 배지가 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)를 포함하는, 방법.

[1553] 178. 실시형태 177에 있어서, 세포 배양 배지가 인간 인터루킨-6(IL-6)을 추가로 포함하는, 방법.

[1554] 179. 실시형태 177 또는 178에 있어서, 세포 배양 배지가 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)를 각각 약 10 ng/mL 내지 약 1000 ng/mL 범위의 농도로 포함하는, 방법.

[1555] 180. 실시형태 179에 있어서, 세포 배양 배지가 트롬보포이에틴(Tpo), Flt3 리간드(Flt-3L) 및 인간 줄기세포 인자(SCF)를 각각 약 50 ng/mL의 농도, 예를 들어 50 ng/mL의 농도로 포함하는, 방법.

[1556] 181. 실시형태 178 내지 180 중 어느 하나에 있어서, 세포 배양 배지가 인간 인터루킨-6(IL-6)을 약 10 ng/mL 내지 약 1000 ng/mL 범위의 농도로 포함하는, 방법.

[1557] 182. 실시형태 181에 있어서, 세포 배양 배지가 인간 인터루킨-6(IL-6)을 약 50 ng/mL의 농도, 예를 들어 50 ng/mL의 농도로 포함하는, 방법.

[1558] 183. 실시형태 173 내지 182 중 어느 하나에 있어서, 세포 배양 배지가 줄기세포 확장제를 약 1 nM 내지 약 1 mM 범위의 농도로 포함하는, 방법.

[1559] 184. 실시형태 183에 있어서, 세포 배양 배지가 줄기세포 확장제를 약 1 uM 내지 약 100 nM 범위의 농도로 포함하는, 방법.

[1560] 185. 실시형태 184에 있어서, 세포 배양 배지가 줄기세포 확장제를 약 500 500 nM 내지 약 750 nM 범위의 농도로 포함하는, 방법.

[1561] 186. 실시형태 185에 있어서, 세포 배양 배지가 줄기세포 확장제를 약 500 nM의 농도, 예를 들어 500 nM의 농도로 포함하는, 방법.

[1562] 187. 실시형태 186에 있어서, 세포 배양 배지가 줄기세포 확장제를 약 750 nM의 농도, 예를 들어 750 nM의 농도로 포함하는, 방법.

[1563] 188. 실시형태 173 내지 187 중 어느 하나에 있어서, 단계 (b)의 배양이 단계 (c)의 도입 전에 배양 기간을 포함하는, 방법.

[1564] 189. 실시형태 188에 있어서, 단계 (c)의 도입 전의 배양 기간이 적어도 12시간, 예를 들어 약 1일 내지 약 12일의 기간 동안이며, 예를 들어 약 1일 내지 약 6일의 기간 동안이며, 예를 들어 약 1일 내지 약 3일의 기간 동안이며, 예를 들어 약 1일 내지 약 2일의 기간이거나, 예를 들어 약 2일의 기간 동안 또는 약 1일의 기간 동안인, 방법.

[1565] 190. 실시형태 173 내지 189 중 어느 하나에 있어서, 단계 (b)의 배양이 단계 (c)의 도입 후에 배양 기간을 포함하는, 방법.

[1566] 191. 실시형태 190에 있어서, 단계 (c)의 도입 후의 배양 기간이 적어도 12시간, 예를 들어 약 1일 내지 약 12일의 기간 동안이며, 예를 들어 약 1일 내지 약 6일의 기간 동안이며, 예를 들어 약 2일 내지 약 4일의 기간 동안이며, 예를 들어 약 2일의 기간 동안이거나 약 3일의 기간 동안이거나, 약 4일의 기간 동안인, 방법.

[1567] 192. 실시형태 173 내지 191 중 어느 하나에 있어서, 세포 집단이 예를 들어 단계 (b)에 따라 배양되지 않은 세포에 비해, 적어도 4배, 예를 들어 적어도 5배, 예를 들어 적어도 10배 확장되는, 방법.

[1568] 193. 실시형태 173 내지 192 중 어느 하나에 있어서, 단계 (c)의 도입이 전기천공을 포함하는, 방법.

[1569] 194. 실시형태 193에 있어서, 전기천공이 1 내지 5 펄스, 예를 들어 1 펄스를 포함하고, 각각의 펄스가 700 볼트 내지 2000 볼트 범위의 펄스 전압에 존재하고, 10 ms 내지 100 ms 범위의 펄스 기간을 갖는, 방법.

[1570] 195. 실시형태 194에 있어서, 전기천공이 1 펄스를 포함하는, 방법.

[1571] 196. 실시형태 194 또는 195에 있어서, 펄스 전압이 1500 내지 1900 볼트 범위이고, 예를 들어 1700 볼트인, 방법.

[1572] 197. 실시형태 194 내지 196 중 어느 하나에 있어서, 펄스 기간이 10 ms 내지 40 ms 범위이고, 예를 들어 20 ms인, 방법.

[1573] 198. 실시형태 173 내지 197 중 어느 하나에 있어서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 인간 세포(예를 들어 인간 세포의 집단)인, 방법.

[1574] 199. 실시형태 198에 있어서, 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 골수, 말초 혈액(예를 들어 동원된 말초 혈액) 또는 제대혈로부터 단리되는, 방법.

[1575] 200. 실시형태 199에 있어서,

[1576] (i) 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 골수로부터 단리되며, 예를 들어 혈색소병증을 앓고 있는 환자의 골수로부터 단리되며, 선택적으로 혈색소병증이 겸상 적혈구 질병 또는 탈라쎄미아이고, 선택적으로 탈라쎄미아가 베타 탈라쎄미아이거나;

[1577] (ii) 단계 (a)에서 제공되는 세포(예를 들어 세포 집단)가 말초 혈액으로부터 단리되며, 예를 들어 혈색소병증을 앓고 있는 환자의 말초 혈액으로부터 단리되며, 선택적으로 혈색소병증이 겸상 적혈구 질병 또는 탈라쎄미아이며, 선택적으로 탈라쎄미아가 베타 탈라쎄미아이고; 선택적으로 말초 혈액이 동원된 말초 혈액이며, 선택적으로 동원된 말초 혈액이 Plerixafor, G-CSF 또는 이들의 조합을 사용하여 동원된 것인, 방법.

[1578] 201. 실시형태 173 내지 200 중 어느 하나에 있어서, 단계 (a)에서 제공되는 세포 집단이 CD34+ 세포에 대해 농화되는, 방법.

[1579] 202. 실시형태 173 내지 201 중 어느 하나에 있어서, 단계 (c)의 도입에 후속하여, 세포(예를 들어 세포 집단)가 동결보존되는, 방법.

[1580] 203. 실시형태 173 내지 202 중 어느 하나에 있어서, 단계 (c)의 도입에 후속하여, 세포(예를 들어 세포 집단)가

[1581] a) 제1 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하고/하거나;

[1582] b) HBG1 프로모터 영역에서 제1 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열과 HBG2 프로모터 영역에서 제1 gRNA 분자의 표적화 도메인에 상보적인(예를 들어 gRNA 표적화 도메인에 적어도 90% 상보적인, 예를 들어 gRNA 표적화 도메인에 완전히 상보적인) 서열 사이의 서열, 예를 들어 실질적으로 모든 서열을 포함하는 결실을 포함하고, 선택적으로 삽입결실, 예를 들어 결실이 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드를 포함하지 않는, 방법.

[1583] 204. 실시형태 173 내지 203 중 어느 하나에 있어서, 단계 (c)의 도입 후, 세포 집단의 세포 중 적어도 약 40%, 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98% 또는 적어도 약 99%가 제1 gRNA 분자의 표적화 도메인에 상보적인 게놈 DNA 서열에서 또는 그 부근에서 삽입결실을 포함하고, 선택적으로 집단의 세포 중 어느 것도 5,250,092 내지 5,249,833, - 가닥(hg38) 사이에 배치된 뉴클레오타이드의 결실을 포함하지 않는, 방법.

[1584] 205. 실시형태 173 내지 204 중 어느 하나의 방법에 의해 수득 가능한 세포(예를 들어 세포 집단).

[1585] 206. 혈색소병증을 치료하는 방법으로서, 실시형태 134 내지 156 중 어느 하나의 세포, 실시형태 157 내지 167 중 어느 하나의 세포 집단, 또는 실시형태 205의 세포(예를 들어 세포 집단)를 포함하는 조성물을 인간 환자에게 투여하는 단계를 포함하는, 방법.

[1586] 207. 인간 환자에서 태아 헤모글로빈 발현을 증가시키는 방법으로서, 실시형태 134 내지 156 중 어느 하나의 세포, 실시형태 157 내지 167 중 어느 하나의 세포 집단, 또는 실시형태 205의 세포(예를 들어 세포 집단)를 포함하는 조성물을 상기 인간 환자에게 투여하는 단계를 포함하는, 방법.

[1587] 208. 실시형태 206에 있어서, 혈색소병증이 베타-탈라쎄미아 또는 겸상 적혈구 질병인, 방법.

[1588] 209. 실시형태 206 내지 208 중 어느 하나에 있어서, 인간 환자에게 인간 환자 체중 1 kg 당 실시형태 205의 적어도 약 1e6개 세포, 예를 들어 인간 환자 체중 1 kg 당 실시형태 205의 적어도 약 1e6개 CD34+ 세포를 포함하는 조성물이 투여되는, 방법.

[1589] 210. 실시형태 209에 있어서, 인간 환자에게 인간 환자 체중 1 kg 당 실시형태 205의 적어도 약 2e6개 세포, 예를 들어 인간 환자 체중 1 kg 당 실시형태 205의 적어도 약 2e6개 CD34+ 세포를 포함하는 조성물이 투여되는, 방법.

[1590] 211. 실시형태 209에 있어서, 인간 환자에게 인간 환자 체중 1 kg 당 실시형태 205의 약 2e6 내지 약 10e6개 세포, 예를 들어 인간 환자 체중 1 kg 당 실시형태 205의 적어도 약 2e6 내지 약 10e6개 CD34+ 세포를 포함하는 조성물이 투여되는, 방법.

[1591] 212. 약제로서 사용하기 위한, 실시형태 1 내지 86 중 어느 하나의 gRNA, 실시형태 87 내지 114 또는 168 내지

169 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 실시형태 113 또는 114의 벡터, 실시형태 134 내지 156 중 어느 하나 또는 205의 세포, 또는 실시형태 157 내지 167 중 어느 하나의 세포 집단.

[1592] 213. 약제의 제조에 사용하기 위한, 실시형태 1 내지 86 중 어느 하나의 gRNA, 실시형태 87 내지 114 또는 168 내지 169 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 실시형태 113 또는 114의 벡터, 실시형태 134 내지 156 중 어느 하나 또는 205의 세포, 또는 실시형태 157 내지 167 중 어느 하나의 세포 집단.

[1593] 214. 질병의 치료에 사용하기 위한, 실시형태 1 내지 86 중 어느 하나의 gRNA, 실시형태 87 내지 114 또는 168 내지 169 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 실시형태 113 또는 114의 벡터, 실시형태 134 내지 156 중 어느 하나 또는 205의 세포, 또는 실시형태 157 내지 167 중 어느 하나의 세포 집단.

[1594] 215. 질병의 치료에 사용하기 위한, 실시형태 1 내지 86 중 어느 하나의 gRNA, 실시형태 87 내지 114 또는 168 내지 169 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 실시형태 113 또는 114의 벡터, 실시형태 134 내지 156 또는 205 중 어느 하나의 세포, 또는 실시형태 157 내지 167 중 어느 하나의 세포 집단으로서, 여기서, 질병은 혈색소병증인, gRNA 분자, 조성물, 핵산, 벡터, 세포, 또는 세포 집단.

[1595] 216. 질병의 치료에 사용하기 위한, 실시형태 1 내지 86 중 어느 하나의 gRNA, 실시형태 87 내지 114 또는 168 내지 169 중 어느 하나의 조성물, 실시형태 105 내지 112 중 어느 하나의 핵산, 실시형태 113 또는 114의 벡터, 실시형태 134 내지 156 또는 205 중 어느 하나의 세포, 또는 실시형태 157 내지 167 중 어느 하나의 세포 집단으로서, 여기서, 혈색소병증은 베타-탈라쎄미아 또는 겸상 적혈구 질병인, gRNA 분자, 조성물, 핵산, 벡터, 세포, 또는 세포 집단.

[1596] **실시예**

[1597] 실시예 1 - 예시적인 일반적 방법

[1598] **가이드 선택 및 설계**

[1599] 관심 영역에서의 PAM을 식별하기 위해, 초기 가이드 선택을, 인간 참조 게놈 및 사용자 정의된 관심 게놈 영역(예를 들어 유전자, 유전자의 엑손, 비-코딩 조절 영역 등)을 사용하여 인실리코(*in silico*)로 수행하였다. 식별된 각각의 PAM에 대해, 분석을 수행하고 통계를 보고하였다. gRNA 분자를 추가로 선택하고, 예를 들어 본원에 기재된 바와 같은 효율 및 효능을 결정하기 위한 다수의 방법을 기초로 순위를 매겼다. 이러한 실시예는 본원에 기재된 본 발명의 CRISPR 시스템, gRNA 및 다른 양태를 검정하는 데 사용될 수 있는 절차에 대한 실험 세부사항을 제공한다. 특정 실험에 이용된 이들 일반적인 절차에 대한 임의의 변형은 해당 실시예에서 주지된다.

[1600] 실시예 전반에 걸쳐, 하기 실험에서, sgRNA 분자 또는 dgRNA 분자 중 어느 하나를 사용하였다. 다르게 지시되지 않는 한, dgRNA 분자를 사용한 경우, gRNA는 하기를 포함한다:

[1601] crRNA: [표적화 도메인]-[SEQ ID NO: 201]

[1602] tracr (trRNA): SEQ ID NO: 224

[1603] 다르게 지시되지 않는 한, sgRNA 분자를 이용하는 실험에서, 하기 서열을 사용하였다:

[1604] [표적화 도메인]-[SEQ ID NO: 195]-UUUU

[1605] **차세대 시퀀싱(NGS) 및 표적-적중 절단 효율 및 삽입결실 형성에 대한 분석**

[1606] 게놈에서 표적 위치를 편집(예를 들어 절단)하는 효율을 결정하기 위해, 딥 시퀀싱(deep sequencing)을 이용하여 비-상동 말단 접합에 의해 도입된 삽입 및 결실의 존재를 확인하였다.

[1607] 요약하자면, PCR 프라이머를 표적 부위 주변으로 설계하였으며, 관심 게놈 영역을 편집된 시료 및 비편집된 시료에서 PCR 증폭시켰다. 생성된 앰플리콘을 Illumina 시퀀싱 라이브러리로 전환시키고 시퀀싱하였다. 시퀀싱 리드를 인간 게놈 참조에 정렬시키고, 변이체 콜링 분석(variant calling analysis)을 받게 하였으며, 이로써 본 발명자들은 관심 표적 영역에서 서열 변이체 및 이들의 빈도를 결정할 수 있었다. 데이터를 다양한 품질 필터를 거치게 하였고, 공지된 변이체 또는 비편집된 시료에서만 식별되는 변이체는 배제하였다. 편집 퍼센트는, 관심 표적-적중 부위에서 발생하는 모든 삽입 또는 결실 사건(즉, 표적-적중 부위에서 리드(야생형 및 돌연변이체 리드)의 총 수에 걸쳐 표적-적중 부위에서의 삽입 및 결실 리드)의 퍼센트로서 정의하였다. NGS 분석 과정에 대한 상세한 사항은 실시예 2.1에 기재되어 있다.

[1608] **RNP 발생**

[1609] Cas9 단백질에의 crRNA 및 trRNA의 첨가는 활성 Cas9 리보핵 단백질 복합체(RNP)의 형성을 야기하며, 이는 crRNA에 의해 특정된 표적 영역에의 결합 및 표적화된 게놈 DNA의 특이적 절단을 매개한다. trRNA 및 crRNA를 Cas9 내로 로딩함으로써 이 복합체를 형성하였으며, 이는 Cas9에 대한 입체 형태 변화를 야기하여 dsDNA에 결합 및 절단 가능하게 하는 것으로 여겨진다.

[1610] crRNA와 trRNA를 95℃에서 2분 동안 개별적으로 변성시키고, 실온이 되게 하였다. Cas9 단백질(10 mg/ml)을 5X CCE 완충제(20 mM HEPES, 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 5% 글리세롤)에 첨가한 다음에, trRNA 및 다양한 crRNA를 첨가하고(별도의 반응으로) 37℃에서 10분 동안 인큐베이션시켜, 활성 RNP 복합체를 형성하였다. 전기천공 및 다른 방법에 의해, HEK-293 및 CD34+ 조혈 세포를 포함하는, 폭넓게 다양한 세포 내로 복합체를 전달하였다.

[1611] **RNP의 CD34+ HSC로의 전달**

[1612] Cas9 RNP를 CD34+ HSC 내로 전달하였다.

[1613] CD34+ HSC를 해동하고 IL12, SCF, TPO, Flt3L 및 페니실린/스트렙토마이신이 첨가된 StemSpan SFEM(StemCell Technologies) 배지에서 밤새 (약 500,000 세포/ml로) 배양하였다. 각 RNP 전달 반응 당 대략 90,000개의 세포를 분취하고 펠렛화 하였다. 그 다음에, 세포를 60 ul P3 핵감염(nucleofection) 완충제(Lonza)에 재현탁하고 활성 RNP를 이어서 첨가하였다. 그 다음에, HSC를 3회 반복실험으로(20 uL/전기천공) 전기천공(예를 들어 Lonza 뉴클레오펙터(Nucleofector)상의 프로그램 CA-137을 사용하여 핵감염)하였다. 전기천공 직후에, StemSpan SFEM 배지(IL12, SCF, TPO, Flt3L 및 페니실린/스트렙토마이신 포함)를 HSC에 첨가하였고, 적어도 24시간 동안 배양하였다. 그 다음에, HSC를 수확하고 T7E1, NGS 및/또는 표면 마커 발현 분석을 하였다.

[1614] **HSC 기능적 검정**

[1615] CD34+ HSC는 유세포 분석 또는 시험관내 콜로니 형성 검정과 같은 공지된 기술을 이용하여 줄기 세포 표현형에 대해 검정될 수 있다. 예로서, 제조사의 프로토콜을 이용하는 Methocult H4034 Optimum 키트(StemCell Technologies)를 사용하여 시험관내 콜로니 형성 검정(CFC)에 의해 세포를 검정하였다. 간략하게, 100 ul 이하의 용량의 500 내지 2000 CD34+ 세포를 1 내지 1.25 ml methocult에 첨가하였다. 혼합물을 4 내지 5초 동안 격렬하게 볼텍싱(vortex)하여 완전히 혼합한 다음, 실온에서 적어도 5분 동안 그대로 있도록 하였다. 시린지를 사용하여, 1 내지 1.25 ml의 MethoCult+ 세포를 35 mm 디쉬 또는 6-웰 플레이트의 웰로 이전하였다. 콜로니 수 및 형태학은 제조사의 프로토콜에 따라 12 내지 14일 후에 평가하였다.

[1616] **생체내 이종간 이식**

[1617] HSC는 자가-갱신 및 다-계통 분화 능력에 의해 기능적으로 정의된다. 이 기능성은 생체내에서만 평가될 수 있다. 인간 HSC 기능을 결정하기 위한 최적 표준은, 일련의 돌연변이를 통해 심하게 면역 약화되어 인간 세포에 대한 수용자로서 작용할 수 있는 NOD-SCID 감마 마우스(NSG)로의 이종간 이식을 통하는 것이다. 유도된 편집이 HSC 기능에 영향을 미치지 않는지 확인하기 위해 편집 후 HSC를 NSG 마우스로 이식할 것이다. 인간 키메라 현상과 계통 발생을 평가하기 위하여 매월 말초 혈액 분석법을 사용할 것이고, 기능성 HSC의 존재를 확립하기 위하여 20주 후의 2차 이식을 사용할 것이다.

[1618] **실시예 2.1** 성인 적혈구계 세포에서 태아 글로빈 빈혈의 탈억제(de-repression)를 위해 CRISPR-Cas9를 사용한 조혈 줄기 및 전구 세포(HSPC)에서 비결실 HPFH 영역 편집.

[1619] 방법:

[1620] 인간 $CD34^+$ 세포 배양물. 인간 CD34+ 세포를 면역선택(immunoselection)(Miltenyi)을 제조업체의 설명에 따라 사용하여 성인 공여자(All세포)의 G-CSF 동원된 말초 혈액으로부터 단리하고, 50 ng/mL의 각각의 트롬보포이에틴(Tpo, Life Technologies, Cat. #PHC9514), Flt3 리간드(Flt-3L, Life Technologies, Cat. #PHC9413), 인간 줄기세포 인자(SCF, Life Technologies, Cat. #PHC2113) 및 인간 인터루킨-6(IL-6, Life Technologies, Cat. #PHC0063), 뿐만 아니라 1x 항생제/항진균제(Gibco, Cat. #10378-016) 및 500 nM 화합물 4가 보충된 StemSpan SFEM(Stemcell Technologies; Cat no. 09650)을 사용하여 4 내지 6일 동안 확장시켰다. 이러한 실시예(이의 하위실시예를 포함함) 전체에서, 세포가 "확장되었음"을 표시하는 프로토콜에서, 이러한 배지를 사용하였다. 이러한 배지는 또한, 이러한 실시예(이의 하위실시예를 포함함) 전체에서, "줄기세포 확장 배지" 또는 "확장 배지"

로 지칭된다.

[1621] Cas9와 가이드 RNA 리보핵단백질(RNP) 복합체의 조립, HSPC의 제조, 및 HSPC 내로의 RNP의 전기천공. 전기천공 직전에, Cas9-가이드 RNA 리보핵단백질 복합체(RNP)를 제조하였다. 이중 가이드 RNA(dgRNA)를 사용한 RNP의 형성을 위해, 6 μg의 각각의 crRNA(4.5 μL 중) 및 tracr(2.52 μL 중)를 우선, 개별 튜브에서 95℃에서 2분 동안 변성시키고, 그 후에, 실온까지 냉각시켰다. Cas9 단백질의 제조를 위해, 12 μg의 CAS9 단백질(2 μL 중)을 1 μL의 10 x CCE 완충제(20 mM HEPES, 100 mM KCL, 5 mM MgCL2, 5% 글리세롤 및 신선하게 첨가된 1 mM DTT)와 혼합하였다. Tracr를 우선, Cas9 조제물과 혼합하고, 37℃에서 5분 동안 인큐베이션하였다. 그 후에, crRNA를 Tracr/CAS9 복합체에 첨가하고, 37℃에서 5분 동안 인큐베이션하였다. 없음/대조군 조건의 경우, crRNA 이외의 비히클을 tracr/CAS9 복합체에 첨가하였다. HSPC를 원심분리에 의해 수합하고, 2.8 x 10^6/mL의 세포 밀도에서 Lonza 전기천공 키트(Cat# V4XP-3032 Lonza Amaxa P3 1차 세포 4-D 핵감염 X Kit S)에 포함된 P3 완충제 +보충제에 재현탁시켰다. RNP를 40 μL의 세포와 함께 위아래로 파이펫팅함으로써 혼합하고, RT에서 2분 동안 인큐베이션하였다. 각각의 복제물에 대해, 21 μL의 the RNP/세포 혼합물을 Lonza Amaxa P3 1차 세포 4-D 핵감염 X 키트 S 내로 옮겼다. 전기천공을 Lonza 형질감염 시스템(4D-Nucleofector X Unit)을 이용하여 프로토콜 CM-137을 사용하여 수행하였다. 2벌 중복 21 μL 전기천공을 수행하였다.

[1622] *시험관내* HbF-함유 적혈구계 세포에 대한 적혈구생성(erythropoiesis) 및 FACS 분석. 전기천공 후, 세포를 IMDM(GE Life Sciences, Cat. #SH30228.01), 330 μg /mL 인간 홀로-트랜스페린(holo-트랜스페린)(Invitria Cat# 777TRF029), 10 μg /mL 재조합 인간 인슐린(Gibco Cat#A1138211), 2 IU/mL 헤파린(Sigma, 파트 #H3393), 5% 인간 AB 혈청(Sigma, Cat# H4522), 2.5 U/mL 인간 에리트로포이에틴(Peprotech #100-064) 및 1x 항생제/항진균제(Gibco, Cat. #10378-016)로 구성된 250 μL 예열된 적혈구계 분화 배지(EDM) 내로 즉시 옮겼다. 7일 이하의 초기 배양 기간 동안, EDM에 1.38 μM 하이드로코티손(Sigma H8672), 100 ng/mL 인간 SCF(Life Technologies, Cat. #PHC2113) 및 5 ng/mL 인간 IL-3(Peprotech #10779-598)을 추가로 보충하여, EDM-I를 제조하였다. 4일 후, 세포 배양물을 신선한 배지에서 희석시켰다. 배양물을 상기 기재된 배양 조건에서 총 7일 동안 유지시켰으며, 이때, 1/2의 세포를 HbF 발현에 대해 세포내 염색에 의해 분석하였다. 간략하게는, 세포를 PBS로 1회 세척하고, LIVE/DEAD® Fixable Violet Dead Cell Stain(ThermoFisher L34963; PBS 중 1:1000)에 재현탁시키고, 30분 동안 인큐베이션하였다. 그 후에, 세포를 세척하고, 항-CD71-BV711(Fisher Scientific Company Llc. BD 563767)과 항-CD235a-APC(BD 551336) 항체의 1/50 희석물로 30분 동안 염색하였다. 그 후에, 세포를 세척하고, 뒤이어 고정 완충제(Biolegend, Cat# 420801)로 고정한 다음, 1x 세포내 염색 침투 세척 완충제(Biolegend, Cat# 421002)를 제조업체의 설명에 따라 이용하여 침투시켰다. 그 후에, 세포를 50 μL의 1x 세포내 염색 침투 세척 완충제 중 항-HbF-PE 항체(Life Technologies, 파트 #MHFH04)의 1/40 희석물과 함께 실온에서 20분 동안 인큐베이션하였다. 세포를 0.2 mL의 1x 세포내 염색 침투 세척 완충제로 2회 세척하고, 염색 완충제에 재현탁시키고, LSRFortessa 유세포분석기(BD Biosciences) 상에서 HbF 발현에 대해 분석하였다. 그 결과를 Flowjo를 사용하여 분석하고, 데이터를 살아 있는 CD71 양성 적혈구계 세포 집단에서의 HbF 양성 세포(F-세포)의 %로 제시하였다.

[1623] 나머지로부터 1개 조건 당 7일차 배양된 세포 80,000개를 11일차까지 추가의 분화를 위해 EDM-II 배양 배지로 옮겼다. EDM-II는 100 ng/mL SCF만 보충된 EDM-I로 구성된다. 11일차에 세포를 계수하고, 1개 조건 당 200,000 세포를 추가의 보충제 없이 EDM으로 옮겼다. 14일차에 세포를 7일차와 유사하게 HbF 발현의 분석을 위해 염색시키지만, 세포의 응집을 방지하기 위해 표면 마커를 염색으로부터 배제하였다.

[1624] 게놈 DNA 제조 및 차세대 시퀀싱(NGS). 전기천공-후 7일째에 Quick Extract DNA 추출 용액(Epicentre Cat # QE09050)을 사용하여 편집된 HSPC 및 비편집된 HSPC로부터 게놈 DNA를 제조하였다. 삽입 및 결실(삽입결실)의 편집 효율 및 패턴을 결정하기 위해, 표적 부위의 측면에 존재하는 프라이머를 사용하여 PCR 생성물을 발생시켰으며, 그 후에 이러한 PCR 생성물을 문헌에 기재된 바와 같이 차세대 시퀀싱(NGS)을 받게 하였다. 비편집된 시료(Cas9 및 Tracr로만 구성된 RNP를 이용하여 전기천공됨) 내 상응하는 서열의 편집 퍼센트는 전형적으로 1% 미만이었고, 결코 3%를 초과하지 않았다.

[1625] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱. 1.8x Agencourt AmpureXP 비드(Beckman Coulter)를 제조업체의 권고사항에 따라 사용하여 PCR 앰플리콘을 정제하였다. Quant-iT PicoGreen dsDNA 검정법(Life Technologies)을 제조업체의 권고사항에 따라 사용하여 앰플리콘을 정량화하였다. Nextera DNA 라이브러리 Prep Kit(Illumina)를 하기 변화를 주면서 제조업체의 권고사항에 따라 사용하여 Illumina 시퀀싱 라이브러리를 발생시켰다. 5 ng의 정제된 PCR 생성물, 0.15 ul의 Nextera 태그화 효소(tagment enzyme) 및 Wang 등(PMID: 24071908; 참조로 본

명세서에 포함됨)에 의해 이전에 기재된 태그화 완충제를 사용하여 5 ul의 최종 부피에서 태그화(tagmentation)를 수행하였다. 그 후에, 태그화된(tagmented) 앰플리콘을, 최종 농도 0.2 mM dNTP(Life Technologies), 0.2 uM Illumina 인덱스 PCR 프라이머(Integrated DNA Technologies), 1x Phusion DNA 중합효소 완충제(New England Biolabs) 및 1 U의 Phusion DNA 중합효소(New England Biolabs)를 사용하여 50 ul의 최종 부피에서 PCR 증폭시켰다. 사용된 PCR 사이클링 조건(cycling condition)은 하기와 같았다: 72℃에서 3분, 98℃에서 2분, 및 98℃에서 10초, 63℃에서 30초 및 72℃에서 3분 동안의 15회 사이클. 그 후에, 1.0x Agencourt AmpureXP 비드(Beckman Coulter)를 제조업체의 권고사항에 따라 사용하여 시퀀싱 라이브러리를 정제하였다. 시퀀싱 라이브러리를, Quant-iT PicoGreen dsDNA 검정법(Life Technologies)을 제조업체의 권고사항에 따라 사용하여 정량화하고, 시퀀싱에 대해 등몰량으로 풀링하였다. MiSeq sequencer(Illumina) 상에서 제조업체의 권고사항에 따라 150개 염기쌍-말단 리드를 이용하여 시퀀싱 라이브러리를 시퀀싱하였다. 1개 앰플리콘 당 최소 1000배 시퀀싱 커버리지(coverage)를 발생시켰다.

[1626] NGS 시퀀싱 데이터 QC 및 변이체 분석. 디폴트 매개변수를 사용하여, Illumina MiSeq 분석 소프트웨어(MiSeq 리포터, 버전 2.6.2, Illumina)를 사용하여, 앰플리콘 특이적 FASTQ 시퀀싱 데이터 파일을 발생시켰다(Cock et al, Nucleic Acids Res. 2010, 38(6):1767-71, PMID: 20015970). 그 후에, FASTQ 파일을, 표준 Perl 스크립트 랩퍼(script wrapper)를 사용하여 함께 접합된 일련의 공개 도메인 소프트웨어 패키지로 구성된 내부적으로 개발된 변이체 분석 파이프라인을 통해 처리하였다. 사용된 작업 흐름을 5개 단계로 나누었다.

[1627] 단계 1, PCR 프라이머 및 표적-적중 및 표적-외 서열 QC: 표적-적중 및 표적-외 부위 둘 다를 위해, 20 뉴클레오타이드 gRNA 표적화 도메인 서열 + PAM 서열 및 표적 특이적 PCR 프라이머 서열(부가적인 Illumina 서열이 없는 좌측 및 우측)을 BLAST 서치(버전 2.2.29+, 문헌[Altschul et al, J Mol Biol., 1990, 215(3):403-10], PMID: 2231712)를 사용하여 인간 게놈 참조 서열(GRCh38을 구축함)에 정렬시켰다. 다수의 게놈 위치를 갖는 표적-적중 및 표적-외 부위를 플래그(flag)하였다.

[1628] 단계 2, 시퀀서 파일 압축 파일 해제: Illumina 시퀀서 발생된 FASTQ.GZ 파일을, gzip 스크립트(버전 1.3.12)를 사용하여 FASTQ 파일로 압축 파일 해제시켰고, 1개 파일 당 리드의 수를 계산하였다. 리드가 없는 파일을 추가의 분석으로부터 배제하였다.

[1629] 단계 3, 서열 리드 정렬 및 품질 트리밍(quality trimming): Illumina 서열의 리드의 3' 말단 및 낮은 품질의 염기를 트리밍하기 위해 '하드-클리핑(hard-clipping)'을 사용하는 BWA-MEM 정렬기(버전 0.7.4-r385, 문헌[Li and Durbin, BIOINFORMATICS, 2009, 25(14):1754-60], PMID: 19451168)를 사용하여, FASTQ 파일에서 시퀀싱 리드를 인간 게놈 참조 서열(GRCh38을 구축함)에 정렬시켰다. BAM 파일 포맷에서의 생성된 정렬 리드(문헌[Li et al, BIOINFORMATICS, 2009 25(16):2078-9], PMID: 19505943)를 SAMtools 스크립트(버전 0.1.19-44428cd, 문헌[Li et al, BIOINFORMATICS, 2009 25(16):2078-9], PMID: 19505943)를 사용하여 FASTQ 파일로 전환시켰다. 그 후에, FASTQ 파일을 BWA-MEM 정렬기를 사용하여 인간 게놈 참조 서열(GRCh38을 구축함)에 다시 정렬시켰으며, 이때 '하드-클리핑'은 없었다.

[1630] 단계 4, 변이체(SNP 및 INDEL) 분석: 변이체(SNP 및 삽입결실)를 식별하기 위해 0.0001 이상에서 설정된 대립유전자 빈도 검출 한계와 함께 VarDict 변이체 콜러(caller)(ZhongWu Lai에 의해 변형된 버전 1.0 'Cas9 aware', 문헌[Lai et al, Nucleic Acids Res., 2016, 44(11):e108], PMID: 27060149)를 사용하여, 정렬된 리드의 BAM 파일을 가공하였다. Cas9 어웨어(aware) VarDict 콜러는 공개 도메인 패키지를 기초로 하지만, PAM 서열의 5'에 3개 염기에 위치한 gRNA 표적화 도메인 서열에서 잠재적인 Cas9 뉴클레아제 절단 부위로, 변이체 사건의 정렬 영역에서 반복적인 서열로 인해 발생된 모호한 변이체 콜(call)로 이동할 수 있다. SAMtools 스트립트를 사용하여, 1개 시료 앰플리콘 당 리드 커버리지를 계산하여, 표적-적중 및 표적-외 부위가 1000배 초과의 서열 커버리지에서 커버되는지 결정하였다. 1000배 미만의 서열 커버리지를 갖는 부위를 플래그하였다.

[1631] 단계 5, dbSNP 여과 및 처리/비처리된 차별적 분석: 식별된 변이체를, dbSNP에서 확인된 공지된 변이체(SNP 및 삽입결실)에 대해 여과하였다(빌드 142, 문헌[Shery et al, Nucleic Acids Res. 2001, 29(1):308-11], PMID: 11125122). 처리된 시료 내의 변이체를 추가로 여과하여, 1) 비편집된 대조군 시료에서 식별된 변이체; 2) 2:1의 VarDict 가닥 바이어스(bias)를 갖는 변이체(참조 서열을 뒷받침하는 포워드 및 리버스 리드 수치가 균형을 이루지만 비-참조 변이체 콜에 대해서는 불균형을 이루는 경우); 3) 잠재적인 Cas9 절단 부위의 어느 한 쪽 면에서 5 bp 초과에 위치한 변이체; 4) 단일 뉴클레오타이드 변이체를 배제하였다.

[1632] 결과:

[1633] 본 발명자들은 본원에서, HBG1 또는 HBG2 프로모터 영역 내의 특이적인 서열의 표적화된 교란(예를 들어 해당 서열에서 또는 그 부근에서의 삽입결실 생성에 의한)이 γ-글로빈 발현의 억제를 경감시켜, 상승된 HbF 단백질을 함유하는 적혈구의 생성을 허용한다(태아 헤모글로빈을 발현하는 세포는 이따금 본원에서 "F-세포"로 지칭됨)는 놀라운 결과를 보여주었다. 상승된 HbF는 탈산소화된 조건 하에 적혈구의 겸상세포 생성을 방지하고, β-탈라쎄미아 및 SCD 둘 모두의 환자에 대한 치료제/치유제가 될 것이다. 본원에서, SCD 환자로부터의 생체외 게놈 편집된 HSC를 이용한 자가 조혈 줄기세포 이식(HSCT)은 또한, 줄기세포 확장 증강 기술, 예를 들어 WO2010/059401(이의 내용이 그 전체가 참조로 포함됨)에 기재된 바와 같은 예를 들어 아릴 탄화수소 수용체(AHR) 저해제, 예를 들어 화합물 4와 조합되어, 전달된 유전자 변형된 HSC의 생체외 확장을 증가시키고 이의 용량을 증가시켰다.

[1634] 프로그램 가능한 뉴클레아제, Cas9를 통한 효율적인 게놈 편집을 위해서는, 표적 세포 및 조직 내로의 가이드 RNA(gRNA) 및 Cas9 단백질의 성공적인 전달이 필수적이다. 본원에서, 본 발명자들은, 전기천공에 의한 예비복합체화된 gRNA/Cas9 리보핵단백질(RNP) 복합체의 전달이 전달 후 거의 즉시 효율적이고 특이적인 게놈 편집을 초래하고 세포에서 분해되어, 표적-외 효과를 감소시킴을 보여준다. 대조적으로, Cas9를 전달하는 데 사용된 플라스미드 및 바이러스 벡터 시스템의 사용은 효소의 연장된 발현을 초래하고, 이는 시스템과 연관된 표적-외 효과를 악화시킬 수 있다. 부가적으로, 표적 세포 내로의 RNP의 전달은, 임상에서 치료 목적을 위해 게놈 편집의 번역을 크게 용이하게 할 어떠한 부가적인 툴도 필요로 하지 않는다.

[1635] 재조합 S. 피오게네스 Cas9 단백질(SEQ ID NO: 236)을 에스케리키아 콜라이로부터 정제하고, 리보핵단백질(RNP) 복합체를 발생시키기 위해 crRNA 및 tracr로 구성된 합성 이중 gRNA(dgRNA)와 복합체화하였다. 연구에 사용된 gRNA 표적화 도메인의 목록 및 서열을 표 1에서 보여준다. 대부분의 gRNA 서열은 HBG1 및 HBG2 프로모터 영역 둘 모두에서 완벽한 표적 또는 단지 1 또는 2개 미스매치만 가진다. 이러한 상황은 인간 베타 글로빈 유전자좌 내에서 HBG 유전자의 중복(duplication)으로 인한 것이다. RNP 복합체를 물질 및 방법 하에 기재된 바와 같이 전기천공을 통해 CD34+ HSPC 내로 전기천공시켰다. RNP 복합체의 전달 전에 세포를 확장시켰다. 이론으로 결부시키고자 하는 것은 아니지만, 활동적으로 분열하는 세포는 전기천공에 의해 전달된 RNP 복합체의 흡수(uptake)를 용이하게 할 수 있다.

[1636] [표 4]

[1637] 본 연구에 사용된 HBG1 및 HBG2 프로모터 영역을 표적화하는 gRNA의 목록. 모든 gRNA 분자를 상기 기재된 dgRNA 포맷에서 2벌 중복으로 시험하였다.

가이드 RNA 표적화 도메인 ID	Av %CD71+ (제7일)	Std %CD71+ (제7일)	% HbF+ (F 세포) 복제물의 평균 제7일	% HbF+ (F 세포), 표준 편차 제7일	% HbF+ (F 세포), 복제물의 평균 제14일	% HbF+ (F 세포), 표준 편차 제14일	% 편집된 HBG1, 복제물의 평균	% 편집된 HBG1, 표준 편차	% 편집된 HBG2, 복제물의 평균	% 편집된 HBG2, 표준 편차
모크	78.95	0.21	16.90	0.71	29.15	2.19	n/a	n/a	n/a	n/a
g8	73.15	1.20	49.25	0.64	66.95	1.48	n/a	n/a	n/a	n/a
GCR-0001	76.05	0.78	50.45	1.63	50.90	1.13	20.77	4.48	0.40	0.11
GCR-0002	73.5	1.41	24.50	0.42	33.85	1.91	1.37	0.04	1.54	0.16
GCR-0003	72.85	3.32	29.05	1.63	36.00	0.42	2.77	0.67	0.21	0.05
GCR-0004	72.4	0.42	24.60	0.14	32.55	0.21	0.82	0.04	0.82	0.04
GCR-0005	71.15	2.76	24.40	0.71	33.00	1.27	0.79	0.02	0.85	0.07
GCR-0006	76.25	2.33	43.40	0.00	43.90	2.40	4.95	0.76	5.92	1.10
GCR-0007	76.85	1.91	29.70	1.98	31.25	0.07	3.94	0.81	0.29	0.01
GCR-0008	70.75	6.15	61.60	2.12	55.75	2.90	57.72	4.53	75.09	1.96
GCR-0009	78.35	0.21	39.95	2.90	41.70	3.54	11.99	2.60	0.24	0.02
GCR-00010	73.1	5.52	46.15	2.19	43.30	0.85	5.42	0.17	9.58	0.63
GCR-0011	72.2	0.28	43.15	0.07	43.65	1.48	33.62	1.92	11.15	1.42
GCR-0012	69.3	0.42	51.00	0.00	49.05	0.07	44.34	4.16	5.39	0.08
GCR-	71.95	3.89	23.95	0.21	32.35	0.78	1.02	0.52	0.99	0.50

[1638]

0013										
GCR-0014	76.55	0.64	28.95	0.49	33.65	2.33	5.26	0.37	0.25	0.04
GCR-0015	76.1	1.56	30.95	0.64	36.45	0.92	5.58	0.01	0.29	0.05
GCR-0016	77.35	0.64	23.75	0.07	30.90	2.12	0.59	0.08	nd	nd
GCR-0017	77.7	2.97	25.35	0.78	29.80	0.42	2.27	0.56	2.00	0.05
GCR-0018	79.3	0.14	28.05	2.05	33.15	1.77	3.32	0.74	nd	nd
GCR-0019	79.2	0.99	28.30	1.70	34.65	0.92	4.47	0.14	6.17	1.19
GCR-0020	78.7	0.57	25.30	1.56	31.05	0.07	1.25	0.64	nd	nd
GCR-0021	76.05	0.07	24.75	0.64	30.70	2.69	1.09	0.01	1.41	0.20
GCR-0022	75.75	1.06	25.00	1.56	32.25	0.92	0.24	0.02	2.01	0.16
GCR-0023	76.95	2.76	23.00	0.57	31.80	0.85	0.31	0.05	nd	nd
GCR-0024	77.7	1.56	22.90	0.28	30.50	1.41	0.47	0.25	nd	nd
GCR-0025	77	0.71	26.70	0.85	32.80	0.71	1.59	0.01	nd	nd
GCR-0026	76.85	2.05	24.15	0.49	30.95	0.35	0.95	0.18	1.43	0.11
GCR-0027	78.65	0.35	25.80	0.00	29.00	1.13	nd	nd	0.78	0.01
GCR-0028	79	2.97	50.60	0.71	47.90	0.85	24.90	2.02	nd	nd
GCR-0029	74.7	0.99	28.90	1.13	32.50	1.70	0.19	0.10	4.09	1.00
GCR-0030	77	3.25	24.95	0.92	32.20	0.57	0.96	0.08	nd	nd
GCR-	79	0.99	30.40	0.71	36.00	1.70	2.87	0.93	2.98	0.84

[1639]

0031										
GCR-0032	78.55	0.21	26.30	0.28	33.00	0.71	4.25	0.56	4.61	0.66
GCR-0033	78.75	0.92	27.20	1.56	32.95	1.91	3.00	0.68	3.58	0.28
GCR-0034	80.05	0.49	50.55	1.48	50.35	1.48	6.31	0.24	12.68	0.64
GCR-0035	79.75	0.35	23.30	0.00	31.30	0.71	3.68	0.16	nd	nd
GCR-0036	79.7	0.57	23.50	0.28	30.95	0.07	0.71	0.21	1.51	0.36
GCR-0037	74.9	2.26	34.85	0.49	34.55	0.64	4.30	0.67	7.95	0.83
GCR-0038	77	0.14	25.20	0.28	30.55	2.19	1.61	0.13	nd	nd
GCR-0039	75.35	1.06	25.50	0.28	34.40	0.99	2.66	0.62	0.43	0.01
GCR-0040	76.35	1.06	28.20	0.71	33.50	0.57	2.66	0.51	2.61	0.21
GCR-0041	74.8	3.68	24.30	0.99	33.80	0.71	0.90	0.14	nd	nd
GCR-0042	73.6	4.10	26.45	0.49	33.65	2.76	1.75	0.22	2.43	0.13
GCR-0043	70.4	0.00	25.45	0.21	33.25	0.21	1.17	0.02	1.66	0.17
GCR-0044	74.7	0.00	26.10	0.85	34.25	2.19	2.40	0.14	3.65	0.01
GCR-0045	75.75	1.91	41.30	1.70	41.55	1.34	7.46	0.71	11.75	2.79
GCR-0046	77.85	1.91	41.30	1.70	42.40	0.71	1.69	0.03	13.10	0.15
GCR-0047	74.6	2.40	49.15	4.03	39.90	2.12	24.89	6.43	45.59	10.58
GCR-0048	76.85	4.88	60.75	1.77	52.95	2.62	28.49	10.88	47.10	15.49
GCR-	78.05	0.21	34.50	0.57	37.45	1.77	3.39	0.55	nd	nd

[1640]

0049										
GCR-0050	76.25	6.72	45.90	2.12	41.20	0.57	21.36	8.74	30.16	1.60
GCR-0051	81.55	2.19	54.55	0.92	51.35	1.77	25.09	0.33	77.84	1.39
GCR-0052	80.6	0.14	37.25	1.77	42.35	0.35	0.31	0.04	20.52	1.70
GCR-0053	80.35	0.64	39.75	1.20	43.45	1.63	6.81	0.57	10.47	2.01
GCR-0054	81.05	0.21	44.55	0.78	46.55	1.48	0.48	0.10	23.91	2.87
GCR-0055	81.4	0.71	22.95	0.35	31.70	0.42	1.79	0.62	2.40	0.47
GCR-0056	81.1	0.85	31.70	0.14	37.00	0.28	2.63	0.46	5.24	0.66
GCR-0057	72.4	6.08	25.25	0.78	30.25	1.20	0.88	0.56	nd	nd
GCR-0058	70.95	3.04	40.00	0.71	39.25	0.49	0.54	0.13	17.84	1.33
GCR-0059	73.85	2.76	29.75	0.64	33.85	0.07	3.36	0.18	5.56	0.91
GCR-0060	67	8.20	24.70	0.14	30.95	0.07	0.99	0.06	1.06	0.38
GCR-0061	71.1	3.82	30.05	0.49	35.80	0.42	2.91	0.67	2.07	nd
GCR-0062	74.9	4.67	43.95	3.46	43.60	2.26	7.77	1.89	13.23	2.01
GCR-0063	77.05	2.62	42.20	1.13	43.80	1.13	11.32	0.84	17.89	1.55
GCR-0064	78.85	0.21	21.85	0.92	31.60	1.41	0.83	0.12	0.81	0.08
GCR-0065	78.15	0.21	26.35	2.05	34.75	0.49	nd	nd	nd	nd
GCR-0066	78.85	1.77	23.85	0.07	31.55	0.64	4.86	0.18	4.89	0.08
GCR-0067	69	7.07	62.75	0.92	53.85	2.19	16.54	4.20	29.61	7.91

[1641]

GCR-0068	78.1	1.84	27.75	1.77	34.40	0.71	3.77	0.12	0.39	0.07
GCR-0069	79.15	2.19	30.00	0.14	35.40	1.56	4.28	0.10	6.23	1.40
GCR-0070	77.9	0.14	30.70	0.42	36.05	1.20	3.47	0.56	nd	nd
GCR-0071	79.2	0.42	36.80	2.55	39.80	2.12	67.16	8.10	nd	nd
GCR-0072	79.25	1.20	26.20	0.28	34.60	1.13	2.89	0.88	3.48	0.64

[1642]

[1643] 게놈 편집된 및 비편집된 HSPC를 형광 염료에 공액된 항체를 사용하여 태아 글로빈 및 적혈구계 세포 표면 마커 트랜스페린 수용체(CD71)의 발현 수준에 대한 유세포분석에 의해 분석하였다. 살아 있는 세포를 라이브 데드 바이올렛(Live Dead Violet)의 배제에 의해 식별하고 게이팅하였다. 배양된 세포가 비편집된 세포와 유사한 적혈모세포와 일관된 CD71+ 세포의 퍼센트를 보여주었으므로, 게놈 편집은 적혈구계 분화에 악영향을 미치지 않았다. HSPC로의 이들 gRNA RNP의 전달은 전기천공 후 7일차에 모크 전기천공된 세포(16.9%)와 비교하여 HbF를 함유하는 자손 적혈구계 세포의 증가된 퍼센트(62.75% 이하)를 초래하였다(표 4). 14일차에 HbF 유도 수준의 부

가적인 평가는 최상으로 수행하는 dsRNA 서열에 대해 높은 유도 수준을 확인시켜 주었지만(표 4), 대조군에서 검출된 백그라운드 HbF 수준은 7일차와 비교하여 더 높았다. BCL11A의 엑손 2를 표적화하는 g8의 표적화 도메인을 포함하는 dgRNA에 의한 HbF 양성 세포의 유도를 또한 병행하여 관찰하였으며, HBG1 또는 HBG2 영역으로의 몇몇 gRNA는 g8보다 더 높은 F 세포 수준 %를 초래하였다. 전기천공 후 7일차에 단리된 HBG1 및 HBG2 프로모터 영역의 게놈 DNA로부터의 PCR 생성물을 또한, 차세대 시퀀싱(NGS)하여, 세포 집단에서 편집된 대립유전자의 퍼센트를 결정하였다. HBG1 및 HBG2 프로모터 영역에서 높은 게놈 편집 퍼센트는 Cas9, 주어진 표적화 도메인의 crRNA 및 tracr를 함유하는 RNP를 이용하여 전기천공된 많은 세포 배양물에서 관찰되었으나(표 4, 도 1), 표적화 도메인이 전달되지 않은 대조군 세포(Cas9 및 tracr만 함유하는 RNP)에서는 관찰되지 않았다. 특히, 7일차에 모크 형질감염된 대조군 백그라운드보다 17% 초과의 HbF+ 세포를 초래한 dgRNA 치료는 HBG1 또는 HBG2 표적 유전자좌에서 5.92% 내지 77.84% 범위의 편집된 대립유전자를 가졌다(표 4, 도 2 및 도 3). HBG1 또는 HBG2 프로모터 영역에 대해 선택적 특이성을 갖는 일부 가이드는, 이들이 특이적인 유전자좌에서 우선적인 편집을 보여주었다(예를 들어 GCR-0001, GCR-0011 및 GCR-0012는 HBG1을 보다 효율적으로 편집하고, GCR-0034, GCR-0046, GCR-0051, GCR-0052, GCR-0054 및 GCR-0058은 HBG2를 보다 효율적으로 편집함). 이러한 상관관계에 대한 주목할 만하고 놀라운 예외는, HBG1 유전자좌의 표적 서열에 특이적이지만 HBG1 및 HBG2 유전자좌 둘 모두를 효율적으로 편집하고 또한 높은 HbF 유도를 초래하는 GCR-0008이다. GCR-0008은 HBG2 유전자좌에서 표적 서열과 단일 미스매치를 가지며, 이는 효율적인 표적-외 편집이 이 부위에서 발생함을 제안한다. 시험된 72개의 gRNA 중 일부의 표적 서열은, 태아 헤모글로빈의 유전적 지속성(HPFH; hereditary percistance of fetal hemoglobin) 발현과 연관된 몇몇 주석이 달린(annotated) 인간 돌연변이를 오버랩핑하거나 그 부근의 영역, 뿐만 아니라 HBG1 및 HBG2 둘 모두의 근접 프로모터 영역에서의 전사 인자에 대한 공지된 결합 부위(도 4 및 도 5)를 맵핑한다. 놀랍게도, 프로모터의 이들 공지된 영역의 외부에서 최상으로 수행하는 gRNA 표적 서열의 클러스터는 HBG1 및 HBG2 유전자에서 작용한다(예를 들어 chr11:5,250,094-5,250,237; hg38 및 chr11:5,255,022-5,255,164; hg38(도 4 및 도 5)에 대해 각각 맵핑함). 이들 gRNA는 GCR-0001, GCR-0006, GCR-0008, GCR-0009, GCR-0010, GCR-0011, GCR-0012, GCR-0034, GCR-0046, GCR-0048, GCR-0051, GCR-0054, GCR-0058 및 GCR-0067을 포함한다. 또 다른 양호하게 수행하는 gRNA(GCR-0028)는, 비결실 HPFH 돌연변이와 연관되지 않은 HBG1 및 HBG2 둘 모두의 근접 프로모터 영역에서 전사 인자 결합 부위(TATA 박스)를 표적화한다(각각 chr11:5,249,833-5,249,927; hg38 및 chr11:5,254,738-5,254,851; hg38(도 4 및 도 5)).

[1644] **실시예 2.2:** 성인 적혈구계 세포에서 태아 글로빈 발현의 탈억제를 위한 HSPC에서 감마 글로빈 프로모터 영역 편집 - sgRNA 포맷과 함께 CRISPR-Cas9를 사용하여 선택 부위에서의 평가

[1645] 방법:

[1646] 방법은 실시예 2.1에서와 같고, 다음의 변형을 갖는다.

[1647] 인간 $CD34^+$ 세포 배양물. 인간 CD34+ 세포는 성인 건강한 공여자 골수(Lonza 카탈로그 # 2M-101D)로부터 유래되었다. 세포를 해동시키고, 그 후에 6일 동안 확장시켰다.

[1648] Cas9와 가이드 RNA 리보핵단백질(RNP) 복합체의 조립, HSPC의 제조, 및 HSPC 내로의 RNP의 전기천공. 단일 가이드 RNA(sgRNA)를 사용한 RNP의 형성을 위해, 12 μg의 각각의 sgRNA, 12 μg의 CAS9 단백질 및 1 μL의 10 x CCE 완충제(20 mM HEPES, 100 mM KCL, 5 mM MgCL2, 5% 글리세롤 및 신선하게 첨가된 1 mM DTT)를 10 ul의 총 부피에서 조합하고, 그 후에 37℃에서 5분 동안 인큐베이션하였다. 없음/대조군 조건의 경우, sgRNA보다는 비히클을 첨가하였다. P3 완충제 +보충제에서의 세포 밀도는 3.9 x 10^6/mL이었다.

[1649] *시험관내* 적혈구생성 및 HbF-함유 적혈구계 세포에 대한 FACS 분석. 11일차에 세포를 추가 보충제가 없는 EDM으로 옮겼다. 14일차 및 18일차에 세포를 펠렛화하고, 추가 보충제가 없는 신선한 EMD에서 재현탁시켰다. 세포 분취물을 14일차 및 21일차에 HbF 발현의 분석을 위해 취하였다. 세포는 7일차와 유사하게 염색되었으나, 세포의 응집을 방지하기 위해 표면 마커를 염색으로부터 배제하였고 항-HbF 항체를 1:20 희석율로 사용하였다.

[1650] 게놈 DNA 제조 및 차세대 시퀀싱(NGS). 전기천공-후 3일째에 Quick Extract DNA 추출 용액(Epicentre Cat # QE09050)을 사용하여 편집된 HSPC 및 비편집된 HSPC로부터 게놈 DNA를 제조하였다. 하기 기재된 NGS 분석은, Cas9 단독으로 전기천공된 대조군 시료에서 어떠한 유의한 편집도 보여주지 않았다.

[1651] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱을 실시예 2.1에서 기재된 바와 같이 수행하였다. NGS 시퀀싱 데이터 QC 및 변이체 분석을 실시예 2.1에 기재된 바와 같이 수행하였다.

[1652] 결과:

[1653] 본 발명자들은 본원에서, F-세포의 생성을 초래하는 HBG1 또는 HBG2 프로모터 영역 내의 특이적인 서열의 표적화된 교란이 또한, sgRNA 포맷의 gRNA를 이용함으로써 달성될 수 있음을 보여주었다.

[1654] [표 5]

[1655] 본 연구에 사용된 HBG1 및 HBG2 프로모터 영역을 표적화하는 gRNA의 목록. 모든 gRNA 분자를 상기 기재된 sgRNA 포맷에서 2벌 중복으로 시험하였다

가이드 RNA 표적화 도메인 ID	% HbF+ (F 세포), 복제물의 평균 7일차	% HbF+ (F 세포), 표준 편차 7일차	% HbF+ (F 세포), 복제물의 평균 14일차	% HbF+ (F 세포), 표준 편차 14일차	% HbF+ (F 세포), 복제물의 평균 21일차	% HbF+ (F 세포), 표준 편차 21일차	% 편집된 HBG1, 복제물의 평균	% 편집된 HBG1, 표준 편차	% 편집된 HBG2, 복제물의 평균	% 편집된 HBG2, 표준 편차
없음/대조군	39.0	1.8	32.7	0.6	14.4	4.9	n/a	n/a	n/a	n/a
GCR-0001	60.8	2.5	49.2	0.4	20.5	7.6	15.8	3.0	0.0	0.0
GCR-0008	67.1	0.8	55.2	2.3	25.6	9.3	55.2	2.0	40.9	2.8
GCR-0010	61.3	0.8	52.3	1.0	14.0	0.8	4.8	0.8	3.0	0.0
GCR-0028	58.7	1.6	49.2	0.1	14.1	2.1	nd	nd	nd	nd
GCR-0047	52.6	1.6	41.2	4.1	16.7	3.3	nd	nd	nd	nd
GCR-0048	64.3	2.3	57.5	2.3	28.8	11.0	53.7	5.3	57.2	5.3
GCR-0051	65.3	3.0	59.1	2.0	20.6	5.7	27.3	0.3	50.6	3.8
GCR-0053	71.7	3.3	73.7	3.3	26.2	2.6	11.9	4.2	28.0	6.1
GCR-0054	67.3	2.0	61.4	4.0	23.4	3.5	nd	nd	18.4	1.9
GCR-0062	61.1	5.2	54.8	1.1	16.7	1.7	2.5	1.8	3.9	2.7
GCR-0063	63.6	3.7	58.0	1.7	24.6	1.3	16.4	0.2	29.9	2.9
GCR-0067	74.6	0.4	70.4	3.7	36.6	0.6	23.7	0.0	37.5	4.4

[1656]

[1657] 성인 골수-유래 HSPC를, 재조합 *S. 피오게네스* Cas9 단백질(SEQ ID NO: 236) 및 sgRNA 포맷의 지시된 gRNA로부터 형성된 RNP 복합체를 이용하여 전기천공시켰다. 생성된 게놈 편집된 및 비편집된 HSPC를 형광 염료에 공액된 항체를 사용하여 태아 글로빈 및 적혈구계 세포 표면 마커 트랜스페린 수용체(CD71)의 발현 수준에 대한 유세포 분석에 의해 분석하였다. 살아 잇는 세포를 라이브 데드 바이올렛(Live Dead Violet)의 배제에 의해 식별하고 게이팅하였다. HSPC로의 이들 gRNA RNP의 전달은 전기천공 후 7일차에 모크 전기천공된 세포(39.0%)와 비교하여 HbF를 함유하는 자손 적혈구계 세포의 증가된 퍼센트(74.6% 이하)를 초래하였다(표 5). 14일차 및 21일차에 HbF 유도 수준의 부가적인 평가는 최상으로 수행하는 sgRNA 서열에 대해 높은 유도 수준을 확인시켜 주었다(표 5). 대조군에서 검출된 백그라운드 HbF 수준은 7일차와 비교하여 더 높았다. 전기천공 후 3일차에 단리된 HBG1 및 HBG2 프로모터 영역의 게놈 DNA로부터의 PCR 생성물을 또한, 차세대 시퀀싱(NGS)하여, 세포 집단에서 편집된 대립유전자의 퍼센트를 결정하였다. HBG1 및 HBG2 프로모터 영역에서 높은 게놈 편집 퍼센트(대형 결실을 배제

함)는 Cas9 및 주어진 표적화 도메인의 sgRNA를 함유하는 RNP를 이용하여 전기천공된 많은 세포 배양물에서 관찰되었으나(표 5), sgRNA가 전달되지 않은 대조군 세포(Cas9만 함유함)에서는 관찰되지 않았다.

[1658] 실시예 2.1에서 7일차에 HbF+ 적혈구계 세포에서 17% 초과의 증가를 갖는 표적화 부위의 선택을 이 연구에 포함시켰다. 실험 설계는 실시예 2.1에서 이전에 기재된 것과 2가지 유의한 방식: gRNA 포맷(dgRNA보다는 sgRNA) 및 HSPCource(상이한 공여자 및 동원된 말초 혈액-유래보다는 골수-유래)에서 상이하였다. 이론으로 결부시키고자 하는 것은 아니지만, 연구 사이에서 비편집된 세포 배양물 내의 HbF+ 세포의 기준선 퍼센트에서의 차이는 HSPC 공급원 사이의 내재적인(inherent) 차이에 의해 유발될 수 있다. 이들 차이에도 불구하고, GCR-47을 제외한 모든 표적화 부위는 7일차에 HbF+ 적혈구계 세포에서 175 초과의 증가와 연관이 있은 채 남아 있었다(도 7). 이러한 증가는 GCR-0067에 대해 21일차까지 유지되었다(도 7). 부가적으로, HBG1, HBG2 또는 둘 모두에서 25% 초과의 삽입결실 형성은 GCR-0008, GCR-0048, GCR-0051, GCR-0053, GCR-0063 및 GCR-0067과 함께 관찰되었다(표 5). dgRNA 포맷(표 4)에서와 같이, sgRNA 포맷에서 GCR-0008은 표적-적중 HBG1 부위 및 표적-외 HBG2 부위 둘 모두에서 효율적인 편집과 연관이 있었다(표 5). 이 연구에서 GCR-0047을 이용한 더 낮은 HbF+ 세포 유도가 감소된 편집과 연관이 있는지는 결정되지 않았다. 주목할 만하게는, 7일차 및 14일차에 17% 초과의 HbF+ 세포 유도는 HBG1 및 HBG2 유전자(예를 들어 각각 chr11:5,250,094-5,250,237; hg38 및 chr11:5,255,022-5,255,164; hg38(도 4 및 도 5)에 대해 맵핑함), 구체적으로 GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051, GCR-0054 및 GCR-0067(도 7)에서 프로모터 기능의 공지된 영역의 외부의 표적화 서열과 연관이 있었다. 이들 중에서, GCR-0008, GCR-0048 및 GCR-0067은 또한, 21일차에 HbF+ 세포에서의 10% 초과의 증가 및 HBG1, HBG2 또는 둘 모두에서 25% 초과의 삽입결실을 가졌다(표 5 및 도 7).

[1659] **실시예 2.3:** 성인 적혈구계 세포에서 태아 글로빈 발현의 탈억제를 위한 HSPC에서 감마 글로빈 프로모터 영역 편집 - sgRNA 포맷과 함께 CRISPR-Cas9를 사용하여 편집 패턴의 부가적인 분석

[1660] 방법:

[1661] 방법은 실시예 2.1에서와 같고, 다음의 변형을 갖는다.

[1662] 인간 $CD34^+$ 세포 배양물. 인간 CD34+ 세포는 성인 건강한 공여자 골수(Lonza 카탈로그 # 2M-101D)로부터 유래되었다. 세포를 해동시키고, 그 후에 6일 동안 확장시켰다.

[1663] Cas9와 가이드 RNA 리보핵단백질(RNP) 복합체의 조립, HSPC의 제조, 및 HSPC 내로의 RNP의 전기천공. 단일 가이드 RNA(sgRNA)를 사용한 RNP의 형성을 위해, 12 μg의 각각의 sgRNA, 12 μg의 CAS9 단백질 및 1 μL의 10 x CCE 완충제(20 mM HEPES, 100 mM KCL, 5 mM MgCL2, 5% 글리세롤 및 신선하게 첨가된 1 mM DTT)를 10 ul의 총 부피에서 조합하고, 그 후에 37℃에서 5분 동안 인큐베이션하였다. 없음/대조군 조건의 경우, sgRNA보다는 비히클을 첨가하였다. P3 완충제 +보충제에서의 세포 밀도는 3.9 x 10^6/mL이었다.

[1664] *시험관내* 적혈구생성 및 HbF-함유 적혈구계 세포에 대한 FACS 분석. 11일차에 세포를 추가 보충제가 없는 EDM으로 옮겼다. 14일차 및 18일차에 세포를 펠렛화하고, 추가 보충제가 없는 신선한 EMD에서 재현탁시켰다. 세포 분취물을 14일차 및 21일차에 HbF 발현의 분석을 위해 취하였다. 세포는 7일차와 유사하게 염색되었으나, 세포의 응집을 방지하기 위해 표면 마커를 염색으로부터 배제하였고 항-HbF 항체를 1:20 희석율로 사용하였다.

[1665] 게놈 DNA 제조 및 차세대 시퀀싱(NGS). 전기천공-후 3일째에 Quick Extract DNA 추출 용액(Epicentre Cat # QE09050)을 사용하여 편집된 HSPC 및 비편집된 HSPC로부터 게놈 DNA를 제조하였다. 하기 기재된 NGS 분석은, Cas9 단독으로 전기천공된 대조군 시료에서 어떠한 유의한 편집도 보여주지 않았다.

[1666] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱을 실시예 2.1에서 기재된 바와 같이 수행하였다. NGS 시퀀싱 데이터 QC 및 변이체 분석을 실시예 2.1에 기재된 바와 같이 수행하였다.

[1667] 역위 및 대형 결실의 검출을 위한 비-정량적 PCR. 하기 프라이머 세트를 사용하여 HBG1 및 HBG2의 영역에서 gDNA를 증폭시켰다. P1: 포워드 프라이머 5'-TGCTGAGATGAAACAGGCGT-3'(SEQ ID NO: 257), 리버스 프라이머 5'-TTAGGCATCCACAAGGGCTG-3'(SEQ ID NO: 258), HBG1과 HBG2 표적/표적-외 부위 사이의 결실에 대해 예상된 약 2.8 kb 생성물, HBG1과 HBG2 표적/표적-외 부위 사이의 역위에 대해 예상된 약 7.7 kb 생성물, 대형 결실 또는 역위가 없는 경우에 대해 예상된 약 7.7 kb 생성물(개별적으로 HBG1 및/또는 HBG2 표적/표적-외 부위에서 비편집된 또는 더 작은 삽입결실). P2: 포워드 프라이머 5'-GCTCTACAAATGGAACCCAACC-3'(SEQ ID NO: 259), 리버스 프라이머 5'-CTGCTCTGATCTCTAACACCTCA-3'(SEQ ID NO: 260), HBG1과 HBG2 표적/표적-외 부위 사이의 역위에 대해 어떠한 생성물도 예상되지 않음, HBG1과 HBG2 표적/표적-외 부위 사이, 대형 결실 또는 역위에 대해 예상된 약 3.8

kb 생성물(개별적으로 HBG1 및/또는 HBG2 표적/표적-외 부위에서 비편집된 또는 더 작은 삽입결실). P3: 포워드 프라이머 5'-GAAGATACAGCTTGCCTCCGA-3'(SEQ ID NO: 261), 리버스 프라이머 5'-TTGCTGAGATGAAACAGGCGT-3'(SEQ ID NO: 262), HBG1과 HBG2 표적/표적-외 부위 사이의 역위에 대해 어떠한 생성물도 예상되지 않음, HBG1과 HBG2 표적/표적-외 부위 사이의 역위에 대해 예상된 약 1.75 kb 생성물, 대형 결실 또는 역위에 대해 어떠한 생성물도 예상되지 않음(개별적으로 HBG1 및/또는 HBG2 표적/표적-외 부위에서 비편집된 또는 더 작은 삽입결실). PCR 생성물을 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8 and 10 kb 밴드를 함유하는 참조 사다리와 함께 아가로스 겔 상에서 시각화하였다(L; New England Biolabs 카탈로그 # N3232L). 선택 생성물을 지시된 바와 같이 단리하고, 실시예 2.1에 기재된 바와 같이 NGS를 받게 하였다.

[1668] 대형 결실의 정량화. HBG1 및 HBG2 프로모터에서 표적화 부위 사이의 대형 결디지털 방울 PCR(ddPCR)에 의해 제조업체의 권고사항에 따라 복사수 결정을 위한 비-경쟁적 검정법 포맷에서 정량화하였다. 간략하게는, gDNA를 프로브용 ddPCR SuperMix(dUTP 없음)(BioRad Cat# 1863024), HindIII-HF 제한 효소(NEB Cat# R3104S) 및 각각의 프라이머 프로브 믹스와 조합하고, 프로브용 방울 발생 오일(BioRad Cat# 1863005)과 함께 DG8 카트리지로 옮기고, QX200 방울 발생기(BioRad)를 이용하여 방울을 발생시켰다. 방울을 96-딥 웰 반응 모듈이 있는 C1000 Touch Thermal Cylcer(BioRad) 상에서 PCR을 받게 하고, 뒤이어 QX200 방울 리더(BioRad)를 이용하여 검출하였다. 1 ul 당 복사물를 QuantaSoft 소프트웨어(BioRad)를 이용한 분석에 의해 결정하였다. 커스텀 프라이머 프로브 세트(Life Technologies Cat# APZW76R, PN4331348, 포워드 프라이머: ACGGATAAGTAGATATTGAGGTAAGC(SEQ ID NO: 263), 리버스 프라이머: GTCTCTTTCAGTTAGCAGTGG(SEQ ID NO: 264), FAM TaqMan 프로브: ACTGCGCTGAAACTGTGGCTTTATAG(SEQ ID NO: 265))를 사용하여, HBG1-HBG2 유전자간 영역 내에서 gDNA를 증폭시켰다. TaqMan 복사수 참조 검정법, 인간, RNase P(Thermo Fisher cat# 4403326)를 참조 앰플리콘으로서 사용하였다. HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물은 제조업체의 보고된 선형 범위 내에 있었다. 퍼센트 결실을 100% x (1 - HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물의 비)로서 보고하였다. 비편집된 대조군 시료는 2.4%의 계산된 퍼센트 결실을 가졌으며, 이는 검정법에서 백그라운드를 반영할 수 있다.

[1669] 결과:

[1670] 본 발명자들은 본원에서, F-세포의 생성을 초래하는 HBG1 및 HBG2 프로모터 영역 내의 특이적인 서열의 표적화된 교란이, F-세포의 생성을 초래하는 HBG1 및 HBG2 프로모터 영역 내의 특이적인 서열의 표적화된 교란이 HBG1 또는 HBG2 표적 또는 표적-외 부위에서의 삽입결실 둘 모두, 뿐만 아니라 개입 영역의 결실 및 역위와 연관이 있음을 보여주었다. 앰플리콘의 NGS 분석은 이러한 관찰을 확인시켜 준다.

[1671] [표 6]

[1672] 제1 독립적인 공여자로부터의 세포를 편집하기 위해 본 연구에 사용된 HBG1 및 HBG2 프로모터 영역을 표적화하는 gRNA의 목록. 모든 gRNA 분자를 상기 기재된 sgRNA 포맷에서 2벌 중복으로 시험하였다.

가이드 RNA 표적화 도메인 ID	% HbF+ (F 세포), 복제물의 평균 7 일차	% HbF+ (F 세포), 표준 편차 7 일차	% HbF+ (F 세포), 복제물의 평균 14 일차	% HbF+ (F 세포), 표준 편차 14 일차	% HbF+ (F 세포), 복제물의 평균 21 일차	% HbF+ (F 세포), 표준 편차 21 일차	% 편집된 HBG1, 복제물의 평균	% 편집된 HBG1, 표준 편차	% 편집된 HBG2, 복제물의 평균	% 편집된 HBG2, 표준 편차
없음/대조군	**37.6**	**3.0**	**17.1**	**0.4**	**32.4**	**5.9**	**n/a**	**n/a**	**n/a**	**n/a**
GCR-0001	**61.3**	**1.1**	**45.6**	**0.4**	**75.1**	**2.1**	**59.1**	**3.9**	**1.0**	**0.4**
GCR-0008	**62.9**	**1.3**	**42.8**	**2.1**	**67.3**	**1.3**	**86.3**	**6.2**	**85.6**	**5.5**
GCR-0010	**69.5**	**3.0**	**53.4**	**0.9**	**75.0**	**0.4**	**19.5**	**6.2**	**25.1**	**1.0**
GCR-0048	**59.2**	**4.2**	**43.8**	**0.8**	**66.2**	**3.6**	**73.5**	**9.2**	**80.4**	**3.6**
GCR-0051	**61.9**	**0.4**	**49.7**	**1.7**	**70.1**	**6.2**	**53.0**	**0.5**	**86.4**	**0.6**
GCR-0067	**72.2**	**1.8**	**58.1**	**1.7**	**77.4**	**4.2**	**68.8**	**1.2**	**77.5**	**1.6**

[1673]

[1674] [표 7]

[1675] 제2 독립적인 공여자로부터의 세포를 편집하기 위해 본 연구에 사용된 HBG1 및 HBG2 프로모터 영역을 표적화하는 gRNA의 목록. 모든 gRNA 분자를 상기 기재된 sgRNA 포맷에서 2벌 중복으로 시험하였다.

가이드 RNA 표적화 도메인 ID	% HbF+ (F 세포), 복제물의 평균 7 일차	% HbF+ (F 세포), 표준 편차 7 일차	% HbF+ (F 세포), 복제물의 평균 14 일차	% HbF+ (F 세포), 표준 편차 14 일차	% HbF+ (F 세포), 복제물의 평균 21 일차	% HbF+ (F 세포), 표준 편차 21 일차	% 편집된 HBG1, 복제물의 평균	% 편집된 HBG1, 표준 편차	% 편집된 HBG2, 복제물의 평균	% 편집된 HBG2, 표준 편차
없음/대조군	**39.3**	**0.2**	**23.1**	**0.8**	**52.7**	**0.3**	**n/a**	**n/a**	**n/a**	**n/a**
GCR-0001	**58.8**	**0.7**	**44.7**	**3.0**	**80.2**	**0.8**	**62.7**	**10.0**	**1.0**	**0.0**
GCR-0008	**61.6**	**2.2**	**44.4**	**0.7**	**75.0**	**0.5**	**92.1**	**5.7**	**86.4**	**5.5**
GCR-0010	**65.5**	**2.5**	**47.7**	**3.0**	**79.4**	**0.0**	**17.7**	**n/a**	**28.5**	**9.0**
GCR-0048	**60.5**	**0.3**	**46.1**	**3.3**	**79.3**	**1.2**	**68.4**	**6.6**	**76.9**	**9.3**
GCR-0051	**64.5**	**1.7**	**47.1**	**4.2**	**78.1**	**0.4**	**62.7**	**7.8**	**86.4**	**3.0**
GCR-0067	**68.8**	**4.6**	**54.9**	**0.4**	**84.5**	**3.1**	**77.4**	**8.7**	**78.0**	**4.2**

[1676]

[1677] 2명의 독립적인 공여자로부터의 성인 골수-유래 HSPC를, 재조합 *S. 피오게네스* Cas9 단백질(SEQ ID NO: 236) 및 sgRNA 포맷의 지시된 gRNA로부터 형성된 RNP 복합체를 이용하여 전기천공시켰다. 평가된 sgRNA는 HBG1 및 HBG2 유전자에서 프로모터 기능의 공지된 영역의 외부에서 표적화 서열을 가졌다(예를 들어 chr11:5,250,094-5,250,237; hg38 및 chr11:5,255,022-5,255,164; hg38(도 4 및 도 5) 각각에 대해 맵핑함). 생성된 게놈 편집된 및 비편집된 HSPC를 형광 염료에 공액된 항체를 사용하여 태아 글로빈 및 적혈구계 세포 표면 마커 트랜스페린 수용체(CD71)의 발현 수준에 대한 유세포분석에 의해 분석하였다. 살아 있는 세포를 라이브 데드 바이올렛의 배제에 의해 식별하고 게이팅하였다. HSPC로의 이들 gRNA RNP의 전달은 전기천공 후 7일차, 14일차 및 21일차에 모크 전기천공된 세포와 비교하여 HbF를 함유하는 자손 적혈구계 세포의 증가된 퍼센트를 초래하였다(표 6 및 표 7). 표적화 부위를 포함하는 모두는 모든 시점 - 7일차, 14일차 및 21일차에 공여자 둘 모두로부터의 HbF+ 적혈구계 세포에서 17% 초과의 증가와 연관이 있었다(도 8). 전기천공 후 3일차에 단리된 HBG1 및 HBG2 프로모터 영역의 게놈 DNA로부터의 PCR 생성물을 또한, 차세대 시퀀싱(NGS)하여, 세포 집단에서 편집된 대립유전자의 퍼센트를 결정하였다. HBG1 및 HBG2 프로모터 영역 둘 모두에서 높은 게놈 편집 퍼센트(53% 내지 92% 삽입결실)는 Cas9 및 표적화 도메인 GCR-0008, GCR-0048, GCR-0051 및 GCR-0067을 갖는 sgRNA를 함유하는 RNP를 이용하여 전기천공된 두 공여자 모두로부터의 세포 배양물에서 관찰되었으나(표 6 및 표 7), sgRNA가 전달되지 않은 대조군 세포(Cas9만 함유함)에서는 관찰되지 않았다. 표적화 도메인 GCR-0010은 두 공여자 모두로부터의 세포 배양물에서 HBG1 및 HBG2 프로모터 영역 둘 모두에서 감소되었으나 제법 큰 편집 퍼센트(17% 내지 28% 삽입결실)와 연관이 있은 반면; GCR-0001은 두 공여자에서 HBG2(1% 및 1%)와 비교하여 HBG1(59% 및 62%)에서 효율적이고 선택적인 편집과 연관이 있었다(표 6 및 표 7).

[1678] GCR-0048 및 GCR-0067 표적화 도메인은 HBG1과 HBG2 둘 모두에 존재하고, 다른 표적화 도메인은 다른 프로모터에서 잠재적인 미스매칭된 표적-외 부위를 갖는 HBG1(GCR-0001, GCR-0008 및 GCR-0010) 또는 HBG2(GCR-0051)에 존재한다. HBG1 및 HBG2 표적/표적-외 부위 둘 모두에서 동시적인 절단은 개입 4.9 kb 게놈 서열의 결실 및/또는 역위를 초래하는 잠재성을 가지며; 본원에서 이는 이따금 '4.9 kb' 또는 'HBG1-HBG2' 또는 '대형' 역위 또는 결실로 지칭된다. 따라서, 3개의 상이한 PCR 반응 세트를 사용하여, 이 영역의 결실 또는 역위를 갖는 게놈의 존재를 검출하였다. 따라서, 3개의 상이한 PCR 반응 세트를 사용하여, 이러한 영역의 결실 또는 역위를 갖는 게놈의 존재를 검출하였다. HBG1-HBG2 결실 또는 역위가 없는 게놈 서열이 생성된 3.8 kb 생성물과 함께 증폭될

것인 반면, HBG1-HBG2 결실 또는 역위를 갖는 서열은 증폭되지 않도록 P2 반응을 설계하였다. 대략 이러한 크기의 밴드를 모든 시료에서 검출하였으며(도 9), 이는 각각의 개별 HBG1 또는 HBG2 표적화 부위에서 더 작은 삽입결실의 검출과 일관되었다(표 6 및 표 7). HBG1-HBG2 역위된(inverted) 게놈 서열이 생성된 1.7 kb 생성물과 함께 증폭될 것인 반면, 이러한 역위가 없거나 HBG1-HBG2 결실이 있는/없는 서열은 증폭되지 않도록 P3 반응을 설계하였다. 대략 이러한 크기의 밴드를, 지시된 표적화 도메인과 함께 Cas9 및 sgRNA를 함유하는 RNP를 이용하여 전기천공된 배양물에서 검출하였으나, 비편집된 대조군 배양물에서는 검출되지 않았다(도 9). 마지막으로, HBG1과 HBG2 표적화 영역 둘 모두를 스팬(span)하여, HBG1-HBG2 결실이 없는 서열뿐만 아니라 HBG1-HBG2 영역이 역위된 서열의 증폭 둘 모두가 7.7 kb 생성물을 생산할 것인 한편, HBG1-HBG2 결실을 함유하는 서열의 증폭이 2.8 kb 생성물을 초래하도록 P1 반응을 설계하였다. 사실상, 비편집된 대조군 시료는 대략 8 kb에서 두드러진 상부 밴드 및 대략 2.8 kb에서 희미하며 잠재적으로 비-특이적인 밴드를 가졌다. 대조적으로, 지시된 표적화 도메인과 함께 Cas9 및 sgRNA를 함유하는 RNP를 이용하여 전기천공된 배양물은 대략 2.8 kb에서 두드러진 밴드 및 대략 8 kb에서 희미한 밴드를 가졌다(도 9). 별표로 지시된 각각의 밴드에 대한 DNA를 단리하고, 차세대 시퀀싱시켜, 예측된 서열(기재된 바와 같이 HBG1-HBG2 역위 또는 결실이 있거나 없음)의 증폭으로서 이의 정체성(identity)을 확인하였다.

[1679] HBG1 및 HBG2 편집 부위를 스패닝하는 영역의 추가의 서열 특징화를 위해, 제4 장범위(long range) PCR 조건(P4)을 개발하여, gRNA GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051 및 GCR-0067에 대한 HBG1 및 HBG2 편집 부위를 포괄하는 6192 bp 영역을 증폭시켰다. 게놈 편집된 시료의 증폭은 2개 앰플리콘 크기(1.2 kb 및 6.2 kb)를 초래하는 한편, 6.2 kb 및 1.2 kb 앰플리콘의 서열 분석은, 1.2 kb 앰플리콘이 HBG1 및 HBG2 절단 부위 사이에서 4.9 kb 결실의 결과인 한편 6.2 kb 앰플리콘이 HBG1 및 HBG2 편집 부위에서 야생형 및 다양한 삽입결실 대립유전자로 구성됨을 보여준다. 시퀀싱 분석은 또한, HBG1과 HBG2 절단 부위 사이에서 낮은 수준의 역위를 보여주며, 여기서, 2개 부위 사이에서 절제된 서열은 반대 방향에서 게놈 내로 재혼입되었다. 이러한 분석은 정량적이지 않지만, HBG1 및 HBG2 절단 부위를 스패닝하는 시퀀싱 리드의 작은 퍼센트(1% 미만)에서만 역위가 검출됨에 따라 이러한 역위는 매우 드문 경향이 있다.

[1680] 이들 결과와 함께, BG1 또는 HBG2 표적/표적-외 부위에서의 삽입결실 외에도, 지시된 표적화 도메인과 함께 Cas9 및 sgRNA를 함유하는 RNP를 이용하여 전기천공된 배양물이 개입 HBG1-HBG2 영역의 결실 또는 역위를 갖는 편집된 대립유전자를 함유함을 제안한다. 따라서, 편집-후 주어진 대립유전자는 HBG1-HBG1 역위, HBG1-HBG2 결실, HBG1 부위로 국소화된 삽입결실, HBG2 부위로 국소화된 삽입결실, 또는 개입 영역의 변형이 없는 HBG2 부위로 국소화된 삽입결실과 함께 HBG1 부위로 국소화된 삽입결실일 수 있을 것이다. gRNA GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051 및 GCR-0067을 이용한 편집에 의해 발생된 가장 보편적인 표적-적중 편집 수선 패턴 변이체를 표 7-2에서 보여준다. 제시된 변이체는 HBG1 및 HBG2 유전자좌에서 발생된 국소화된 삽입결실 및 HBG1 및 HBG2 절단 부위 사이의 서열의 절제에 의해 유발된 대형 4.9 kb 결실이다. 국소화된 삽입결실을 실시예 2.1에 기재된 바와 같이 PCR 및 NGS 분석을 사용하여 특징화하였다. 정량적 ddPCR 검정법을 개발하여, HBG2 영역으로 개입 HBG1의 대형 4.9 kb 결실을 갖는 대립유전자의 빈도를 결정하였다(도 10). 간략하게는, HBG1 프로모터(도 10에 제시된 5.2 kb Fwd와 Rev 프라이머 사이)의 업스트림 영역의 복사수는 이러한 방법에 기재된 바와 같이 RPPH1 유전자좌에서의 복사수와의 관계에서 정의되었다. 국소화된 더 작은 삽입결실에 대한 대립유전자 빈도는 총 삽입결실 빈도에 비하지 않는데, 왜냐하면 NGS에 대한 HBG1 또는 HBG2 부위의 증폭이 4.9 kb 역위 또는 결실을 함유하는 대립유전자에서 발생하지 않을 것이기 때문이다. 예를 들어, gRNA GCR-0001의 경우 4.9 kb 결실의 빈도는 35.2%이며, 나머지 64.8% (100% - 35.2%, 역위는 무시함)는 야생형과 더 작은 국소화된 삽입결실의 혼합이다. 따라서, 9% 단일 염기쌍 A 염기 결실은 총 삽입결실 빈도의 5.8%, 즉, 64.8%의 9%일 것이다. 표 7-2에 제시된 대립유전자 빈도는 실험 사이에서 약간 다양하고, 절대값으로 여겨져서는 안 된다.

[1681] [표 7-2]

[1682] 제1 공여자로부터의 세포를 gRNA GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051 및 GCR-0067을 이용하여 편집함으로써 발생된 상부 표적-적중 편집 수선(top on-target editing repair pattern)(수합적으로 임의의 gRNA 분자에 대해, 본원에서 "삽입결실 패턴"으로도 지칭됨) 변이체. 염색체 11 참조 게놈 구축(build) 38에 비한 변이체 크기, 변이체 유형(Ins = 삽입, Del = 결실), 참조 대립유전자, 변이체 대립유전자, 변이체 개시 및 종결 위치, 및 대립유전자 빈도가 제시된다. gRNA GCR-0001을 이용하여 편집된 HBG2 유전자좌는 유의한 국소화된 indel을 보여주지 않지만, qPCR 결과를 기초로, 상기 편집된 HBG2 유전자좌는 4.9 kb 결실을 초래한다. *미스매치를 갖는 HBG1/2 표적 부위.

[1683]

gRNA 명칭	크기 (bp)	유형	참조 대립유전자	변이체 대립유전자	변이체 개시 및 종결 위치	대립유전자 빈도
GCR-0001 (HBG1)	4928	Del	제시되지 않음	제시되지 않음	5250172-5255100	35.2%
	-1	Del	CA	C	5250171-5250172	9.0%
	1	Ins	A	AT	5250172-5250173	8.1%
	-1	Del	AT	A	5250172-5250173	8.0%
	-2	Del	ATA	A	5250172-5250174	5.1%
	1	Ins	A	AA	5250172-5250172	2.6%
GCR-0008 (HBG1)	4928	Del	제시되지 않음	제시되지 않음	5250150-5255078	40.4%
	1	Ins	T	TA	5250150-5250151	50.5%
	-2	Del	AGC	A	5250151-5250153	3.4%
	-1	Del	TA	T	5250150-5250151	3.3%
	-1	Del	AG	A	5250151-5250152	2.4%
	-4	Del	TAGCT	T	5250150-5250154	2.2%
GCR-0008* (HBG2)	4928	Del	제시되지 않음	제시되지 않음	5250150-5255078	40.4%
	1	Ins	T	TA	5255078-5255079	55.8%
	-19	Del	CCCTTTAGCTAGTTTTCTTC (SEQ ID NO: 266)	A	5255073-5255092	5.6%
	-2	Del	AGC	A	5255079-5255081	2.9%
	-6	Del	TTTAGCT	T	5255076-5255082	2.8%
	-1	Del	TA	T	5255078-5255079	2.6%
GCR-0010 (HBG1)	4928	Del	제시되지 않음	제시되지 않음	5250151-5255079	43.6%
	1	Ins	A	AG	5250151-5250152	5.6%
	-4	Del	AGCTA	A	5250151-5250155	1.8%
	1	Ins	A	AA	5250151-5250152	1.6%
GCR-0010* (HBG2)	4928	Del	제시되지 않음	제시되지 않음	5250151-5255079	43.6%
	1	Ins	A	AG	5255079-5255080	5.7%
	-19	Del	CCCTTTAGCTAGTTTTCTTC (SEQ ID NO: 266)	A	5255073-5255092	2.6%
	-52	Del	GCCTTGTT.....TTCCCTTTA	G	5255027-5255079	2.0%
	-4	Del	TAGCT	T	5255078-5255082	1.5%
	-6	Del	TTTAGCT	T	5255076-5255082	1.5%
GCR-0048 (HBG1)	4928	Del	제시되지 않음	제시되지 않음	5250187-5255115	43.5%
	-1	Del	AC	A	5250187-5250188	40.0%
	-2	Del	ACT	A	5250187-5250189	10.1%
	-4	Del	ACTTC	A	5250187-5250191	5.2%
	1	Ins	A	AA	5250187-5250188	1.9%
	-1	Del	GA	G	5250186-5250187	2.1%

[1684]

[1685]

GCR-0048 (HBG2)	4928	Del	제시되지 않음	제시되지 않음	5250187-5255115	43.5%
	-1	Del	AC	A	5255115-5255116	42.4%
	-2	Del	ACT	A	5255115-5255117	8.9%
	-3	Del	ACTT	A	5255115-5255118	2.7%
	-4	Del	ACTTC	A	5255115-5255119	2.5%
	-15	Del	GGACTTCTTTTGTCAG (SEQ ID NO: 267)	G	5255113-5255128	2.5%
GCR-0051* (HBG1)	4928	Del	제시되지 않음	제시되지 않음	5250150-5255078	44.9%
	1	Ins	T	TA	5250150-5250151	39.4%
	-1	Del	TA	A	5250150-5250151	1.3%
	-2	Del	AGC	A	5250151-5250153	0.7%
	-14	Del	TTAGCTAGTTTCCTT (SEQ ID NO: 268)	T	5250149-5250163	0.9%
	-6	Del	TTTAGCT	T	5250148-5250154	0.6%
GCR-0051 (HBG2)	4928	Del	제시되지 않음	제시되지 않음	5250150-5255078	44.9%
	1	Ins	T	TA	5255078-5255079	52.5%
	-19	Del	CCCTTTAGCTAGTTTTCTTC (SEQ ID NO: 269)	C	5255073-5255092	4.9%
	-2	Del	AGC	A	5255079-5255081	3.2%
	-4	Del	TAGCT	T	5255078-5255082	2.4%
	-1	Del	TA	T	5255078-5255079	2.2%
GCR-0067 (HBG1)	4928	Del	제시되지 않음	제시되지 않음	5250099-5255027	55.3%
	-6	Del	GCCTTTG	G	5250093-5250099	11.3%
	-1	Del	TG	T	5250098-5250099	9.4%
	-1	Del	GC	G	5250099-5250100	7.0%
	-5	Del	GCCTTG	G	5250099-5250104	5.4%
	-2	Del	GCC	G	5250099-5250101	3.4%
GCR-0067 (HBG2)	4928	Del	제시되지 않음	제시되지 않음	5250099-5255027	55.3%
	-6	Del	GCCTTTG	G	5255021-5255027	36.1%
	-5	Del	GCCTTG	G	5255027-5255032	15.7%

[1686]

	-1	Del	TG	T	5255026-5255027	4.0%
	-1	Del	GC	G	5255027-5255028	2.4%
	-2	Del	GCC	G	5255027-5255029	2.1%

[1687] HBG1 및 HBG2 프로모터 표적화 부위 사이에서 대형 4.9 kb 결실의 높은 빈도를 포함하여 감마 글로빈 유전자좌에서의 높은 전반적인 편집 빈도는 지시된 표적화 도메인의 sgRNA를 함유하는 RNP를 이용하여 세포를 전기천공시킨 후 관찰되었다. 제시된 표적-적중 편집 패턴은, 기재된 바와 같이 적혈구계 분화 후 증가된 F 세포를 발생시킨 세포에서 식별되었다. 실시형태에서, 본원에 기재된 임의의 gRNA 분자에 대한 삽입결실 패턴은 HBG1 유전자좌에서 검출된 가장 빈번한 1, 2, 3, 4, 5 또는 6개 삽입결실을 포함한다. 실시형태에서, 본원에 기재된 임의의 gRNA 분자에 대한 삽입결실 패턴은 HBG2 유전자좌에서 검출된 가장 빈번한 1, 2, 3, 4, 5 또는 6개 삽입결실을 포함한다. 실시형태에서, 본원에 기재된 임의의 gRNA 분자에 대한 삽입결실 패턴은 예를 들어 표 7-2에 기재된 바와 같이(그러나 대략 4.9 kb 결실을 이중-계수하지 않음) HBG1 유전자좌에서 검출된 가장 빈번한 1, 2, 3, 4, 5 또는 6개 삽입결실 및 HBG2 유전자좌에서 검출된 가장 빈번한 1, 2, 3, 4, 5 또는 6개 삽입결실을 포함한다.

[1688] **실시예 2.4:** 감마 글로빈 프로모터 영역 편집 후, HSPC의 단리된 하위-집단에서의 예시적인 편집 패턴

[1689] 방법:

[1690] 방법은 실시예 2.1에서와 같고, 다음의 변형을 갖는다.

[1691] 인간 $CD34^+$ 세포 배양물. 인간 CD34+ 세포를 성인 건강한 공여자 골수(Hemacare 카탈로그 # M001F-GCSF-3)로부터의 G-CSF 동원된 말초 혈액으로부터 면역선택(Miltenyi)을 제조업체의 설명에 따라 사용하여 단리하고, RNP 복합체를 이용한 전기천공 전에 2일 동안 확장시켰다.

[1692] Cas9와 가이드 RNA 리보핵단백질(RNP) 복합체의 조립, HSPC의 제조, 및 HSPC 내로의 RNP의 전기천공. 단일 가이드 RNA(sgRNA)를 사용한 RNP의 형성을 위해, 12 μg의 각각의 sgRNA, 12 μg의 CAS9 단백질 및 1 μL의 10 x CCE 완충제(20 mM HEPES, 100 mM KCL, 5 mM MgCL2, 5% 글리세롤 및 신선하게 첨가된 1 mM DTT)를 10 ul의 총 부피에서 조합하고, 그 후에 37℃에서 5분 동안 인큐베이션하였다. 없음/대조군 조건의 경우, sgRNA보다는 비히클을 첨가하였다. P3 완충제 +보충제에서의 세포 밀도는 1.3 x 10^6/mL이었다. 7개 복제 전기천공을 GCR-0067을 사용하여 수행하고, 전기천공-후에 세포를 조합하였다. 없음/대조군 조건의 경우, 전기천공 또는 Cas9의 첨가 없이, 세포를 P3 완충제로부터 확장 배지로 직접 옮겼다. RNP 복합체를 이용한 전기천공 후, 유세포분석 세포 소팅 전에 부가적인 3일 동안 세포를 확장시켰다.

[1693] 조혈 줄기 및 전구체 하위집단에서 편집 효율의 분석을 위한 유세포분석 세포 소팅. 전기천공-후 3일째에 편집된 세포 배양물을 수합하고, 2% FBS(Omega Scientific, Cat. #FB-11) 및 2mM EDTA(Corning Cat. #46-034-CL)가 보충된 HBSS(GE Life Sciences, Cat. #SH30588.01)로 구성된 FACS 염색 완충제 중 항-CD34(BD Biosciences, Cat# 348057), 항-CD38(BD Biosciences, Cat# 560677), 항-CD90(BD Biosciences, Cat# 559869), 항-CD45RA(BD Biosciences, Cat# 563963), 항-CD49f(BD Biosciences, Cat# 562598)와 함께 인큐베이션하였다. 세포를 FACS 염색 완충제로 세척하고, DAPI(4',6-디아미디노-2-페닐인돌)의 첨가에 의해 세포 생존력을 결정하였다. 다색 FACS 분석을 FACS Aria 세포 소터(BD Biosciences) 상에서 수행하였다. 음성과 양성 세포 집단 사이의 구별을, 각각의 항체를 개별적으로 이소형 및 형광크롬-공액 매칭된 비-특이적 대조군 항체(BD Biosciences Cat# 550854, 554680, 563437, 557872 및 562602)로 대체한 대조군 염색된 세포 배양물을 사용하여 설정된 게이트에 의해 결정하였다. 소팅된 세포의 순도를 소트-후 순도 체크에 의해 확인하였다.

[1694] 게놈 DNA 제조 및 차세대 시퀀싱(NGS). 게놈 DNA를 비편집된 HSPC로부터 제조하고, 전기천공-후 3일째에 편집된 HSPC의 하위집단을 DNeasy 혈액 & 조직 키트(Qiagen Cat# 69504)를 사용하여 소팅하였다. 하기 기재된 NGS 분석은 대조군 비편집된 시료에서 어떠한 유의한 편집도 보여주지 않았다.

[1695] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱을 실시예 2.1에서 기재된 바와 같이 수행하였다. NGS 시퀀싱 데이터 QC 및 변이체 분석을 실시예 2.1에 기재된 바와 같이 수행하였다.

[1696] 대형 결실의 정량화. HBG1 및 HBG2 프로모터에서 표적화 부위 사이의 대형 결디지털 방울 PCR(ddPCR)에 의해 제조업체의 권고사항에 따라 복사수 결정을 위한 비-경쟁적 검정법 포맷에서 정량화하였다. 간략하게는, gDNA를 프로브용 ddPCR SuperMix(dUTP 없음)(BioRad Cat# 1863024), HindIII-HF 제한 효소(NEB Cat# R3104S) 및 각각의 프라이머 프로브 믹스와 조합하고, 프로브용 방울 발생 오일(BioRad Cat# 1863005)과 함께 DG8 카트리지로 옮기고, QX200 방울 발생기(BioRad)를 이용하여 방울을 발생시켰다. 방울을 96-딥 웰 반응 모듈이 있는 C1000 Touch Thermal Cylcer(BioRad) 상에서 PCR을 받게 하고, 뒤이어 QX200 방울 리더(BioRad)를 이용하여 검출하였다. 1 ul 당 복사물를 QuantaSoft 소프트웨어(BioRad)를 이용한 분석에 의해 결정하였다. 커스텀 프라이머 프로브 세트(Life Technologies Cat# APZW76R, PN4331348, 포워드 프라이머: ACGGATAAGTAGATATTGAGGTAAGC(SEQ ID NO: 270), 리버스 프라이머: GTCTCTTTCAGTTAGCAGTGG(SEQ ID NO: 271), FAM TaqMan 프로브: ACTGCGCTGAAACTGTGGCTTTATAG(SEQ ID NO: 272))를 사용하여, HBG1-HBG2 유전자간 영역 내에서 gDNA를 증폭시켰다. TaqMan 복사수 참조 검정법, 인간, RNase P (Thermo Fisher cat# 4403326)를 참조 앰플리콘으로서 사용하였다. HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물은 제조업체의 보고된 선형 범위 내에 있었다. 퍼센트 결실을 100% x (1 - HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물의 비)로서 보고하였다. 비편집된 대조군 시료는 13% 이하의 계산된 퍼센트 결실을 가졌으며, 이는 검정법에서 백그라운드를 반영할 수 있다.

[1697] 결과:

[1698] 이론으로 결부시키고자 하는 것은 아니지만, CD34+ HSPC 집단은 생착, 자가-재생 및 세포 운명에 대한 다양한 잠재성의 세포를 함유하는 것으로 생각되며, 이때 장기간 조혈 줄기세포를 생착시키는 것을 포함하여 공유된 특성을 갖는 세포에서 부가적인 마커를 추가로 농화시킨다(문헌[Notta, *Science*, 2011 Jul 8;333(6039):218-21. doi: 10.1126/science.1201219; Huntsman, 혈액. 2015 Sep 24;126(13):1631-3. doi: 10.1182/혈액-2015-07-

660670]; 각각은 그 전체가 참조로 본 명세서에 포함됨). 이론으로 결부시키고자 하는 것은 아니지만, 이러한 하위집단은 이식-후 상이한 시기에 조혈 시스템을 재구성할 수 있으므로 유전자 편집된 세포 이식에서 특정한 치료적 이득일 수 있다. 따라서, 세포를 유세포분석 받게 하여, HSPC 하위집단에서 편집 효율을 특징화하였다. 5개 세포 집단을 단리하였으며, 첫번째는 CD34+ 세포이었으며, 뒤이어 제시된 전략을 사용하여 CD34+CD45RA-CD38+, CD34+CD45RA-CD38-CD90-CD49f+, CD34+CD45RA-CD38-CD90+CD49f+ 및 CD34+CD45RA-CD38-CD90-CD49f- 세포에 대해 4-방식 소팅이었다(도 11). 이들 소팅된 집단을 각각, HBG1 및 HBG2 표적 부위 국소화된 영역을 개별적으로 시퀀싱함으로써, 뿐만 아니라 개입 HBG1-HBG2 영역의 결실의 검출에 의해 편집에 대해 분석하였다. HBG1-HBG2 결실(즉, 절제) 빈도는 하위집단에 걸쳐 69 내지 81 퍼센트 범위로 일반적으로 높았다(도 12). 개별적으로 HBG1 및 HBG2 프로모터에서의 국소화된 증폭 및 NGS를 사용하여, 각각의 부위에서 편집 %(소형 삽입결실)를 결정하였다. 두 프로모터 모두에서 표적화함으로써 유발된 역위 또는 결실을 갖는 대립유전자는 검정법에서 증폭되지 않을 것이다. HBG1-HBG2 결실 또는 역위가 없는 대립유전자 중에서, 편집은 하위집단에 걸쳐, HBG1에 대해 53 내지 73 퍼센트 및 HBG2에 대해 75 내지 91 퍼센트의 범위로 유사하게 높았다. HBG1-HBG2 결실된 게놈의 퍼센트를 HBG2에서 편집을 갖는 비-HBG1-HBG2 결실된 게놈의 퍼센트와 조합함으로써 최소 총 편집을 추정하였다(최소 총 편집 = % 결실 + % 비결실[100 - % 역위를 무시한 결실] x HBG2에서 % 편집). HBG1에서의 국소화된 소형 삽입결실은 포함시키지 않았는데, 본 발명자들이 HBG1에서의 % 편집 중 어느 정도가 HBG1 단독으로 국소화된 삽입결실을 갖는 대립유전자인지, 어느 정도가 HBG1과 HBG2 둘 모두에서 공존 삽입결실을 갖는 대립유전자인지 결정할 수 없었기 때문이다. 최소 총 편집은 하위집단에 걸쳐 92 내지 97 퍼센트 범위인 것으로 추정되었다. HBG1 및 HBG2 표적 부위로 국소화된 삽입결실에 대한 편집 패턴은 또한, 하위집단에 걸쳐 유사하였으며, 이때 가장 빈번하게 관찰된 편집 패턴은 하위집단에 걸쳐 보편적이었다. 모든 하위집단 및 각각의 부위 HBG1 또는 HBG2에 대해, 상위 3개 삽입결실 패턴은 6 bp 결실, 5 bp 결실 및 1 bp 결실을 포함하였다(도 13a(HBG1 유전자좌); 도 13b(HBG2 유전자좌)). 요약하자면, 총 HSPC와 정의된 CD34+CD45RA-CD38-CD90+CD49f+ 장기간 조혈 줄기세포를 포함하여 하위분획 사이에서 일관된 편집 빈도 및 패턴은 HSPC 이식에 대해 본원에서 기재된 유전자 편집 접근법의 유용성을 뒷받침한다.

[1699] **실시예 2.5:** 감마 글로빈 프로모터 영역 편집 후 HSPC의 콜로니 형성 능력

[1700] 방법:

[1701] 방법은 실시예 2.1에서와 같고, 다음의 변형을 갖는다.

[1702] 인간 $CD34^+$ 세포 배양물. 인간 CD34+ 세포는 성인 건강한 공여자 골수(Lonza 카탈로그 # 2M-101D)로부터 유래되었다. 세포를 해동시키고, 그 후에 RNP 복합체를 이용한 전기천공 전에 2일 동안 확장시켰다.

[1703] Cas9와 가이드 RNA 리보핵단백질(RNP) 복합체의 조립, HSPC의 제조, 및 HSPC 내로의 RNP의 전기천공. 단일 가이드 RNA(sgRNA)를 사용한 RNP의 형성을 위해, 12 μg의 각각의 sgRNA, 12 μg의 CAS9 단백질 및 1 μL의 10 x CCE 완충제(20 mM HEPES, 100 mM KCL, 5 mM MgCL2, 5% 글리세롤 및 신선하게 첨가된 1 mM DTT)를 10 ul의 총 부피에서 조합하고, 그 후에 37℃에서 5분 동안 인큐베이션하였다. P3 완충제 +보충제에서의 세포 밀도는 6.64 x 10^6/mL이었다. 2개의 전기천공 복제물을 2명의 독립적인 공여자 각각으로부터의 세포를 이용하여 수행하였다. 없음/대조군 조건의 경우, sgRNA보다는 비히클을 첨가하였다. RNP 복합체를 이용한 전기천공 후, 세포를 확장 배지로 복귀시켰다.

[1704] 게놈 DNA 제조 및 차세대 시퀀싱(NGS). 게놈 DNA를 비편집된 HSPC로부터 제조하고, 전기천공-후 3일째에 편집된 HSPC의 하위집단을 DNeasy 혈액 & 조직 키트(Qiagen Cat# 69504)를 사용하여 소팅하였다. 하기 기재된 NGS 분석은, Cas9 단독으로 전기천공된 대조군 시료에서 어떠한 유의한 편집도 보여주지 않았다.

[1705] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱을 실시예 2.1에서 기재된 바와 같이 수행하였다. NGS 시퀀싱 데이터 QC 및 변이체 분석을 실시예 2.1에 기재된 바와 같이 수행하였다.

[1706] 대형 결실의 정량화. HBG1 및 HBG2 프로모터에서 표적화 부위 사이의 대형 결디지털 방울 PCR(ddPCR)에 의해 제조업체의 권고사항에 따라 복사수 결정을 위한 비-경쟁적 검정법 포맷에서 정량화하였다. 간략하게는, gDNA를 프로브용 ddPCR SuperMix(dUTP 없음)(BioRad Cat# 1863024), HindIII-HF 제한 효소(NEB Cat# R3104S) 및 각각의 프라이머 프로브 믹스와 조합하고, 프로브용 방울 발생 오일(BioRad Cat# 1863005)과 함께 DG8 카트리지로 옮기고, QX200 방울 발생기(BioRad)를 이용하여 방울을 발생시켰다. 방울을 96-딥 웰 반응 모듈이 있는 C1000 Touch Thermal Cylcer(BioRad) 상에서 PCR을 받게 하고, 뒤이어 QX200 방울 리더(BioRad)를 이용하여 검출하였다. 1 ul 당 복사물를 QuantaSoft 소프트웨어(BioRad)를 이용한 분석에 의해 결정하였다. 커스텀 프라이머 프로

브 세트(Life Technologies Cat# APZW76R, PN4331348, 포워드 프라이머: ACGGATAAGTAGATATTGAGGTAAGC(SEQ ID NO: 273), 리버스 프라이머: GTCTCTTTCAGTTAGCAGTGG(SEQ ID NO: 274), FAM TaqMan 프로브: ACTGCGCTGAAACTGTGGCTTTATAG(SEQ ID NO: 275))를 사용하여, HBG1-HBG2 유전자간 영역 내에서 gDNA를 증폭시켰다. TaqMan 복사수 참조 검정법, 인간, RNase P (Thermo Fisher cat# 4403326)를 참조 앰플리콘으로서 사용하였다. HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물은 제조업체의 보고된 선형 범위 내에 있었다. 퍼센트 결실을 100% x (1 - HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물의 비)로서 보고하였다. 비편집된 대조군 시료는 9% 이하의 계산된 퍼센트 결실을 가졌으며, 이는 검정법에서 백그라운드를 반영할 수 있다.

[1707] 콜로니 형성 단위 세포 검정법. RNP 전달 후 2일째에, 살아 있는 세포를 LSRFortessa(BD Biosciences) 상에서 TruCount 튜브(BD Biosciences Cat# 340334)를 제조업체의 권고사항에 따라 사용한 유세포분석에 의해 나열하였다. 살아 있지 않은 세포를 DAPI(4',6-디아미디노-2-페닐인돌)를 사용하여 구별하였다. 분석을 FlowJo 소프트웨어(Tree Star)를 사용하여 수행하였다. 콜로니 형성 단위(CFU) 검정법을 위해, 세포 및 1x 항생제/항진균제(Gibco, Cat. #10378-016)를 MethoCult H4034 Optimum(Stemcell Technologies) 메틸셀룰로스 배지(Stemcell Technologies)에 첨가하고, 1 mL를 SmartDish 플레이트(Stemcell Technologies)에서 3벌 중복으로 평판배양하였다. 배양 디쉬를 37℃에서 습도화된 인큐베이터에서 인큐베이션하였다. 배양물을 평판배양-후 14일차에 StemVision(Stemcell Technologies)을 사용하여 이미지화하였다. 콜로니를 Colony 마커 소프트웨어(Stemcell Technologies)를 사용하여 수동으로 점수를 매겼다. 1개 웰 당 콜로니 수(평균 3개 웰)를 Methocult 1 ml 당 평판배양된 세포의 수(226 내지 318의 범위였음)로 나누고, 1000으로 곱하여, 1000개 세포 당 CFU 빈도를 수득하였다.

[1708] 결과:

[1709] 본 발명자들은 본원에서, HSPC가 F-세포의 생산과 연관된 HBG1 및 HBG2 프로모터 영역 내에서 특이적인 서열의 표적화된 교란 후 콜로니 형성 검정법에서 작용함을 보여주었다. 게놈 편집된 및 비편집된 HSPC를, 콜로니 형성 단위 검정법을 사용하여 전구 세포 조성물 및 분화 잠재성에 대해 평가하였다. 콜로니를 계수하고, 적혈구계 전구 세포(CFU-적혈구계[CFU-E] 및 버스트-형성 단위(burst-forming unit)-적혈구계[BFU-E]), 과립구 및/또는 대식세포 전구 세포(CFU-과립구, 대식세포[CFU-GM]; CFU-과립구[CFU-G]; 및 CFU-대식세포[CFU-M]), 도는 다중-잠재적인 전구 세포(CFU-과립구, 적혈구, 대식세포, 거핵구[CFU-GEMM])로부터 유래하는 것으로 분류하였다.

[1710] [표 8]

[1711] 제1 독립적인 공여자로부터의 세포를 편집하기 위해 본 연구에 사용된 HBG1 및 HBG2 프로모터 영역을 표적화하는 선택 gRNA의 목록.

가이드 RNA 표적화 도메인 ID	% 편집된 HBG1	% 편집된 HBG2	% HBG1-HBG2 결실	% 대략적인 최소 편집	1000 개 세포 당 BFU-E/CFU-E	1000 개 세포 당 CFU-G/M/GM	1000 개 세포 당 CFU-GEMM	1000 개 세포 당 총 콜로니
없음/대조군	n/a	n/a	8.6	n/a	48	162	3	213
GCR-0008	80.7	78.9	42.6	88.9	18	134	2	154
GCR-0010	34.9	18.0	29.5	54.1	25	118	4	148
GCR-0048	81.4	84.0	40.6	90.5	28	130	0	158
GCR-0051	51.0	74.9	41.2	85.2	28	133	2	164
GCR-0067	82.9	88.6	51.3	94.4	25	123	5	153

[1712]

[1713] [표 9]

[1714] 제2 독립적인 공여자로부터의 세포를 편집하기 위해 본 연구에 사용된 HBG1 및 HBG2 프로모터 영역을 표적화하는 선택 gRNA의 목록.

가이드 RNA 표적화 도메인 ID	% 편집된 HBG1	% 편집된 HBG2	% HBG1-HBG2 결실	% 대략적인 최소 편집	1000 개 세포 당 BFU-E/CFU-E	1000 개 세포 당 CFU-G/M/GM	1000 개 세포 당 CFU-GEMM	1000 개 세포 당 총 콜로니
없음/대조군	n/a	n/a	8.9	n/a	90	241	12	343
GCR-0008	82.3	85.2	47.0	92.1	51	151	4	206
GCR-0010	42.3	35.1	43.0	67.1	56	162	5	223
GCR-0048	68.7	76.1	41.7	86.0	49	129	7	186
GCR-0051	55.2	82.7	41.3	89.9	32	157	6	196
GCR-0067	81.2	88.9	47.1	94.1	39	169	9	217

[1715]

[1716] 지시된 gRNA와 함께 RNP를 이용하여 전기천공된 두 공여자 모두는 효율적인 편집, HBG1 및 HBG2 프로모터 표적/표적-외 부위로 국소화된 두 삽입결실 모두, 뿐만 아니라 개입 영역의 결실을 가졌다(표 8 및 표 9). 편집된 배양물은 전반적인 콜로니 형성 능력에서의 하강과 연관이 있었으며(표 8 및 표 9), 이는 가능하게는 편집을 겪은 세포의 저하된 적응도(fitness)를 표시한다. 총 콜로니 수에서의 이러한 감소에도 불구하고, 적혈구계, 과립구/대식세포 및 다중-잠재적인 콜로니는 모두 적어도 하나의 공여자에서 관찰되었다(표 8 및 표 9). 더욱이, 비편집된 시료와 편집된 시료 사이에서 콜로니 유형의 비율에서 최소 차이가 존재하였으며(도 14), 이는 이들 표적 부위에서 편집된 세포 배양물은 왜곡된(skewed) 분화 능력을 갖지 않았음을 표시한다.

[1717] **실시예 2.6:** 감마 글로빈 프로모터 영역 편집 후 시험관내에서 HSPC의 세포 증식

[1718] 방법:

[1719] 방법은 실시예 2.1에서와 같고, 다음의 변형을 갖는다.

[1720] 인간 $CD34^+$ 세포 배양물. 인간 CD34+ 세포를 성인 건강한 공여자(Lonza 카탈로그 # 2M-101D 및 Hemacare 카탈로그 # M001F-GCSF-3)의 골수로부터 유래하였다. 그 후에, 세포를 RNP 복합체를 이용한 전기천공 전에 2일 동안 확장시켰다.

[1721] Cas9와 가이드 RNA 리보핵단백질(RNP) 복합체의 조립, HSPC의 제조, 및 HSPC 내로의 RNP의 전기천공. 단일 가이드 RNA(sgRNA)를 사용한 RNP의 형성을 위해, 12 μg의 각각의 sgRNA, 12 μg의 CAS9 단백질 및 1 μL의 10 x CCE 완충제(20 mM HEPES, 100 mM KCL, 5 mM MgCL2, 5% 글리세롤 및 신선하게 첨가된 1 mM DTT)를 10 ul의 총 부피에서 조합하고, 그 후에 37℃에서 5분 동안 인큐베이션하였다. P3 완충제 + 보충제에서의 세포 밀도는 1 x 10^7 내지 2.5 x 10^7/mL이었다. 1개의 전기천공 복제물을 3명의 독립적인 공여자 각각으로부터의 세포를 이용하여 수행하였다. 없음/대조군 조건의 경우, sgRNA보다는 비히클을 첨가하였다. RNP 복합체를 이용한 전기천공 후, 세포를 확장 배지로 복귀시켰다.

[1722] 게놈 DNA 제조 및 차세대 시퀀싱(NGS). 전기천공-후 3일째에 Quick Extract DNA 추출 용액(Epicentre Cat # QE09050) 또는 DNeasy 혈액 & 조직 키트(Qiagen Cat# 69504)을 사용하여 편집된 HSPC 및 비편집된 HSPC로부터 게놈 DNA를 제조하였다. 하기 기재된 NGS 분석은, Cas9 단독으로 전기천공된 대조군 시료에서 어떠한 유의한 편집도 보여주지 않았다.

[1723] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱을 실시예 2.1에서 기재된 바와 같이 수행하였다. NGS 시퀀싱 데이터 QC 및 변이체 분석을 실시예 2.1에 기재된 바와 같이 수행하였다.

[1724] 세포 증식 및 표현형검사(phenotyping). RNP 전달 후 2일째에, 살아 있는 세포를 LSRFortessa(BD Biosciences) 상에서 TruCount 튜브(BD Biosciences Cat# 340334)를 제조업체의 권고사항에 따라 사용한 유세포분석에 의해 나열하였다. 살아 있지 않은 세포를 DAPI(4',6-디아미디노-2-페닐인돌)를 사용하여 구별하고, CD34+ 및 CD34+CD90+ 세포를 결정하였다. 2% FBS(Omega Scientific, Cat. #FB-11) 및 2mM EDTA(Corning Cat. #46-034-CL)가 보충된 HBSS(GE Life Sciences, Cat. #SH30588.01)로 구성된 FACS 염색 완충제 중 항-CD34(BD Biosciences Cat# 348057, BD Biosciences Cat# 340666 또는 eBioscience Cat# 25-0349-425) 및 항-CD90(BD Biosciences Cat# 559869)의 포함에 의해 함량을 결정하였다. FlowJo 소프트웨어(Tree Star)를 사용하여 분석을 수행하였다. 그 후에, 세포를 2.0 x 10^4 또는 1.0 x 10^5개 살아 있는 세포/ml로 확장 배지 내로 접종하고, 7일 동안 배양하였다. 1.0 x 10^5/ml로 접종된 배양물에 대해, 배양 기간 동안 3 또는 4배 희석을 위해 신선한 배지를 첨가하였다. 7일 배양 후, 세포를 일단 상기와 같이 다시 나열하였다. 7일 배양 후, 세포를 부가적으로, 하기 항체 패널을 이용한 염색 후 표면 마커 발현에 대해 분석하였다. 패널 1: CD38(FITC-공액체, BD Biosciences #340926, 클론 HB7), CD133 에피토프 1(PE-공액체, Miltenyi #130-080-801, 클론 AC133), CD34(PerCP-공액체, BD Biosciences #340666, 클론 8G12), CD90(APC-공액체, BD Biosciences # 559869, 클론 5E10), CD45RA(Pe-Cy7-공액체, eBioscience #25-0458-42, 클론 HI100). 패널 2: CD34(PerCP-공액체, BD Biosciences #340666, 클론 8G12), CD33(PE-Cy7-공액체, BD Biosciences #333946, 클론 P67.6), CD14(APC-H7-공액체, BD Biosciences #560270, 클론 MΦP9), CD15(PE-공액체, Biolegend #301905, 클론 HI98). 패널 3: CD34(PerCP-공액체, BD Biosciences #340666, 클론 8G12), CD41a(APC-H7-공액체, BD Biosciences #561422, 클론 HIP8), CD71(FITC-공액체, BD Biosciences #555536, 클론 M-A712), CD19(PE-공액체, BD Biosciences #340720, 클론 SJ25C1), CD56(APC-공액체, Biolegend #318310, 클론 HCD56)에 특이적인 항체. 상응하는 이소형 대조군 항체 패널을 사용하여, 염색물을 병행하여 염색시켰다. 살아 있지 않은 세포를 DAPI(4',6-디아미디노-2-페닐인돌) 염색에 의해 구별하였다. 염색된 시료를 LSRFortessa 유세포분석기(BD Biosciences) 상에서 세포 표면 단백질 발현에 대해 분석하였다. 결과를 Flowjo를 사용하여 분석하고, 데이터를 DAPI 음성 살아 있는 세포 집단의 %로서 제시하였다.

[1725] 결과:

[1726] 본 발명자들은 본원에서, HSPC가 F-세포의 생산과 연관된 HBG1 및 HBG2 프로모터 영역 내에서 특이적인 서열의 표적화된 교란 후 시험관내에서 전형적인 세포성 조성물과 함께 확장될 수 있음을 보여주었다. 게놈 편집된 및 비편집된 HSPC를, HSPC의 확장을 촉진하는 배양 조건에서 증식 능력 및 세포 조성에 대해 평가하였다. GCR-0067을 함유하는 RNP를 이용하여 전기천공된 복제물 독립적인 공여자는 효율적인 편집을 가졌다(표 9-2).

[1727] [표 9-2]

[1728] 지시된 gRNA를 사용하여 3명의 독립적인 공여자로부터의 세포에서 HBG1 또는 HBG2 프로모터 영역의 편집된 앰플리콘 퍼센트.

가이드 RNA 표적화 도메인 ID	% 편집된 HBG1, 공여자 A	% 편집된 HBG2, 공여자 A	% 편집된 HBG1, 공여자 B	% 편집된 HBG2, 공여자 B	% 편집된 HBG1, 공여자 C	% 편집된 HBG2, 공여자 C
GCR-0067	62.7	64.5	57.0	73.4	53.1	56.9

[1729]

[1730] 편집된 배양물은 전반적인 증식 능력에서의 하강과 연관이 있었으며, 이는 가능하게는 편집을 겪은 세포의 약간 저하된 적응도를 가리키지만, 이는 3명의 독립적인 세포 공여자에 걸쳐 유의성에 도달하지 않았다(도 15). 유사한 감소는 총 CD34+ 집단에서와 같이 조혈 줄기세포 농화된 CD34+CD90+ 집단 내에서 관찰되었으며, 이는 이러한 집단이 상이하게 효과를 받지 않음을 표시한다(도 15). 이들 배양 조건 하에, HPSC와 분화된 자손 둘 모두가 존재하는 것으로 예상되며, 따라서, 세포성 조성물을 세포 표면 마커의 보다 포괄적인 패널의 발현에 대해 추가로 분석하였다. 이들 세포 집단의 구별을 도 16에 예시한다. 세포성 조성물은 3명의 독립적인 세포 공여자에 걸쳐 게놈 편집된 및 비편집된 배양물 사이에서 유사하였으며, 이때, 독립 표본 t-검정에 의해 주어진 집단에서 유의한 차이는 존재하지 않았다(도 17). CD33+ 집단에 대한 대형 오차 막대는 게놈 편집된 및 비편집된 배양물 둘

모두에서 무시할 만한 CD33+ 세포를 갖는 단일 공여자로 인한 것이다(도 17).

[1731] **실시예 3: Cas9 변이체의 평가**

[1732] CD34+ 조혈 줄기 세포에서의 평가

[1733] 일차 인간 조혈 줄기 세포(HSC)에서 베타-2-마이크로글로불린(B2M) 유전자를 넉아웃하는 효율을 측정함으로써 14개의 정제된 스트렙토코커스 피오게네스 Cas9(SPyCas9) 단백질을 평가하였다. 이들 단백질을 3개의 그룹으로 나누었다: 제1 그룹은 개선된 선택성을 갖는 SPyCas9 변이체로 구성되었다(문헌[Slaymaker et al. 2015, Science 351: 84(e1.0, e1.1 and K855A); Kleinstiver et al. 2016, Nature 529: 490(HF)]). 제2 그룹은 절단 가능한 TEV 부위가 있거나 없는 상이한 수 및/또는 위치의 SV40 핵 국재화 신호(NLS) 및 6x히스티딘(His6) 태그(SEQ ID NO: 247) 또는 8x히스티딘(His8) 태그(SEQ ID NO: 248)를 갖는 야생형 SPyCas9, 및 구조 연구를 위해 Cas9를 안정화시키는 것으로 보고된(문헌[Nishimasu et al. 2014, Cell 156:935]) 2개의 시스테인 치환(C80L, C574E)을 갖는 SpyCas9 단백질로 구성되었다. 제3 그룹은 상이한 공정에 의해 생산된 동일한 재조합 SPyCas9로 구성되었다(도 6). B2M 넉아웃은 FACS 및 차세대 시퀀싱(NGS)으로 결정하였다.

[1734] 방법

[1735] 재료

[1736] 1. 네온 전기천공기(Invitrogen, MPK5000)

[1737] 2. 네온 전기천공 키트(Invitrogen, MPK1025)

[1738] 3. crRNA(SEQ ID NO: 201에 융합된, B2M 유전자 내의 서열에 상보적인, GGCCACGGAGCGAGACAUCU의 표적화 도메인 서열(SEQ ID NO: 276))

[1739] 4. tracrRNA(SEQ ID NO: 224)

[1740] 5. Cas9 저장 완충제: 20 mM Tris-Cl, pH 8.0, 200 mM KCl, 10 mM $MgCl_2$

[1741] 6. 골수 유래 CD34+ HSC(Lonza, 2M-101C)

[1742] 7. 세포 배양 배지(Stemcell Technologies, StemSpam CC-100있는 StemSpam SFEM II)

[1743] 8. FACS 세척 완충제: PBS 중의 2% FCS

[1744] 9. FACS 블록 완충제: PBS mL 당, 0.5 ug 마우스 IgG, 150 ug Fc 블록, 20 uL FCS 첨가

[1745] 10. Chelex 현탁액: H_2O 중 10% Chelex 100(bioRad, Cat# 142-1253)

[1746] 11. 항-B2M 항체: Biolegend, cat# 316304

[1747] 과정

[1748] Lonza의 권장 사항에 따라 세포를 해동 및 성장시키고, 2 내지 3일마다 배지를 추가한다. 5일차에, 15분 동안 200 xg에서 세포를 펠렛화하고, PBS로 한 번 세척하고, NEON 키트로부터의 T-완충제로 세포를 2 x 10^4 /uL로 재현탁하고, 얼음 위에 놓는다. Cas9 저장 완충제로 Cas9 단백질을 5 mg/ml로 희석한다. H_2O로 crRNA 및 tracrRNA를 100 uM로 재구성한다. 리보핵 단백질(RNP) 복합체를 CAS9 단백질, crRNA 및 tracrRNA 각각 0.8 uL을 0.6 uL의 Cas9 저장 완충제와 혼합함으로써 제조하고, 실온에서 10분 동안 인큐베이션한다. 2분 동안 RNP 복합체와 7 uL의 HSC를 혼합하고, 네온 파이펫 팁 내로 전체 10 uL를 옮기고, 1700 v, 20 ms 및 1 펄스에서 전기천공한다. 전기천공 후, 세포를 37℃, 5% CO_2에서 사전 보정된 1 ml 배지를 함유하는 24-웰 플레이트의 웰로 즉시 옮긴다. 전기천공 후 72시간에 FACS 및 NGS 분석을 위해 세포를 수확한다.

[1749] FACS: 250 uL의 세포를 24-웰 플레이트의 각 웰로부터 96-웰 U-바닥 플레이트의 웰로 이동시키고 세포를 펠렛화한다. 2% FCS(태아 송아지 혈청)-PBS로 한 번 세척한다. 50 uL FACS 블록 완충제를 세포에 첨가하고 10분 동안 얼음 위에서 인큐베이션하고, 1 uL FITC 표지된 B2M 항체를 첨가하고 30분 동안 인큐베이션한다. 150 uL FACS 세척 완충제로 한 번 세척한 후 재차 200 uL FACS 세척 완충제로 한 번 세척한다. 세포를 200 uL FACS 완충제 FACS 분석에 재현탁시켰다.

[1750] NGS 시료 준비: 250 uL의 세포 현탁액을 24-웰 플레이트의 각 웰로부터 1.5 ml Eppendorf 튜브로 옮기고, 1 mL

PBS를 첨가하고 세포를 펠렛화한다. Chelex 현탁액 100 uL를 첨가하고 99℃에서 8분 동안 인큐베이션하고 10초 볼텍싱 후 99℃에서 8분 동안 인큐베이션, 10초 볼텍싱한다. 10,000 xg에서 3분간 원심 분리하여 수지를 펠렛화하여 침전시키고 상등액 용해물을 PCR에 사용한다. 4 uL 용해물을 취하여 Titanium 키트(Clontech, cat# 639208)를 사용하여 b2m 프라이머(b2mg67F: CAGACAGCAAACTCACCCAGT(SEQ ID NO: 277), b2mg67R: CTGACGCTTATCGACGCCCT(SEQ ID NO: 278))로 PCR 반응를 수행하고 제조사의 지시를 따른다. 다음의 PCR 조건을 사용한다: 1 사이클 동안 98℃에서 5분; 30 사이클 동안 95℃에서 15초, 62℃에서 15초, 및 72℃에서 1분; 마지막으로 1 사이클 동안 72℃에서 3분. PCR 산물을 NGS에 사용하였다.

[1751] NGS 라이브러리 제조 및 앰플리콘의 시퀀싱을 실시예 2.1에 기재된 바와 같이 수행하였다. NGS 시퀀싱 데이터 QC 및 변이체 분석을 실시예 2.1에 기재된 바와 같이 수행하였다.

[1752] 통계: FACS에 의한 B2M KO 세포의 백분율 및 NGS에 의한 삽입결실의 백분율을 사용하여 CAS9 절단 효율을 평가하였다. Cas9를 고정 효과로 하여 실험을 설계하였다. 각 실험은 내포된(nested) 임의 효과로서 공여자 내에 내포되어 있다. 따라서 혼합 선형 모델을 FACS 및 NGS 데이터 분석법에 적용하였다.

[1753] 결과

[1754] 실험적 및 공여자 변이를 정규화하기 위하여, 야생형 SPyCas9 양 옆에 배치된 2개의 SV40 NLS 및 단백질의 C-말단의 His6 태그(SEQ ID NO: 247)를 갖는 원래의 설계인, iProt105026에 대한 각 단백질의 상대적 활성을 그래프로 나타냈다(도 6). 통계 분석법은, 참조 Cas9 단백질 iProt105026와 비교하여, iProt106331, iProt106518, iProt106520 및 iProt106521은 HSC에서 B2M을 넉아웃하는 것에 있어서 유의미하게 상이하지 않은 반면, 테스트된 다른 변이체들(PID426303, iProt106519, iProt106522, iProt106545, iProt106658, iProt106745, iProt106746, iProt106747, iProt106884)은 HSC에서 B2M을 넉아웃하는 것에 있어서 참조 iProt105026와 매우 유의미하게 상이하다는 것을 나타낸다. 본 발명자들은 His6 태그(SEQ ID NO: 247)를 C-말단으로부터 N-말단으로 이동시키는 것(iProt106520)이 단백질의 활성에 영향을 미치지 않는다는 것을 확인하였다(도 6). 하나의 NLS는 단백질의 C 말단에 위치했을 때에만 활성을 유지하기에 충분하였다(iProt106521 대 iProt106522, 도 6). 공정 1에서 정제된 단백질은 공정 2 및 3보다 일관되게 더 높은 넉아웃 효율을 가졌다(iProt106331 대 iProt106545 및 PID426303, 도 6). 일반적으로, 보고된 개선된 선택성을 갖는 SPyCas9 변이체는 야생형 SPyCas9만큼 활성적이지 않았다(iProt106745, iProt106746 및 iProt106747, 도 6). 흥미롭게도 iProt106884는 표적화 부위를 절단하지 않았다. 이는 이 변이체가 포유류 세포에서 적당한 표적화 부위의 20%까지 절단하지 못하였다는 Kleinstiver 등에 의한 보고와 일치한다(문헌Kleinstiver et al. 2016, Nature 529: 490]). 최종적으로, 2개의 시스테인 치환(iProt106518)을 갖는 Cas9 변이체는 높은 수준의 효소 활성을 유지하였다(도 6).

[1755] **실시예 4. 모세관 전기영동 질량 분광법에 의한 유전자-편집 시 헤모글로빈 하위단위 변화의 정량화**

[1756] HSC의 유전자-편집 시, 유세포분석에 의한 태아 헤모글로빈 생산의 포착 외에도, 적혈구계 세포에서 개별 헤모글로빈 하위단위에서의 변화를 또한, 모세관 전기영동 질량 분광법(CE-MS)에 의해 측정하였다. 겸상 적혈구 환자의 말초 혈액으로부터 유래된 CD34+ 세포를 Cas9 단백질(2개의 NLS 및 1개의 His 태그를 가짐, iProt106331) 및 가이드 RNA의 존재 하에 모크 편집하거나, Cas9 단백질과 가이드 RNA sg0067 둘 모두의 존재 하에 유전자-편집하였고, 이전에 기재된 바와 같이 적혈구계 계통으로 분화시켰다. 적혈구계 분화 14일차째에, 각각의 조건으로부터의 동일한 수의 세포를 태아 헤모글로빈(HbF)을 인지하는 FITC_공액 항체로 염색시키고, 세포를 유세포분석 받게 하여, 유전자 편집 시 HbF 유도를 정량화하였다. 유세포분석은 HbF+ 세포의 30% 내지 50% 상향조절을 보고하였다(표 10). 병행하여, 1개 조건 당 다양한 양의 적혈구계 세포(6, 12, 66 및 99 x 10^6개)를 수확하고, CE-MS를 받게 하여 알파-글로빈, 정상 베타-글로빈(HbB), 겸상 베타-글로빈(HbS) 및 태아 감마-글로빈(HbF) 하위단위를 정량화하고, 투입 세포의 양을 이용하여 데이터를 정규화하였다. CE-MS는 600개 세포/ul만큼 낮은 농도에서 모든 글로빈 하위단위를 검출할 수 있었다. 결과는, 편집된 겸상 적혈구 환자 시료가 감마-글로빈에서 약 40% 증가 및 겸상 베타-글로빈에서 병발하는 50% 저하를 가졌음을 보여주었다(도 18).

[1757] [표 10]

[1758] 글로빈 하위단위의 CE-MS 정량화를 받은 시료의 처리 및 세포 양의 설명. 이들 시료의 태아 헤모글로빈 발현을 또한, 형광단에 공액된 항-HbF 항체를 사용한 유세포분석에 의해 측정하였다.

공여자 ID	시료 ID	시료 설명	세포의 수 (1 x 10e6)	유세포분석에 의해 측정된 HbF+ 세포(%)
겸상 적혈구 환자 1	1	모크 편집됨, sgRNA 없음, 적혈구 분화됨	99.2	55.5
	2	sg0067 을 이용하여 편집된 유전자, 적혈구 분화됨	66.6	81.15
겸상 적혈구 환자 2	3	모크 편집됨, sgRNA 없음, 적혈구 분화됨	12.6	59.70
	4	sg0067 을 이용하여 편집된 유전자, 적혈구 분화됨	5.5	81.85

[1759]

[1760] **실시예 5. 유전자-편집된 조혈 줄기세포(HSC)는 장기간 생착을 할 수 있었고 다계통 재구성을 뒷받침한다**

[1761] 유전자-편집된 HSC의 줄기세포 기능을 평가하기 위해, 편집된 인간 HSC를 면역약화된 마우스에게 이식하여, 장기간 조혈 재생을 검사하였다. $5x10^5$개 인간 골수-유래 CD34+ 세포를 0일차에 해동시키고, 3일차에 모크 RNP 복합체(Cas9 단독 및 gRNA 없음), 또는 Cas9 및 CR001128의 표적화 도메인을 포함하는 sgRNA(본원에서 이따금 sg1128로 지칭됨)와 함께 형성된 RNP 복합체를 이용하여 전기천공시켰다.

[1762] (AUCAGAGGCCAAACCCUUCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 181)의 표적화 도메인 서열을 포함함. 후속해서, 세포를 실시예 2.1에 기재된 동일한 농도 및 배지에서 유지시키고, 배양 6일차에, 각각의 처리로부터의 모든 세포를 2 그레이(Gray) 아치사 방사선 조사된 NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ (NSG) 마우스에게 이식하였다(도 19). 본 발명자들은 이식-후 16주째에 수여자의 골수에서 대략 30% 내지 40% 인간 세포 생착을 관찰하였다. 더욱이, 이러한 생착 수준은 모크 치료 대조군과 비교할 만하였으며, 이는 유전자-편집이 HSC의 장기간 생착 능력을 손상시키지 않음을 제안한다(도 20). 유전자-편집된, 이식된 HSC는 이식-후 4, 8, 12, 16주째에 정상 골수계, B 림프계 및 T 림프계 세포 재생을 지속시켰다(도 21). 이러한 데이터는, 유전자-편집된 HSC가 이식-후 16주까지를 포함하여 장기간 생착을 할 수 있었고 정상 다계통 재구성을 뒷받침함을 실증한다.

[1763] **실시예 6.**

[1764] **유전자-편집된, 장기간 생착된 HSC는 지속된 태아 헤모글로빈(HbF) 생산을 할 수 있었다**

[1765] 개별 연구에서, 본 발명자들은 BCL11A 유전자의 적혈구계-특이적 인핸서 영역으로부터의 gRNA(sg-G1128)와 비교하여 감마 글로빈 프로모터 영역으로부터의 gRNA(표 11에서 sg-G0008, sg-G0051, sg-G0010, sg-G0048, sg-G0067)를 이용한 HSC-유전자 편집물의 줄기세포 기능을 평가하였다. 이들 gRNA를 이용하여 편집된 인간 HSC를 면역약화된 NSG 마우스 내에 이식하여, 조혈 재생을 검사하였다(도 22).

[1766] [표 11]

[1767] 감마 글로빈 프로모터 영역을 표적화하기 위해 설계된, 이식 연구에 사용된 가이드 RNA.

sgRNA 명칭	표적 도메인 ID	crRNA ID	gRNA 표적 도메인 서열	sgRNA 100량체 서열
sg-G0008	GCR-0008	CR005821	GGAGAAGGAAACUAGCUAAA (SEQ ID NO: 8)	GGAGAAGGAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 84)
sg-G0010	GCR-0010	CR005823	GGGAGAAGGAAACUAGCUAA (SEQ ID NO: 10)	GGGAGAAGGAAACUAGCUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 94)
sg-G0048	GCR-0048	CR005811	ACGGCUGACAAAAGAAGUCC (SEQ ID NO: 48)	ACGGCUGACAAAAGAAGUCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 134)
sg-G0051	GCR-0051	CR005813	GGAGAAGAAACUAGCUAAA (SEQ ID NO: 51)	GGAGAAGAAAACUAGCUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 144)
sg-G0067	GCR-0067	CR005820	ACUGAAUCGGAACAAGGCAA (SEQ ID NO: 67)	ACUGAAUCGGAACAAGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 174)

[1768]

[1769] **실험 절차**

[1770] **CD34+ 해동 및 배양**

[1771] 골수 CD34+ 세포를 해동하고 배양하였다. 1 x 10^6개 골수 CD34+ 세포의 각각의 바이얼을 액화 질소로부터 제거하고, 70% 에탄올을 분무하고, 소형 얼음 펠렛이 남을 때까지 37℃ 수조에서 급속 해동시켰다. 바이얼에 70% 에탄올을 분무하고, 닦아내었다. 5 ml 파이펫을 사용하여, 바이얼의 내용물을 50 ml 팔콘 튜브로 옮겼다. 동결바이얼을 1 ml 예열된 IMDM, 10% FBS로 1회 헹구고, 50 ml 팔콘 튜브에 적가하였으며; 25 ml의 예열된 IMDM, 10% FBS를 세포에 2분에 걸쳐 서서히 첨가하고, 부드럽게 흔들어서 혼합시켰다. 세포를 300 g에서 10분 동안 회전시켰다. 28x10^6개 CD34+ 세포를 StemSpan SFEM + 100ng/ml SCF/IL6/Flt3L/TPO + 500 nM 화합물 4 + 1X Pen/Strep 내에서 0.2-0.5 x 10e6 세포/ml로 배양하였다.

[1772] **RNP 전기천공**

[1773] 해동-후 48시간째에, 세포를 1) Cas9 단독; 또는 2) sg1128(상기 기재된 바와 같음), 또는 표 11의 sgRNA 중 하나를 함유하는 Cas9 RNP 복합체를 이용하여 전기천공시켰다. 간략하게는, 1:2 몰비의 Cas9 단백질 : sgRNA를 사용하여 RNP를 제조하여, 10 uM Cas9 및 20 uM sgRNA와의 RNP 혼합물을 초래하였다. 5 ul의 10 uM RNP를 20 ul P3 완충제(Lonza)에서 1x10^6개 CD34+ 세포에 첨가하였다. 22 ul의 전기천공 믹스를 96웰 전기천공 플레이트의 1개 웰로 옮기고, Lonza Amaxa 96-웰 Shuttle 또는 Lonza Nucleofector 시스템, 프로그램 CA-137을 사용하여 전기천공시켰다. 전기천공 후 즉시, 80 ul의 배양 배지(StemSpan SFEM + 100 ng/ml SCF/IL6/Flt3L/TPO + 500 nM 화합물 4 + 1X Pen/Strep)를 전기천공된 시료에 첨가하였다.

[1774] 전기천공-후 24시간째에, 500 K 출발 세포 등가물을 1 마리의 마우스 당 이식하였다. 30,000 내지 100,000개 세포를 적혈구계 분화에 사용하여, 편집-후 HbF 유도를 평가하였다. 편집 빈도의 NGS 분석을 위해 잔여 세포를 배양물에 부가적인 24시간 동안(전기천공-후 총 48시간) 방치하였다.

[1775] **이식**

[1776] 1개 조건 당 10 마리의 NSG 마우스를 이식 전 4 내지 24시간째에 RadSource X-Ray 방사선 조사 장치를 사용하여 200 Rad로 또는 $Cesium^{138}$ 방사선 조사 장치를 사용하여 2 그레이로 방사선 조사하였다. $5x10^5$개의 출발 세포 등가물을 1 마리의 마우스 당 꼬리 정맥 주사를 통해 이식하였다. 이식 후, 마우스를 4 내지 8주 동안 항생제 투약 방식에 놓았다. 마우스를 위원회의 동물 관리 절차에 따르고 승인된 IACUC 프로토콜에 따라 처리하였다. 4, 8, 12, 16 및 20주째에, 생착 및 계통 분석을 위해 말초 혈액을 꼬리 정맥 닉(nick)을 통해 수합하였다. 8 내지 9주 및 20주째에, 분석(이전의 프로토콜, 예를 들어 실시예 2.1에 기재된 바와 같이 유세포분석, Taqman qPCR and NGS)을 위해서, 뿐만 아니라 이전의 프로토콜에 기재된 바와 같이 적혈구계 분화에 대한 hCD45+CD34+ 세포의 소팅을 위해 골수를 수합하였다. 20주째에, hCD45+CD34+ 세포 소팅 및 적혈구계 분화를 수행하였다.

[1777] 도 22는 BCL11A 유전자의 적혈구계-특이적 인핸서 영역으로부터의 gRNA(sg-G1128; sg1128로도 지칭됨)와 비교하여 감마 글로빈 프로모터 영역으로부터의 sgRNA(sg-G0008, sg-G0051, sg-G0010, sg-G0048, sg-G0067)를 이용하여 편집된 HSC의 줄기세포 기능을 평가하기 위한 이식 연구의 개략도를 보여준다. $5x10^5$개 인간 CD34+ 세포를 0일차에 해동시키고, 3일차에 모크 RNP 복합체(Cas9 단독 및 gRNA 없음), 또는 Cas9(NLS-Cas9-NLS-His6(SEQ ID NO: 247로서 개시된 "His6")) 및 다양한 gRNA를 이용하여 형성된 RNP 복합체를 이용하여 전기천공시켰다. 6일차에, 각각의 조건으로부터의 모든 세포를 수확하고, 2 그레이 아치사 방사선 조사된 NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ(NSG) 마우스에게 이식하였다(도 22). 이식-후 4, 8, 12, 16 및 20주째에 마우스를 양육시켰다. 이식-후 8 및 20주째에, 동물로부터의 골수 세포를 또한 수확하여, 골수에서의 인간 세포 생착을 검사하였다.

[1778] **결과**

[1779] **편집 HSC는 NSG 마우스에서 장기간 생착을 할 수 있었고 다계통 분화를 뒷받침한다**

[1780] 결과는, 감마 글로빈 프로모터 영역으로부터의 sgRNA가 sg-G1128과 비교할 만하게 이식 후 8주째에 평균 20% 내지 40%의 골수 생착을 달성하였음을 보여주었다. 조기 시점으로부터의 생착이 단수명 조혈 전구 세포에 의해 기여될 수 있는 한편, 더 이후의 시점(20주)에서의 생착은 장기간 HSC에 의해 재구성의 증거이다. 시험된 sgRNA는 이식-후 20주째에 5% 내지 22% 골수 생착을 보여주었다(도 23a 내지 도 23f 및 도 24a 및 도 24b). 본 발명자들이 단지 약 500,000개 모크 또는 게놈-편집된 CD34+ 세포를 각각의 마우스 내에 주사하여, 이러한 수준의 생착을 달성하였음을 주지하는 것이 중요하다. 이러한 생착 수준은 $1x10^6$개 게놈-편집된 CD34+ 세포를 이식한 다른 연구와 비교할 만하였으며, 이는 본 세포의 고도로 효율적인 생착을 표시한다. 또한, 본 발명자들은 모든 유전자-편집된 그룹을 모크 편집 대조군과 비교할 때 이식-후 4, 8, 12, 16, 20주째에 골수계, B 림프계, and T 림프계 세포의 정상 회수를 관찰하였다(도 23a 내지 도 23f 및 도 25). 이들 데이터는, 임의의 공지된 HPFH로 맵핑되지 않는 gRNA를 이용하여 감마 글로빈 프로모터 영역의 표적화함으로써, 게놈 편집 전략이 장기간 HSC의 생착 능력에 영향을 미치지 않거나 새로운 숙주에 이식되었을 때 이들의 다계통 재구성 기능을 변경시키지 않음을 실증한다. 대조적으로, 본 발명자들은 다른 보고된 전략과 비교하여 50% 적은 세포를 주사함으로써 강력한(robust) 생착을 달성할 수 있다. 요약하자면, 이들 sgRNA에 의해 편집된 HSC는 장기간 조혈을 지속시키기 위해 조혈 줄기세포 자리(niche)에서 강력한 다계통 재구성과 함께 장기간 생착을 할 수 있다.

[1781] **높은 편집 효율은 이식-전 및 이식-후에 유지되었다**

[1782] 본 발명자들은 실험 절차에 의해 기재된 바와 같이 시험된 sgRNA의 편집 효율을 NGS에 의해 검사하였다. 유전자-편집 후 3일째에, 모크 편집 CD34+ 세포와 함께 100,000개의 유전자-편집된 인간 CD34+ 세포를 NGS 분석을 받게 하였다. 잔여 세포를 이전에 기재된 바와 같이 NSG 수여자에게 이식하였다. 이식-후 8주 및 20주째에, 이식된 NSG 마우스로부터의 골수 세포를 수확하고, 100,000개의 소팅된 인간 CD45+ 세포를 NGS를 받게 하여 편집 효율을 측정하였다. 결과는, sgRNA sg-G1128, sg-G0048 및 sg-G0067이 80-90% 편집을 달성한 한편, sg-G0010은 약 45% 편집을 실증하였음을 보여준다(도 26). 이식-전과 이식 후 8주 또는 20주째에 편집 효율을 비교할 때, 결과는, 조혈 줄기 및 전구체 집단에서의 편집 사건이 이식 기간 전체를 통틀어서 장기간 유지되었음을 보여주었다. 이러한 데이터는 이식된 개체에서 편집된 세포의 내구성(durability)을 실증한다. 본 발명자들은 심지어,

이식-후 20주째에 sg-G1128에 대해 약간 증가된 편집 효율을 관찰하며, 이는, 편집된 세포가 골수 미세환경에 의해 선택되었거나 이식 시 생존 이점을 가짐을 내포한다.

[1783] 또 다른 독립적인 이식 반복에서, NGS 분석은 시험된 gRNA를 이용하여 달성된 편집 범위를 드러내었으며, 이때, sg-G1128은 전기천공-후 2일째에 90% 초과의 편집을 초래하였다(도 8a). 편집은 또한, 이식-후 9주 및 20주째에 단리된 인간 CD34+ 세포에서 검출되었다(도 27b 및 도 27c).

[1784] **유전자-편집, 장기간 생착된 HSC는 태아 헤모글로빈의 증가된 생산을 지속시켰다**

[1785] NSG 마우스는, 이러한 마우스가 적혈구계 성숙을 할 수 있게 하는 올바른(right) 인간 사이토카인이 결여되어 있기 때문에 적혈구계 발달을 뒷받침하지 않는다. 그러나, 생착된 인간 HSC를 마우스 골수로부터 수확하고, 적혈구계 분화 배지(이들 실시예 어디에서나 기재된 바와 같음)에 놓아, 적혈구계 분화를 유도할 수 있다. 데이터는, 유전자-편집된, 20주 생착된 HSC가 모크 편집 및 이식된 HSC와 비교하여 더 높은 수준의 HbF를 생산할 수 있었음을 보여준다(도 28).

[1786] **실시예 7. 표적-외 삽입결실 패턴 분석**

[1787] **잠재적인 gRNA 표적-외 유전자좌의 *인 실리코* 식별**

[1788] HGB1/HBG2 영역 gRNA GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051 및 GCR-0067에 대한 잠재적 표적-외 유전자좌를 다음과 같이 확인하였다. 각각의 gRNA에 대해, 5 뉴클레오타이드 미스매치까지 허용하는 표준 매개변수를 사용하는, BFAST 서열 정렬기(버전 0.6.4f, 문헌[Homer et al, PLoS One, 2009, 4(11), e7767, PMID: 19907642])를 사용하여 20 뉴클레오타이드 gRNA 프로토스페이서(protospacer) 서열을 인간 게놈 참조 서열(build GRCh38)에 정렬하였다. 확인된 유전자좌를 Cas9 정준 5'에 인접한 5'-NGG-3' PAM 서열(즉, 5'-표적-외 유전자좌-PAM-3')인 부위만을 함유하도록 필터링하였다. 5 뉴클레오타이드 미스매치를 갖는 부위는 BEDTools 스크립트(버전 2.11.2, 문헌[Quinlan and Hall, Bioinformatics, 2010 26(6):841-2, PMID: 20110278])를 사용하여RefSeq 유전자 주석(문헌[Pruitt et al, Nucleic Acids Res., 2014 42(Database issue):D756-63, PMID: 24259432]) 대비 엑손으로 주석 표시된 유전자좌만 함유하도록 추가로 필터링하였다. BCL11A+ 58 영역 및 HPFH 영역 gRNA에 대해 확인된 잠재적 표적-외 유전자좌의 수를 표 12에 나타내었다.

[1789] [표 12]

[1790] RefSeq 엑손 내에 0, 1, 2, 3 및 4개의 뉴클레오타이드 미스매치 및 5개의 뉴클레오타이드 미스매치를 갖는 HBG1/HBG2 영역 gRNA GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051 및 GCR-0067에 대해 식별된 인 실리코 표적-외 유전자좌의 수치가 제시된다. *gRNA GCR-0048 및 GCR-0067은 1개 HBG 유전자좌 당 2개의 완벽한 매치 표적-적중 부위를 가진다. + 는 1 또는 2개의 미스매치 상동성 HBG1 또는 HBG2 표적 부위를 포함한다.

[1791]

gRNA 명칭	N 미스매치를 갖는 표적-외 개수						
	0	1	2	3	4	5 RefSeq 엑손	총 부위
GCR-0001	0	0	4+	17	323	28	372
GCR-0008	0	1+	0	27	267	60	355
GCR-0010	0	1+	0	17	216	64	298
GCR-0048*	0	0	0	7	89	38	134
GCR-0051	0	1+	3	25	369	66	464
GCR-0067*	0	0	0	11	95	38	144

[1792] CD34+ HSPC에서의 표적-외 분석

[1793] **게놈 DNA 추출**

[1794] RNP로 처리된 및 처리되지 않은 HSPC로부터 제조사의 권장 사항에 따라 DNeasy Blood & Tissue 키트(Qiagen, Cat # 69506)를 사용하여 게놈 DNA를 단리하였다.

[1795] **잠재적 표적-외 부위의 표적화된 증폭을 위한 PCR 프라이머 설계**

[1796] 앰플리콘의 중심에 위치하는 gRNA 프로토스페이서 서열을 갖는, 대략 160 내지 300 염기쌍 길이 범위의 앰플리콘 크기를 목표로 하는 디폴트 매개변수를 사용하는, Primer3(버전 2.3.6, 문헌[Untergasser et al, Nucleic

Acids Res., 2012 40(15):e115, PMID: 22730293])를 사용하여 BCL11A +58 영역 gRNA(CR00309, CR00311, CR00312, CR00316, CR001125, CR001126, CR001127 및 CR001128) 및 HPFH 영역 gRNA(CR001028, CR001030, CR001137, CR001221, CR003035 및 CR003085)에 대해 확인된 잠재적 표적-외 유전자좌(및 표적 적중 유전자좌)를 표적화하는 PCR 앰플리콘을 설계하였다. gRNA CR001127에 대해서는 0 내지 4 뉴클레오타이드 미스매치를 갖는 부위에 대해서만 PCR 프라이머를 설계하였으며, RefSeq 엑손 내에 5 뉴클레오타이드 미스매치를 갖는 부위에 대해서는 프라이머를 설계하지 않았다. 생성된 PCR 프라이머 쌍 및 앰플리콘 서열을 인간 게놈 참조 서열(build GRCh38) 대비 BLAST 서열 검색(버전 2.2.19, 문헌[Altschul et al, J Mol Biol., 1990, 215(3):403-10, PMID: 2231712])하여 유일성을 체크하였다. 하나를 초과하는 앰플리콘 서열을 야기하는 프라이머 쌍을 폐기하고 재설계하였다. 표 3 및 표 5는 BCL11A +58 영역 및 HPFH 영역 gRNA 각각에 대해 설계된 성공적인 PCR 프라이머 쌍의 수를 나타낸다.

[1797] **Illumina 시퀀싱 라이브러리 제조, 정량화 및 시퀀싱**

[1798] 제조사의 권장 사항을 이용하는 Quant-iT PicoGreen dsDNA 키트(Thermo Fisher, Cat # P7581)를 사용하여 RNP 처리된(gRNA 당 2 반복실험) 및 처리되지 않은(gRNA 당 1 반복실험) HSPC 시료로부터의 게놈 DNA를 정량화하였다. 2개의 순차적인 PCR 반응을 사용하여 각 시료에 대하여 개개의 표적-외 유전자좌(및 표적 적중 유전자좌)를 표적화하는 Illumina 시퀀싱 라이브러리를 생성하였다. 제1 PCR은 보편적인 Illumina 시퀀싱 호환 서열로 테일링(tailing)된 표적 특이적 PCR 프라이머(상기 설계된)를 사용하여 표적 유전자좌를 증폭하였다. 제2 PCR은 시퀀싱 동안 멀티플렉싱을 가능하게 하는 시료 바코드를 포함하는, 추가적인 Illumina 시퀀싱 호환 서열을 제1 PCR 앰플리콘에 첨가하였다. PCR 1은, 각 반응이 3 내지 6 ng의 gDNA(약 500 내지 1000 세포에 상응), 최종 농도 0.25 μM의 PCR 1 프라이머 쌍(Integrated DNA Technologies) 및 1x 최종 농도의 Q5 Hot Start 마스터 믹스(New England BioLabs, Cat # 102500-140)을 함유하는, 10 μL의 최종 부피로 수행하였다. PCR 1 좌측 프라이머는 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3' 서열로 5' 테일링되고(즉, 5'-테일-표적 특이적 좌측 프라이머-3'), 우측 프라이머는 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3' 서열로 5' 테일링되었다(즉, 5'-테일-표적 특이적 우측 프라이머-3'). PCR 1은 다음의 사이클 조건을 사용하여 써모사이클러(thermocycler) 상에서 수행하였다: 1분 동안 98℃ 1 사이클; 10초 동안 98℃, 20초 동안 63℃, 30초 동안 72℃ 25사이클; 2분 동안 72℃ 1 사이클. 그런 다음, PCR 1을 뉴클레아제 비함유 물(Ambion, Cat# AM9932)을 사용하여 100중 1로 희석하고 PCR 2로의 투입물로 사용하였다. PCR 2는, 각 반응이 2 μL의 희석된 PCR 1 산물, 최종 농도 0.5 μM의 PCR 2 프라이머 쌍(Integrated DNA Technologies) 및 1x 최종 농도의 Q5 Hot Start 마스터 믹스(New England BioLabs, Cat # 102500-140)를 함유하는, 10 μL의 최종 부피로 수행하였다. 사용된 PCR 2 좌측 프라이머 서열은 5'-AATGATACGGCGACCACCGAGATCTACACNNNNNNNNTCGTCGGCAGCGTC-3'였고, 사용된 PCR 2 우측 프라이머 서열은 5'-CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTCTCGTGGGCTCGG-3'였으며, 여기서 NNNNNNNN은 표준 Illumina 시퀀싱 과정의 부분으로서 시료 멀티플렉싱에 사용되는 8 뉴클레오타이드 바코드 서열을 의미한다. PCR 2는 다음의 사이클 조건을 사용하여 써모사이클러(thermocycler) 상에서 수행하였다: 3분 동안 72℃ 1 사이클; 2분 동안 98℃ 1 사이클; 10초 동안 98℃, 30초 동안 63℃, 2분 동안 72℃ 15사이클. 현재 실행 가능한 Illumina 시퀀싱 라이브러리인, PCR 2 앰플리콘을 제조사의 권장 사항에 따라 Agencourt AMPure XP 비드(Beckman Coulter, Cat # A63882)를 이용하여 정화하였다. 그 다음에, 정화된 Illumina 시퀀싱 라이브러리를 Power SYBR Green PCR 마스터 믹스(Life Technologies, Cat# 4367660) 및 Illumina 시퀀싱 라이브러리 말단에 특이적인 프라이머(포워드 프라이머 서열 5'-CAAGCAGAAGACGGCATACGA-3' 및 리버스 프라이머 서열 5'-AATGATACGGCGACCACCGAGA-3')를 사용하는 표준 qPCR 정량화 방법을 사용하여 정량화하였다. 그 다음에, Illumina 시퀀싱 라이브러리들을 등몰 푸울링하고, 제조사의 권장 사항에 따라 MiSeq Reagent Kit v3(Illumina, Cat# MS-102-3003)를 사용하여 300 염기의 쌍-말단 리드로 MiSeq 기기(Illumina, Cat# SY-410-1003) 상에서 Illumina 시퀀싱을 하였다. 적어도 하나의 처리된 반복실험에서 각 유전자좌에 대해 최소 1000 배의 서열 커버리지(coverage)가 생성되었다. 처리된 및 처리되지 않은 시료의 PCR, 정화, 푸울링 및 시퀀싱을 개별적으로 수행하여 처리된 및 처리되지 않은 시료 또는 그로부터 생성된 PCR 앰플리콘 사이에 교차 오염의 가능성을 피하였다.

[1799] **NGS 시퀀싱 데이터 QC 및 변이체 분석**

[1800] 방법은 실시예 2.1에서와 같고, 다음의 변형을 갖는다. 분석의 단계 5를 위해, 2% 초과의 조합된 삽입결실 빈도(대략 10 내지 20개 초과의 세포에서의 편집)를 갖는 부위를 고려하였으며, 잠재적인 활성 편집 부위를, 게놈 참조 서열에 대한 리드 정렬의 시각적인 조사를 허용하는 Integrative 게놈 Viewer(IGV 버전 2.3, 문헌[Robinson et al, Nat Biotechnol. 2011, 9(1):24-6], PMID: 21221095)를 사용하여 리드 정렬 수준에서 추가로 검사하였다.

[1801] 대형 결실의 정량화.

[1802] HBG1 및 HBG2 프로모터에서 표적화 부위 사이의 대형 결디지털 방울 PCR(ddPCR)에 의해 제조업체의 권고사항에 따라 복사수 결정을 위한 비-경쟁적 검정법 포맷에서 정량화하였다. 간략하게는, gDNA를 프로브용 ddPCR SuperMix(dUTP 없음)(BioRad Cat# 1863024), HindIII-HF 제한 효소(NEB Cat# R3104S) 및 각각의 프라이머 프로브 믹스와 조합하고, 프로브용 방울 발생 오일(BioRad Cat# 1863005)과 함께 DG8 카트리지로 옮기고, QX200 방울 발생기(BioRad)를 이용하여 방울을 발생시켰다. 방울을 96-딥 웰 반응 모듈이 있는 C1000 Touch Thermal Cylcer(BioRad) 상에서 PCR을 받게 하고, 뒤이어 QX200 방울 리더(BioRad)를 이용하여 검출하였다. 1 ul 당 복사물를 QuantaSoft 소프트웨어(BioRad)를 이용한 분석에 의해 결정하였다. 커스텀 프라이머 프로브 세트(Life Technologies Cat# APZW76R, PN4331348, 포워드 프라이머: ACGGATAAGTAGATATTGAGGTAAGC(SEQ ID NO: 285), 리버스 프라이머: GTCTCTTTCAGTTAGCAGTGG(SEQ ID NO: 286), FAM TaqMan 프로브: ACTGCGCTGAAACTGTGGCTTTATAG(SEQ ID NO: 287))를 사용하여, HBG1-HBG2 유전자간 영역 내에서 gDNA를 증폭시켰다. TaqMan 복사수 참조 검정법, 인간, RNase P (Thermo Fisher cat# 4403326)를 참조 앰플리콘으로서 사용하였다. HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물은 제조업체의 보고된 선형 범위 내에 있었다. 퍼센트 결실을 100% x (1 - HBG1-HBG2 및 RNase P 앰플리콘에 대한 1 ul 당 복사물의 비)로서 보고하였다. 비편집된 대조군 시료는 2% 이하의 계산된 퍼센트 결실을 가졌으며, 이는 검정법에서 백그라운드를 반영할 수 있다.

[1803] **HBG1/HBG2 영역 *인 실리코* 표적-외 분석 결과**

[1804] 표 13은 성공적으로 특징화된 표적-외 부위의 수를 보여준다. 비특징화된 부위는 PCR 프라이머 설계 또는 PCR 증폭에 실패하였고, 평가되는 것으로 남아 있다.

[1805] **gRNA GCR-0001**: HBG1 표적-적중 부위는 대략 58%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여준 반면, gRNA 표적화 도메인 서열에 비해 2개의 미스매치(5'-AGTCCTGGTATCtTCTATGg - PAM-3', PAM = TGG(SEQ ID NO: 288))를 갖는 상동성 HBG2 표적 부위(본원 및 하기에서 소문자로 제시됨)는 최소의 국소화된 편집을 보여주었다. 그러나, 4.9 kb 결실의 ddPCR 분석은 이를 대략 34%의 빈도로 발생하는 것을 보여주었다. 추가의 분석은 2개 복제물 모두에서 대략 26%의 평균 삽입결실 빈도를 갖는 1개의 양성 표적-외 부위를 식별하였다. 부위는 gRNA 표적화 도메인 서열에 비해 3개의 미스매치(5'-AtTCCcaGTATCCTCTATGA-PAM-3', PAM = TGG(SEQ ID NO: 289))를 갖고, Y 염색체 상에서 염기쌍 위치 21,470,475 내지 21,470,497에서 유전자간에 위치한다. 이러한 표적-외 부위에서의 편집이 유전자 발현 또는 세포 생존력에 임의의 유해 효과를 갖는지 불분명하며, 추가의 분석이 필요해진다.

[1806] **gRNA GCR-0008**: HBG1 표적-적중 부위는 대략 88%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. gRNA 표적화 도메인 서열에 비해 1개의 미스매치를 갖는 상동성 HBG2 표적 부위(5'-GGAGAAGaAAACTAGCTAAA- PAM-3', PAM = GGG(SEQ ID NO: 290))는 대략 85%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. 4.9 kb 결실의 ddPCR 분석은 이를 대략 50%의 빈도로 발생하는 것을 보여주었다. 다른 어떠한 부위도 편집을 보여주지 않았다.

[1807] **gRNA GCR-0010**: HBG1 표적-적중 부위는 대략 32%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. gRNA 표적화 도메인 서열에 비해 1개의 미스매치를 갖는 상동성 HBG2 표적 부위(5'-GGGAGAAGaAAACTAGCTAA- PAM-3', PAM = AGG(SEQ ID NO: 291))는 대략 27%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. 4.9 kb 결실의 ddPCR 분석은 이를 대략 33%의 빈도로 발생하는 것을 보여주었다. 다른 어떠한 부위도 편집을 보여주지 않았다.

[1808] **gRNA GCR-0048**: HBG1 및 HBG2 표적-적중 부위는 두 부위 모두에 대해 대략 86%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. 4.9 kb 결실의 ddPCR 분석은 이를 대략 45%의 빈도로 발생하는 것을 보여주었다. 다른 어떠한 부위도 편집을 보여주지 않았다.

[1809] **gRNA GCR-0051**: HBG2 표적-적중 부위는 대략 88%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. gRNA 표적화 도메인 서열에 비해 1개의 미스매치를 갖는 상동성 HBG1 표적 부위(5'-GGAGAAGgAAACTAGCTAAA- PAM-3', PAM = GGG(SEQ ID NO: 292))는 대략 59%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. 4.9 kb 결실의 ddPCR 분석은 이를 대략 40%의 빈도로 발생하는 것을 보여주었다. 다른 어떠한 부위도 편집을 보여주지 않았다.

[1810] **gRNA GCR-0067:** HBG1 및 HBG2 표적-적중 부위는 각각 대략 74 및 78%의 평균 삽입결실 빈도를 갖는 강력한 국소화된 편집을 보여주었다. 4.9 kb 결실의 ddPCR 분석은 이를 대략 62%의 빈도로 발생하는 것을 보여주었다. 다른

어떠한 부위도 편집을 보여주지 않았다.

[1811] 상기 기재된 국소화된 INDEL 빈도는 총 삽입결실 빈도에 비하지 않고, 대형 4.9 kb 결실의 빈도를 고려하지 않는다.

[1812] [표 13]

[1813] HBG1/HBG2 영역 gRNA GCR-0001, GCR-0008, GCR-0010, GCR-0048, GCR-0051 및 GCR-0067에 대해 식별된 인 실리코 0-3 미스매치 표적-외 부위의 수치, 게놈-편집된 HSPC에서 성공적으로 특징화된 부위의 수치, 및 편집을 보여주는 부위의 수치가 제시된다.

[1814]

gRNA 명칭	0-3 미스매치 *인 실리코* 표적-외 부위의 수	성공적으로 특징화된 *인 실리코* 부위의 수	식별된 활성 *인 실리코* 표적-외 부위의 수
GCR-0001	21	19	1
GCR-0008	28	28	1 (HBG2)
GCR-0010	18	18	1 (HBG2)
GCR-0048	7	7	0
GCR-0051	29	29	1 (HBG1)
GCR-0067	11	11	0

[1815] 불편파(unbiased) 표적-외 분석

[1816] 올리고 삽입-기초 검정법(예를 들어 문헌[Tsai et al., *Nature Biotechnology*. **33**, 187-197; 2015] 참조)을 사용하여, HBG1 및/또는 HBG2를 표적화하는 Cas9에 의해 절단된 잠재적인 표적-외 게놈 부위를 결정하였다. 이들 실험에서, Cas9GFP-발현 HEK293 세포(HEK-293_Cas9GFP)를 Lipofectamine® RNAiMAX를 사용하여 gRNA(15 nM crRNA:tracr) 및 삽입 올리고(10 nM)로 형질감염시켰다. 검정법은 게놈 내의 이중 가닥 절단부 내에 혼입된 올리고의 식별에 의존하며, 이는 Cas9에 의한 절단으로 인한 것일 수 있거나 그렇지 않을 수 있다.

[1817] 하나의 실험에서, HBG1 및/또는 HBG2를 표적화하는 gRNA(도 29에서 지시된 표적화 도메인을 포함하는 이중 가이드 RNA)를 HEK-293_Cas9GFP 세포에서 스크리닝하고, 그 결과를 도 29에 플롯화한다. 별개의 실험에서, 동일한 방법을 사용하여, HEK-293_Cas9GFP 세포에서 HBG1 및/또는 HBG2를 표적화하는 동일한 gRNA뿐만 아니라 부가적인 gRNA(도 30에서 지시된 표적화 도메인을 포함하는 이중 가이드 및 단일 가이드 RNA를 포함함) 중 일부를 스크리닝하였으며, 그 결과를 도 30에 플롯화한다. 도 30에 도시된 데이터를 갖는 실험은 균형잡힌 G/C 함량을 갖는 대안적인 삽입 올리고를 사용하였고, 인간 게놈과의 상보성, 뿐만 아니라 Tsai 등의 방법과 비교하여 PCR 단계에 대한 다른 변형을 최소화하였고, 이는 민감성을 개선하고 위양성을 줄였다. 두 실험 모두에서, 검정법은 예상된 표적 서열 및 하나 이상의 잠재적인 표적-외 부위에서 높은-효율 편집을 검출하였다.

[1818] 표적-적중 부위 이외의 게놈 내의 부위에서 삽입 올리고의 검출이 Cas9의 잠재적인 표적-외 효과를 식별하긴 하지만, 잠재적인 표적-외 부위의 표적화된 딥(deep) 시퀀싱을 사용하여, 잠재적인 부위가 진실로 Cas9에 의해 절단되는 표적-외 부위인지 결정할 수 있다. 이를 위해, HEK-293_Cas9GFP 세포를, 삽입 올리고가 없이 이전의 2개 실험에서 사용된 gRNA(25 nM crRNA:tracr에서)로 유사하게 형질감염시킨 실험을 수행하였다. 도 29 및 도 30에서 도시된 식별된 잠재적인 표적-외 부위 각각의 앰플리콘 딥 시퀀싱을 사용하여, HEK-293_Cas9GFP 세포주 내의 잠재적인 표적-외 부위에 삽입결실(Cas9 절단 사건을 가리킴)이 존재하는지 식별하였다. 이러한 실험의 결과는, 올리고 삽입 검정법에 의해 식별된 잠재적인 표적-외 부위 대부분이 gRNA의 형질감염 후 검출 가능한 삽입결실을 갖지 않았으며, 이때 하기 표 14에 제공된 바와 같이 몇 가지 예외가 있었음을 실증하였다. 주지하자면, 잠재적인 표적-외를 검출하기 위해 이들 실험에 사용된 HEK-293_Cas9GFP 세포는 Cas9를 구성적으로 과발현하며, 이는 다양한 관심 세포 유형(예를 들어 CD34+ HSC)에서 일시적인 전달 양상(예를 들어 RNP 전달)과 비교하여 더 높은 수의 잠재적인 표적-외 "히트(hit)"를 초래하는 경향이 있다.

[1819] [표 14]

[1820] HEK-293_Cas9GFP 세포주에서 입증된 표적-외 부위.

가이드 ID	표적-외 좌표	평균 Indel 빈도
CR005813 (dgRNA; GCR-0051)	chr13:95591406-95591426	2%
G000690 (sgRNA GCR-0051)	chr13:95591406-95591426	2%
CR005821 (dgRNA GCR008)	chr20:10409602-10409622	27.60%
G000692 (sgRNA GCR008)	chr20:10409602-10409622	20.10%

[1821]

[1822] 본원에 기재된 방법을 사용하여, 잠재적인 표적-외 부위를 sgRNA/Cas9 편집된 CD34+ HSPC에서 검사하였고, CD34+ 세포 유형에서 어떠한 편집도 보여주지 않았다.

[1823] **실시예 8**

[1824] **실험 절차**

[1825] 건강한 공여자의 동원된 말초 혈액(mPB)으로부터 수확된 CD34+ 세포를 실시예 6에서 이전에 기재된 바와 같이 동일한 조건에서 배양하고 유전자 편집하였다. 세포를 이전의 단락에 기재된 동일한 방법을 사용하여 이들 세포의 생물학적 기능에 대해 특징화하였다.

[1826] **동원된 말초 혈액으로부터 유래된 CD34+ 세포는 편집 시 이들 세포의 CD34+ 세포 수치, 확장 능력, 및 생존력을 유지시킨다**

[1827] 환자에게 이식된 CD34+ 세포의 수는 골수 이식의 성공과 직접적으로 상관관계가 있다. 따라서, 본 발명자들은 임상적으로 허용 가능한 방법 ISHAGE를 사용하여 유전자-편집 시 10-일의 기간에 걸쳐 CD34+ 세포의 퍼센트를 나열하였다. 결과는, sg1128 또는 sg0067을 이용한 어떠한 편집도 모크 편집 대조군과 비교하였을 때 CD34+ 세포 수치(도 31)에 영향을 주지 않았음을 보여준다. 세포가 전기천공 후 3일 동안 배양물에서 유지만 되는 본 발명자들의 세포 과정 총 기간 동안, CD34+ 세포의 퍼센트는 대략 90%에서 유지되었다(도 31a). 유전자 편집은 또한, CD34+ 세포가 확장되는 능력에 영향을 주지 않았고(도 31b), 전기천공-후 이들 세포의 생존력은 70% 내지 90%의 범위였다(도 31c). 본 발명자들은 전기천공 후 3일째에 약 2배의 세포 확장, 전기천공-후 7일째에 약 10배의 확장, 및 전기천공 후 10일째에 15 내지 25배의 확장을 관찰하였다(도 31b).

[1828] **건강한 개체의 동원된 말초 혈액으로부터 유래된 CD34+ 세포는 겸상 적혈구 질병 환자로부터의 골수 세포 공급원 또는 CD34+ 세포와 비교하여 유사한 편집 효율을 실증한다**

[1829] Sg1128은 건강한 공여자의 mPB로부터 유래된 CD34+ 세포에서 약 70% 내지 75% 편집 효율을 실증한 반면, sg0067은 95% 초과의 총 편집 효율(대형 5 kb 결실 및 소형 삽입결실을 포함함)을 보여주었다(도 32). 이들 sgRNA로부터의 편집 효율은 건강한 공여자로부터의 골수 유래 CD34+ 세포 또는 겸상 적혈구 질병 환자로부터의 말초 혈액 유래 CD34+ 세포를 포함하여 상이한 공급원으로부터의 세포에서 유사하였다(다른 실시예 참조). 각각의 sgRNA의 편집 효율과 편집 패턴 둘 모두가 시험된 모든 공여자에 걸쳐 고도로 일관되었음을 주지하는 것이 중요하다.

[1830] **BCL11A의 CRISPR 넉다운 또는 g-글로빈 유전자 클러스터의 돌연변이는 g-글로빈 전사체 및 F-세포 생산을 증가시킨다**

[1831] sg1128에 의한 BCL11A의 넉다운, 또는 sg0067에 의한 g-글로빈 유전자 클러스터에서 잠재적인 BCL11A 결합 부위의 돌연변이 둘 모두는 g-글로빈 전사체를 유의하게 증가시켰으며(도 33), 이는 모크 편집 대조군과 비교하여 F-세포 생산의 15% 내지 20% 상향조절을 초래하였다(도 34). 요약하자면, sg1128 및 sg0067은 CD34+ 세포를 높은 효율로 편집하고, 고도로 일관된 편집 패턴을 발생시킬 수 있다. 편집된 세포는 본 발명자들의 6일 세포 가공 절차의 종료까지 대략 90% CD34+ 세포 수치를 유지시킨다. 세포의 확장 능력 및 생존력은 유전자 편집 절차에 의해 손상되지 않았다. 편집된 CD34+ 세포는 적혈구로 분화되었을 때, 유의하게 더 높은 수준의 g-글로빈 전사체를 발현시켰으며, 이는 증가된 수의 F-세포 생산으로 번역된다. 이러한 개선된 헤모글로빈 발현 및 F-세포 수는 겸상 적혈구 질병의 혈액학적 특징을 구제할 수 있다.

[1832] **실시예 9: 겸상 적혈구 질병 환자로부터 유래된 유전자 편집된 HSPC의 생체내 생착 및 특징화**

[1833] **실험 절차**

[1834] 겸상 적혈구 질병 환자로부터의 CD34+ 세포를 실시예 8에서 이전에 기재된 바와 동일한 조건에서 배양하고 유전자 편집하였다. 세포를 이전의 실시예에 기재된 동일한 방법을 사용하여 이들 세포의 생물학적 기능에 대해 특징화하였다.

[1835] **겸상 적혈구 질병 개체로부터 유래된 CD34+ 세포는 편집 시 이들 세포의 CD34+ 면역형 및 생존력을 유지시킨다.**

[1836] 원시 세포 상태를 유지하는 조혈 줄기 및 전구 세포(HSPC)는 CD34를 발현해야 한다. 이는, 이식 시 생착 가능한 조혈 줄기세포와 연관된 가장 중요한 임상 마커 중 하나이고, 이식물의 성공은 이식된 CD34+ 세포의 수와 직접적으로 상관관계가 있다. 따라서, 본 발명자들은 임상적으로 허용 가능한 방법 ISHAGE를 사용하여 유전자-편집 시 겸상 개체의 말초 혈액으로부터 수득된 CD34+ 세포의 수를 나열하였다. 결과는, sg1128 또는 sg0067을 이용한 어떠한 편집도 모크 편집 대조군과 비교하였을 때 CD34+ 세포 수치(도 35a 및 도 35b)에 영향을 주지 않았음을 보여준다. CD34+ 세포의 퍼센트는 편집 시 3일째에 대략 80%에서 유지되었다(도 35b). 유전자 편집은 또한, 환자 유래의 CD34+ 세포가 확장되는 능력에 영향을 주지 않았고(도 35c), 전기천공-후(D0에서의 전기천공) 이들 세포의 생존력은 10-일 기간에 걸쳐 측정된 모든 시점에서 75% 초과였다(도 35d).

[1837] **겸상 적혈구 질병 개체로부터 유래된 CD34+ 세포는 건강한 공여자로부터의 CD34+ 세포와 비교하여 유사한 편집 효율을 실증한다**

[1838] Sg1128은 겸상 적혈구 환자로부터 유래된 CD34+ 세포에서 약 35% 편집 효율을 실증한 반면, sg0067은 환자 시료에서 80% 내지 95% 편집 효율을 보여주었다(도 36). 편집 효율의 수준은 건강한 공여자로부터의 골수 또는 동원된 말초 혈액으로부터 유래된 CD34+ 세포를 사용하여 수득된 편집 효율의 수준과 유사하다. NGS(실시예 2.1에 기재된 바와 같음)에 의해 측정된 바와 같은 각각의 가이드 RNA의 편집 패턴 또한, 상이한 환자에 걸쳐 고도로 일관되었다.

[1839] **유전자 편집은 겸상 적혈구 질병 개체로부터 유래된 CD34+ 세포의 *시험관내* 다계통 분화 능력을 약화시키지 않는다**

[1840] 유전자 편집된 조혈 줄기 및 전구 세포는 콜로니-형성 단위 검정법에 의해 측정된 바와 같이 적혈구계, 과립구계, 단핵구계 및 거핵구계 계통으로 완전히 분화할 수 있었다(도 37).

[1841] **BCL11A의 CRISPR 넉다운 또는 g-글로빈 유전자 클러스터의 돌연변이는 g-글로빈 전사체, F-세포 생산, 및 태아 헤모글로빈 발현을 증가시킨다**

[1842] sg1128에 의한 BCL11A의 넉다운, 또는 sg0067에 의한 g-글로빈 클러스터에서 삽입결실/결실의 생성은 유세포분석에 의해 측정된 바와 같이 g-글로빈 전사체를 유의하게 증가시켰고(도 38), F-세포의 수를 증가시켰다(도 39). 또한, F-세포의 태아 헤모글로빈 발현 강도 또한, 유세포분석에 의해 측정된 바와 같이 세포 당 기준(per cell basis)으로 개선되었다(도 40).

[1843] **CD34+ 세포로부터 유래된 환자의 유전자 편집은 겸상 적혈구 수치에서 유의한 저하 및 정상 세포 수에서 증가를 초래하였다**

[1844] 마지막으로, 유전자 편집이 겸상 적혈구 형태를 구제할 수 있는지 이해하기 위해, 본 발명자들은 겸상 적혈구 질병 환자 유래의 CD34+ 세포를 CRISPR-편집하고, 이들 세포를 적혈구로 분화시키고, 이들 세포를 4일 동안 저산소증 챔버 내를 거치게 하고, 겸상 적혈구 형태를 유도하였다. 세포를 항-HbF-FITC 항체로 공동-염색하고, 챔버 내에서 고정하고, 이미징 유세포분석을 받게 하여, 1개 세포 당 1개의 단일 이미지를 포착하였다. 단일 세포 이미징 유세포분석은 높은 처리량 방식에서 세포 길이 및 HbF의 발현을 기초로 겸상 적혈구 대 정상 세포를 구분할 수 있다(각각의 환자로부터의 40,000개 단일 세포 이미지를 데이터 분석에 사용하였음). 결과는, sg1128 또는 sg0067을 이용한 유전자 편집이 겸상 적혈구 수치를 대략 40% 저하시키고(도 41a), 병발적으로 정상 세포 수치를 1.2배 증가시킬 수 있었음(도 41b)을 보여준다. 겸상 적혈구 수치를 저하시키고 병발적으로 정상 적혈구 수치를 증가시키는 화합물 효과는 임상으로 번역되었을 때 환자에게 큰 이득이 될 것이다.

[1845] 요약하자면, sg1128 및 sg0067은 겸상 적혈구 질병 환자로부터의 CD34+ 세포를 높은 효율로 편집하고 고도로 일관된 편집 패턴을 발생시킬 수 있다. CD34+ 세포의 퍼센트, 세포 확장 능력, 및 세포의 생존력은 편집된 세포를 모크 편집 대조군과 비교하였을 때 유전자 편집 절차에 의해 손상되지 않았다. 편집된 CD34+ 세포는 적혈구로

분화되었을 때, 유의하게 더 높은 수준의 g-글로빈 전사체를 발현하였다. 이는 증가된 수의 F-세포로 번역되고, 또한 1개 세포 당 HbF 발현을 증가시켰다. 높은 HbF-발현 F-세포의 수의 증가 및 경상 적혈구 수의 감소는 본 발명자들의 단일 세포 이미징 유세포분석에서 반영되었다. 본 발명자들은 모크 편집 대조군과 비교하였을 때 유전자 편집된 그룹에서 겸상 적혈구의 50% 저하 및 동시적으로 편집된 그룹에서 높은 HbF-발현 정상 적혈구의 1.2배 증가를 관찰하였다. 유전자 편집 시 높은 HbF-발현 정상 적혈구에서의 병발하는 증가와 더불어 겸상 적혈구 수를 감소시키는 이러한 화합물 효과는 임상으로 번역되었을 때 환자에게 유의한 이득을 주어야 한다. 더불어, 이들 데이터는 b-글로빈병증을 치료하기 위한 세포 요법의 수단으로서 CRISPR/Cas-매개 게놈 편집의 발달을 뒷받침한다.

[1846] 임의의 서열 목록과 명세서에서 나열된 임의의 서열 사이에 임의의 불일치가 존재하는 정도까지, 명세서에서 나열된 서열은 올바른 서열인 것으로 여겨져야 한다. 다르게 지시되지 않는 한, 모든 게놈 위치는 hg38에 따른 것이다.

[1847] 참조로서 포함

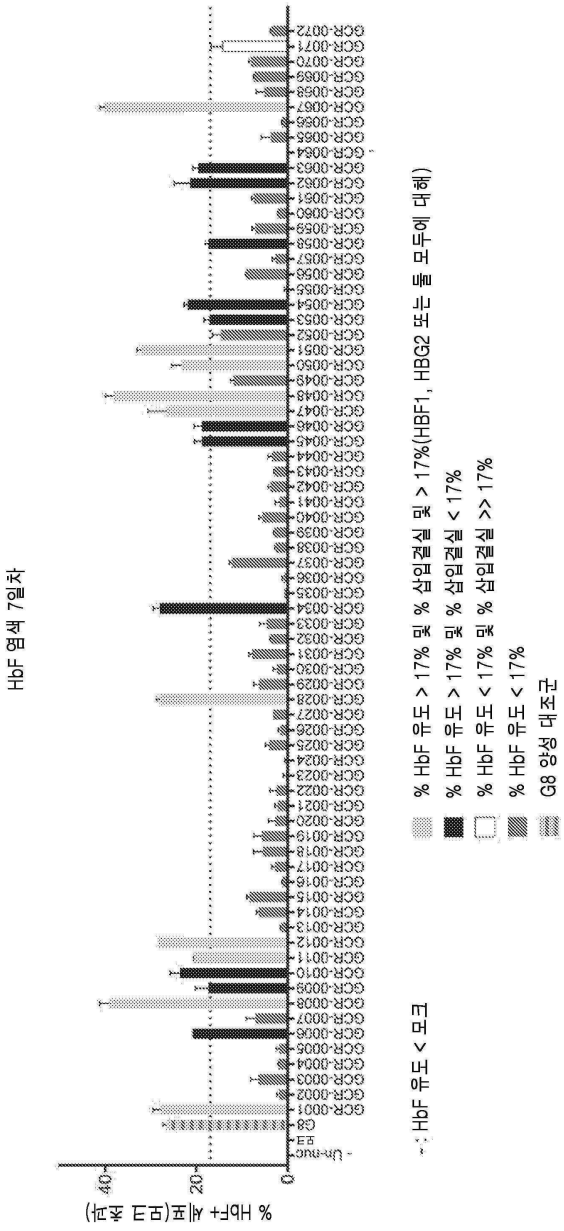
[1848] 본원에 언급된 모든 간행물, 특허, 및 특허 출원은, 각 개별 간행물, 특허 또는 특허 출원이 구체적이고 개별적으로 참조로 포함되는 것으로 기재된 것과 같이, 그 전문이 참조로 본원에 포함된다. 분쟁의 경우에, 본원에서의 임의의 정의를 포함하여, 본 출원이 우선한다.

[1849] 등가물

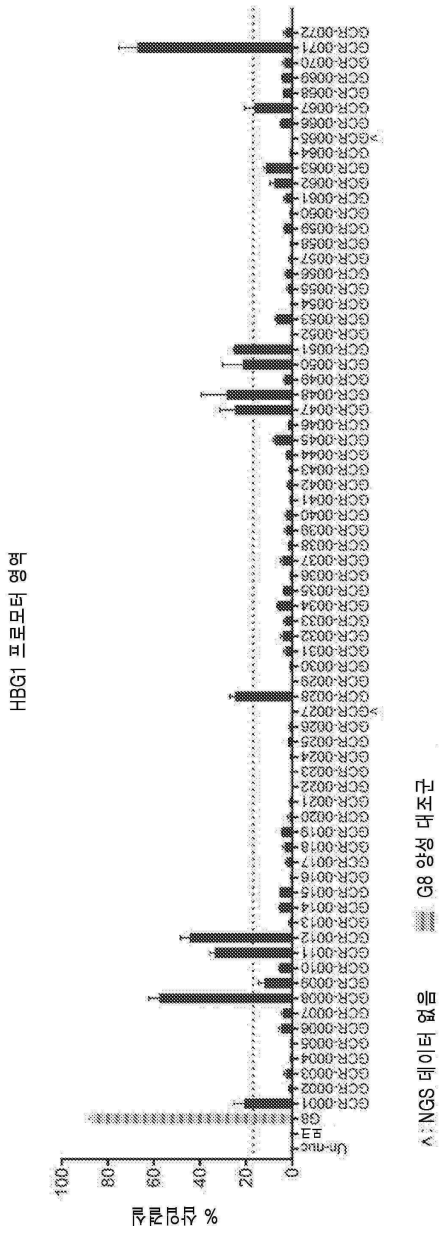
[1850] 당업자는 본원에 기재된 본 발명의 특정 실시형태의 많은 등가물을 단지 일상적인 실험을 이용하여 인식하거나, 확인할 수 있을 것이다. 이 발명은 특정 양태를 참조로 하여 개시되었으나, 본 발명의 다른 양태 및 변형이 본 발명의 진정한 사상 및 범위로부터 벗어나지 않고 당업자에 의해 고안될 수 있음은 자명하다. 첨부된 청구범위는 이러한 양태 및 등가물 변형 모두를 포함하는 것으로 해석되도록 의도된다.

도면

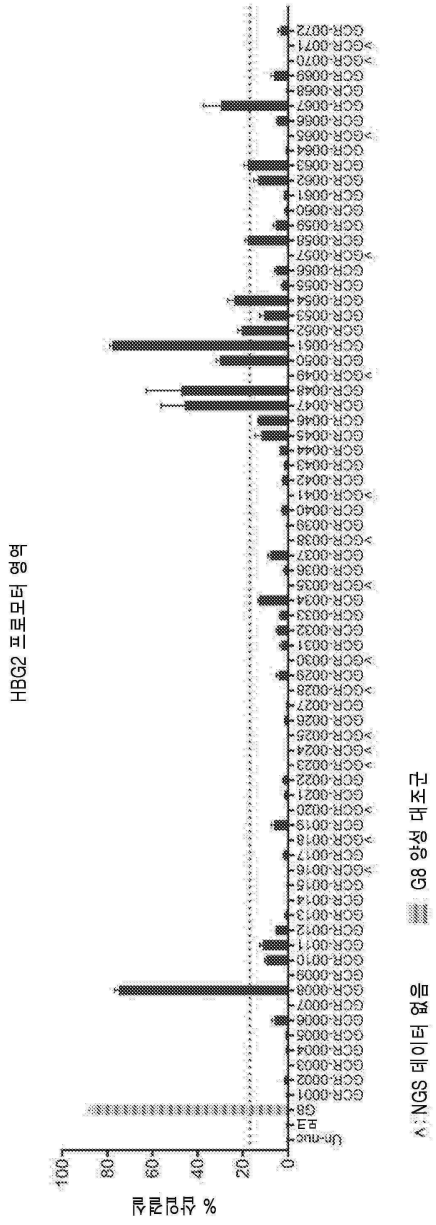
도면1


HbF 염색 7일차
% HbF+ 세포(모든 세포)
40
20
0
Un-nuc
E3
G8
GCR-0001
GCR-0002
GCR-0003
GCR-0004
GCR-0005
GCR-0006
GCR-0007
GCR-0008
GCR-0009
GCR-0010
GCR-0011
GCR-0012
GCR-0013
GCR-0014
GCR-0015
GCR-0016
GCR-0017
GCR-0018
GCR-0019
GCR-0020
GCR-0021
GCR-0022
GCR-0023
GCR-0024
GCR-0025
GCR-0026
GCR-0027
GCR-0028
GCR-0029
GCR-0030
GCR-0031
GCR-0032
GCR-0033
GCR-0034
GCR-0035
GCR-0036
GCR-0037
GCR-0038
GCR-0039
GCR-0040
GCR-0041
GCR-0042
GCR-0043
GCR-0044
GCR-0045
GCR-0046
GCR-0047
GCR-0048
GCR-0049
GCR-0050
GCR-0051
GCR-0052
GCR-0053
GCR-0054
GCR-0055
GCR-0056
GCR-0057
GCR-0058
GCR-0059
GCR-0060
GCR-0061
GCR-0062
GCR-0063
GCR-0064
GCR-0065
GCR-0066
GCR-0067
GCR-0068
GCR-0069
GCR-0070
GCR-0071
GCR-0072
-: HbF 유도 < 모크
% HbF 유도 > 17% 및 % 삽입결실 및 > 17%(HBF1, HBG2 또는 둘 모두에 대해)
% HbF 유도 > 17% 및 % 삽입결실 < 17%
% HbF 유도 < 17% 및 % 삽입결실 >> 17%
% HbF 유도 < 17%
G8 양성 대조군

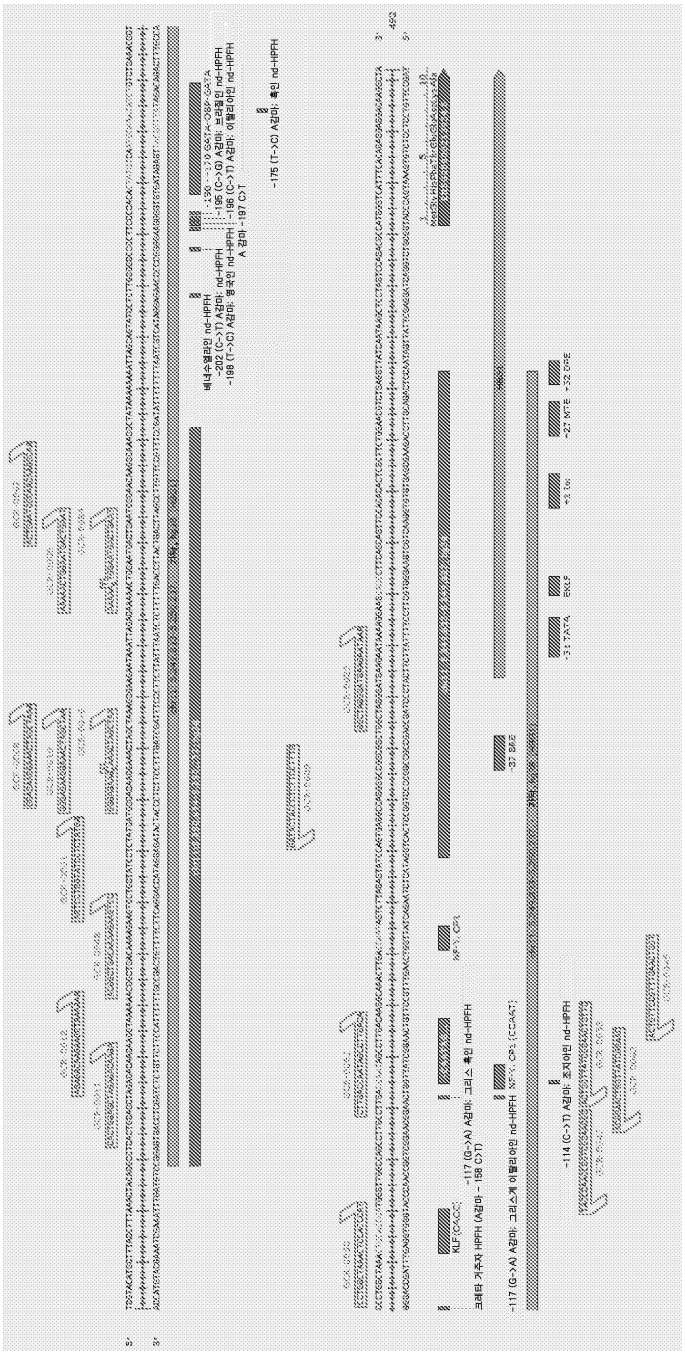
도면2


HBG1 프로모터 영역
편집효율 %
100
80
60
40
20
0
Unedited
G8
GCR-0001
GCR-0002
GCR-0003
GCR-0004
GCR-0005
GCR-0006
GCR-0007
GCR-0008
GCR-0009
GCR-0010
GCR-0011
GCR-0012
GCR-0013
GCR-0014
GCR-0015
GCR-0016
GCR-0017
GCR-0018
GCR-0019
GCR-0020
GCR-0021
GCR-0022
GCR-0023
GCR-0024
GCR-0025
GCR-0026
GCR-0027
GCR-0028
GCR-0029
GCR-0030
GCR-0031
GCR-0032
GCR-0033
GCR-0034
GCR-0035
GCR-0036
GCR-0037
GCR-0038
GCR-0039
GCR-0040
GCR-0041
GCR-0042
GCR-0043
GCR-0044
GCR-0045
GCR-0046
GCR-0047
GCR-0048
GCR-0049
GCR-0050
GCR-0051
GCR-0052
GCR-0053
GCR-0054
GCR-0055
GCR-0056
GCR-0057
GCR-0058
GCR-0059
GCR-0060
GCR-0061
GCR-0062
GCR-0063
GCR-0064
GCR-0065
GCR-0066
GCR-0067
GCR-0068
GCR-0069
GCR-0070
GCR-0071
GCR-0072
* : NGS 데이터 없음
G8 양성 대조군

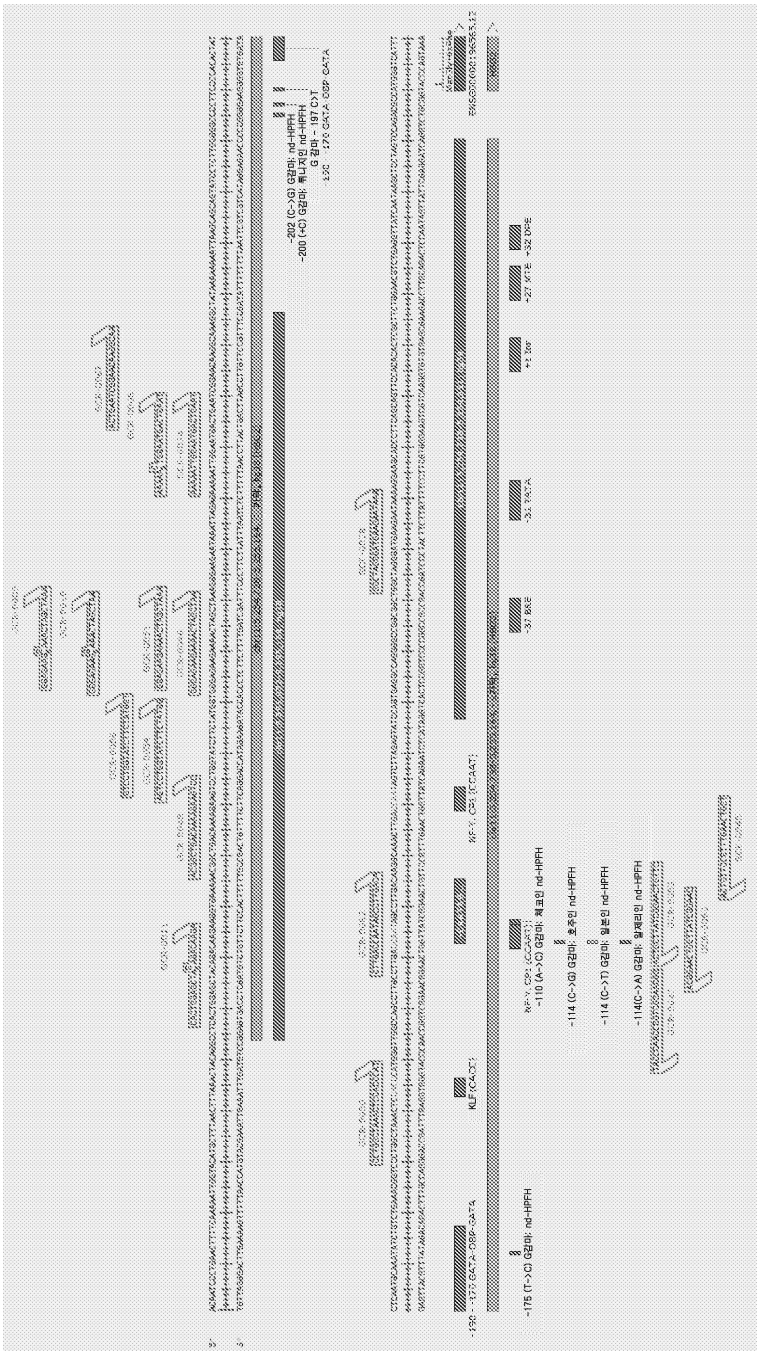
도면3


HBG2 프로모터 영역
% 상동분석
100
80
60
40
20
0
Un-nuc
G8
Ea
GCR-0001
GCR-0002
GCR-0003
GCR-0004
GCR-0005
GCR-0006
GCR-0007
GCR-0008
GCR-0009
GCR-0010
GCR-0011
GCR-0012
GCR-0013
GCR-0014
GCR-0015
>GCR-0016
>GCR-0018
GCR-0017
GCR-0019
>GCR-0020
GCR-0021
GCR-0022
>GCR-0023
>GCR-0024
>GCR-0025
GCR-0026
GCR-0027
>GCR-0028
GCR-0029
>GCR-0030
GCR-0031
GCR-0032
GCR-0033
GCR-0034
>GCR-0035
GCR-0036
GCR-0037
>GCR-0038
GCR-0039
GCR-0040
>GCR-0041
GCR-0042
GCR-0043
GCR-0044
GCR-0045
GCR-0046
GCR-0047
GCR-0048
>GCR-0049
GCR-0050
GCR-0051
GCR-0052
GCR-0053
GCR-0054
GCR-0055
GCR-0056
>GCR-0057
GCR-0058
GCR-0059
GCR-0060
GCR-0061
GCR-0062
GCR-0063
GCR-0064
>GCR-0065
GCR-0066
GCR-0067
GCR-0068
GCR-0069
>GCR-0070
>GCR-0071
GCR-0072
^ : NGS 데이터 없음
G8 양성 대조군

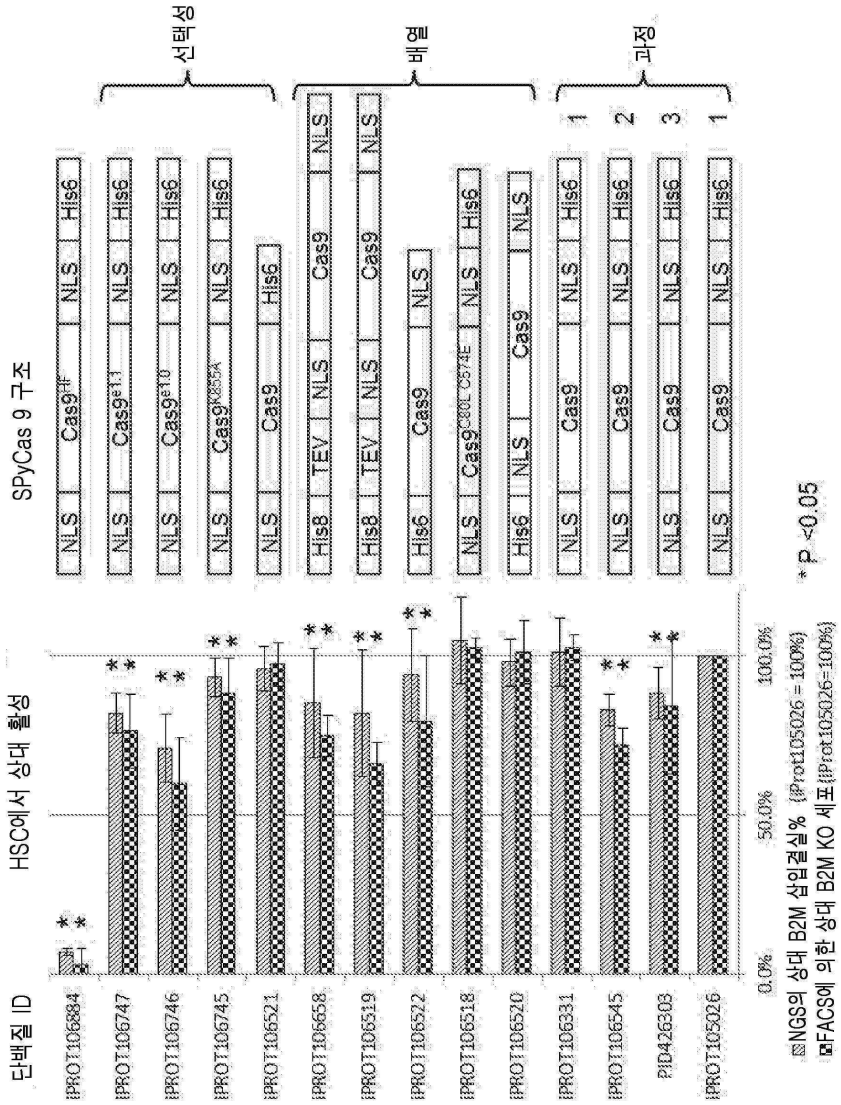
도면4



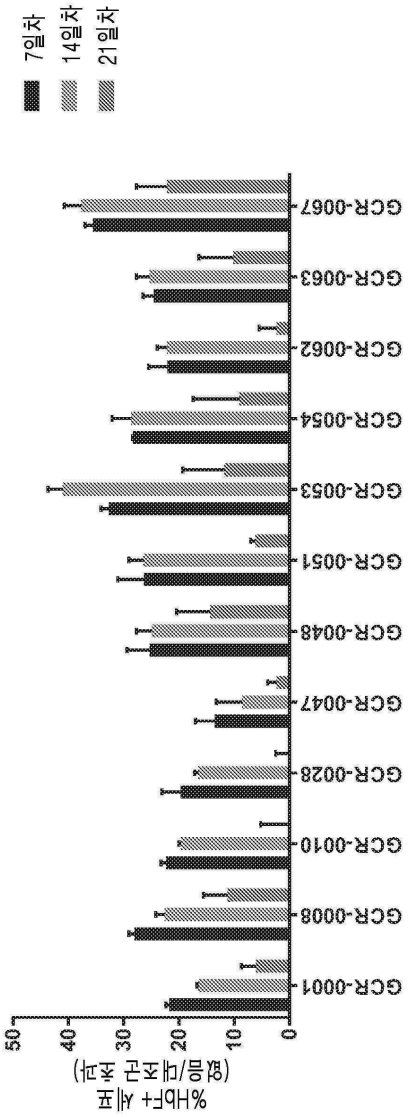
도면5



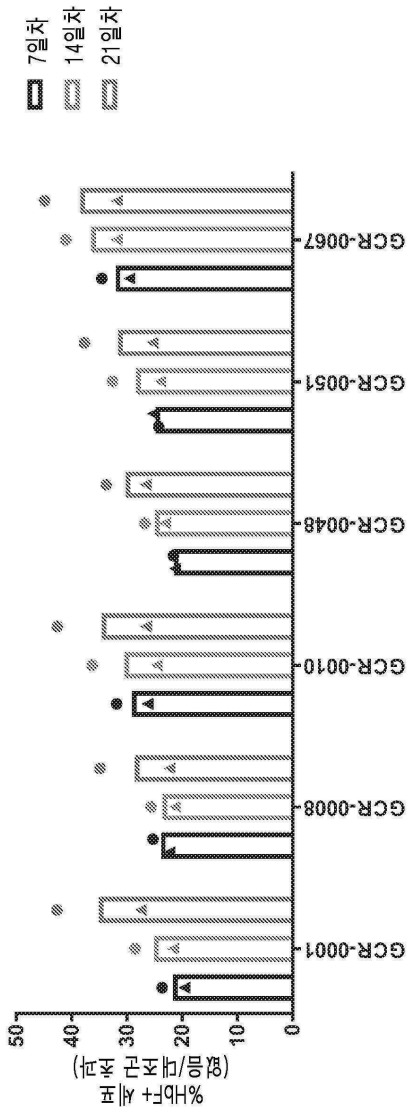
도면6


선택성
융합
과정
SPyCas 9 구조
HSC에서 상대 활성
단백질 ID
iPROT106884
iPROT106747
iPROT106746
iPROT106745
iPROT106521
iPROT106658
iPROT106519
iPROT106522
iPROT106518
iPROT106520
iPROT106331
iPROT106545
PID426303
iPROT105026
0.0%
50.0%
100.0%
▨NGS의 상대 B2M 삽입결실% (iProt105026 = 100%)
■FACS에 의한 상대 B2M KO 세포(iProt105026=100%)
* P <0.05

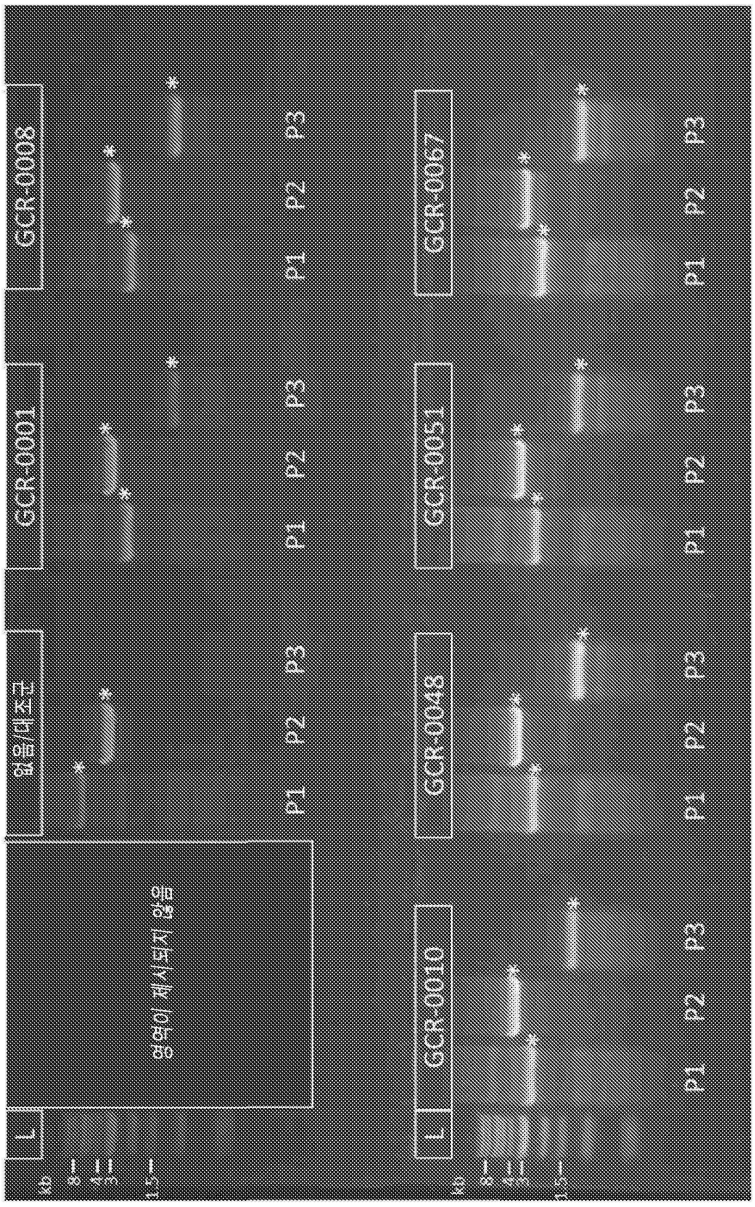
도면7


7일차
14일차
21일차
GCR-0067
GCR-0063
GCR-0062
GCR-0054
GCR-0053
GCR-0051
GCR-0048
GCR-0047
GCR-0028
GCR-0010
GCR-0008
GCR-0001
0
10
20
30
40
50
%HbF+ 세포
(없음/대조군 초과)

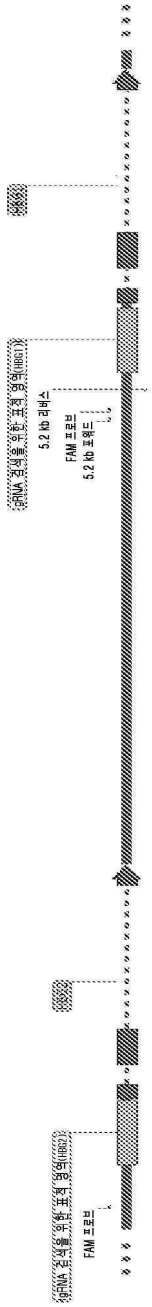
도면8


7일차
14일차
21일차
GCR-0067
GCR-0051
GCR-0048
GCR-0010
GCR-0008
GCR-0001
0
10
20
30
40
50
%HbF+ 세포
(없음/대조군 추출)

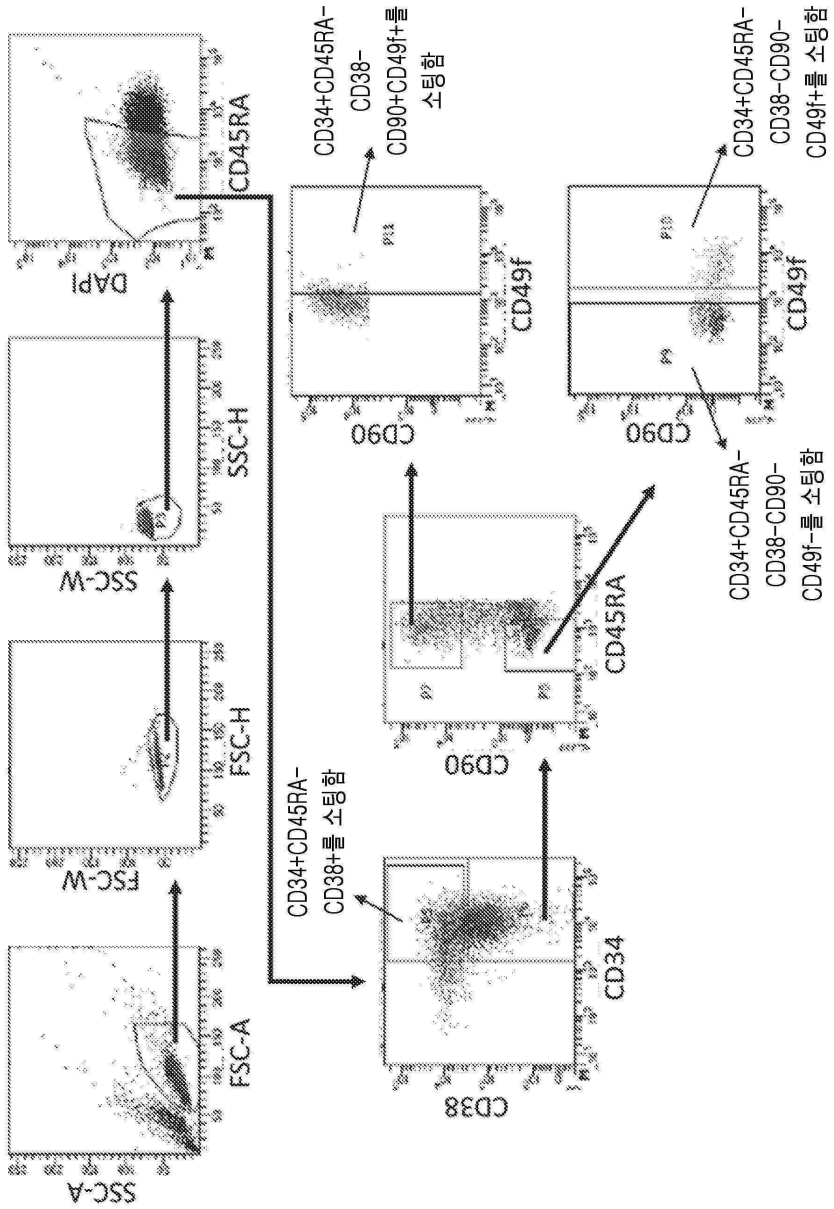
도면9



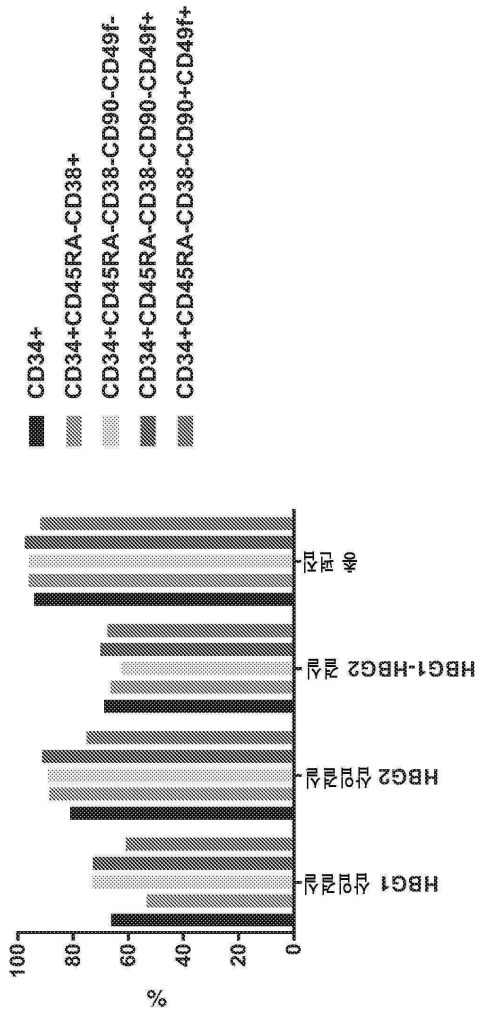
도면10


gRNA 결합 부위 표적 DNA(HBG1)
5.2 kb 리버스
FAM 프로브
5.2 kb 포워드
gRNA 결합 부위 표적 DNA(HBG2)
FAM 프로브

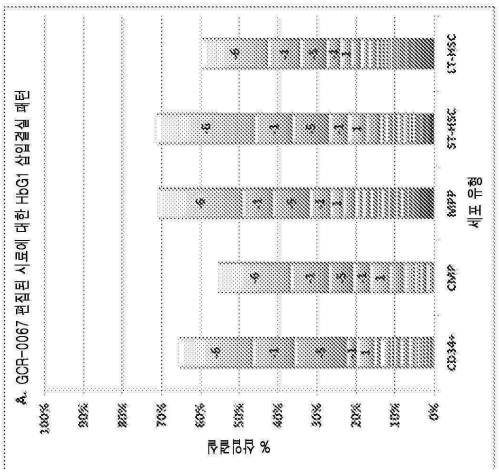
도면11


SSC-A
FSC-A
FSC-W
FSC-H
SSC-W
SSC-H
DAPI
CD45RA
CD38
CD34
CD34+CD45RA-
CD38+를 소팅함
CD90
CD45RA
CD34+CD45RA-
CD38-
CD90+CD49f+를
소팅함
CD49f
CD34+CD45RA-
CD38-CD90-
CD49f-를 소팅함
CD34+CD45RA-
CD38-CD90-
CD49f+를 소팅함

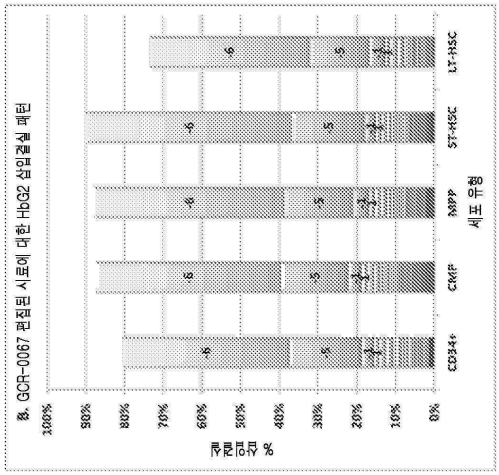
도면12


CD34+
CD34+CD45RA-CD38+
CD34+CD45RA-CD38-CD90-CD49f-
CD34+CD45RA-CD38-CD90-CD49f+
CD34+CD45RA-CD38-CD90+CD49f+
HBG1 삽입결실
HBG2 삽입결실
HBG1-HBG2 결실
야생형
100
80
60
40
20
0
%

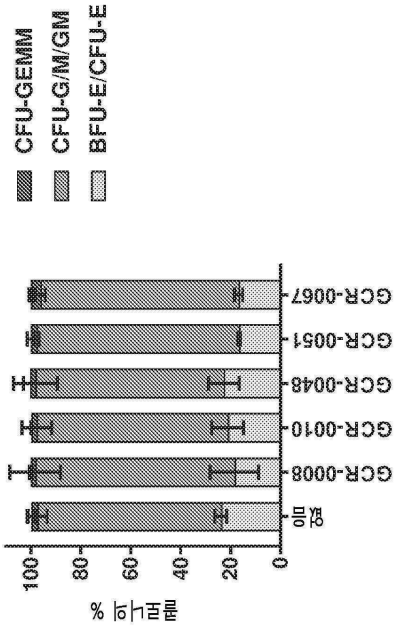
도면13a


A. GCR-0067 편집된 시료에 대한 HbG1 삽입결실 패턴
% 삽입결실
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
CD34+
CMP
MPP
ST-HSC
LT-HSC
세포 유형

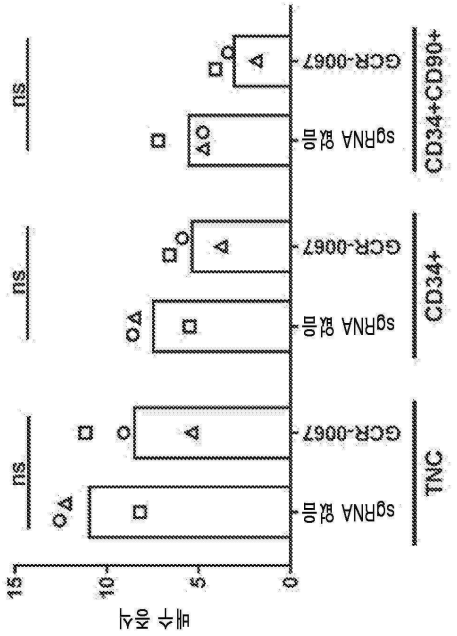
도면13b


8. GCR-0067 편집된 시료에 대한 HbG2 삽입결실 패턴
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% 대립유전자
CD34+
CMP
MPP
ST-HSC
LT-HSC
세포 유형

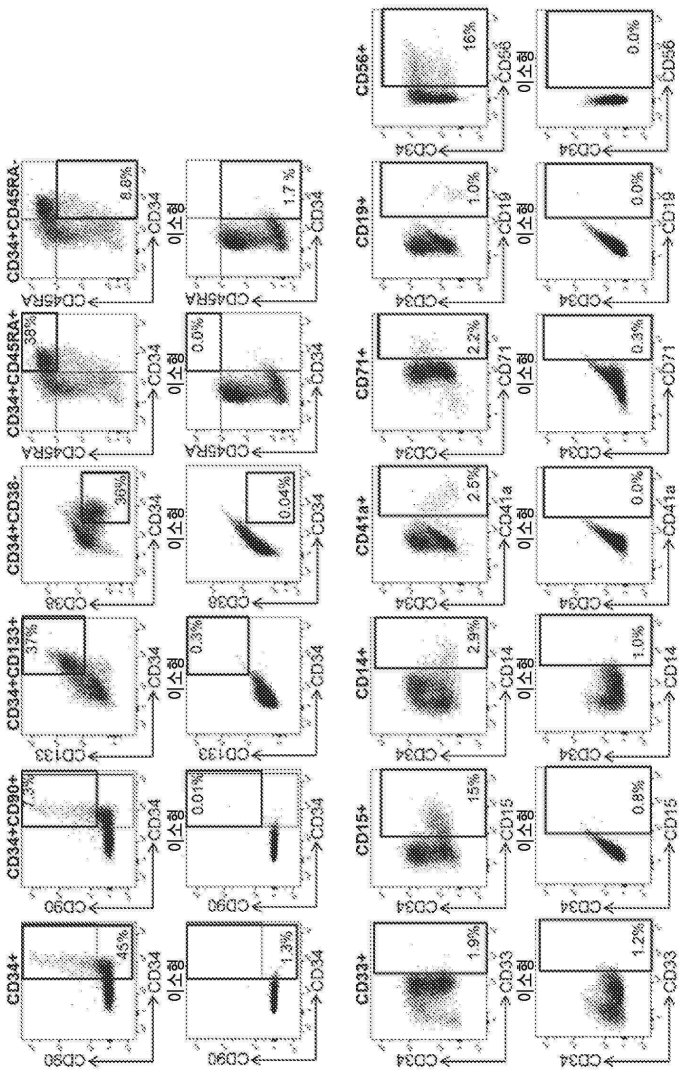
도면14


CFU-GEMM
CFU-G/M/GM
BFU-E/CFU-E
100
80
60
40
20
0
콜로니의 %
없음
GCR-0008
GCR-0010
GCR-0048
GCR-0051
GCR-0067

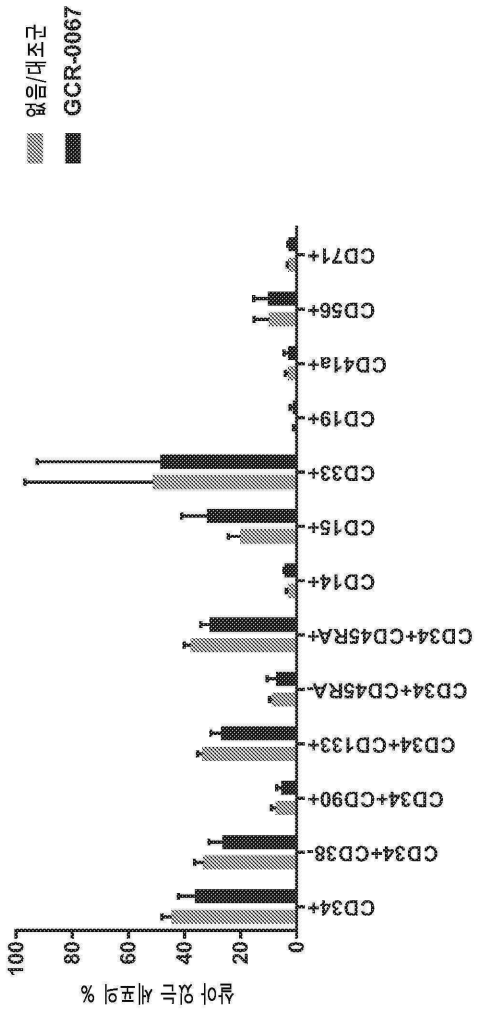
도면15


빈도 (%)
15
10
5
0
ns
ns
ns
sgRNA 없음
GCR-0067
sgRNA 없음
GCR-0067
sgRNA 없음
GCR-0067
TNC
CD34+
CD34+CD90+

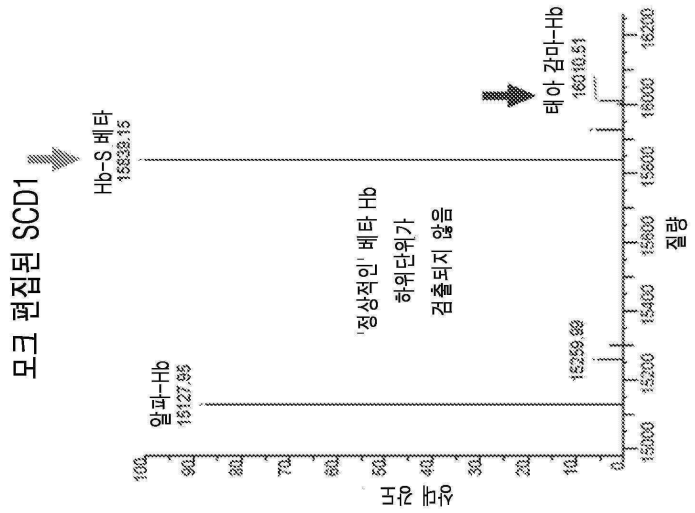
도면16


CD34+
CD34+CD90+
CD34+CD123+
CD34+CD38-
CD34+CD45RA+
CD34+CD45RA-
CD33+
CD15+
CD14+
CD41a+
CD71+
CD19+
CD56+
43%
1.3%
7.3%
0.01%
37%
0.3%
36%
0.04%
38%
0.0%
8.8%
1.7 %
1.9%
15%
2.9%
1.0%
2.5%
0.0%
2.2%
0.3%
1.0%
0.0%
16%
0.0%
1.2%
0.8%

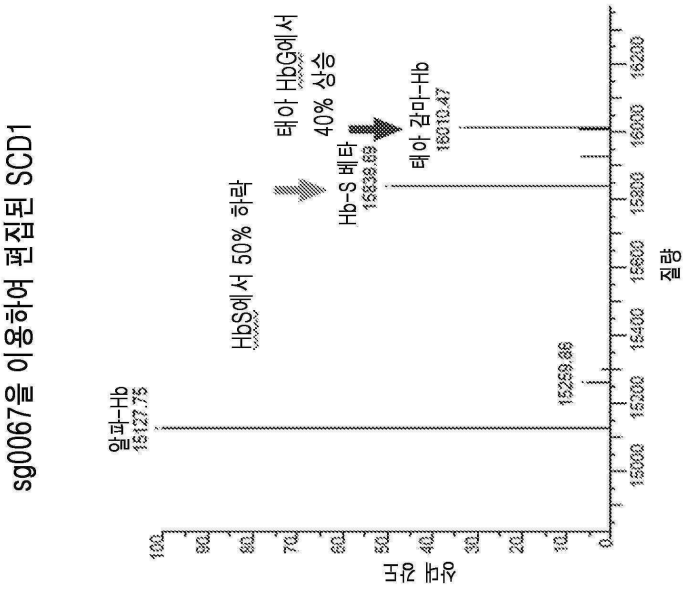
도면17


없음/대조군
GCR-0067
살아 있는 세포의 %
100
80
60
40
20
0
CD34+
CD34+CD38-
CD34+CD90+
CD34+CD133+
CD34+CD45RA-
CD34+CD45RA+
CD14+
CD15+
CD33+
CD19+
CD41a+
CD56+
CD71+

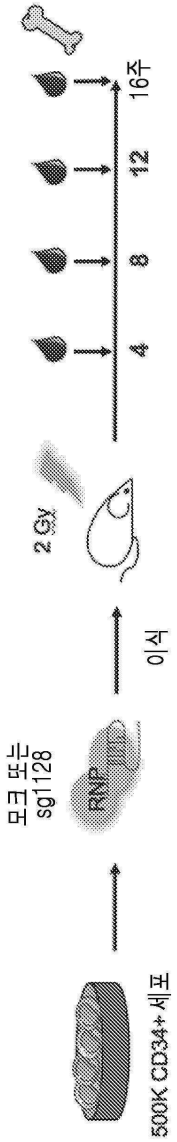
도면18a


모크 편집된 SCD1
Hb-S 베타
15838.15
알파-Hb
16127.95
'정상적인' 베타 Hb
하위단위가
검출되지 않음
15259.99
태아 감마-Hb
16010.51
상대 강도
100
90
80
70
60
50
40
30
20
10
0
15000
15200
15400
15600
15800
16000
16200
질량

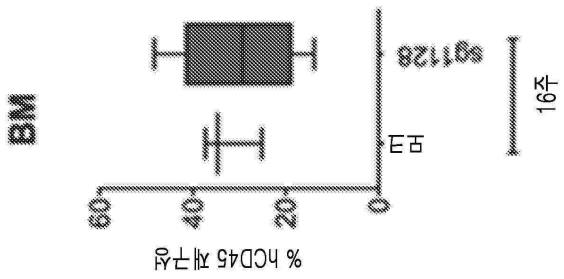
도면18b


sg0067을 이용하여 편집된 SCD1
알파-Hb
15127.75
HbS에서 50% 하락
태아 HbG에서
40% 상승
Hb-S 베타
15839.69
태아 감마-Hb
16010.47
15259.86
상대 강도
질량
100
90
80
70
60
50
40
30
20
10
0
15000
15200
15400
15600
15800
16000
16200

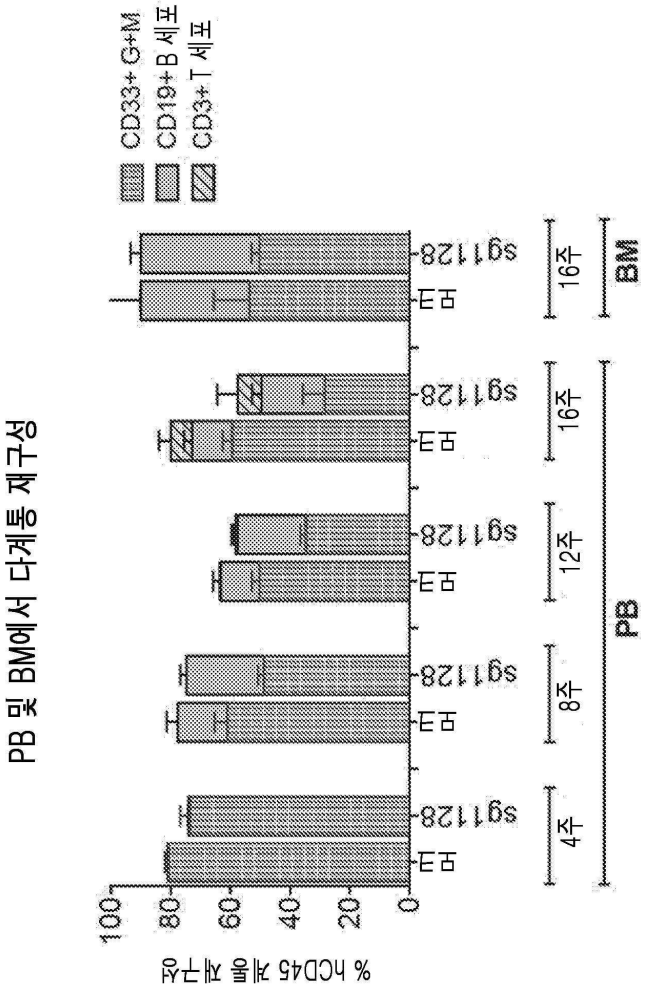
도면19


500K CD34+ 세포
모크 또는
sg1128
RNP
이식
2 Gy
4
8
12
16주

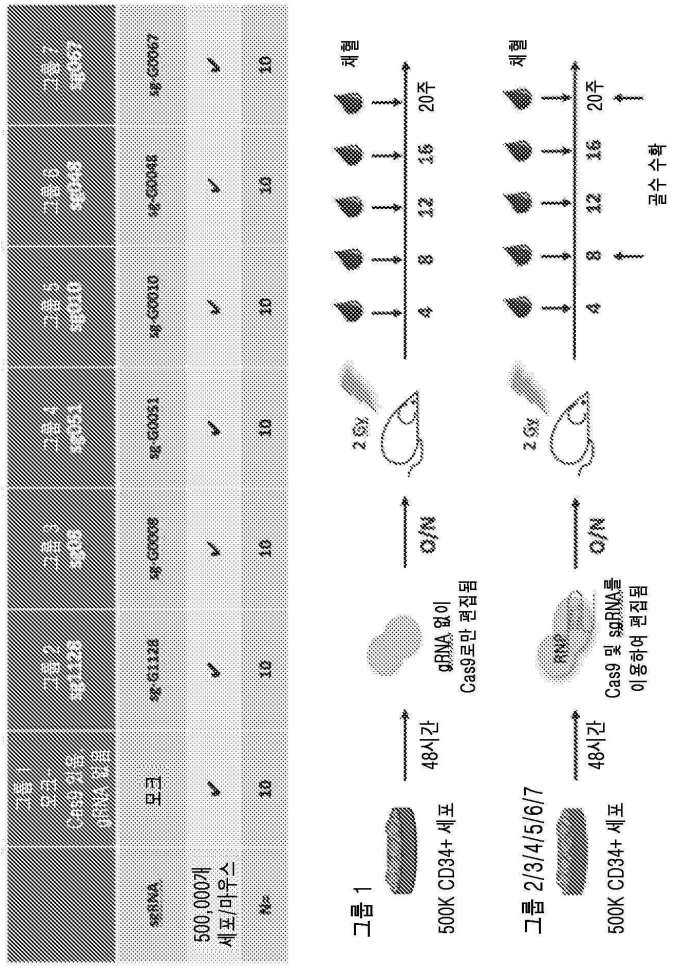
도면20


BM
60
40
20
0
% hCD45 세포성
모크
sg1128
16주

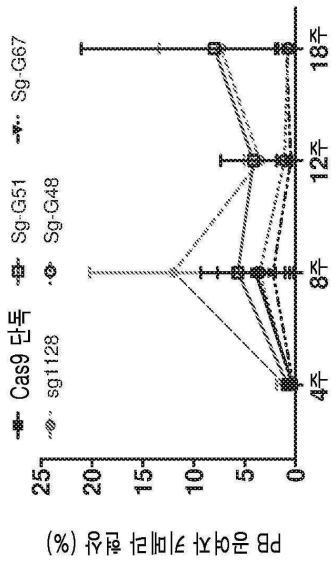
도면21


PB 및 BM에서 다계통 재구성
CD33+ G+M
CD19+ B 세포
CD3+ T 세포
% hCD45 계통 재구성
100
80
60
40
20
0
대조
sg1128
4주
8주
12주
16주
PB
BM

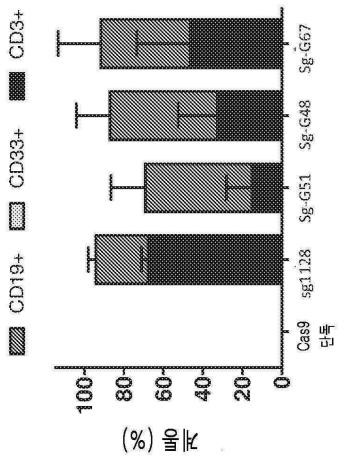
도면22


그룹 1
그룹 2
그룹 3
그룹 4
그룹 5
그룹 6
그룹 7
sgRNA
모크
sg-G1128
sg-G0008
sg-G0051
sg-G0010
sg-G0048
sg-G0067
500,000개 세포/마우스
10
그룹 1
48시간
500K CD34+ 세포
gRNA 없이 Cas9로만 편집됨
2 Gy
체혈
4
8
12
16
20주
그룹 2/3/4/5/6/7
RNP
Cas9 및 sgRNA를 이용하여 편집됨
골수 수확

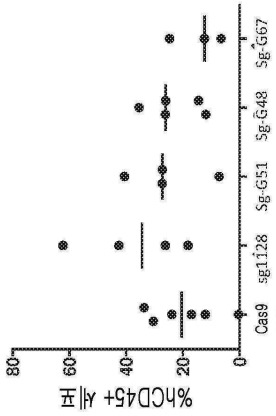
도면23a


Cas9 단독
sg1128
Sg-G51
Sg-G48
Sg-G67
PB 유전자 기물라 편집 (%)
25
20
15
10
5
0
4주
8주
12주
18주

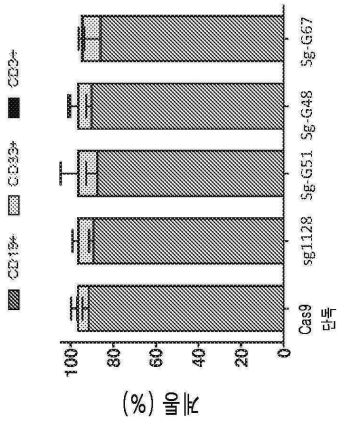
도면23b



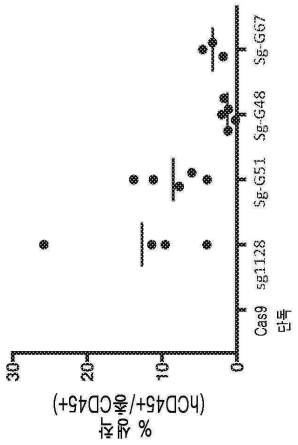
도면23c



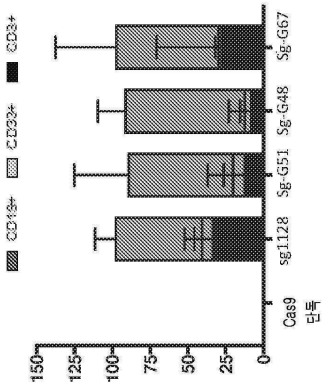
도면23d



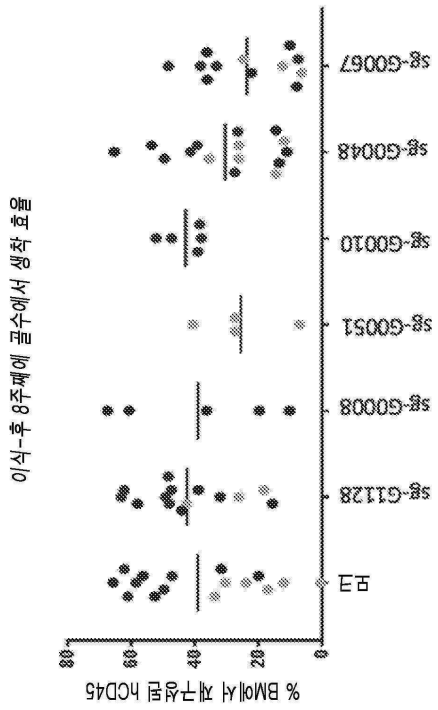
도면23e



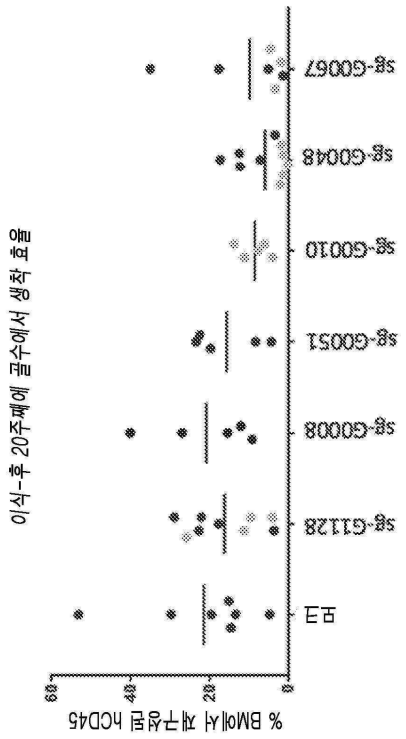
도면23f



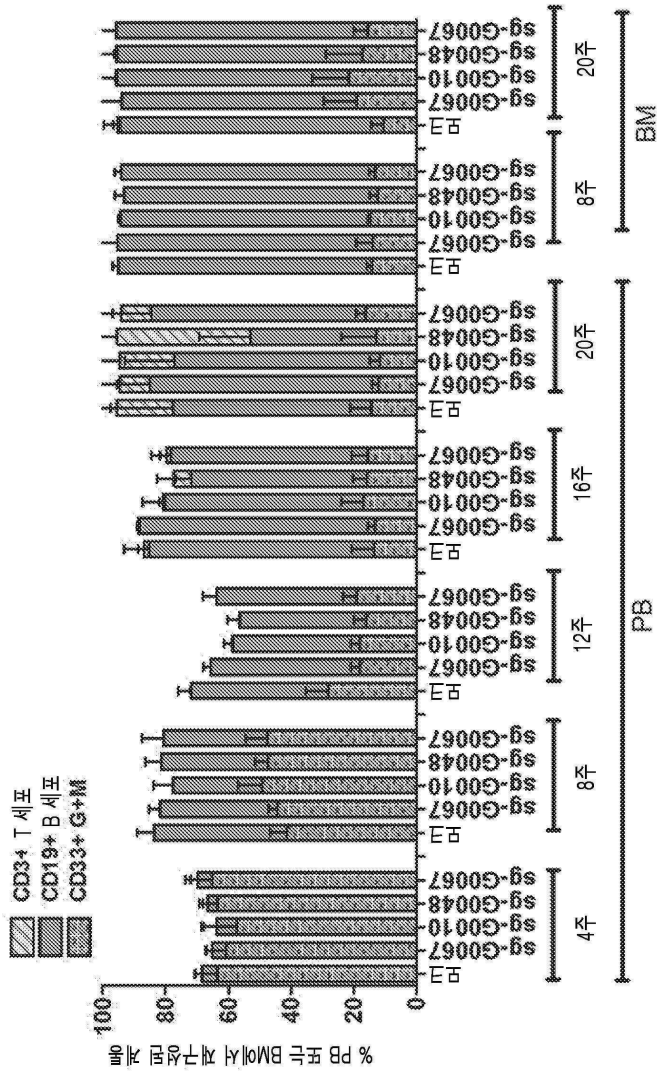
도면24a



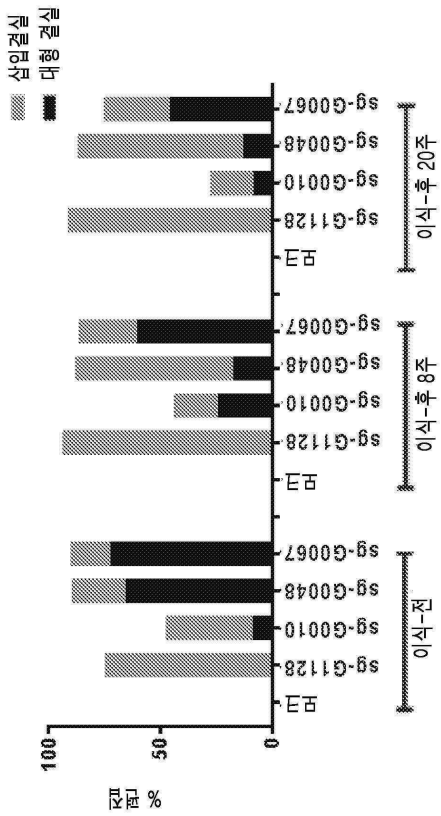
도면24b


% BM에서 재구성된 hCD45
이식-후 20주째에 골수에서 생착 효율
0
20
40
60
모크
sg-G1128
sg-G0008
sg-G0051
sg-G0010
sg-G0048
sg-G0067

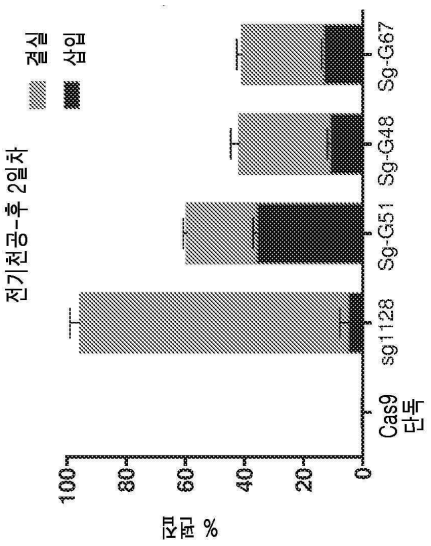
도면25


% PB 또는 BM에서 재구성된 계통
100
80
60
40
20
0
CD3+ T 세포
CD19+ B 세포
CD33+ G+M
4주
8주
12주
16주
20주
20주
8주
PB
BM
모크
sg-G0067
sg-G0010
sg-G0048
sg-G0067

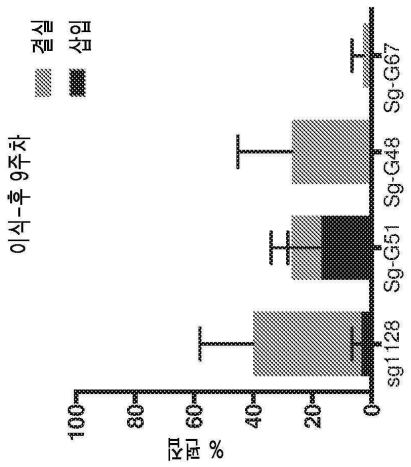
도면26


삽입결실
대형 결실
sg-G0067
sg-G0048
sg-G0010
sg-G1128
무처
이식-후 20주
이식-후 8주
이식-전
100
50
0
% 편집

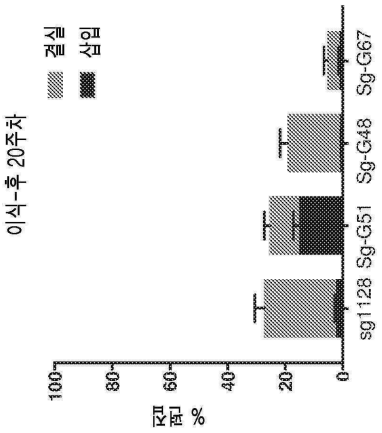
도면27a


전기천공-후 2일차
결실
삽입
Cas9 단독
sg1128
Sg-G51
Sg-G48
Sg-G67
100
80
60
40
20
0
% 편집

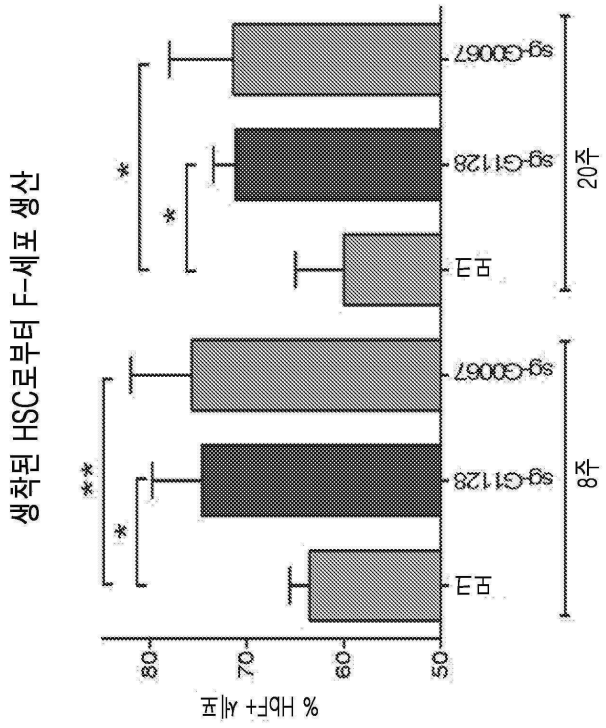
도면27b


이식-후 9주차
결실
삽입
편집 %
100
80
60
40
20
0
sg1128
Sg-G51
Sg-G48
Sg-G67

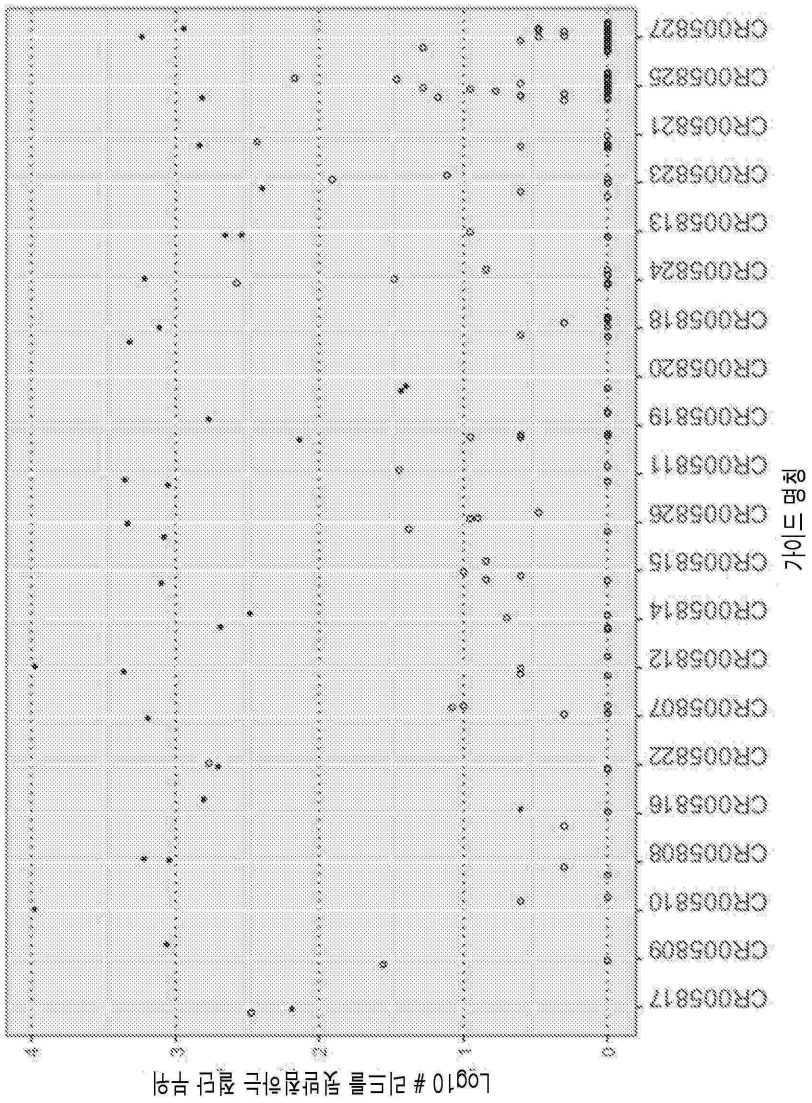
도면27c


이식-후 20주차
결실
삽입
편집 %
100
80
60
40
20
0
sg1128
Sg-G51
Sg-G48
Sg-G67

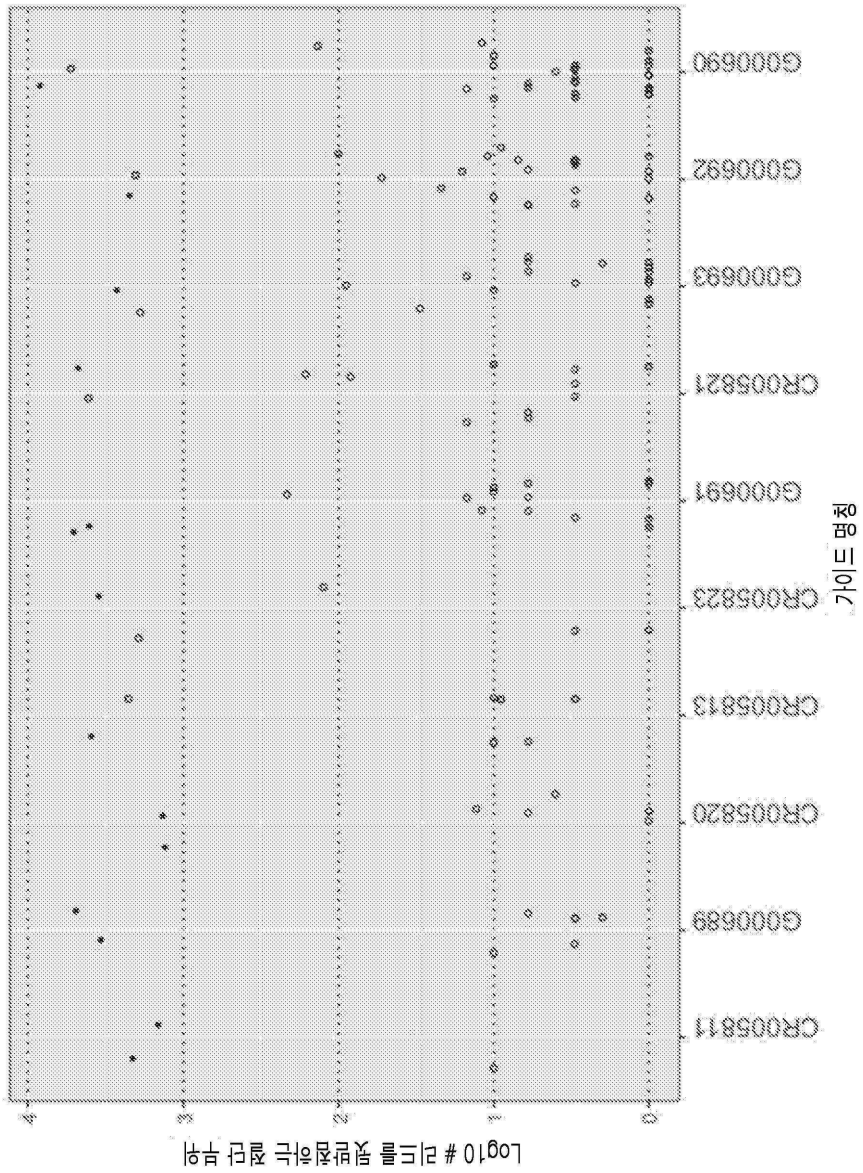
도면28



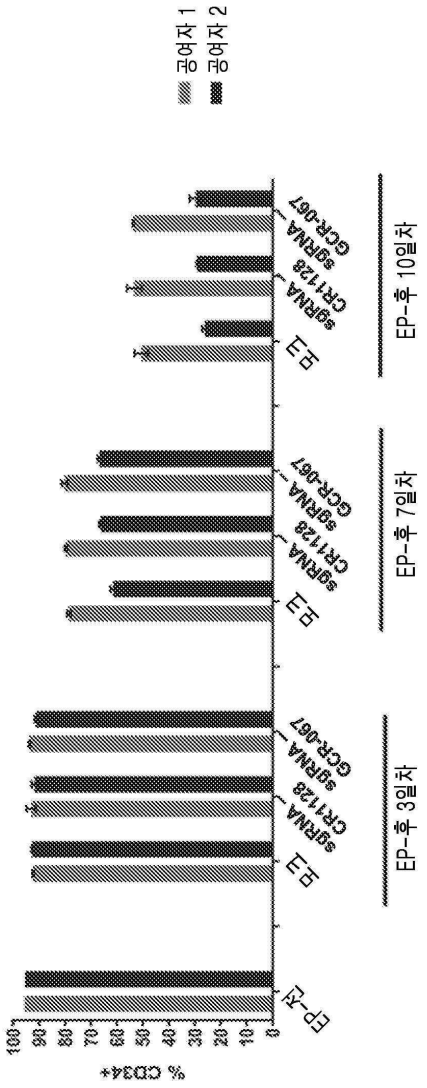
도면29


CR0058I7
CR005809
CR005810
CR005808
CR005816
CR005822
CR005807
CR005812
CR005814
CR005815
CR005826
CR005811
CR005819
CR005820
CR005818
CR005824
CR005813
CR005823
CR005821
CR005825
CR005827
유전자 멤칭
Log10 # 리드를 뒷받침하는 절단 부위
0
1
2
3
4

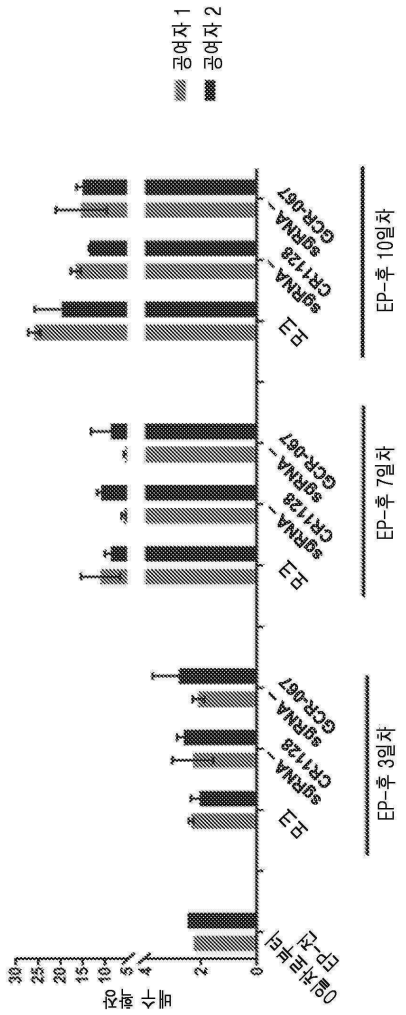
도면30


가이드 명칭
G0006890
G0006892
G0006893
CR0058271
G0006891
CR0058233
CR0058133
CR0058200
G0006889
CR0058111
Log10 # 리드를 뒷받침하는 절단 부위

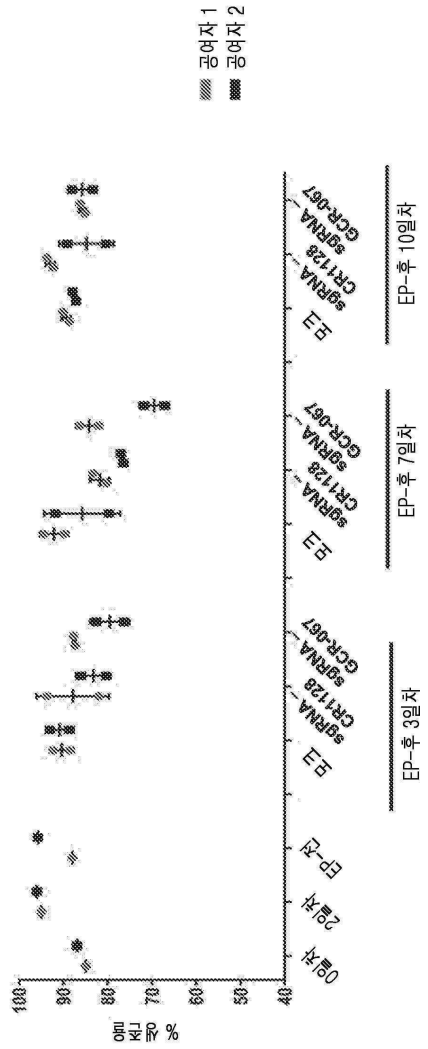
도면31a


공여자 1
공여자 2
% CD34+
100
90
80
70
60
50
40
30
20
10
0
EP-전
EP-후 3일차
EP-후 7일차
EP-후 10일차
모의
sgRNA CR1128
sgRNA GCR-067

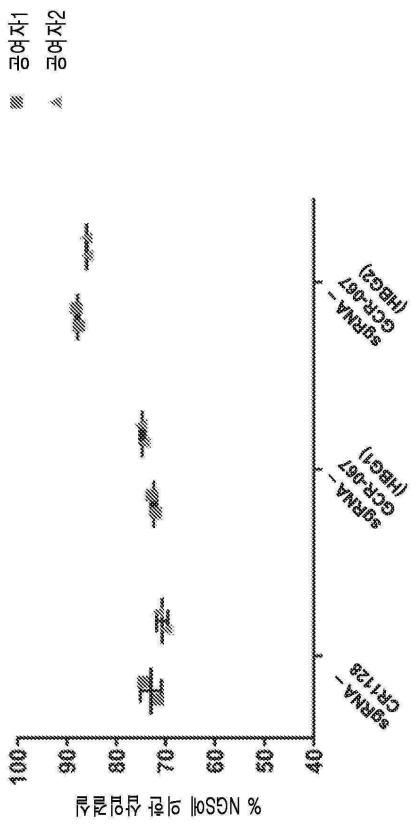
도면31b


공여자 1
공여자 2
EP-후 10일차
EP-후 7일차
EP-후 3일차
sgRNA-GCR-067
sgRNA-CR1128
모크
오염되지 않은 EP-군
빈도 수 확인
30
25
20
15
10
5
4
2
0

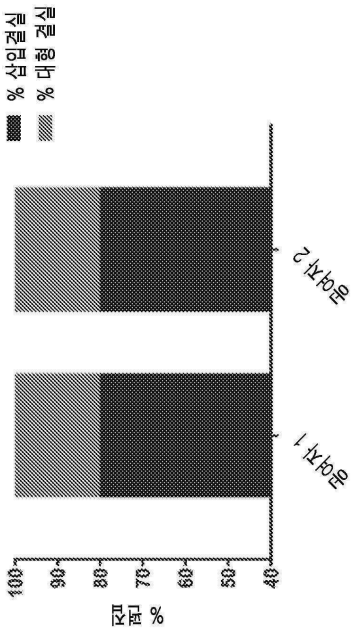
도면31c


공여자 1
공여자 2
%
100
90
80
70
60
50
40
0일차
2일차
EP-전
모두
sgRNA_CR1128_sgRNA_GCR-067
EP-후 3일차
EP-후 7일차
EP-후 10일차

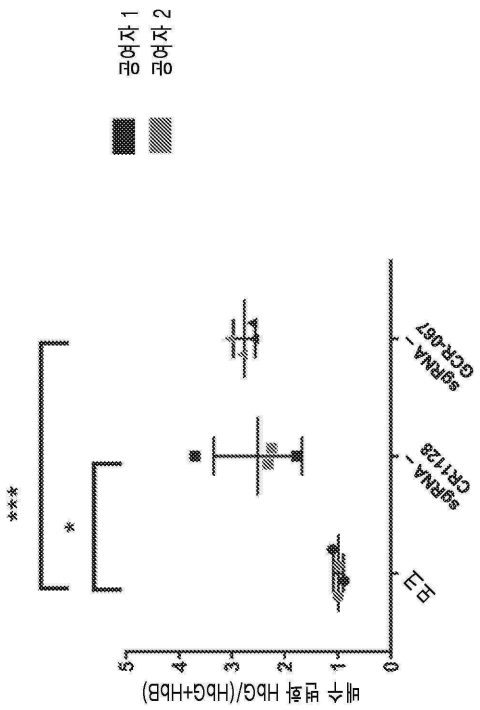
도면32a


공여자1
공여자2
% NGS에 의한 삽입결실
100
90
80
70
60
50
40
sgRNA-CR1128
sgRNA-GCR-067 (HBG1)
sgRNA-GCR-067 (HBG2)

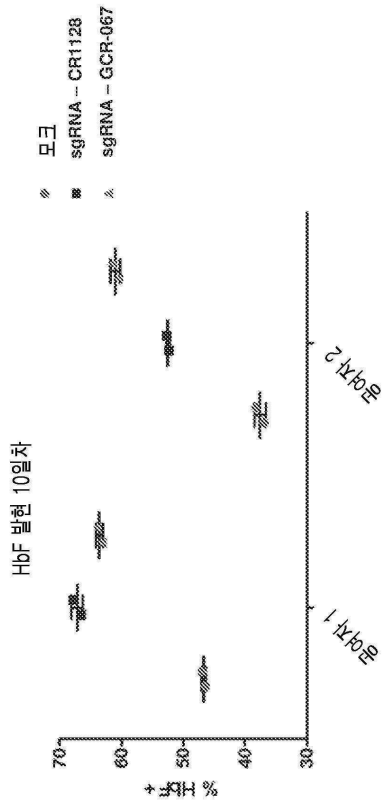
도면32b


% 삽입결실
% 대립형질
빈도 %
100
90
80
70
60
50
40
공여자 1
공여자 2

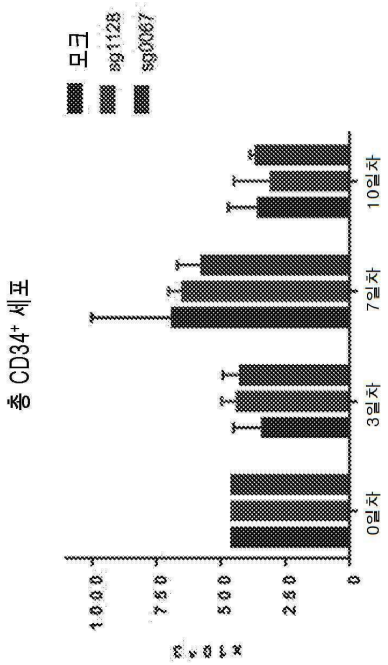
도면33


공여자 1
공여자 2
세포 분화 HbG/(HbG+HbB)
0
1
2
3
4
5
*
***
대조
sgRNA – CR1128
sgRNA – GCR-067

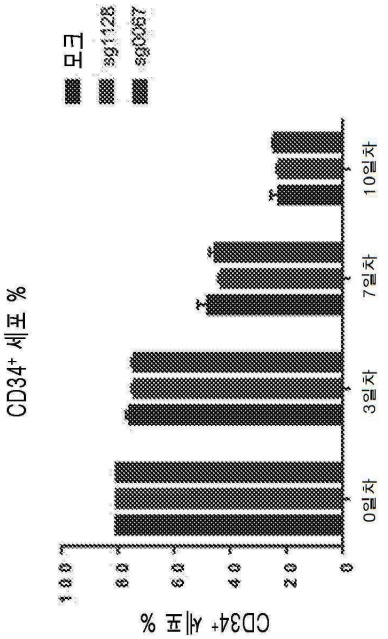
도면34


HbF 발현 10일차
% HbF+
30
40
50
60
70
모크
sgRNA – CR1128
sgRNA – GCR-067
공여자 1
공여자 2

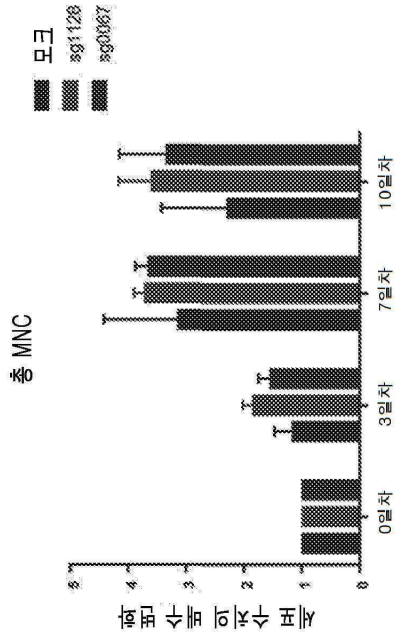
도면35a


모크
sg1128
sg0067
총 CD34+ 세포
1000
750
500
250
0
0일차
3일차
7일차
10일차

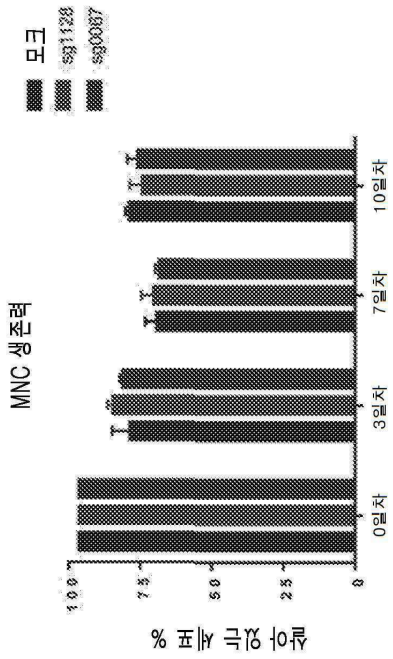
도면35b


모크
sg1128
sg0067
CD34+ 세포 %
100
80
60
40
20
0
0일차
3일차
7일차
10일차

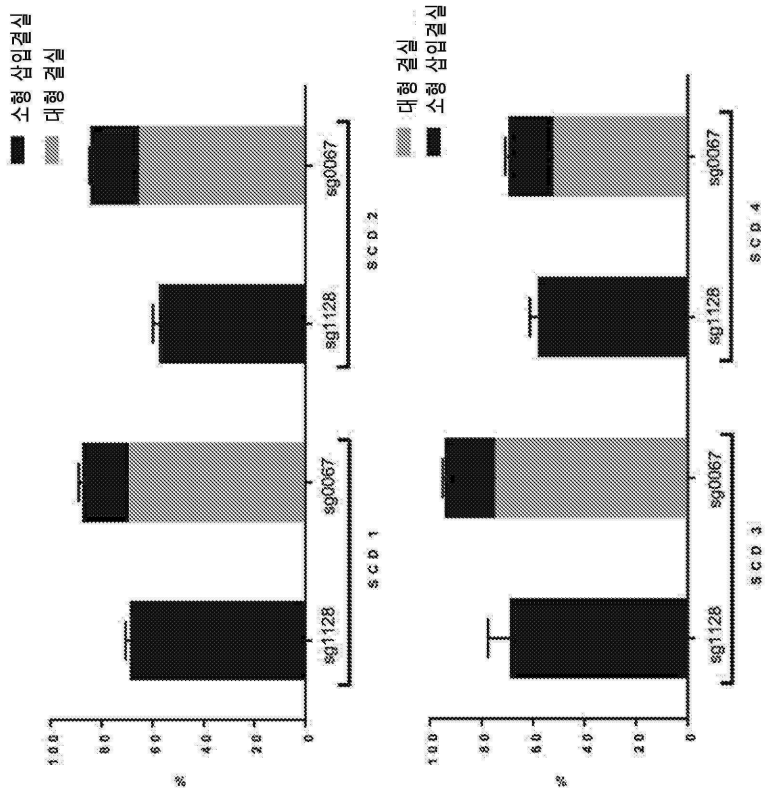
도면35c


모크
sg11128
sg00087
총 MNC
세포 수치의 배수 변화
0일차
3일차
7일차
10일차

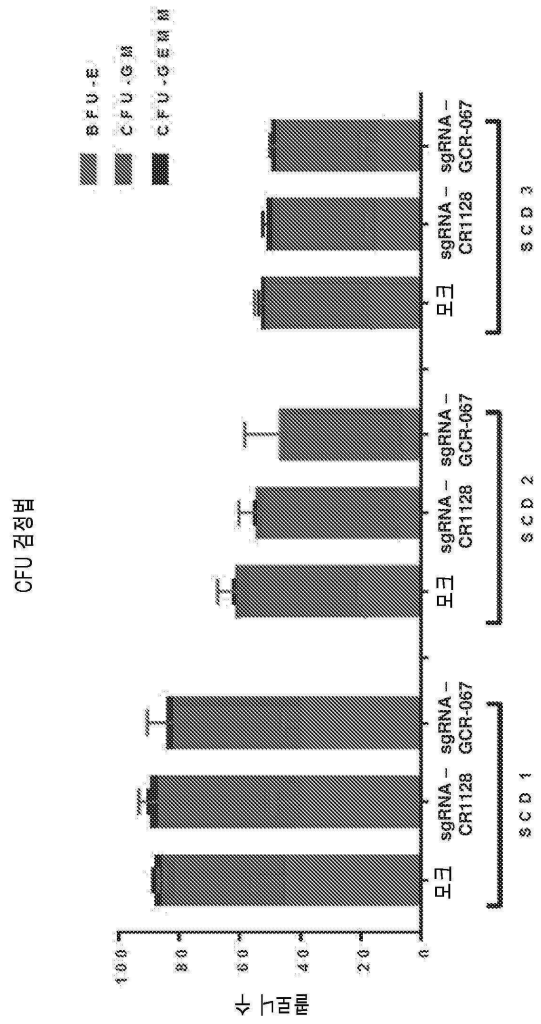
도면35d


모크
sg11128
sg00087
MNC 생존력
살아 있는 세포 %
100
75
50
25
0
0일차
3일차
7일차
10일차

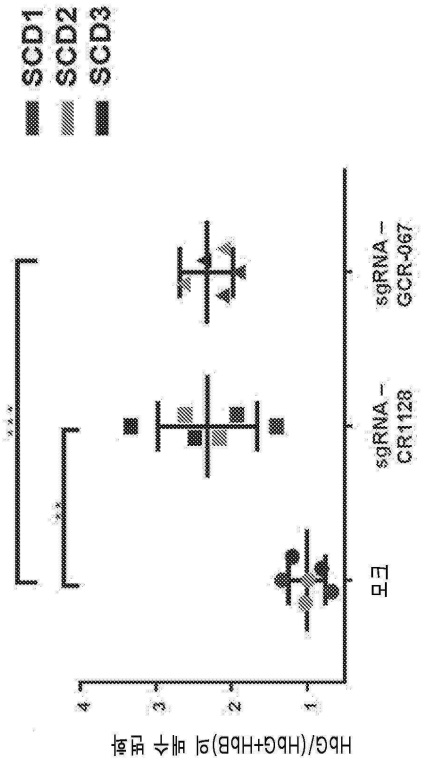
도면36


소형 삽입결실
대형 결실
sg1128
sg0067
sgD1
sgD2
sgD3
sgD4
%
100
80
60
40
20
0

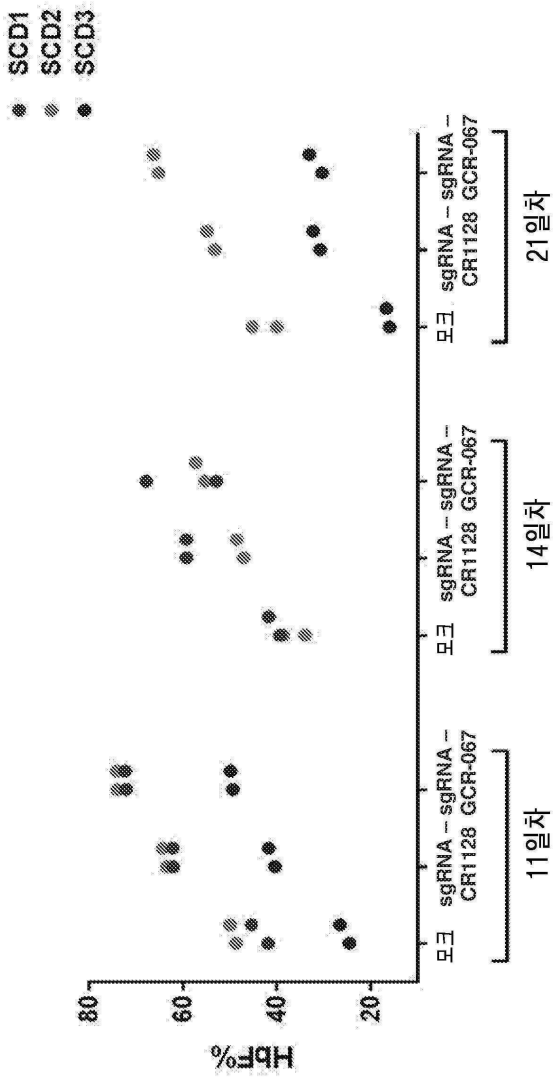
도면37


CFU 검정법
BFU-E
CFU-GM
CFU-GEMM
콜로니 수
100
80
60
40
20
0
모크
sgRNA – CR1128
sgRNA – GCR-067
SCD 1
SCD 2
SCD 3

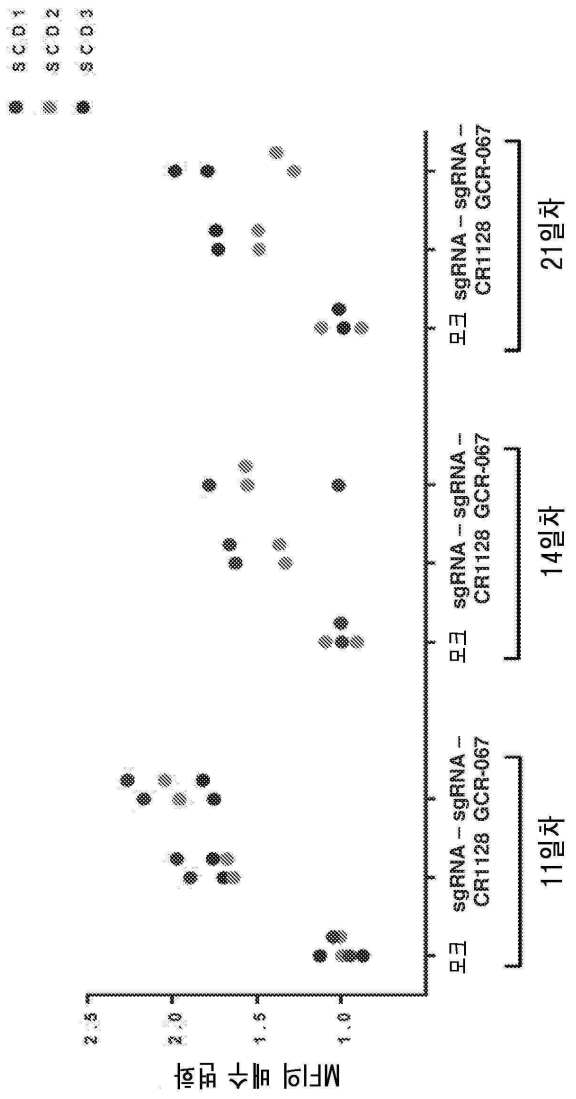
도면38


SCD1
SCD2
SCD3
HbG/(HbG+HbB)의 백분 변화
4
3
2
1
모크
sgRNA – CR1128
sgRNA – GCR-067

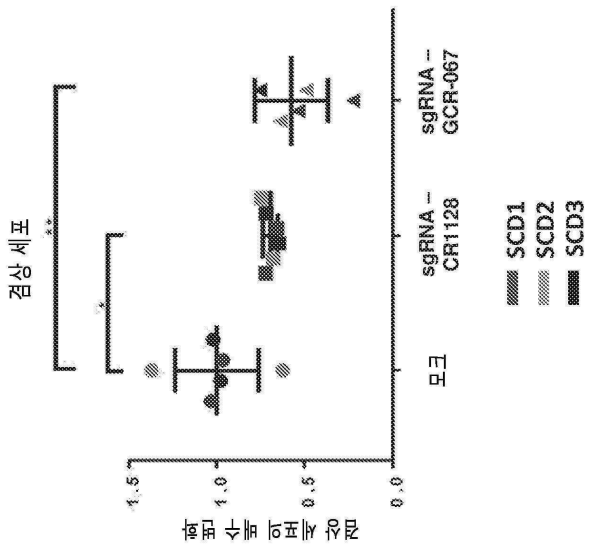
도면39


SCD1
SCD2
SCD3
HbF%
80
60
40
20
모크
sgRNA – sgRNA – CR1128 GCR-067
11일차
모크
sgRNA – sgRNA – CR1128 GCR-067
14일차
모크
sgRNA – sgRNA – CR1128 GCR-067
21일차

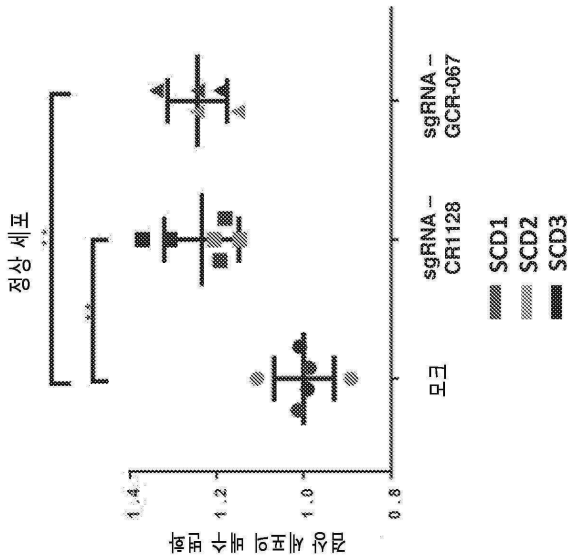
도면40


SCD1
SCD2
SCD3
MFI의 배수 변화
2.5
2.0
1.5
1.0
모크
sgRNA – sgRNA – CR1128 GCR-067
11일차
14일차
21일차

도면41a


검상 세포
**
*
경상 세포의 배수 변화
1.5
1.0
0.5
0.0
모크
sgRNA – CR1128
sgRNA – GCR-067
SCD1
SCD2
SCD3

도면41b



서 열 목 록

SEQUENCE LISTING

<110> NOVARTIS AG

INTELLIA THERAPEUTICS, INC.

<120> COMPOSITIONS AND METHODS FOR THE TREATMENT OF HEMOGLOBINOPATHIES

<130> PAT057603-WO-PCT

<140><141><150> 62/455,464

<151> 2017-02-06

<160> 332

<170> PatentIn version 3.5

<210> 1

<211> 20

<212> RNA

<213> Homo sapiens

<400> 1

aguccuggua uccucuauga 20

<210> 2

<211> 20

<212> RNA

<213> Homo sapiens

<400> 2

aauuagcagu auccucuugg 20

<210> 3

<211> 20

<212> RNA

<213> Homo sapiens

<400> 3

agaauaaauu agagaaaaac 20

<210> 4

<211> 20

<212> RNA

<213> Homo sapiens

<400> 4

aaaaauuagc aguauccucu 20

<210> 5

<211> 20

<212> RNA

<213> Homo sapiens

<400> 5

aaaauuagca guauccucuu 20

<210> 6

<211> 20

<212> RNA

<213> Homo sapiens

<400> 6

aaaaacugga augacugaau 20

<210> 7

<211> 20

<212> RNA

<213> Homo sapiens

<400> 7

cucccaucau agaggauacc 20

<210> 8

<211> 20

<212> RNA

<213> Homo sapiens

<400> 8

ggagaaggaa acuagcuaaa 20

<210> 9

<211> 20

<212> RNA

<213> Homo sapiens

<400> 9

guuuccuucu cccaucauag 20

<210> 10

<211> 20

<212> RNA

<213> Homo sapiens

<400> 10

gggagaagga aacuagcuaa 20

<210> 11

<211> 20

<212> RNA

<213> Homo sapiens

<400> 11

cacuggagcu agagacaaga 20

<210> 12

<211> 20

<212> RNA

<213> Homo sapiens

<400> 12

agagacaaga agguaaaaaa 20

<210> 13

<211> 20

<212> RNA

<213> Homo sapiens

<400> 13

aaauuagcag uauccucuug 20

<210> 14

<211> 20

<212> RNA

<213> Homo sapiens

<400> 14

guccugguau ccucuaugau 20

<210> 15

<211> 20

<212> RNA

<213> Homo sapiens

<400> 15

guauccucua ugaugggaga 20

<210> 16

<211> 20

<212> RNA

<213> Homo sapiens

<400> 16

ccuagccagc cgccggcccc 20

<210> 17

<211> 20

<212> RNA

<213> Homo sapiens

<400> 17

auuaagcagc aguauccucu 20

<210> 18

<211> 20

<212> RNA

<213> Homo sapiens

<400> 18

uauccaguga ggccaggggc 20

<210> 19

<211> 20

<212> RNA

<213> Homo sapiens

<400> 19

cauugagaua gugugggaa 20

<210> 20

<211> 20

<212> RNA

<213> Homo sapiens

<400> 20

ccagugaggc caggggccgg 20

<210> 21

<211> 20

<212> RNA

<213> Homo sapiens

<400> 21

gugggaagg ggcccccaag 20

<210> 22

<211> 20

<212> RNA

<213> Homo sapiens

<400> 22

agaauaaauu agagaaaaau 20

<210> 23

<211> 20

<212> RNA

<213> Homo sapiens

<400> 23

ccaggggccg gcggcuggcu 20

<210> 24

<211> 20

<212> RNA

<213> Homo sapiens

<400> 24

ugaggccagg ggccggcggc 20

<210> 25

<211> 20

<212> RNA

<213> Homo sapiens

<400> 25

caguuccaca cacucgcuuc 20

<210> 26

<211> 20

<212> RNA

<213> Homo sapiens

<400> 26

ccgccggccc cuggccucac 20

<210> 27

<211> 20

<212> RNA

<213> Homo sapiens

<400> 27

guuugccuug ucaaggcuau 20

<210> 28

<211> 20

<212> RNA

<213> Homo sapiens

<400> 28

ggcuagggau gaagaauaaa 20

<210> 29

<211> 20

<212> RNA

<213> Homo sapiens

<400> 29

agaaguccug guaucuucua 20

<210> 30

<211> 20

<212> RNA

<213> Homo sapiens

<400> 30

caggggccgg cggcuggcua 20

<210> 31

<211> 20

<212> RNA

<213> Homo sapiens

<400> 31

acuggauacu cuaagacuau 20

<210> 32

<211> 20

<212> RNA

<213> Homo sapiens

<400> 32

uuaagcagca guauccucuu 20

<210> 33

<211> 20

<212> RNA

<213> Homo sapiens

<400> 33

cccuggcuaa acuccaccca 20

<210> 34

<211> 20

<212> RNA

<213> Homo sapiens

<400> 34

aaaaauugga augacugaau 20

<210> 35

<211> 20

<212> RNA

<213> Homo sapiens

<400> 35

uuagaguauc cagugaggcc 20

<210> 36

<211> 20

<212> RNA

<213> Homo sapiens

<400> 36

cccaugggug gaguuuagcc 20

<210> 37

<211> 20

<212> RNA

<213> Homo sapiens

<400> 37

aggcaaggcu ggccaaccca 20

<210> 38

<211> 20

<212> RNA

<213> Homo sapiens

<400> 38

uagaguaucc agugaggcca 20

<210> 39

<211> 20

<212> RNA

<213> Homo sapiens

<400> 39

uaucugucug aaacgguccc 20

<210> 40

<211> 20

<212> RNA

<213> Homo sapiens

<400> 40

auugagauag uguggggaag 20

<210> 41

<211> 20

<212> RNA

<213> Homo sapiens

<400> 41

cuucaucccu agccagccgc 20

<210> 42

<211> 20

<212> RNA

<213> Homo sapiens

<400> 42

gcuauugguc aaggcaaggc 20

<210> 43

<211> 20

<212> RNA

<213> Homo sapiens

<400> 43

augcaaauau cugucugaaa 20

<210> 44

<211> 20

<212> RNA

<213> Homo sapiens

<400> 44

gcauugagau agugugggga 20

<210> 45

<211> 20

<212> RNA

<213> Homo sapiens

<400> 45

uggucaaguu ugccuuguca 20

<210> 46

<211> 20

<212> RNA

<213> Homo sapiens

<400> 46

gggagaagaa aacuagcuaa 20

<210> 47

<211> 20

<212> RNA

<213> Homo sapiens

<400> 47

ggcaaggcug gccaacccau 20

<210> 48

<211> 20

<212> RNA

<213> Homo sapiens

<400> 48

acggcugaca aaagaagucc 20

<210> 49

<211> 20

<212> RNA

<213> Homo sapiens

<400> 49

cgagugugug gaacugcuga 20

<210> 50

<211> 20

<212> RNA

<213> Homo sapiens

<400> 50

ccuggcuaaa cuccacccau 20

<210> 51

<211> 20

<212> RNA

<213> Homo sapiens

<400> 51

ggagaagaaa acuagcuaaa 20

<210> 52

<211> 20

<212> RNA

<213> Homo sapiens

<400> 52

cucccaccau agaagauacc 20

<210> 53

<211> 20

<212> RNA

<213> Homo sapiens

<400> 53

cuugucaagg cuauugguca 20

<210> 54

<211> 20

<212> RNA

<213> Homo sapiens

<400> 54

aguccuggua ucuucuaugg 20

<210> 55

<211> 20

<212> RNA

<213> Homo sapiens

<400> 55

auauuugcau ugagauagug 20

<210> 56

<211> 20

<212> RNA

<213> Homo sapiens

<400> 56

gcuaaacucc acccaugggu 20

<210> 57

<211> 20

<212> RNA

<213> Homo sapiens

<400> 57

acguuccaga agcgagugug 20

<210> 58

<211> 20

<212> RNA

<213> Homo sapiens

<400> 58

guccugguau cuucuauggu 20

<210> 59

<211> 20

<212> RNA

<213> Homo sapiens

<400> 59

uauuugcauu gagauagugu 20

<210> 60

<211> 20

<212> RNA

<213> Homo sapiens

<400> 60

uaagcagcag uauccucuug 20

<210> 61

<211> 20

<212> RNA

<213> Homo sapiens

<400> 61

ggaaugacug aaucggaaca 20

<210> 62

<211> 20

<212> RNA

<213> Homo sapiens

<400> 62

cuugaccaau agccuugaca 20

<210> 63

<211> 20

<212> RNA

<213> Homo sapiens

<400> 63

caaggcuauu ggucaaggca 20

<210> 64

<211> 20

<212> RNA

<213> Homo sapiens

<400> 64

aaggcuggcc aacccauggg 20

<210> 65

<211> 20

<212> RNA

<213> Homo sapiens

<400> 65

acucgcuucu ggaacgucug 20

<210> 66

<211> 20

<212> RNA

<213> Homo sapiens

<400> 66

auuugcauug agauagugug 20

<210> 67

<211> 20

<212> RNA

<213> Homo sapiens

<400> 67

acugaaucgg aacaaggcaa 20

<210> 68

<211> 20

<212> RNA

<213> Homo sapiens

<400> 68

ccaugggugg aguuuagcca 20

<210> 69

<211> 20

<212> RNA

<213> Homo sapiens

<400> 69

aagcagcagu auccucuugg 20

<210> 70

<211> 20

<212> RNA

<213> Homo sapiens

<400> 70

agaguaucca gugaggccag 20

<210> 71

<211> 20

<212> RNA

<213> Homo sapiens

<400> 71

gagugugugg aacugcugaa 20

<210> 72

<211> 20

<212> RNA

<213> Homo sapiens

<400> 72

uagucuuaga guauccagug 20

<210> 73

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 73

aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag 60

ucggugcuuu uuuu 74

<210> 74

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"
<400> 74
aguccuggua uccucuauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100
<210> 75
<211> 100
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"
<400> 75
aguccuggua uccucuauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100
<210> 76
<211> 100
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"
<400> 76
aguccuggua uccucuauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 77
<211> 42
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 77

aguccuggua uccucuauga guuuuagagc uaugcuguuu ug 42

<210> 78

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 78

aguccuggua uccucuauga guuuuagagc uaugcuguuu ug 42

<

210> 79

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 79

aaaaacugga augacugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 80

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 80

aaaaacugga augacugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 81

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 81

aaaaacugga augacugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 82

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 82

aaaaacugga augacugaau guuuuagagc uaugcuguuu ug 42

<210> 83

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 83

aaaaacugga augacugaau guuuuagagc uaugcuguuu ug 42

<210> 84

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 84

ggagaaggaa acuagcuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 85

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 85

ggagaaggaa acuagcuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 86

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 86

ggagaaggaa acuagcuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 87

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 87

ggagaaggaa acuagcuaaa guuuuagagc uaugcuguuu ug 42

<210> 88

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 88

ggagaaggaa acuagcuaaa guuuuagagc uaugcuguuu ug 42

<210> 89

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 89

guuuccuucu cccaucauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 90

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 90

guuuccuucu cccaucauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 91

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 91

guuuccuucu cccaucauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 92

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 92

guuuccuucu cccaucauag guuuuagagc uaugcuguuu ug 42

<210> 93

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 93

guuuccuucu cccaucauag guuuuagagc uaugcuguuu ug 42

<210> 94

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 94

gggagaagga aacuagcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 95

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 95

gggagaagga aacuagcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 96

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 96

gggagaagga aacuagcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 97

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 97

gggagaagga aacuagcuaa guuuuagagc uaugcuguuu ug 42

<210> 98

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 98

gggagaagga aacuagcuaa guuuuagagc uaugcuguuu ug 42

<210> 99

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 99

cacuggagcu agagacaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 100

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 100

cacuggagcu agagacaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 101

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 101

cacuggagcu agagacaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 102

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 102

cacuggagcu agagacaaga guuuuagagc uaugcuguuu ug 42

<210> 103

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 103

cacuggagcu agagacaaga guuuuagagc uaugcuguuu ug 42

<210> 104

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 104

agagacaaga agguaaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 105

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 105

agagacaaga agguaaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 106

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 106

agagacaaga agguaaaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 107

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 107

agagacaaga agguaaaaaa guuuuagagc uaugcuguuu ug 42

<210> 108

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 108

agagacaaga agguaaaaaa guuuuagagc uaugcuguuu ug 42

<210> 109

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 109

ggcuagggau gaagaauaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 110

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 110

ggcuagggau gaagaauaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 111

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 111

ggcuagggau gaagaauaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 112

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 112

ggcuagggau gaagaauaaa guuuuagagc uaugcuguuu ug 42

<210> 113

<211> 42

<212> RNA

<213> Artificial Sequence

<220

><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 113

ggcuagggau gaagaauaaa guuuuagagc uaugcuguuu ug 42

<210> 114

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 114

aaaaauugga augacugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 115

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 115

aaaaauugga augacugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 116

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 116

aaaaauugga augacugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 117

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 117

aaaaauugga augacugaau guuuuagagc uaugcuguuu ug 42

<210> 118

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 118

aaaaauugga augacugaau guuuuagagc uaugcuguuu ug 42

<210> 119

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 119

uggucaaguu ugccuuguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 120

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 120

uggucaaguu ugccuuguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 121

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 121

uggucaaguu ugccuuguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 122

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 122

uggucaaguu ugccuuguca guuuuagagc uaugcuguuu ug 42

<210> 123

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 123

uggucaaguu ugccuuguca guuuuagagc uaugcuguuu ug 42

<210> 124

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 124

gggagaagaa aacuagcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 125

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 125

gggagaagaa aacuagcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 126

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 126

gggagaagaa aacuagcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 127

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 127

gggagaagaa aacuagcuaa guuuuagagc uaugcuguuu ug 42

<210> 128

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 128

gggagaagaa aacuagcuaa guuuuagagc uaugcuguuu ug 42

<210> 129

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 129

ggcaaggcug gccaacccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 130

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 130

ggcaaggcug gccaacccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 131

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 131

ggcaaggcug gccaacccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 132

<211> 42

<

212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 132

ggcaaggcug gccaacccau guuuuagagc uaugcuguuu ug 42

<210> 133

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 133

ggcaaggcug gccaacccau guuuuagagc uaugcuguuu ug 42

<210> 134

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 134

acggcugaca aaagaagucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 135

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 135

acggcugaca aaagaagucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 136

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 136

acggcugaca aaagaagucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 137

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 137

acggcugaca aaagaagucc guuuuagagc uaugcuguuu ug 42

<210> 138

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 138

acggcugaca aaagaagucc guuuuagagc uaugcuguuu ug 42

<210> 139

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 139

ccuggcuaaa cuccacccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 140

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 140

ccuggcuaaa cuccacccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 141

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 141

ccuggcuaaa cuccacccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 142

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 142

ccuggcuaaa cuccacccau guuuuagagc uaugcuguuu ug 42

<210> 143

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 143

ccuggcuaaa cuccacccau guuuuagagc uaugcuguuu ug 42

<210> 144

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 144

ggagaagaaa acuagcuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 145

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 145

ggagaagaaa acuagcuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 146

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 146

ggagaagaaa acuagcuaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 147

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 147

ggagaagaaa acuagcuaaa guuuuagagc uaugcuguuu ug 42

<210> 148

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 148

ggagaagaaa acuagcuaaa guuuuagagc uaugcuguuu ug 42

<210> 149

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 149

cuugucaagg cuauugguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 150

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 150

cuugucaagg cuauugguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 151

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 151

cuugucaagg cuauugguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 152

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 152

cuugucaagg cuauugguca guuuuagagc uaugcuguuu ug 42

<210> 153

<211> 42

<212> RNA

<213> Artificial Sequence

<220

><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 153

cuugucaagg cuauugguca guuuuagagc uaugcuguuu ug 42

<210> 154

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 154

aguccuggua ucuucuaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 155

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 155

aguccuggua ucuucuaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 156

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 156

aguccuggua ucuucuaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 157

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 157

aguccuggua ucuucuaugg guuuuagagc uaugcuguuu ug 42

<210> 158

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 158

aguccuggua ucuucuaugg guuuuagagc uaugcuguuu ug 42

<210> 159

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 159

guccugguau cuucuauggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 160

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 160

guccugguau cuucuauggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 161

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 161

guccugguau cuucuauggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 162

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 162

guccugguau cuucuauggu guuuuagagc uaugcuguuu ug 42

<210> 163

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 163

guccugguau cuucuauggu guuuuagagc uaugcuguuu ug 42

<210> 164

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 164

cuugaccaau agccuugaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 165

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 165

cuugaccaau agccuugaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 166

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 166

cuugaccaau agccuugaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 167

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 167

cuugaccaau agccuugaca guuuuagagc uaugcuguuu ug 42

<210> 168

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 168

cuugaccaau agccuugaca guuuuagagc uaugcuguuu ug 42

<210> 169

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 169

caaggcuauu ggucaaggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 170

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 170

caaggcuauu ggucaaggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 171

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 171

caaggcuauu ggucaaggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 172

<211> 42

<

212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 172

caaggcuauu ggucaaggca guuuuagagc uaugcuguuu ug 42

<210> 173

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 173

caaggcuauu ggucaaggca guuuuagagc uaugcuguuu ug 42

<210> 174

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 174

acugaaucgg aacaaggcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 175

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 175

acugaaucgg aacaaggcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 176

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 176

acugaaucgg aacaaggcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 177

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 177

acugaaucgg aacaaggcaa guuuuagagc uaugcuguuu ug 42

<210> 178

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 178

acugaaucgg aacaaggcaa guuuuagagc uaugcuguuu ug 42

<210> 179

<400> 179

000

<210> 180

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 180

aucagaggcc aaacccuucc 20

<210> 181

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 181

aucagaggcc aaacccuucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 182

<211> 12

<212> RNA

<213> Artificial Sequence

<

220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 182

guuuuagagc ua 12

<210> 183

<211> 12

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 183

guuuaagagc ua 12

<210> 184

<211> 5

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 184

ugcug 5

<210> 185

<211> 5

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 185

uuuug 5

<210> 186

<211> 4

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 186

gaaa 4

<210> 187

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 187

uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc gagucggugc 60

<210> 188

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 188

uagcaaguuu aaauaaggcu aguccguuau caacuugaaa aaguggcacc gagucggugc 60

<210> 189

<211> 5

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 189

cagca 5

<210> 190

<211> 5000

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220><221> MISC_FEATURE

<222> (1)..(5000)

<223> /note="This sequence may encompass 50-5000 nucleotides"

<220>

<221> source

<223> /note="See specification as filed for detailed description of
substitutions and preferred embodiments"

<400> 190

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 120

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 180

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 300

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 360

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 420

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 480

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 540

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 600

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 660

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 720

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 780

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 840

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 960

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1020

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1080

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1140

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1200

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1320

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1380

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1440

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1500

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1560

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1620

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1680

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1740

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1860

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1980

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2040

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2100

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2160

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2220

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2280

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2340

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2400

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2460

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2520

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2580

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2640

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2700

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2760

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2820

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2880

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2940

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3000

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3060

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3120

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3180

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3240

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3300

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3360

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3420

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3480

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3540

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3600

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3660

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3720

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3780

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3840

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3900

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3960

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4020

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4080

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4140

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4200

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4260

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4320

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4380

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4440

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4500

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4560

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4620

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4680

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4740

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4800

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4860

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4920

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4980

aaaaaaaaaa aaaaaaaaaa 5000

<210> 191

<211> 13

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 191

uagcaaguua aaa 13

<210> 192

<211> 13

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 192

uagcaaguuu aaa 13

<210> 193

<211> 47

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 193

uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag ucggugc 47

<210> 194

<211> 51

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 194

uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag ucggugcuuu u 51

<210> 195

<211> 76

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 195

guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugc 76

<210> 196

<211> 76

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 196

guuuaagagc uagaaauagc aaguuuaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugc 76

<210> 197

<211> 86

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 197

guuuuagagc uaugcuggaa acagcauagc aaguuaaaau aaggcuaguc cguuaucaac 60

uugaaaaagu ggcaccgagu cggugc 86

<210> 198

<211> 86

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 198

guuuaagagc uaugcuggaa acagcauagc aaguuuaaau aaggcuaguc cguuaucaac 60

uugaaaaagu ggcaccgagu cggugc 86

<210> 199

<211> 17

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 199

guuuuagagc uaugcug 17

<210> 200

<211> 17

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 200

guuuaagagc uaugcug 17

<210> 201

<211> 22

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 201

guuuuagagc uaugcuguuu ug 22

<210> 202

<211> 22

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 202

guuuaagagc uaugcuguuu ug 22

<210> 203

<211> 65

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 203

cagcauagca aguuaaaaua aggcuagucc guuaucaacu ugaaaaagug gcaccgaguc 60

ggugc 65

<210> 204

<211> 65

<212> RNA

<213> Artificial Sequence

<

220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 204

cagcauagca aguuuaaaua aggcuagucc guuaucaacu ugaaaaagug gcaccgaguc 60

ggugc 65

<210> 205

<211> 1368

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 205

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

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

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

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

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

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

325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe

340 345 350

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

355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp

370 375 380

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

385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu

405 410 415

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

420 425 430

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

435 440 445

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

450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

465 470 475 480

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

485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser

500 505 510

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

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

595 600 605

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

610 615 620

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

625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr

645 650 655

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

660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe

675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe

690 695 700

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

705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

725 730 735

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

740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

755 760 765

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

770 775 780

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

785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu

805 810 815

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

820 825 830

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

835 840 845

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

850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

980 985 990

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

995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala

1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr

1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

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

1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys

1265 1270 1275

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

1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn

1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala

1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser

1325 1330 1335

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

1340 1345 1350

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

1355 1360 1365

<210> 206

<211> 7

<212> PRT

<213> Simian virus 40

<400> 206

Pro Lys Lys Lys Arg Lys Val

1 5

<210> 207

<211> 16

<212> PRT

<213> Unknown

<220

><221> source

<223> /note="Description of Unknown:

Nucleoplasmin bipartite NLS sequence"

<400> 207

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

1 5 10 15

<210> 208

<211> 9

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

C-myc NLS sequence"

<400> 208

Pro Ala Ala Lys Arg Val Lys Leu Asp

1 5

<210> 209

<211> 11

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

C-myc NLS sequence"

<400> 209

Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro

1 5 10

<210> 210

<211> 38

<212> PRT

<213> Homo sapiens

<400> 210

Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly

1 5 10 15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro

20 25 30

Arg Asn Gln Gly Gly Tyr

35

<210> 211

<211> 42

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

IBB domain from importin-alpha sequence"

<400> 211

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

1 5 10 15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys

20 25 30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val

35 40

<210> 212

<211> 8

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Myoma T protein sequence"

<400> 212

Val Ser Arg Lys Arg Pro Arg Pro

1 5

<210> 213

<211> 8

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Myoma T protein sequence"

<400> 213

Pro Pro Lys Lys Ala Arg Glu Asp

1 5

<210> 214

<211> 8

<212> PRT

<213> Homo sapiens

<400> 214

Pro Gln Pro Lys Lys Lys Pro Leu

1 5

<210> 215

<211> 12

<212> PRT

<213> Mus musculus

<400> 215

Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro

1 5 10

<210> 216

<211> 5

<212> PRT

<213> Influenza virus

<400> 216

Asp Arg Leu Arg Arg

1 5

<210> 217

<211> 7

<212> PRT

<213> Influenza virus

<400> 217

Pro Lys Gln Lys Lys Arg Lys

1 5

<210> 218

<211> 10

<212> PRT

<213> Hepatitis delta virus

<400> 218

Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu

1 5 10

<210> 219

<211> 10

<212> PRT

<213> Mus musculus

<400> 219

Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg

1 5 10

<210> 220

<211> 20

<212> PRT

<213> Homo sapiens

<400> 220

Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys

1 5 10 15

Lys Ser Lys Lys

20

<210> 221

<211> 17

<212> PRT

<213> Homo sapiens

<400> 221

Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys

1 5 10 15

Lys

<210> 222

<211> 4107

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 222

atggataaaa agtacagcat cgggctggac atcggtacaa actcagtggg gtgggccgtg 60

attacggacg agtacaaggt accctccaaa aaatttaaag tgctgggtaa cacggacaga 120

cactctataa agaaaaatct tattggagcc ttgctgttcg actcaggcga gacagccgaa 180

gccacaaggt tgaagcggac cgccaggagg cggtatacca ggagaaagaa ccgcatatgc 240

tacctgcaag aaatcttcag taacgagatg gcaaaggttg acgatagctt tttccatcgc 300

ctggaagaat cctttcttgt tgaggaagac aagaagcacg aacggcaccc catctttggc 360

aatattgtcg acgaagtggc atatcacgaa aagtacccga ctatctacca cctcaggaag 420

aagctggtgg actctaccga taaggcggac ctcagactta tttatttggc actcgcccac 480

atgattaaat ttagaggaca tttcttgatc gagggcgacc tgaacccgga caacagtgac 540

gtcgataagc tgttcatcca acttgtgcag acctacaatc aactgttcga agaaaaccct 600

ataaatgctt caggagtcga cgctaaagca atcctgtccg cgcgcctctc aaaatctaga 660

agacttgaga atctgattgc tcagttgccc ggggaaaaga aaaatggatt gtttggcaac 720

ctgatcgccc tcagtctcgg actgacccca aatttcaaaa gtaacttcga cctggccgaa 780

gacgctaagc tccagctgtc caaggacaca tacgatgacg acctcgacaa tctgctggcc 840

cagattgggg atcagtacgc cgatctcttt ttggcagcaa agaacctgtc cgacgccatc 900

ctgttgagcg atatcttgag agtgaacacc gaaattacta aagcacccct tagcgcatct 960

atgatcaagc ggtacgacga gcatcatcag gatctgaccc tgctgaaggc tcttgtgagg 1020

caacagctcc ccgaaaaata caaggaaatc ttctttgacc agagcaaaaa cggctacgct 1080

ggctatatag atggtggggc cagtcaggag gaattctata aattcatcaa gcccattctc 1140

gagaaaatgg acggcacaga ggagttgctg gtcaaactta acagggagga cctgctgcgg 1200

aagcagcgga cctttgacaa cgggtctatc ccccaccaga ttcatctggg cgaactgcac 1260

gcaatcctga ggaggcagga ggatttttat ccttttctta aagataaccg cgagaaaata 1320

gaaaagattc ttacattcag gatcccgtac tacgtgggac ctctcgcccg gggcaattca 1380

cggtttgcct ggatgacaag gaagtcagag gagactatta caccttggaa cttcgaagaa 1440

gtggtggaca agggtgcatc tgcccagtct ttcatcgagc ggatgacaaa ttttgacaag 1500

aacctcccta atgagaaggt gctgcccaaa cattctctgc tctacgagta ctttaccgtc 1560

tacaatgaac tgactaaagt caagtacgtc accgagggaa tgaggaagcc ggcattcctt 1620

agtggagaac agaagaaggc gattgtagac ctgttgttca agaccaacag gaaggtgact 1680

gtgaagcaac ttaaagaaga ctactttaag aagatcgaat gttttgacag tgtggaaatt 1740

tcaggggttg aagaccgctt caatgcgtca ttggggactt accatgatct tctcaagatc 1800

ataaaggaca aagacttcct ggacaacgaa gaaaatgagg atattctcga agacatcgtc 1860

ctcaccctga ccctgttcga agacagggaa atgatagaag agcgcttgaa aacctatgcc 1920

cacctcttcg acgataaagt tatgaagcag ctgaagcgca ggagatacac aggatgggga 1980

agattgtcaa ggaagctgat caatggaatt agggataaac agagtggcaa gaccatactg 2040

gatttcctca aatctgatgg cttcgccaat aggaacttca tgcaactgat tcacgatgac 2100

tctcttacct tcaaggagga cattcaaaag gctcaggtga gcgggcaggg agactccctt 2160

catgaacaca tcgcgaattt ggcaggttcc cccgctatta aaaagggcat ccttcaaact 2220

gtcaaggtgg tggatgaatt ggtcaaggta atgggcagac ataagccaga aaatattgtg 2280
atcgagatgg cccgcgaaaa ccagaccaca cagaagggcc agaaaaatag tagagagcgg 2340
atgaagagga tcgaggaggg catcaaagag ctgggatctc agattctcaa agaacacccc 2400
gtagaaaaca cacagctgca gaacgaaaaa ttgtacttgt actatctgca gaacggcaga 2460
gacatgtacg tcgaccaaga acttgatatt aatagactgt ccgactatga cgtagaccat 2520
atcgtgcccc agtccttcct gaaggacgac tccattgata acaaagtctt gacaagaagc 2580
gacaagaaca ggggtaaaag tgataatgtg cctagcgagg aggtggtgaa aaaaatgaag 2640

aactactggc gacagctgct taatgcaaag ctcattacac aacggaagtt cgataatctg 2700
acgaaagcag agagaggtgg cttgtctgag ttggacaagg cagggtttat taagcggcag 2760
ctggtggaaa ctaggcagat cacaaagcac gtggcgcaga ttttggacag ccggatgaac 2820
acaaaatacg acgaaaatga taaactgata cgagaggtca aagttatcac gctgaaaagc 2880
aagctggtgt ccgattttcg gaaagacttc cagttctaca aagttcgcga gattaataac 2940
taccatcatg ctcacgatgc gtacctgaac gctgttgtcg ggaccgcctt gataaagaag 3000
tacccaaagc tggaatccga gttcgtatac ggggattaca aagtgtacga tgtgaggaaa 3060

atgatagcca agtccgagca ggagattgga aaggccacag ctaagtactt cttttattct 3120
aacatcatga attttttaa gacggaaatt accctggcca acggagagat cagaaagcgg 3180
ccccttatag agacaaatgg tgaaacaggt gaaatcgtct gggataaggg cagggatttc 3240
gctactgtga ggaaggtgct gagtatgcca caggtaaata tcgtgaaaaa aaccgaagta 3300
cagaccggag gattttccaa ggaaagcatt ttgcctaaaa gaaactcaga caagctcatc 3360
gcccgcaaga aagattggga ccctaagaaa tacgggggat ttgactcacc caccgtagcc 3420
tattctgtgc tggtggtagc taaggtggaa aaaggaaagt ctaagaagct gaagtccgtg 3480

aaggaactct tgggaatcac tatcatggaa agatcatcct ttgaaaagaa ccctatcgat 3540
ttcctggagg ctaagggtta caaggaggtc aagaaagacc tcatcattaa actgccaaaa 3600
tactctctct tcgagctgga aaatggcagg aagagaatgt tggccagcgc cggagagctg 3660
caaaagggaa acgagcttgc tctgccctcc aaatatgtta attttctcta tctcgcttcc 3720
cactatgaaa agctgaaagg gtctcccgaa gataacgagc agaagcagct gttcgtcgaa 3780
cagcacaagc actatctgga tgaaataatc gaacaaataa gcgagttcag caaaagggtt 3840
atcctggcgg atgctaattt ggacaaagta ctgtctgctt ataacaagca ccgggataag 3900

cctattaggg aacaagccga gaatataatt cacctcttta cactcacgaa tctcggagcc 3960

cccgccgcct tcaaatactt tgatacgact atcgaccgga aacggtatac cagtaccaaa 4020

gaggtcctcg atgccaccct catccaccag tcaattactg gcctgtacga aacacggatc 4080

gacctctctc aactgggcgg cgactag 4107

<210> 223

<211> 4164

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 223

atggctccga agaaaaagcg taaagtggat aaaaaataca gcattggtct ggacattggc 60

acgaactcag tgggttgggc ggtcatcacg gatgaatata aggtcccgtc aaaaaagttc 120

aaagtgctgg gcaacaccga tcgccattcg attaaaaaga atctgatcgg cgcgctgctg 180

tttgatagcg gtgaaaccgc ggaagcaacg cgtctgaaac gtaccgcacg tcgccgttac 240

acgcgccgta aaaatcgtat tctgtatctg caggaaatct ttagcaacga aatggcgaaa 300

gttgatgact catttttcca ccgcctggaa gaatcgtttc tggtcgaaga agacaaaaag 360

catgaacgtc acccgatttt cggtaatatc gttgatgaag tcgcgtacca tgaaaaatat 420

ccgacgattt accatctgcg taaaaaactg gtggattcaa ccgacaaagc cgatctgcgc 480

ctgatttacc tggcactggc tcatatgatc aaatttcgtg gccacttcct gattgaaggt 540

gacctgaacc cggataactc tgacgttgat aagctgttca tccagctggt ccaaacctat 600

aatcagctgt tcgaagaaaa cccgatcaat gcaagtggcg ttgatgcgaa ggccattctg 660

tccgctcgcc tgagtaaatc ccgccgtctg gaaaacctga ttgcacaact gccgggcgaa 720

aagaaaaacg gcctgtttgg taatctgatc gctctgtcac tgggtctgac gccgaacttt 780

aaatcgaatt tcgacctggc agaagatgct aagctgcagc tgagcaaaga tacctacgat 840

gacgatctgg acaacctgct ggcgcaaatt ggtgaccagt atgccgacct gtttctggcg 900

gccaaaaatc tgtcagatgc cattctgctg tcggacatcc tgcgcgtgaa caccgaaatc 960

acgaaagcgc cgctgtcagc ctcgatgatt aaacgctacg atgaacatca ccaggacctg 1020

accctgctga aagcactggt tcgtcagcaa ctgccggaaa agtacaagga aattttcttt 1080

gaccaatcta agaacggcta tgcaggttac atcgatggcg gtgctagtca ggaagaattc 1140

tacaagttca tcaagccgat cctggaaaaa atggatggca cggaagaact gctggtgaaa 1200

ctgaatcgtg aagatctgct gcgtaaacaa cgcacctttg acaacggcag cattccgcat 1260

cagatccacc tgggtgaact gcatgcgatt ctgcgccgtc aggaagattt ttatccgttc 1320

ctgaaagaca accgtgaaaa aattgaaaag atcctgacgt ttcgcatccc gtattacgtt 1380

ggcccgctgg cgcgtggtaa tagccgcttc gcctggatga cccgcaaatc tgaagaaacc 1440

attacgccgt ggaactttga agaagtggtt gataaaggtg caagcgctca gtcttttatc 1500

gaacgtatga ccaatttcga taaaaacctg ccgaatgaaa aggtcctgcc gaaacatagc 1560

ctgctgtatg aatactttac cgtgtacaac gaactgacga aagtgaagta tgttaccgaa 1620

ggcatgcgca aaccggcgtt tctgtctggt gaacagaaaa aagccattgt ggatctgctg 1680

ttcaagacca atcgtaaagt tacggtcaaa cagctgaagg aagattactt caaaaagatc 1740

gaagaattcg acagcgtgga aatttctggc gttgaagatc gtttcaacgc cagtctgggt 1800

acctatcatg acctgctgaa gatcatcaag gacaaggatt ttctggataa cgaagaaaat 1860

gaagacattc tggaagatat cgtgctgacc ctgacgctgt tcgaagatcg tgaaatgatt 1920

gaagaacgcc tgaaaacgta cgcacacctg tttgacgata aagttatgaa gcagctgaaa 1980

cgccgtcgct ataccggctg gggtcgtctg tctcgcaaac tgattaatgg catccgcgat 2040

aagcaaagtg gtaaaacgat tctggatttc ctgaaatccg acggctttgc caaccgtaat 2100

ttcatgcagc tgatccatga cgatagtctg acctttaagg aagacattca gaaagcacaa 2160

gtgtcaggcc agggtgattc gctgcatgaa cacattgcga acctggccgg ctccccggct 2220

attaaaaagg gtatcctgca gaccgtcaaa gtcgtggatg aactggtgaa ggttatgggc 2280

cgtcacaaac cggaaaacat tgtgatcgaa atggcgcgcg aaaatcagac cacgcaaaag 2340

ggtcagaaaa actcacgtga acgcatgaag cgcattgaag aaggcatcaa agaactgggt 2400

tcgcagattc tgaaagaaca tccggttgaa aacacccagc tgcaaaatga aaaactgtac 2460

ctgtattacc tgcaaaatgg ccgtgacatg tatgtcgatc aggaactgga catcaaccgc 2520

ctgagcgact atgatgtcga ccacattgtg ccgcagagct ttctgaagga cgattctatc 2580

gataataaag tgctgacccg ttctgataag aaccgcggta aaagcgacaa tgttccgtct 2640

gaagaagttg tcaaaaagat gaagaactac tggcgtcaac tgctgaatgc gaagctgatt 2700

acgcagcgta aattcgataa cctgaccaag gcggaacgcg gcggtctgag tgaactggat 2760

aaggccggct ttatcaaacg tcaactggtg gaaacccgcc agattacgaa acatgttgcc 2820

cagatcctgg attcccgcat gaacacgaaa tatgacgaaa atgataagct gattcgtgaa 2880

gtcaaagtga tcaccctgaa gagtaagctg gtgtccgatt tccgtaagga ctttcagttc 2940

tacaaagttc gcgaaattaa caattaccat cacgcacacg atgcttatct gaatgcagtg 3000

gttggcaccg ctctgatcaa aaagtatccg aaactggaaa gcgaatttgt gtatggtgat 3060
tacaaagtct atgacgtgcg caagatgatt gcgaaaagtg aacaggaaat cggcaaggcg 3120
accgccaagt actttttcta ttccaacatc atgaactttt tcaagaccga aatcacgctg 3180
gcaaatggcg aaattcgtaa acgcccgctg atcgaaacca acggcgaaac gggtgaaatt 3240
gtgtgggata aaggtcgtga cttcgcgacc gttcgcaaag tcctgtcaat gccgcaagtg 3300
aatatcgtta aaaagaccga agttcagacg ggcggtttta gtaaagaatc catcctgccg 3360
aagcgtaact cggataaact gattgcgcgc aaaaaggatt gggacccgaa aaagtacggc 3420

ggttttgata gtccgaccgt tgcatattcc gtcctggtcg tggctaaagt tgaaaaaggc 3480
aagagtaaaa agctgaagtc cgtcaaagaa ctgctgggta ttaccatcat ggaacgtagc 3540
tcttttgaaa agaacccgat tgacttcctg gaagccaagg gctacaaaga agtgaaaaag 3600
gatctgatta tcaagctgcc gaaatattcg ctgttcgaac tggaaaacgg tcgtaaacgc 3660
atgctggcaa gcgctggcga actgcagaag ggtaatgaac tggcactgcc gtctaaatat 3720
gtgaactttc tgtacctggc tagccattat gaaaaactga agggttctcc ggaagataac 3780
gaacagaagc aactgttcgt tgaacaacat aaacactacc tggatgaaat catcgaacag 3840

atctcagaat tctcgaaacg cgtcattctg gcggatgcca atctggacaa agtgctgagc 3900
gcgtataaca agcatcgtga taaaccgatt cgcgaacagg ccgaaaatat tatccacctg 3960
tttaccctga cgaacctggg cgcaccggca gcttttaaat acttcgatac cacgatcgac 4020
cgtaagcgct ataccagcac gaaagaagtt ctggatgcta ccctgattca tcagtcaatc 4080
accggtctgt atgaaacgcg tattgacctg agccaactgg gcggtgatag ccgtgccgac 4140
catcaccatc accatcacta atag 4164

<210> 224
<211> 74
<212> RNA
<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 224
aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag 60
ucggugcuuu uuuu 74

<210> 225

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 225

aacagcauag caaguuuaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag 60

ucggugcuuu uuuu 74

<210> 226

<211> 16

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 226

guuuuagagc uaugcu 16

<210> 227

<211> 90

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 227

guuuaagagc uaugcuggaa acagcauagc aaguuuaaau aaggcuaguc cguuaucaac 60

uugaaaaagu ggcaccgagu cggugcuuuu 90

<210> 228

<211> 67

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 228

agcauagcaa guuaaaauaa ggcuaguccg uuaucaacuu gaaaaagugg caccgagucg 60

gugcuuu 67

<210> 229

<211> 85

<

212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 229

guuggaacca uucaaaacag cauagcaagu uaaaauaagg cuaguccguu aucaacuuga 60

aaaaguggca ccgagucggu gcuuu 85

<210> 230

<211> 7

<212> PRT

<213> Tobacco etch virus

<400> 230

Glu Asn Leu Tyr Phe Gln Gly

1 5

<210> 231

<211> 80

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 231

guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugcuuuu 80

<210> 232

<211> 67

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 232

aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag 60

ucggugc 67

<210> 233

<211> 1393

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 233

Met Ala Pro Lys Lys Lys Arg Lys Val Asp Lys Lys Tyr Ser Ile Gly
1 5 10 15

Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu
20 25 30

Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg
35 40 45

His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly
50 55 60

Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr
65 70 75 80

Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn
85 90 95

Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser
100 105 110

Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly
115 120 125

Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr
130 135 140

His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg
145 150 155 160

Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe
165 170 175

Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu
180 185 190

Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro
195 200 205

Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu
210 215 220

Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu
225 230 235 240

Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu
245 250 255

Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu
260 265 270

Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala
275 280 285

Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu
290 295 300

Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile
305 310 315 320

Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His
325 330 335

His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro
340 345 350

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

355 360 365

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

370 375 380

Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys

385 390 395 400

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

405 410 415

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

420 425 430

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

435 440 445

Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala

450 455 460

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr

465 470 475 480

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

485 490 495

Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn

500 505 510

Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val

515 520 525

Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys

530 535 540

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

545 550 555 560

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

565 570 575

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu

580 585 590

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

595 600 605

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

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

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

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

850 855 860

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

Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp
930 935 940
Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu
945 950 955 960
Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys
965 970 975
Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala
980 985 990

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

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

Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe
1100 1105 1110

Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
1115 1120 1125
Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp
1130 1135 1140
Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu
1145 1150 1155
Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly
1160 1165 1170

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

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

Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys
1295 1300 1305
Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu
1310 1315 1320
Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr

1325 1330 1335

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

1340 1345 1350

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

1355 1360 1365

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

1370 1375 1380

Lys Arg Lys Val His His His His His His

1385 1390

<210> 234

<211> 1393

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 234

Met Ala Pro Lys Lys Lys Arg Lys Val Asp Lys Lys Tyr Ser Ile Gly

1 5 10 15

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

20 25 30

Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg

35 40 45

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

50 55 60

Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr

65 70 75 80

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

85 90 95

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

100 105 110

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

115 120 125

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

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

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

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

Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile
370 375 380

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

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

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

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

610 615 620

Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile
625 630 635 640

Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met
645 650 655

Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg
660 665 670

Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu
675 680 685

Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu
690 695 700

Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln
705 710 715 720

Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala
725 730 735

Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val
740 745 750

Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val
755 760 765

Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn
770 775 780

Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly
785 790 795 800

Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn
805 810 815

Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val
820 825 830

Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His
835 840 845

Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val
850 855 860

Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser
865 870 875 880
Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn
885 890 895

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

Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys
965 970 975
Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala
980 985 990
His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys
995 1000 1005
Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val
1010 1015 1020

Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly
1025 1030 1035
Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe
1040 1045 1050
Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg
1055 1060 1065
Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp
1070 1075 1080

Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro
1085 1090 1095
Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe

1100 1105 1110

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

1115 1120 1125

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp

1130 1135 1140

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

1145 1150 1155

Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly

1160 1165 1170

Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

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

1310 1315 1320

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

1325 1330 1335

Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala
1340 1345 1350
Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1355 1360 1365
Asp Leu Ser Gln Leu Gly Gly Asp Ser Arg Ala Asp Pro Lys Lys
1370 1375 1380

Lys Arg Lys Val His His His His His His His
1385 1390

<210> 235
<211> 1399
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 235
Met Gly Ser Ser His His His His His His His His Glu Asn Leu Tyr
1 5 10 15
Phe Gln Gly Ser Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly
20 25 30

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

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys
100 105 110
Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu
115 120 125

Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp
130 135 140
Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys
145 150 155 160

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

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

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

Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys
355 360 365
Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp

370 375 380

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

385 390 395 400

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

405 410 415

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His

420 425 430

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

435 440 445

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

450 455 460

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

465 470 475 480

Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp

485 490 495

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

500 505 510

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

515 520 525

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

530 535 540

Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu

545 550 555 560

Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn

565 570 575

Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile

580 585 590

Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn

595 600 605

Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys

610 615 620

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

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

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

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

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

865 870 875 880

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

885 890 895

Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile

900 905 910

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

915 920 925

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

930 935 940

Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn

945 950 955 960

Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile

965 970 975

Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe

980 985 990

Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr

995 1000 1005

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

1010 1015 1020

Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg

1085 1090 1095

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

1100 1105 1110

Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu
1115 1120 1125

Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys
1130 1135 1140

Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr
1145 1150 1155

Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys
1160 1165 1170

Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile
1175 1180 1185

Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu
1190 1195 1200

Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu
1205 1210 1215

Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met
1220 1225 1230

Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu
1235 1240 1245

Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu
1250 1255 1260

Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe
1265 1270 1275

Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile
1280 1285 1290

Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp
1295 1300 1305

Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg
1310 1315 1320

Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu
1325 1330 1335

Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg

1340 1345 1350

Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile
1355 1360 1365
His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser
1370 1375 1380
Gln Leu Gly Gly Asp Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys
1385 1390 1395
Val

<210> 236
<211> 1396
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"
<400> 236
Met Ala His His His His His His Gly Gly Ser Pro Lys Lys Lys Arg
1 5 10 15
Lys Val Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser
20 25 30
Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys
35 40 45
Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu
50 55 60

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

115 120 125

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

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

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

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

Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala
370 375 380

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

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

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

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

610 615 620

Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu
625 630 635 640

Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr
645 650 655

Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg
660 665 670

Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg
675 680 685

Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly
690 695 700

Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr
705 710 715 720

Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser
725 730 735

Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys
740 745 750

Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met
755 760 765

Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn
770 775 780

Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg
785 790 795 800

Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His
805 810 815

Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr
820 825 830

Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn
835 840 845

Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu
850 855 860

Lys Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn
865 870 875 880
Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met
885 890 895

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

Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys
965 970 975
Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val
980 985 990
Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala
995 1000 1005
Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser
1010 1015 1020

Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met
1025 1030 1035
Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr
1040 1045 1050
Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr
1055 1060 1065
Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn
1070 1075 1080

Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala
1085 1090 1095
Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys

1100 1105 1110

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

1115 1120 1125

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

1130 1135 1140

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

1145 1150 1155

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

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

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

1235 1240 1245

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

1250 1255 1260

Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln

1265 1270 1275

His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe

1280 1285 1290

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

1295 1300 1305

Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala

1310 1315 1320

Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro

1325 1330 1335

Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr
1340 1345 1350
Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser
1355 1360 1365
Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly
1370 1375 1380

Gly Asp Ser Arg Ala Asp Pro Lys Lys Lys Arg Lys Val
1385 1390 1395

<210> 237
<211> 1386
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 237

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

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

Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser
100 105 110
Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly
115 120 125

Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr
130 135 140
His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg
145 150 155 160

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

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

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

Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala
355 360 365
Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile

370 375 380

Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys
385 390 395 400

Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
405 410 415

Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg
420 425 430

Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile
435 440 445

Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala
450 455 460

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr
465 470 475 480

Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala
485 490 495

Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn
500 505 510

Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val
515 520 525

Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys
530 535 540

Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu
545 550 555 560

Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr
565 570 575

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu
580 585 590

Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile
595 600 605

Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu
610 615 620

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

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

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

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

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

865 870 875 880

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys

965 970 975

Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala

980 985 990

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

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
1115 1120 1125
Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp
1130 1135 1140
Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu
1145 1150 1155
Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly
1160 1165 1170

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

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

Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys
1295 1300 1305
Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu
1310 1315 1320
Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr
1325 1330 1335
Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala

1340 1345 1350

Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1355 1360 1365
Asp Leu Ser Gln Leu Gly Gly Asp Ser Arg Ala Asp His His His
1370 1375 1380
His His His
1385

<210> 238
<211> 1389
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 238

Met Ala His His His His His His Gly Gly Ser Asp Lys Lys Tyr Ser
1 5 10 15
Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr
20 25 30
Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr
35 40 45
Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp
50 55 60
Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg
65 70 75 80
Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe
85 90 95
Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu
100 105 110
Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile
115 120 125

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

130 135 140

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

195 200 205

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

260 265 270

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

325 330 335

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

340 345 350

Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly
355 360 365
Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys

370 375 380

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

385 390 395 400

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

405 410 415

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

420 425 430

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

435 440 445

Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro

450 455 460

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

465 470 475 480

Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala

485 490 495

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

500 505 510

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

515 520 525

Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met

530 535 540

Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp

545 550 555 560

Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu

565 570 575

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

580 585 590

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

595 600 605

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

610 615 620

Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu
625 630 635 640
Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys

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

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

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

835 840 845
Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn
850 855 860
Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val

865 870 875 880

Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

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

980 985 990

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

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly

1100 1105 1110

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

1115 1120 1125

Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly

1130 1135 1140

Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys

1145 1150 1155

Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu

1160 1165 1170

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

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

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

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

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

1310 1315 1320

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

1325 1330 1335

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

1340 1345 1350

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

1355 1360 1365

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

1370 1375 1380

Lys Lys Lys Arg Lys Val

1385

<210> 239

<211> 1399

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 239

Met Gly Ser Ser His His His His His His His His Glu Asn Leu Tyr

1 5 10 15

Phe Gln Gly Ser Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly

20 25 30

Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro

35 40 45

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

50 55 60

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

65 70 75 80

Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys

85 90 95

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys

100 105 110

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

115 120 125

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

130 135 140

Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys

145 150 155 160

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

165 170 175

Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly

180 185 190

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

195 200 205

Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser

210 215 220

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

225 230 235 240

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

245 250 255

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

260 265 270

Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

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

340 345 350

Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys

355 360 365

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

370 375 380

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

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

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

Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn
565 570 575
Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile
580 585 590
Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn
595 600 605
Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys
610 615 620

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

625 630 635 640

Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu

645 650 655

Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys

660 665 670

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

675 680 685

Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys

690 695 700

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

705 710 715 720

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

725 730 735

Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala

740 745 750

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

755 760 765

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

770 775 780

Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg

785 790 795 800

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

805 810 815

Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr

820 825 830

Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu

835 840 845

Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln

850 855 860

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

865 870 875 880

Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val
885 890 895

Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile
900 905 910

Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu
915 920 925

Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr
930 935 940

Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn
945 950 955 960

Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile
965 970 975

Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe
980 985 990

Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr
995 1000 1005

Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys
1010 1015 1020

Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val
1025 1030 1035

Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr
1040 1045 1050

Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr
1055 1060 1065

Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile
1070 1075 1080

Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg
1085 1090 1095

Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn
1100 1105 1110

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

1115 1120 1125

Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys
1130 1135 1140
Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr
1145 1150 1155
Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys
1160 1165 1170
Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile
1175 1180 1185

Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu
1190 1195 1200
Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu
1205 1210 1215
Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met
1220 1225 1230
Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu
1235 1240 1245

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

Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg
1310 1315 1320
Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu
1325 1330 1335
Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg
1340 1345 1350

Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile
1355 1360 1365

His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser
1370 1375 1380
Gln Leu Gly Gly Asp Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys
1385 1390 1395
Val

<210> 240
<211> 1393
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 240
Met Ala Pro Lys Lys Lys Arg Lys Val Asp Lys Lys Tyr Ser Ile Gly

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

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

Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr

130 135 140

His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg

145 150 155 160

Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe

165 170 175

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

180 185 190

Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro

195 200 205

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

210 215 220

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

225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His

325 330 335

His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro

340 345 350

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

355 360 365

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

370 375 380

Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys

385 390 395 400

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

405 410 415

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

420 425 430

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

435 440 445

Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala

450 455 460

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr

465 470 475 480

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

485 490 495

Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn

500 505 510

Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val

515 520 525

Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys

530 535 540

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

545 550 555 560

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

565 570 575

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu

580 585 590

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

595 600 605

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

610 615 620

Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile
625 630 635 640
Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met

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

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

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

835 840 845
Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Ala Val
850 855 860
Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser

865 870 875 880

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys

965 970 975

Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala

980 985 990

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

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp

1130 1135 1140

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

1145 1150 1155

Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly

1160 1165 1170

Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

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

1310 1315 1320

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

1325 1330 1335

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

1340 1345 1350

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

1355 1360 1365

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

1370 1375 1380

Lys Arg Lys Val His His His His His His

1385 1390

<210> 241

<211> 1393

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 241

Met Ala Pro Lys Lys Lys Arg Lys Val Asp Lys Lys Tyr Ser Ile Gly

1 5 10 15

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

20 25 30

Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg

35 40 45

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

50 55 60

Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr

65 70 75 80

Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn

85 90 95

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

100 105 110

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

115 120 125

Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr

130 135 140

His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg
145 150 155 160
Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe
165 170 175

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

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

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

Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile
370 375 380

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

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

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

Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr
565 570 575
Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu
580 585 590
Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile
595 600 605
Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu
610 615 620

Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile

625 630 635 640

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

645 650 655

Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg

660 665 670

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

675 680 685

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

690 695 700

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

705 710 715 720

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

725 730 735

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

740 745 750

Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val

755 760 765

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

770 775 780

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

785 790 795 800

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

805 810 815

Glu Ala Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val

820 825 830

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

835 840 845

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

850 855 860

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

865 870 875 880

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

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

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

Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp
1070 1075 1080
Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro
1085 1090 1095
Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe
1100 1105 1110
Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile

1115 1120 1125

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp
1130 1135 1140
Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu
1145 1150 1155
Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly
1160 1165 1170
Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp
1175 1180 1185

Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile
1190 1195 1200
Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg
1205 1210 1215
Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu
1220 1225 1230
Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser
1235 1240 1245

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

Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu
1310 1315 1320
Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr
1325 1330 1335
Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala
1340 1345 1350

Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1355 1360 1365

Asp Leu Ser Gln Leu Gly Gly Asp Ser Arg Ala Asp Pro Lys Lys
1370 1375 1380
Lys Arg Lys Val His His His His His His
1385 1390

<210> 242
<211> 1393
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 242
Met Ala Pro Lys Lys Lys Arg Lys Val Asp Lys Lys Tyr Ser Ile Gly
1 5 10 15

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

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

130 135 140

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

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

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

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

Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys
385 390 395 400

Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
405 410 415

Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg
420 425 430

Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile
435 440 445

Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala
450 455 460

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr
465 470 475 480

Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala
485 490 495

Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn
500 505 510

Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val
515 520 525

Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys
530 535 540

Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu
545 550 555 560

Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr
565 570 575

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu
580 585 590

Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile
595 600 605

Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu
610 615 620

Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile

625 630 635 640

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

645 650 655

Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg

660 665 670

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

675 680 685

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

690 695 700

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

705 710 715 720

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

725 730 735

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

740 745 750

Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val

755 760 765

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

770 775 780

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

785 790 795 800

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

805 810 815

Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val

820 825 830

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

835 840 845

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

850 855 860

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

865 870 875 880

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn
885 890 895
Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu
900 905 910

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

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

Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe
1040 1045 1050
Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Ala
1055 1060 1065
Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp
1070 1075 1080
Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro
1085 1090 1095

Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe
1100 1105 1110
Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile

1115 1120 1125

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp

1130 1135 1140

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

1145 1150 1155

Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly

1160 1165 1170

Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp

1175 1180 1185

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

1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

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

1310 1315 1320

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

1325 1330 1335

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

1340 1345 1350

Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1355 1360 1365
Asp Leu Ser Gln Leu Gly Gly Asp Ser Arg Ala Asp Pro Lys Lys
1370 1375 1380
Lys Arg Lys Val His His His His His His
1385 1390

<210> 243

<211> 1393

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 243

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

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

Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly
115 120 125
Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr
130 135 140

```
His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg
145                 150                 155                 160
Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe
                165                 170                 175

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

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

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

Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile
    370                 375                 380
Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys
```

385 390 395 400

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

405 410 415

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

420 425 430

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

435 440 445

Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala

450 455 460

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr

465 470 475 480

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

485 490 495

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

500 505 510

Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val

515 520 525

Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys

530 535 540

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

545 550 555 560

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

565 570 575

Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu

580 585 590

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

595 600 605

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

610 615 620

Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile

625 630 635 640

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

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

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

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

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys

965 970 975

Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala

980 985 990

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

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp
1130 1135 1140
Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu
1145 1150 1155
Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly
1160 1165 1170
Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp
1175 1180 1185

Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile
1190 1195 1200
Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg
1205 1210 1215
Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu
1220 1225 1230
Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser
1235 1240 1245

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

Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu
1310 1315 1320
Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr
1325 1330 1335
Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala
1340 1345 1350
Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile

1355 1360 1365

Asp Leu Ser Gln Leu Gly Gly Asp Ser Arg Ala Asp Pro Lys Lys
1370 1375 1380
Lys Arg Lys Val His His His His His His
1385 1390

<210> 244
<211> 1386
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 244

Met Ala Pro Lys Lys Lys Arg Lys Val Asp Lys Lys Tyr Ser Ile Gly
1 5 10 15

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

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

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

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

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

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

385 390 395 400

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

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

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

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

Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met
645 650 655

Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg
660 665 670

Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu
675 680 685

Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu
690 695 700

Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln
705 710 715 720

Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala
725 730 735

Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val
740 745 750

Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val
755 760 765

Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn
770 775 780

Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly
785 790 795 800

Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn
805 810 815

Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val
820 825 830

Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His
835 840 845

Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val
850 855 860

Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser
865 870 875 880

Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys

965 970 975

Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala

980 985 990

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

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

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

1115 1120 1125

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp
1130 1135 1140
Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu
1145 1150 1155

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

Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu
1220 1225 1230
Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser
1235 1240 1245
His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys
1250 1255 1260
Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile
1265 1270 1275

Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala
1280 1285 1290
Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys
1295 1300 1305
Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu
1310 1315 1320
Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr
1325 1330 1335

Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala
1340 1345 1350
Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile

1355 1360 1365

Asp Leu Ser Gln Leu Gly Gly Asp Ser Arg Ala Asp His His His

1370 1375 1380

His His His

1385

<210> 245

<211> 835

<212> PRT

<213> Homo sapiens

<400> 245

Met Ser Arg Arg Lys Gln Gly Lys Pro Gln His Leu Ser Lys Arg Glu

1 5 10 15

Phe Ser Pro Glu Pro Leu Glu Ala Ile Leu Thr Asp Asp Glu Pro Asp

20 25 30

His Gly Pro Leu Gly Ala Pro Glu Gly Asp His Asp Leu Leu Thr Cys

35 40 45

Gly Gln Cys Gln Met Asn Phe Pro Leu Gly Asp Ile Leu Ile Phe Ile

50 55 60

Glu His Lys Arg Lys Gln Cys Asn Gly Ser Leu Cys Leu Glu Lys Ala

65 70 75 80

Val Asp Lys Pro Pro Ser Pro Ser Pro Ile Glu Met Lys Lys Ala Ser

85 90 95

Asn Pro Val Glu Val Gly Ile Gln Val Thr Pro Glu Asp Asp Asp Cys

100 105 110

Leu Ser Thr Ser Ser Arg Gly Ile Cys Pro Lys Gln Glu His Ile Ala

115 120 125

Asp Lys Leu Leu His Trp Arg Gly Leu Ser Ser Pro Arg Ser Ala His

130 135 140

Gly Ala Leu Ile Pro Thr Pro Gly Met Ser Ala Glu Tyr Ala Pro Gln

145 150 155 160

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

165 170 175

Gln Pro Phe Thr Ser Ala Trp Phe Leu Leu Gln His Ala Gln Asn Thr
180 185 190
His Gly Leu Arg Ile Tyr Leu Glu Ser Glu His Gly Ser Pro Leu Thr

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

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

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

385 390 395 400
Gly Glu Lys Pro Tyr Lys Cys Asn Leu Cys Asp His Ala Cys Thr Gln
405 410 415
Ala Ser Lys Leu Lys Arg His Met Lys Thr His Met His Lys Ser Ser

420 425 430

Pro Met Thr Val Lys Ser Asp Asp Gly Leu Ser Thr Ala Ser Ser Pro

435 440 445

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

450 455 460

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

465 470 475 480

Glu Asn Gly Asp Glu Glu Glu Glu Glu Asp Asp Glu Glu Glu Glu Glu

485 490 495

Glu Glu Glu Glu Glu Glu Glu Glu Leu Thr Glu Ser Glu Arg Val Asp

500 505 510

Tyr Gly Phe Gly Leu Ser Leu Glu Ala Ala Arg His His Glu Asn Ser

515 520 525

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

530 535 540

Asp Val Met Gln Gly Met Val Leu Ser Ser Met Gln His Phe Ser Glu

545 550 555 560

Ala Phe His Gln Val Leu Gly Glu Lys His Lys Arg Gly His Leu Ala

565 570 575

Glu Ala Glu Gly His Arg Asp Thr Cys Asp Glu Asp Ser Val Ala Gly

580 585 590

Glu Ser Asp Arg Ile Asp Asp Gly Thr Val Asn Gly Arg Gly Cys Ser

595 600 605

Pro Gly Glu Ser Ala Ser Gly Gly Leu Ser Lys Lys Leu Leu Leu Gly

610 615 620

Ser Pro Ser Ser Leu Ser Pro Phe Ser Lys Arg Ile Lys Leu Glu Lys

625 630 635 640

Glu Phe Asp Leu Pro Pro Ala Ala Met Pro Asn Thr Glu Asn Val Tyr

645 650 655

Ser Gln Trp Leu Ala Gly Tyr Ala Ala Ser Arg Gln Leu Lys Asp Pro

660 665 670

Phe Leu Ser Phe Gly Asp Ser Arg Gln Ser Pro Phe Ala Ser Ser Ser
675 680 685

Glu His Ser Ser Glu Asn Gly Ser Leu Arg Phe Ser Thr Pro Pro Gly
690 695 700

Glu Leu Asp Gly Gly Ile Ser Gly Arg Ser Gly Thr Gly Ser Gly Gly

705 710 715 720

Ser Thr Pro His Ile Ser Gly Pro Gly Pro Gly Arg Pro Ser Ser Lys
725 730 735

Glu Gly Arg Arg Ser Asp Thr Cys Glu Tyr Cys Gly Lys Val Phe Lys
740 745 750

Asn Cys Ser Asn Leu Thr Val His Arg Arg Ser His Thr Gly Glu Arg
755 760 765

Pro Tyr Lys Cys Glu Leu Cys Asn Tyr Ala Cys Ala Gln Ser Ser Lys

770 775 780

Leu Thr Arg His Met Lys Thr His Gly Gln Val Gly Lys Asp Val Tyr
785 790 795 800

Lys Cys Glu Ile Cys Lys Met Pro Phe Ser Val Tyr Ser Thr Leu Glu
805 810 815

Lys His Met Lys Lys Trp His Ser Asp Arg Val Leu Asn Asn Asp Ile
820 825 830

Lys Thr Glu
835

<210> 246

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 246

aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag 60

ucggugcuuu uuuu 74

<210> 247

<211> 6

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
6xHis tag"

<400> 247

His His His His His His

1 5

<210> 248

<211> 8

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
8xHis tag"

<400> 248

His His His His His His His His

1 5

<210> 249

<211> 10

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> MISC_FEATURE

<222> (1)..(10)

<223> /note="This sequence may encompass 1-10 nucleotides"

<400> 249

uuuuuuuuuu 10

<210> 250

<211> 10

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> MISC_FEATURE

<222> (1)..(10)

<223> /note="This sequence may encompass 1-10 nucleotides"

<400> 250

aaaaaaaaaa 10

<210> 251

<211> 36

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 251

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcu 36

<210> 252

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 252

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu ug 42

<210> 253

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 253

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu ug 42

<210> 254

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 254

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 255

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 255

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 256

<211> 100

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 256

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 257

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 257

tgctgagatg aaacaggcgt 20

<210> 258

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 258

ttaggcatcc acaagggctg 20

<210> 259

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 259

gctctacaaa tggaacccaa cc 22

<210> 260

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 260

ctgctctgat ctctaacacc tca 23

<210> 261

<211

> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 261

gaagatacag cttgcctccg a 21

<210> 262

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 262

ttgctgagat gaaacaggcg t 21

<210> 263

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 263

acggataagt agatattgag gtaagc 26

<210> 264

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 264

gtctctttca gttagcagtg g 21

<210> 265

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic probe"

<400> 265

actgcgctga aactgtggct ttatag 26

<210> 266

<211> 20

<212> DNA

<213> Homo sapiens

<400> 266

ccctttagct agtttttcttc 20

<210> 267

<211> 16

<212> DNA

<213> Homo sapiens

<400> 267

ggacttcttt tgtcag 16

<210> 268

<211> 15

<212> DNA

<213

> Homo sapiens

<400> 268

ttagctagtt tcctt 15

<210> 269

<211> 20

<212> DNA

<213> Homo sapiens

<400> 269

ccctttagct agtttttcttc 20

<210> 270

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 270

acggataagt agatattgag gtaagc 26

<210> 271

<

211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 271

gtctctttca gttagcagtg g 21

<210> 272

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

probe"

<400> 272

actgcgctga aactgtggct ttatag 26

<210> 273

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 273

acggataagt agatattgag gtaagc 26

<210> 274

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 274

gtctctttca gttagcagtg g 21

<210> 275

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic probe"

<400> 275

actgcgctga aactgtggct ttatag 26

<210> 276

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 276

ggccacggag cgagacaucu 20

<210> 277

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 277

cagacagcaa actcacccag t 21

<210> 278

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 278

ctgacgctta tcgacgccct 20

<210> 279

<211> 33

<212> DNA

<213> Artificial Sequence

<220><221> source

<223

> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 279

tcgtcggcag cgtcagatgt gtataagaga cag 33

<210> 280

<211> 34

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 280

gtctcgtggg ctcggagatg tgtataagag acag 34

<210> 281

<211> 51

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<220><221> modified_base

<222> (30)..(37)

<223> a, c, t, g, unknown or other

<400> 281

aatgatacgg cgaccaccga gatctacacn nnnnnnntcg tcggcagcgt c 51

<210> 282

<211> 47

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<220><221> modified_base

<222> (25)..(32)

<223> a, c, t, g, unknown or other

<400> 282

caagcagaag acggcatacg agatnnnnnn nngtctcgtg ggctcgg 47

<210> 283

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 283

caagcagaag acggcatacg a 21

<210> 284

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 284

aatgatacgg cgaccaccga ga 22

<210> 285

<211

> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 285

acggataagt agatattgag gtaagc 26

<210> 286

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 286

gtctctttca gttagcagtg g 21

<210> 287

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

probe"

<400> 287

actgcgctga aactgtggct ttatag 26

<210> 288

<211> 23

<212> DNA

<213> Homo sapiens

<400> 288

agtcctggta tcttctatgg tgg 23

<210> 289

<211> 23

<212> DNA

<213> Homo sapiens

<400> 289

attcccagta tcctctatga tgg 23

<210> 290

<211> 23

<212> DNA

<213> Homo sapiens

<400> 290

ggagaagaaa actagctaaa ggg 23

<210> 291

<211> 23

<212> DNA

<213> Homo sapiens

<400> 291

gggagaagaa aactagctaa agg 23

<210> 292

<211> 23

<212> DNA

<213> Homo sapiens

<400> 292

ggagaaggaa actagctaaa ggg 23

<210> 293

<211> 20

<212> DNA

<213> Homo sapiens

<400> 293

ggagaaggaa actagctaaa 20

<210> 294

<211> 20

<212> DNA

<213> Homo sapiens

<400> 294

actgaatcgg aacaaggcaa 20

<210> 295

<211> 20

<212> DNA

<213> Homo sapiens

<400> 295

gggagaagga aactagctaa 20

<210> 296

<211> 20

<212> DNA

<213> Homo sapiens

<400> 296

aaaaactgga atgactgaat 20

<210> 297

<211> 20

<212> DNA

<213> Homo sapiens

<400> 297

agagacaaga aggtaaaaaa 20

<210> 298

<211> 20

<212> DNA

<213> Homo sapiens

<400> 298

agtcctggta tcctctatga 20

<210> 299

<211> 20

<212> DNA

<213> Homo sapiens

<400> 299

cactggagct agagacaaga 20

<210> 300

<211> 20

<212> DNA

<213> Homo sapiens

<400> 300

acggctgaca aaagaagtcc 20

<210> 301

<211> 20

<212> DNA

<213> Homo sapiens

<400> 301

gggagaagaa aactagctaa 20

<210> 302

<211> 20

<212> DNA

<213> Homo sapiens

<400> 302

aaaaattgga atgactgaat 20

<210> 303

<211> 492

<212> DNA

<213> Homo sapiens

<400> 303

tggtacatgc tttagcttta aactacaggc ctcactggag ctagagacaa gaaggtaaaa 60

aacggctgac aaaagaagtc ctggtatcct ctatgatggg agaaggaaac tagctaaagg 120

gaagaataaa ttagagaaaa actggaatga ctgaatcgga acaaggcaaa ggctataaaa 180

aaaattagca gtatcctctt gggggcccct tccccacact atctcaatgc aaatatctgt 240

ctgaaacggt ccctggctaa actccaccca tgggttggcc agccttgcct tgaccaatag 300

ccttgacaag gcaaacttga ccaatagtct tagagtatcc agtgaggcca ggggccggcg 360

gctggctagg gatgaagaat aaaaggaagc acccttcagc agttccacac actcgcttct 420

ggaacgtctg aggttatcaa taagctccta gtccagacgc catgggtcat ttcacagagg 480

aggacaaggc ta 492

<210> 304

<211> 20

<212> DNA

<213> Homo sapiens

<400> 304

gtttccttct cccatcatag 20

<210> 305

<211> 20

<212> DNA

<213> Homo sapiens

<400> 305

cctggctaaa ctccacccat 20

<210> 306

<211> 20

<212> DNA

<213> Homo sapiens

<400> 306

cttgaccaat agccttgaca 20

<210> 307

<211> 20

<212> DNA

<213> Homo sapiens

<400> 307

ggctagggat gaagaataaa 20

<210> 308

<211> 10

<212> PRT

<213> Homo sapiens

<400> 308

Met Gly His Phe Thr Glu Glu Asp Lys Ala

1 5 10

<210> 309

<211> 20

<212> DNA

<213> Homo sapiens

<400> 309

ggcaaggctg gccaacccat 20

<210> 310

<211> 20

<212> DNA

<213> Homo sapiens

<400> 310

cttgtcaagg ctattggtca 20

<210> 311

<211> 20

<212> DNA

<213> Homo sapiens

<400> 311

caaggctatt ggtcaaggca 20

<210> 312

<211> 20

<212> DNA

<213> Homo sapiens

<400> 312

tggtcaagtt tgccttgtca 20

<210> 313

<211> 20

<212> DNA

<213> Homo sapiens

<400> 313

ggagaaggaa actagctaaa 20

<210> 314

<211> 20

<212> DNA

<213> Homo sapiens

<400> 314

gggagaagga aactagctaa 20

<210> 315

<211> 20

<212> DNA

<213> Homo sapiens

<400> 315

actgaatcgg aacaaggcaa 20

<210> 316

<211> 20

<212> DNA

<213> Homo sapiens

<400> 316

gtcctggtat cttctatggt 20

<210> 317

<211> 20

<212> DNA

<213> Homo sapiens

<400> 317

agtcctggta tcttctatgg 20

<210> 318

<211> 20

<212> DNA

<213> Homo sapiens

<400> 318

ggagaagaaa actagctaaa 20

<210> 319

<211> 20

<212> DNA

<213> Homo sapiens

<400> 319

aaaaactgga atgactgaat 20

<210> 320

<211> 20

<212> DNA

<213> Homo sapiens

<400> 320

cactggagct agagacaaga 20

<210> 321

<211> 20

<212> DNA

<213> Homo sapiens

<400> 321

acggctgaca aaagaagtcc 20

<210> 322

<211> 20

<212> DNA

<213> Homo sapiens

<400> 322

gggagaagaa aactagctaa 20

<210> 323

<211> 20

<212> DNA

<213> Homo sapiens

<400> 323

aaaaattgga atgactgaat 20

<210> 324

<211> 500

<212> DNA

<213> Homo sapiens

<400> 324

acaatccctg aacttttcaa aaattggtac atgctttaac tttaaactac aggcctcact 60

ggagctacag acaagaaggt gaaaaacggc tgacaaaaga agtcctggta tcttctatgg 120

tgggagaaga aaactagcta aagggaagaa taaattagag aaaaattgga atgactgaat 180

cggaacaagg caaaggctat aaaaaaaatt aagcagcagt atcctcttgg gggccccttc 240

cccacactat ctcaatgcaa atatctgtct gaaacggtcc ctggctaaac tccacccatg 300

ggttggccag ccttgccttg accaatagcc ttgacaaggc aaacttgacc aatagtctta 360

gagtatccag tgaggccagg ggccggcggc tggctaggga tgaagaataa aaggaagcac 420

ccttcagcag ttccacacac tcgcttctgg aacgtctgag gttatcaata agctcctagt 480

ccagacgcca tgggtcattt 500

<210> 325

<211> 20

<212> DNA

<213> Homo sapiens

<400> 325

cctggctaaa ctccacccat 20

<210> 326

<211> 20

<212> DNA

<213> Homo sapiens

<400> 326

cttgaccaat agccttgaca 20

<210> 327

<211> 20

<212> DNA

<213> Homo sapiens

<400> 327

ggctagggat gaagaataaa 20

<210> 328

<211> 4

<212> PRT

<213> Homo sapiens

<400> 328

Met Gly His Phe

1

<210> 329

<211> 20

<212> DNA

<213> Homo sapiens

<400> 329

ggcaaggctg gccaacccat 20

<210> 330

<211> 20

<212> DNA

<213> Homo sapiens

<400> 330

cttgtcaagg ctattggtca 20

<210> 331

<211> 20

<212> DNA

<213> Homo sapiens

<400> 331

caaggctatt ggtcaaggca 20

<210> 332

<211> 20

<212> DNA

<213> Homo sapiens

<400> 332

tggtcaagtt tgccttgtca 20