



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0139911
(43) 공개일자 2022년10월17일

- (51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01)
- (52) CPC특허분류
C12N 15/86 (2013.01)
C12N 2740/15043 (2013.01)
- (21) 출원번호 10-2022-7029922
- (22) 출원일자(국제) 2021년02월04일
심사청구일자 없음
- (85) 번역문제출일자 2022년08월29일
- (86) 국제출원번호 PCT/GB2021/050247
- (87) 국제공개번호 WO 2021/160993
국제공개일자 2021년08월19일
- (30) 우선권주장
2001996.4 2020년02월13일 영국(GB)
2017820.8 2020년11월11일 영국(GB)
- (71) 출원인
옥스포드 바이오메디카(유케이) 리미티드
영국, 오엑스4 6엘티 옥스포드셔 옥스포드, 트랜스포트 웨이, 윈드러쉬 코트
- (72) 발명자
팔리, 다니엘
영국, 오엑스4 6엘티 옥스포드셔 옥스포드, 트랜스포트 웨이, 윈드러쉬 코트, 옥스포드 바이오메디카(유케이) 리미티드 내
- 라이트, 조던
영국, 오엑스4 6엘티 옥스포드셔 옥스포드, 트랜스포트 웨이, 윈드러쉬 코트, 옥스포드 바이오메디카(유케이) 리미티드 내
- (74) 대리인
안소영

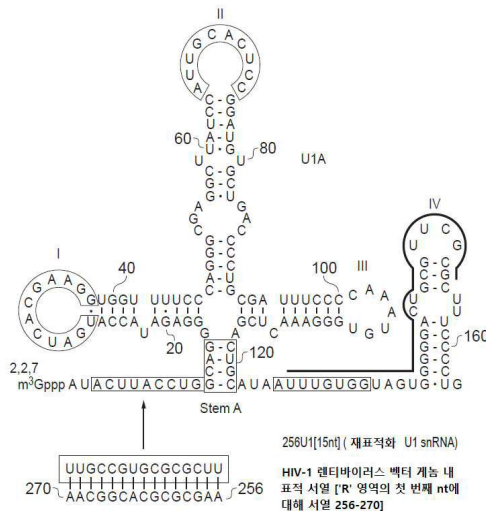
전체 청구항 수 : 총 54 항

(54) 발명의 명칭 렌티바이러스 벡터의 생산

(57) 요약

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화된, 뉴클레오타이드 서열.

대표도 - 도1



(52) CPC특허분류

C12N 2740/15052 (2013.01)

C12N 2740/16043 (2013.01)

C12N 2740/16052 (2013.01)

C12N 2840/44 (2013.01)

명세서

청구범위

청구항 1

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위(major splice donor site)가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위(cryptic splice donor site)가 불활성화된, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 2

제1항에 있어서, 상기 렌티바이러스 벡터가 3세대 렌티바이러스 벡터인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 3

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 tat-독립적 렌티바이러스 벡터에서 사용하기 위한 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 4

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 tat 부재하에서 생산되는 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 5

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 tat와는 독립적으로 전사된 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 6

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 U3-독립적 렌티바이러스 벡터에서 사용하기 위한 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 7

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 U3 프로모터와는 독립적으로 전사된 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 8

렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열로서, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위가 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위가 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 이종(heterologous) 프로모터에 의해 전사된

것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 9

선행하는 청구항들 중 어느 한 항에 있어서, 잠복 스플라이스 공여자 부위가 주요 스플라이스 공여자 부위의 3'의 첫 번째 잠복 스플라이스 공여자 부위인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 10

선행하는 청구항들 중 어느 한 항에 있어서, 상기 잠복 스플라이스 공여자 부위가 주요 스플라이스 공여자 부위의 6개 뉴클레오타이드 내에 있는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 11

선행하는 청구항들 중 어느 한 항에 있어서, 주요 스플라이스 공여자 부위 및 잠복 스플라이스 공여자 부위가 돌연변이되거나 결실된, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 12

선행하는 청구항들 중 어느 한 항에 있어서, 스플라이스 부위의 불활성화 전의 뉴클레오타이드 서열이 서열번호 1, 3, 4, 9, 10 및/또는 13 중 어느 하나에 기재된 서열을 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 13

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오타이드 서열이 서열번호 1, 3, 4, 9, 10 및/또는 13 중 어느 하나에 기재된 서열에 비해 돌연변이 또는 결실이 있는 서열을 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 14

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오타이드 서열이 불활성화되지 않았다면 서열번호 1의 뉴클레오타이드 13 및 14에 상응하는 뉴클레오타이드 사이에 절단 부위를 가지는 불활성화된 주요 스플라이스 공여자 부위를 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 15

선행하는 청구항들 중 어느 한 항에 있어서, 불활성화 전의 주요 스플라이스 공여자 부위의 뉴클레오타이드 서열이 서열번호 4에 기재된 서열을 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 16

선행하는 청구항들 중 어느 한 항에 있어서, 불활성화 전의 잠복 스플라이스 공여자 부위의 뉴클레오타이드 서열이 서열번호 10에 기재된 서열을 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 17

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오타이드 서열이 불활성화되지 않았다면 서열번호 1의 뉴클레오타이드 17 및 18에 상응하는 뉴클레오타이드 사이에 절단 부위를 가지는 불활성화된 잠복 스플라이스 공여자 부위를 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 18

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오타이드 서열이 서열번호 2, 5, 6, 7, 8, 11, 12 및/또는 14 중 어느 하나에 기재된 서열을 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 19

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오타이드 서열이 서열번호 9에 기재된 서열을 포함하지 않는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열.

청구항 20

선행하는 청구항들 중 어느 한 항에 있어서, 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및 잠복 스플라이스 공여자 부위로부터의 스플라이싱 활성이 억제되거나 제거된, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 21

선행하는 청구항들 중 어느 한 항에 있어서, 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및 잠복 스플라이스 공여자 부위로부터의 스플라이싱 활성이 형질감염된 세포 또는 형질도입된 세포에서 억제되거나 제거된, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 22

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오티드 서열이 패키징(packaging) 서열의 SL4 루프 내의 잠복 스플라이스 공여자 부위의 돌연변이를 추가로 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 23

제22항에 있어서, 패키징 서열의 SL4 루프 내의 상기 잠복 스플라이스 공여자 부위의 GT 디뉴클레오티드가 GC로 돌연변이되는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 24

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오티드 서열이 관심 뉴클레오티드를 추가로 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 25

제24항에 있어서, 관심 뉴클레오티드가 치료 효과를 일으키는 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 26

선행하는 청구항들 중 어느 한 항에 있어서, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열이 벡터 전이유전자(transgene) 발현 카세트인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 27

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오티드 서열이 변형된 U1 snRNA를 암호화하는 뉴클레오티드 서열을 추가로 포함하고, 여기서 상기 변형된 U1 snRNA는 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 결합하도록 변형된 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 28

제27항에 있어서, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열이 변형된 U1 snRNA를 암호화하는 뉴클레오티드 서열에 작동가능하게 연결된 것인, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 29

선행하는 청구항들 중 어느 한 항에 있어서, 뉴클레오티드 서열이 트립토판 RNA-결합 감소 단백질(tryptophan RNA-binding attenuation protein, TRAP) 결합 부위를 추가로 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 30

제27항에 있어서, 상기 뉴클레오티드 서열이 관심 뉴클레오티드를 포함하고, 여기서 (i) 상기 TRAP 결합 부위는

상기 관심 뉴클레오티드의 시작 코돈 ATG와 중첩되고; 및/또는 상기 핵산 서열은 또한 Kozak 서열을 포함하며, 여기서 상기 TRAP 결합 부위는 Kozak 서열과 중첩되는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 31

제29항에 있어서, 상기 뉴클레오티드 서열이 관심 뉴클레오티드를 포함하고, 여기서 (i) 상기 TRAP 결합 부위는 상기 관심 뉴클레오티드의 시작 코돈 ATG의 일부를 포함하거나 또는 여기서 ATG 시작 코돈은 TRAP 결합 부위의 일부를 포함하고; 및/또는 뉴클레오티드 서열은 Kozak 서열을 추가로 포함하며, 여기서 상기 Kozak 서열은 TRAP 결합 부위의 일부를 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 32

제29항에 있어서, 뉴클레오티드 서열이 다중 클로닝 부위(multiple cloning site) 및 Kozak 서열을 추가로 포함하고, 여기서 상기 다중 클로닝 부위는 TRAP 결합 부위의 3' KAGN2-3 반복부(repeat)의 하류 및 Kozak 서열의 상류와 중첩되거나 위치하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 33

제29항 내지 제32항 중 어느 한 항에 있어서, TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부가 시작 코돈 ATG의 적어도 첫 번째 뉴클레오티드와 중첩되는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 34

제29항 내지 제33항 중 어느 한 항에 있어서, TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부가 시작 코돈 ATG의 처음 두 개의 뉴클레오티드와 중첩되는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 35

제29항 내지 제34항 중 어느 한 항에 있어서, TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부가 코어 Kozak 서열 내의 시작 코돈 ATG의 첫 번째 뉴클레오티드와 중첩되는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 핵산 서열.

청구항 36

제29항 내지 제35항 중 어느 한 항에 있어서, 핵산 서열이 서열번호 114 또는 서열번호 116에 정의된 서열을 포함하는, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 핵산 서열.

청구항 37

제29항 내지 제36항 중 어느 한 항에 있어서, 관심 뉴클레오티드가 TRAP 결합 부위 또는 그의 일부에 작동가능하게 연결된, 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열.

청구항 38

선행하는 청구항들 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 포함하는 발현 카세트.

청구항 39

뉴클레오티드 서열 세트를 포함하는 바이러스 벡터 생산 시스템으로서, 여기서 뉴클레오티드 서열은 gag-pol, env, 선택적으로 rev 및 선행하는 청구항들 중 어느 한 항에 정의된 렌티바이러스 벡터의 RNA 게놈을 포함하는 벡터 구성요소를 암호화하는 것인, 바이러스 벡터 생산 시스템.

청구항 40

제1항 내지 제37항 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열, 제36

항에 따른 발현 카세트 또는 제39항에 따른 바이러스 벡터 생산 시스템을 포함하는 세포.

청구항 41

하기를 포함하는 렌티바이러스 벡터를 생산하기 위한 세포:

(iv) a) gag-pol 및 env, 및 선택적으로 rev를 포함하는 벡터 구성요소를 암호화하는 뉴클레오타이드 서열, 및 제 1항 내지 제37항 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열 또는 제 38항에 따른 발현 카세트; 또는

b) 제39항에 따른 바이러스 벡터 생산 시스템; 및

(v) 선택적으로 변형된 U1 snRNA를 암호화하는 뉴클레오타이드 서열 및/또는 선택적으로 TRAP를 암호화하는 뉴클레오타이드 서열.

청구항 42

제40항 또는 제41항에 있어서, 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및/또는 스플라이스 공여자 영역으로부터의 스플라이싱 활성이 억제되거나 제거된, 세포.

청구항 43

제42항에 있어서, 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및/또는 스플라이스 공여자 영역으로부터의 스플라이싱 활성이 렌티바이러스 벡터 생산 동안 억제되거나 제거된, 세포.

청구항 44

제40항 내지 제43항 중 어느 한 항에 있어서, 관심 뉴클레오타이드의 번역이 렌티바이러스 벡터 생산 동안 억제되는, 세포.

청구항 45

하기 단계를 포함하는 렌티바이러스 벡터의 생산 방법:

(iv) a) gag-pol 및 env, 및 선택적으로 rev를 포함하는 벡터 구성요소를 암호화하는 뉴클레오타이드 서열, 및 제 1항 내지 제37항 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열 또는 제 38항에 따른 발현 카세트; 또는

b) 제39항에 따른 바이러스 벡터 생산 시스템을

세포 내로 도입하는 단계;

(v) 선택적으로, 벡터 구성요소 및 렌티바이러스 벡터의 RNA 게놈을 암호화하는 상기 뉴클레오타이드 서열을 포함하는 세포를 선택하는 단계;

(vi) 렌티바이러스 벡터의 생산에 적합한 조건 하에 세포를 배양하는 단계.

청구항 46

제45항에 있어서, 단계 (i)이 TRAP를 암호화하는 뉴클레오타이드 서열을 세포 내로 도입하는 것을 추가적으로 포함하는, 방법.

청구항 47

제45항 또는 제46항에 있어서, 단계 (i)이 변형된 U1 snRNA를 암호화하는 뉴클레오타이드 서열을 도입하는 것을 추가적으로 포함하는, 방법.

청구항 48

제45항 내지 제47항 중 어느 한 항에 따른 방법에 의해 생산되는 렌티바이러스 벡터.

청구항 49

제48항에 있어서, 렌티바이러스 벡터가 제1항 내지 제37항 중 어느 한 항에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 포함하는 것인, 렌티바이러스 벡터.

청구항 50

렌티바이러스 벡터를 생산하기 위한, 제1항 내지 제37항 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열, 제38항에 따른 발현 카세트, 제39항에 따른 바이러스 벡터 생산 시스템, 또는 제40항 내지 제44항 중 어느 한 항에 따른 세포의 용도.

청구항 51

렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및/또는 스플라이스 공여자 영역으로부터의 스플라이싱 활성을 억제하거나 제거하기 위한, 제1항 내지 제37항 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열, 제38항에 따른 발현 카세트, 제39항에 따른 바이러스 벡터 생산 시스템, 또는 제40항 내지 제44항 중 어느 한 항에 따른 세포의 용도.

청구항 52

제48항에 있어서, 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및/또는 스플라이스 공여자 영역으로부터의 스플라이싱 활성이 형질감염된 세포 또는 형질도입된 세포에서 억제되거나 제거되는 것인, 용도.

청구항 53

관심 뉴클레오티드의 번역을 억제하기 위한, 제1항 내지 제37항 중 어느 한 항에 따른 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열, 제38항에 따른 발현 카세트, 제39항에 따른 바이러스 벡터 생산 시스템, 또는 제40항 내지 제44항 중 어느 한 항에 따른 세포의 용도.

청구항 54

제1항 내지 제37항 중 어느 한 항에 정의된 바와 같은 RNA 게놈을 포함하는 렌티바이러스 벡터.

발명의 설명

기술 분야

[0001]

본 발명은 진핵 세포에서 렌티바이러스 벡터의 생산에 관한 것이다. 보다 구체적으로, 본 발명은 벡터 게놈에서 주요 스플라이스 공여자 부위 및 인접한 잠복 스플라이스 공여자 부위의 불활성화에 관한 것이다. 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 제공하며, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화된다. 이러한 뉴클레오티드 서열을 포함하는 방법 및 용도도 본 발명에 포함된다.

배경 기술

[0002]

지난 수십 년 동안 백신 및 인간 유전자 치료를 위한 바이러스 벡터의 개발 및 제조는 과학 저널과 특허에 잘 기록되어 있다. 치료 효과를 위한 전이유전자(transgene)를 전달하기 위한 조작된 바이러스의 사용은 광범위하다. γ -레트로바이러스 및 렌티바이러스와 같은 RNA 바이러스(Muhlebach, M.D. *et al.*, 2010, *Retroviruses: Molecular Biology, Genomics and Pathogenesis*, 13:347-370; Antoniou, M.N., Skipper, K.A. & Anakok, O., 2013, *Hum. Gene Ther.*, 24:363-374), 및 아데노바이러스(Capasso, C. *et al.*, 2014, *Viruses*, 6:832-855) 및 아데노-관련 바이러스(AAV)(Kotterman, M.A. & Schaffer, D.V., 2014, *Nat. Rev. Genet.*, 15:445-451)와 같은 DNA 바이러스를 기반으로 하는 현대 유전자 치료 벡터는 점점 더 많은 인간 질병 적응증(indication)에서 가능성을 보여 왔다. 여기에는 혈액학적 병태에 대한 환자 세포의 생체외(*ex vivo*) 변형(Morgan, R.A. & Kakarla, S., 2014, *Cancer J.*, 20:145-150; Touzot, F. *et al.*, 2014, *Expert Opin. Biol. Ther.*, 14:789-798), 및 안과 질환(Balaggan, K.S. & Ali, R.R., 2012, *Gene Ther.*, 19:145-153), 심혈관 질환(Katz, M.G. *et al.*, 2013, *Hum. Gene Ther.*, 24:914-927), 신경퇴행성 질환(Coune, P.G., Schneider, B.L. & Aebischer, P., 2012, *Cold Spring Harb. Perspect. Med.*, 4:a009431) 및 종양 요법(Pazarentzos, E. & Mazarakis, N.D., 2014, *Adv. Exp. Med Biol.*, 818:255-280)의 생체내(*in vivo*) 치료가 포함된다.

[0003] 임상 시험에서 이러한 접근법의 성공이 품목 허가 및 상업화를 향해 구축되기 시작함에 따라, 예를 들어 백신 접종 및 유전자 요법의 맥락에서 환자에게 바이러스 벡터를 투여하는 것과 관련된 안전성 측면을 고려하는 것이 중요하다.

[0004] 개선된 안전성 프로파일을 갖는 바이러스 벡터에 대한 당업계의 지속적인 요구가 존재한다.

발명의 내용

[0005] 발명의 요약

[0006] 본 발명은 렌티바이러스 벡터 게놈의 패키징 영역에서 주요 스플라이스 공여자 부위(major splice donor site) 및 인접한 잠복 스플라이스 공여자 부위(cryptic splice donor site) 둘 모두의 불활성화에 기초한다. 렌티바이러스 벡터 게놈 내에 존재하는 주요 스플라이스 공여자 부위(MSD)는 전형적으로 RNA의 5' 영역을 향하여 고도로 구조화된 RNA를 포함하는 패키징 영역 내에 포함된다.

[0007] 레트로바이러스 게놈 내의 스플라이스 공여자 부위는 생산 세포 내에서 바이러스 RNA(vRNA) 안정성에 중요한 것으로 나타났다. 그러나, 본 발명자들은 렌티바이러스 벡터 게놈 발현 카세트 내의 MSD의 활성이 매우 난잡(promiscuous)할 수 있고, 전형적으로 >1350 bp 하류에 위치하는 내부 발현 카세트 내의 강력하거나 심지어 약한 잠복 스플라이스 수용체(acceptor) 부위에 매우 효율적으로 스플라이싱할 수 있음을 보여준다. 놀랍게도, vRNA 생산을 구동하는 외부 프로모터로부터 유래된 검출가능한 세포질 mRNA의 최대 95%가 내부 서열에 따라 스플라이싱된다.

[0008] 효율적인 벡터 생산을 위해, 가장 바람직한 생성물은 스플라이싱되지 않은(unsplliced) 패키징가능한 vRNA이며, 이 벡터 구성요소가 전형적으로 일시적인 형질감염 벡터 생산 설정 및 안정적인 형질감염 벡터 생산 설정 모두에서 제한 요소이다. 더욱이, 이 비정상적으로 스플라이싱된 mRNA가 관심 전이유전자를 암호화하는 경우(예를 들어 내부 프로모터-utr 서열로 스플라이싱), 이 mRNA는 내보내어지고(exported) 벡터 생산 동안 효율적으로 번역될 수 있을 것이다; 이것은 내부 프로모터가 약한/침묵(조직 특이적) 프로모터인지 여부와 독립적으로 발생한다.

[0009] 또한, 형질도입된 (환자) 세포에 전달된 벡터 백본에 MSD의 존재는 스플라이싱 기구에 의해 활용되는 것을 다른 사람들이 밝혔으며, 상류의 세포성 프로모터로부터의 리드-스루(read-through) 전사가 발생할 때(렌티바이러스 벡터는 활성 전사 부위를 표적으로 함), 이는 세포성 엑손이 있는 잠재적인 비정상적 스플라이스-생성물로 이어진다. 따라서, 렌티바이러스 벡터 게놈에서 MSD 부위를 기능적으로 돌연변이시키는 것이 왜 바람직한 것인지 몇 가지 이유가 있다.

[0010] 다른 사람들은 MSD 부위 내에 돌연변이를 보유하는 초기 세대 렌티바이러스 벡터를 생성하였지만 이러한 벡터는 vRNA의 전사를 구동하는 내재적인 U3 프로모터를 포함하였고, 따라서 트랜스로(*in trans*) 공급되는 tat에 의한 트랜스활성화(transactivation)에 의존하였다. 3세대 렌티바이러스 벡터는 U3 프로모터를 이종(heterologous) 프로모터 요소로 대체하였으며, 전사를 위해 tat를 필요로 하지 않는다. tat는 세포성 유전자의 트랜스활성화제이고 종양 발생에서 역할을 할 수 있기 때문에, 3세대 벡터의 U3/tat-독립성은 조절자에 의한 안전성에 있어서 중요한 발전으로 간주된다.

[0011] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열을 제공하고, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화된다.

[0012] 본 실시예에서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위와 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위 양자 모두의 불활성화는 둘 중 하나만 돌연변이시키는 것보다 개선을 제공한다는 것이 입증되어 있다. 또한 본 실시예에서 입증된 바와 같이, 본 발명은 표적 세포에서 통합된 렌티바이러스 벡터로의 감소된 전사 리드-스루(transcriptional read-through)를 예를 들어 적어도 2배 촉진한다.

[0013] 일 측면에서 본 발명에 따른 뉴클레오타이드 서열은 U3 또는 tat-독립적 렌티바이러스 벡터 시스템에서 사용하기 위한 것이다. 일 측면에서 렌티바이러스 벡터 시스템은 본원에 기재된 바와 같은 3세대 렌티바이러스 벡터 시스템일 수 있다.

[0014] 잠복 스플라이스 공여자 부위는 주요 스플라이스 공여자 부위의 3'의 첫 번째 잠복 스플라이스 공여자 부위 또는 서열이다. 일 측면에서 잠복 스플라이스 공여자 부위 또는 서열은 주요 스플라이스 공여자 부위의 6개 뉴클

레오티드 내에 있다. 주요 스플라이스 공여자 부위와 잠복 스플라이스 공여자 부위는 돌연변이되거나 결실될 수 있다.

- [0015] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 제공하며, 여기서 스플라이스 부위의 불활성화 전의 뉴클레오티드 서열은 서열번호 1, 3, 4, 9, 10 및/또는 13 중 어느 하나에 기재된 서열을 포함한다. 뉴클레오티드 서열은 서열번호 1, 3, 4, 9, 10 및/또는 13 중 어느 하나에 기재된 서열에 비해 돌연변이 또는 결실이 있는 서열을 포함할 수 있다. 일 측면에서 서열은 서열번호 13을 포함한다.
- [0016] 일 측면에서 뉴클레오티드 서열은 불활성화되지 않았다면 주요 스플라이스 공여자 영역(서열번호 13)의 뉴클레오티드 1의 바로 상류에 절단 부위를 가지는 불활성화된 주요 스플라이스 공여자 부위를 포함한다.
- [0017] 일 측면에서 뉴클레오티드 서열은 불활성화되지 않았다면 주요 스플라이스 공여자 영역(서열번호 13)의 뉴클레오티드에 상응하는 뉴클레오티드 1의 바로 상류 뿐만 아니라 뉴클레오티드 4와 5 사이에 절단 부위를 가지는 불활성화된 주요 스플라이스 공여자 부위 및 불활성화된 잠복 스플라이스 공여자 부위를 포함한다.
- [0018] 일 측면에서 뉴클레오티드 서열은 불활성화되지 않았다면 서열번호 1의 뉴클레오티드 13 및 14에 상응하는 뉴클레오티드 사이에 절단 부위를 가지는 불활성화된 주요 스플라이스 공여자 부위를 포함한다.
- [0019] 또 다른 측면에서 불활성화 전의 주요 스플라이스 공여자 부위의 뉴클레오티드 서열은 서열번호 4에 기재된 서열을 포함한다. 일 측면에서 불활성화 전의 잠복 스플라이스 공여자 부위는 서열번호 10에 기재된 서열을 포함한다.
- [0020] 뉴클레오티드 서열은 불활성화되지 않았다면 서열번호 1의 뉴클레오티드 17 및 18에 상응하는 뉴클레오티드 사이에 절단 부위를 가지는 불활성화된 잠복 스플라이스 공여자 부위를 포함할 수 있다.
- [0021] 일 측면에서 본 발명에 따른 뉴클레오티드 서열은 서열번호 2, 5, 6, 7, 8, 11, 12 및/또는 14 중 어느 하나에 기재된 서열을 포함한다.
- [0022] 바람직한 측면에서 뉴클레오티드 서열은 서열번호 14에 기재된 서열을 포함한다.
- [0023] 추가의 측면에서 뉴클레오티드 서열은 서열번호 9에 기재된 서열을 포함하지 않는다.
- [0024] 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및 잠복 스플라이스 공여자 부위로부터의 스플라이싱 활성은 예를 들어 형질감염된 세포 또는 형질도입된 세포에서 억제되거나 제거될 수 있다.
- [0025] 일 측면에서 뉴클레오티드 서열은 벡터 생산을 위해 U3 또는 tat-독립적 시스템에서 렌티바이러스 벡터에 사용하기에 적합할 수 있다. 본원에 기재된 바와 같이, 3세대 렌티바이러스 벡터는 U3/tat-독립적이며, 본 발명에 따른 뉴클레오티드 서열은 3세대 렌티바이러스 벡터의 맥락에서 사용될 수 있다. 본 발명의 일 측면에서 tat는 렌티바이러스 벡터 생산 시스템에 제공되지 않으며, 예를 들어 tat는 트랜스로 제공되지 않는다. 일 측면에서 본원에 기재된 바와 같은 세포 또는 벡터 또는 벡터 생산 시스템은 tat 단백질을 포함하지 않는다. 본 발명의 일 측면에서 HIV-1 U3은 렌티바이러스 벡터 생산 시스템에 존재하지 않으며, 예를 들어 HIV-1 U3은 벡터 게놈 발현 카세트의 전사를 구동하기 위해 시스로(in cis) 제공되지 않는다.
- [0026] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 제공하며, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화되고, 여기서 상기 뉴클레오티드 서열은 tat-독립적 렌티바이러스 벡터에서 사용하기 위한 것이다.
- [0027] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 제공하며, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화되고, 여기서 상기 뉴클레오티드 서열은 tat의 부재하에서 생산된 것이다.
- [0028] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 제공하며, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화되고, 여기서 상기 뉴클레오티드 서열은 tat와는 독립적으로 전사된 것이다.
- [0029] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열을 제공하며, 여기서 렌

티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 U3-독립적 렌티바이러스 벡터에서 사용하기 위한 것이다.

- [0030] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열을 제공하며, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 U3 프로모터와는 독립적으로 전사된 것이다.
- [0031] 일 측면에서 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열을 제공하며, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화되고, 여기서 상기 뉴클레오타이드 서열은 이중 프로모터에 의해 전사된 것이다.
- [0032] 일 측면에서, 본원에 기재된 뉴클레오타이드 서열의 전사는 U3의 존재에 의존하지 않는다. 뉴클레오타이드 서열은 U3-독립적인 전사 이벤트로부터 유래될 수 있다. 뉴클레오타이드 서열은 이중 프로모터로부터 유래될 수 있다. 본원에 기재된 뉴클레오타이드 서열은 천연 U3 프로모터를 포함하지 않을 수 있다.
- [0033] 일 측면에서, 본원에 기재된 뉴클레오타이드 서열은 패키징 서열의 SL4 루프 내의 잠복 스플라이스 공여자 부위의 돌연변이를 선택적으로 추가로 포함할 수 있다. 일 측면에서 패키징 서열의 SL4 루프 내의 상기 잠복 스플라이스 공여자 부위의 GT 디뉴클레오타이드는 GC로 돌연변이된다.
- [0034] 일 측면에서 뉴클레오타이드 서열은 치료 효과를 일으킬 수 있는 관심 뉴클레오타이드를 추가로 포함한다.
- [0035] 추가의 측면에서 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열은 벡터 전이유전자 발현 카세트이다.
- [0036] 본 발명의 또 다른 측면에서 뉴클레오타이드 서열은 변형된 U1 snRNA를 암호화하는 뉴클레오타이드 서열을 추가로 포함할 수 있고, 여기서 상기 변형된 U1 snRNA는 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열에 결합하도록 변형된 것이다. 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열은 변형된 U1 snRNA를 암호화하는 뉴클레오타이드 서열에 작동가능하게 연결될 수 있다. 일 측면에서 변형된 U1 snRNA를 암호화하는 뉴클레오타이드 서열은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열과는 상이한 뉴클레오타이드 서열 상에, 예를 들어 상이한 플라스미드 상에 있다.
- [0037] 또 다른 측면에서 뉴클레오타이드 서열은 트립토판 RNA-결합 감쇠 단백질(tryptophan RNA-binding attenuation protein, TRAP) 결합 부위를 추가로 포함하고, 또한 Kozak 서열을 포함할 수 있으며, 여기서 상기 TRAP 결합 부위는 Kozak 서열과 중첩되거나, 또는 여기서 상기 Kozak 서열은 TRAP 결합 부위의 일부를 포함한다. 뉴클레오타이드 서열은 또한 다중 클로닝 부위(multiple cloning site) 및 Kozak 서열을 추가로 포함할 수 있고, 여기서 상기 다중 클로닝 부위는 TRAP 결합 부위의 3' KAGN2-3 반복부(repeat)의 하류 및 Kozak 서열의 상류와 중첩되거나 위치한다. 관심 뉴클레오타이드는 TRAP 결합 부위 또는 그의 일부에 작동 가능하게 연결될 수 있다.
- [0038] 일 측면에서 본 발명은 관심 뉴클레오타이드(전이유전자) 및 트립토판 RNA-결합 감쇠 단백질(TRAP) 결합 부위를 포함하는 핵산 서열을 제공하며, 여기서 TRAP 결합 부위는 전이유전자 시작 코돈 ATG와 중첩된다.
- [0039] Kozak 서열/중첩되는 Kozak 서열에 관한 본 명세서의 임의의 개시는 시작 코돈의 ATG를 지칭하는 등가 측면 및 그와의 중첩에 동등하게 적용가능하다(적절한 경우).
- [0040] 또 다른 측면에서, 본 발명은 관심 뉴클레오타이드 및 Kozak 서열을 포함하는 핵산 서열을 제공하며, 여기서 상기 Kozak 서열은 트립토판 RNA-결합 감쇠 단백질(TRAP) 결합 부위의 일부를 포함한다.
- [0041] 일 측면에서 본 발명은 관심 뉴클레오타이드(전이유전자) 및 TRAP 결합 부위를 포함하는 핵산 서열을 제공하며, 여기서 TRAP 결합 부위는 전이유전자 시작 코돈 ATG의 일부를 포함하거나 그 반대도 마찬가지이다.
- [0042] 본 발명은 또한 본원에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열을 포함하는 발현 카세트를 제공한다.
- [0043] 본 발명은 또한 뉴클레오타이드 서열 세트를 포함하는 바이러스 벡터 생산 시스템을 제공하며, 여기서 뉴클레오타이드 서열은 gag-pol, env, 선택적으로 rev, 및 본원에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 포함하는 벡터 구성요소를 암호화한다.

- [0044] 일 측면에서 본 발명은 또한 본원에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열, 발현 카세트 또는 바이러스 벡터 생산 시스템을 포함하는 세포를 제공한다.
- [0045] 렌티바이러스 벡터를 생산하기 위한 세포는 다음을 포함할 수 있다:
- [0046] (i) a) gag-pol 및 env, 및 선택적으로 rev를 포함하는 벡터 구성요소를 암호화하는 뉴클레오티드 서열, 및 본원에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열 또는 본원에 정의된 바와 같은 발현 카세트; 또는
- [0047] b) 본원에 정의된 바와 같은 바이러스 벡터 생산 시스템; 및
- [0048] (ii) 선택적으로 변형된 U1 snRNA를 암호화하는 뉴클레오티드 서열; 및
- [0049] (iii) 선택적으로 TRAP를 암호화하는 뉴클레오티드 서열.
- [0050] 세포에서, 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및/또는 스플라이스 공여자 영역으로부터의 스플라이싱 활성은 예를 들어 렌티바이러스 벡터 생산 동안 억제되거나 제거될 수 있다. 일 측면에서 관심 뉴클레오티드의 번역은 렌티바이러스 벡터 생산 동안 억제된다.
- [0051] 본 발명은 또한 하기 단계를 포함하는 렌티바이러스 벡터의 생산 방법을 제공한다:
- [0052] (i) a) gag-pol 및 env, 및 선택적으로 rev를 포함하는 벡터 구성요소를 암호화하는 뉴클레오티드 서열, 및 본원에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열 또는 본원에 정의된 바와 같은 발현 카세트; 또는
- [0053] b) 본원에 정의된 바와 같은 바이러스 벡터 생산 시스템을
- [0054] 세포 내로 도입하는 단계;
- [0055] (ii) 선택적으로, 벡터 구성요소 및 렌티바이러스 벡터의 RNA 게놈을 암호화하는 상기 뉴클레오티드 서열을 포함하는 세포를 선택하는 단계;
- [0056] (iii) 렌티바이러스 벡터의 생산에 적합한 조건 하에 세포를 배양하는 단계.
- [0057] 방법은 TRAP를 암호화하는 뉴클레오티드 서열을 세포 내로 도입하는 것을 추가적으로 포함할 수 있다.
- [0058] 방법은 변형된 U1 snRNA를 암호화하는 뉴클레오티드 서열을 도입하는 것을 추가적으로 포함할 수 있다.
- [0059] 본 발명은 또한 본원에 기재된 임의의 방법에 의해 생산되는 렌티바이러스 벡터로 확장된다.
- [0060] 일 측면에서 본 발명은 렌티바이러스 벡터를 생산하기 위한, 또는, 예를 들어 형질감염된 세포 또는 형질도입된 세포에서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위 및/또는 스플라이스 공여자 영역으로부터의 스플라이싱 활성을 억제하거나 제거하기 위한, 본원에 정의된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오티드 서열, 본원에 정의된 바와 같은 발현 카세트, 본원에 정의된 바와 같은 바이러스 벡터 생산 시스템, 또는 본원에 정의된 바와 같은 세포의 용도를 제공한다.

도면의 간단한 설명

- [0061] **도 1:** U1 snRNA 분자의 개략도 및 본 발명에서 사용하기 위한 표적화 서열을 변형하는 방법의 예시. 내인성 비-코딩(non-coding) RNA인 U1 snRNA는 인트론 스플라이싱의 초기 단계에서 5'-(AC)UUACCUG-3'(회색으로 강조 표시됨) 천연 스플라이스 공여자 표적화 서열을 통해 컨센서스(consensus) 스플라이스 공여자 부위(5'-MAGGURR-3')에 결합한다. 스템 루프 I은 polyA 억제에 중요한 것으로 밝혀진 U1A-70K 단백질에 결합한다. 스템 루프 II는 U1A 단백질에 결합하고, 5'-AUUUGUGG-3' 서열은 Sm 단백질에 결합하는데, 이는 스템 루프 IV와 함께 U1 snRNA 프로세싱에 중요하다. 본 발명에서, 변형된 U1 snRNA는 천연 스플라이스 공여자 표적화 서열의 부위에서 벡터 게놈 vRNA 분자 내의 표적 서열에 상보적인 이중 서열을 도입하도록 변형되는데; 이 도면에서 주어진 예는 변형된 U1 snRNA를 표준 HIV-1 렌티바이러스 벡터 게놈의 15개 뉴클레오티드(벡터 게놈 분자의 첫 번째 뉴클레오티드에 대해 256-270, 256U1)에 대해 지정한다(패키징 신호인 경우 SL1 루프에 위치).

도 2: HIV-1 기반 렌티바이러스 벡터 내 주요 스플라이스 공여자 부위(MSD)로부터의 비정상적인 스플라이싱의 의미. **A.** 패키징 신호의 스템 루프(SL2) 내에 내장된 기능적 주요 스플라이스 공여자를 포함하는 3세대(자가-불활성화(Self-inactivating)(SIN)) 렌티바이러스 벡터 발현 카세트의 전형적인 구성과 렌티바이러스 벡터 생산

동안 생성된 mRNA 유형을 보여주는 개략도. '표준' 렌티바이러스 벡터(LV) DNA 카세트 및 MSD로부터의 난잡한(promiscuous) 활성을 억제하거나 제거하는 MSD 영역의 기능적 돌연변이(들)을 가지는 렌티바이러스 벡터 DNA 카세트('MSD-KO LV DNA 카세트')로부터 생성된 mRNA 유형을 나타내었다. 두 카세트 모두의 경우에서, 전장의('스플라이싱되지 않은(Unspliced)') 벡터 RNA(vRNA)는 rev의 공동-발현의 결과로서, rev는 rev 반응 요소(rev response element, RRE)에 결합하며 일반적으로 MSD에서부터 RRE 서열과 함께 포함된 스플라이스 수용체 7(sa 7)까지의 스플라이싱을 억제하는 것으로 믿어진다. 표준 렌티바이러스 벡터 DNA 카세트의 경우, rev가 없는 경우, 일반적으로 모든 인트론의 스플라이싱-아웃(splicing-out)이 효율적으로 발생한다고 믿어진다('스플라이싱된'). 그러나 '비정상적인(aberrant)' 스플라이스 생성물이 렌티바이러스 벡터 생산 중에 만들어질 수 있으며, 여기서 MSD는 스플라이스 수용체 부위 또는 잠복 스플라이스 수용체 부위에 매우 효율적으로 스플라이싱하며('비정상적' 스플라이싱된'), 전형적으로 RRE가 포함된 인트론을 '간과(over-looking)'함으로써 rev가 MSD의 이러한 활성에 최소한의 영향을 미치도록 한다. 렌티바이러스 벡터 생산은 또한 MSD-돌연변이된 렌티바이러스 벡터 DNA 카세트의 패키징 영역으로 재지정된 변형된 U1 snRNA의 공동-발현으로 수행될 수 있다. (기호 풀이: Pro, 프로모터; 5'R에서 gag까지의 영역은 패키징 요소 {Ψ}를 포함함; msd, 주요 스플라이스 공여자; cppt, 중앙 폴리퓨린 트랙(central polypurine tract); Int, 인트론; sd/sa, 스플라이스 공여자/수용체; GOI, 관심 유전자; 회색 화살표는 3세대 렌티바이러스 벡터 생산 동안 생성된 스플라이싱되지 않은 vRNA의 비율을 평가하기 위한 정방향 {f} 및 역방향 {r} 프라이머의 위치를 나타냄; 전사 후 조절 요소(Post-transcriptional regulatory element) {PRE}는 명확성을 위해 표시하지 않음). **B.** 표준 3세대 렌티바이러스 벡터 생산을 HEK293T 세포에서 +/- rev로 수행하였으며 생산 후 세포에서 총 RNA를 추출하였다. 총 RNA에 대하여 2개의 프라이머 세트(A에 표시된 위치)를 사용하여 qPCR(SYBR green)을 수행하였다: f+rT는 렌티바이러스 벡터 발현 카세트로부터 생성된 전체 전사체를 증폭하였고, f+rUS는 스플라이싱되지 않은 전사체를 증폭하였다; 따라서 스플라이싱되지 않은-대-전체(Unspliced-to-Total) vRNA 전사체의 비율을 계산하고 플롯팅하였다. 데이터는 표준 3세대 렌티바이러스 벡터 생산 동안 전체 대비 스플라이싱되지 않은 vRNA의 비율이 미미(modest)하고 내부 전이유전자 카세트(이 경우 서로 다른 프로모터 및 GFP 유전자를 포함함)에 따라 다르다는 것을 나타낸다; 게다가, 이 비율은 rev의 작용에 의해 최소한으로만 증가한다.

도 3: 세 가지 다른 프로모터-GFP 발현 카세트(EF1a, EFS 및 CMV)를 포함하는 HIV-1 렌티바이러스 벡터 계통을 MSD를 기능적으로 돌연변이시키도록 변형시켜 'MSD-2KO' 렌티바이러스 벡터 계통 또는 벡터를 생산하였다(돌연변이에 대한 설명은 도 10A 참조). 표준 프로토콜에 따라 HEK293T 세포에서 벡터를 생산하였고 적정하였다. 데이터는 MSD의 기능적 돌연변이('MSD-2KO')가 렌티바이러스 벡터 역가를 최대 100배까지 감소시킨다는 것을 보여준다.

도 4: **A** EF1a-GFP 내부 발현 카세트를 암호화하는 표준 또는 MSD-돌연변이된 렌티바이러스 벡터 발현 카세트의 구성과 렌티바이러스 벡터 생산 동안 생성된 mRNA 유형을 보여주는 개략도. (기호 풀이: Pro, 프로모터; 5'R에서 gag까지의 영역은 패키징 요소 {Ψ}를 포함함; msd, 주요 스플라이스 공여자; cppt, 중앙 폴리퓨린 트랙; Int, 인트론; sd/sa, 스플라이스 공여자/수용체; GOI, 관심 유전자; 회색 화살표는 3세대 렌티바이러스 벡터 생산 동안 생성된 스플라이싱되지 않은 vRNA의 비율을 평가하기 위한 정방향 {f} 및 역방향 {r} 프라이머의 위치를 나타냄; 전사 후 조절 요소 {PRE}는 명확성을 위해 표시하지 않음). **B i** 표준 렌티바이러스 벡터 또는 MSD-2KO 렌티바이러스 벡터를 HEK293T 세포 +/- tat, 또는 179U1, 또는 305U1에서 생산하였으며, 적정하였다. **ii** 총 세포질 mRNA를 생산 후 세포로부터 추출하였고 SL2 스플라이싱 영역 SL2 스플라이싱 영역에서부터 EF1a 스플라이스 수용체까지의 주요 '비정상적인' 스플라이스 생성물을 검출할 수 있는 프라이머 (f+rG)를 사용하여 RT-PCR/겔 전기영동으로 분석하였다. 데이터는 MSD-2KO 렌티바이러스 벡터 계통(vRNA)의 5' 패키징 영역으로 재지정된 변형된 U1 snRNA가 tat와 유사한 방식으로 표준 및 MSD-2KO 렌티바이러스 벡터 둘 다의 역가를 증가시킬 수 있음을 보여준다. MSD-2KO 돌연변이는 SL2 스플라이싱 영역에서부터 EF1a 스플라이스 수용체까지의 '비정상적인' 스플라이스 생성물의 검출을 없앴다(도 4A 참조). 중요하게도, 변형된 U1 snRNA에 의한 역가의 증가는, tat의 사용과는 대조적으로, '비정상적인' 스플라이스 생성물이 사실상 검출되지 않는 상태(maintenance)를 동반하였다.

도 5: EF1a, EFS 또는 CMV 프로모터에 의해 구동되는 GFP 내부 카세트를 암호화하는 표준 렌티바이러스 벡터 또는 MSD-돌연변이된 렌티바이러스 벡터를 HEK293T 세포 +/- 256U1에서 생산하였고, 적정하였다. 5' 패키징 영역으로 재지정된 변형된 U1 snRNA의 사용에 의한 향상된 렌티바이러스 벡터 역가는 전이유전자 카세트 내에서 사용되는 프로모터와 무관하다. 데이터는, 대부분에서, MSD-2KO 돌연변이의 감쇠(attenuating) 표현형이 변형된 U1 snRNA의 공동-발현에 의해 구제되고, 따라서 이는 놀랍게도 표준 렌티바이러스 벡터 계통과 비교하여 불균형

적으로 MSD-돌연변이된 렌티바이러스 벡터 게놈의 역가를 증가시킨다는 것을 보여준다.

도 6: 5' 패키징 영역으로 재지정된 변형된 U1 snRNA의 사용에 의한 MSD-돌연변이된 렌티바이러스 벡터 역가의 향상은 5'LTR 내의 5'polyA 신호의 잠재적 활성의 억제와 연결되지 않는다. 이전 보고서는 MSD의 돌연변이가 HIV-1 프로바이러스 및 '미니-리포터' 카세트의 5'LTR의 5'R 서열 내의 polyA 신호를 활성화하여 전사의 조기 종료를 유발할 수 있음을 보여준다; 내인성 U1 snRNA와 심지어 재지정된 U1 snRNA의 결합이 이 polyA 활성을 차단할 수 있다. **A** 두 개의 polyA 신호 돌연변이(pAM1 = AAUAAA > AACAAA; pAKO = AAUAAA의 결실)와 야생형 polyA 신호(wt pA = AAUAAA)가 전사 리드-스루 HIV-1 polyA 부위에 미치는 영향을 평가하도록 GFP-polyA-Gluciferase 리포터 카세트를 설계하였다. 리드-스루 HIV-1 polyA 신호를 무시되라게 활성화에 의해 측정할 수 있었고, 이를 GFP 발현에 의해 정규화하였다. **B** 변형된 U1 snRNA가 유사한 방식으로 작용하는지 테스트하기 위해, 5'polyA 신호의 기능적 돌연변이(pAm1)를 EF1a-GFP 또는 CMV-GFP 발현 카세트를 포함하는 MSD-돌연변이된 렌티바이러스 벡터 게놈에 도입하였다. EF1a-GFP 또는 CMV-GFP 발현 카세트를 포함하는 표준 및 MSD-돌연변이된 렌티바이러스 벡터 게놈도 사용하였다. 렌티바이러스 벡터를 HEK293T 세포 +/- 305U1에서 생산하였고, 적정하였다. 데이터는 5'polyA 신호의 기능적 제거가 렌티바이러스 벡터 역가의 매우 미미한 증가만을 초래하였으며, 따라서 변형된 U1 snRNA에 의해 제공되는 관찰된 렌티바이러스 벡터 역가의 증가가 - 특히 MSD-2KO/polyA-돌연변이된 렌티바이러스 벡터 게놈에서의 그러한 증가가 - 5'polyA 활성의 억제에 기인할 수 없다는 점을 나타낸다.

도 7: 벡터 게놈의 SL1에 대한 U1-70K 단백질 결합, SL2에 대한 U1A 단백질 결합 또는 SL4에 대한/근처의 Sm 단백질 결합을 제거하는 것으로 알려진 여러 개의 돌연변이를 305U1 및 256U1 변형된 U1 snRNA에 도입하였다. EF1a-GFP 내부 카세트를 암호화하는 표준 또는 MSD-돌연변이된 렌티바이러스 벡터를 이러한 돌연변이된 변형된 U1 snRNA의 존재 하에 생산하고 적정하였고, 역가 값을 변형된 U1 snRNA 없이 생산된 표준 렌티바이러스 벡터에 대해 정규화하였다. 데이터는 변형된 U1 snRNA에 의한 MSD-2KO 렌티바이러스 벡터 역가의 향상이 U1-70K 단백질 또는 U1A 단백질 결합에 의존하지 않고 Sm 단백질 결합 부위에 의존한다는 것을 보여준다. 따라서, 5' 패키징 영역으로 재지정된 변형된 U1 snRNA를 사용하여 MSD-2KO 렌티바이러스 벡터 역가를 향상시키는 것은 U1 snRNA의 알려진 기능과 연결되지 않는다.

도 8: 다양한 길이의 표적화 서열을 포함하는 변형된 U1 snRNA의 사용에 의한 MSD-돌연변이된 렌티바이러스 벡터 역가의 향상. EF1a-GFP 카세트를 함유하는 MSD-2KO 렌티바이러스 벡터를 '305' 영역을 표적으로 하는 변형된 U1 snRNA의 존재 하에 HEK293T 세포에서 생산하였으며, 여기서 각각의 변형된 U1 snRNA는 상보성에서 상이한 길이의 재표적화 서열을 포함하였다. 7-15개 뉴클레오타이드의 상보성 길이를 포함하는 변형된 U1 snRNA를 사용할 때 역가 증가가 관찰되었으며, 10개 이상의 뉴클레오타이드에서 최대 효과가 관찰되었다.

도 9: 변형된 U1 snRNA를 벡터 게놈 RNA의 패키징 영역으로 표적화할 때 MSD-돌연변이된 렌티바이러스 벡터의 최대 역가 회복/부스트가 관찰된다. EF1-GFP 카세트를 포함하는 MSD-2KO 렌티바이러스 벡터를 상보성의 15개 뉴클레오타이드 길이(또는 표시된 경우 9개의 뉴클레오타이드)를 포함하는 벡터 게놈 vRNA 분자의 5'말단의 길이에 따른 부위에 대한 표적화 서열을 가지는 변형된 U1 snRNA의 존재 하에 생산하였다. 변형된 U1 snRNA는 벡터 게놈 vRNA 분자의 5' 말단의 길이를 따라 표적화 서열 부위의 첫 번째 뉴클레오타이드에 따라 명명하였다. 각각의 변형된 U1 snRNA에 대한 데이터 막대는 벡터 게놈 vRNA의 5'말단 내에서 각 공지된 기능적 서열의 대략적인 표지된 위치 아래에 정렬하였다(축척이 아님).

도 10: 기능적 주요 스플라이스 공여자 돌연변이, 렌티바이러스 벡터 역가에 대한 이들의 영향 및 변형된 U1 snRNA에 의한 회복에 대한 설명. **A** '야생형' HIV-1(NL4-3; 현재 렌티바이러스 벡터 게놈 내의 '표준' 서열)의 스템 루프 2(SL2) 영역의 서열을 상단에 나타내었다. 서열은 주요 스플라이스 공여자 부위(MSD: 컨센서스 = CTGGT) 및 잠복 스플라이스 공여자 부위(이는 MSD 부위가 자체적으로 돌연변이될 때 활용됨(crSD: 컨센서스 = TGAGT))를 포함한다. 스플라이스 공여자 부위가 사용될 때 스플라이싱 위치의 뉴클레오타이드를 **볼드체** 및 화살표로 식별하였다. MSD 및 crSD 부위 스플라이싱 활성을 모두 제거하는 4가지 기능적 MSD 돌연변이가 기재되어 있다: MSD 및 crSD 부위로부터 2개의 'GT' 모티프를 돌연변이시키고 (대부분의 실시예에서 널리 사용되는) MSD-2KO; MSD 및 crSD 부위를 모두 제거하는 돌연변이를 또한 포함하는 MSD-2KOv2; 스플라이스 공여자 부위가 전혀 없는 완전히 새로운 스템-루프 구조를 도입한 MSD-2KOm5; 및 SL2 서열을 완전히 결실시키는 ΔSL2. MSD-2KO, MSD-2KOv2 및 MSD-2KOm5 돌연변이에서 SL2 서열에 도입된 치환은 소문자 *이탤릭체*로 표시하였다. **B** 기능적 MSD 돌연변이를 포함하는 4개의 렌티바이러스 벡터 게놈 변이체(도 10A에 기재됨)를 EFS-GFP 내부 카세트를 사용하여 클로닝하였고, MSD-2KO 또는 MSD-2KOm5 변이체는 EF1a-, CMV- 또는 huPGK-GFP 내부 카세트를 사용하여 추가로 클로닝하였다. 표준 및 MSD-돌연변이된 LV를 HEK293T 세포 +/- 256U1에서 생산하였고, 적정하였다. 데이터는 렌티바이러스 벡터 역가의 감쇠 정도가 특정 돌연변이에 따라 달라질 수 있으며, MSD-2KOm5 변이체가 일반적인

로 덜 감쇠된 표현형을 생성함을 나타낸다. 변형된 U1 snRNA는 생산 동안 공동 발현될 때 기능적 MSD 돌연변이를 포함하는 4개의 렌티바이러스 벡터 게놈 변이체에 대하여 렌티바이러스 벡터 역가를 증가시킬 수 있었다. 역가 증가는 256U1이 MSD-2KOm5 서열을 보유하는 MSD-돌연변이된 LV 게놈과 함께 발현되었을 때 가장 컸다.

도 11: 변형된 U1 snRNA 발현 카세트는 일시적 형질감염 프로토콜에서 사용하기 쉽도록 렌티바이러스 벡터 게놈 플라스미드 백본에 위치할 수 있다. 많은 실시예에서는 렌티바이러스 벡터의 생산에서 렌티바이러스 벡터 구성 요소 플라스미드와의 공동-형질감염으로 별도의 변형된 U1 snRNA를 발현하는 플라스미드를 사용한다. 일시적 형질감염 동안 시스로(in cis) 변형된 U1 snRNA 카세트를 제공할 수 있도록 렌티바이러스 벡터 게놈 플라스미드 백본에서 '허용' 부위를 식별하기 위해, 3개의 변이체를 클로닝하였다. **A** 일시적 형질감염 동안 시스로 변형된 U1 snRNA 카세트를 제공하는 렌티바이러스 벡터 게놈 변이체의 개략도. 버전 1('[cis] ver1') 및 버전 3('[cis] ver3')은 변형된 U1 snRNA 카세트가 렌티바이러스 벡터 게놈 카세트에 대해 역전(inverted)되도록 내성 마커와 복제 기점 사이에 변형된 U1 snRNA 카세트를 배치하였고(내성 마커 방향은 ver1과 ver3간에 다름), 버전 2('[cis] ver2')는 변형된 U1 snRNA 카세트를 렌티바이러스 벡터 게놈 카세트의 동일한 방향으로 업스트림에 배치하였다. (기호 풀이: Pro, 프로모터; 5'R에서 gag까지의 영역은 패키징 요소 {Ψ}를 포함함; msd, 주요 스플라이스 공여자 {여기에서는 MSD-2KO로 표시됨}; RRE, rev 반응 요소; cppt, 중앙 폴리퓨린 트랙; Transgenic, 치료 페이로드를 포함하는 이중 서열; U1-Pro, U1 프로모터; Term[3'box], U1 전사 종결자). **B** EF1a-GFP 카세트를 포함하는 MSD-2KO 렌티바이러스 벡터 게놈 플라스미드의 세 가지 '시스' 버전을 사용하여 '트랜스(trans)' 접근법과 병행하여 HEK293T 세포에서 렌티바이러스 벡터를 생산하였고, 상기 '트랜스' 접근법에서는 동일한 MSD-2KO 렌티바이러스 벡터 게놈(백본에 변형된 U1 snRNA 카세트가 삽입되지 않음)을 별도의 플라스미드의 공동-형질감염에 의해 공급되는 변형된 U1 snRNA +/-에서 생산하였다. 데이터는 별도의 변형된 U1 snRNA를 암호화하는 플라스미드의 공동-형질감염과 유사하게 MSD-2KO 렌티바이러스 벡터 역가가 '시스' 렌티바이러스 벡터 게놈의 사용에 의해 증가할 수 있음을 보여준다.

도 12: 렌티바이러스 벡터 생산 동안 전이유전자를 발현하는 비정상적으로 스플라이싱된 mRNA가 MSD-2KO 렌티바이러스 벡터에서 제거되어, TRiP 시스템을 사용할 때 TRAP에 의해 표적화되는데 필요한 전이유전자 mRNA의 양을 감소시킨다. **A** EF1a-GFP 전이유전자 카세트를 암호화하는 'TRiP' 렌티바이러스 벡터 게놈의 개략도로서, 여기서 TRAP 결합 부위(tbs)는 카세트의 5'UTR 내에 위치함(벡터 생산 동안 TRAP의 공급은 전이유전자 발현 수준을 감소시킴). MSD-2KO 렌티바이러스 벡터를 생산하는 동안, 전장의 스플라이싱되지 않은 패키징가능한 vRNA 및 전이유전자 mRNA가 렌티바이러스 벡터 카세트로부터 생산된 세포질 RNA의 주요 형태이다 (i)(전이유전자 프로모터가 생산 중 활성화된 경우). 그러나, 표준 렌티바이러스 벡터 게놈에서 MSD의 난잡한 활성화로 인해 전이유전자를 암호화할 수 있는 추가적인 '비정상적인' 스플라이스 산물이 생성된다 (ii); 이것은 내부 전이유전자 프로모터, 즉 조직 특이적 프로모터와 독립적으로 발생할 수 있다. (기호 풀이: Pro, 프로모터; 5'R에서 gag까지의 영역은 패키징 요소 {Ψ}를 포함함; msd, 주요 스플라이스 공여자; cppt, 중앙 폴리퓨린 트랙; Int, 인트론; sd/sa, 스플라이스 공여자/수용체; GOI, 관심 유전자; 회색 화살표는 3세대 렌티바이러스 벡터 생산 동안 생성된 스플라이싱되지 않은 vRNA의 비율을 평가하기 위한 정방향 {f} 및 역방향 {r} 프라이머의 위치를 나타냄; 전사 후 조절 요소 {PRE}는 명확성을 위해 표시하지 않음). **B** EF1a-GFP 카세트를 포함하는 표준 또는 MSD-2KO 렌티바이러스 벡터 게놈 플라스미드를 사용하여 HEK293T 세포에서 렌티바이러스 벡터를 생성하고, GFP 발현 점수를 생성하였다(%GFP × MFI). 표준 렌티바이러스 벡터 생산 동안 배양물에서 생산된 GFP의 총량에 비해, MSD-2KO 변형은 TRAP가 없는 경우에도 생산된 GFP의 양을 감소시키는 실질적인 효과가 있었다. 따라서, TRAP의 억제 효과는 MSD-2KO 렌티바이러스 벡터 게놈의 사용에 의해 증대되었고, 이는 배양물에서 훨씬 더 낮은 수준의 GFP로 이어졌다.

도 13: 표준 또는 MSD-2KO 렌티바이러스 벡터의 증가를 가능하게 하는 변형된 U1 snRNA를 안정적으로 발현하는 HEK293T 세포의 성공적인 분리는 변형된 U1 snRNA 카세트가 렌티바이러스 벡터 패키징 및 생산자 세포주에 도입될 수 있음을 보여준다. EFS-GFP 카세트를 포함하는 표준 또는 MSD-2KO 렌티바이러스 벡터 게놈을 HEK293T 또는 HEK293T.305U1(9nt 변이체) 세포 +/- 추가의 305U1 플라스미드에서 생산하였다. 데이터는 변형된 U1 snRNA를 발현하는 안정적인 카세트가 독성 없이 세포에 도입될 수 있음을 나타낸다.

도 14. tbs를 ATG 개시 코돈에 더 가깝게 배치하기 위해 전이유전자 5'UTR 내의 tbs의 3' 말단과 중첩되는 최적의 Kozak 서열의 식별.

A. Kozak이 tbs의 3'말단과 중첩되어 KAGNN 반복부(들)이 유지되도록 배치된 코어 컨센서스 'RVVATG' 및 더 넓은 컨센서스 'GNNRVVATG'를 준수하는 조각된 변이체에서 Kozak 서열의 위치와 서열을 나타내는 개략도. 이것은 tbs가 ATG 개시 코돈에 더 가깝게 위치하도록 하여 TRAP-tbs 복합체 내에 ATG 개시 코돈을 '숨김'으로써 개선된

전이유전자 억제 수준(+TRAP)을 가능하게 하는 tbs-Kozak 접합 변이체를 식별할 수 있게 한다. 컨센서스 Kozak 서열의 유지는 전이유전자 비억제 수준(TRAP 없음)이 높은 수준에서 유지될 수 있도록 한다(즉, 벡터-형질도입된 세포에서의 발현을 모델링).

B. 각각 pBlueScript(TRAP 없음) 또는 pEF1 α -TRAP(TRAP)와 함께 리포터 플라스미드를 HEK293T 세포에 공동-형질감염시켜 GFP 발현의 비억제 또는 억제 수준에 대해 리포터를 테스트하였다. 형질감염된 세포를 형질감염 2일 후 유세포 분석에 의해 분석하고, GFP 발현 점수(% GFP $\times$ 중앙 형광 강도(median fluorescence intensity))를 생성하고 log-10으로 전환하였다. 모든 tbs-Kozak 접합(junction) 변이체 리포터는 원래 구성과 비교하여 동일한 비억제 GFP 수준을 유지하였다. 변이체 '0', '2' 및 '3'은 원래 구성에 비해 개선된 억제 수준을 나타내었다(표준 편차 막대, n=3).

도 15. 중첩되는 tbs-Kozak 변이체를 사용하여 AAV 벡터 게놈 플라스미드에서 전이유전자 억제의 개선. 2개의 'tbs-Kozak' 변이체(0 및 3)를 EFS 또는 huPGK 프로모터 GFP 리포터 카세트에 클로닝하였고, 추가로 L33 또는 L12 개선된 리더 서열을 포함하였다. 중첩되지 않는 tbs/Kozak 변이체도 EFS/huPGK-L33 카세트에 클로닝하였다; 이들은 tbs-Kozak 영역에서만 달랐다(Original = [tbs]-ACAGCCACCATG; HpaI 변이체 = [tbs-GAGTT]AACGCCACCATG). 각각 pBlueScript(TRAP 없음) 또는 pEF1 α -TRAP(TRAP)와 함께 리포터 플라스미드를 공동-형질감염시켜 GFP 발현의 비억제 또는 억제 수준에 대해 리포터를 테스트하였다. 형질감염된 HEK293T 세포(현탁액, 무혈청)를 형질감염 2일 후 유세포 분석에 의해 분석하고, GFP 발현 점수(% GFP $\times$ 중앙 형광 강도)를 생성하고 log-10으로 전환하였다. 데이터는 tbs를 Kozak 서열과 중첩시키면 비중첩 tbs/Kozak 변이체와 비교하여 TRAP에 의한 전이유전자 발현의 개선된 억제를 허용한다는 것을 입증한다. (표준 편차 막대, n=3).

도 16. 전장의 EF1a 프로모터의 맥락에서 TRAP-매개 전이유전자 억제의 개선.

A. 3개의 'tbs-Kozak' 변이체(0, 2 및 3)를 EF1a 프로모터 GFP 리포터 카세트에 클로닝하였다. 스플라이싱 후 리더 서열은 L33 서열(엑손1)과 tbs의 바로 상류에 있는 엑손2로부터의 짧은 12 nt 서열을 포함한다.

B. 각각 pBlueScript(TRAP 없음) 또는 pEF1 α -TRAP(TRAP)와 함께 리포터 플라스미드를 공동-형질감염시켜 GFP 발현의 비억제 또는 억제 수준에 대해 리포터를 테스트하였다. 형질감염된 HEK293T 세포(현탁액, 무혈청)를 형질감염 2일 후 유세포 분석에 의해 분석하고, GFP 발현 점수(% GFP $\times$ 중앙 형광 강도)를 생성하고 log-10으로 전환하였다.

C. GFP 전이유전자 카세트를 HIV-1 렌티바이러스 벡터 게놈에 클로닝하고, B에 설명된 바와 같이 GFP 발현의 비억제 또는 억제 수준에 대해 테스트하였다. 데이터는 tbs를 Kozak 서열과 중첩시키는 것이 TRAP에 의한 전이유전자 발현의 개선된 억제를 허용한다는 것을 입증한다. (표준 편차 막대, n=3).

도 17. 현탁액(무혈청) HEK293T 세포에서 중첩되는 tbs-Kozak 변이체의 TRAP-매개 전이유전자 억제 테스트. 표 IV의 중첩되는 tbs-Kozak 변이체를 pEF1a-GFP 리포터 플라스미드 내로 클로닝하고 HEK293T 세포 +/- pTRAP에 형질감염시키고, 형질감염 2일 후에 유세포 분석을 수행하였다. **A.** GFP 발현 점수(% GFP 양성 $\times$ MFI)를 +/- TRAP에 대해 생성하고, 배수 억제 값을 생성하고 플롯팅하였다. 변이체를 x축을 따라 표시하였으며, 3' tbs KAGNN 반복부와 코어 Kozak 서열의 상대적 중첩에 따라 그룹화하였다('중첩 그룹' - KAGatg, KAGNatg, KAGNNatg); KAGNN은 검은색 대괄호로 표시하였고 코어 Kozak 뉴클레오티드는 회색 선으로 표시하였다. 하기의 중첩 그룹을 비교하여 통계 분석을 수행하였다(중첩 그룹 내의 등분산은 F-Test에 의해 확인하였다): 배수-억제는 양측 T-검정을 사용하여 KAGNatg보다 KAGatg에 대해 통계적으로 더 컸고(*p= 0.0293); KAGNNatg보다 KAGNatg에 대해 더 컸으며(**p=0.00000482); 비-중첩 tbs보다 KAGNNatg에 대해 더 컸다(***p= 0.000259). **B.** 비-억제 GFP 발현 점수를 최고에서 최저(왼쪽에서 오른쪽)로 플롯팅하였고, 두 개의 KAGatg 중첩 그룹 변이체 tbskzkV0.G 및 tbskzkV0.T(A에서 모든 변이체 중 가장 큰 억제를 나타냄)를 강조표시하였는데 'G' 변이체가 'T' 변이체보다 선호됨을 보여주며 이는 전자가 더 나은 'ON'(비-억제) 수준을 가지기 때문이다.

도 18. 최적의 중첩되는 tbs-Kozak 변이체를 사용한 인트론-함유 프로모터의 개선된 억제.

A. 중첩되지 않는 tbs-Kozak 변이체와 비교하여 중첩되는 tbs-Kozak 변이체의 사용을 예시하는 데 사용되는 발현 카세트의 개략도. 널리 사용되는 EF1a 프로모터 서열은 널리 사용되는 CAG 프로모터와 마찬가지로 자체 인트론을 포함한다(도 10 및 실시예 5 참조). CAG 프로모터는 CMV 인핸서, 코어 프로모터 및 닭 β -액틴 유전자로부터의 엑손1/인트론 서열 및 토끼 β -글로빈 유전자로부터의 스플라이스 수용체/엑손 서열을 포함하는 매우 강력한 인공 프로모터이다. 이 작업에서, EF1a 프로모터(엑손 1[L33] 포함)로부터의 'EF1a-INT' 서열, 모든 EF1a 인트론 및 스플라이스 수용체, 및 EF1a 엑손 2로부터의 12개 뉴클레오티드를 CAG 프로모터 내로 클로닝하여, CAG

엑손/인트론 서열을 대체하였다. 'EF1a-INT' 서열은 또한 CMV 프로모터 작제물 내로 클로닝하였다. B. 바이러스 벡터 생산 동안 전이유전자 발현을 모델링하기 위해 현탁액(무혈청) HEK293T 세포에서 GFP 발현 및 TRAP에 의한 억제에 대해 작제물을 평가하였다. GFP 발현 점수(%GFP × MFI)를 생성하고 플롯팅하였으며, 뿐만 아니라 TRAP의 존재 하에서 배수 억제 점수를 표시하였다.

도 19. tbs의 하류의 5'UTR 서열에 대한 개선 사항의 개요.

A. TRAP-tbs 억제가 전이유전자 카세트의 5'UTR 암호화 영역의 DNA 발현 카세트를 보여주는 개략도로써, 다중 클로닝 부위(MCS)가 tbs와 전이유전자의 개시 코돈 사이에 삽입되어 있음(TRAP은 도넛 모양으로 표시됨). 본 발명은 tbs의 말단 KAGNN 반복부에서/상에서 시작하고 전이유전자 개시 코돈의 상류에 최대 5개의 클로닝 부위를 함유하는 선호되는 중첩되는 제한 효소 부위를 기술한다.

B. 전이유전자의 Kozak 서열이 tbs의 3' KAGNN 반복부와 대부분 또는 부분적으로 중첩되도록 위치할 수 있는 방법을 보여주는 개략도; 그렇게 하면 TRAP-tbs 복합체 내에 주요 개시 코돈을 효과적으로 '숨길' 수 있어 번역 기계(translation machinery)에 훨씬 더 접근하기 어렵게 만들어, 전이유전자 발현의 '억제된' 수준을 훨씬 더 낮출 수 있다.

C. 선호되는 중첩되는 tbs 및 Kozak 컨센서스 서열을 요약하는 표. tbs의 3' KAGNN 반복부는 박스로 표시하였고 코어 Kozak 서열은 굵게 표시하였다.

도 20: A. 렌티바이러스 벡터의 HIV-1 패키징 영역에서 스플라이스 공여자 부위의 돌연변이를 추가로 예시하기 위해 사용된 DNA 서열. 패키징 서열의 SL2 루프 영역(MSD 및 crSD1 포함), 뿐만 아니라 SL4 루프(두 개의 추가적인 마이너한 잠복 스플라이스 공여자 부위, crSD2 및 crSD3 포함)를 상자로 표시하였다. 기호폴이 - 대문자로 표기된 서열은 'wt SL2-MSD-SL4(STD)'에 따라 야생형(또는 이와 정렬됨)이며, 굵은 검정색으로 스플라이스 공여자 뉴클레오티드를 나타낸다. 'GT' 디뉴클레오티드(기능적 스플라이스 공여자 부위에 중요한 것으로 간주됨)는 존재하는 경우 굵은 회색으로 표시하였다. 대쉬는 서열의 갭을 나타내지 않으며 더 나은 명확성을 위해 정렬된 서열의 간격을 지정하는데 사용된다; 포워드 슬래시는 SL2 및 SL4 루프 사이에 표시되지 않는 서열을 나타낸다(모든 '누락된' 서열은 야생형 HIV-1에서 유래된다). 밑줄 친 서열은 SL의 스템에 있는 뉴클레오티드를 나타낸다. 소문자 이탤릭체 서열은 변형된 서열이다. '1/2/3/4KO'는 각 변이체에서 돌연변이된 스플라이스 공여자의 수를 나타낸다. B. 표준 렌티바이러스 벡터 게놈 RNA의 야생형 주요 스플라이스 공여자 부위, 뿐만 아니라 'MSD-2KO' 및 'MSD-2KOm5' 변이체에 대한 내인성 U1 snRNA의 예측된 어닐링을 보여주는 개략도. 'MSD-2KOm5' 변이체는 'MSD-2KO'(또는 실제로 야생형)보다 더 큰 안정성(더 많은 상보성)으로 내인성 U1 snRNA에 어닐링하도록 설계되었지만 여전히 스플라이스 공여자 부위에서 기능적으로 돌연변이된다. 또한, 'MSD-2KOm5'는 패키징 2차 구조에 대한 영향을 최소화하도록 스템 루프를 형성할 수 있도록 하는 플랭킹 서열을 포함한다(A 참조). C. EF1a-GFP 전이유전자 카세트 및 표준(STD) MSD 영역(야생형 서열), 또는 MSD 단독(MSD-1KO) 또는 MSD/crSD1(MSD-2KO)의 돌연변이를 포함하는 패키징 영역을 암호화하는 렌티바이러스 벡터 게놈을 작제하였다. 트랜스로 공급되는 256U1 변형된 snRNA가 있거나 없이 일시적인 형질감염에 의해 현탁액(무혈청) HEK293T 세포에서 LV를 생산하였다. 생산 후 세포에서 PolyA로 선택된 총 mRNA를 추출하고 RT-PCR/아가로스 겔 분석을 실시하여, MSD의 상류 및 EF1a 스플라이스 하류의 프라이머를 사용하여 MSD 영역으로부터의 비정상적인 스플라이싱의 전체적인 효과를 확인하였다(프라이머 위치는 도 23 참조). 겔 이미지는 SL2/SL4 영역에서부터 EF1a 스플라이스 수용체까지의 비정상적인 스플라이스 생성물의 위치/유형을 나타낸다(스플라이스 접합의 시퀀싱으로 확인됨[데이터는 표시되지 않음]). D. MSD-1KO 및 MSD-2KO LV 양자 모두의 렌티바이러스 벡터 역가는 표준 LV에 비해 감소되었지만 돌연변이된 LV의 패키징 영역을 표적으로 하는 변형된 U1 snRNA의 사용에 의해 구제되었다(log10 스케일의 Y 축).

도 21: 패키징 서열의 SL4에 있는 crSD2 부위로부터의 비정상적인 스플라이싱을 제거하기 위한 추가적 돌연변이를 SL2에 MSD-2KO 또는 MSD-2KOm5 돌연변이를 보유하는 LV 게놈에 추가하였다(도 20 참조). 트랜스로 공급되는 256U1 변형된 snRNA가 있거나 없이 일시적인 형질감염에 의해 현탁액(무혈청) HEK293T 세포에서 LV를 생산하였다. A. 생산 후 세포에서 PolyA로 선택된 총 mRNA를 추출하고 RT-PCR/아가로스 겔 분석을 실시하여, MSD의 상류 및 EF1a 스플라이스 수용체의 하류의 프라이머를 사용하여 SL2의 MSD 영역으로부터의 뿐만 아니라 SL4로부터의 비정상적인 스플라이싱의 전체적인 효과를 확인하였다(프라이머 위치는 도 23 참조). 겔 이미지는 SL2/SL4 영역에서부터 EF1a 스플라이스 수용체까지의 비정상적인 스플라이스 생성물의 위치/유형을 나타낸다(스플라이스 접합의 시퀀싱으로 확인됨[데이터는 표시되지 않음]). B. 모든 MSD-1/2/3/4KO 게놈의 렌티바이러스 벡터 역가는 표준 LV에 비해 감소되었지만 돌연변이된 LV의 패키징 영역을 표적으로 하는 변형된 U1 snRNA의 사용에 의해

구제되었다(log10 스케일의 Y 축).

도 22: 통합된 MSD-돌연변이된 LV는 표준 LV에 비해 전사 리드-스루 이벤트의 비율이 더 낮다. 렌티바이러스 벡터는 전사적으로 활성인 유전자를 선호하여 표적 세포 DNA에 반무작위로 삽입한다. 이것은 전형적으로 통합된 벡터가 세포 유전자 전사 단위 내부에 위치할 것임을 의미한다. 통합된 LV가 세포성 프로모터의 하류에 있는 경우, 일부 전사 리드-스루('리드-인(read-in)')가 LV 유닛 내로 발생할 수 있다. 현재 표준 3세대 LV는 손상되지 않은 MSD를 포함하며, 이론적으로 5'LTR 및 패키징 영역에 대한 리드-인은 RNA에 대한 내인성 U1 snRNA의 동원(recruitment)을 허용할 수 있다. 이러한 상황에서, 동원된 U1 snRNA가 5' polyA 부위에서 폴리아데닐화를 억제하여, LV 유닛으로의 추가 신장에 이르게 할 것이라고 생각할 수 있다. 또한, MSD로부터의 스플라이싱이 LV 유닛 내의 하류 스플라이스 수용체로 발생하거나 또는 심지어 트랜스-스플라이싱에 의한 세포성 전사체로 발생할 수 있다. MSD-돌연변이된 LV의 사용은 스플라이싱 능력이 있는 전구체로 이어지는 내인성 U1 snRNA를 동원할 수 없을 것으로 예상되며, 따라서 5' polyA 부위에서 폴리아데닐화를 억제하는 능력이 떨어질 수 있다. 따라서, 이 새로운 유형의 벡터에 대해 더 적은 전사 리드-인이 예상된다. 통합된 LV 카세트의 상류로부터의 리드-인 전사를 평가하기 위해 형질도입된 세포에서 추출된 총 RNA의 RT-qPCR에 사용되는 정방향(f) 및 역방향 프라이머(r)의 위치를 회색 화살표로 나타내었다.

도 23: 통합된 MSD-돌연변이된 LV가 표준 LV에 비해 전사 리드-스루 이벤트의 더 낮은 비율을 나타낸다는 증거. 실시예 S2에서 생성된 표준 LV 벡터 및 MSD-돌연변이 변이체 벡터를 사용하여 일치된 MOI에서 HEK293T 세포 또는 1차 세포 92BR을 형질도입하였다. 256U1의 존재 하에 생성된 단지 MSD 만-돌연변이된 벡터 제제를 사용하였는데, 이는 이것이 표준 LV 게놈 제제에 필적할만한 역가를 제공했기 때문이다(도 21B 참조). 형질도입된 세포를 10일 동안 계대하여 통합되지 않은 cDNA를 제거하고, 통합되지 않은 벡터 cDNA로부터 발현되었을 수 있는 벡터 RNA의 신호를 감소시켰다. 숙주 세포 RNA를 추출하고 DNase 처리한 후 RT-qPCR 분석하여 리드-인 전사체를 검출하였다(프라이머 위치는 도 22 참조). 검출된 HIV Psi RNA 카피를 GAPDH 신호(로딩)의 검출 및 통합된 벡터의 DNA 카피 수에 대해 정규화하였으며, 이는 별도로 수행하였다. 256U1 없이 생성된 표준 LV에 대해 검출된 HIV Psi DNA 카피를 1로 설정하였고, 다른 모든 데이터 포인트를 이것에 대해 상대적으로 설정하였다. 모든 MSD-돌연변이된 벡터와 표준 벡터의 통계적 비교를 수행하였다; F-Test(critical one-tail)로 평균 분산을 평가한 후 t-Test(등분산[HEK293T] 또는 이분산[92 BR]을 가정하는 critical two-tail)를 수행하여 두 그룹 간의 유의한 차이를 나타내었다(* p= 0.00012 및 ** p= 0.000073).

발명을 실시하기 위한 구체적인 내용

[0062] RNA 스플라이싱

[0063] 본원에 기재된 바와 같이, 본 발명은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열을 제공하고, 여기서 렌티바이러스 벡터의 RNA 게놈의 주요 스플라이스 공여자 부위는 불활성화되고, 여기서 주요 스플라이스 공여자 부위의 3'의 잠복 스플라이스 공여자 부위는 불활성화된다.

[0064] RNA 스플라이싱은 5개의 작은 핵 리보핵단백질(small nuclear ribonucleoprotein, snRNP)로 구성된 스플라이세오솜(spliceosome)이라고 하는 큰 RNA-단백질 복합체에 의해 촉매된다. 인트론과 엑손 사이의 경계는 스플라이싱이 발생할 위치를 나타내는 pre-mRNA 내의 특정 뉴클레오타이드 서열로 표시된다. 이러한 경계를 "스플라이스 부위"라고 지칭한다. 용어 "스플라이스 부위"는 절단 및/또는 다른 스플라이스 부위에 결합하기에 적합한 것으로 진핵 세포의 스플라이싱 기계(splicing machinery)에 의해 인식될 수 있는 폴리뉴클레오타이드를 지칭한다.

[0065] 스플라이스 부위는 pre-mRNA 전사체에 존재하는 인트론의 절제를 허용한다. 전형적으로, 5' 스플라이스 경계는 "스플라이스 공여자 부위" 또는 "5' 스플라이스 부위"라고 지칭되고, 3' 스플라이스 경계는 "스플라이스 수용체 부위" 또는 "3' 스플라이스 부위"라고 지칭된다. 스플라이스 부위는, 예를 들어, 자연 발생 스플라이스 부위, 조작된 또는 합성 스플라이스 부위, 표준(canonical) 또는 컨센서스 스플라이스 부위, 및/또는 비-표준(non-canonical) 스플라이스 부위, 예를 들어 잠복 스플라이스 부위를 포함한다.

[0066] 스플라이스 수용체 부위는 일반적으로 분기점 또는 분기 부위, 폴리피리미딘 관(polypyrimidine tract) 및 수용체 컨센서스 서열의 세 가지 개별 서열 요소로 구성된다. 진핵생물의 분기점 컨센서스 서열은 YNYTRAC(여기서 Y는 피리미딘이고, N은 임의의 뉴클레오타이드이고, R은 퓨린임)이다. 3' 수용체 스플라이스 부위 컨센서스 서열은 YAG(여기서 Y는 피리미딘임)이다 (예를 들어, Griffiths et al., eds., *Modern Genetic Analysis*, 2nd edition, W.H. Freeman and Company, New York (2002) 참조). 3' 스플라이스 수용체 부위는 전형적으로 인트론의 3' 말단에 위치한다.

- [0067] "표준 스플라이스 부위" 또는 "컨센서스 스플라이스 부위"라는 용어는 상호교환적으로 사용될 수 있으며 종에 걸쳐 보존되는 스플라이스 부위를 지칭한다.
- [0068] 진핵생물 RNA 스플라이싱에 사용되는 5' 공여자 스플라이스 부위 및 3' 수용체 스플라이스 부위에 대한 컨센서스 서열은 당업계에 잘 알려져 있다. 이러한 컨센서스 서열은 인트론의 각 말단에 거의 변하지 않는 디뉴클레오티드를 포함한다: 인트론의 5' 말단에 GT, 및 인트론의 3' 말단에 AG.
- [0069] 표준 스플라이스 공여자 부위 컨센서스 서열은 (DNA의 경우) AG/GTRAGT(여기서 A는 아데노신이고, T는 티민이고, G는 구아닌이고, C는 시토신이고, R은 퓨린이고, "/"는 절단 부위를 나타냄)일 수 있다. 이것은 본원에 기재된 보다 일반적인 스플라이스 공여자 컨센서스 서열 MAGGURR과 합치한다. 스플라이스 공여자가 이러한 컨센서스에서 벗어날 수 있다는 것이 당업계에 잘 알려져 있는데, 이는 특히 예를 들어 vRNA 패키징 영역 내에서 2차 구조와 같이 다른 제약 조건(constraint)이 동일한 서열에 적용되는 바이러스 게놈에서 그러하다. 비-표준 스플라이스 부위는 표준 스플라이스 공여자 컨센서스 서열과 비교하여 드물게 발생하기는 하지만 이들 또한 당업계에 잘 알려져 있다.
- [0070] "주요 스플라이스 공여자 부위"는 바이러스 벡터 뉴클레오티드 서열의 5' 영역에 전형적으로 위치하는 천연 바이러스 RNA 패키징 서열 내에 암호화되고 내장된, 바이러스 벡터 게놈의 첫 번째(우세한) 스플라이스 공여자 부위를 의미한다.
- [0071] 일 측면에서 뉴클레오티드 서열은 활성인 주요 스플라이스 공여자 부위를 함유하지 않으며, 즉 상기 뉴클레오티드 서열의 주요 스플라이스 공여자 부위로부터 스플라이싱이 발생하지 않고, 주요 스플라이스 공여자 부위로부터의 스플라이싱 활성이 제거된다.
- [0072] 주요 스플라이스 공여자 부위는 렌티바이러스 게놈의 5' 패키징 영역에 있다.
- [0073] HIV-1 바이러스의 경우, 주요 스플라이스 공여자 컨센서스 서열은 (DNA의 경우) TG/GTRAGT(여기서 A는 아데노신이고, T는 티민이고, G는 구아닌이고, C는 시토신이고, R은 퓨린이고, "/" 절단 부위를 나타냄)이다.
- [0074] 본 발명의 일 측면에서, 스플라이스 공여자 영역, 즉 돌연변이 이전의 주요 스플라이스 공여자 부위를 포함하는 벡터 게놈의 영역은 하기 서열을 가질 수 있다:
- [0075] GGGGCGGCGACTGGTGAGTACGCCAAAAAT (서열번호 1)
- [0076] 본 발명의 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0077] GGGGCGGCGACTGCAGACAACGCCAAAAAT (서열번호 2 - MSD-2K0)
- [0078] 본 발명의 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0079] GGGGCGGCGAGTGGAGACTACGCCAAAAAT (서열번호 11 - MSD-2K0v2)
- [0080] 본 발명의 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0081] GGGGAAGGCAACAGATAAATATGCCTTAAAAAT (서열번호 12 - MSD-2K0m5)
- [0082] 본 발명의 일 측면에서 변형 전의 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0083] GGCGACTGGTGAGTACGCC (서열번호 9)
- [0084] 이 서열은 또한 본원에서 "스텝 루프 2" 영역(SL2)으로 지칭된다. 이 서열은 벡터 게놈의 스플라이스 공여자 영역에서 스텝 루프 구조를 형성할 수 있다. 본 발명의 일 측면에서 이 서열(SL2)은 본원에 기재된 바와 같이 본 발명에 따른 뉴클레오티드 서열로부터 결실될 수 있다.
- [0085] 이와 같이, 본 발명은 SL2를 포함하지 않는 뉴클레오티드 서열을 포함한다. 본 발명은 서열번호 9에 따른 서열을 포함하지 않는 뉴클레오티드 서열을 포함한다.
- [0086] 본 발명의 일 측면에서 주요 스플라이스 공여자 부위는 하기 컨센서스 서열을 가질 수 있으며, 여기서 R은 퓨린이고 "/"는 절단 부위이다:
- [0087] TG/GTRAGT (서열번호 3)
- [0088] 일 측면에서, R은 구아닌(G)일 수 있다.

- [0089] 본 발명의 일 측면에서, 주요 스플라이스 공여자 및 잠복 스플라이스 공여자 영역은 하기 코어 서열을 가질 수 있으며, 여기서 "/"는 주요 스플라이스 공여자 및 잠복 스플라이스 공여자 부위에서의 절단 부위이다:
- [0090] /GTGA/GTA (서열번호 13).
- [0091] 본 발명의 일 측면에서 MSD-돌연변이된 벡터 계통은 주요 스플라이스 공여자 및 잠복 스플라이스 공여자 '영역'(서열번호 13)에서 적어도 2개의 돌연변이를 가질 수 있으며, 여기서 첫 번째 및 두 번째 'GT' 뉴클레오티드는 각각 주요 스플라이스 공여자 및 잠복 스플라이스 공여자 뉴클레오티드의 3' 바로이다.
- [0092] 본 발명의 일 측면에서 주요 스플라이스 공여자 컨센서스 서열은 CTGGT(서열번호 4)이다. 주요 스플라이스 공여자 부위는 CTGGT 서열을 포함할 수 있다.
- [0093] 일 측면에서, 스플라이스 부위의 불활성화 이전의, 뉴클레오티드 서열은 서열번호 1, 3, 4, 9, 10 및/또는 13 중 어느 하나에 기재된 서열을 포함한다.
- [0094] 일 측면에서 뉴클레오티드 서열은 불활성화되지 않았다면 서열번호 1의 뉴클레오티드 13 및 14에 상응하는 뉴클레오티드 사이에 절단 부위를 가지는 불활성화된 주요 스플라이스 공여자 부위를 포함한다.
- [0095] 본원에 기재된 바와 같은 본 발명에 따르면, 뉴클레오티드 서열은 불활성인 잠복 스플라이스 공여자 부위를 또한 함유한다. 일 측면에서 뉴클레오티드 서열은 주요 스플라이스 공여자 부위(의 3')에 인접한 활성인 잠복 스플라이스 공여자 부위를 함유하지 않으며, 이는 다시 말하면 인접한 잠복 스플라이스 공여자 부위로부터 스플라이싱이 발생하지 않고, 잠복 스플라이스 공여자 부위로부터의 스플라이싱이 제거된다.
- [0096] "잠복 스플라이스 공여자 부위"라는 용어는 인접한 서열 맥락(예를 들어, 가까운 '선회되는' 스플라이스 공여자의 존재)으로 인해 스플라이스 공여자 부위로서 정상적으로 기능하지 않거나 스플라이스 공여자 부위로서 덜 효율적으로 활용되지만, 인접한 서열의 돌연변이(예를 들어, 가까운 '선회되는' 스플라이스 공여자의 돌연변이)에 의해 더 효율적으로 기능하는 스플라이스 공여자 부위가 되도록 활성화될 수 있는 핵산 서열을 지칭한다.
- [0097] 일 측면에서 잠복 스플라이스 공여자 부위는 주요 스플라이스 공여자의 3'의 첫 번째 잠복 스플라이스 공여자 부위이다.
- [0098] 일 측면에서 잠복 스플라이스 공여자 부위는 주요 스플라이스 공여자 부위의 3' 측에서 주요 스플라이스 공여자 부위의 6개 뉴클레오티드 내에 있다. 바람직하게는 잠복 스플라이스 공여자 부위는 주요 스플라이스 공여자 절단 부위의 4개 또는 5개, 바람직하게는 4개 뉴클레오티드 내에 있다.
- [0099] 본 발명의 일 측면에서 잠복 스플라이스 공여자 부위는 컨센서스 서열 TGAGT(서열번호 10)를 가진다.
- [0100] 일 측면에서 뉴클레오티드 서열은 불활성화되지 않았다면 서열번호 1의 뉴클레오티드 17 및 18에 상응하는 뉴클레오티드 사이에 절단 부위를 가지는 불활성화된 잠복 스플라이스 공여자 부위를 포함한다.
- [0101] 본 발명의 일 측면에서 주요 스플라이스 공여자 부위 및/또는 인접한 잠복 스플라이스 공여자 부위는 "GT" 모티프를 함유한다. 본 발명의 일 측면에서 주요 스플라이스 공여자 부위 및 인접한 잠복 스플라이스 공여자 부위 양자 모두는 돌연변이된 "GT" 모티프를 함유한다. 돌연변이된 GT 모티프는 주요 스플라이스 공여자 부위와 인접한 잠복 스플라이스 공여자 부위 양자 모두로부터의 스플라이스 활성을 불활성시킬 수 있다. 이러한 돌연변이의 예는 본원에서 "MSD-2KO"로 지칭된다.
- [0102] 일 측면에서 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0103] CAGACA (서열번호 5)
- [0104] 예를 들어, 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0105] GGCGACTGCAGACAACGCC (서열번호 6)
- [0106] 불활성화 돌연변이의 추가의 예는 본원에서 "MSD-2KOv2"로 지칭된다.
- [0107] 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0108] GTGGAGACT (서열번호 7)
- [0109] 예를 들어, 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0110] GGCGAGTGGAGACTACGCC (서열번호 8)

- [0111] 예를 들어, 일 측면에서 돌연변이된 스플라이스 공여자 영역은 하기 서열을 포함할 수 있다:
- [0112] AAGGCAACAGATAAATATGCCTT (서열번호 14)
- [0113] 일 측면에서 상기 기재한 바와 같은 스템 루프 2 영역은 스플라이스 공여자 영역으로부터 결실될 수 있고, 그 결과 주요 스플라이스 공여자 부위와 인접한 잠복 스플라이스 공여자 부위 양자 모두가 불활성화된다. 이러한 결실은 본원에서 " Δ SL2"로 지칭된다.
- [0114] 주요 및 인접한 잠복 스플라이스 공여자 부위를 불활성화시키기 위해 다양한 서로 다른 유형의 돌연변이가 핵산 서열에 도입될 수 있다.
- [0115] 일 측면에서 돌연변이는 스플라이스 영역에서 스플라이싱 활성을 제거하거나 억제하는 기능적 돌연변이이다. 본원에 기재된 바와 같은 뉴클레오티드 서열은 서열번호 1, 3, 4, 9, 10 및/또는 13 중 어느 하나의 임의의 뉴클레오티드에서 돌연변이 또는 결실을 함유할 수 있다. 적합한 돌연변이는 당업자에게 공지되어 있을 것이며, 본원에 기재되어 있다.
- [0116] 예를 들어, 점 돌연변이가 핵산 서열에 도입될 수 있다. 본원에서 사용되는 용어 "점 돌연변이"는 단일 뉴클레오티드에 대한 임의의 변화를 지칭한다. 점 돌연변이에는, 예를 들어, 결실, 전이(transition) 및 전환(transversion)이 포함되며; 이들은 단백질 코딩 서열 내에 존재할 때 넨센스(nonsense) 돌연변이, 미스센스(missense) 돌연변이, 또는 침묵(silent) 돌연변이로 분류될 수 있다. "넨센스" 돌연변이는 정지 코돈을 생성한다. "미스센스" 돌연변이는 다른 아미노산을 암호화하는 코돈을 생성한다. "침묵" 돌연변이는 동일한 아미노산 또는 단백질의 기능을 변경하지 않는 다른 아미노산을 암호화하는 코돈을 생성한다. 하나 이상의 점 돌연변이가 잠복 스플라이스 공여자 부위를 포함하는 핵산 서열 내로 도입될 수 있다. 예를 들어, 잠복 스플라이스 부위를 포함하는 핵산 서열은 그 안에 2개 이상의 점 돌연변이를 도입함으로써 돌연변이될 수 있다.
- [0117] 적어도 2개의 점 돌연변이가 스플라이스 공여자 영역으로부터의 스플라이싱의 감쇠를 달성하기 위해 주요 스플라이스 공여자 및 잠복 스플라이스 공여자 부위를 포함하는 핵산 서열 내의 여러 위치에 도입될 수 있다. 일 측면에서 돌연변이는 스플라이스 공여자 절단 부위에서 4개의 뉴클레오티드 내에 있을 수 있고; 표준 스플라이스 공여자 컨센서스 서열에서 이것은 A^1G^2/G^3T^4 이며, 여기서 "/"가 절단 부위이다. 스플라이스 공여자 절단 부위가 이러한 컨센서스에서 벗어날 수 있다는 것이 당업계에 잘 알려져 있는데, 이는 특히 예를 들어 vRNA 패키징 영역 내의 2차 구조와 같이 다른 제약 조건(constraint)이 동일한 서열에 적용되는 바이러스 계놈에서 그러하다. G^3T^4 디뉴클레오티드는 일반적으로 표준 스플라이스 공여자 컨센서스 서열 내에서 가장 가변성이 적은 서열이며, G^3 및/또는 T^4 에 대한 돌연변이가 가장 큰 감쇠 효과를 달성할 가능성이 가장 높다는 것이 잘 알려져 있다. 예를 들어, HIV-1 바이러스 벡터 계놈의 주요 스플라이스 공여자 부위의 경우 이것은 T^1G^2/G^3T^4 일 수 있으며, 여기서 "/"가 절단 부위이다. 예를 들어, HIV-1 바이러스 벡터 계놈의 잠복 스플라이스 공여자 부위의 경우 이것은 G^1A^2/G^3T^4 일 수 있으며, 여기서 "/"가 절단 부위이다. 추가로, 점 돌연변이(들)는 스플라이스 공여자 부위에 인접하게 도입될 수 있다. 예를 들어, 점 돌연변이는 스플라이스 공여자 부위의 상류 또는 하류에 도입될 수 있다. 주요 및/또는 잠복 스플라이스 공여자 부위를 포함하는 핵산 서열이 그 안에 다중 점 돌연변이를 도입함으로써 돌연변이되는 구현예에서, 점 돌연변이는 잠복 스플라이스 공여자 부위의 상류 및/또는 하류에 도입될 수 있다.
- [0118] 본원에 기재된 바와 같이, 그리고 실시예에 나타난 바와 같이, 본 발명에 따른 렌티바이러스 벡터의 RNA 계놈을 암호화하는 뉴클레오티드 서열은 패키징 서열의 SL4 루프 내의 잠복 스플라이스 공여자 부위의 돌연변이를 선택적으로 추가로 포함할 수 있다. 일 측면에서 패키징 서열의 SL4 루프 내의 상기 잠복 스플라이스 공여자 부위의 GT 디뉴클레오티드는 GC로 돌연변이된다.
- [0119] **스플라이스 부위 돌연변이의 삭제**
- [0120] 본 발명의 스플라이스 부위 돌연변이는 다양한 기술을 사용하여 삭제될 수 있다. 예를 들어, 천연 서열의 단편에 결찰을 가능하게 하는 제한 부위가 플랭킹된 돌연변이 서열을 함유하는 올리고뉴클레오티드를 합성함으로써 돌연변이를 특정 유전자좌에 도입할 수 있다. 결찰 후, 생성된 재구성된 서열은 원하는 뉴클레오티드 삽입, 치환 또는 결실을 가지는 유도체를 포함한다.
- [0121] DNA 서열의 변경을 허용하는 다른 알려진 기술은 김슨(Gibson) 어셈블리, 골든-게이트(Golden-gate) 클로닝 및

인-퓨전과 같은 재조합 접근법을 포함한다.

- [0122] 대안적으로, 올리고뉴클레오타이드-지정 부위-특이적(또는 절편(segment) 특이적) 돌연변이유발 절차를 사용하여 요구되는 치환, 결실 또는 삽입에 따라 변경된 서열을 제공할 수 있다. 스플라이스 부위 돌연변이의 결실 또는 절단 유도체는 또한 원하는 결실에 인접한 편리한 제한 엔도뉴클레아제 부위를 이용하여 삭제될 수 있다.
- [0123] 제한(restriction) 후, 돌출부(overhang)가 채워지고, DNA가 다시 연결될 수 있다.
- [0124] 위에서 설명한 변경을 수행하는 예시적인 방법은 Sambrook et al.에 의해 공개된다(Molecular cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, 1989).
- [0125] 스플라이스 부위 돌연변이는 또한 PCR 돌연변이유발, 화학적 돌연변이유발, 강제 뉴클레오타이드 오혼입(misincorporation)(예: Liao and Wise, 1990)에 의한 화학적 돌연변이유발(Drinkwater and Klinesinst, 1986) 기술, 또는 무작위로 돌연변이된 올리고뉴클레오타이드(Horwitz et al., 1989)의 사용에 의해 삭제될 수 있다.
- [0126] 본 발명은 또한 하기 단계를 포함하는 렌티바이러스 벡터 뉴클레오타이드 서열을 생산하는 방법을 제공한다:
- [0127] (i) 본원에 기재된 바와 같은 렌티바이러스 벡터의 RNA 게놈을 암호화하는 뉴클레오타이드 서열을 제공하는 단계; 및
- [0128] (ii) 상기 뉴클레오타이드 서열에서 본원에 기재된 바와 같이 주요 스플라이스 공여자 부위 및 잠복 스플라이스 공여자 부위를 돌연변이시키는 단계.
- [0129] **벡터/발현 카세트**
- [0130] 벡터는 한 환경에서 다른 환경으로 엔티티(entity)를 전달할 수 있도록 하거나 용이하게 하는 도구이다. 본 발명에 따르면, 그리고 예로서, 재조합 핵산 기술에 사용되는 일부 벡터는 핵산의 절편(예: 이중 DNA 절편, 예컨대 이중 cDNA 절편)과 같은 엔티티가 표적 세포로 전달되어 표적 세포에 의해 발현되도록 한다. 벡터는 바이러스 벡터 구성요소를 암호화하는 뉴클레오타이드 서열의 통합을 촉진하여 표적 세포 내에서 바이러스 벡터 구성요소를 암호화하는 뉴클레오타이드 서열과 그의 발현을 유지할 수 있다.
- [0131] 벡터는 발현 카세트(발현 작제물(expression construct)로도 지칭됨)일 수 있거나 이를 포함할 수 있다. 본원에 기재된 발현 카세트는 전사될 수 있는 서열을 함유하는 핵산의 영역을 포함한다. 따라서, mRNA, tRNA 및 rRNA를 암호화하는 서열이 이 정의에 포함된다.
- [0132] 벡터는 하나 이상의 선택 가능한 마커 유전자(예: 네오마이신 내성 유전자) 및/또는 추적 가능한 마커 유전자(들)(예: 녹색 형광 단백질(GFP)를 암호화하는 유전자)를 포함할 수 있다. 벡터는 예를 들어 표적 세포를 감염 및/또는 형질도입하기 위해 사용될 수 있다. 벡터는 벡터가 해당 숙주 세포에서 복제할 수 있게 하는 뉴클레오타이드 서열을 추가로 포함할 수 있으며, 예컨대 조건부 복제 종양 용해 벡터(conditionally replicating oncolytic vector)이다.
- [0133] 용어 "카세트"는 "진주계이트", "작제물" 및 "하이브리드"와 같은 용어와 동의어로서, 프로모터에 직접 또는 간접적으로 부착된 폴리뉴클레오타이드 서열을 포함한다. 본 발명에서 사용하기 위한 발현 카세트는 바이러스 벡터 구성요소를 암호화하는 뉴클레오타이드 서열의 발현을 위한 프로모터 및 선택적으로 바이러스 벡터 구성요소를 암호화하는 뉴클레오타이드 서열의 조절자(regulator)를 포함한다. 바람직하게는 카세트는 프로모터에 작동가능하게 연결된 폴리뉴클레오타이드 서열을 적어도 포함한다.
- [0134] 발현 카세트, 예를 들어 플라스미드, 코스미드(cosmid), 바이러스 또는 파지 벡터의 선택은 종종 그것이 도입되는 숙주 세포에 의존할 것이다. 발현 카세트는 DNA 플라스미드(슈퍼코일화된(supercoiled), 닉(nicked) 또는 선형화된(linearised)), 미니서클 DNA(선형(linear) 또는 슈퍼코일화된), 제한 효소 분해 및 정제에 의해 플라스미드 백본을 제거하여 관심 영역만을 포함하는 플라스미드 DNA, 효소적 DNA 증폭 플랫폼, 예를 들어 doggybone DNA(dbDNA™)를 사용하여 생성된 DNA일 수 있으며, 여기서 사용되는 최종 DNA는 폐쇄된 결찰 형태(closed ligated form)이거나 개방 절단 말단(open cut end)을 가지도록 제조된(예: 제한 효소 소화) 것이다.
- [0135] **렌티바이러스 벡터 생산 시스템 및 세포**
- [0136] 렌티바이러스 벡터 생산 시스템은 렌티바이러스 벡터의 생산에 필요한 구성요소를 암호화하는 뉴클레오타이드 서열 세트를 포함한다. 따라서, 벡터 생산 시스템은 렌티바이러스 벡터 입자를 생성하는 데 필요한 바이러스 벡터 구성요소를 암호화하는 뉴클레오타이드 서열 세트를 포함한다.

- [0137] "바이러스 벡터 생산 시스템" 또는 "벡터 생산 시스템" 또는 "생산 시스템"은 렌티바이러스 벡터 생산을 위한 필요한 구성요소를 포함하는 시스템으로 이해되어야 한다.
- [0138] 일 구현예에서, 바이러스 벡터 생산 시스템은 Gag 및 Gag/Pol 단백질, 및 Env 단백질 및 벡터 게놈 서열을 암호화하는 뉴클레오타이드 서열을 포함한다. 생산 시스템은 선택적으로 Rev 단백질, 또는 이의 기능적 대체물을 암호화하는 뉴클레오타이드 서열을 포함할 수 있다.
- [0139] 일 구현예에서, 바이러스 벡터 생산 시스템은 모듈식(modular) 핵산 작제물(모듈식 작제물)을 포함한다. 모듈식 작제물은 렌티바이러스 벡터의 생산에 사용되는 2개 이상의 핵산을 포함하는 DNA 발현 작제물이다. 모듈식 작제물은 렌티바이러스 벡터의 생산에 사용되는 2개 이상의 핵산을 포함하는 DNA 플라스미드일 수 있다. 플라스미드는 박테리아 플라스미드일 수 있다. 핵산은 예를 들어 gag-pol, rev, env, 벡터 게놈을 암호화할 수 있다. 또한, 패키징 및 생산자 세포주의 생성을 위해 설계된 모듈식 작제물은 전사 조절 단백질(예: TetR, CymR) 및/또는 번역 억제 단백질(예: TRAP) 및 선택 가능한 마커(예: Zeocin™, 하이그로마이신(hygromycin), 블라스티시딘(blasticidin), 퓨로마이신(puromycin), 네오마이신 내성 유전자)를 추가로 암호화해야 할 수 있다. 본 발명에 사용하기에 적합한 모듈식 작제물은 그 전체가 참고로 본원에 포함된 EP 3502260에 기재되어 있다.
- [0140] 본 발명에 따라 사용하기 위한 모듈식 작제물은 하나의 작제물에 2개 이상의 레트로바이러스 구성요소를 암호화하는 핵산 서열을 포함하기 때문에, 이러한 모듈식 작제물의 안전성 프로파일이 고려되었고 추가적인 안전성 구성이 작제물에 직접 조작되었다. 이러한 구성에는 레트로바이러스 벡터 구성요소의 다중 오픈 리딩 프레임에 의한 인슐레이터(insulator)의 사용 및/또는 모듈식 작제물에서 레트로바이러스 유전자의 특정 방향 및 배열이 포함된다. 이러한 구성을 사용함으로써 복제 가능한(replication-competent) 바이러스 입자를 생성하는 직접적인 리드-스루가 방지될 것으로 믿어진다.
- [0141] 바이러스 벡터 구성요소를 암호화하는 핵산 서열은 모듈식 작제물에서 역방향 및/또는 교대 전사 배향일 수 있다. 따라서, 바이러스 벡터 구성요소를 암호화하는 핵산 서열이 동일한 5'에서 3' 방향으로 제시되지 않으므로, 바이러스 벡터 구성요소가 동일한 mRNA 분자로부터 생산될 수 없다. 역방향 배향은 서로 다른 벡터 구성요소에 대한 적어도 2개의 코딩 서열이 '머리-대-머리(head-to-head)' 및 '꼬리-대-꼬리(tail-to-tail)' 전사 방향으로 제시됨을 의미할 수 있다. 이것은 모듈식 작제물의 한 가닥에는 하나의 벡터 구성요소, 예를 들어 env에 대한 코딩 서열을 제공하고, 모듈식 작제물의 반대 가닥에는 다른 벡터 구성요소, 예를 들어 rev에 대한 코딩 서열을 제공함으로써 달성될 수 있다. 바람직하게는, 2개 이상의 벡터 구성요소에 대한 코딩 서열이 모듈식 작제물에 존재할 때, 코딩 서열 중 적어도 2개는 역 전사 배향으로 존재한다. 따라서, 2개 이상의 벡터 구성요소에 대한 코딩 서열이 모듈식 작제물에 존재할 때, 각 구성요소는 그것이 인접하고 있는 다른 벡터 구성요소에 대한 모든 인접 코딩 서열(들)에 대해 반대 5'에서 3' 배향으로 존재하도록 배향될 수 있으며, 즉 각 코딩 서열에 대해 5'에서 3'(또는 전사) 배향을 교대로 사용할 수 있다.
- [0142] 본 발명에 따라 사용하기 위한 모듈식 작제물은 하기 벡터 구성요소들 중 2개 이상을 암호화하는 핵산 서열을 포함할 수 있다: gag-pol, rev, env, 벡터 게놈. 모듈식 작제물은 벡터 구성요소의 임의의 조합을 암호화하는 핵산 서열을 포함할 수 있다. 일 구현예에서, 모듈식 작제물은 하기를 암호화하는 핵산 서열을 포함할 수 있다:
- [0143] i) 레트로바이러스 벡터의 RNA 게놈 및 rev, 또는 이의 기능적 대체물
- [0144] ii) 레트로바이러스 벡터의 RNA 게놈 및 gag-pol;
- [0145] iii) 레트로바이러스 벡터의 RNA 게놈 및 env;
- [0146] iv) gag-pol 및 rev, 또는 이의 기능적 대체물;
- [0147] v) gag-pol 및 env;
- [0148] vi) env 및 rev, 또는 이의 기능적 대체물;
- [0149] vii) 레트로바이러스 벡터의 RNA 게놈, rev, 또는 이의 기능적 대체물, 및 gag-pol;
- [0150] viii) 레트로바이러스 벡터의 RNA 게놈, rev, 또는 이의 기능적 대체물, 및 env;
- [0151] ix) 레트로바이러스 벡터의 RNA 게놈, gag-pol 및 env; 또는
- [0152] x) gag-pol, rev, 또는 이의 기능적 대체물, 및 env,
- [0153] 여기서 핵산 서열은 역방향 및/또는 교대 배향이다.

- [0154] 일 구현예에서, 레트로바이러스 벡터를 생산하기 위한 세포는 상기 조합 i) 내지 x) 중 어느 하나를 암호화하는 핵산 서열을 포함할 수 있으며, 여기서 핵산 서열은 동일한 유전적 유전자좌(genetic locus)에 위치하고 역방향 및/또는 교대 배향이다. 동일한 유전적 유전자좌는 세포 내 단일 염색체외(extrachromosomal) 유전자좌를 지칭할 수 있으며, 예를 들어 단일 플라스미드, 또는 세포의 게놈 내 단일 유전자좌(즉, 단일 삽입 부위)일 수 있다. 세포는 레트로바이러스 벡터, 예를 들어 렌티바이러스 벡터를 생산하기 위한 안정하거나 일시적인 세포일 수 있다. 일 측면에서 세포는 tat를 포함하지 않는다.
- [0155] DNA 발현 작제물은 DNA 플라스미드(슈퍼코일화된, Nick 또는 선형화된), 미니서클 DNA(선형 또는 슈퍼코일화된), 제한 효소 분해 및 정제에 의해 플라스미드 백본을 제거하여 관심 영역만을 포함하는 플라스미드 DNA, 효소적 DNA 증폭 플랫폼, 예를 들어 doggybone DNA(dbDNA™)를 사용하여 생성된 DNA일 수 있으며, 여기서 사용되는 최종 DNA는 폐쇄된 결찰 형태이거나 개방 절단 말단을 가지도록 제조된(예: 제한 효소 소화) 것이다.
- [0156] 일 구현예에서, 렌티바이러스 벡터는 HIV-1, HIV-2, SIV, FIV, BIV, EIAV, CAEV 또는 비스나 렌티바이러스로부터 유래된다.
- [0157] "바이러스 벡터 생산 세포", "벡터 생산 세포" 또는 "생산 세포"는 렌티바이러스 벡터 또는 렌티바이러스 벡터 입자를 생산할 수 있는 세포로 이해되어야 한다. 렌티바이러스 벡터 생산 세포는 "생산자 세포" 또는 "패키징 세포"일 수 있다. 바이러스 벡터 시스템의 하나 이상의 DNA 작제물은 바이러스 벡터 생산 세포 내에서 안정적으로 통합되거나 에피솜적으로(episomally) 유지될 수 있다. 대안적으로, 바이러스 벡터 시스템의 모든 DNA 구성 요소는 바이러스 벡터 생산 세포 내로 일시적으로 형질감염될 수 있다. 또 다른 대안에서, 구성요소 중 일부를 안정적으로 발현하는 생산 세포를 벡터 생산에 필요한 나머지 구성요소로 일시적으로 형질감염시킬 수 있다.
- [0158] 본원에서 사용되는 바와 같이, 용어 "패키징 세포"는 렌티바이러스 벡터 입자의 생산에 필요한 요소를 포함하지만 벡터 게놈이 결여된 세포를 지칭한다. 선택적으로, 이러한 패키징 세포는 바이러스 구조 단백질(예컨대 *gag*, *gag/pol* 및 *env*)을 발현할 수 있는 하나 이상의 발현 카세트를 함유한다.
- [0159] 생산자 세포/패키징 세포는 임의의 적합한 세포 유형일 수 있다. 생산자 세포는 일반적으로 포유동물 세포이지만, 예를 들어 곤충 세포일 수 있다.
- [0160] 본원에서 사용되는 바와 같이, 용어 "생산자/생산 세포" 또는 "벡터 생산/생산자 세포"는 렌티바이러스 벡터 입자의 생산에 필요한 모든 요소를 함유하는 세포를 지칭한다. 생산자 세포는 안정한 생산자 세포주이거나 일시적으로 유래될 수 있거나 또는 레트로바이러스 게놈이 일시적으로 발현되는 안정한 패키징 세포일 수 있다.
- [0161] 본 발명의 방법에서, 벡터 구성요소는 바이러스 벡터가 렌티바이러스 벡터인 경우 *gag*, *env*, *rev* 및/또는 렌티바이러스 벡터의 RNA 게놈을 포함할 수 있다. 벡터 구성요소를 암호화하는 뉴클레오티드 서열은 동시에 또는 임의의 순서로 순차적으로 세포 내로 도입될 수 있다.
- [0162] 벡터 생산 세포는 조직 배양 세포주와 같이 시험관 내에서 배양된 세포일 수 있다. 본 발명의 방법 및 용도의 일부 구현예에서, 렌티바이러스 벡터를 생산하기 위한 적절한 생산 세포 또는 세포는 적절한 조건 하에서 배양될 때 바이러스 벡터 또는 바이러스 벡터 입자를 생산할 수 있는 세포이다. 따라서, 세포는 전형적으로 벡터 구성요소를 암호화하는 뉴클레오티드 서열을 포함하며, 이는 *gag*, *env*, *rev* 및 렌티바이러스 벡터의 RNA 게놈을 포함할 수 있다. 적합한 세포주는 무린 섬유아세포 유래 세포주 또는 인간 세포주와 같은 포유동물 세포를 포함하지만 이에 제한되지는 않는다. 이들은 일반적으로 인간 세포를 포함하는 포유동물, 예를 들어 HEK293T, HEK293, CAP, CAP-T 또는 CHO 세포이지만, 예를 들어 SF9 세포와 같은 곤충 세포일 수 있다. 바람직하게는, 벡터 생산 세포는 인간 세포주로부터 유래된다. 따라서, 이러한 적합한 생산 세포는 본 발명의 임의의 방법 또는 용도에 사용될 수 있다.
- [0163] 뉴클레오티드 서열을 세포 내로 도입하는 방법은 당업계에 잘 알려져 있고 이전에 기술되었다. 따라서, 분자 및 세포 생물학의 통상적인 기술을 사용하여, *gag*, *env*, *rev* 및 렌티바이러스 벡터의 RNA 게놈을 포함하는 벡터 구성요소를 암호화하는 뉴클레오티드 서열을 세포 내로 도입하는 것은 당업자의 능력 범위 내에 있다.
- [0164] 안정적인 생산 세포는 패키징 세포 또는 생산자 세포일 수 있다. 패키징 세포로부터 생산자 세포를 생성하기 위해 벡터 게놈 DNA 작제물이 안정적으로 또는 일시적으로 도입될 수 있다. 패키징/생산자 세포는 벡터의 구성요소, 즉 WO 2004/022761에 기술된 바와 같이 게놈, *gag-pol* 구성요소 및 외피(엔벨로프) 중 하나를 발현하는 레트로바이러스 벡터를 적합한 세포주에 형질도입함으로써 생성될 수 있다.
- [0165] 대안적으로, 뉴클레오티드 서열은 세포 내로 형질감염될 수 있고, 이어서 생산 세포 게놈으로의 통합이 드물게

무작위로 일어난다. 형질감염 방법은 당업계에 잘 알려진 방법을 사용하여 수행할 수 있다. 예를 들어, 안정적인 형질감염 과정은 연쇄동일화(concatemerisation)를 돕기 위해 조작된 작제물을 사용할 수 있다. 다른 예에서, 형질감염 과정은 인산칼슘 또는 상업적으로 입수가 가능한 제형, 예컨대 Lipofectamine™ 2000CD(Invitrogen, CA), FuGENE® HD 또는 폴리에틸렌이민(PEI)을 사용하여 수행될 수 있다. 대안적으로, 뉴클레오티드 서열은 전기천공(electroporation)을 통해 생산 세포 내로 도입될 수 있다. 당업자는 생산 세포 내로 뉴클레오티드 서열의 통합을 촉진하는 방법을 알고 있을 것이다. 예를 들어, 핵산 작제물이 원래 환형인 경우 이를 선형화하면 도움이 될 수 있다. 덜 무작위적인 통합 방법론은 내인성 게놈 내의 선택된 부위로의 통합을 가이드하기 위해 포유동물 숙주 세포의 내인성 염색체와 공유되는 상동성 영역을 포함하는 핵산 작제물을 수반할 수 있다. 또한, 재조합 부위가 작제물에 존재하는 경우 이들은 표적화된 재조합에 사용될 수 있다. 예를 들어, 핵산 작제물은 *loxP* 부위를 포함할 수 있으며 이는 Cre 재조합효소(recombinase)와 조합될 때 표적화된 통합을 허용한다(즉, P1 박테리오파지에서 유래된 Cre/*lox* 시스템을 사용). 대안적으로 또는 추가적으로, 재조합 부위는 *att* 부위(예를 들어, λ 파지로부터)이고, 여기서 *att* 부위는 람다 인테그라제(lambda integrase)의 존재 하에 부위-지정된 통합을 허용한다. 이것은 렌티바이러스 유전자가 높은 및/또는 안정적인 발현을 허용하는 숙주 세포 게놈 내의 유전자좌를 표적으로 할 수 있게 한다.

- [0166] 표적화된 통합의 다른 방법이 당업계에 잘 알려져 있다. 예를 들어, 게놈 DNA의 표적화된 절단을 유도하는 방법을 사용하여 선택된 염색체 유전자좌에서 표적화된 재조합을 촉진할 수 있다. 이러한 방법은 종종 이중 가닥 절단(double strand break, DSB)을 유도하는 방법 또는 시스템의 사용을 수반하는데, 예를 들어 비-상동 말단 연결(non-homologous end joining, NHEJ)과 같은 생리학적 메커니즘에 의해 절단의 복구를 유도하기 위한 내인성 게놈의 Nick(nick)이다. 특정의 뉴클레아제, 예컨대 조작된 아연 핑거 뉴클레아제(ZFN), 전사-활성화제 유사 이펙터 뉴클레아제(TALEN), 특정 절단을 가이드하기 위해 조작된 crRNA/tracr RNA('단일 가이드 RNA')와 함께 CRISPR/Cas9 시스템의 사용, 및/또는 Argonaute 시스템(예: *T. thermophilus*로부터)에 기초한 뉴클레아제 사용을 통해 절단이 발생할 수 있다.
- [0167] 패키징/생산자 세포주는 단지 렌티바이러스 형질도입 또는 단지 핵산 형질감염의 방법을 사용하여 뉴클레오티드 서열의 통합에 의해 생성될 수 있거나, 이 둘의 조합이 사용될 수 있다.
- [0168] 생산 세포로부터 레트로바이러스 벡터를 생성하는 방법 및 특히 레트로바이러스 벡터의 처리는 WO 2009/153563에 기재되어 있다.
- [0169] 일 구현예에서, 생산 세포는 RNA-결합 단백질(예를 들어, 트립토판 RNA-결합 감쇠 단백질, TRAP) 및/또는 Tet Repressor(TetR) 단백질 또는 대안적인 조절 단백질(예를 들어, CymR)을 포함할 수 있다.
- [0170] 생산 세포로부터 렌티바이러스 벡터의 생산은 형질감염 방법을 통해, 유도 단계(예: 독시사이클린 유도)를 포함할 수 있는 안정한 세포주로부터의 생산으로부터, 또는 이 둘의 조합을 통해 이루어질 수 있다. 형질감염 방법은 당업계에 잘 알려진 방법을 사용하여 수행할 수 있으며, 그 예는 이전에 설명되었다.
- [0171] 패키징 또는 생산자 세포주 또는 렌티바이러스 벡터를 암호화하는 구성요소로 일시적으로 형질감염된 생산 세포는 세포 및 바이러스 수 및/또는 바이러스 역가를 증가시키기 위해 배양된다. 세포 배양은 세포가 본 발명에 따른 관심 바이러스 벡터를 대사, 및/또는 성장 및/또는 분열 및/또는 생산할 수 있게 하기 위해 수행된다. 이것은 당업자에게 잘 알려진 방법에 의해 달성될 수 있고, 예를 들어 적절한 배양 배지에서 세포에 영양분을 제공하는 것을 포함하지만 이에 제한되지는 않는다. 상기 방법은 표면에 부착된 성장, 현탁액에서의 성장, 또는 이들의 조합을 포함할 수 있다. 배양은, 예를 들어 배치식(batch), 유가식(fed-batch), 연속식 시스템 등을 사용하여, 조직 배양 플라스크, 조직 배양 멀티웰 플레이트, 접시(dish), 롤러 병(roller bottle), 웨이브 백(wave bag) 또는 생물반응기에서 수행될 수 있다. 세포 배양을 통해 바이러스 벡터의 대규모 생산을 달성하기 위해, 현탁액에서 성장할 수 있는 세포를 가지는 것이 당업계에서 바람직하다. 세포 배양에 적합한 조건이 알려져 있다(예를 들어, Tissue Culture, Academic Press, Kruse and Paterson, editors (1973), 및 R.I. Freshney, Culture of animal cells: A manual of basic technique, fourth edition (Wiley-Liss Inc., 2000, ISBN 0-471-34889-9 참조).
- [0172] 바람직하게는 세포는 조직 배양 플라스크 또는 생물반응기에서 초기에 '벌크 업'되고, 이후에 다층 배양 용기 또는 대형 생물반응기(50 L 초과)에서 성장하여 본 발명에서 사용하기 위한 벡터 생산 세포를 생성한다.
- [0173] 바람직하게는 세포를 현탁 방식으로 성장시켜 본 발명에서 사용하기 위한 벡터 생산 세포를 생성한다.

[0174] **렌티바이러스 벡터**

[0175] 렌티바이러스는 더 큰 레트로바이러스 그룹의 일부이다. 렌티바이러스의 상세한 목록은 Coffin et al (1997) "Retroviruses" Cold Spring Harbour Laboratory Press Eds: JM Coffin, SM Hughes, HE Varmus pp 758-763 에서 찾을 수 있다. 간단히 말해서, 렌티바이러스는 영장류 그룹과 비영장류 그룹으로 나눌 수 있다. 영장류 렌티바이러스의 예에는 인간 자가면역결핍 증후군(AIDS)의 원인 인자인 인간 면역결핍 바이러스(HIV), 및 원숭이 면역결핍 바이러스(SIV)가 포함되지만 이에 국한되지는 않는다. 비영장류 렌티바이러스 그룹에는 프로토타입 "느린 바이러스(slow virus)" 비스나/메디 바이러스(visna/maedi virus)(VMV), 뿐만 아니라 관련 염소 관절염-뇌염 바이러스(CAEV), 말 감염성 빈혈 바이러스(EIAV), 고양이 면역결핍 바이러스(FIV), 메디 비스나(Maedi visna) 바이러스(MVV) 및 소 면역결핍 바이러스(BIV)가 포함된다. 일 구현예에서, 렌티바이러스 벡터는 HIV-1, HIV-2, SIV, FIV, BIV, EIAV, CAEV 또는 비스나 렌티바이러스로부터 유래된다.

[0176] 렌티바이러스 패밀리는 렌티바이러스가 분열 세포와 비분열 세포 모두를 감염시키는 능력이 있다는 점에서 레트로바이러스와 다르다(Lewis et al (1992) EMBO J 11(8):3053-3058 및 Lewis and Emerman (1994) J Virol 68 (1):510-516). 대조적으로, MLV와 같은 다른 레트로바이러스는 예를 들어 근육, 뇌, 폐 및 간 조직을 구성하는 것과 같은 비분열 또는 천천히 분열하는 세포를 감염시킬 수 없다.

[0177] 본원에서 사용되는 렌티바이러스 벡터는 렌티바이러스로부터 유래할 수 있는 적어도 하나의 구성요소 일부를 포함하는 벡터이다. 바람직하게는, 그 구성요소 일부는 벡터가 표적 세포를 감염시키거나 형질도입하고 NOI를 발현하는 생물학적 기전에 관여한다.

[0178] 렌티바이러스 벡터는 시험관 내에서 양립가능한(compatible) 표적 세포에서 NOI를 복제하는 데 사용될 수 있다. 따라서, 본원은 시험관 내에서 양립가능한 표적 세포 내로 본 발명의 벡터를 도입하고 NOI의 발현을 초래하는 조건하에서 표적 세포를 성장시킴으로써 시험관 내에서 단백질을 제조하는 방법을 기술한다. 단백질 및 NOI는 당업계에서 잘 알려진 방법에 의해 표적 세포로부터 회수될 수 있다. 적합한 표적 세포는 포유동물 세포주 및 기타 진핵 세포주를 포함한다.

[0179] 일부 측면에서 벡터는 프로모터와 인핸서 사이의 상호작용을 차단하고 인접 유전자로부터의 리드-스루를 감소시키는 장벽으로 작용하는 유전자 서열인 "인슐레이터"를 가질 수 있다.

[0180] 일 구현예에서 인슐레이터는 프로모터 간섭 및 인접 유전자로부터의 리드-스루를 방지하기 위해 하나 이상의 렌티바이러스 핵산 서열 사이에 존재한다. 인슐레이터가 하나 이상의 렌티바이러스 핵산 서열 사이의 벡터에 존재하는 경우, 이들 인슐레이터된 유전자 각각은 개별 발현 단위로서 배열될 수 있다.

[0181] 레트로바이러스 및 렌티바이러스 게놈의 기본 구조는 5' LTR 및 3' LTR과 같은 많은 공통적인 구성을 공유하며, 이들 사이 또는 이들 내에 게놈이 패키징되게 하는 패키징 신호, 프라이머 결합 부위, 표적 세포 게놈으로의 통합을 가능하게 하는 통합 부위 및 패키징 구성요소를 암호화하는 *gag/pol* 및 *env* 유전자가 위치하며, 이들은 바이러스 입자의 어셈블리에 필요한 폴리펩티드다. 렌티바이러스는 HIV의 *rev* 유전자 및 RRE 서열과 같은 추가 구성을 가지고 있고, 이는 통합된 프로바이러스의 RNA 전사체를 핵으로부터 감염된 표적 세포의 세포질로 효율적으로 내보낼 수 있다.

[0182] 프로바이러스에서, 이들 유전자는 양쪽 끝에서 긴 말단 반복부(long terminal repeat, LTR)라고 하는 영역에 의해 플랭크된다. LTR은 프로바이러스 통합 및 전사를 담당한다. LTR은 또한 인핸서-프로모터 서열의 역할을 하며 바이러스 유전자의 발현을 제어할 수 있다.

[0183] LTR 자체는 U3, R 및 U5라고 하는 세 가지 요소로 나눌 수 있는 동일한 서열이다. U3는 RNA의 3' 말단에 대해 고유한 서열로부터 유래한다. R은 RNA의 양 말단에서 반복되는 서열로부터 유래하고 U5는 RNA의 5' 말단에 대해 고유한 서열로부터 유래한다. 세 가지 요소의 크기는 서로 다른 레트로바이러스에 따라 상당히 다를 수 있다.

[0184] 본원에 기재된 바와 같은 전형적인 렌티바이러스 벡터에서, 복제에 필수적인 하나 이상의 단백질 코딩 영역의 적어도 일부가 바이러스로부터 제거될 수 있고; 예를 들어, *gag/pol* 및 *env*가 없거나 작동하지 않을 수 있다. 이것은 바이러스 벡터를 복제-결합으로 만든다.

[0185] 렌티바이러스 벡터는 영장류 렌티바이러스(예: HIV-1) 또는 비영장류 렌티바이러스(예: EIAV)에서 유래할 수 있다.

[0186] 일반적으로, 전형적인 레트로바이러스 벡터 생산 시스템은 필수적인 바이러스 패키징 기능으로부터 바이러스 게놈의 분리를 수반한다. 이러한 바이러스 벡터 구성요소는 일반적으로 별도의 DNA 발현 카세트(다르게는 플라스

미드, 발현 플라스미드, DNA 작제물 또는 발현 작제물로 알려짐) 상에서 생산 세포에 제공된다.

- [0187] 벡터 계놈은 NOI를 포함한다. 벡터 계놈은 전형적으로 패키징 신호(ψ), NOI를 포함하는 내부 발현 카세트, (선택적으로) 전사 후 요소(PRE), 전형적으로 중앙 폴리퓨린 트랙(cppt), 3'-ppu 및 자가-불활성화(SIN) LTR을 필요로 한다. R-U5 영역은 벡터 계놈 RNA와 NOI mRNA 양자 모두의 올바른 폴리아데닐화뿐만 아니라 역전사 과정에 필요하다. 벡터 계놈은, WO 2003/064665에 기술된 바와 같이, rev 부재 시 벡터 생성을 허용하는 오픈 리딩 프레임 선택적으로 포함할 수 있다.
- [0188] 패키징 기능에는 *gag/pol* 및 *env* 유전자가 포함된다. 이들은 생산 세포에 의한 벡터 입자 생산에 필요하다. 계놈에 트랜스로 이러한 기능을 제공하는 것은 복제-결합 바이러스 벡터의 생산을 용이하게 한다.
- [0189] 감마-레트로바이러스 벡터의 생산 시스템은 전형적으로 계놈, *gag/pol* 및 *env* 발현 작제물을 필요로 하는 3성분 시스템이다. HIV-1 기반 렌티바이러스 벡터의 생산 시스템은 추가로 액세서리 유전자 *rev*가 제공되고 벡터 계놈에 *rev*-반응 요소(RRE)를 포함하도록 요구할 수 있다. EIAV 기반 렌티바이러스 벡터는 오픈 리딩 프레임(ORF)이 계놈 내에 존재하는 경우 트랜스로 *rev*를 제공할 필요가 없다(WO 2003/064665 참조).
- [0190] 일반적으로 벡터 계놈 카세트 내에 암호화된 "외부" 프로모터(벡터 계놈 카세트를 구동함) 및 "내부" 프로모터(NOI 카세트를 구동함) 양자 모두는 다른 벡터 시스템 구성요소를 구동하는 프로모터와 마찬가지로 강력한 진핵 생물 또는 바이러스 프로모터이다. 이러한 프로모터의 예는 CMV, EF1 α , PGK, CAG, TK, SV40 및 유비퀴틴 프로모터를 포함한다. DNA 라이브리리에 의해 생성된 것(예: JeT 프로모터)과 같은 강력한 '합성' 프로모터도 전사를 유도하는 데 사용할 수 있다. 대안적으로, 로돕신(Rho)과 같은 조직-특이적 프로모터, 로돕신 키나제(RhoK), 원추형(cone-rod) 호메오박스 함유 유전자(CRX), 신경 망막-특이적 류신 지퍼 단백질(NRL), 난황 황반 이영양증 2(Vitelliform Macular Dystrophy 2, VMD2), 티로신 하이드록실라제, 뉴런-특이적 뉴런-특이적 에놀라제(NSE) 프로모터, 성상교세포-특이적 신경교 섬유성 산성 단백질(astrocyte-specific glial fibrillary acidic protein, GFAP) 프로모터, 인간 α 1-항트립신(hAAT) 프로모터, 포스포에놀피루베이트 카르복시키나제(phosphoenolpyruvate carboxykinase, PEPCK), 간 지방산 결합 단백질 프로모터, Flt-1 프로모터, INF- β 프로모터, Mb 프로모터, SP-B 프로모터, SYN1 프로모터, WASP 프로모터, SV40 / hA1b 프로모터, SV40 / CD43, SV40 / CD45, NSE / RU5' 프로모터, ICAM-2 프로모터, GPIIb 프로모터, GFAP 프로모터, 피브로넥틴 프로모터, 엔도글린 프로모터, 엘라스타제-1 프로모터, 데스민(Desmin) 프로모터, CD68 프로모터, CD14 프로모터 및 B29 프로모터를 사용하여 전사를 구동할 수 있다.
- [0191] 레트로바이러스 벡터의 생산은 생산 세포를 이러한 DNA 구성요소로 일시적으로 동시-형질감염시키거나 모든 구성요소가 생산 세포 계놈 내에 안정적으로 통합된 안정적인 생산 세포주의 사용을 포함한다(예: Stewart HJ, Fong-Wong L, Strickland I, Chipchase D, Kelleher M, Stevenson L, Thoree V, McCarthy J, Ralph GS, Mitrophanous KA and Radcliffe PA. (2011). *Hum Gene Ther.* Mar; **22** (3):357-69). 대안적인 접근법은 안정적인 패키징 세포(이 안에 패키징 구성요소가 안정적으로 통합됨)를 사용한 다음 필요에 따라 벡터 계놈 플라스미드에 일시적으로 형질감염시키는 것이다(예: Stewart, H. J., M. A. Leroux-Carlucci, C. J. Sion, K. A. Mitrophanous and P. A. Radcliffe (2009). *Gene Ther.* Jun; **16** (6):805-14). 또한 완전하지 않은 대안적인 패키징 세포주가 생성될 수 있고(단 하나 또는 두 개의 패키징 구성요소가 세포주에 안정적으로 통합됨) 벡터를 생성하기 위해 누락된 구성요소를 일시적으로 형질감염시키는 것도 가능하다. 생산 세포는 또한 전사 조절자의 tet 억제인자(TetR) 단백질 그룹의 구성원(예: T-Rex, Tet-On 및 Tet-Off), 전사 조절자의 큐메이트(cumate) 유도성 스위치 시스템 그룹의 구성원(예: 큐메이트 억제인자(CymR) 단백질), 또는 RNA-결합 단백질(예: TRAP - 트립토판-활성화된 RNA-결합 단백질)과 같은 조절 단백질을 발현할 수 있다.
- [0192] 본 발명의 일 구현예에서, 바이러스 벡터는 EIAV로부터 유래된다. EIAV는 렌티바이러스 중 가장 단순한 계놈 구조를 가지며 본 발명에서 사용하기에 특히 바람직하다. *gag/pol* 및 *env* 유전자 외에도, EIAV는 *tat*, *rev* 및 *S2*의 세 가지 다른 유전자를 암호화한다. *Tat*는 바이러스 LTR의 전사 활성제로 작용하고(Derse and Newbold (1993) *Virology* 194(2):530-536 및 Maury et al (1994) *Virology* 200(2):632-642), *rev*는 *rev*-반응 요소(RRE)를 통해 바이러스 유전자의 발현을 조절하고 조정한다(Martarano et al. (1994) *J Virol* 68(5):3102-3111). 이들 두 단백질의 작용 기전은 영장류 바이러스의 유사한 기전과 대체로 유사한 것으로 생각된다(Martarano et al. (1994) *J Virol* 68(5):3102-3111). *S2*의 기능은 알려져 있지 않다. 또한, EIAV 단백질인 Ttm이 확인되었는데, 이는 막관통 단백질의 시작부분의 *env* 코딩 서열에 스플라이싱되는 *tat*의 첫 번째 엑손에 의해 암호화된다. 본 발명의 대안적인 구현예에서 바이러스 벡터는 HIV로부터 유래된다: HIV는 *S2*를 암호화하지 않는다는 점에서 EIAV와 상이하지만 EIAV와 달리 *vif*, *vpr*, *vpu* 및 *nef*를 암호화한다.

- [0193] 용어 "재조합 레트로바이러스 또는 렌티바이러스 벡터"(RRV)는 패키징 구성요소의 존재 하에 표적 세포를 형질 도입할 수 있는 바이러스 입자로의 RNA 게놈의 패키징을 허용하기에 충분한 레트로바이러스 유전 정보를 가지는 벡터를 지칭한다. 표적 세포의 형질도입은 역전사 및 표적 세포 게놈으로의 통합을 포함할 수 있다. RRV는 벡터에 의해 표적 세포로 전달될 비-바이러스성 코딩 서열을 운반한다. RRV는 표적 세포 내에서 감염성 레트로바이러스 입자를 생성하기 위한 독립적인 복제가 불가능하다. 일반적으로 RRV에는 기능적 *gag/pol* 및/또는 *env* 유전자, 및/또는 복제에 필수적인 기타 유전자가 없다.
- [0194] 바람직하게는 본 발명의 RRV 벡터는 최소 바이러스 게놈을 가진다.
- [0195] 본원에서 사용되는 용어 "최소 바이러스 게놈"은 바이러스 벡터가 감염, 형질도입 및 NOI를 표적 세포에 전달하는 데 필요한 기능을 제공하는 데 필수적인 요소를 유지하면서 비필수 요소를 제거하도록 조작되었음을 의미한다. 이 전략에 대한 추가의 자세한 내용은 WO 1998/17815 및 WO 99/32646에서 찾을 수 있다. 최소 EIAV 벡터는 *tat*, *rev* 및 *S2* 유전자를 결여하며, 이러한 유전자가 생산 시스템에서 트랜스로 제공되지도 않는다. 최소 HIV 벡터는 *vif*, *vpr*, *vpu*, *tat* 및 *nef*를 결여한다.
- [0196] 생산 세포 내에서 벡터 게놈을 생산하기 위해 사용되는 발현 플라스미드는 생산 세포/패키징 세포에서 게놈의 전사를 지시하기 위해 레트로바이러스 게놈에 작동가능하게 연결된 전사 조절 제어 서열을 포함할 수 있다. 모든 3세대 렌티바이러스 벡터는 5' U3 인핸서-프로모터 영역에서 결실되고, 벡터 게놈 RNA의 전사는 하기에서 논의되는 바와 같이 다른 바이러스 프로모터, 예를 들어 CMV 프로모터와 같은 이종 프로모터에 의해 구동된다. 이 구성은 *tat*와는 독립적으로 벡터 생성을 가능하게 한다. 일부 렌티바이러스 벡터 게놈은 효율적인 바이러스 생산을 위해 추가 서열이 필요하다. 예를 들어, 특히 HIV의 경우, RRE 서열이 포함될 수 있다. 그러나 (별도의) GagPol 카세트 상에서 RRE에 대한 요구(및 트랜스로 제공되는 *rev*에 대한 의존성)는 GagPol ORF의 코돈 최적화에 의해 감소되거나 제거될 수 있다. 이 전략에 대한 추가의 자세한 내용은 WO 2001/79518에서 찾을 수 있다.
- [0197] *rev*/RRE 시스템과 동일한 기능을 수행하는 대체 서열도 알려져 있다. 예를 들어, *rev*/RRE 시스템의 기능적 유사체는 Mason Pfizer 원숭이 바이러스에서 발견된다. 이것은 구성적 수송 요소(constitutive transport element, CTE)로 알려져 있으며 감염된 세포의 인자와 상호작용하는 것으로 여겨지는 RRE-유형 서열을 게놈에 포함한다. 세포성 인자는 *rev* 유사체로 생각할 수 있다. 따라서, CTE는 *rev*/RRE 시스템의 대안으로 사용될 수 있다. 알려져 있거나 이용가능하게 된 Rev 단백질의 다른 기능적 등가물은 본 발명과 관련될 수 있다. 예를 들어, HTLV-1의 Rex 단백질이 HIV-1의 Rev 단백질을 기능적으로 대체할 수 있다는 것도 알려져 있다. *Rev* 및 RRE는 본 발명의 방법에 사용하기 위한 벡터에서 부재하거나 비기능적일 수 있으며; 대체 *rev* 및 RRE, 또는 기능적으로 동등한 시스템이 존재할 수 있다.
- [0198] 본원에서 사용되는 용어 "기능적 대체물"은 다른 단백질 또는 서열과 동일한 기능을 수행하는 대체 서열을 가지는 단백질 또는 서열을 의미한다. "기능적 대체물"이라는 용어는 동일한 의미로 본원에서 "기능적 등가물" 및 "기능적 유사체"와 상호교환적으로 사용된다.
- [0199] **SIN 벡터**
- [0200] 본원에 기재된 바와 같은 렌티바이러스 벡터는 바이러스 인핸서 및 프로모터 서열이 결실된 자가-불활성화(SIN) 구성으로 사용될 수 있다. SIN 벡터는 비-SIN 벡터와 유사한 효능으로 생성되고 생체내, 생체외 또는 시험관내에서 비분열 표적 세포를 형질도입할 수 있다. SIN 프로바이러스에서 긴 말단 반복부(LTR)의 전사 불활성화는 vRNA의 이동을 방지해야 하며, 이는 복제-가능 바이러스의 형성 가능성을 더욱 감소시키는 구성이다. 이것은 또한 LTR의시스-작용 효과를 제거함으로써 내부 프로모터로부터 유전자의 조절된 발현을 가능하게 해야 한다.
- [0201] 예를 들어, 자가-불활성화 레트로바이러스 벡터 시스템은 3' LTR의 U3 영역에서 전사 인핸서 또는 인핸서 및 프로모터를 결실시킴으로써 작제되었다. 벡터 역전사 및 통합 라운드 후, 이러한 변화는 5' 및 3' LTR 양자 모두 내로 복사되며 전사적으로 불활성인 프로바이러스를 생성한다. 그러나, 이러한 벡터에서 LTR 내부의 모든 프로모터(들)은 여전히 전사적으로 활성일 것이다. 이 전략은 내부에 배치된 유전자의 전사에 대한 바이러스 LTR의 인핸서 및 프로모터의 영향을 제거하는 데 사용되었다. 이러한 효과에는 전사 증가 또는 전사 억제제가 포함된다. 이 전략은 3' LTR에서 게놈 DNA로의 다운스트림 전사를 제거하는 데에도 사용할 수 있다. 이것은 내인성 종양유전자(oncogene)의 우발적 활성화를 방지하는 것이 중요한 인간 유전자 치료에서 특히 중요하다. Yu *et al.*, (1986) *PNAS* 83: 3194-98; Marty *et al.*, (1990) *Biochimie* 72: 885-7; Naviaux *et al.*, (1996) *J. Virol.* 70: 5701-5; Iwakuma *et al.*, (1999) *Virology* 261: 120-32; Deglon *et al.*, (2000) *Human Gene Therapy* 11: 179-90. SIN 렌티바이러스 벡터는 US 6,924,123 및 US 7,056,699에 기재되어 있다.

[0202] **복제-결합 렌티바이러스 벡터**

[0203] 복제-결합 렌티바이러스 벡터의 게놈에서 *gag/pol* 및/또는 *env*의 서열은 돌연변이되고/되거나 기능하지 않을 수 있다.

[0204] 본원에 기재된 바와 같은 전형적인 렌티바이러스 벡터에서, 바이러스 복제에 필수적인 단백질에 대한 하나 이상의 코딩 영역의 적어도 일부는 벡터로부터 제거될 수 있다. 이것은 바이러스 벡터를 복제-결합으로 만든다. 비-분할 표적 세포를 형질도입하고/하거나 표적 세포 게놈 내로 그의 게놈을 통합할 수 있는 NOI를 포함하는 벡터를 생성하기 위해 바이러스 게놈의 일부는 또한 NOI로 치환될 수 있다.

[0205] 일 구현예에서 렌티바이러스 벡터는 WO 2006/010834 및 WO 2007/071994에 기재된 바와 같은 비-통합 벡터이다.

[0206] 추가의 구현예에서 벡터는 바이러스 RNA가 없거나 결여된 서열을 전달하는 능력을 갖는다. 추가의 구현예에서, 전달될 RNA 상에 위치한 이중 결합 도메인(*gag*에 대해 이중) 및 Gag 또는 GagPol 상의 동족(cognate) 결합 도메인이 전달될 RNA의 패키징을 보장하기 위해 사용될 수 있다. 이들 벡터 모두는 WO 2007/072056에 기재되어 있다.

[0207] **NOI 및 폴리뉴클레오티드**

[0208] 본 발명의 폴리뉴클레오티드는 DNA 또는 RNA를 포함할 수 있다. 이들은 단일 가닥 또는 이중 가닥일 수 있다. 수많은 서로 다른 폴리뉴클레오티드가 유전자 코드의 축퇴의 결과로서 동일한 폴리펩티드를 암호화할 수 있다는 것이 통상의 기술자에 의해 이해될 것이다. 또한, 통상의 기술자가 일상적인 기술을 사용하여 본 발명의 폴리뉴클레오티드에 의해 암호화되는 폴리펩티드 서열에 영향을 미치지 않는 뉴클레오티드 치환을 수행하여 본 발명의 폴리펩티드가 발현될 임의의 특정 숙주 유기체의 코돈 사용을 반영할 수 있다는 것을 이해하여야 한다.

[0209] 폴리뉴클레오티드는 당업계에서 이용가능한 임의의 방법에 의해 변형될 수 있다. 이러한 변형은 본 발명의 폴리뉴클레오티드의 생체내 활성 또는 수명을 향상시키기 위해 수행될 수 있다.

[0210] DNA 폴리뉴클레오티드와 같은 폴리뉴클레오티드는 재조합적으로, 합성적으로 또는 당업자에게 이용가능한 임의의 수단에 의해 생성될 수 있다. 이들은 표준 기술에 의해 클로닝될 수도 있다.

[0211] 더 긴 폴리뉴클레오티드는 일반적으로 재조합 수단, 예를 들어 중합효소 연쇄 반응(PCR) 클로닝 기술을 사용하여 생산될 것이다. 이것은 클로닝하고자 하는 표적 서열을 플랭킹하는 한 쌍의 프라이머(예: 약 15-30개 뉴클레오티드)를 만들고, 프라이머를 동물 또는 인간 세포로부터 수득한 mRNA 또는 cDNA와 접촉시키고, 원하는 영역을 증폭시키는 조건에서 PCR을 수행하고, 증폭된 단편을 분리하고(예: 반응 혼합물을 아가로스 겔로 정제함으로써), 그리고 증폭된 DNA를 회수하는 것을 포함한다. 프라이머는 증폭된 DNA가 적합한 벡터 내로 클로닝될 수 있도록 적절한 제한 효소 인식 부위를 포함하도록 설계될 수 있다.

[0212] **일반적인 레트로바이러스 벡터 구성요소**

[0213] **프로모터 및 인핸서**

[0214] NOI 및 폴리뉴클레오티드의 발현은 조절 서열, 예를 들어 전사 조절 요소 또는 번역 억제 요소를 사용하여 제어할 수 있으며, 여기에는 프로모터, 인핸서 및 기타 발현 조절 신호(예: tet 억제인자(TetR) 시스템) 또는 벡터 생산 세포에서 전이유전자 억제 시스템(Transgene Repression In vector Production cell system, TRiP) 또는 본원에 기재된 NOI의 다른 조절자가 포함된다.

[0215] 원핵생물 프로모터 및 진핵생물 세포에서 기능하는 프로모터가 사용될 수 있다. 조직-특이적 또는 자극-특이적 프로모터가 사용될 수 있다. 2개 이상의 상이한 프로모터로부터의 서열 요소를 포함하는 키메라 프로모터가 또한 사용될 수 있다.

[0216] 적합한 프로모팅(promoting) 서열은 강력한 프로모터로서, 폴리오마 바이러스, 아데노바이러스, 계두 바이러스, 소 유두종 바이러스, 조류 육종 바이러스, 사이토메갈로바이러스(CMV), 레트로바이러스 및 유인원 바이러스 40(SV40)과 같은 바이러스의 게놈으로부터 유래된 것들, 또는 액틴 프로모터, EF1 α , CAG, TK, SV40, 유비퀴틴, PGK 또는 리보솜 단백질 프로모터와 같은 이중 포유동물 프로모터로부터 유래된 것들을 포함한다. 대안적으로, 조직-특이적 프로모터, 예컨대 로돕신(Rho), 로돕신 키나제(RhoK), 원추형(cone-rod) 호메오박스 함유 유전자(CRX), 신경 망막-특이적 류신 지퍼 단백질(NRL), 난황 황반 이영양증 2(VMD2), 티로신 하이드록실라제, 뉴런-특이적 뉴런-특이적 에놀라제(NSE) 프로모터, 성상교세포-특이적 신경교 섬유성 산성 단백질(GFAP) 프로모터, 인간 α 1-항트립신(hAAT) 프로모터, 포스포에놀피루베이트 카르복시키나제(PEPCK), 간 지방산 결합 단백질 프로

모터, Flt-1 프로모터, INF- β 프로모터, Mb 프로모터, SP-B 프로모터, SYN1 프로모터, WASP 프로모터, SV40 / hAlb 프로모터, SV40 / CD43, SV40 / CD45, NSE / RU5' 프로모터, ICAM-2 프로모터, GPIIb 프로모터, GFAP 프로모터, 피브로넥틴 프로모터, 엔도글린 프로모터, 엘라스타제-1 프로모터, 데스민(Desmin) 프로모터, CD68 프로모터, CD14 프로모터 및 B29 프로모터를 사용하여 전사를 구동할 수 있다.

[0217] 인헨서 서열을 벡터에 삽입함으로써 NOI의 전사를 더욱 증가시킬 수 있다. 인헨서는 상대적으로 방향- 및 위치-독립적이다; 그러나, SV40 인헨서 및 CMV 초기 프로모터 인헨서와 같은 진핵 세포 바이러스로부터의 인헨서를 사용할 수 있다. 인헨서는 프로모터에 대해 5' 또는 3' 위치에서 벡터 내로 스플라이싱될 수 있지만, 바람직하게는 프로모터로부터 5' 부위에 위치한다.

[0218] 프로모터는 적절한 표적 세포에서 발현을 보장하거나 증가시키기 위한 구성을 추가로 포함할 수 있다. 예를 들어, 구성은 보존된 영역, 예를 들어 프리브나우(Pribnow) 박스 또는 TATA 박스일 수 있다. 프로모터는 뉴클레오타이드 서열의 발현 수준에 영향을 미치는(예컨대 유지, 향상 또는 감소시키는) 다른 서열을 함유할 수 있다. 적합한 다른 서열은 Sh1-인트론 또는 ADH 인트론을 포함한다. 다른 서열에는 온도, 화학, 빛 또는 응력 유도 요소와 같은 유도 요소가 포함된다. 또한, 전사 또는 번역을 향상시키기에 적합한 요소가 존재할 수 있다.

[0219] **NOI의 조절자**

[0220] 레트로바이러스 패키징/생산자 세포주의 생성 및 레트로바이러스 벡터 생산에 있어서 복잡한 요인은 특정 레트로바이러스 벡터 구성요소와 NOI의 구성적 발현이 세포독성을 일으켜 이러한 구성요소를 발현하는 세포를 사멸시켜 벡터를 생산할 수 없다는 것이다. 따라서, 이러한 구성요소(예: gag-pol 및 VSV-G와 같은 외피 단백질)의 발현이 조절될 수 있다. 다른 비-세포독성 벡터 구성요소, 예를 들어 rev의 발현 또한 세포에 대한 대사 부담을 최소화하기 위해 조절될 수 있다. 본원에 기재된 바와 같은 모델식 작제물 및/또는 세포는 적어도 하나의 조절 요소와 관련된 세포독성 및/또는 비-세포독성 벡터 구성요소를 포함할 수 있다. 본원에서 사용되는 용어 "조절 요소"는 관련 유전자 또는 단백질의 발현을 증가시키든 또는 감소시키든 영향을 미칠 수 있는 임의의 요소를 지칭한다. 조절 요소는 유전자 스위치 시스템, 전사 조절 요소 및 번역 억제 요소를 포함한다.

[0221] 여러 원핵생물 조절 시스템이 포유류 세포에서 유전자 스위치를 생성하도록 조정되었다. 많은 레트로바이러스 패키징 및 생산자 세포주를 유전자 스위치 시스템(예: 테트라사이클린 및 큐메이트 유도성 스위치 시스템)을 사용하여 제어함으로써 하나 이상의 레트로바이러스 벡터 구성요소의 발현이 벡터 생산 시에 스위치 온(switch on)될 수 있도록 하였다. 유전자 스위치 시스템에는 전사 조절자의 (TetR) 단백질 그룹의 것들(예: T-Rex, Tet-On 및 Tet-Off), 전사 조절자의 큐메이트 유도성 스위치 시스템 그룹의 것들(예: CymR 단백질) 및 RNA-결합 단백질에 관련된 것들(예: TRAP)이 포함된다.

[0222] 이러한 테트라사이클린-유도성 시스템 중 하나는 T-REX™ 시스템에 기초한 테트라사이클린 억제인자(TetR) 시스템이다. 예로써, 이러한 시스템에서 테트라사이클린 오퍼레이터(TetO2)는 첫 번째 뉴클레오타이드가 인간 사이토메갈로바이러스 주요 즉시 초기 프로모터(hCMVp)의 TATATAA 요소의 마지막 뉴클레오타이드의 3' 말단으로부터 10 bp가 되는 위치에 배치되고 그러면 TetR 단독으로 억제인자 역할을 할 수 있다(Yao F, Svensjo T, Winkler T, Lu M, Eriksson C, Eriksson E. Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. 1998. *Hum Gene Ther*; 9: 1939-1950). 이러한 시스템에서 NOI의 발현은 TetO2 서열의 2개 카피가 나란히(in tandem) 삽입된 CMV 프로모터에 의해 제어될 수 있다. TetR 동종이량체는, 유도제(테트라사이클린 또는 그 유사체 독시사이클린[dox]) 부재 시, TetO2 서열에 결합하고 상류 CMV 프로모터로부터의 전사를 물리적으로 차단한다. 존재하는 경우, 유도제는 TetR 동종이량체에 결합하여 알로스테릭 변화를 일으켜 TetO2 서열에 더 이상 결합할 수 없도록 하여 유전자 발현을 유발한다. TetR 유전자는 코돈 최적화될 수 있으며 이는 이것이 번역 효율을 개선하여 TetO2 제어된 유전자 발현을 더 엄격하게 제어할 수 있기 때문이다.

[0223] TRiP 시스템은 WO 2015/092440에 기술되어 있으며 벡터 생산 동안 생산 세포에서 NOI의 발현을 억제하는 또 다른 방법을 제공한다. 조직-특이적 프로모터를 포함한 구성적 및/또는 강력한 프로모터가 전이유전자를 구동하는 것이 선호되는 경우 그리고 특히 생산 세포에서 전이유전자 단백질의 발현이 벡터 역가의 감소를 초래하고/하거나 전이유전자-유래 단백질의 바이러스 벡터 전달로 인해 생체내 면역 반응을 유도하는 경우, TRAP-결합 서열(예: TRAP-tbs) 상호작용은 레트로바이러스 벡터 생산을 위한 전이유전자 단백질 억제 시스템의 기초를 형성한다(Maunders *et al*, Nat Commun. (2017) Mar 27; 8).

[0224] 간단히 말해서, TRAP-tbs 상호작용은 번역 블록을 형성하여 전이유전자 단백질의 번역을 억제한다(Maunders *et*

al, Nat Commun. (2017) Mar 27; 8). 번역 블록은 생산 세포에서만 효과적이며 따라서 DNA- 또는 RNA-기반 벡터 시스템을 방해하지 않는다. TRiP 시스템은 전이유전자 단백질이 단일 또는 다중 시스트론(cistronic) mRNA로부터 조직-특이적 프로모터를 포함한 구성적 및/또는 강력한 프로모터로부터 발현될 때 번역을 억제할 수 있다. 전이유전자 단백질의 조절되지 않은 발현이 벡터 역가를 감소시키고 벡터 생성물 품질에 영향을 미칠 수 있음이 입증되었다. 일시적 및 안정적 PaCL/PCL 벡터 생산 시스템 양자 모두에 대한 전이유전자 단백질의 억제는 생산 세포가 벡터 역가의 감소를 방지하는 데 유익하며: 이는 독성 또는 분자 부담 문제가 세포성 스트레스로 이어질 수 있는 경우; 전이유전자-유래 단백질의 바이러스 벡터 전달로 인해 전이유전자 단백질이 생체내 면역 반응을 유도하는 경우; 유전자-편집 전이유전자의 사용이 on/off 표적 영향을 초래할 수 있는 경우; 전이유전자 단백질이 벡터 및/또는 외피 당단백질 배제에 영향을 미칠 수 있는 경우 그러하다.

[0225] **외피 및 위형화(Pseudotyping)**

[0226] 하나의 바람직한 측면에서, 본원에 기재된 렌티바이러스 벡터는 위형화(슈도타입핑, pseudotype)된 것이다. 이와 관련하여, 위형화는 하나 이상의 이점을 제공할 수 있다. 예를 들어, HIV 기반 벡터의 *env* 유전자 산물은 이들 벡터가 CD4라고 하는 단백질을 발현하는 세포만을 감염시키는 것으로 제한한다. 그러나 이러한 벡터의 *env* 유전자가 다른 외피 바이러스로부터의 *env* 서열로 치환된 경우, 이들은 더 넓은 감염 스펙트럼을 가질 수 있다 (Verma and Somia (1997) Nature 389(6648):239-242). 예를 들어, 연구자들은 VSV로부터의 당단백질을 사용하여 HIV 기반 벡터를 위형화하였다(Verma and Somia (1997) Nature 389(6648):239-242).

[0227] 다른 대안에서, Env 단백질은 돌연변이 또는 조작된 Env 단백질과 같은 변형된 Env 단백질일 수 있다. 표적화 능력을 도입하거나 독성을 감소시키기 위해 또는 다른 목적을 위해 변형이 이루어지거나 선택될 수 있다 (Valsesia-Wittman et al 1996 J Virol 70: 2056-64; Nilson et al (1996) Gene Ther 3(4):280-286; 및 Fielding et al (1998) Blood 91(5):1802-1809 및 거기에 인용된 참조문헌).

[0228] 벡터는 선택한 분자로 위형화될 수 있다.

[0229] 본원에서 사용된 "env"는 본원에 기재된 바와 같이 내인성 렌티바이러스 외피 또는 이중 외피를 의미할 것이다.

[0230] **VSV-G**

[0231] 라브도바이러스(rhabdovirus)인 수포성 구내염 바이러스(vesicular stomatitis virus, VSV)의 외피 당단백질(G)은 특정 외피 바이러스 및 바이러스 벡터 비리온을 위형화할 수 있는 것으로 밝혀진 외피 단백질이다.

[0232] 레트로바이러스 외피 단백질이 없는 상태에서 MoMLV-기반 레트로바이러스 벡터를 위형화하는 이의 능력은 Emi et al. (1991) Journal of Virology 65:1202-1207에 의해 최초로 개시되었다. WO 1994/294440은 레트로바이러스 벡터가 VSV-G를 사용하여 성공적으로 위형화될 수 있음을 교시하고 있다. 이러한 위형화된 VSV-G 벡터는 광범위한 포유동물 세포를 형질도입하는데 사용될 수 있다. 보다 최근에, Abe et al. (1998) J Virol 72(8) 6356-6361은 비-감염성 레트로바이러스 입자가 VSV-G의 첨가에 의해 감염성으로 만들어질 수 있다고 교시하고 있다.

[0233] Burns et al. (1993) Proc. Natl. Acad. Sci. USA 90:8033-7은 VSV-G를 사용하여 레트로바이러스 MLV를 성공적으로 위형화하였으며, 이는 원래 형태의 MLV와 비교하여 변경된 숙주 범위를 가지는 벡터를 생산하였다. VSV-G 위형화된 벡터는 포유동물 세포뿐만 아니라 어류, 파충류 및 곤충에서 유래한 세포주도 감염시키는 것으로 나타났다(Burns et al. (1993) ibid). 이들은 또한 다양한 세포주에 대해 기존의 양쪽성(amphotropic) 외피보다 더 효율적인 것으로 나타났다(Yee et al., (1994) Proc. Natl. Acad. Sci. USA 91:9564-9568, Emi et al. (1991) Journal of Virology 65:1202-1207). VSV-G 단백질은 그의 세포질 꼬리가 레트로바이러스 코어와 상호작용할 수 있기 때문에 특정 레트로바이러스를 위형화하는데 사용할 수 있다.

[0234] VSV-G 단백질과 같은 비-레트로바이러스 위형화 외피의 제공은 벡터 입자가 감염성의 손실 없이 높은 역가로 농축될 수 있다는 이점을 제공한다(Akkina et al. (1996) J. Virol. 70:2581-5). 레트로바이러스 외피 단백질은 초원심분리 동안 전단력을 분명히 견딜 수 없고, 이는 아마도 이들이 2개의 비-공유적으로 연결된 서브유닛으로 구성되어 있기 때문일 것이다. 서브유닛 간의 상호작용이 원심분리에 의해 방해될 수 있다. 이에 비해 VSV 당단백질은 단일 단위로 구성된다. 따라서 VSV-G 단백질 위형화는 효율적인 표적 세포 감염/형질도입 및 제조 공정 동안 모두에 대하여 잠재적인 이점을 제공할 수 있다.

[0235] WO 2000/52188은 안정적인 생산자 세포주로부터, 수포성 구내염 바이러스-G 단백질(VSV-G)을 막-연관 바이러스 외피 단백질로 가지는 위형화된 레트로바이러스 벡터의 생성을 기재하고 있으며, VSV-G 단백질에 대한 유전자

서열을 제공한다.

[0236] **로스 리버 바이러스(Ross River Virus)**

[0237] 로스 리버 바이러스 외피는 비영장류 렌티바이러스 벡터(FIV)를 위형화하는 데 사용되었으며 전신 투여 후 주로 간에 형질도입하였다(Kang *et al.*, 2002, *J. Virol.*, 76:9378-9388). 효율성은 VSV-G 위형화된 벡터로 얻은 것보다 20배 더 큰 것으로 보고되었으며, 간독성(hepatotoxicity)을 시사하는 간 효소의 혈청 수준으로 측정했을 때 세포독성이 더 적었다.

[0238] **배큘로바이러스(Baculovirus) GP64**

[0239] 배큘로바이러스 GP64 단백질은 임상 및 상업적 적용에 필요한 고역가 바이러스의 대규모 생산에 사용되는 바이러스 벡터에 대해서 VSV-G의 대안인 것으로 밝혀졌다(Kumar M, Bradow BP, Zimmerberg J (2003) *Hum Gene Ther.* 14(1):67-77). VSV-G-위형화된 벡터와 비교하여, GP64-위형화된 벡터는 유사한 넓은 지향성(tropism) 및 유사한 고유 역가를 가진다. GP64 발현은 세포를 죽이지 않기 때문에, GP64를 구성적으로 발현하는 HEK293T 기반 세포주가 생성될 수 있다.

[0240] **대체 외피**

[0241] EIAV를 위형화하는데 사용될 때 합리적인 역가를 제공하는 다른 외피로는 모콜라(Mokola), 광견병(Rabies), 에볼라 및 LCMV(림프구성 맥락수막염 바이러스)가 포함된다. 4070A로 위형화된 렌티바이러스의 마우스로의 정맥내 주입은 간에서 최대 유전자 발현을 유도하였다.

[0242] **패키징 서열**

[0243] 본 발명의 맥락에서 사용되는 바와 같이, "패키징 서열" 또는 "psi"로 상호교환적으로 지칭되는 "패키징 신호"라는 용어는 바이러스 입자 형성 동안 레트로바이러스 RNA 가닥의 캡슐화(encapsidation)에 필요한 비-코딩, 시스-작용 서열과 관련하여 사용된다. HIV-1에 있어서, 이 서열은 주요 스플라이스 공여자 부위(SD)의 상류에서부터 적어도 *gag* 시작 코돈까지 연장되는 유전자좌에 매핑되어 있다(*gag*의 5' 서열의 일부 또는 전부에서부터 뉴클레오타이드 688까지가 포함될 수 있음). EIAV에서 패키징 신호는 *Gag*의 5' 코딩 영역 내로의 R 영역을 포함한다.

[0244] 본원에서 사용되는 용어 "연장된 패키징 신호" 또는 "연장된 패키징 서열"은 *gag* 유전자 내로 추가로 확장된 psi 서열 주위의 서열의 사용을 지칭한다. 이러한 추가 패키징 서열의 포함은 벡터 RNA를 바이러스 입자에 삽입하는 효율성을 증가시킬 수 있다.

[0245] 고양이 면역결핍 바이러스(FIV) RNA 캡슐화 결정인자는 게놈 mRNA의 5' 말단(R-U5)에 하나의 영역을 포함하고 *gag*의 근위(proximal) 311 nt 내에 매핑된 다른 영역을 포함하여, 분리되고 불연속적인 것으로 나타났다(Kaye *et al.*, *J Virol.* Oct;69(10):6588-92 (1995)).

[0246] **내부 리보솜 진입 부위(IRES)**

[0247] IRES 요소의 삽입은 단일 프로모터로부터 다중 코딩 영역의 발현을 허용한다(Adam *et al* (상기와 같이); Koo *et al* (1992) *Virology* 186:669-675; Chen *et al* 1993 *J. Virol* 67:2142-2148). IRES 요소는 피코르나바이러스(picornavirus)의 번역되지 않은 5' 말단에서 처음 발견되었는데 여기서 이들은 바이러스 단백질의 cap-독립적 번역을 촉진한다(Jang *et al* (1990) *Enzyme* 44: 292-309). RNA의 오픈 리딩 프레임 사이에 위치하는 경우, IRES 요소는 IRES 요소에서 리보솜의 진입을 촉진한 후 하류 번역 개시를 촉진함으로써 하류 오픈 리딩 프레임의 효율적인 번역을 허용한다.

[0248] IRES에 대한 리뷰는 Mountford와 Smith가 제공한다(TIG May 1995 vol 11, No 5:179-184). 다수의 상이한 IRES 서열이 알려져 있으며 뇌심근염 바이러스(encephalomyocarditis virus, EMCV)로부터의 서열(Ghattas, I.R., *et al.*, *Mol. Cell. Biol.*, 11:5848-5859 (1991); BiP 단백질 [Macejak and Sarnow, *Nature* 353:91 (1991)]; 드로소필라(*Drosophila*)의 Antennapedia 유전자(엑손 d 및 e) [Oh, *et al.*, *Genes & Development*, 6:1643-1653 (1992)] 뿐만 아니라 소아마비 바이러스(polio virus, PV)의 서열 [Pelletier and Sonenberg, *Nature* 334: 320-325 (1988); see also Mountford and Smith, TIG 11, 179-184 (1985)]을 포함한다.

[0249] PV, EMCV 및 돼지 수포성 질병 바이러스로부터의 IRES 요소는 이전에 레트로바이러스 벡터에 사용되었다(Coffin *et al.*, as above).

- [0250] "IRES"라는 용어는 IRES로 작용하거나 IRES의 기능을 향상시키는 임의의 서열 또는 서열의 조합을 포함한다. IRES(들)는 바이러스 기원(예컨대 EMCV IRES, PV IRES, 또는 FMDV 2A-유사 서열) 또는 세포 기원(예컨대 FGF2 IRES, NRF IRES, Notch 2 IRES 또는 EIF4 IRES)일 수 있다.
- [0251] IRES가 각 폴리뉴클레오티드의 번역을 개시할 수 있으려면 모듈식 작제물에서 폴리뉴클레오티드 사이 또는 그 앞에 위치해야 한다.
- [0252] 안정적인 세포주의 개발에 사용되는 뉴클레오티드 서열은 안정적인 통합이 발생한 세포를 선택하기 위한 선택 가능한 마커의 추가를 필요로 한다. 이러한 선택 가능한 마커는 뉴클레오티드 서열 내에서 단일 전사 단위로서 발현될 수 있거나 또는 폴리스리스트론 메시지에서 선택 가능한 마커의 번역을 개시하기 위해 IRES 요소를 사용하는 것이 바람직할 수 있다(Adam et al 1991 J.Virol. 65, 4985).
- [0253] **유전자 배향 및 인슐레이터**
- [0254] 핵산은 방향성이 있으며 이는 궁극적으로 세포에서 전사 및 복제와 같은 메커니즘에 영향을 미친다는 것이 잘 알려져 있다. 따라서 유전자는 동일한 핵산 작제물의 일부일 때 서로에 대해 상대적인 배향을 가질 수 있다.
- [0255] 본 발명의 특정 구현예에서, 세포 또는 작제물의 동일한 유전자좌에 존재하는 적어도 2개의 핵산 서열은 역방향 및/또는 교대(alternating) 배향일 수 있다. 다시 말해서, 본 발명의 특정 구현예에서, 이 특정 유전자좌에서, 순차적 유전자 쌍은 동일한 배향을 갖지 않을 것이다. 이것은 영역이 숙주 세포의 동일한 물리적 위치 내에서 발현될 때 전사 및 번역 리드-스루를 모두 방지하는 데 도움이 될 수 있다.
- [0256] 벡터 생산에 필요한 핵산이 세포 내 동일한 유전적 유전자좌에 기반을 둘 때, 교대 배향을 가지는 것은 레트로 바이러스 벡터 생산에 이점이 있다. 이것은 차례로 또한 복제 가능한 레트로바이러스 벡터의 생성을 방지하는바 생성된 작제물의 안전성을 향상시킬 수 있다.
- [0257] 핵산 서열이 역방향 및/또는 교대 배향일 때 인슐레이터의 사용은 NOI의 유전자 환경으로부터 NOI의 부적절한 발현 또는 사일런싱(침묵)을 방지할 수 있다.
- [0258] "인슐레이터"라는 용어는 인슐레이터-결합 단백질에 결합될 때 주변 조절자 신호로부터 유전자를 보호하는 능력을 가지는 DNA 서열 요소의 부류를 지칭한다. 인슐레이터에는 인핸서 차단 기능과 염색질 장벽 기능의 두 가지 유형이 있다. 인슐레이터가 프로모터와 인핸서 사이에 위치할 때, 인슐레이터의 인핸서-차단 기능은 인핸서의 전사-강화 영향으로부터 프로모터를 보호한다(Geyer and Corces 1992; Kellum and Schedl 1992). 염색질 장벽 인슐레이터는 근처의 응축된 염색질의 진행을 막음으로써 기능하며, 이는 전사적으로 활성인 염색질 영역이 전사적으로 불활성인 염색질 영역으로 전환되게 하고 결과적으로 유전자 발현을 사일런싱시킨다. 이종염색질(heterochromatin)의 확산을 억제하여 따라서 유전자를 사일런싱하는 인슐레이터는 이 과정을 방지하기 위해 히스톤 변형에 관여하는 효소를 리크루트한다(Yang J, Corces VG. 2011;110:43-76; Huang, Li et al. 2007; Dhillon, Raab et al. 2009). 인슐레이터는 이러한 기능 중 하나 또는 둘 다를 가질 수 있으며 닭 β-글로빈 인슐레이터(cHS4)가 하나의 그러한 예이다. 이 인슐레이터는 가장 광범위하게 연구된 척추동물 인슐레이터이며, G+C가 매우 풍부하며 인핸서-차단 및 이종염색질 장벽 기능을 모두 가지고 있다(Chung J H, Whitely M, Felsenfeld G. Cell. 1993;74:505-514). 인핸서 차단 기능을 가지는 다른 이러한 인슐레이터는 인간 β-글로빈 인슐레이터 5(HS5), 인간 β-글로빈 인슐레이터 1(HS1) 및 닭 β-글로빈 인슐레이터(cHS3)를 포함하지만 이에 제한되지 않는다(Farrell CM1, West AG, Felsenfeld G., Mol Cell Biol. 2002 Jun;22(11):3820-31; J Ellis et al. EMBO J. 1996 Feb 1; 15(3): 562-568). 원치 않는 원위(distal) 상호작용을 줄이는 것 외에도 인슐레이터는 인접한 레트로바이러스 핵산 서열 사이의 프로모터 간섭(즉, 하나의 전사 단위의 프로모터가 인접한 전사 단위의 발현을 손상시키는 경우)을 방지하는 데 도움이 된다. 각각의 레트로바이러스 벡터 핵산 서열 사이에 인슐레이터를 사용하는 경우, 직접적인 리드-스루의 감소가 복제 가능한 레트로바이러스 벡터 입자의 형성을 방지하는 데 도움이 된다.
- [0259] 인슐레이터는 각각의 레트로바이러스 핵산 서열 사이에 존재할 수 있다. 일 구현예에서, 인슐레이터의 사용은 벡터 구성요소를 암호화하는 뉴클레오티드 서열에서 하나의 NOI 발현 카세트가 또 다른 NOI 발현 카세트와 상호 작용하는 프로모터-인핸서 상호작용을 방지한다.
- [0260] 벡터 계놈과 gag-pol 서열 사이에 인슐레이터가 존재할 수 있다. 따라서 이는 복제 가능한 레트로바이러스 벡터 및 RNA 전사체와 같은 '야생형'의 생성 가능성을 제한함으로써, 작제물의 안전성 프로파일을 개선한다. 안정적으로 통합된 다중 유전자 벡터의 발현을 개선하기 위한 인슐레이터 요소의 사용은 Moriarity et al, Nucleic

Acids Res. 2013 Apr;41(8):e92에 인용되어 있다.

- [0261] **백터 역가**
- [0262] 당업자는 렌티바이러스 백터의 역가를 결정하는 다수의 상이한 방법이 있음을 이해할 것이다. 역가는 종종 형질 도입 단위/mL(TU/mL)로 설명된다. 백터 입자의 수를 증가시키고 백터 제제의 비활성(specific activity)을 증가 시킴으로써 역가를 증가시킬 수 있다.
- [0263] **치료 용도**
- [0264] 본원에 기재된 렌티바이러스 백터 또는 본원에 기재된 렌티바이러스 백터로 형질도입된 세포 또는 조직은 의학에 사용될 수 있다.
- [0265] 또한, 본원에 기재된 렌티바이러스 백터, 본 발명의 생산 세포 또는 본원에 기재된 렌티바이러스 백터로 형질도입된 세포 또는 조직은 관심 뉴클레오티드를 이를 필요로 하는 표적 부위에 전달하기 위한 약제의 제조에 사용될 수 있다. 본 발명의 렌티바이러스 백터 또는 형질도입된 세포의 이러한 용도는 전술한 바와 같이 치료 또는 진단 목적을 위한 것일 수 있다.
- [0266] 따라서, 본원에 기재된 바와 같은 렌티바이러스 백터에 의해 형질도입된 세포가 제공된다.
- [0267] "바이러스 백터 입자에 의해 형질도입된 세포"는 바이러스 백터 입자에 의해 운반되는 핵산이 전달된 세포, 특히 표적 세포로 이해되어야 한다.
- [0268] 본 발명의 일 구현예에서, 관심 뉴클레오티드는 TRAP가 걸여된 표적 세포에서 번역된다.
- [0269] "표적 세포"는 NOI를 발현하고자 하는 세포로 이해되어야 한다. NOI는 본 발명의 바이러스 백터를 사용하여 표적 세포 내로 도입될 수 있다. 표적 세포로의 전달은 생체내, 생체의 또는 시험관내에서 수행될 수 있다.
- [0270] 바람직한 구현예에서, 관심 뉴클레오티드는 치료 효과를 일으킨다.
- [0271] NOI는 치료 또는 진단 용도를 가질 수 있다. 적합한 NOI는 효소, 보조인자(co-factor), 사이토카인, 케모카인, 호르몬, 항체, 항산화 분자, 조작된 면역글로불린-유사 분자, 단일 사슬 항체, 융합 단백질, 면역 공동-자극 분자, 면역조절 분자, 키메라 항원 수용체, 표적 단백질의 트랜스도메인 음성 돌연변이, 독소, 조건부 독소, 항원, 전사 인자, 구조 단백질, 리포터 단백질, 세포내(subcellular) 국소화 신호, 종양 억제 단백질, 성장 인자, 막 단백질, 수용체, 혈관 활성(vasoactive) 단백질 또는 펩티드, 항바이러스성 단백질 또는 리보자임(ribozyme), 또는 이의 유도체(예컨대 관련 리포터 그룹을 가지는 유도체)를 암호화하는 서열을 포함하지만 이에 제한되지는 않는다. NOI는 또한 마이크로-RNA를 암호화할 수 있다. 이론에 얽매이지 않음, 마이크로-RNA의 처리는 TRAP에 의해 억제될 것으로 믿어진다.
- [0272] 일 구현예에서, NOI는 신경퇴행성 장애의 치료에 유용할 수 있다.
- [0273] 또 다른 구현예에서, NOI는 파킨슨병의 치료에 유용할 수 있다.
- [0274] 또 다른 구현예에서, NOI는 도파민 합성에 관여하는 효소 또는 효소들을 암호화할 수 있다. 예를 들어, 효소는 티로신 히드록실라제, GTP-시클로히드록실라제 I 및/또는 방향족 아미노산 도파 데카르복실라제 중 하나 이상일 수 있다. 세 가지 유전자의 서열은 모두 이용가능하다(각각 GenBank® Accession Nos. X05290, U19523 및 M76180).
- [0275] 또 다른 구현예에서, NOI는 소포성 모노아민 수송체 2(vesicular monoamine transporter 2, VMAT2)를 암호화할 수 있다. 대안적인 구현예에서 바이러스 계통은 방향족 아미노산 도파 데카르복실라제를 암호화하는 NOI 및 VMAT2를 암호화하는 NOI를 포함할 수 있다. 이러한 계통은, 특히 L-DOPA의 말초 투여와 함께, 파킨슨병의 치료에 사용될 수 있다.
- [0276] 또 다른 구현예에서 NOI는 치료 단백질 또는 치료 단백질의 조합을 암호화할 수 있다.
- [0277] 또 다른 구현예에서, NOI는 신경교 세포 유래 신경영양 인자(glial cell derived neurotrophic factor, GDNF), 뇌 유래 신경영양 인자(BDNF), 모양체 신경영양 인자(ciliary neurotrophic factor, CNTF), 뉴로트로핀-3(NT-3), 산성 섬유아세포 성장 인자(aFGF), 염기성 섬유아세포 성장 인자(bFGF), 인터루킨-1 베타(IL-1 β), 종양 괴사 인자 알파(TNF- α), 인슐린 성장 인자-2, VEGF-A, VEGF-B, VEGF-C/VEGF-2, VEGF-D, VEGF-E, PDGF-A, PDGF-B, PDGF-A 및 PDGF-B의 이종이량체 및 동종이량체로 이루어진 군으로부터 선택되는 단백질 또는 단백질들을

암호화할 수 있다.

- [0278] 또 다른 구현예에서, NOI는 안지오스타틴, 엔도스타틴, 혈소관 인자 4, 색소 상피 유래 인자(PEDF), 태반 성장 인자, 레스틴(restin), 인터페론- α , 인터페론-유도성 단백질, 그로-베타(gro-beta) 및 튜브다운-1(tubedown-1), 인터류킨(IL)-1, IL-12, 레티노산, 항-VEGF 항체 또는 이의 단편/변이체, 예컨대 애플리버셉트(aflibercept), 트롬보스폰딘, VEGF 수용체 단백질, 예컨대 US 5,952,199 및 US 6,100,071에 기재된 것들, 및 항-VEGF 수용체 항체로 이루어진 군으로부터 선택되는 항-혈관신생 단백질 또는 항-혈관신생 단백질들을 암호화할 수 있다.
- [0279] 또 다른 구현예에서, NOI는 NF-kB 억제제, IL1베타 억제제, TGF베타 억제제, IL-6 억제제, IL-23 억제제, IL-18 억제제, 종양 괴사 인자 알파 및 종양 괴사 인자 베타, 림프독소 알파 및 림프독소 베타, LIGHT 억제제, 알파 시누클레인(synuclein) 억제제, Tau 억제제, 베타 아밀로이드 억제제, IL-17 억제제로 이루어진 군으로부터 선택되는 항염증성 단백질, 항체 또는 단백질 또는 항체의 단편/변이체를 암호화할 수 있다.
- [0280] 다른 구현예에서 NOI는 낭포성 섬유증 막관통 전도도 조절인자(cystic fibrosis transmembrane conductance regulator, CFTR)를 암호화할 수 있다.
- [0281] 다른 구현예에서 NOI는 안구 세포에서 정상적으로 발현되는 단백질을 암호화할 수 있다.
- [0282] 다른 구현예에서, NOI는 광수용체 세포 및/또는 망막 색소 상피 세포에서 정상적으로 발현되는 단백질을 암호화할 수 있다.
- [0283] 또 다른 구현예에서, NOI는 RPE65, 아틸탄화수소-상호작용 수용체 단백질 유사 1(AIPL1), CRB1, 레티날 레티날 아세틸트랜스퍼라제(LRAT), 광수용체-특이적 호메오 박스(CRX), 레티날 구아닐레이트 사이클리제(retinal guanylate cyclase, GUCY2D), RPGR 상호작용 단백질 1(RPGRIP1), LCA2, LCA3, LCA5, 디스트로핀(dystrophin), PRPH2, CNTF, ABCR/ABCA4, EMP1, TIMP3, MERTK, ELOVL4, MYO7A, USH2A, VMD2, RLBP1, COX-2, FPR, 하모닌(harmonin), Rab 에스코트 단백질 1, CNGB2, CNGA3, CEP 290, RPGR, RS1, RP1, PRELP, 글루타티온 경로 효소 및 옵티신(opticin)으로 이루어진 군으로부터 선택되는 단백질을 암호화할 수 있다.
- [0284] 다른 구현예에서, NOI는 인간 응고 인자 VIII 또는 인자 IX를 암호화할 수 있다.
- [0285] 다른 구현예에서, NOI는 페닐알라닌 히드록실라제(PAH), 메틸말로닐 CoA 뮤타제, 프로피오닐 CoA 카르복실라제, 이소발레릴 CoA 탈수소효소(dehydrogenase), 분지쇄 케토산(ketoacid) 탈수소효소 복합체, 글루타릴 CoA 탈수소 효소, 아세틸 CoA 카르복실라제, 프로피오닐 CoA 카르복실라제, 3 메틸 크로토닐(crotonyl) CoA 카르복실라제, 피루브산 카르복실라제, 카르바모일-포스페이트 신타제(synthase) 암모니아, 오르니틴 트랜스카르바밀라제(transcarbamylase), 글루코실세라미다제 베타, 알파 갈락토시다제 A, 글루코실세라미다제 베타, 시스티노신(cystinosin), 글루코사민(N-아세틸)-6-설파타제(sulfatase), N-아세틸-알파-글루코사미니다제, N-설포글루코사민 설포하이드롤라제, 갈락토스아민-6 설파타제, 아틸설파타제 A, 시토크롬 B-245 베타, ABCD1, 오르니틴 카르바모일트랜스퍼라제, 아르기니노숙시네이트 신타제, 아르기니노숙시네이트 리아제(lyase), 아르기나제 1, 알라닌 글리코실레이트(glycochylate) 아미노 트랜스퍼라제, ATP-결합 카세트, 서브-패밀리 B 멤버를 포함하는 군으로부터 선택되는 대사에 관여하는 단백질 또는 단백질들을 암호화할 수 있다.
- [0286] 다른 구현예에서, NOI는 키메라 항원 수용체(CAR) 또는 T 세포 수용체(TCR)를 암호화할 수 있다. 일 구현예에서, CAR은 항-5T4 CAR이다. 다른 구현예에서, NOI는 B-세포 성숙 항원(BCMA), CD19, CD22, CD20, CD138, CD30, CD33, CD123, CD70, 전립선 특이적 막 항원(PSMA), 루이스 Y 항원(LeY), 티로신-단백질 키나제 막관통 수용체(ROR1), Mucin 1, 세포 표면 관련(cell surface associated, Muc1), 상피 세포 부착 분자(EpCAM), 내피 성장 인자 수용체(EGFR), 인슐린, 단백질 티로신 포스파타제, 비-수용체 유형 22, 인터류킨 2 수용체 알파, 헬리카제 C 도메인 1로 유도된 인터페론, 인간 표피 성장 인자 수용체(HER2), 글리피칸 3(glypican 3, GPC3), 디시알로강글리오시드(disialoganglioside, GD2), 메시오텔린(mesothelin), 수포 내피 성장 인자 수용체 2(VEGFR2)를 암호화할 수 있다.
- [0287] 다른 구현예에서, NOI는 ULBP1, 2 및 3, H60, Rae-1a, b, g, d, MICA, MICB를 포함하는 군으로부터 선택되는 NKG2D 리간드에 대한 키메라 항원 수용체(CAR)를 암호화할 수 있다.
- [0288] 추가의 구현예에서 NOI는 SGSH, SUMF1, GAA, 공통 감마 사슬(CD132), 아데노신 데아미나제(deaminase), WAS 단백질, 글로빈, 알파 갈락토시다제 A, δ -아미노레불린산(aminolevulinate, ALA) 신타제, δ -아미노레불린산 탈수효소(dehydratase)(ALAD), 하이드록시메틸빌란(Hydroxymethylbilane, HMB) 신타제, 우로포르피리노젠

(Uroporphyrinogen, URO) 신타제, 우로포르피리노겐(URO) 데카르복실라제, 코프로포르피리노겐(Coproporphyrinogen, COPRO) 옥시다제, 프로토포르피리노겐(Protoporphyrinogen, PROTO) 옥시다제, 페로킬라타제(Ferrochelataze), α -L-이두로니다제(α -L-iduronidase), 이두로네이트 설패타제(Iduronate sulfatase), 헤파란 설패미다제(Heparan sulfamidase), N-아세틸글루코사미니다제, 헤파란- α -글루코사미나이드 N-아세틸트랜스퍼라제, 3 N-아세틸글루코사민 6-설패타제, 갈락토스-6-설패이트 설패타제, β -갈락토시다제, N-아세틸갈락토스아민-4-설패타제, β -글루쿠로니다제 및 히알루로니다제를 암호화할 수 있다.

[0289] NOI 외에도 벡터는 또한 siRNA, shRNA, 또는 조절된 shRNA를 포함하거나 암호화할 수 있다. (Dickins et al. (2005) Nature Genetics 37: 1289-1295, Silva et al. (2005) Nature Genetics 37:1281-1288).

[0290] **적용중**

[0291] 본 발명에 따른 레트로바이러스 및 AAV 벡터를 포함하는 벡터는 WO 1998/05635, WO 1998/07859, WO 1998/09985에 열거된 장애의 치료에 유용한 하나 이상의 NOI(들)을 전달하는데 사용될 수 있다. 관심 뉴클레오티드는 DNA 또는 RNA일 수 있다. 그러한 질병의 예는 다음과 같다:

[0292] 사이토카인 및 세포 증식/분화 활성화에 반응하는 장애; 면역억제제 또는 면역자극제 활성화(예: 인간 면역결핍 바이러스에 의한 감염을 포함한 면역 결핍 치료, 림프구 성장 조절, 암 및 여러 자가면역 질환 치료, 및 이식 거부반응 방지 또는 중앙 면역 유도); 조혈 조절(예: 골수성 또는 림프계 질환의 치료); 뼈, 연골, 힘줄, 인대 및 신경 조직의 성장 촉진(예: 상처 치유용, 화상, 궤양 및 치주 질환 및 신경변성(neurodegeneration) 치료); 난포-자극 호르몬의 억제 또는 활성화(가임력 조절); 화학주성/화학동역학적(chemotactic/chemokinetic) 활성화(예: 특정 세포 유형을 손상 또는 감염 부위로 이동시키기 위한); 지혈(haemostatic) 및 혈전 용해(thrombolytic) 활성화(예: 혈우병 및 뇌졸중(stroke) 치료); 항염증 활성화(예를 들어, 패혈성 쇼크 또는 크론병 치료용); 대식세포 억제 및/또는 T 세포 억제 활성화 및 이에 따른 항염증 활성화; 항-면역 활성화(즉, 염증과 관련되지 않은 반응을 포함하는 세포성 및/또는 체액성 면역 반응에 대한 억제 효과); 대식세포 및 T 세포가 세포외 기질 성분 및 피브로넥틴에 부착하는 능력의 억제, 뿐만 아니라 T 세포에서 상향 조절된 fas 수용체 발현.

[0293] 암, 백혈병, 양성 및 악성 종양 성장, 침윤 및 확산, 혈관신생, 전이, 복수(ascites) 및 악성 흉막 삼출(pleural effusion)을 포함하는 악성 장애.

[0294] 류마티스 관절염을 비롯한 관절염, 과민증, 알레르기 반응, 천식, 전신성 홍반성 루푸스, 콜라겐 질환 및 기타 질환을 포함하는 자가면역 질환.

[0295] 동맥경화증, 동맥경화성 심장질환, 재관류 손상, 심정지, 심근경색증, 혈관염증질환, 호흡곤란증후군(respiratory distress syndrome), 심혈관계 영향, 말초혈관질환, 편두통 및 아스피린 의존성 항혈전증, 뇌졸중, 뇌허혈(cerebral ischaemia), 허혈성 심장질환 또는 기타 질병을 포함하는 혈관 질환.

[0296] 소화성 궤양(peptic ulcer), 궤양성 대장염, 크론병 및 기타 질병을 포함하는 위장관 질환.

[0297] 간 섬유증, 간경화를 포함하는 간 질환.

[0298] 페닐케톤뇨증 PKU, 윌슨병, 유기산혈증(organic acidemias), 요소 순환 장애, 담즙정체(cholestasis), 및 기타 질병을 포함하는 유전성 대사 장애.

[0299] 갑상선염 또는 기타 선(glandular) 질환, 사구체신염 또는 기타 질병을 포함하는 신장 및 비뇨기과 질환.

[0300] 중이염 또는 기타 이비인후과 질환을 포함하는 귀, 코 및 인후 장애, 피부염 또는 기타 피부 질환.

[0301] 치주 질환, 치주염, 치은염 또는 기타 치과/구강 질환을 포함하는 치과 및 구강 질환.

[0302] 고환염 또는 부고환-고환염(epididymo-orchitis), 불임, 고환 외상(orchidal trauma) 또는 기타 고환 질환을 포함하는 고환 질환.

[0303] 태반 기능 장애, 태반 기능 부전, 습관성 유산, 자간증(eclampsia), 전자간증, 자궁내막증 및 기타 부인과 질환을 포함하는 부인과 질환.

[0304] 안과 장애, 예컨대 LCA10을 포함하는 레버 선천적 흑암증(Leber Congenital Amaurosis, LCA), 후방포도막염, 중간포도막염, 전방포도막염, 결막염, 맥락망막염(chorioretinitis), 포도막망막염(uveoretinitis), 시신경염, 개방각 녹내장 및 청소년 선천성 녹내장을 포함하는 녹내장, 안내 염증, 예를 들어 망막염 또는 낭포성 황반부종, 교감성 안과(sympathetic ophthalmia), 공막염(scleritis), 색소성 망막염, 연령 관련 황반 변성(AMD) 및

청소년 황반 변성을 포함하는 황반 변성, 베스트병(Best Disease), 베스트 난황 황반변성(Best vielliform macular degeneration), 스타가르트병(Stargardt's Disease), 어서 증후군, Doyme's 벌집형 망막 이영양증(Doyme's honeycomb retinal dystrophy), Sorby's 황반 이영양증(Sorby's Macular Dystrophy), 소아 망막 분열증(Juvenile retinoschisis), 원추간상 이영양증(Cone-Rod Dystrophy), 각막 이영양증(Corneal Dystrophy), Fuch's 이영양증(Fuch's Dystrophy), Leber's 선천적 흑암증(Leber's congenital amaurosis), Leber's 유전성 시신경병증(Leber's hereditary optic neuropathy, LHON), Adie 증후군, 오구치병(Oguchi disease), 퇴행성 폰 더스병, 안구 외상, 감염에 의한 안구 염증, 증식성 유리체-망막병증(vitreo-retinopathies), 급성 허혈성 시신경병증, 과도한 흉터, 예를 들어 녹내장 여과 수술 후, 안구 이식물에 대한 반응, 각막 이식편 이식 거부, 및 기타 안과 질환, 예컨대 당뇨병성 황반 부종, 망막 정맥 폐쇄, RBP1 관련 망막 이영양증, 맥락막 혈증(choroideremia) 및 색맹.

[0305] 파킨슨병, 파킨슨병의 치료로 인한 합병증 및/또는 부작용, AIDS 관련 치매 복합성 HIV 관련 뇌병증(encephalopathy), 데빅병(Devic's disease), 시드넘 무도병(Sydenham chorea), 알츠하이머병 및 기타 퇴행성 질환, 중추신경계(CNS)의 병태 또는 장애, 뇌졸중, 소아마비 후 증후군, 정신 질환, 척수염, 뇌염(encephalitis), 아급성 경화성 범뇌염(subacute sclerosing pan-encephalitis), 뇌척수염, 급성 신경병증, 아급성 신경병증, 만성 신경병증, 파브리병(Fabry disease), 고셔병(Gaucher disease), 시스틴증(Cystinosis), 폼페병(Pompe disease), 이색성 백혈이영양증(metachromatic leukodystrophy), 비스콧 알드리치 증후군(Wiscott Aldrich Syndrome), 부신백질이영양증(adrenoleukodystrophy), 베타 지중해 빈혈(beta-thalassemia), 겸상적혈구병, 길랑-바레 증후군(Guillain-Barre syndrome), 시드넘 무도병, 중증 근무력증(myasthenia gravis), 가성 뇌종양(pseudo-tumour cerebri), 다운 증후군, 헌팅턴병, 중추신경계(CNS) 압박 또는 중추신경계 외상 또는 중추신경계의 감염, 근위축 및 이영양증, 중추 및 말초 신경계의 질병, 병태 또는 장애, 근위축성 측삭 경화증, 척수성 근위축증, 척수 및 박리 손상(spinal cord and avulsion injury)을 포함하는 운동 신경 질환을 포함하는 신경학적 및 신경퇴행성 장애.

[0306] 기타 질병 및 병태, 예컨대 낭포성 섬유증, 산필리포(Sanfilipo) 증후군 A, 산필리포 증후군 B, 산필리포 증후군 C, 산필리포 증후군 D를 포함하는 점액다당류증(mucopolysaccharidosis), 헌터 증후군, 헐러-샤이 증후군(Hurler-Scheie syndrome), 모르퀴오 증후군(Morquio syndrome), ADA-SCID, X-연관 SCID, X-연관 만성 육아종성(granulomatous) 질환, 포르피린증(porphyrria), 혈우병 A, 혈우병 B, 외상 후 염증, 출혈, 응고 및 급성기 반응, 악액질(cachexia), 식욕부진(anorexia), 급성 감염, 패혈성 쇼크, 감염성 질환, 당뇨병, 수술의 합병증 또는 부작용, 골수 이식 또는 기타 이식의 합병증 및/또는 부작용, 유전자 치료의 합병증 및 부작용, 예를 들어 바이러스 매개체의 감염 또는 AIDS로 인한 것, 체액성 및/또는 세포성 면역 반응 억제 또는 저해, 천연 또는 인공 세포, 조직 및 장기, 예컨대 각막, 골수, 장기, 렌즈, 심박 조율기(pacemaker), 천연 또는 인공 피부 조직의 이식의 경우 이식 거부의 예방 및/또는 치료.

[0307] **siRNA, 마이크로-RNA 및 shRNA**

[0308] 특정 다른 구현예에서, NOI는 마이크로-RNA를 포함한다. 마이크로-RNA는 유기체에서 자연적으로 생성되는 small RNA의 매우 큰 그룹으로, 그 중 적어도 일부는 표적 유전자의 발현을 조절한다. 마이크로-RNA 패밀리의 창립 멤버는 *let-7*과 *lin-4*이다. *let-7* 유전자는 벌레 발달 동안 내인성 단백질-코딩 유전자의 발현을 조절하는 작고 고도로 보존된 RNA 종을 암호화한다. 활성 RNA 종은 초기에 ~70 nt 전구체로 전사되며, 이는 전사 후 성숙한 ~21 nt 형태로 처리된다. *let-7*과 *lin-4* 양자 모두 헤어핀 RNA 전구체로 전사되고 이는 Dicer 효소에 의해 이들의 성숙한 형태로 처리된다.

[0309] NOI 외에도 벡터는 또한 siRNA, shRNA 또는 조절된 shRNA를 포함하거나 암호화할 수 있다(Dickins et al. (2005) Nature Genetics 37: 1289-1295, Silva et al. (2005) Nature Genetics 37:1281-1288).

[0310] 이중 가닥 RNA(dsRNA)에 의해 매개되는 전사 후 유전자 사일런싱(PGTS)은 외래 유전자의 발현을 제어하기 위한 보존된 세포 방어 메커니즘이다. 트랜스포존 또는 바이러스와 같은 요소의 무작위 통합은 상동 단일 가닥 mRNA 또는 바이러스 게놈 RNA의 서열 특이적 분해를 활성화하는 dsRNA의 발현을 유발하는 것으로 생각된다. 사일런싱 효과는 RNA 간섭(RNAi)으로 알려져 있다(Ralph et al. (2005) Nature Medicine 11:429-433). RNAi의 메커니즘은 긴 dsRNA를 프로세싱하여 약 21-25개 뉴클레오티드(nt) RNA의 이중체(duplex)가 되게하는 것을 포함한다. 이러한 산물을 작은 간섭 또는 사일런싱 RNA(siRNA)라고 하며 이는 mRNA 분해의 서열-특이적 매개체이다. 분화된 포유동물 세포에서, dsRNA >30 bp는 인터페론 반응을 활성화하여 단백질 합성 및 비-특이적 mRNA 분해를 중단(shut-down)시키는 것으로 밝혀졌다(Stark et al., Annu Rev Biochem 67:227-64 (1998)). 그러나 이 반응은 21

nt siRNA 이중체를 사용하여 우회할 수 있으며(Elbashir et al., EMBO J. Dec 3;20(23):6877-88 (2001), Hutvagner et al., Science. Aug 3, 293(5531):834-8. Eupub Jul 12 (2001)) 배양된 포유동물 세포에서 유전자 기능 분석을 가능하게 한다.

[0311] 약제학적 조성물

[0312] 본 개시내용은 본원에 기재된 바와 같은 렌티바이러스 벡터 또는 본원에 기재된 바와 같은 바이러스 벡터로 형질도입된 세포 또는 조직을 약제학적으로 허용되는 담체, 희석제 또는 부형제와 조합하여 포함하는 약제학적 조성물을 제공한다.

[0313] 본 개시내용은 치료적 유효량의 렌티바이러스 벡터를 포함하는, 유전자 요법에 의해 개체를 치료하기 위한 약제학적 조성물을 제공한다. 약제학적 조성물은 인간 또는 동물용일 수 있다.

[0314] 조성물은 약제학적으로 허용가능한 담체, 희석제, 부형제 또는 어주반트(adjuvant)를 포함할 수 있다. 약제학적 담체, 부형제 또는 희석제의 선택은 의도된 투여 경로 및 표준 약제학적 관행과 관련하여 이루어질 수 있다. 약제학적 조성물은 담체, 부형제 또는 희석제를 포함하거나, 또는 이에 더하여 임의의 적합한 결합제(들), 윤활제(들), 현탁제(들), 코팅제(들), 가용화제(들) 및 표적 부위의 벡터 진입을 돕거나 증가시킬 수 있는 기타 담체 제제(예컨대 예를 들어 지질 전달 시스템)를 포함할 수 있다.

[0315] 적절한 경우, 조성물은 흡입; 좌약 또는 pessary(의)의 형태로; 로션, 용액, 크림, 연고 또는 더스팅 분말 형태로 국소적으로; 피부 패치의 사용에 의해; 전분 또는 락토스와 같은 부형제를 함유하는 정제 형태로, 또는 단독으로 또는 부형제와 혼합하여 캡슐 또는 난자(ovules)로, 또는 향미제 또는 착색제를 함유하는 엘릭시르, 용액 또는 현탁액 형태로 경구적으로 중 임의의 하나 이상에 의해 투여될 수 있고; 또는 이들은 비경구적으로, 예를 들어 해면체내(intracavernosally), 정맥내, 근육내, 두개내(intracranially), 안구내, 복강내 또는 피하로 주사될 수 있다. 비경구 투여의 경우, 조성물은 멸균 수용액의 형태로 가장 잘 사용될 수 있으며 이는 다른 물질, 예를 들어 용액을 혈액과 등장성으로 만들기엔 충분한 염 또는 당당류를 함유할 수 있다. 협측(buccal) 또는 설하 투여의 경우, 조성물은 통상적인 방식으로 제형화될 수 있는 정제 또는 로젠지(lozenge)의 형태로 투여될 수 있다.

[0316] 본원에 기재된 바와 같은 렌티바이러스 벡터는 또한 생체외에서 표적 세포 또는 표적 조직을 형질도입시킨 후 상기 표적 세포 또는 조직을 이를 필요로 하는 환자에게 전달하는데 사용될 수 있다. 이러한 세포의 예는 자가(autologous) T 세포일 수 있고 이러한 조직의 예는 공여자 각막일 수 있다.

[0317] 변이체, 유도체, 유사체(analogue), 상동체(homologue) 및 단편

[0318] 본원에 언급된 특정 단백질 및 뉴클레오티드에 더하여, 본 발명은 또한 그의 변이체, 유도체, 유사체, 상동체 및 단편의 사용을 포함한다.

[0319] 본 발명의 맥락에서, 임의의 주어진 서열의 변이체란 해당 폴리펩티드 또는 폴리뉴클레오티드가 내인성 기능 중 적어도 하나를 유지하는 방식으로 잔기(아미노산 잔기이든 핵산 잔기이든)의 특정 서열이 변형된 서열이다. 변이체 서열은 천연 발생 단백질에 존재하는 적어도 하나의 잔기의 부가, 결실, 치환, 변형, 대체 및/또는 변이에 의해 수득될 수 있다.

[0320] 본 발명의 단백질 또는 폴리펩티드와 관련하여, 본원에서 사용되는 용어 "유도체"는 생성된 단백질 또는 폴리펩티드가 그의 내인성 기능 중 적어도 하나를 유지한다면, 서열로부터 또는 서열에 대하여 하나(또는 그 이상)의 아미노산 잔기의 임의의 치환, 변이, 변형, 대체, 결실 및/또는 부가를 포함한다.

[0321] 폴리펩티드 또는 폴리뉴클레오티드와 관련하여, 본원에서 사용되는 용어 "유사체"는 임의의 모방체(mimetic), 즉 모방하는 폴리펩티드 또는 폴리뉴클레오티드의 내인성 기능 중 적어도 하나를 보유하는 화학적 화합물을 포함한다.

[0322] 전형적으로, 아미노산 치환은 변형된 서열이 필요한 활성 또는 능력을 유지한다면, 예를 들어 1, 2 또는 3개에서 10 또는 20개까지의 치환으로 이루어질 수 있다. 아미노산 치환은 비-천연 발생 유사체의 사용을 포함할 수 있다.

[0323] 본 발명에서 사용되는 단백질은 침묵 변화를 생성하고 기능적으로 동등한 단백질을 생성하는 아미노산 잔기의 결실, 삽입 또는 치환을 또한 가질 수 있다. 의도적인 아미노산 치환은 내인성 기능이 유지되는 한, 잔기의 극성, 전하, 용해도, 소수성, 친수성 및/또는 양친매성 성질의 유사성에 기초하여 이루어질 수 있다. 예를 들어,

음전하를 띤 아미노산에는 아스파르트산 및 글루탐산이 포함되고; 양전하를 띤 아미노산에는 라이신 및 아르기닌이 포함되고; 유사한 친수성 값을 갖는 전하를 띠지 않는 극성 헤드 그룹을 가지는 아미노산에는 아스파라긴, 글루타민, 세린, 트레오닌 및 티로신이 포함된다.

- [0324] 보존적 치환이, 예를 들어 하기 표에 따라 이루어질 수 있다. 두 번째 컬럼의 동일한 블록, 바람직하게는 세 번째 컬럼의 동일한 라인에 있는 아미노산은 서로 치환될 수 있다:

지방족	비-극성	G A P
		I L V
	극성 - 전하를 띠지 않는	C S T M
		N Q
방향족	극성 - 전하를 띠는	D E
		K R H
		F W Y

[0325]

- [0326] "상동체(homologue)"라는 용어는 야생형 아미노산 서열 및 야생형 뉴클레오타이드 서열과 일정한 상동성을 가지는 엔터티를 의미한다. "상동성"이라는 용어는 "동일성"과 동일시될 수 있다.

- [0327] 본 발명의 맥락에서, 상동성 서열은 대상 서열과 적어도 50%, 55%, 65%, 75%, 85% 또는 90% 동일할 수 있는, 바람직하게는 적어도 95%, 97 또는 99% 동일할 수 있는 아미노산 서열을 포함하는 것으로 간주된다. 전형적으로, 상동체는 대상 아미노산 서열과 동일한 활성 부위 등을 포함할 것이다. 상동성은 유사성(즉, 유사한 화학적 특성/기능을 가지는 아미노산 잔기) 측면에서도 고려될 수 있지만, 본 발명의 맥락에서는 서열 동일성 측면에서 상동성을 표현하는 것이 바람직하다.

- [0328] 본 발명의 맥락에서, 상동성 서열은 대상 서열과 적어도 50%, 55%, 65%, 75%, 85% 또는 90% 동일할 수 있는, 바람직하게는 적어도 95%, 97%, 98% 또는 99% 동일할 수 있는 뉴클레오타이드 서열을 포함하는 것으로 간주된다. 상동성은 유사성 측면에서도 고려될 수 있지만, 본 발명의 맥락에서는 서열 동일성 측면에서 상동성을 표현하는 것이 바람직하다.

- [0329] 상동성 비교는 육안으로, 또는 보다 일반적으로 쉽게 이용 가능한 서열 비교 프로그램의 도움으로 수행할 수 있다. 이러한 상업적으로 이용 가능한 컴퓨터 프로그램은 2개 이상의 서열 간의 퍼센트 상동성 또는 동일성을 계산할 수 있다.

- [0330] 퍼센트 상동성은 연속적인 서열에 대해 계산될 수 있다. 즉, 하나의 서열이 다른 서열과 정렬되고 한 서열의 각 아미노산은 한 번에 한 잔기씩 다른 서열의 상응하는 아미노산과 직접 비교된다. 이것을 "ungapped(갭이 없는)" 정렬이라고 한다. 전형적으로, 이러한 갭이 없는 정렬은 상대적으로 적은 수의 잔기에 대해서만 수행된다.

- [0331] 이것은 매우 간단하고 일관된 방법이지만, 예를 들어, 그렇지 않으면 동일한 쌍의 서열에 있어서, 뉴클레오타이드 서열의 1개의 삽입 또는 결실로 인해 다음 코돈들을 정렬에서 벗어나게 하고, 따라서 전역(global) 정렬이 수행될 때 퍼센트 상동성이 크게 감소하는 결과를 잠재적으로 초래할 수 있다는 점을 고려하지 못한다. 결과적으로, 대부분의 서열 비교 방법은 전체 상동성 점수에 부당한 불이익을 주지 않으면서 가능한 삽입 및 결실을 고려하는 최적의 정렬을 생성하도록 설계된다. 이는 로컬 상동성을 최대화하기 위해 서열 정렬에 "갭"을 삽입함으로써 달성된다.

- [0332] 그러나, 이러한 더 복잡한 방법은 정렬에서 발생하는 각 갭에 "갭 패널티"를 할당함으로써, 동일한 아미노산이 동일한 개수인 경우, 2개의 비교된 서열 간의 더 높은 관련성을 반영하는 가능한 한 적은 갭을 가지는 서열 정렬이 많은 갭을 가지는 서열 정렬에 비하여 더 높은 점수를 얻게 한다. "아핀 갭 비용(Affine gap cost)"은 갭의 존재에 대해서는 상대적으로 높은 비용을 부과하고 갭 내의 각 후속 잔기에 대해서는 더 작은 패널티를 부과하는 데 전형적으로 사용된다. 이것은 가장 일반적으로 사용되는 갭 스코어링 시스템이다. 높은 갭 패널티는 당연히 더 적은 갭을 가지는 최적화된 정렬을 생성한다. 대부분의 정렬 프로그램은 갭 패널티를 수정하는 것을 허용한다. 그러나, 이러한 소프트웨어를 서열 비교에 사용할 때는 디폴트 값을 사용하는 것이 바람직하다. 예를 들어 GCG Wisconsin Bestfit 패키지를 사용하는 경우 아미노산 서열에 대한 디폴트 갭 패널티는 갭에 대해 -12이고 각 확장에 대해 -4이다.

- [0333] 따라서 최대 퍼센트 상동성을 계산하기 위해서는 먼저 갭 패널티를 고려하여 최적의 정렬을 생성하는 것이 필요하다. 이러한 정렬을 수행하는 데 적합한 컴퓨터 프로그램은 GCG Wisconsin Bestfit 패키지이다(University of Wisconsin, U.S.A.; Devereux et al. (1984) Nucleic Acids Research 12:387). 서열 비교를 수행할 수 있는 다른 소프트웨어의 예에는 BLAST 패키지(Ausubel et al. (1999) ibid - Ch. 18 참조), FASTA(Atschul et al. (1990) J. Mol. Biol. 403-410) 및 GENEWORKS 비교 도구 제품군(suite)이 포함되지만 이에 국한되지는 않는다. BLAST 및 FASTA는 양자 모두 오프라인 및 온라인 검색에 사용할 수 있다(Ausubel et al. (1999) ibid, 페이지 7-58 내지 7-60 참조). 그러나, 일부 응용의 경우, GCG Bestfit 프로그램을 사용하는 것이 바람직하다. BLAST 2 Sequences라고 하는 또 다른 도구는 단백질과 뉴클레오타이드 서열을 비교하는 데도 사용할 수 있다(FEMS Microbiol Lett (1999) 174(2):247-50; FEMS Microbiol Lett (1999) 177(1):187-8 참조).
- [0334] 최종 퍼센트 상동성은 동일성의 측면에서 측정할 수 있지만, 정렬 프로세스 자체는 전형적으로 all-or-nothing 쌍 비교를 기반으로 하지 않는다. 대신, 화학적 유사성 또는 진화적 거리를 기반으로 각 쌍별 비교에 점수를 할당하는 척도화된 유사성 점수 매트릭스(scaled similarity score matrix)가 일반적으로 사용된다. 일반적으로 사용되는 이러한 매트릭스의 예는 BLOSUM62 매트릭스로, 이는 BLAST 프로그램 제품군의 디폴트 매트릭스이다. GCG Wisconsin 프로그램은 일반적으로 공개 디폴트 값을 사용하거나 또는 제공되는 경우 사용자 지정 기호 비교 테이블을 사용한다(자세한 내용은 사용자 설명서 참조). 일부 응용의 경우, GCG 패키지에 대해 공개 디폴트 값을 사용하거나, 또는 다른 소프트웨어의 경우, BLOSUM62와 같은 디폴트 매트릭스를 사용하는 것이 바람직하다.
- [0335] 소프트웨어가 최적의 정렬을 생성하면, 퍼센트 상동성, 바람직하게는 퍼센트 서열 동일성을 계산할 수 있다. 소프트웨어는 전형적으로 서열 비교의 일부로 이 작업을 수행하고 수치 결과를 생성한다.
- [0336] "단편" 또한 변이체이고 이 용어는 전형적으로 기능적으로 관심있거나 또는 예를 들어 아세이에서 관심 있는 폴리펩티드 또는 폴리뉴클레오타이드의 선택된 영역을 지칭한다. 따라서 "단편"은 전장 폴리펩티드 또는 폴리뉴클레오타이드의 일부인 아미노산 또는 핵산 서열을 지칭한다.
- [0337] 이러한 변이체는 부위-지정 돌연변이유발과 같은 표준 재조합 DNA 기술을 사용하여 제조할 수 있다. 삽입을 수행하고자 하는 경우, 삽입 부위의 어느 한 쪽의 자연 발생 서열에 상응하는 5' 및 3' 플랭킹 영역과 함께 삽입을 암호화하는 합성 DNA를 만들 수 있다. 플랭킹 영역은 자연 발생 서열의 부위에 상응하는 편리한 제한 부위를 포함함으로써 서열이 적절한 효소(들)로 절단되고 합성 DNA가 절단부에 결합될 수 있다. 그런 다음 DNA는 암호화된 단백질을 만들기 위해 본 발명에 따라 발현된다. 이들 방법은 DNA 서열의 조작을 위해 당업계에 공지된 수많은 표준 기술의 예시일 뿐이며 다른 공지된 기술도 사용될 수 있다.
- [0338] 본 발명의 세포 및/또는 모듈식 작제물에서 사용하기에 적합한 조절 단백질의 모든 변이체, 단편 또는 상동체는 NOI의 동족 결합 부위에 결합하는 능력을 유지하여 NOI의 번역이 바이러스 벡터 생산 세포에서 억제되거나 방지되게 할 것이다.
- [0339] 결합 부위의 모든 변이체, 단편 또는 상동체는 동족 RNA-결합 단백질에 결합하는 능력을 유지하여, NOI의 번역이 바이러스 벡터 생산 세포에서 억제되거나 방지되게 할 것이다.
- [0340] **코돈 최적화**
- [0341] 본 발명에서 사용되는 폴리뉴클레오타이드(NOI 및/또는 벡터 생산 시스템의 구성요소 포함)는 코돈 최적화될 수 있다. 코돈 최적화는 WO 1999/41397 및 WO 2001/79518에서 이전에 설명되었다. 서로 다른 세포는 특정 코돈의 사용법이 다르다. 이러한 코돈 편향(bias)은 세포 유형에서 특정 tRNA의 상대적 풍부도의 편향에 대응한다. 서열의 코돈을 변경하여 상응하는 tRNA의 상대적 풍부도와 일치하도록 맞춤화함으로써, 발현을 증가시키는 것이 가능하다. 같은 이유로, 상응하는 tRNA가 특정 세포 유형에서 드문 것으로 알려진 코돈을 의도적으로 선택하여 발현을 감소시키는 것이 가능하다. 따라서, 추가 수준의 번역 제어를 사용할 수 있다.
- [0342] 레트로바이러스를 포함한 많은 바이러스는 많은 수의 희귀 코돈을 사용하며 일반적으로 사용되는 포유류 코돈에 해당하도록 이들을 변경하여 포유동물 생산 세포에서 관심 유전자, 예를 들어, NOI 또는 패키징 구성요소의 증가된 발현을 달성할 수 있다. 코돈 사용법 표는 포유동물 세포 뿐만 아니라 다양한 기타 유기체에 대해 당업계에 공지되어 있다.
- [0343] 바이러스 벡터 구성요소의 코돈 최적화는 여러 가지 다른 이점을 가진다. 그들의 서열의 변경으로 인해, 생산자 세포/패키징 세포에서 바이러스 입자의 어셈블리에 필요한 바이러스 입자의 패키징 구성요소를 암호화하는 뉴클레오타이드 서열은 RNA 불안정성 서열(INS)이 이들로부터 제거된다. 동시에, 패키징 구성요소에 대한 서열을 코딩

하는 아미노산 서열은 서열에 의해 암호화된 바이러스 구성요소가 동일하게 유지되거나 또는 패키징 구성요소의 기능이 손상되지 않도록 적어도 충분히 유사하도록 유지된다. 렌티바이러스 벡터에서 코돈 최적화는 내보내기 (export)에 대한 Rev/RRE 요구사항을 극복하여, 최적화된 서열이 Rev-독립적이 되게 한다. 코돈 최적화는 또한 벡터 시스템 내에서 서로 다른 작제물 간의(예를 들어 *gag-pol* 및 *env* 오픈 리딩 프레임에서 중첩되는 영역 간의) 상동 재조합을 감소시킨다. 따라서 코돈 최적화의 전반적인 효과는 바이러스 역가의 현저한 증가와 개선된 안전성이다.

[0344] 일 구현예에서 INS와 관련된 코돈만이 코돈 최적화된다. 그러나, 훨씬 더 바람직하고 실용적인 구현예에서는, 서열은 전체적으로 코돈 최적화되며, 일부 예외, 예를 들어 *gag-pol*의 프레임시프트 부위를 포함하는 서열은 제외한다(하기 참조).

[0345] 렌티바이러스 벡터의 *gag-pol* 유전자는 *gag-pol* 단백질을 암호화하는 2개의 중첩되는 리딩 프레임을 포함한다. 두 단백질의 발현은 번역 동안 프레임시프트에 의존한다. 이 프레임시프트는 번역 동안 리보솜 "미끄러짐 (slippage)"의 결과로 발생한다. 이 미끄러짐은 적어도 부분적으로는 리보솜을 지연시키는 RNA 2차 구조에 의해 발생하는 것으로 생각된다. 이러한 2차 구조는 *gag-pol* 유전자에서 프레임시프트 부위의 하류에 존재한다. HIV의 경우, 중첩 영역은 *gag*의 시작부분의 하류의 뉴클레오티드 1222(여기서 뉴클레오티드 1은 *gag* ATG의 A이다)에서부터 *gag*의 끝(nt 1503)까지 확장된다. 결과적으로, 프레임시프트 부위와 2개의 리딩 프레임의 중첩 영역에 걸쳐 있는 281 bp 단편은 바람직하게는 코돈 최적화되지 않는다. 이 단편을 유지하면 Gag-Pol 단백질을 보다 효율적으로 발현할 수 있다. EIAV의 경우 중첩의 시작부분은 nt 1262(여기서 뉴클레오티드 1은 *gag* ATG의 A이다)이고 중첩의 끝 부분은 nt 1461로 간주되었다. 프레임시프트 부위와 *gag-pol* 중첩이 보존되는 것을 보장하기 위해, 야생형 서열은 nt 1156에서 1465까지 유지되었다.

[0346] 예를 들어, 편리한 제한 부위를 수용하기 위해, 최적의 코돈 사용법으로부터 파생(derivation)이 적용될 수 있으며, 보존적 아미노산 변화가 Gag-Pol 단백질에 도입될 수 있다.

[0347] 일 구현예에서, 코돈 최적화는 약하게 발현된 포유동물 유전자에 기초한다. 세 번째 및 때로는 두 번째 및 세 번째 염기가 변경될 수 있다.

[0348] 유전자 코드의 축퇴 특성으로 인해, 다수의 *gag-pol* 서열이 숙련된 작업자에 의해 달성될 수 있음이 이해될 것이다. 또한 코돈-최적화된 *gag-pol* 서열을 생성하기 위한 시작점으로 사용할 수 있는 많은 레트로바이러스 변이체가 설명되어 있다. 렌티바이러스 계통은 매우 다양할 수 있다. 예를 들어, 여전히 기능하는 많은 유사 종(quasi-species)의 HIV-1이 있다. 이는 EIAV의 경우에도 마찬가지다. 이러한 변이체는 형질도입 과정의 특정 부분을 향상시키는 데 사용될 수 있다. HIV-1 변이체의 예는 <http://hiv-web.lanl.gov> 에서 Los Alamos National Security, LLC에서 운영하는 HIV 데이터베이스에서 찾을 수 있다. EIAV 클론의 세부사항은 <http://www.ncbi.nlm.nih.gov>에 있는 National Center for Biotechnology Information (NCBI) 데이터베이스에서 찾을 수 있다.

[0349] 코돈-최적화된 *gag-pol* 서열에 대한 전략은 모든 레트로바이러스와 관련하여 사용될 수 있다. 이는 EIAV, FIV, BIV, CAEV, VMR, SIV, HIV-1 및 HIV-2를 포함한 모든 렌티바이러스에 적용된다. 또한 이 방법은 HTLV-1, HTLV-2, HFV, HSRV 및 인간 내인성 레트로바이러스(HERV), MLV 및 기타 레트로바이러스로부터 유전자 발현을 증가시키는 데 사용될 수 있다.

[0350] 코돈 최적화는 *gag-pol* 발현을 Rev-독립적으로 만들 수 있다. 그러나, 렌티바이러스 벡터에서 항-*rev* 또는 RRE 인자의 사용을 가능하게 하기 위해서는, 바이러스 벡터 생성 시스템을 완전히 Rev/RRE-독립적으로 만드는 것이 필요할 것이다. 따라서, 계통 또한 변형되어야 한다. 이것은 벡터 계통 구성요소를 최적화함으로써 달성된다. 유리하게는, 이러한 변형은 또한 생산자 및 형질도입된 세포 모두에서 모든 추가적인 단백질이 없는 보다 안전한 시스템의 생산으로 이어진다.

[0351] **벡터 역가를 개선하기 위한 변형된 U1과의 조합**

[0352] MSD-돌연변이된 렌티바이러스 벡터는 LV 생산 동안 및 표적 세포 모두에서 비정상적인 스플라이싱 이벤트에 참여하는 능력이 감소하기 때문에 유전자 치료 벡터로 사용하기에 현재의 표준 렌티바이러스 벡터보다 바람직하다. 그러나, 본 발명이 있기까지, MSD-돌연변이된 벡터의 생산은 HIV-1 *tat* 단백질의 공급에 의존하거나(1세대 및 2세대 U3-의존적 렌티바이러스 벡터) 또는 벡터 RNA 수준에 대한 MSD 돌연변이의 불안정한 효과로 인해 효율성이 낮았다(3세대 벡터에서). 안전상의 이유로 *tat*를 현대의 3세대 U3-독립적 LV 시스템에 다시 '재도입'하려는 요구나 정당화가 없으며, 결과적으로 현재 임상 사용을 위한 MSD-돌연변이된 벡터의 생산 역가의

감소에 대한 해결책이 없다.

- [0353] 본 발명자들은 MSD-돌연변이된 3세대(즉, U3/tat-독립적인) LV가 생산 동안 벡터 게놈 RNA의 5' 패키징 영역에 결합하도록 지시된 변형된 U1 snRNA의 공동-발현에 의해 높은 역가로 생산될 수 있음을 보여준다. 놀랍게도 이러한 변형된 U1 snRNA가 5'R 영역 내의 5'polyA 신호의 존재와 무관한 방식으로 MSD-돌연변이된 LV의 생산 역가를 향상시킬 수 있다는 것이 밝혀졌으며, 이는 다른 사람들이 폴리아데닐화를 억제하기 위해 변형된 U1 snRNA 사용하는 것(소위 U1 간섭, [Ui])에 비해 새로운 메커니즘을 나타낸다. 놀랍게도 변형된 U1 snRNA를 패키징 영역의 중요한 서열에 표적화하는 것이 MSD-돌연변이된 LV 역가에서 가장 큰 향상을 생성한다는 것이 밝혀졌다. 본 발명자들은 또한 MSD-돌연변이된 LV의 역가의 감소를 덜 뚜렷하게 하고, 변형된 U1 snRNA에 의한 이러한 MSD-돌연변이된 LV 변이체의 역가의 증가가 최대가 되게 하는, 주요 스플라이스 공여자 영역 내의 신규한 서열 돌연변이를 개시한다.
- [0354] 본 발명자들은 놀랍게도 더 이상 내인성 서열(스플라이스 공여자 부위)을 표적으로 하지 않고 이제 vRNA 분자 내의 서열을 표적화하도록 변형된 U1 snRNA를 기반으로 하는 비-코딩 RNA를 공동-발현시킴으로써 렌티바이러스 벡터의 출력(output) 역가가 향상될 수 있다는 것을 발견하였다. 본 실시예에서 입증되는 바와 같이, 본 발명자들은 상기 변형된 U1 snRNA에 의한 주요 스플라이스 공여자 영역(주요 스플라이스 공여자 및 잠복 스플라이스 공여자 부위를 함유함) 내에 약독화 돌연변이를 보유하는 렌티바이러스 벡터의 출력 역가의 상대적인 향상이 돌연변이되지 않은 주요 스플라이스 공여자 영역을 포함하는 표준 렌티바이러스 벡터보다 크다는 것을 보여준다.
- [0355] 본 실시예에서 입증되는 바와 같이, 감소된 역가를 유도하는 주요 스플라이스 공여자 영역 내에 광범위한 돌연변이 유형(점 돌연변이, 영역 결실 및 서열 교체)을 보유하는 벡터 게놈은 변형된 U1 snRNA와 조합하여 사용될 수 있다. 이 접근법은 벡터 생산 동안 다른 벡터 구성요소와 함께 변형된 U1 snRNA의 공동-발현을 포함할 수 있다. 변형된 U1 snRNA는 벡터 게놈 vRNA 내의 표적 서열에 상보적인 이중 서열로 대체함으로써 컨센서스 스플라이스 공여자 부위에 대한 결합이 제거되도록 설계된다. 본 발명은 표적 서열 및 상보성 길이, 설계 및 발현 방식을 포함하는 변형된 U1 snRNA의 다양한 적용 방식 및 최적 특성을 설명한다.
- [0356] 본원에 기재된 바와 같은 본 발명의 일 측면에서, 벡터는 변형된 U1 snRNA와 조합하여 사용될 수 있다. 이것은 아래에서 더 논의된다.
- [0357] 변형된 U1 snRNA 분자의 존재/풍부도는 전체 RNA를 추출한 후 DNA 프라이머를 사용한 RT-PCR 또는 RT-qPCR(정량적)을 통해 벡터 생산 세포 추출물 또는 벡터 비리온 내에서 정량화할 수 있다. 중요하게는, 정방향 프라이머는 변형된 U1 snRNA 분자의 표적화 서열에 상보성을 가지도록 설계되어 qPCR 동안 변형된 U1 snRNA만 증폭되고 내인성 U1 snRNA는 그렇지 아니하다.
- [0358] 일 측면에서 본 발명은 본원에 기재된 바와 같은 본 발명에 따른 변형된 U1을 포함하는 벡터 비리온을 제공한다.
- [0359] 스플라이싱 및 폴리아데닐화는 특히 대부분의 단백질 코딩 전사체가 다중 인트론을 포함하는 고등 진핵생물에서 mRNA 성숙을 위한 핵심 프로세스이다. 스플라이싱에 필요한 pre-mRNA 내의 요소로는 5' 스플라이스 공여자 신호, 분기점을 둘러싼 서열 및 3' 스플라이스 수용체 신호가 포함된다. 이 세 가지 요소와 상호 작용하는 스플라이세오솜은 U1 snRNA를 포함한 5개의 작은 핵 RNA(snRNA) 및 관련 핵 단백질(snRNP)에 의해 형성된다. U1 snRNA는 폴리머라제 II 프로모터에 의해 발현되며 대부분의 진핵 세포에 존재한다(Lund *et al.*, 1984, *J. Biol. Chem.*, 259:2013-2021). 인간 U1 snRNA(작은 핵 RNA)는 4개의 스템-루프로 이루어진 잘 정의된 구조를 가지는 164 nt 길이이다(West, S., 2012, *Biochemical Society Transactions*, 40:846-849). U1 snRNA는 엑손-인트론 접합부에서 5' 스플라이스 공여자 부위에 광범위하게 상보적인 짧은 서열을 그의 5'-말단에 포함한다. U1 snRNA는 5' 스플라이스 공여자 부위에 염기쌍을 형성하여 스플라이스-부위 선택 및 스플라이세오솜 어셈블리에 참여한다. 스플라이싱 외의 U1 snRNA에 대한 알려진 기능은 3'-말단 mRNA 처리 조절에 있다; 이는 초기 polyA 신호에서(특히 인트론 내의) 조기(premature) 폴리아데닐화(polyA)를 억제한다.
- [0360] 인간 U1 snRNA(작은 핵 RNA)는 4개의 스템-루프로 이루어진 잘 정의된 구조를 가지는 164 nt 길이이다(도 1 참조). 내인성 비-코딩 RNA인 U1 snRNA는 인트론 스플라이싱의 초기 단계에서 천연 스플라이스 공여자 어닐링 서열(예: 5'-ACUUAACUG-3')을 통해 컨센서스 5' 스플라이스 공여자 부위(예: 5'-MAGGURR-3' 여기서 M은 A 또는 C이고, R은 A 또는 G이다)에 결합한다. 스템 루프 I은 polyA 억제에 중요한 것으로 밝혀진 U1A-70K 단백질에 결합한다. 스템 루프 II는 U1A 단백질에 결합하고, 5'-AUUUGUGG-3' 서열은 Sm 단백질에 결합하는데, 이는 스템 루프 IV와 함께 U1 snRNA 프로세싱에 중요하다. 본원에 정의된 바와 같이, 본 발명에 따라 사용하기 위한 변형된

U1 snRNA는 천연 스플라이스 공여자 표적화/어닐링 서열의 부위에서 벡터 게놈 vRNA 분자 내의 표적 서열에 상보적인 이중 서열을 도입하도록 변형된다(도 1 참조).

- [0361] 본원에서 사용되는 바와 같이, 용어 "변형된 U1 snRNA", "재지정된(re-directed) U1 snRNA", "재표적화된 U1 snRNA", "재목적화된(re-purposed) U1 snRNA" 및 "돌연변이 U1 snRNA"는 표적 유전자의 스플라이싱 과정을 개시하기 위해 사용하는 컨센서스 5' 스플라이스 공여자 부위 서열(예: 5'-MAGGURR-3')에 더 이상 상보적이지 않도록 변형된 U1 snRNA를 의미한다. 따라서, 변형된 U1 snRNA는 스플라이스 공여자 부위 서열(예: 5'-MAGGURR-3')에 더 이상 상보적이지 않도록 변형된 U1 snRNA이다. 대신에, 변형된 U1 snRNA는 MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 패키징 영역(표적 부위) 내의 고유한 RNA 서열을 가지는 뉴클레오티드 서열, 즉 vRNA의 스플라이싱과 관련이 없는 서열을 표적으로 하거나 이에 상보적하도록 설계된다. MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 패키징 영역 내의 뉴클레오티드 서열은 미리 선택될 수 있다. 따라서, 변형된 U1 snRNA는 5' 말단이 MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 패키징 영역 내의 뉴클레오티드 서열에 상보적하도록 변형된 U1 snRNA이다. 그 결과, 이론에 얽매이지 않아서, 변형된 U1 snRNA는 변형된 U1 snRNA의 5' 말단에 있는 짧은 서열과 표적 부위 서열의 상보성에 기초하여 표적 부위 서열에 결합하고, 따라서 vRNA 안정화시켜 MSD-돌연변이된 렌티바이러스 벡터의 출력 벡터 역할을 증가시키는 것으로 생각된다.
- [0362] 본원에서 사용되는 바와 같이, "천연 스플라이스 공여자 어닐링 서열" 및 "천연 스플라이스 공여자 표적화 서열"이라는 용어는 인트론의 컨센서스 5' 스플라이스 공여자 부위에 대해 광범위하게 상보적인 내인성 U1 snRNA의 5'-말단에 있는 짧은 서열을 의미한다. 천연 스플라이스 공여자 어닐링 서열은 5'-ACUUACCUG-3'일 수 있다.
- [0363] 본원에서 사용되는 바와 같이, 용어 "컨센서스 5' 스플라이스 공여자 부위"는 스플라이스 부위 선택에 사용되는 인트론의 5' 말단에 있는 컨센서스 RNA 서열을 의미하며, 예를 들어 서열 5'-MAGGURR-3'을 가진다.
- [0364] 본원에서 사용되는 바와 같이, "MSD-돌연변이된 렌티바이러스 벡터 게놈 서열의 패키징 영역 내의 뉴클레오티드 서열", "표적 서열" 및 "표적 부위"라는 용어는 MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 패키징 영역 내에 특정 RNA 서열을 가지는 부위를 의미하며 이는 변형된 U1 snRNA에 결합/어닐링하기 위한 표적 부위로 사전 선택된다.
- [0365] 본원에서 사용되는 바와 같이, "MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 패키징 영역" 및 "MSD-돌연변이된 렌티바이러스 벡터 게놈 서열의 패키징 영역"이라는 용어는 5' U5 도메인의 시작부분에서부터 *gag* 유전자로부터 유래된 서열의 말단까지의 MSD-돌연변이된 렌티바이러스 벡터 게놈의 5' 말단에 있는 영역을 의미한다. 따라서, MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 패키징 영역은 5' U5 도메인, PBS 요소, 스템 루프(SL) 1 요소, SL2 요소, SL3 ψ 요소, SL4 요소 및 *gag* 유전자로부터 유래된 서열을 포함한다. 복제-결합 바이러스 벡터 입자의 생산을 가능하게 하기 위해 렌티바이러스 벡터 생산 동안 게놈에 트랜스로 완전한 *gag* 유전자를 제공하는 것은 당업계에서 일반적이다. 트랜스로 제공되는 *gag* 유전자의 뉴클레오티드 서열은 야생형 뉴클레오티드에 의해 암호화될 필요는 없지만 코돈 최적화될 수 있으며; 중요하게도 트랜스로 제공되는 *gag* 유전자의 주요 속성은 그것이 *gag* 및 *gagpol* 단백질의 발현을 암호화하고 지시한다는 것이다. 따라서, 완전한 *gag* 유전자가 렌티바이러스 벡터 생산 동안 트랜스로 제공되는 경우, "렌티바이러스 벡터 게놈 분자의 패키징 영역"이라는 용어는 5' U5 도메인의 시작부분에서부터 SL3 ψ 요소에 있는 '코어' 패키징 신호까지의 MSD-돌연변이된 렌티바이러스 벡터 게놈 분자의 5' 말단에 있는 영역, 그리고 ATG 코돈(SL4 내에 존재)에서부터 벡터 게놈에 존재하는 나머지 *gag* 뉴클레오티드 서열의 말단까지의 천연 *gag* 뉴클레오티드 서열을 의미할 수 있음을 당업자는 이해할 것이다.
- [0366] 본원에서 사용되는 바와 같이, 용어 "*gag* 유전자로부터 유래된 서열"은 벡터 게놈에 존재할 수 있는, 예를 들어 남아있는, ATG 코돈에서부터 뉴클레오티드 688까지로부터 유래된 *gag* 유전자의 임의의 천연 서열을 의미한다 (Kharytonchyk, S. *et. al.*, 2018, *J. Mol. Biol.*, 430:2066-79).
- [0367] 본원에서 사용되는 바와 같이, "천연 스플라이스 공여자 어닐링 서열을 포함하는 U1 snRNA의 처음 11개 뉴클레오티드 내에 이중 서열을 도입한다", "위치 3 내지 11에서 9개의 뉴클레오티드 내에 상기 이중 서열을 도입한다" 및 "U1 snRNA의 5' 말단에 있는 처음 11개의 뉴클레오티드 내에 이중 서열을 도입한다"라는 용어는 U1 snRNA의 처음 11개의 뉴클레오티드 또는 위치 3 내지 11에서 9개의 뉴클레오티드의 전부 또는 일부를 상기 이중 서열로 치환하거나 U1 snRNA의 처음 11개의 뉴클레오티드 또는 위치 3 내지 11에서 9개의 뉴클레오티드를 상기 이중 서열과 동일한 서열을 가지도록 변형시키는 것을 포함한다.
- [0368] 본원에서 사용되는 바와 같이, "천연 스플라이스 공여자 어닐링 서열 내에 이중 서열을 도입한다" 및 "U1 snRNA

의 5' 말단의 천연 스플라이스 공여자 어닐링 서열 내에 이중 서열을 도입한다"라는 용어는 천연 스플라이스 공여자 어닐링 서열의 전부 또는 일부를 상기 이중 서열로 치환하거나 천연 스플라이스 공여자 어닐링 서열을 상기 이중 서열과 동일한 서열을 가지도록 변형시키는 것을 포함한다.

- [0369] 본원에서 사용되는 용어 "렌티바이러스 벡터 역가를 향상시킨다"는 "렌티바이러스 벡터 역가를 증가시킨다", "렌티바이러스 벡터 역가를 회복한다" 및 "렌티바이러스 벡터 역가를 개선한다"를 포함한다.
- [0370] 따라서, 일 구현예에서, 변형된 U1 snRNA는 MSD-돌연변이된 렌티바이러스 벡터 게놈 서열의 패키징 영역 내의 뉴클레오티드 서열에 결합하도록 변형된다.
- [0371] 일부 구현예에서, 변형된 U1 snRNA는 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열을 도입하도록 내인성 U1 snRNA에 비해 5' 말단에서 변형된다.
- [0372] 일부 구현예에서, 변형된 U1 snRNA는 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열을 천연 스플라이스 공여자 어닐링 서열 내에 도입하도록 내인성 U1 snRNA에 비해 5' 말단에서 변형된다.
- [0373] 변형된 U1 snRNA는 천연 스플라이스 공여자 어닐링 서열을 포함하는 서열을 상기 뉴클레오티드 서열에 상보적인 이중 서열로 치환하도록 내인성 U1 snRNA에 비해 5' 말단에서 변형될 수 있다.
- [0374] 변형된 U1 snRNA는 변형된 U1 snRNA 변이체일 수 있다. 본 발명에 따라 변형된 U1 snRNA 변이체는 천연 발생 U1 snRNA 변이체, U1-70K 단백질 결합을 제거하는 스템 루프 I 영역 내에 돌연변이를 함유하는 U1 snRNA 변이체, 또는 U1A 단백질 결합을 제거하는 스템 루프 II 영역 내 돌연변이를 함유하는 U1 snRNA 변이체일 수 있다. U1-70K 단백질 결합을 제거하는 스템 루프 I 영역 내에 돌연변이를 함유하는 U1 snRNA 변이체는 U1_m1 또는 U1_m2, 바람직하게는 U1A_m1 또는 U1A_m2일 수 있다.
- [0375] 일부 구현예에서, 본원에 기재된 바와 같은 변형된 U1 snRNA는 본원에 기재된 바와 같은 U1_256 서열의 주요 U1 snRNA 서열 [클로버 잎](nt 410-562)과 적어도 70% 동일성(적합하게는 적어도 75%, 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 동일성)을 가지는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, 본 발명의 변형된 U1 snRNA는 본원에 기재된 바와 같은 U1_256 서열의 주요 U1 snRNA 서열 [클로버 잎](nt 410-562)을 포함한다. U1_256 서열(서열번호 15)의 주요 U1 snRNA 서열 [클로버 잎](nt 410-562)은 다음과 같다:
- [0376] GCAGGGGAGATACCATGATCACGAAGGTGGTTTCCAGGGCGAGGCTTATCCATTGCACTCCGGATGTGCTGACCCCTGCGATTTCCTCCCAATGTGGGAAACCTCGACTGCATAATTGTGGTAGTGGGGGACTGCGTTCGCGCTTTCCCTG. (서열번호 66)
- [0377] 일부 바람직한 구현예에서, 천연 스플라이스 공여자 어닐링 서열을 포함하는 U1 snRNA의 처음 11개 뉴클레오티드는 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열로 전부 또는 부분적으로 치환될 수 있다. 적합하게는, U1 snRNA의 처음 11개 뉴클레오티드의 1-11(적합하게는 2-11, 3-11, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 11) 핵산이 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열로 치환된다.
- [0378] 일부 구현예에서, 천연 스플라이스 공여자 어닐링 서열은 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열로 전부 또는 부분적으로 치환될 수 있다. 적합하게는, 천연 스플라이스 공여자 어닐링 서열의 1-11(적합하게는 2-11, 3-11, 5-11, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 11) 핵산이 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열로 치환된다. 바람직한 구현예에서, 전체 천연 스플라이스 공여자 어닐링 서열이 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열로 치환되며, 즉 천연 스플라이스 공여자 어닐링 서열(예를 들어, 5'-ACUUACCU-3')이 본원에 기재된 바와 같은 이중 서열로 전부 치환된다.
- [0379] 일부 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열은 상기 뉴클레오티드 서열에 대해 상보성인 적어도 7개의 뉴클레오티드를 포함한다. 일부 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오티드 서열에 상보적인 이중 서열은 상기 뉴클레오티드 서열에 대해 상보성인 적어도 9개의 뉴클레오티드를 포함한다. 바람직하게는, 본 발명에서 사용하기 위한 이중 서열은 상기 뉴클레오티드 서열에 상보성인 15개의 뉴클레오티드를 포함한다.
- [0380] 적합하게는, 본 발명에서 사용하기 위한 이중 서열은 7-25(적합하게는 7-20, 7-15, 9-15, 7, 8, 9, 10, 11,

12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25)개의 뉴클레오타이드를 포함할 수 있다.

[0381] 적합하게는, 본 발명에서 사용하기 위한 이중 서열은 25개의 뉴클레오타이드를 포함할 수 있다.

[0382] 일부 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 5' U5 도메인, PBS 요소, SL1 요소, SL2 요소, SL3 ψ 요소, SL4 요소 및/또는 *gag* 유전자로부터 유래된 서열 내에 위치한다. 적합하게는, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 SL1, SL2 및/또는 SL3 ψ 요소(들) 내에 위치한다. 일부 바람직한 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 SL1 및/또는 SL2 요소(들) 내에 위치한다. 일부 특히 바람직한 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 SL1 요소 내에 위치한다.

[0383] 일부 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 적어도 7개의 뉴클레오타이드를 포함한다. 일부 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 적어도 9개의 뉴클레오타이드를 포함한다. 적합하게는, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 7-25(적합하게는 7-20, 7-15, 9-15, 7, 8, 9, 10, 11, 12, 13, 14 또는 15)개의 뉴클레오타이드를 포함한다.

[0384] 적합하게는, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열은 15개의 뉴클레오타이드를 포함한다.

[0385] MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열에 대한 본원에 기재된 바와 같은 변형된 U1 snRNA의 결합은 본원에 기재된 바와 같은 변형된 U1 snRNA의 부재 하에서의 렌티바이러스 벡터 생산에 비해 렌티바이러스 벡터 생산 동안 렌티바이러스 벡터 역가를 향상시킬 수 있다. 따라서, 본원에 기재된 바와 같은 변형된 U1 snRNA의 존재 하에서의 렌티바이러스 벡터의 생산은 본원에 기재된 바와 같은 변형된 U1 snRNA의 부재 하에서의 렌티바이러스 벡터 생산에 비해 렌티바이러스 벡터 역가를 향상시킨다. 렌티바이러스 벡터 역가의 측정을 위한 적합한 아세이는 본원에 기재된 바와 같다. 적합하게는, 렌티바이러스 벡터 생산은 *gag*, *env*, *rev* 및 렌티바이러스 벡터의 RNA 게놈을 포함하는 벡터 구성요소와 상기 변형된 U1 snRNA의 공동-발현을 포함한다. 렌티바이러스 벡터의 RNA 게놈은 MSD-2KO RNA 게놈일 수 있다. 일부 구현예에서, 렌티바이러스 벡터 역가의 향상은 기능적 5'LTR polyA 부위의 존재 또는 부재 하에 발생한다. 일부 구현예에서, 본 발명의 변형된 U1 snRNA에 의해 매개되는 렌티바이러스 벡터 역가의 향상은 벡터 게놈의 5'LTR에서 polyA 부위 억제와 무관하다.

[0386] 일부 구현예에서, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열에 대한 본원에 기재된 바와 같은 변형된 U1 snRNA의 결합은 본원에 기재된 바와 같은 변형된 U1 snRNA의 부재 하에서의 렌티바이러스 벡터 생산에 비해 렌티바이러스 벡터 생산 동안 렌티바이러스 벡터 역가를 적어도 30% 증가시킬 수 있다. 적합하게는, MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내의 뉴클레오타이드 서열에 대한 본원에 기재된 바와 같은 변형된 U1 snRNA의 결합은 본원에 기재된 바와 같은 변형된 U1 snRNA의 부재 하에서의 MSD-돌연변이된 렌티바이러스 벡터 생산에 비해 생산 동안 MSD-돌연변이된 렌티바이러스 벡터 역가를 적어도 35%(적합하게는 적어도 40%, 45%, 50%, 60%, 70%, 100%, 150%, 200%, 250%, 300%, 350%, 400%, 450%, 500%, 550%, 600%, 650%, 700%, 750%, 800%, 850%, 900%, 950%, 1,000%, 2,000%, 5,000%, 또는 10,000%) 증가시킬 수 있다.

[0387] 본원에 기재된 바와 같은 변형된 U1 snRNA는 (a) 변형된 U1 snRNA에 결합을 위한 MSD-돌연변이된 렌티바이러스 벡터 게놈의 패키징 영역 내 표적 부위(미리 선택된 뉴클레오타이드 부위)를 선택하는 단계; 및 (b) 단계 (a)에서 선택된 미리 선택된 뉴클레오타이드 부위에 상보적인 이중 서열을 U1 snRNA의 5' 말단에 있는 천연 스플라이스 공여자 어닐링 서열(예: 5'-ACUUACCUG-3') 내에 도입하는 단계에 의해 설계될 수 있다.

[0388] 분자 생물학의 통상적인 기술을 사용하여 내인성 U1 snRNA의 5' 말단에 있는 천연 스플라이스 공여자 어닐링 서열(예: 5'-ACUUACCUG-3') 내에, 또는 그 대신에, 표적 부위에 상보적인 이중 서열을 도입하는 것은 당업자의 능력 범위 내에 있다. 일반적으로 말해서, 적절한 일상적인 방법은 지시된(directed) 돌연변이유발 또는 상동 재조합을 통한 치환을 포함한다.

[0389] 분자 생물학의 통상적인 기술을 사용하여 내인성 U1 snRNA의 5' 말단에 있는 천연 스플라이스 공여자 어닐링 서열(예: 5'-ACUUACCUG-3')을 표적 부위에 상보적인 이중 서열과 동일한 서열을 가지도록 변형시키는 것은 당업자의 능력 범위 내에 있다. 예를 들어, 적합한 방법은 지시된 돌연변이유발 또는 무작위 돌연변이유발에 이어서

본원에 기재된 바와 같은 변형된 U1 snRNA를 제공하는 돌연변이에 대한 선택을 포함한다.

- [0390] 본원에 기재된 바와 같은 변형된 U1 snRNA는 당업계에 일반적으로 알려진 방법에 따라 제조될 수 있다. 예를 들어, 변형된 U1 snRNA는 화학적 합성 또는 재조합 DNA/RNA 기술에 의해 제조될 수 있다.
- [0391] 일 측면에서 변형된 U1 snRNA를 암호화하는 뉴클레오티드 서열은 상이한 뉴클레오티드 서열, 예를 들어 상이한 플라스미드 상에 있을 수 있다.
- [0392] 통상적인 분자 및 세포 생물학 기술을 사용하여 본원에 기재된 바와 같은 변형된 U1 snRNA를 암호화하는 뉴클레오티드 서열을 세포 내로 도입하는 것은 당업자의 능력 범위 내에 있다.
- [0393] **개선된 TRAP 결합 부위와 Kozak 서열의 조합**
- [0394] 상기 논의된 바와 같이, TRIP 시스템은 WO 2015/092440에 기재되어 있으며 벡터 생산 동안 생산 세포에서 NOI의 발현을 억제하는 방법을 제공한다. 조직-특이적 프로모터를 포함한 구성적 및/또는 강력한 프로모터가 전이유전자를 구동하는 것이 선호되는 경우 그리고 특히 생산 세포에서 전이유전자 단백질의 발현이 벡터 역가의 감소를 초래하고/하거나 전이유전자-유래 단백질의 바이러스 벡터 전달로 인해 생체내 면역 반응을 유도하는 경우, TRAP-결합 서열(예: TRAP-tbs) 상호작용은 레트로바이러스 벡터 생산을 위한 전이유전자 단백질 억제 시스템의 기초를 형성한다(Maunders *et al*, Nat Commun. (2017) Mar 27; 8).
- [0395] 일 구현예에서, 뉴클레오티드 서열은 트립토판 RNA-결합 감쇠 단백질(TRAP) 결합 부위 또는 그의 일부를 추가로 포함한다.
- [0396] 일 측면에서 본 발명은 관심 뉴클레오티드(전이유전자) 및 트립토판 RNA-결합 감쇠 단백질(TRAP) 결합 부위를 포함하는 핵산 서열을 제공하고, 여기서 TRAP 결합 부위는 전이유전자 시작 코돈 ATG와 중첩된다.
- [0397] Kozak 서열/중첩되는 Kozak 서열에 관한 본 명세서의 임의의 개시는 시작 코돈의 ATG를 지칭하는 등가 측면 및 그와의 중첩에 동등하게 적용가능하다(적절한 경우).
- [0398] 또 다른 측면에서, 본 발명은 관심 뉴클레오티드 및 Kozak 서열을 포함하는 핵산 서열을 제공하며, 여기서 상기 Kozak 서열은 트립토판 RNA-결합 감쇠 단백질(TRAP) 결합 부위의 일부를 포함한다.
- [0399] 일 측면에서 본 발명은 관심 뉴클레오티드(전이유전자) 및 TRAP 결합 부위를 포함하는 핵산 서열을 제공하고, 여기서 TRAP 결합 부위는 전이유전자 시작 코돈 ATG의 일부를 포함하거나 그 반대도 마찬가지이다.
- [0400] 일 구현예에서, 뉴클레오티드 서열은 본원에 기재된 tbs 또는 그의 일부, 다중 클로닝 부위(MCS) 및 본원에 기재된 Kozak 서열을 추가로 포함하며, 여기서 상기 MCS는 tbs 또는 그의 일부의 하류 및 Kozak 서열의 상류에 위치한다. 적합하게는, tbs 또는 그의 일부와 Kozak 서열은 중첩되지 않는다.
- [0401] 본 발명의 일부 구현예에서, 관심 뉴클레오티드는 tbs 또는 그의 일부에 작동가능하게 연결된다. 일부 구현예에서, 관심 뉴클레오티드는 TRAP가 결합된 표적 세포에서 번역된다.
- [0402] tbs 또는 그의 일부는 관심 뉴클레오티드의 번역이 바이러스 벡터 생산 세포에서 억제되거나 방지되도록 TRAP와 상호작용할 수 있다.
- [0403] 따라서, 본 발명의 뉴클레오티드 서열 및 TRAP를 암호화하는 핵산 서열을 바이러스 벡터 생산 세포에 도입하는 것을 포함하는, 바이러스 벡터 생산 세포에서 NOI의 번역을 억제하는 방법이 제공되며, 여기서 TRAP는 TRAP 결합 부위 또는 그의 일부에 결합하여 NOI의 번역을 억제한다.
- [0404] 표 1은 본 발명에서 사용될 수 있는 서열을 나타내고, 여기서 K는 T 또는 G일 수 있고, "R"은 서열의 해당 위치에서 퓨린(즉, A 또는 G)을 지정하는 것으로 이해되어야 하고, "V"는 G, A, 또는 C로부터의 임의의 뉴클레오티드를 지정하는 것으로 이해되어야 하고, "N"은 서열의 해당 위치에서 임의의 뉴클레오티드를 지정하는 것으로 이해되어야 한다. 예를 들어, 이것은 G, A, T, C 또는 U일 수 있다. TRAP 결합 부위(tbs) 서열 또는 3' tbs 서열은 이탤릭체로 표시하였고, 다중 클로닝 부위(MCS)는 밑줄로 표시하였으며, Kozak 서열은 굵게 표시하였다.

[0405] 표 1:

서열 번호	설명	서열
69	3' <i>tbs</i> - Kozak 접합 변이체	GAGAATG
70	3' <i>tbs</i> - Kozak 접합 변이체	GAGCATG
71	3' <i>tbs</i> - Kozak 접합 변이체	GAGGATG
72	3' <i>tbs</i> - Kozak 접합 변이체	TAGAATG
73	3' <i>tbs</i> - Kozak 접합 변이체	TAGCATG
74	3' <i>tbs</i> - Kozak 접합 변이체	TAGGATG
75	3' <i>tbs</i> - Kozak 접합 변이체	GAGAAATG
76	3' <i>tbs</i> - Kozak 접합 변이체	GAGACATG
77	3' <i>tbs</i> - Kozak 접합 변이체	GAGAGATG
78	3' <i>tbs</i> - Kozak 접합 변이체	GAGCAATG
79	3' <i>tbs</i> - Kozak 접합 변이체	GAGCCATG
80	3' <i>tbs</i> - Kozak 접합 변이체	GAGCGATG
81	3' <i>tbs</i> - Kozak 접합 변이체	GAGGAATG
82	3' <i>tbs</i> - Kozak 접합 변이체	GAGGCATG

[0406]

83	3' <i>tbs</i> -Kozak 접합 변이체	GAGGGATG
84	3' <i>tbs</i> -Kozak 접합 변이체	TAGAAATG
85	3' <i>tbs</i> -Kozak 접합 변이체	TAGACATG
86	3' <i>tbs</i> -Kozak 접합 변이체	TAGAGATG
87	3' <i>tbs</i> -Kozak 접합 변이체	TAGCAATG
88	3' <i>tbs</i> -Kozak 접합 변이체	TAGCCATG
89	3' <i>tbs</i> -Kozak 접합 변이체	TAGCGATG
90	3' <i>tbs</i> -Kozak 접합 변이체	TAGGAATG
91	3' <i>tbs</i> -Kozak 접합 변이체	TAGGCATG
92	3' <i>tbs</i> -Kozak 접합 변이체	TAGGGATG
93	스플라이싱 이벤트로부터 유래된 L33 개선된 리더	<u>CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT</u> GTCGTGAAAA
94	<i>Bacillus subtilis</i> 로부터의 TRAP	MNQKHSSDFVVIKAVEDGVNVIGLTRGTDTKFHHSEK LDKGEVIAQFTEHTSAIKVRGEALIQTAYGEMKSEKK
95	<i>Aminomonas</i> <i>paucivorans</i> 로부터의 TRAP	MKEGEEAKTSVLSDYVVVKALENGVTIIGLTRGQETK FAHTEKLDGGEVWIAQFTEHTSAIKVRGASEIHTKHGM LFSGRGRNEKG
96	<i>Desulfotomaculum</i> <i>hydrothermale</i> 로부터의 TRAP	MNPMTDRSDITGDYVVVKALENGVTIIGLTRGGVTKFH HTEKLDKGEIMIAQFTEHTSAIKIRGRAELLTKHGKIRTE VDS
97	<i>B. stearothermophilus</i> 로부터의 TRAP	MYTNSDFVVIKALEDGNVNIGLTRGADTRFHHSEKLDK GEVLIAQFTEHTSAIKVRGKAYIQTRHGVIESEGKK

[0407]

98	<i>B. stearotherophilus</i> S72N 로부터의 TRAP	MYTNSDFVVIKALEDGVNVIGLTRGADTRFHHSEKLDK GEVLIAQFTEHTSAIKVRGKAYIQTRHGVIEGKK
99	<i>B. halodurans</i> 로부터의 TRAP	MNVGDNSNFFVIKAKENG VNVFGMTRGTDTRFHHSE KLDKGEVMIAQFTEHTSAVKIRGKAIQTSYGTLDEKDE
100	<i>Carboxydothemus</i> <i>hydrogenoformans</i> 로부터의 TRAP	MVCDNFAFSSAINAEYIVVKALENGVTIMGLTRGKDTK FHHTEKLDKGEVMVAQFTEHTSAIKIRGKAEIYTKHGVI KNE
101	L33 개선된 리더	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAG
102	L12 개선된 리더	CTTTTTCGCAAC
103	확장된 Kozak 컨센서스 서열	GNNRVVATG
104	코어 Kozak 컨센서스 서열	RVVATG
105	EMCV M 루프	CGTGGTTTTCTTTGAAAAACACGATGATACC
106	최적의 (중첩되는) <i>tbs</i> -Kozak	<i>GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC</i> <i>CTAGCAGAGACGAGCCGAGATG</i>
107	최적의 (중첩되는) <i>tbs</i> - <u>MCS</u> -Kozak	<i>GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC</i> <i>CTAGCAGAGACGAGAAGAGCTCTAGACCATG</i>
108	3' <i>tbs</i> -Kozak 접합 변이체	TAGATG
109	3' <i>tbs</i> -Kozak 접합 변이체	GAGATATG
110	3' <i>tbs</i> -Kozak 접합 변이체	GAGCTATG
111	3' <i>tbs</i> -Kozak 접합 변이체	TAGATATG
112	3' <i>tbs</i> -Kozak 접합 변이체	TAGCTATG
113	3' <i>tbs</i> -Kozak 컨센서스 서열	KAGNNG

[0408]

114	3' <i>tbs-Kozak</i> 컨센서스 서열	KAGATG
115	3' <i>tbs-Kozak</i> 컨센서스 서열	KAGNNTG
116	3' <i>tbs-Kozak</i> 컨센서스 서열	KAGNATG
117	tbs-kzkV0.G 변이체가 있는 닭 β-액틴 / 토끼 β-글로빈 키메라 5'UTR-인트론; 엑손 서열은 굵게 표시(함께 스플라이싱되어 5'UTR 리더가 됨), tbskzkV0.G 는 이탤릭체로 표시	CGGCGGGCGGGAACGTTGCCTTCGCCCCGTGCC CGCTCCGCGCCGCCTCGCGCCGCCCGCCCCGGCT CTGACTGACCGCGTTACTCCCACAGGTGAGCGGG CGGGACGGCCCTTCTCCCTCCGGGCTGTAATTAGC GCTTGTTTAATGACGGCTCGTTTCTTTCTGTGGC TGCGTGAAAGCCTAAAGGGCTCCGGGAGGGCCTT TGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGC GTGTGTGTGCGTGGGGAGCGCCGCGTGC GGCC CGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGC GGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGA GGGGAGCGCGGGCCGGGGCGGTGCCCCGCGGT GCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTG CGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTG TGGGCGCGGCGGTGCGGCTGTAACCCCCCTGG CACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGG CTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGC GGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA GGTGGGGGTGCCGGGCGGGGCGGGGCCGCCTCG GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGG CCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGA GCCGCAGCCATTGCCTTTATGGTAATCGTGCGAGA GGGCGCAGGGAATTCCTTTGTCCCAAATCTGGCGG AGCCGAAATCTGGGAGGCGCCGCCGACCCCCCTCT AGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCA GGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTC GCCGCGCCGCGTCCCCTTCTCCATCTCCAGCCTC GGGGCTGCCGCAGGGGGACGGCTGCCTTCGGGGG GGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTG TGACCGGCGGCTTTAGAGCCTCTGCTAACCATGTTC

[0409]

		ATGCCTTCTTCTTTTCTCCTACAGCTCCTGGGCAAAG AGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCC TAGCAGAGACGAGCCGAGATG
118	tbskzkV0.G 변이체가 있는 EF1a 5'UTR- 인트론; 엑손 서열은 굵게 표시 (함께 스플라이싱되어 5'UTR 리더가 됨), tbskzkV0.G 는 이탤릭체로 표시	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT AAGTGCCGTGTGTGGTTCCTCGGGCCTGGCCTCT TTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTT CCACCTGGCTGCAGTACGTGATTCTTGATCCCGAG CTTCTGGGTTGGAAGTGGGTGGGAGAGTTCTGTGGCC TTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGTGGCCTGGCCTGGGCGCTGGGGCCGCCGCGT GCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTG CTTTCGATAAGTCTCTAGCCATTTAAATTTTGTATG ACCTGCTGCGACGCTTTTTTCTGGCAAGATAGTCT TGTAATGCGGGCCAAGATCAGCACACTGGTATTTT GGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGT GCGTCCCAGCGCACATGTTCTGGCGAGGCGGGGCC TGCGAGCGCGGCCACCGAGAATCGGACGGGGGTA GTCTCAAGCTGCCCGGCCTGCTCTGGTGCCTGGCC TCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGC AAGGCTGGCCCGGTCGGCACCAAGTTGCGTGAGCG GAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGA GCACAAAATGGAGGACGCGGCGCTCGGGAGAGCG GGCGGGTGAGTACCCACACAAAGGAAAAGGGCCT TTCCGTCCTCAGCCGTGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTT CTCCAGCTTTTGGAGTACGTGCTTTAGGTTGGGG GGAGGGGTTTTATGCGATGGAGTTTCCCCACACTG AGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCAC TTGATGTAATTCTCCTTGGAATTTGCCCTTTTGTAGT TTGGATCTTGTTTCACTCTCAAGCCTCAGACAGTGG TTCAAAGTTTTTTCTTCCATTTCAGGTGTCGTGAAA AGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAG CCTAGCAGAGACGAGCCGAGATG
119	서열번호 117 에 상응하는 스플라이싱된 서열; 5'UTR 리더 서열은 굵게 표시,	CGGCGGGCGGGAACGTTGCCTTCGCCCCGTGCC CGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCT CTGACTGACCGCGTTACTCCACAGCTCCTGGGCA

[0410]

	tbskzkV0.G 는 이탤릭체로 표시	AAGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGA GCCTAGCAGAGACGAGCCGAGATG
120	서열번호 118 에 상응하는 스플라이싱된 서열; 5'UTR 리더 서열은 굵게 표시, tbskzkV0.G 는 이탤릭체로 표시	CTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT GTCGTGAAAAGAGTTTAGCGGAGTGGAGAAGAGC GGAGCCGAGCCTAGCAGAGACGAGCCGAGATG
121	닭 β-액틴 / 토끼 β- 글로빈 키메라 5'UTR- 인트론(엑손 서열은 굵게 표시(함께 스플라이싱되어 5'UTR 리더가 됨))	CGGCGGGCGGGAACGTTGCCTTCGCCCCGTGCCC CGCTCCGCGCCGCCTCGCGCCGCCCGCCCCGGCT CTGACTGACCGCGTTACTCCCACAGGTGAGCGGG CGGGACGGCCCTTCTCCCTCCGGGCTGTAATTAGC GCTTGCTTTAATGACGGCTCGTTTCTTTTCTGTGGC TGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCTT TGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGC GTGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCC CGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGC GGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGA GGGGAGCGCGGGCCGGGGGCGGTGCCCGCGGT GCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTG CGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTG TGGGCGCGGCGGTGCGGCTGTAACCCCCCTGG CACCCCCCTCCCCGAGTTGCTGAGCACGGCCCCG CTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCG GGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG GGCCGGGGAGGGCTCGGGGAGGGGCGCGGCGG CCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGA GCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAGA GGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGG AGCCGAAATCTGGGAGGCGCCGCCACCCCCCTCT AGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCA GGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTC GCCGCGCCGCGTCCCCTTCTCCATCTCCAGCCTC GGGGCTGCCGCAGGGGGACGGCTGCCTTCGGGG GGACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTG

[0411]

		TGACCGGCGGCTTTAGAGCCTCTGCTAACCATGTTCTGCTTCTTCTTTTCTTACAGCTCCTGGGCAAA
122	EF1a 5'UTR-인트론 (엑손 서열은 굵게 표시 (함께 스플라이싱되어 5'UTR 리더가 됨))	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT AAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCT TTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTT CCACCTGGCTGCAGTACGTGATTCTTGATCCCGAG CTTCGGGTGGAAGTGGGTGGGAGAGTTCTGTGGCC TTGCGCTTAAGGAGCCCTTCGCCTCGTGTGAG TTGTGGCCTGGCCTGGGCGCTGGGGCCGCCGCGT GCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTG CTTTCGATAAGTCTCTAGCCATTTAAAATTTTGATG ACCTGCTGCGACGCTTTTTTCTGGCAAGATAGTCT TGTAATGCGGGCCAAGATCAGCACACTGGTATTTT GGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGT GCGTCCCAGCGCACATGTTGCGCGAGGCGGGGCC TGCGAGCGCGGCCACCGAGAATCGGACGGGGGTA GTCTCAAGCTGCCCCGGCCTGCTCTGGTGCCTGGCC TCGCGCCGCGGTGTATCGCCCCGCCCTGGGCGGC AAGGCTGGCCCGGTGCGCACCAAGTTGCGTGAGCG GAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGA GCACAAAATGGAGGACGCGGCGCTCGGGAGAGCG GGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCT TTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTT CTCCAGCTTTTGGAGTACGTCGTCTTTAGGTTGGGG GGAGGGGTTTTATGCGATGGAGTTTCCCCACACTG AGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCAC TTGATGTAATTCTCCTTGAATTTGCCCTTTTGAGT TTGATCTTGGTTCATTCTCAAGCCTCAGACAGTGG TTCAAAGTTTTTTCTTCCATTT CAGGTGTCGTGAAA A
123	닭 β-액틴 / 토끼 β- 글로빈 키메라 5'UTR- 인트론-tbs 컨센서스 (엑손 서열은 굵게 표시)	CGGCGGGCGGGAACGTTGCCTTCGCCCCGTGCC CGCTCCGCGCCGCCTCGCGCCGCCCGCCCGGCT CTGACTGACCGCGTTACTCCCACAG GTGAGCGGG CGGGACGGCCCTTCTCCCTCCGGGCTGTAATTAGC GCTTGGTTAATGACGGCTCGTTTCTTTCTGTGCG

[0412]

	(함께 스플라이싱되어 5'UTR 리더가 됨), tbs 컨센서스는 이탤릭체로 표시)	TGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCTT TGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGC GTGTGTGTGTGCGTGCGGAGCGCCGCGTGCGGCC CGCGCTGCCCCGGCGGCTGTGAGCGCTGCGGGCGC GGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGA GGGAGCGCGGGCCGGGGGCGGTGCCCCGCGGT GCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTG CGGGGTGTGTGCGTGCGGGGGGTGAGCAGGGGGTG TGGGCGCGGCGGTGCGGCTGTAACCCCCCTGG CACCCCCCTCCCCGAGTTGCTGAGCACGGCCCCG CTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGC GGGGCTCGCCGTGCCGGGCGGGGGGTGGCGCA GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGG CCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGA GCCGCAGCCATTGCCTTTATGGTAATCGTGCGAGA GGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGG AGCCGAAATCTGGGAGGCGCCGCCGACCCCCCTCT AGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCA GGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTC GCCGCGCCGCCGTCCCCTTCTCCATCTCCAGCCTC GGGGCTGCCGCAGGGGGACGGCTGCCTTCGGGGG GGACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTG TGACCGGCGGCTTTAGAGCCTCTGCTAACCATGTTT ATGCCCTTCTTCTTTTCTACAGCTCCTGGGCAAA/K <i>AGN₂₋₃</i>¹⁰⁻¹¹
124	EF1a 5'UTR-인트론-tbs 컨센서스 (엑손 서열은 굵게 표시 (함께 스플라이싱되어 5'UTR 리더가 됨), tbs 컨센서스는 이탤릭체로 표시)	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT AAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCT TTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTT CCACCTGGCTGCAGTACGTGATTCTTGATCCCGAG CTTCGGGTTGGAAGTGGGTGGGAGAGTTCGTGGCC TTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGTGGCCTGGCCTGGGCGCTGGGGCCGCCGCGT GCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTG CTTTCGATAAGTCTCTAGCCATTTAAATTTTGTATG ACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCT

[0413]

		TGTAAATGCGGGCCAAGATCAGCACACTGGTATTC GGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGT GCGTCCCAGCGCACATGTTGCGCGAGGCGGGGCC TGCAGCGCGGCCACCGAGAATCGGACGGGGGTA GTCTCAAGCTGCCCCGGCCTGCTCTGGTGCCTGGCC TCGCGCCCGCGTGTATCGCCCCGCCCTGGGCGGC AAGGCTGGCCCCGGTCGGCACCAAGTTGCGTGAGCG GAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGA GCACAAAATGGAGGACGCGGCGCTCGGGAGAGCG GGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCT TTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTT CTCCAGCTTTTGGAGTACGTCGCTTTAGGTTGGGG GGAGGGGTTTTATGCGATGGAGTTTCCCCACACTG AGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCAC TTGATGTAATTCTCCTTGAATTTGCCCTTTTTGAGT TTGGATCTTGTTTCATTCTCAAGCCTCAGACAGTGG TTCAAAGTTTTTTCTTCCATTTCAAGTGTCTGTA AA[KAGN₂₋₃]₁₀₋₁₁
125	TRAP 결합 부위 변이체 - [KAGNN] ₁₁	GAGUUUAGCGGAGUGGAGAAGAGCGGAGCCGAGC CUAGCAGAGACGAGUGGAGCU
126	TRAP 결합 부위 변이체 - [KAGNN] ₁₁	GAGUUUAGCGGAGUGGAGAAGAGCGGAGCCGAGC CUAGCAGAGACGAGAAGAGCU
127	TRAP 결합 부위 변이체 - [KAGNN] ₆	UAGUUUAGUUUAGUUUAGUUUAGUUUAGUU
128	TRAP 결합 부위 변이체 - [KAGNN] ₆	UAGUUUAGUUGAGUUUAGUUGAGUUUAGUU
129	TRAP 결합 부위 변이체 - [KAGNN] ₆	GAGUUUAGUUGAGUUGAGUUGAGUUUAGUUGAGUU
130	TRAP 결합 부위 변이체 - [KAGNN] ₆	UAGUUUAGUUGAGUUUAGUUGAGUUUAGUUGAGUU
131	TRAP 결합 부위 변이체 - [KAGNN] ₁₁	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC CTAGCAGAGACGAGTGGAGCT
132	TRAP 결합 부위 변이체 - [KAGNN] ₁₀	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC CTAGCAGAGACGAGAA

[0414]

133	TRAP 결함 부위 변이체 - [KAGNN] ₉	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC CTAGCAGAGAC
134	TRAP 결함 부위 변이체 - [KAGNN] ₈	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC CTAGCA
135	TRAP 결함 부위 변이체 - [KAGNN] ₇	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC C
136	TRAP 결함 부위 변이체 - [KAGNN] ₆	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCC
137	TRAP 결함 부위 변이체 - [KAGNNN] ₃ [KAGNN] ₈	GAGTTTAGCGGAGTGGAGAAAGAGACGGAGCCGA GACCTAGCAGAGACGAGAAGAGCT
138	TRAP 결함 부위 변이체 - [KAGNNN] ₁ [KAGNN] ₁₀	GAGTTTAGCGGAGTGGAGAAGAGACGGAGCCGAG CCTAGCAGAGACGAGAAGAGCT
139	TRAP 결함 부위 변이체 - [KAGNN] ₁₁	GAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGC CTAGCAGAGACGAGAAGAGCT
140	TRAP 결함 부위 변이체 - [KAGNNN] ₁ [KAGNN] ₇	GAGTTTAGCGGAGTGGAGAAAGAGCGGAGCCGAG CCTAGCA
141	TRAP 결함 부위 변이체 - [KAGNNN] ₁ [KAGNN] ₆	GAGTTTAGCGGAGTGGAGAAAGAGCGGAGCCGAG CC
142	3' <i>tbs</i> -Kozak 접합 변이체	GAGATG
143	3' <i>tbs</i> -Kozak 접합 변이체	KAGVATG
144	3' <i>tbs</i> -Kozak 접합 변이체	KAGVVATG
145	3' <i>tbs</i> -Kozak 접합 변이체	KAGRWATG
146	3' <i>tbs</i> -Kozak 접합 변이체	KAGNRVATG

[0415]

147	최적의 3' <i>tbs</i> - Kozak 접합 변이체	KAGCCGAGATG
148	최적의 3' <i>tbs</i> - Kozak 접합 변이체	KAGNGGAGCCATG
149	최적의 3' <i>tbs</i> - Kozak 접합 변이체	KAGNNGAGACCATG
150	3' <i>tbs</i> - Kozak 접합 변이체	KAGGCGAGCATG
151	스페이스 변이체	ATAGCAGAGACGGCT
152	스페이스 변이체	ATAGCAGAGA
153	스페이스 변이체	ATAGC
154	스페이스 변이체	ATATCAGAGACGGCTAGCGTATACCA
155	스페이스 변이체	ATATCAGAGACGGCT
156	스페이스 변이체	AGAGACGGCT
157	스페이스 변이체	TACCA
158	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAG</u> AVVATG
159	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCGTCG</u> ACVATG
160	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCGAATTCGA</u> AVVATG
161	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAGACGTCG</u> ACVATG
162	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAGAATTCGA</u> AVVATG
163	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAGATATCGAT</u> RVVATG
164	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	KAGACTAGTACTTAAGCTTRVVATG
165	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAG</u> ACCATG

[0416]

166	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCGTCG</u> ACCATG
167	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCGAATT</u> CGAACCATG
168	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAGACGTCG</u> ACCATG
169	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAGAATT</u> CGAACCATG
170	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>GAGCTCTAGATATCGAT</u> ACCATG
171	3' <i>tbs</i> - <u>MCS</u> - Kozak 변이체	<u>KAGACTAGTACTTAAGCTT</u> ACCATG
172	L33 개선된 리더, 최적의 (중첩되는) <i>tbs</i> ([KAGNN] ₈)- Kozak 접합을 포함하는 예시적인 핵산 서열 1	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGA GTTTAGCGGAGTGGAGAAGAGCGGAGCCGAG GCCG AGATG
173	L33 개선된 리더, 최적의 (중첩되는) <i>tbs</i> ([KAGNN] ₁₁)- Kozak 접합을 포함하는 예시적인 핵산 서열 2	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGA GTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTA GCAGAGACGAG GCCGAGATG
174	L12 개선된 리더, 최적의 (중첩되는) <i>tbs</i> ([KAGNN] ₁₁)- Kozak 접합을 포함하는 예시적인 핵산 서열 3	CTTTTTCGCAACGAGTTTAGCGGAGTGGAGAAGAG CGGAGCCGAGCCTAGCAGAGACGAG GCCGAGATG
175	L33, 최적의 (중첩되는) <i>tbs</i> ([KAGNN] ₁₁)- Kozak 접합을 포함하는 스플라이싱된 리더를	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT GTCGTGAAAAGAGTTTAGCGGAGTGGAGAAGAGCG GAGCCGAGCCTAGCAGAGACGAGCCGAGATG

[0417]

	생성하는 인트론 함유 5'UTR 에 대한 예시적인 핵산 서열 4	
176	개선된 스페이서, 최적의 (중첩되는) <i>tbs</i> ([KAGNN] ₈)- Kozak 접합을 포함하는 예시적인 핵산 서열 5	ATAGCAGAGACGGCTGAGTTTAGCGGAGTGGAGAA GAGCGGAGCCGAG CCGAGATG
177	L33 개선된 리터 <i>tbs</i> ([KAGNN] ₁₁)- MCS- Kozak 을 포함하는 예시적인 핵산 서열 6	CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGA GTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTA GCAGAGACGAGAAGAGCTCTAG ACCATG
178	개선된 스페이서, <i>tbs</i> ([KAGNN] ₁₁)- MCS- Kozak 을 포함하는 예시적인 핵산 서열 7	ATAGCAGAGACGGCTGAGTTTAGCGGAGTGGAGAA GAGCGGAGCCGAGCCTAGCAGAGACGAGAAGAGC <u>TCTAGACCATG</u>

[0418]

[0419]

트립토판 RNA-결합 감쇠 단백질(TRAP)은 바실러스 서브틸리스(*Bacillus subtilis*)에서 광범위하게 특성화된 박테리아 단백질이다. TRAP은 WO 2015/092440에 기재되어 있다.

[0420]

TRAP 오픈 리딩 프레임은 포유류(예: 호모 사피엔스) 세포에서의 발현을 위해 코돈 최적화될 수 있는데, 그 이유는 박테리아 유전자 서열이 포유류 세포에서의 발현에 최적이지 아닐 가능성이 있기 때문이다. 잠재적인 불안정한 서열 및 스플라이싱 부위를 제거하여 서열을 최적화할 수도 있다. TRAP 단백질에서 C-말단에 발현되는 HIS-태그의 사용은 번역 억제체의 측면에서 이점을 제공하는 것으로 보이며 선택적으로 사용될 수 있다. 이 C-말단 HIS-태그는 개선된 기능적 이점을 배제할 수 없지만 진핵 세포 내에서 TRAP의 용해도 또는 안정성을 개선할 수 있다. 그럼에도 불구하고, HIS-태깅된 TRAP 및 태깅되지 않은 TRAP 양자 모두 전이유전자 발현의 강력한 억제를 허용하였다. TRAP 전사 단위 내의 특정 시스-작용 서열도 최적화될 수 있는데; 예를 들어, EF1α 프로모터-구동 작제물은 일시적 형질감염의 맥락에서 CMV 프로모터-구동 작제물에 비해 TRAP 플라스미드의 낮은 입력(input)으로 더 나은 억제를 가능하게 한다.

[0421]

일 구현예에서, 본 발명에서 사용하기 위한 TRAP는 박테리아로부터 유래된다.

[0422]

본 발명의 일 구현예에서, TRAP는 *바실러스* 종, 예를 들어 *바실러스 서브틸리스*로부터 유래된다. 예를 들어, TRAP는 서열번호 94에 기재된 서열을 포함할 수 있다.

[0423]

본 발명의 바람직한 구현예에서, 서열번호 94는 6개의 히스티딘 아미노산으로 C-말단에 태깅된다(HISx6 태그).

[0424]

대안적인 구현예에서, TRAP는 *Aminomonas paucivorans*로부터 유래된다. 예를 들어, TRAP는 서열번호 95에 기재된 서열을 포함할 수 있다.

[0425]

대안적인 구현예에서, TRAP는 *Desulfotomaculum hydrothermale*로부터 유래된다. 예를 들어, TRAP는 서열번호 96에 기재된 서열을 포함할 수 있다.

[0426]

대안적인 구현예에서, TRAP는 *B. stearothermophilus*로부터 유래된다. 예를 들어, TRAP는 서열번호 97에 기재된 서열을 포함할 수 있다.

[0427]

대안적인 구현예에서, TRAP는 *B. stearothermophilus* S72N으로부터 유래된다. 예를 들어, TRAP는 서열번호 98에 기재된 서열을 포함할 수 있다.

[0428]

대안적인 구현예에서, TRAP는 *B. halodurans*로부터 유래된다. 예를 들어, TRAP는 서열번호 99에 기재된 서열을 포함할 수 있다.

[0429]

대안적인 구현예에서, TRAP는 *Carboxydotherrmus hydrogenoformans*로부터 유래된다. 예를 들어, TRAP는 서열번호

호 100에 기재된 서열을 포함할 수 있다.

- [0430] 일 구현예에서, TRAP는 트립토판 RNA-결합 감쇠 단백질 유전자 패밀리 *mtrB*(TrpBP 슈퍼패밀리, 예를 들어 NCBI 보존 도메인 데이터베이스 # c103437)에 의해 암호화된다.
- [0431] 바람직한 구현예에서, TRAP는 6개의 히스티딘 아미노산으로 C-말단에 태깅된다(HISx6 태그).
- [0432] 바람직한 구현예에서, TRAP는 서열번호 94 내지 100 중 어느 하나에 대해 50%, 60%, 70%, 80%, 90%, 95%, 99% 또는 100% 동일성을 가지고, 작동가능하게 연결된 NOI의 발현이 바이러스 벡터 생산 세포에서 변형, 예를 들어 억제 또는 방지되도록 RNA-결합 부위와 상호작용할 수 있는 아미노산 서열을 포함한다.
- [0433] 바람직한 구현예에서, TRAP는 서열번호 94 내지 100 중 어느 하나에 대해 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99% 또는 100% 동일성을 가지고, 작동가능하게 연결된 NOI의 발현이 바이러스 벡터 생산 세포에서 변형, 예를 들어 억제 또는 방지되도록 RNA-결합 부위와 상호작용할 수 있는 아미노산 서열을 포함한다.
- [0434] 또 다른 구현예에서, TRAP는 작동가능하게 연결된 NOI의 발현이 바이러스 벡터 생산 세포에서 변형, 예를 들어 억제 또는 방지되도록 RNA-결합 부위와 상호작용할 수 있는 단백질을 암호화하는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드에 의해 암호화될 수 있다. 예를 들어, TRAP는 서열번호 94 내지 100의 단백질을 암호화하는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드에 의해 암호화될 수 있다.
- [0435] 본 발명에서 사용하기 위한 TRAP의 모든 변이체, 단편 또는 동족체는 NOI(마커 유전자일 수 있음)의 번역이 바이러스 벡터 생산 세포에서 억제 또는 방지되도록 본원에 기재된 바와 같은 TRAP 결합 부위에 결합하는 능력을 보유할 것이다.
- [0436] **TRAP 결합 부위**
- [0437] 용어 "결합 부위"는 특정 단백질과 상호작용할 수 있는 핵산 서열로 이해되어야 한다.
- [0438] TRAP에 결합할 수 있는 컨센서스 TRAP 결합 부위 서열은 [KAGNN]이 여러 번 반복되는데(예: 6, 7, 8, 9, 10, 11, 12회 이상); 이러한 서열은 천연 *trp* 오페론에서 발견된다. 자연적 맥락에서, 때때로 AAGNN이 허용되고 때때로 추가적인 "스페이싱" N 뉴클레오티드가 기능적 서열을 생성한다. 시험관 내 실험은 TRAP-RNA 결합에 적어도 6개 이상의 컨센서스 반복부가 필요함을 입증하였다(Babitzke P, Y. J., Campanelli D. (1996) *Journal of Bacteriology* **178**(17): 5159-5163). 따라서, 바람직하게는 일 구현예에서 *tbs* 내에 6개 이상의 연속적인 [KAGN_{≥2}] 서열이 존재하며, 여기서 K는 DNA에서 T 또는 G이고 RNA에서 U 또는 G일 수 있다.
- [0439] RNA-결합 단백질로서 TRAP의 경우, 바람직하게는 TRIP 시스템은 적어도 8개의 KAGNN 반복부를 함유하는 *tbs* 서열로 최대로 작동하지만, 7개의 반복부는 여전히 강력한 전이유전자 억제를 얻는 데 사용될 수 있고, 6개의 반복부는 벡터 역가를 구제할 수 있는 수준으로 전이유전자를 충분히 억제하는데 사용될 수 있다. KAGNN 컨센서스 서열은 TRAP 매개 억제를 유지하기 위해 다양할 수 있지만, 바람직하게는 선택된 정확한 서열은 억제되지 않은 상태에서 높은 수준의 번역을 보장하도록 최적화될 수 있다. 예를 들어, *tbs* 서열은 TRAP의 부재에서(즉, 표적 세포에서) mRNA의 번역 효율을 방해할 수 있는 스플라이싱 부위, 불안정한 서열 또는 스템-루프를 제거함으로써 최적화될 수 있다. 주어진 *tbs*의 KAGNN 반복부의 구성과 관련하여, KAG 반복부 사이의 N "스페이싱" 뉴클레오티드의 수는 바람직하게는 2이다. 그러나, 적어도 2개의 KAG 반복부 사이에 2개를 초과하는 N 스페이서를 함유하는 *tbs*가 허용될 수 있다(3개의 N을 포함하는 반복부의 50%는 시험관 내 결합 연구에 의해 판단되는 바와 같이 기능적 *tbs*를 생성할 수 있다; Babitzke P, Y. J., Campanelli D. (1996) *Journal of Bacteriology* **178**(17): 5159-5163). 실제로, 11x KAGNN *tbs* 서열은 KAGNN 반복부로의 최대 3번의 교체를 허용할 수 있으며 여전히 TRAP 결합과 협력하여 잠재적으로 유용한 번역 차단 활성을 유지하는 것으로 나타났다.
- [0440] 본 발명의 일 구현예에서, TRAP 결합 부위 또는 그의 일부는 서열 KAGN_{≥2}(예를 들어, KAGN₂₋₃)를 포함한다. 이에 대한 의심의 소지를 없애기 위해, 이 *tbs* 또는 그의 일부는 예를 들어 UAGNN, GAGNN, TAGNN, UAGNNN, GAGNNN 또는 TAGNNN과 같은 반복부 서열 중 하나를 포함한다.
- [0441] "N"은 서열의 해당 위치에서 임의의 뉴클레오티드를 지정하는 것으로 이해되어야 한다. 예를 들어, 이것은 G, A, T, C 또는 U일 수 있다. 이러한 뉴클레오티드의 수는 바람직하게는 2이지만 최대 3, 예를 들어 1, 2 또는 3이다. 11x 반복부 *tbs*의 KAG 반복부 또는 그의 일부는 3개의 스페이싱 뉴클레오티드에 의해 분리될 수 있으며 번역 억제를 유도하는 일부 TRAP-결합 활성을 여전히 유지한다. 번역 억제를 유도하는 최대 TRAP-결합 활성을

유지하기 위해서는 바람직하게는 1개 이하의 N_3 스페이서가 11x 반복부 tbs 또는 그의 일부에 사용될 것이다.

- [0442] 또 다른 구현예에서, tbs 또는 그의 일부는 $KAGN_{\geq 2}$ 의 다중 반복부(예를 들어, $KAGN_{2-3}$ 의 다중 반복부)를 포함한다.
- [0443] 또 다른 구현예에서, tbs 또는 그의 일부는 서열 $KAGN_2$ 의 다중 반복부를 포함한다.
- [0444] 또 다른 구현예에서, tbs 또는 그의 일부는 $KAGN_{\geq 2}$ 의 적어도 6개의 반복부(예를 들어, $KAGN_{2-3}$ 의 적어도 6개의 반복부)를 포함한다.
- [0445] 또 다른 구현예에서, tbs 또는 그의 일부는 $KAGN_2$ 의 적어도 6개의 반복부를 포함한다. 예를 들어, tbs 또는 그의 일부는 $KAGN_2$ 의 6, 7, 8, 9, 10, 11, 12개 이상의 반복부를 포함할 수 있다. 예를 들어, tbs 또는 그의 일부는 서열번호 125-136 또는 139 중 어느 하나를 포함할 수 있다.
- [0446] 또 다른 구현예에서, tbs 또는 그의 일부는 $KAGN_{\geq 2}$ 의 적어도 8개의 반복부(예를 들어, $KAGN_{2-3}$ 의 적어도 8개의 반복부)를 포함한다. 예를 들어, tbs 또는 그의 일부는 서열번호 125, 126, 131-134, 137-24 중 어느 하나를 포함할 수 있다.
- [0447] 바람직하게는, tbs 또는 그의 일부에 존재하는 $KAGNNN$ 반복부의 수는 1 이하이다. 예를 들어, tbs 또는 그의 일부는 서열번호 125, 126, 131-134, 136-141 중 어느 하나를 포함할 수 있다.
- [0448] 또 다른 구현예에서, tbs 또는 그의 일부는 $KAGN_{\geq 2}$ 의 11개의 반복부(예를 들어, $KAGN_{2-3}$ 의 11개의 반복부)를 포함한다. 바람직하게는, 이 tbs 또는 그의 일부에 존재하는 $KAGNNN$ 반복부의 수는 3개 이하이다. 예를 들어, tbs 또는 그의 일부는 서열번호 125, 126, 131, 137-139 중 어느 하나를 포함할 수 있다.
- [0449] 또 다른 구현예에서, tbs 또는 그의 일부는 $KAGN_{\geq 2}$ 의 12개의 반복부(예를 들어, $KAGN_{2-3}$ 의 12개의 반복부)를 포함한다.
- [0450] 바람직한 구현예에서, tbs 또는 그의 일부는 $KAGN_2$ 의 8-11개의 반복부(예를 들어, $KAGN_2$ 의 8, 9, 10 또는 11개의 반복부)를 포함한다. 예를 들어, tbs 또는 그의 일부는 서열번호 125, 126, 131-134, 137-141 중 어느 하나를 포함할 수 있다.
- [0451] 일 구현예에서, TRAP 결합 부위 또는 그의 일부는 서열번호 125-141 중 임의의 것을 포함할 수 있다.
- [0452] 예를 들어, TRAP 결합 부위 또는 그의 일부는 서열번호 125 또는 서열번호 126에 기재된 서열을 포함할 수 있다.
- [0453] " $KAGN_{\geq 2}$ 의 반복부"는 일반적인 $KAGN_{\geq 2}$ (예: $KAGN_{2-3}$) 모티프가 반복되는 것으로 이해되어야 한다. 이 모티프의 기준을 만족하는 상이한 $KAGN_{\geq 2}$ 서열이 tbs 또는 그의 일부를 구성하기 위해 결합될 수 있다. 수득되는 tbs 또는 그의 일부가 이 모티프의 요구 사항을 충족하는 단 하나의 서열의 반복으로 제한되는 것으로 의도되지 않으며, 비록 이 가능성이 정의에 포함되어 있긴 하지만 그러하다. 예를 들어, " $KAGN_{\geq 2}$ 의 6개의 반복부"에는 서열번호 127-130에 기재된 서열이 포함되지만 이에 국한되지는 않는다.
- [0454] 1개의 $KAGNNN$ 반복부 및 7개의 $KAGNN$ 반복부를 포함하는 8개-반복부 tbs 또는 그의 일부는 TRAP 매개 억제 활성을 유지한다. 1개 이상의 $KAGNNN$ 반복부를 포함하는 8개 미만-반복부 tbs 서열 또는 그의 일부(예: 7개- 또는 6개-반복부 tbs 서열 또는 그의 일부)는 TRAP 매개 억제 활성이 더 낮을 수 있다. 따라서, 8개 미만-반복부가 존재할 때, tbs 또는 그의 일부는 $KAGNN$ 반복부만을 포함하는 것이 바람직하다.
- [0455] $KAGNN$ 반복부 컨센서스에서 사용하기에 바람직한 뉴클레오티드는 NN 스페이서 위치 중 적어도 하나에 피리미딘; NN 스페이서 위치의 첫 번째 위치에 피리미딘; NN 스페이서 위치 모두에 피리미딘; K 위치에 G이다.
- [0456] 또한 NN 스페이서 위치가 AA일 때 K 위치에서 G를 사용하는 것이 바람직하다(즉, 컨센서스 서열에서 TAGAA가 반복부로 사용되지 않는 것이 바람직하다).
- [0457] "상호작용할 수 있는"이란 핵산 결합 부위(예: tbs 또는 그의 일부)가 세포, 예를 들어 진핵생물 바이러스 벡터 생산 세포에서 마주치는 조건 하에서 단백질, 예를 들어 TRAP에 결합할 수 있음으로 이해되어야 한다. TRAP과 같은 RNA-결합 단백질과의 이러한 상호작용은 핵산 결합 부위(예: tbs 또는 그의 일부)가 작동가능하게 연결된

NOI의 번역의 억제 또는 방지를 초래한다.

- [0458] "작동가능하게 연결된"이란 기재된 구성요소가 의도된 방식으로 기능하도록 허용하는 관계에 있으므로 이해되어야 한다. 따라서 NOI에 작동가능하게 연결된 본 발명에서 사용하기 위한 tbs 또는 그의 일부는 TRAP가 tbs 또는 그의 일부에 결합할 때 NOI의 번역이 변형되는 방식으로 위치된다.
- [0459] 주어진 오픈 리딩 프레임(ORF)의 NOI 번역 개시 코돈의 상류에서 TRAP와 상호작용할 수 있는 tbs 또는 그의 일부의 배치는 해당 ORF로부터 유래된 mRNA의 특이적 번역 억제를 허용한다. tbs 또는 그의 일부와 번역 개시 코돈을 분리하는 뉴클레오티드의 수는 억제의 정도에 영향을 미치지 않으면서 예를 들어 0 내지 34개의 뉴클레오티드로 다양할 수 있다. 추가의 예로서, 0 내지 13개의 뉴클레오티드를 사용하여 TRAP-결합 부위 또는 그의 일부와 번역 개시 코돈을 분리할 수 있다.
- [0460] tbs 또는 그의 일부는 내부 리보솜 진입 부위(IRES)의 하류에 배치되어 멀티시스트론 mRNA에서 NOI의 번역을 억제할 수 있다. 실제로, 이것은 tbs-결합 TRAP가 40S 리보솜의 통과를 차단할 수 있다는 추가의 증거를 제공한다; IRES 요소는 전체 번역 복합체가 형성되기 전에 CAP-독립적인 방식으로 40S 리보솜 서브유닛을 mRNA에 격리시키는 기능을 한다(IRES 번역 개시에 대한 검토를 위해 Thompson, S. (2012) *Trends in Microbiology* 20(11): 558-566 참조). 따라서, TRIP 시스템은 바이러스 벡터 계놈으로부터 발현되는 단일 mRNA로부터의 다중 오픈 리딩 프레임을 억제하는 것이 가능하다. 여러 치료 유전자를 암호화하는 벡터를 생성할 때, 특히 모든 전이유전자 산물이 벡터 역가에 어느 정도 부정적인 영향을 미칠 수 있는 경우에, 이것은 TRIP 시스템의 유용한 특징이 될 것이다.
- [0461] 일 구현예에서, 핵산 서열은 IRES와 tbs 또는 그의 일부 사이에 스페이서 서열을 포함한다. IRES는 "내부 리보솜 진입 부위"라는 소제목으로 본원에 기술된 바와 같은 IRES일 수 있다. 스페이서 서열은 0 내지 30개 뉴클레오티드 길이, 바람직하게는 15개 뉴클레오티드 길이일 수 있다. 스페이서는 서열번호 151-157 중 어느 하나에 정의된 서열을 포함할 수 있고, 바람직하게는 스페이서는 서열번호 151에 정의된 서열을 포함한다.
- [0462] 일 구현예에서, IRES와 tbs 또는 그의 일부 사이의 스페이서 서열은 tbs 또는 그의 일부의 3' 말단 및 NOI의 하류 개시 코돈으로부터 3 또는 9개의 뉴클레오티드이다.
- [0463] 일 구현예에서, tbs 또는 그의 일부에는 II형 제한 효소 부위가 결합되어 있다. 바람직한 구현예에서, tbs 또는 그의 일부에는 SapI 제한 효소 부위가 결합되어 있다.
- [0464] 일부 구현예에서, 핵산 서열은 RRE 서열 또는 그의 기능적 대체물을 추가로 포함한다.
- [0465] **중첩되는 Kozak 서열과 TRAP 결합 부위 서열**
- [0466] 본 발명자들은 놀랍게도 중첩되지 않는 tbs 및 Kozak 서열의 사용에 비해 tbs 또는 그의 일부의 3' 말단 내에 Kozak 서열을 '숨김'으로써(중첩되는 tbs 및 Kozak 서열의 사용; 도 19B 및 19C 참조) 개선된 수준의 억제가 달성될 수 있음을 발견하였다. 또한, 테스트된 모든 중첩되는 Kozak 및 tbs 서열은 예기치 않게 효율적인 번역 개시 수준을 지시하였다. 즉, 테스트된 중첩되는 서열은 TRAP의 부재 하에 중첩되지 않는 Kozak 및 tbs 서열과 유사한 수준의 전이유전자 발현을 나타내었다. 이론에 얽매이고자 함이 없이, 개선된 수준의 억제는 tbs 또는 그의 일부가 Kozak 서열과 중첩될 때 TRAP-tbs 복합체에 의한 전이유전자 개시 코돈의 개선된 폐색에 기인할 수 있다.
- [0467] "Kozak 서열"이라는 용어는 번역 개시 부위로서 리보솜에 의해 인식되는 진핵생물 mRNA의 컨센서스 서열로 이해되어야 한다. Kozak 서열은 DNA에서 ATG(mRNA에서는 AUG) 개시(시작) 코돈을 포함한다. 진핵생물 mRNA에 존재하는 정확한 Kozak 서열은 번역 개시의 효율성을 결정한다. 즉, 특정 Kozak 서열은 효율적인 번역 개시로 이어지지 않는다.
- [0468] 전체 Kozak 서열은 전형적으로 DNA의 경우 (gcc)gccRccATGG 및 RNA의 경우 (gcc)gccRccAUGG인 컨센서스 서열을 갖는 것으로 이해되며, 여기서; 소문자는 이 위치에서 가장 일반적인 염기를 나타내고 이 위치의 염기는 변할 수 있으며; 대문자는 이 위치에서 고도로 보존된 염기를 나타내고; "R"은 퓨린(즉, A 또는 G)이 전형적으로 이 위치에서 최적임을 나타내고; 괄호 안의 서열 (gcc)는 불확실한 의미를 가진다. T/U는 일반적으로 개시 코돈의 상류에 있는 Kozak 서열 컨센서스의 모든 위치에서 가장 덜 선호되는 뉴클레오티드이다.
- [0469] 전체 Kozak 서열의 처음 3개 염기가 불확실한 의미를 가지므로, Kozak 서열은 본원에서 "확장된 Kozak 서열"로 지칭되는 컨센서스 서열을 가지는 것으로 또한 이해될 수 있고, DNA의 경우 GNNRVVATGG(서열번호 103) 및 RNA의 경우 GNNRVVAUGG이며, 여기서 "R"은 서열의 해당 위치에서 퓨린(즉, A 또는 G)을 지정하는 것으로 이해되어야

하고, "V"는 G, A, 또는 C로부터의 임의의 뉴클레오타드를 지정하는 것으로 이해되어야 하며, "N"은 서열의 해당 위치에서 임의의 뉴클레오타드를 지정하는 것으로 이해되어야 한다. 예를 들어, "N"은 G, A, T, C 또는 U일 수 있다. 위치 -1의 "R" 및 위치 +3의 "G"(위치 0에 해당하는 ATG의 "A"에 대하여)는 Kozak 강도의 측면에서 가장 중요한 위치로 간주됨을 유의해야 한다. 그러나, 전이유전자 서열의 +3 위치에 "G"가 있는지 여부는 암호화되는 ORF에 따라 다르므로, 사양의 목적상 +3 위치는 '코어' Kozak 서열의 일부로 간주되지 않는다.

- [0470] 전체 Kozak 서열의 처음 6개 위치에서 발견된 염기는 모든 염기가 해당 위치에서 발견될 수 있을 정도로 다양하다(위에서 gccgcc로 표시됨). 따라서, 전체 Kozak 컨센서스 서열은 위의 RccAUG로 표시된 감소된 가변성을 가지는 전체 Kozak 서열의 일부로 이루어진 '코어' Kozak 서열을 포함하는 것으로 간주될 수 있다. '코어' Kozak 컨센서스 서열은 본원에서 mRNA의 경우 RVVAUG 및 DNA의 경우 RVVATG로 정의되며, 여기서 "R"은 서열의 해당 위치에서 퓨린(즉, A 또는 G)을 지정하는 것으로 이해되어야 하고 "V"는 G, A 또는 C로부터의 임의의 뉴클레오타드를 지정하는 것으로 이해되어야 한다.
- [0471] 본 발명의 바람직한 일 구현예에서, Kozak 서열은 서열 RVVATG(서열번호 104)를 포함하고; 여기서 "R"은 서열의 해당 위치에서 퓨린(즉, A 또는 G)을 지정하는 것으로 이해되어야 하고 "V"는 G, A 또는 C로부터의 임의의 뉴클레오타드를 지정하는 것으로 이해되어야 한다.
- [0472] 본 발명의 일 구현예에서, Kozak 서열은 서열 RNNATG를 포함하고; 여기서 "R"은 서열의 해당 위치에서 퓨린(즉, A 또는 G)을 지정하는 것으로 이해되어야 하고 "N"은 G, A, T/U 또는 C로부터의 임의의 뉴클레오타드를 지정하는 것으로 이해되어야 하며, "T/U"의 사용은 TRAP가 없을 때 발현 수준을 감소시킬 수 있음을 인식해야 한다.
- [0473] 일부 구현예에서, Kozak 서열은 TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부와 중첩된다. 따라서, 코어 Kozak 서열은 TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부와 중첩될 수 있다.
- [0474] 바람직한 중첩되는 tbs 및 Kozak 컨센서스 서열의 요약을 도 19C에 나타내었다.
- [0475] 바람직한 구현예에서, TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부는 코어 Kozak 서열 내의 ATG 트리플렛의 적어도 첫 번째 또는 처음 두 개의 뉴클레오타드와 중첩된다.
- [0476] 본원에 기재된 바와 같이, 본 발명의 일 측면에서, TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부는 관심 뉴클레오타드(전이유전자 ORF)의 ATG 시작 코돈과 중첩된다. 일 측면에서 TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부는 관심 뉴클레오타드(전이유전자 ORF)의 ATG 시작 코돈의 처음 1개 또는 2개와 중첩된다.
- [0477] 일 측면에서 중첩되는 tbs-Kozak 서열은 컨센서스 서열 KAGNNG(서열번호 113)를 가질 수 있으며, 여기서 "NN"은 Kozak 서열의 ATG 트리플렛 내의 처음 2개의 뉴클레오타드이다.
- [0478] 컨센서스 서열은 KAGATG(서열번호 114)일 수 있고; 여기서 "K"는 G 또는 T/U이다.
- [0479] 일 측면에서, 중첩되는 tbs-Kozak 서열은 도 19C에 도시된 바와 같이 GAGATG(서열번호 142)일 수 있다.
- [0480] 일 구현예에서, TRAP 결합 부위 또는 그의 일부의 3' 말단 KAGNN 반복부는 관심 뉴클레오타드 내의 ATG 트리플렛의 첫 번째 뉴클레오타드와 중첩된다.
- [0481] 일 측면에서 서열은 서열 KAGNNTG(서열번호 115)를 포함할 수 있고, 여기서 두 번째 "N"은 ATG 트리플렛 내의 첫 번째 뉴클레오타드이다. 컨센서스 서열은 KAGNATG(서열번호 116)일 수 있고; 여기서 "K"는 서열의 해당 위치에서 G 또는 T/U를 지정하는 것으로 이해되어야 하고, "N"은 G, A, T, U 또는 C로부터의 임의의 뉴클레오타드를 지정하는 것으로 이해되어야 하지만 바람직하게는 "V"이고, 즉, G, A 또는 C로부터이다. 예를 들어, 중첩되는 서열은 도 19C에 도시된 바와 같이 KAGVATG(서열번호 143)일 수 있다.
- [0482] 일 구현예에서, 본 발명에서 사용하기 위한 중첩되는 Kozak 서열 및 TRAP 결합 부위 또는 그의 일부는 서열번호 142-146에 기재된 서열을 포함한다.
- [0483] 일 구현예에서, 본 발명의 뉴클레오타드 서열은 서열번호 142-146에 기재된 서열을 포함한다. 일 측면에서 본 발명의 서열은 서열 KAGATG를 포함한다.
- [0484] 본 발명의 핵산에 사용하기 위한 컨센서스 서열 GAGATG(서열번호 142), KAGVATG(서열번호 143) 및 KAGVVATG(서열번호 144)에 해당하는 바람직한 중첩되는 tbs 또는 그의 일부 및 코어 Kozak 서열(본원에 정의된 바와 같은 KAGNN의 컨센서스 tbs 반복부 서열 및 본원에 정의된 바와 같은 컨센서스 '코어' Kozak 서열 RVVATG를 기반으로

함)은 서열번호 142 및 69-92에 기재된 서열을 포함한다.

- [0485] 일부 구현예에서, 뉴클레오타이드 서열은 서열번호 147-150에 기재된 서열을 포함한다. 바람직하게는, 뉴클레오타이드 서열은 서열번호 147 또는 서열번호 148에 기재된 서열을 포함한다.
- [0486] 바람직한 구현예에서, 본 발명의 뉴클레오타이드 서열은 서열번호 106에 기재된 중첩되는 tbs 및 Kozak 서열을 포함한다.
- [0487] TRIP 시스템의 취급성(tractability)을 개선하기 위해 여러 다른 제한 효소(RE)의 옵션을 통해, 즉 tbs와 Kozak 서열 사이에 다중 클로닝 부위(MCS)를 통합하여 프로모터-5'UTR-tbs 서열을 포함하는 발현 카세트에 NOI를 직접 클로닝하는 능력을 가질 수 있는 것이 바람직하다(도 19A 참조). 본 발명자들은 여러 상이한 MCS가 TRIP 시스템에 의해 허용될 수 있다는 것을, 즉 MCS가 사용될 때 전이유전자 억제에 여전히 얻어진다는 것을 입증하였다. 이는 5'UTR 리더 서열이 TRAP 매개 억제의 정도를 조절할 수 있고, ATG 개시 코돈에 대한 tbs의 근접성이 중요하고, 효율적인 번역 개시를 보장하기 위해서는 효율적인 Kozak 서열이 MCS와 NOI의 개시 코돈과 함께 또는 그 사이에서 유지되어야 한다는 점을 감안할 때 예상치 못한 일이었다. 또한, TRAP 매개 억제를 유지하면서 사용할 수 있는 RE 부위의 수 및/또는 조합을 예측할 수 없었다.
- [0488] 이와 같이, 여러 (중첩되는) RE 부위가 tbs에서부터 ATG 개시 코돈까지 가능한 한 짧은 거리에 통합될 수 있도록 하는 한편(ATG에 대한 tbs의 근접성을 유지하기 위해) RVVATG의 효율적인 코어 Kozak 서열을 유지하도록 하는 서열 '압축'이 필요하였다.
- [0489] 일 구현예에서, 뉴클레오타이드 서열은 본원에 기재된 바와 같은 tbs 또는 그의 일부, 다중 클로닝 부위(MCS) 및 본원에 기재된 바와 같은 Kozak 서열을 추가로 포함하며, 여기서 상기 MCS는 tbs 또는 그의 일부의 하류 및 Kozak 서열의 상류에 위치한다. 적당하게는, tbs 또는 그의 일부와 Kozak 서열은 중첩되지 않는다.
- [0490] 본원에서 사용되는 바와 같이, "다중 클로닝 부위"는 서로 매우 근접한 여러 제한 효소 인식 부위(제한 효소 부위)를 포함하는 DNA 영역으로 이해되어야 한다. 일 구현예에서, RE 부위는 본 발명에서 사용하기 위한 MCS에서 중첩될 수 있다.
- [0491] 본원에서 사용되는 바와 같이, "제한 효소 부위" 또는 "제한 효소 인식 부위"는 제한 효소에 의해 인식되는 4-8 개 뉴클레오타이드 길이의 뉴클레오타이드의 특정 서열을 함유하는 DNA 분자 상의 위치이다. 제한 효소는 특정 RE 부위(즉, 특정 서열)를 인식하고 RE 부위 내부 또는 근처에서 DNA 분자를 절단한다.
- [0492] 일 구현예에서, 뉴클레오타이드 서열은 서열번호 158-171에 기재된 서열을 포함한다.
- [0493] 일 구현예에서, 뉴클레오타이드 서열은 서열번호 165-171에 기재된 서열을 포함한다.
- [0494] 일 구현예에서, 뉴클레오타이드 서열은 서열번호 165, 168 또는 171에 기재된 서열을 포함한다.
- [0495] 바람직한 구현예에서, 본 발명의 뉴클레오타이드 서열은 서열번호 107에 기재된 중첩되는 tbs-MCS-Kozak 서열을 포함한다.
- [0496] 일 구현예에서, Kozak 서열은 서열 RNNATG를 포함하고; 여기서 "R"은 서열의 해당 위치에서 퓨린(즉, A 또는 G)을 지정하는 것으로 이해되어야 하고, "N"은 G, A, T/U 또는 C로부터의 임의의 뉴클레오타이드를 지정하는 것으로 이해되어야 한다.
- [0497] NOI 번역의 억제 또는 방지는 바이러스 벡터 생산 동안 번역되는 NOI 산물(예: 단백질)의 양을 동등한 시점에서 본 발명의 뉴클레오타이드 서열이 없을 때 발현되는 양과 비교하여 변경하는 것으로 이해되어야 한다. 번역의 이러한 변경은 결과적으로 NOI에 의해 암호화되는 단백질 발현의 억제 또는 방지를 초래한다.
- [0498] 일 구현예에서, 본 발명의 뉴클레오타이드 서열은 관심 뉴클레오타이드의 번역이 바이러스 벡터 생산 세포에서 억제되거나 방지되도록 TRAP와 상호작용할 수 있다.
- [0499] 벡터 생산 동안 임의의 주어진 시간에 NOI의 번역은 벡터 생산의 동일한 시점에서 본 발명의 뉴클레오타이드 서열의 부재 하에 번역된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1%로 감소될 수 있다.
- [0500] 벡터 생산 동안 임의의 주어진 시간에 NOI의 번역은 벡터 생산의 동일한 시점에서 본 발명의 뉴클레오타이드 서열의 부재 하에 번역된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1% 미만으로 감소될 수 있다.

- [0501] 본 발명의 맥락에서, 벡터 생산 동안 임의의 주어진 시간에 NOI의 번역은 벡터 생산의 동일한 시점에서 (본 발명의 뉴클레오티드 서열과 대조적으로) 중첩되지 않는 Kozak 및 tbs 서열을 포함하는 핵산 서열의 존재 하에 번역된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1%로 감소될 수 있다.
- [0502] 본 발명의 맥락에서 벡터 생산 동안 임의의 주어진 시간에 NOI의 번역은 벡터 생산의 동일한 시점에서 (본 발명의 뉴클레오티드 서열과 대조적으로) 중첩되지 않는 Kozak 및 tbs 서열을 포함하는 핵산 서열의 존재 하에 번역된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1% 미만으로 감소될 수 있다.
- [0503] NOI의 번역을 방지하는 것은 번역의 양을 실질적으로 0으로 줄이는 것으로 이해되어야 한다.
- [0504] 벡터 생산 동안 임의의 주어진 시간에 NOI로부터 단백질의 발현은 벡터 생산의 동일한 시점에서 본 발명의 뉴클레오티드 서열의 부재 하에 발현된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1%로 감소될 수 있다.
- [0505] 벡터 생산 동안 임의의 주어진 시간에 NOI로부터 단백질의 발현은 벡터 생산의 동일한 시점에서 본 발명의 뉴클레오티드 서열의 부재 하에 발현된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1% 미만으로 감소될 수 있다.
- [0506] 본 발명의 맥락에서 벡터 생산 동안 임의의 주어진 시간에 NOI로부터 단백질의 발현은 벡터 생산의 동일한 시점에서 (본 발명의 뉴클레오티드 서열과 대조적으로) 중첩되지 않는 Kozak 및 tbs 서열을 포함하는 핵산 서열의 존재 하에 발현된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1%로 감소될 수 있다.
- [0507] 본 발명의 맥락에서 벡터 생산 동안 임의의 주어진 시간에 NOI로부터 단백질의 발현은 벡터 생산의 동일한 시점에서 (본 발명의 뉴클레오티드 서열과 대조적으로) 중첩되지 않는 Kozak 및 tbs 서열을 포함하는 핵산 서열의 존재 하에 발현된 양의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1% 미만으로 감소될 수 있다.
- [0508] NOI로부터 단백질의 발현을 방지하는 것은 발현되는 단백질의 양을 실질적으로 0으로 감소시키는 것으로 이해되어야 한다.
- [0509] NOI 번역의 분석 및/또는 정량화 방법은 당업계에 잘 알려져 있다.
- [0510] 용해된 세포로부터의 단백질 산물은 쿠마시(Coomassie) 또는 은 염색에 의한 시각화와 함께 SDS-PAGE 분석과 같은 방법을 사용하여 분석할 수 있다. 대안적으로 단백질 산물은 단백질 산물에 결합하는 항체 프로브를 사용한 웨스턴 블롯팅 또는 효소 결합 면역흡착 분석법(ELISA)을 사용하여 분석할 수 있다. 손상되지 않은 세포의 단백질 산물은 면역형광법으로 분석할 수 있다.
- [0511] 일 구현예에서, 전사 시작 부위/프로모터의 말단에서 tbs 또는 그의 일부의 시작 사이의 거리는 34개 뉴클레오티드 미만이다.
- [0512] 일 구현예에서, 전사 시작 부위/프로모터의 말단에서 tbs 또는 그의 일부의 시작 사이의 거리는 13개 뉴클레오티드 미만이다.
- [0513] 일 구현예에서, 뉴클레오티드 서열은 벡터 전이유전자 발현 카세트이다.
- [0514] 일 구현예에서, 본 발명의 핵산 서열은 프로모터를 추가로 포함한다. 전형적으로, 프로모터의 전사는 생성된 mRNA 전사체에 암호화된 5' UTR을 초래한다. 프로모터는 당해 분야에 공지되어 있고 관심 뉴클레오티드의 발현을 제어하기에 적합한 임의의 프로모터일 수 있다. 예를 들어, 프로모터는 EF1 α , EFS, CMV 또는 CAG일 수 있다.
- [0515] 바람직한 구현예에서, 본원에 기재된 바와 같은 중첩되는 tbs 및 Kozak 서열은 프로모터의 5' UTR 내에 위치하며, 여기서 5' UTR은 연관된 프로모터로부터의 천연 서열을 포함할 수 있거나, 보다 바람직하게는 5' UTR은 본원에 기재된 5' UTR 서열로 구성된다.
- [0516] 바람직한 구현예에서, 본원에 기재된 바와 같은 tbs와 Kozak 서열 사이에 압축된/중첩되는 MCS를 포함하는 서열은 상기 5' UTR 내에 위치한다.

- [0517] 본원에 기재된 바와 같은 중첩되는 tbs 및 Kozak 서열 또는 본원에 기재된 바와 같은 tbs와 Kozak 서열 사이에 압축된/중첩되는 MCS를 포함하는 서열은 상기 5' UTR의 3' 말단에 위치할 수 있다.
- [0518] 바람직하게는, 5' UTR은 하기 서열 중 하나를 포함한다: 서열번호 29-37, 45-58, 69-92 및 108-116. 더욱 바람직하게는, 5' UTR은 서열번호 29 또는 서열번호 108을 포함한다. 훨씬 더 바람직하게는 5' UTR은 서열번호 29를 포함한다.
- [0519] 프로모터-5' UTR 영역은 인트론을 포함할 수 있다. 인트론은 천연 인트론 또는 이중 인트론일 수 있다. 예를 들어, 프로모터는 EF1 α 또는 CAG일 수 있다.
- [0520] 프로모터는 인트론이 없는 바이러스 벡터 계통에서 전형적으로 사용되는 프로모터, 예를 들어 CMV일 수 있다.
- [0521] 바람직한 구현예에서, 프로모터-5' UTR 영역은 이중 인트론을 포함하는 인공 5' UTR을 포함하도록 조작되었다. 따라서, 프로모터-5' UTR 영역은 이중 엑손-인트론-엑손 서열을 포함하도록 조작되었으며, 여기서 mRNA 전사체 내에 암호화된 성숙한 5' UTR은 인트론의 스플라이싱-아웃으로부터 발생한다. 프로모터-5' UTR 서열은 당업계에 공지된 방법을 사용하여 조작될 수 있다. 예를 들어, 프로모터는 본원에 기재된 바와 같이 조작될 수 있다.
- [0522] 바람직하게는, 인트론 또는 이중 인트론의 스플라이싱-아웃으로 인한 성숙한 mRNA로부터의 전이유전자 단백질의 발현은 TRAP에 의해 효율적으로 억제된다. 적합하게는, 인트론 또는 이중 인트론은 서열번호 122에 따른 EF1 α 인트론 서열일 수 있다.
- [0523] 인트론 또는 이중 인트론은 본원에 기재된 바와 같은 중첩되는 tbs 및 Kozak 서열 또는 본원에 기재된 바와 같은 tbs와 Kozak 서열 사이에 MCS를 포함하는 서열의 상류, 즉 5'에 위치할 수 있다.
- [0524] 5' UTR은 하기 서열을 포함할 수 있다(닭 β -액틴/토끼 β -글로빈 키메라 5'UTR-인트론, 엑손 서열은 굵게 표시 (함께 스플라이싱되어 5'UTR 리더가 됨)):
- [0525] 5' UTR은 서열번호 121에 기재된 서열을 포함할 수 있다.
- [0526] 5' UTR은 서열번호 122에 기재된 서열을 포함할 수 있다.
- [0527] 일 구현예에서, 프로모터는 서열번호 123에 기재된 서열을 포함한다.
- [0528] 일 구현예에서, 프로모터는 서열번호 124에 기재된 서열을 포함한다.
- [0529] 일 구현예에서, 프로모터는 서열번호 117에 기재된 서열을 포함한다.
- [0530] 일 구현예에서, 프로모터는 서열번호 118에 기재된 서열을 포함한다.
- [0531] 서열번호 117에 상응하는 스플라이싱된 서열은 서열번호 119에 기재되어 있다.
- [0532] 서열번호 118에 상응하는 스플라이싱된 서열은 서열번호 120에 기재되어 있다.
- [0533] **개선된 리더 서열**
- [0534] TRIP 시스템을 다양한 프로모터(서로 다른 길이 및 조성의 다양한 천연 5'UTR을 함유함)에 적용할 때, TRAP에 의한 효율적인 억제를 제공하면서도 또한 TRAP 없이 우수한 수준의 발현을 유지하기 위해서는 프로모터-UTR 컨텍스트 내에서 tbs 서열을 간단하게 적용할 수 있는 것이 바람직하다. 이 작업의 시초에는, tbs가 다양한 구성적 프로모터의 천연 UTR에 삽입될 때 TRAP-tbs에 의해 매개되는 달성 가능한 억제 수준이 무엇인지 알려지지 않았다. 이상적으로는, 천연 5'UTR 서열에 의해 지시될 수 있는 억제 수준의 잠재적인 변동성을 피하기 위해 선택한 프로모터를 수정할 때 tbs와 함께 단일의 보존된 5'UTR 리더 서열을 제공할 수 있는 것이 유리할 것이다. 놀랍게도, EF1 α 프로모터의 첫 번째 엑손(서열번호 101)이 천연의 리더 서열을 포함하는 5'UTR 리더와 비교하여 TRAP에 의해 일관되게 우수한 수준의 전이유전자 억제를 제공하고, 이 리더는 또한 TRAP가 없는 경우 우수한 수준의 전이유전자 발현을 제공한다는 것이 발견되었다.
- [0535] 일부 구현예에서, 뉴클레오타이드 서열은 tbs 또는 그의 일부의 상류에 5' 리더 서열을 포함한다. 리더 서열은 TRAP 결합 부위 또는 그의 일부의 바로 상류에 있을 수 있으며, 즉 리더 서열과 TRAP 결합 부위 또는 그의 일부를 분리하는 추가의 서열이 없을 수 있다. 5' 리더가 스플라이싱 이벤트에서 파생된 경우, 엑손/엑손 접합에서 tbs까지의 서열은 최소 길이(바람직하게는 $\leq 12\text{nt}$)로 유지되어야 한다. 리더 서열은 비-코딩 EF1 α 엑손 1 영역으로부터 유래된 서열을 포함할 수 있다. 바람직한 구현예에서, 리더 서열은 서열번호 101, 서열번호 102 또는 서열번호 93에 정의된 바와 같은 서열을 포함한다.

- [0536] **예시적인 뉴클레오타이드 서열**
- [0537] 본 발명의 예시적인 뉴클레오타이드 서열을 하기에 나타내었다.
- [0538] 서열번호 172 - L33 개선된 리더, 최적의 (중첩되는) *tbs* ([KAGNN]₈)-**Kozak** 집합을 포함하는 예시적인 뉴클레오타이드 서열 1
- [0539] CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCGAGATG
- [0540] 서열번호 173 - L33 개선된 리더, 최적의 (중첩되는) *tbs* ([KAGNN]₁₁)-**Kozak** 집합을 포함하는 예시적인 뉴클레오타이드 서열 2
- [0541] CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTAGCAGAGACGAGCCGAGATG
- [0542] 서열번호 174 - L12 개선된 리더, 최적의 (중첩되는) *tbs* ([KAGNN]₁₁)-**Kozak**을 포함하는 예시적인 뉴클레오타이드 서열 3
- [0543] CTTTTTCGCAACGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTAGCAGAGACGAGCCGAGATG
- [0544] 서열번호 175 - L33, 최적의 (중첩되는) *tbs* ([KAGNN]₁₁)-**Kozak** 집합을 포함하는 스플라이싱된 리더를 생성하는 인트론 함유 5'UTR에 대한 예시적인 뉴클레오타이드 서열 4
- [0545] CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTGCTGAAAAAGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTAGCAGAGACGAGCCGAGATG
- [0546] 서열번호 176 - 개선된 스페이서, 최적의 (중첩되는) *tbs* ([KAGNN]₈)-**Kozak** 집합을 포함하는 예시적인 뉴클레오타이드 서열 5
- [0547] ATAGCAGAGACGGCTGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCGAGATG
- [0548] 서열번호 177 - L33 개선된 리더 *tbs* ([KAGNN]₁₁)-MCS-**Kozak**을 포함하는 예시적인 뉴클레오타이드 서열 6
- [0549] CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTAGCAGAGACGAGAAGAGCTCTAGACCATG
- [0550] 서열번호 178 - 개선된 스페이서, *tbs* ([KAGNN]₁₁)-MCS-**Kozak**을 포함하는 예시적인 뉴클레오타이드 서열 7
- [0551] ATAGCAGAGACGGCTGAGTTTAGCGGAGTGGAGAAGAGCGGAGCCGAGCCTAGCAGAGACGAGAAGAGCTCTAGACCATG
- [0552] 일 구현예에서, 뉴클레오타이드 서열은 서열번호 172-178 중 어느 하나를 포함한다.
- [0553] 일 구현예에서, 뉴클레오타이드 서열은 다음을 포함한다:
- [0554] (a) (i) 서열번호 101 또는 102; 및/또는
- [0555] (ii) 서열번호 151-157 중 어느 하나;
- [0556] (b) 서열번호 127-130, 132-136, 140, 141 중 어느 하나; 및
- [0557] (c) (i) 서열번호 142-149 중 어느 하나, 바람직하게는 서열번호 147-149 중 어느 하나; 또는
- [0558] (ii) 서열번호 158-171 중 어느 하나, 바람직하게는 서열번호 165-171 중 어느 하나.
- [0559] 일 구현예에서, 뉴클레오타이드 서열은 다음을 포함한다:
- [0560] (a) 서열번호 101 또는 102;
- [0561] (b) 서열번호 127-130, 133-136, 140, 141 중 어느 하나; 및
- [0562] (c) 서열번호 142-149 중 어느 하나, 바람직하게는 서열번호 147-149 중 어느 하나.
- [0563] 일 구현예에서, 뉴클레오타이드 서열은 다음을 포함한다:
- [0564] (a) 서열번호 151-157 중 어느 하나;
- [0565] (b) 서열번호 127-130, 133-136, 140, 141 중 어느 하나; 및
- [0566] (c) 서열번호 142-149 중 어느 하나, 바람직하게는 서열번호 147-149 중 어느 하나.

- [0567] 일 구현예에서, 뉴클레오티드 서열은 다음을 포함한다:
- [0568] (a) 서열번호 101 또는 102;
- [0569] (b) 서열번호 127-130, 132-136, 140, 141 중 어느 하나; 및
- [0570] (c) 서열번호 157-171 중 어느 하나, 바람직하게는 서열번호 165-171 중 어느 하나.
- [0571] 일 구현예에서, 뉴클레오티드 서열은 다음을 포함한다:
- [0572] (a) 서열번호 157-171 중 어느 하나;
- [0573] (b) 서열번호 127-130, 132-136, 140, 141 중 어느 하나; 및
- [0574] (c) 서열번호 157-171 중 어느 하나, 바람직하게는 서열번호 165-171 중 어느 하나.
- [0575] 본 발명의 실시는 달리 명시되지 않는 한, 당업자의 능력 범위 내에 있는 화학, 분자 생물학, 미생물학 및 면역학의 통상적인 기술을 사용할 것이다. 이러한 기술은 문헌에 설명되어 있다. 예를 들어, 하기 문헌 참조: J. Sambrook, E. F. Fritsch, and T. Maniatis (1989) *Molecular Cloning: A Laboratory Manual*, Second Edition, Books 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F. M. et al. (1995 and periodic supplements) *Current Protocols in Molecular Biology*, Ch. 9, 13, and 16, John Wiley & Sons, New York, NY; B. Roe, J. Crabtree, and A. Kahn (1996) *DNA Isolation and Sequencing: Essential Techniques*, John Wiley & Sons; J. M. Polak and James O'D. McGee (1990) *In Situ Hybridization: Principles and Practice*; Oxford University Press; M. J. Gait (ed.) (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press; and, D. M. J. Lilley and J. E. Dahlberg (1992) *Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology*, Academic Press. 이들 일반 교과서 각각은 본원에 참고로 포함된다.
- [0576] 본 개시내용은 본원에 개시된 예시적인 방법 및 물질에 의해 제한되지 않으며, 본원에 기재된 것과 유사하거나 동등한 임의의 방법 및 물질이 본 개시내용의 구현예의 실시 또는 테스트에서 사용될 수 있다. 수치 범위 (numeric range)에는 범위를 한정하는 수치가 포함된다. 달리 명시되지 않는 한, 모든 핵산 서열은 5'에서 3' 방향으로 왼쪽에서 오른쪽으로 기재되며; 아미노산 서열은 각각 아미노에서 카르복시 방향으로 왼쪽에서 오른쪽으로 기재된다.
- [0577] 값(value)의 범위가 제공되는 경우, 문맥에서 명확하게 달리 지시하지 않는 한, 해당 범위의 상한과 하한 사이의 각 중간 값도 하한 단위의 10분의 1까지 구체적으로 개시되는 것으로 이해된다. 명시된 범위의 임의의 명시된 값 또는 중간 값과 해당 명시된 범위의 임의의 다른 명시된 값 또는 중간 값 사이의 각각의 더 작은 범위는 본 개시내용 내에 포함된다. 이들 더 작은 범위의 상한 및 하한은 독립적으로 범위에 포함되거나 배제될 수 있으며, 어느 한 한계, 둘 다 포함되지 않거나, 둘 모두가 더 작은 범위에 포함되는 각 범위는 또한 본 개시내용에 포함되며, 명시된 범위에서 구체적으로 제외된 한계의 적용을 받는다. 명시된 범위가 한계 중 하나 또는 둘 다를 포함하는 경우, 포함된 한계 중 하나 또는 둘 모두를 배제하는 범위도 본 개시내용에 포함된다.
- [0578] 본 명세서 및 첨부된 청구범위에서 사용되는 단수형 "a", "an", 및 "the"는 문맥이 명백하게 달리 지시하지 않는 한 복수의 지시대상을 포함한다는 점에 유의해야 한다.
- [0579] 본원에서 사용되는 용어 "포함하는(comprising)", "포함하다(comprise)" 및 "~를 포함하는(comprised of)"은 "포함하는(including)", "포함하다(include)" 또는 "함유하는(containing)", "함유하다(contain)"와 동의어이며, 포괄적이거나 개방형(open-ended)이며 추가의 명시되지 않은 구성원, 요소 또는 방법 단계를 배제하지 않는다. 용어 "포함하는(comprising)", "포함하다(comprise)" 및 "~를 포함하는(comprised of)"은 또한 "~로 구성되는(consisting of)"이라는 용어를 포함한다.
- [0580] 본원에 논의된 간행물은 본 출원의 출원일 이전에 공개된 점을 위해서만 제공된다. 본원에 있는 어떤 것도 그러한 간행물이 여기에 첨부된 청구범위에 대한 선행 기술을 구성한다는 것을 인정하는 것으로 해석되어서는 안된다.
- [0581] 본 발명은 이제 실시예를 통해 추가로 설명될 것이며, 이는 당업자가 본 발명을 수행하는 데 도움을 주기 위한 것이며 어떤 식으로든 본 발명의 범위를 제한하도록 의도되지 않는다.
- [0582] **실시예**

- [0583] 일반 분자/세포 생물학 기술 및 아세이
- [0584] 변형된 U1 snRNA 발현 작제물
- [0585] 변형된 U1 snRNA에 대한 DNA 기반 발현 작제물은 RNA 전사 및 종결을 구동하는 내인성 U1 snRNA 유전자의 보존된 서열을 포함하며, 이는 256U1(U1_256이라고도 함) snRNA의 비제한적인 예에서 아래에 강조표시되어 있다:
- [0586] TAAGGACCAGCTTCTTTGGGAGAGAACAGACGACGGGGCGGGAGGGAAAAAGGGAGAGGCAGACGTCACCTTCCCCTTGGCGGCTCTGGCAGCAGATTGGTCG GTTGAAGTGGCAGAAAGGCAGACGGGGACTGGGCAAGGCACTGTCGGTGACATCACGGACAGGGCGACTTCTATGTAGATGAGGCAGCGCAGAGGCTGCTGCT TCGCCACTTGCTGCTTACCACGAAGGAGTTCCCGTGCCCTGGGAGCGGGTTTCAGGACCGCTGATCGGAAGTGAGAATCCCAGCTGTGTGTCAGGGCTGGAA AGGGCTCGGGAGTGC GCGGGGCAAGTGACCGTGTGTGTAAGAGTGAGGCGTATGAGGCTGTGTCGGGGCAGAGGCCAAGATCTCat **ttgccgtgcgcgct** **tGCAGGGGAGATACCATGATCACGAAGGTGGTTTTCCAGGGCGAGGCTTATCCATTGCACTCCGGATGTGCTGACCCCTGCGATTTCCTCCAAATGTGGGAA ACTCGACTGCATAATTTGTGGTAGTGGGGGACTGCGTTCGCGCTTTCCCTGTTTCAAAAGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTTTGGT TTGTGCTTTGGTTGGCGCTCTAAATGTTAA** (서열번호 15)
- [0587] 기호 풀이(Key): 대문자만 = U1 PolII 프로모터(nt1-392); 소문자 = 재표적화 영역(nt393-409); 소문자 볼드체 = 재표적화 서열[이 예에서 야생형 HIV-1 패키징 신호의 nt256-270을 표적화](nt395-409); 대문자 이탤릭체 = 주요 U1 snRNA 서열[클로버잎](nt410-562); 대문자 밑줄 = 전사 종결 영역(nt563-652).
- [0588] 연구에 사용된 초기의 변형된 U1 snRNA 및 대조군에 대한 요약이 하기 표에 제시되어 있으며, 이는 새로운 어닐 링 서열 및 표적 부위 서열을 나타낸다(서열은 5'에서 3' 방향으로 표시됨).
- [0589] **표 I.** 초기 연구에 사용된 테스트 변형된 U1 snRNA 및 대조군 U1 snRNA 내의 표적-어닐링 서열(표적 서열에 상보적인 이중 서열) 및 이들의 표적 서열을 설명하는 서열 목록. 뉴클레오티드는 '재표적화 영역'에서 각각의 발현 카세트 내에서 암호화되는 DNA로 표시된다. (AT) 모티프는 모든 초기 작제물에 존재하였으며, 이는 각 경우에 U1 snRNA 분자의 처음 두 개의 뉴클레오티드를 형성한다. 표적 서열 번호는 표시된 곳에서 HIV-1의 NL4-3(GenBank: M19921.2) 또는 HXB2(GenBank: K03455.1) 균주 내의 표적을 지칭하는데, 이는 이 연구의 렌티바이러스 벡터 계통에는 이 두 가지 고도로 보존된 균주로 구성된 하이브리드 패키징 신호가 포함되어 있었기 때문이다(이 연구에 사용된 패키징 서열은 GenBank: MH782475.1의 벡터 서열과 가장 유사하다)

변형된 U1 snRNA*	HIV-1 표적 서열 [NL4-3]**	U1 snRNA 표적-어닐링 서열
U1_16	16-GACCAGATCTGAGCC-30 (서열번호 16)	(AT) GGCTCAGATCTGGTC (서열번호 17)
U1_31	31-TGGGAGCTCTCTGGC-45 (서열번호 18)	(AT) GCCAGAGAGCTCCCA (서열번호 19)
U1_76	76-TAAAGCTTGCCTTGA-90 (서열번호 20)	(AT) TCAAGGCAAGCTTTA (서열번호 21)
U1_136	136-TAGAGATCCCTCAGA-150 (서열번호 22)	(AT) TCTGAGGGATCTCTA (서열번호 23)
U1_179(9nt)	179-GCAGTGGCG-187 (서열번호 24)	(AT) CGCCACTGC (서열번호 25)
U1_181	181-AGTGGCGCCCGAACA-195 (서열번호 26)	(AT) TGTTCGGGCGCCACT (서열번호 27)
U1_196	196-GGGACTTGAAAGCGA-210 (서열번호 28)	(AT) TCGCTTTCAAGTCCC (서열번호 29)
U1_211	211-AAGggAAaCCAGAGG-225 (서열번호 30)	(AT) CCTCTGGTTTCCCTT (서열번호 31)
U1_226	226-AGcTCTCTCGACGCA-240 (서열번호 32)	(AT) TGCCTCGAGAGAGCT (서열번호 33)
U1_241	241-GGACTCGGCTTGCTG-255 (서열번호 34)	(AT) CAGCAAGCCGAGTCC (서열번호 35)
U1_256	256-AAGCGCGCACGGCAA-270 (서열번호 36)	(AT) TTGCCGTGCGCGCTT (서열번호 37)
U1_271	271-GAGGCGAGGGGCGGC-285 (서열번호 38)	(AT) GCCGCCCCCTCGCCTC (서열번호 39)
U1_286	286-GACTGGTGAGTACGC-300 (서열번호 40)	(AT) GCGTACTCACCAGTC (서열번호 41)
U1_305(9nt)	305-AATTTTGAC (TA) -313/5 (서열번호 42)	(AT) GTCAAAATT (서열번호 43)
U1_305	305-AATTTTGACTAGCGG-319 (서열번호 44)	(AT) CCGCTAGTCAAAATT (서열번호 45)
U1_316	316-GCGGAGGCTAGAAGG-330 (서열번호 46)	(AT) CCTTCTAGCCTCCGC (서열번호 47)
U1_331	331-AGAGAGATGGGTGCG-345 (서열번호 48)	(AT) CGCACCCTATCTCTCT (서열번호 49)
U1_346	346-AGAGCGTCgGTATTA-360 (서열번호 50)	(AT) TAATACTGACGCTCT (서열번호 51)
U1_361	361-AGCGGGGAGAATTA-375 (서열번호 52)	(AT) TAATTCTCCCCGCT (서열번호 53)
U1_376	376-GATCGCGATGGGAAA-390 (서열번호 54)	(AT) TTTCCCATCGCGATC (서열번호 55)

[0590]

U1_391	389-AAATTCGGTTAAGGC-403 (서열번호 54)	(AT) GCCTTAACCGAATTT (서열번호 55)
U1_690	7159-GATCTTCAGACCTGG-7173 (서열번호 56)	(AT) CCAGGTCTGAAGATC (서열번호 57)
U1_1203	7672-TTACACAAGCTTAAT-7686 (서열번호 58)	(AT) ATTAAGCTTGTGTAA (서열번호 59)
U1_1546	4375-TAGTAGACATAATAG-4389 (서열번호 60)	(AT) CTATTATGTCTACTA (서열번호 61)
대조군 U1 snRNA	표적 서열	U1 snRNA 표적-어닐링 서열
U1_LacZ1	388-CTACAGGAA-396 (서열번호 62)	(AT) TTCCTGTAG (서열번호 63)
U1_LacZ2	438-TCATCTGTG-446 (서열번호 64)	(AT) CACAGATGA (서열번호 65)

* 벡터 게놈 RNA 서열에 상대적인 넘버링

** 소문자 표적 서열은 (HXB2)에 대한 것이고, 밑줄 친 표적 서열은 gag ORF(U1

376)에서 AA>CGCG 프레임시프트이다.

[0591]

[0592]

부착 세포 배양, 형질감염 및 렌티바이러스 벡터 생산

[0593]

HEK293T 세포를 37°C, 5% CO₂에서 완전 배지(10% 열-불활성화(FBS)(Gibco), 2 mM L-글루타민(Sigma) 및 1% 비필수 아미노산(NEAA)(Sigma)이 보충된 돌베코 변형된 이글 배지(DMEM)(Sigma))에서 유지하였다.

[0594]

부착 모드에서 HIV-1 벡터의 표준 규모 생산은 하기 조건에서 10 cm 접시에서 이루어졌다(모든 조건은 다른 형식으로 수행할 때 면적에 따라 조정됨): HEK293T 세포는 10 mL 완전 배지에서 ml 당 3.5×10^5 개 세포로 시딩하였고, 약 24시간 후 10 cm 플레이트당 하기 질량비의 플라스미드를 사용하여 세포를 형질감염시켰다: 4.5 µg 게놈, 1.4 µg Gag-Pol, 1.1 µg Rev, 0.7 µg VSV-G 및 0.01 내지 2 µg의 변형된 U1 snRNA 플라스미드.

[0595]

형질감염은 제조업체의 프로토콜(Life Technologies)에 따라 Opti-MEM에서 Lipofectamine 2000CD와 DNA를 혼합하여 매개하였다. 소듐 부티레이트(Sigma)를 ~18시간 이후에 10 mM 최종 농도까지 5-6시간 동안 첨가한 후, 10 ml의 새로운 무혈청 배지로 형질감염 배지를 대체하였다. 전형적으로, 벡터 상청액을 20-24 시간 후에 수확한 다음, 여과(0.22 µm)하고 -20/-80°C에서 동결하였다. 뉴클레아제 처리에 대한 양성 대조군으로서, 전형적으로 Benzonase®를 여과 전에 1시간 동안 5 U/mL로 수확물에 첨가하였다.

[0596]

현탁 세포 배양, 형질감염 및 렌티바이러스 벡터 생산

[0597]

HEK293T.1-65s 현탁 세포를 37°C, 5% CO₂에서, 진탕 인큐베이터(190 RPM으로 설정된 25 mm 궤도)에서 Freestyle + 0.1% CLC(Gibco)에서 성장시켰다. 현탁액을 사용한 모든 벡터 생산은 24-웰 플레이트(1 mL 부피, 진탕 플랫폼 상에서), 25 mL 진탕 플라스크 또는 생물반응기(≤5L)에서 수행하였다. HEK293Ts 세포를 무혈청 배지에서 ml 당 8×10^5 개 세포로 시딩하였고 벡터 생성 전체에 걸쳐 진탕하면서 37°C, 5% CO₂에서 인큐베이션하였다. 시딩 후 약 24 시간에 형질감염 시의 유효 최종 배양 부피당 하기 질량비의 플라스미드를 사용하여 세포를 형질감염시켰다: 0.95 µg/mL 게놈, 0.1 µg/mL Gag-Pol, 0.6 µg/mL Rev, 0.7 µg/mL VSV-G, 및 0.01 내지 0.2 µg/mL의 변형된 U1 snRNA 플라스미드.

[0598]

형질감염은 제조업체의 프로토콜(Life Technologies)에 따라 Opti-MEM에서 Lipofectamine 2000CD와 DNA를 혼합하여 매개하였다. 소듐 부티레이트(Sigma)를 ~18시간 이후에 10 mM 최종 농도까지 첨가하였다. 전형적으로, 벡터 상청액을 20-24 시간 후에 수확한 다음, 여과(0.22 µm)하고 -20/-80°C에서 동결하였다. 뉴클레아제 처리에 대한 양성 대조군으로서, 전형적으로 Benzonase®를 여과 전에 1시간 동안 5 U/mL로 수확물에 첨가하였다.

[0599]

렌티바이러스 벡터 적정(titration) 아세이

[0600]

GFP 마커-함유 카세트에 의한 렌티바이러스 벡터 적정을 위해, HEK293T 세포를 96-웰 플레이트에 1.2×10^4 세포/웰로 시딩하였다. GFP-암호화 바이러스 벡터를 사용하여 8 mg/ml 폴리브렌 및 $1 \times$ 페니실린 스트렙토마이신

을 포함하는 완전 배지에서 약 5-6시간 동안 세포를 형질도입한 후 새로운 배지를 추가하였다. 형질도입된 세포를 37℃, 5% CO₂에서 2일 동안 인큐베이션하였다. 그런 다음 Attune-NxT(Thermofisher)를 사용하여 유세포 분석(flow cytometry)을 위해 배양물을 준비하였다. GFP 발현 퍼센트를 측정하고 형질도입 당시 2×10^4 개 세포의 예측된 세포 수(전형적인 성장 속도 기준), 벡터 샘플의 희석 인자, 양성 GFP 모집단의 백분율 및 형질도입 시 종 부피를 사용하여 벡터 역가를 계산하였다.

[0601] 통합 분석에 의한 렌티바이러스 벡터 적정을 위해, 0.5 mL 부피의 순수(neat)에서 1:5로 희석된 벡터 상청액을 사용하여 8 µg/mL 폴리브렌의 존재 하에 12-웰 규모에서 1×10^5 HEK293T 세포를 형질도입하였다. 배양물을 10일 동안 계대한(2-3일마다 1:5 분할) 후 1×10^6 세포 펠릿으로부터 숙주 DNA를 추출하였다. HIV 패키징 신호(Ψ) 및 RRP1에 대한 FAM 프라이머/프로브 세트를 사용하여 이중 정량적 PCR을 수행하였으며, 하기 인자를 사용하여 벡터 역가(TU/mL)를 계산하였다: 형질도입 부피, 벡터 희석, 반응 당 검출된 RRP1-정규화된 HIV-1 Ψ 카피.

[0602] **LV-형질도입된 세포의 전사 리드-스루('리드-인') 분석**

[0603] 표시된 HIV 벡터를 24웰 플레이트 규모에서 256_U1 snRNA의 존재 또는 부재 하에 현탁액 모드 1.65s 세포의 일시적 형질감염에 의해 생산하였다. 2일 후에 상청액을 수확하여 HEK293T 세포에 대한 유세포 분석을 사용하여 GFP 발현에 의해 적정하였다. 그런 다음 GFP 역가를 사용하여 감염 다중도(multiplicity of infection, MOI) 1에서 4.5×10^4 HEK293T 또는 92BR 세포(당나귀 1차 섬유아세포)를 형질도입하였다. 형질도입된 세포를 10일에 걸쳐 3회 계대한 후 수확하여 2개의 분취량으로 분할하였다: 하나는 총 RNA 추출을 위해 처리하였고 다른 하나는 게놈 DNA 추출을 위해 처리하였다. 총 RNA의 200 ng(293T) 또는 80 ng(92BR)을 DNase I 처리하고 SSIV VILO RT 시스템(Life Technologies)을 이용하여 RT-PCR을 수행하였다. cDNA를 1 ng/ul로 희석하고, 5 ul를 HIV Psi 및 세포성 GAPDH에 대한 프라이머 세트를 사용하는 SYBR qPCR에 적용하였다. 샘플 간의 발현 점수를 생성하기 위해 Delta Ct 방법을 사용하여 HIV Psi 카피를 GAPDH 카피로 정규화하였다. 게놈 DNA 추출물은 Qiacube 추출 시스템(Qiagen)을 사용하여 준비하고 용출된 DNA 5 ul를 qPCR을 사용하여 HIV 벡터 통합 아세이를 수행하였다. 세포당 통합된 벡터 게놈의 평균을 계산하기 위해 HIV Psi 카피를 세포 표적 RPPH1에 대해 정규화하였다. 그런 다음 형질도입 효율을 설명하기 위해 샘플의 HIV RNA 발현 점수를 세포당 통합된 벡터 카피의 수로 정규화하였다. 이 최종 값이 상대적인 HIV Psi RNA 발현 점수이며, 테스트된 각 세포주에 대해 모든 값을 256_U1 snRNA가 없는 상태에서 만들어진 표준 벡터(온전한 MSD 및 crSD)로 정규화하였다.

[0604] **실시예 1 - MSD로부터의 난잡한 스플라이싱, MSD-2KO 렌티바이러스 벡터의 감소된 역가 및 재지정된 U1 snRNA에 의한 역가 회복/부스트.**

[0605] 렌티바이러스 벡터 게놈의 일반적인 구조는 모든 3개 세대의 벡터 시스템에서 일관되게 유지되었는데(Lentiviral vectors: basic to translational. Toshie SAKUMA, Michael A. BARRY and Yasuhiro IKEDA. Biochem. J. (2012) 443, 603-618), 패키징 서열을 함유하는 HIV-1 프로바이러스의 5' 영역과 전이유전자 카세트의 상류에 있는 RRE의 위치는 유지하였지만, 다른 측면에서는 달랐고, 후 세대는 tat-독립적이 되고, 3'LTR에서 자가-불활성화되며, cppt 및 wPRE의 사용을 통합하였다. 5' 패키징 서열의 엔지니어링에 대한 예시가 명백하게 부족한 것은 그의 복잡한 구조와 그 안에 암호화된 HIV-1 복제의 여러 측면, 즉, 전사, 스플라이싱의 균형, GagPol의 번역, 게놈 이량체화, 어셈블리, 역전사 및 통합에 필요한 압축된 정보 때문일 수 있다. 이 복잡한 영역 내에서, 주요 스플라이스 공여자(MSD)는 SL1(이량체화 루프)과 SL3(Gag에 결합) 사이의, 스텝-루프 2(SL2) 영역에 내장되어 있다. RRE 서열(및 외피 영역 내의 관련 스플라이스 수용체 7(sa7))을 렌티바이러스 벡터 게놈에서 패키징 영역의 바로 하류에 재배치하면 rev 부재 시 MSD에 스플라이스 수용체(sa7)를 '제공'하는 것으로 생각되었다. 렌티바이러스 벡터 생산 동안 rev의 공급은 RRE에 대한 rev 결합 및 sa7에 대한 MSD 스플라이싱의 억제를 초래하고, 결과적으로 스플라이싱되지 않은 전장 렌티바이러스 벡터 게놈 vRNA의 생산을 초래한다고 추정되었다(도 2A). 그러나, 본 발명자들(도 4Bii) 및 다른 사람들(예: Cui *et al.* (1999), *J. Virol.*, 73: 6171-6176)은 MSD에서부터 전이유전자(트랜스제닉) 서열 내의 스플라이스 수용체 부위까지의 비정상적인 스플라이싱이 상당할 수 있어서, 벡터 비리온으로 패키징할 수 있는 스플라이싱되지 않은 vRNA가 상대적으로 미미한(modest) 수준에 이르게 되고(총 대비 - 도 2B 참조), 일부 경우에는 5% 미만임을 보여주었다.

[0606] 많은 수의 벡터 게놈 플라스미드가 세포로 전달되기 때문에, 패키징을 위한 vRNA 생성의 비효율성이 일시적 형질감염 방법에서 생성된 렌티바이러스 벡터 역가에서 항상 직접적으로 관찰되는 것은 아니다; 이러한 유형의 비

정상적인 스플라이싱이 발생하더라도 1×10^7 TU/mL 이상의 표준 3세대 벡터의 역가는 통상적으로 가능하다. 그러나, 본 발명자들은 훨씬 더 적은 수의 통합된 벡터 게놈 카세트가 존재할 수 있는 안정한 생산자 세포주의 개발을 위해서는, 비정상적인 스플라이싱의 문제가 사실상 더 중요할 것으로 예상한다. 실제로, 본 발명자들은 일반적으로 게놈 구성요소가 안정한 생산자 클론에서 제한적이라는 것을 발견하고, MSD 활성이 이러한 제한에 실질적으로 기여할 수 있다고 가정한다.

[0607] 전이유전자 카세트 내로의 MSD로 인한 비정상적으로 스플라이싱된 mRNA를 생성하는 데에는 아마도 덜 분명한 결과가 더 있을 수 있다: 생산 동안 전이유전자 카세트의 (증가된) 발현. 이전에, TRiP 시스템은 렌티바이러스 벡터 생산 동안 전이유전자 발현을 억제하기 위해 개발되었으며(WO 2015/092440에 기술됨), 이는 벡터 생산에 대한 특정 전이유전자 단백질의 부정적인 영향에 비례하여 연결된 벡터 역가의 회복을 가능하게 한다. 본 발명자들은 효율적인 비정상적인 스플라이싱(예를 들어, EF1a 구동 카세트를 포함하는 표준 렌티바이러스 벡터에서 -도 4 및 12 참조)이 전형적으로 전이유전자를 암호화하는 mRNA를 생성한다는 것을 발견하였다. 이론에 얽매이지 않고도, EF1a 구동 카세트의 경우 MSD는 강한 EF1a 스플라이스 수용체에 스플라이스하지만 다른 프로모터-UTR 서열의 경우 MSD는, rev의 존재하에서도, 약한 잠복 스플라이스 수용체를 '선택'한다. MSD는 벡터 게놈에 더 중심적인 다른 부위, 즉 전이유전자 프로모터 영역을 선호하여 RRE-sa7 서열을 '지나간 것(look past)'으로 보인다. 이것은 HIV-1 5' 패키징 영역의 '잔여(residual)' 속성일 수 있는데 왜냐하면 야생형 HIV-1에서 MSD는 전형적으로 중앙에 및 게놈의 3' 말단에 배치된 스플라이스 수용체에 스플라이스하기 때문이다. MSD가 벡터 서열 하류 내의 많은 위치에서 비정상적으로 스플라이싱하지만 넨센스-매개 붕괴 규칙(nonsense-mediated decay rules)을 통과하는 mRNA(즉, 이들은 단백질[전이유전자 단백질]을 암호화하기 때문에 합법적인 mRNA로 보인다)만이 세포질로 운반되고(및/또는 세포질에서 안정하며) 그런 다음 번역된다는 것도 가능하다. 이것은 mRNA를 암호화하는 전이유전자의 억제를 유지하기 위해 TRiP 시스템에 추가 부담을 만들어, 더 큰 mRNA 풀에 대한 억제 제어가 덜 되게 한다(도 12 참조). 더욱이, 조직 특이적 프로모터의 사용(부분적으로 렌티바이러스 벡터 생산 동안 전이유전자 발현을 피하기 위한)은 이러한 비정상적인 스플라이싱 메커니즘에 의해 전이유전자를 암호화하는 번역 가능한 mRNA의 세포질 출현에 의해 '취소(undone)'될 가능성이 있다. 본질적으로, 전이유전자는 벡터 게놈 vRNA의 발현을 구동하는 (전형적으로 강력한) 구성적 프로모터에 의해 발현될 것이다.

[0608] 따라서, MSD-돌연변이된 렌티바이러스 벡터를 생성하는 데는 여러 가지 이유가 있으며, 실제로 다른 사람들은 tat-독립적인 렌티바이러스 벡터에서 그렇게 하려고 시도했지만 성공하지 못했다. 본 발명자들은 HIV-1 tat-독립적 3세대 벡터에서 MSD의 돌연변이가 SL2 내의 인접한 잠복 스플라이스 공여자 부위를 활성화시켜 상당한 수준의 스플라이싱을 초래한다는 것을 발견하였다. 이러한 이유로 본 발명자들은 MSD와 근처의 잠복 스플라이스 공여자(crSD) 모두에서 돌연변이를 사용했으며(도 10 참조), 이 변형을 'MSD-2KO' 또는 'MSD2KO' 또는 'MSD의 기능적 변형'으로 지칭한다. 이 이중 돌연변이가 렌티바이러스 벡터 생산 동안 SL2의 스플라이싱 영역(MSD 및 CrSD 모두 포함)에서부터 강력한 EF1a 스플라이스 수용체까지의 비정상적인 스플라이싱을 제거하는 데 매우 효과적이라는 것이 도 4에 나타나 있다. 또한 다른 사람들이 유사하게 보고한 바와 같이, 3개의 서로 다른 프로모터-GFP 전이유전자 카세트를 포함하는 MSD-2KO 렌티바이러스 벡터 게놈이 감소된 벡터 역가를 유도하는 것으로 나타났다(도 3). 도 4Bi에서, HIV-1 tat를 트랜스로 제공할 경우 관찰된 MSD-2KO 렌티바이러스 벡터 게놈의 역가 감소를 구제할 수 있음이 또한 입증되었지만, (SL4의 마이너한 잠복 스플라이스 공여자로부터의) '비정상적인' 스플라이스 생성물의 양이 증가한다(도 4Bi). 중요하게도, 벡터 패키징 신호의 다른 영역으로 재지정된 변형된 U1 snRNA는 마이너한 스플라이스 생성물의 존재를 증가시키지 않으면서 MSD-2KO 렌티바이러스 벡터 게놈 역가를 증가시키는 것으로 나타났다(도 4).

[0609] 실시예 2 - MSD-돌연변이된 렌티바이러스 벡터 역가의 향상은 벡터 게놈 카세트 내의 5'polyA 부위의 억제로 인한 것이 아니다.

[0610] 본 발명이 5'polyA 부위를 억제하도록 작용하는지 평가하기 위해, 5'polyA 부위에 대한 기능적 돌연변이를 EF1a 또는 CMV 구동 GFP 전이유전자 카세트를 포함하는 MSD-2KO 렌티바이러스 벡터 게놈 내에 도입하였다(도 6); 'pAm1' polyA 돌연변이에 대한 설명은 도 6A에 있으며, 이것은 완전한 폐지 폴리아데닐화 활성을 보여준다. 놀랍게도, 본 발명자들은 5'polyA 신호의 돌연변이가 MSD-2KO 렌티바이러스 벡터 게놈을 함유하는 EF1a-GFP에서 역가를 단지 부분적으로만 증가시키고, CMV-GFP MSD2KO 렌티바이러스 벡터 게놈에서는 실질적으로 효과가 없다는 것을 발견하였고, 이는 MSD-2KO 돌연변이의 감쇠 효과 정도에 상응하는 것으로 보인다(EF1a 함유 게놈의 경우, MSD2KO 돌연변이가 덜 두드러짐). 중요하게도, 이 실험에서 305U1 분자의 공급은 표준 렌티바이러스 벡터 게놈(여기서 내인성 U1 snRNA가 잔류하는 5'polyA 활성을 완전히 억제할 수 있는 것으로 추정됨)과 가능한 5'polyA 활성이 없는 MSD2KO/pAm1 렌티바이러스 벡터 게놈 양자 모두의 역가를 증가시켰다. 이것은 MSD-돌연변

이된 렌티바이러스 벡터의 역가를 회복하기 위해 공급되는 변형된 U1 snRNA가 이전에 설명되지 않은 전사 후 단계에서 작용한다는 강력한 증거를 제공한다.

[0611] 그 다음, 본 발명자들은 305U1 및 256U1의 70K 및 U1A 결합 루프를 돌연변이시키려고 하였으며, 이는 이것이 MSD2KO 렌티바이러스 벡터 게놈의 역가의 관찰된 증가에 영향을 미치는지 평가하기 위해서였다. 도 7은 변형된 U1 snRNA 내의 SL1 또는 SL2의 기능적 돌연변이가 생산 동안 공동-발현되는 경우 MSD-2KO 렌티바이러스 벡터 역가를 증가시키는 이들 분자의 능력에 아무런 영향을 미치지 않고; Sm 단백질 결합 돌연변이만이 이 활성을 차단하였다는 것을 보여준다. 이것은 polyA 활성을 억제하는 데 있어 재지정된 U1 snRNA의 이전에 필수적인 70K-결합 특성이 본 발명에서 중요하지 않다는 것을 보여주었고, MSD-돌연변이된 렌티바이러스 벡터 역가를 증가시키는 데 사용되는 변형된 U1 snRNA가 신규한 메커니즘에 의해 기능한다는 추가의 증거를 제공한다.

[0612] MSD-2KO 렌티바이러스 벡터에 적용될 때 변형된 U1 snRNA에 의해 매개되는 역가 증가가 표준 렌티바이러스 벡터와 비교하여 '선택되는' 표적 부위에 있어서 차이가 있는지를 평가하기 위해, MSD-2KO 렌티바이러스 벡터 게놈의 5' 영역을 따라 서로 다른 부위를 표적으로 하는 변형된 U1 snRNA의 패널(표 I 참조)을 스크리닝하였다(도 9). 이 스크리닝은 패키징 영역에 대한 표적화가 선호되고(SL1-3), 아마도 SL3 내에 '핫스팟(hotspot)'이 있을 수 있음을 나타낸다. 표적 부위에 대해 상보적인 15개의 뉴클레오타이드를 가지는(또는 9개의 뉴클레오타이드를 기 포함) 변형된 U1 snRNA로 스크리닝을 수행하였고, 본 발명자들이 이전에 표준 렌티바이러스 벡터에 대해 입증했듯이, 9개 이상의 뉴클레오타이드의 상보성 길이를 사용할 때 역가의 증가가 더 강력할 수 있다. 실제로, 본 발명자들은 MSD-2KO 렌티바이러스 벡터의 경우, 변형된 U1 snRNA('305' 서열을 표적화)로 관찰된 역가 부스트가 상보성의 단 7개의 뉴클레오타이드에서도 관찰될 수 있다는 것을 보여주지만, 선호되는 사용에 있어서는 역가 부스트의 증가로 인해 10개 내지 15개의 상보성의 뉴클레오타이드를 사용하는 것이 더 나을 것이며(도 8), 이렇게 하면 변형된 U1 snRNA에 의한 가능한 '표적-이탈(off-target)' 효과도 최소화할 수 있기 때문이다.

[0613] **실시예 3 - 변형된 U1 snRNA에 의한 MSD-돌연변이된 렌티바이러스 벡터 역가의 향상은 스플라이스 공여자 돌연변이의 유형에 의존하지 않는다**

[0614] 도 10A는 MSD와 하류에 위치한 잠복 스플라이스 공여자를 모두 돌연변이시키는 MSD-돌연변이된 렌티바이러스 벡터 게놈 패키징 영역의 'MSD-2KO' 변이체의 SL2 루프에 대한 유전자 변형을 도시한다(MSD-2KO 변이체는 본 명세서의 많은 비제한적인 예에서 사용되었다). 변형된 U1 snRNA를 사용하여 역가를 부스팅하는 효과가 어쨌든 MSD2-KO 변이체에 대해 수행된 특정 변경에 의존한 것인지를 평가하기 위하여, 본 발명자들은 3개의 다른 스플라이스 공여자 영역 돌연변이를 만들었다: [1] MSD 및 잠복 공여자 서열 내에 두 가지 특정 변화를 또한 도입한 'MSD2KOv2'; [2] 전체 SL2 루프를 인공 스템 루프로 대체한 'MSD-2KOm5'; 및 [3] 완전한 SL2 결실, 따라서 전체 스플라이스 공여자 영역(스플라이싱 영역이라고도 함) 제거. 그런 다음 본 발명자들은 HEK293T 세포 +/- 변형된 U1 snRNA(256U1)에서 표준 또는 MSD-돌연변이된 렌티바이러스 벡터 변이체(EFS-GFP 발현 카세트 포함)를 생산하였고, 벡터 상청액을 적정하였다(도 10B). 결과는 4개의 MSD-돌연변이된 렌티바이러스 벡터 변이체 모두가 표준 벡터와 비교하여 감소되었지만, 4개 모두가 렌티바이러스 벡터 생산 동안 제공된 변형된 U1 snRNA의 사용에 의해 향상될 수 있음을 보여주며, 이는 변형된 U1 snRNA가 스플라이스 공여자 영역 돌연변이의 특정 서열 의존성이 없음을 나타낸다. 흥미롭게도, MSD-2KOm5 변이체가 가장 적게 감소되었으며, 256U1 분자의 존재 하에 생성되었을 때 사용된 내부 프로모터의 정체에 관계없이 출력 역가가 가장 크게 증가하였다(EFS, EF1a, CMV 및 인간 PGK 프로모터 비교).

[0615] **실시예 4 - 시스에서(in cis) 렌티바이러스 벡터 게놈 플라스미드 DNA 백본 내에 암호화된 변형된 U1 snRNA 카세트의 사용.**

[0616] 본원의 이전 실시예는 HEK293T 세포를 렌티바이러스 벡터 구성요소 플라스미드 및 변형된 U1 snRNA를 암호화하는 플라스미드로 일시적 공동-형질감염시킴으로써 *트랜스에서* 렌티바이러스 벡터 생산 동안 변형된 U1 snRNA 분자의 사용을 개시하였다. MSD-돌연변이된 렌티바이러스 벡터 게놈 카세트 및 변형된 U1 snRNA 카세트가 동일한 플라스미드 DNA 분자 내에 적절하게 암호화될 수 있는지 평가하기 위해, 3개의 변이체 작제물을 클로닝하였다(도 11A). 256U1 발현 카세트가 렌티바이러스 벡터 게놈 카세트 및/또는 기능적 플라스미드 백본 서열과 관련하여 3가지 다른 배열로 삽입되도록 MSD-돌연변이된 렌티바이러스 벡터 게놈 카세트(MSD-2KO 변이체)를 변형시켰다. 이러한 '시스(cis)' 버전 플라스미드를 HEK293T 세포에서 MSD-돌연변이된 렌티바이러스 벡터를 만드는 데 사용하였으며, 변형된 U1 snRNA 플라스미드를 비변형 MSD-돌연변이된 렌티바이러스 벡터 게놈과 공동-형질감염시킨 '트랜스(trans)' 모드와 비교하였다(도 11B). 결과는 이러한 '시스(cis)' 버전 플라스미드의 역가가 공동-형질감염으로 제공된 비변형 MSD-2KO 렌티바이러스 벡터 게놈 +256U1과 유사하였음을 보여준다.

- [0617] 실시예 5 - 표준 또는 MSD-돌연변이된 렌티바이러스 벡터 양자 모두의 생산을 향상시키기 위한 변형된 U1 snRNA를 안정적으로 발현하는 세포주의 사용
- [0618] 305U1 발현 카세트를 HEK293T 세포에 안정적으로 통합시켰고, 일시적 형질감염 +/- 추가의 305U1 플라스미드 DNA에 의해 표준 또는 MSD-2KO 렌티바이러스 벡터를 생산하였다. 안정적인 세포의 성공적인 생성은 변형된 U1 snRNA가 세포독성 효과 없이 세포 내에서 내인성으로 발현될 수 있음을 최초로 보여주며, 변형된 U1 snRNA가 U1 snRNA 합성 또는 스플라이세오솜에 관여하는 세포 인자를 적정-아웃(titrate-out)시키지 않으며, 표적을 벗어나는(off-targeting) 현상이 발생하지 않거나 표적을 벗어나는 효과가 정상적인 세포 생존에 영향을 미치지 않음을 나타낸다. 렌티바이러스 벡터의 출력 역가는 표준 및 MSD-2KO 렌티바이러스 벡터 양자 모두에 대해 변형된 U1 snRNA에 의해 매개되는 역가 증가가 변형된 U1 snRNA의 안정적인 제공에서 가능함을 입증한다(도 13). 이는 변형된 U1 snRNA를 렌티바이러스 벡터 패키징 및 생산자 세포주에 쉽게 통합할 수 있게 할 것이다.
- [0619] 실시예 6 - MSD-돌연변이된 렌티바이러스 벡터는 생산 동안 더 적은 전이유전자 단백질을 생산한다
- [0620] 렌티바이러스 벡터 생산 동안 비정상적인 스플라이싱을 제거하는 추가 이점은 전이유전자 단백질 생산에 이르게 하는 전이유전자를 암호화하는 mRNA의 양을 줄인다는 것이다. 전이유전자 발현은 렌티바이러스 벡터 생산에 상당한 영향을 미칠 수 있으며, 이로 인해 바이러스 벡터 생산 동안 전이유전자 번역을 억제하기 위해 TRiP 시스템을 이전에 개발하게 되었다(WO 2015/092440에 기술됨). 간략하게, 박테리아 단백질 'TRAP'이 벡터 생산 동안 공동-발현되어 5'UTR 내에서 전이유전자 ORF의 상류에 삽입된 'TRAP 결합 서열'(tbs)에 결합함으로써, 스캐닝 리보솜을 차단한다.
- [0621] 이 작업 과정에서, 본 발명자들은 SL2의 주요 공여자 스플라이스 영역에서부터 내부 스플라이스 수용체 부위까지의 스플라이싱-아웃으로 인해 벡터 게놈 카세트를 구동하는 '외부' (CMV) 프로모터로부터 전이유전자를 암호화하는 mRNA가 효과적으로 생산된다는 것을 예기치 않게 발견하였다. 이것이 발생하는 정도는 cppt와 전이유전자 ORF 사이의 내부 서열(즉, 프로모터-5'UTR 서열)에 따라 다르다. 전이유전자 카세트에서 EF1a 프로모터(매우 강력한 스플라이스 수용체를 포함)를 사용하면 외부 프로모터로부터 유래하는 전체 전사체의 95% 이상에서 MSD로부터의 비정상적인 스플라이싱을 초래한다(도 2 참조). 표준 또는 MSD-2KO 렌티바이러스 벡터 생산 배양물에서 총 GFP 발현을 비교함으로써(도 12B), 본 발명자들은 생산 중에 발현된 전이유전자 단백질의 최대 80%가 비정상적인 스플라이스 생성물에서 유래한다는 것을 보여준다. 본 발명자들은 MSD-2KO 유전자형을 TRiP 시스템과 결합하면 생산되는 전이유전자 단백질의 감소를 증대시킨다는 것을 발견하였다.
- [0622] 실시예 7: tbs-Kozak 접합의 최적화.
- [0623] 4개의 tbs-Kozak 변이체를 설계하였고(표 II-패널 I 참조) 도 14A에 표시된 대로 GFP 리포터 작제물에 클로닝하였다. 변이체 0, 1 및 2는 확장된 Kozak 서열(컨센서스 GNNRVVATG 준수)이 마지막 2개의 tbs 반복부 서열과 중첩되도록 설계한 반면, 변이체 3은 단지 마지막 tbs 반복부 서열과 중첩되도록 설계하였다. Kozak 서열이 마지막 tbs 반복부 서열의 3개 뉴클레오티드 하류인 원래의 리포터(즉, 중첩 없음)와 함께 이들 변이체를 HEK293T 세포 +/- TRAP-발현 플라스미드에 개별적으로 공동-형질감염시키고, GFP 발현을 형질감염 2일 후에 유세포 분석에 의해 측정하였다. GFP 발현 점수(%GFP 양성 세포 × MFI)를 유세포 분석 데이터로부터 생성하고 플롯팅하였다(도 14B).
- [0624] 이러한 tbs-Kozak 변이체 중 2개를 scAAV2 벡터 게놈 발현 카세트 내에서 2개의 상이한 프로모터(EFS 및 huPGK)와 함께 테스트하고, 중첩되는 tbs/Kozak 서열을 가지지 않는 tbs 함유 카세트와 비교하였다. 이러한 GFP 리포터 벡터 게놈 플라스미드를 HEK293T 세포 +/- TRAP-발현 플라스미드에 개별적으로 공동-형질감염시키고, GFP 발현을 형질감염 2일 후에 유세포 분석에 의해 측정하였다. GFP 발현 점수(%GFP 양성 세포 × MFI)를 유세포 분석 데이터로부터 생성하고 플롯팅하였다(도 15). 중첩되지 않는 tbs/Kozak 변이체(원래 변이체 및 tbs와 Kozak 사이에 HpaI 부위를 포함하는 두 번째 변이체)는 50배 내지 100배 억제 가능한 반면, 새로운 tbs-Kozak 변이체(tbs_V0 및 tbs_V3)는 이의 적어도 10배(500배 내지 3500배)만큼 억제되었다. 이들 tbs-Kozak 변이체는 또한 L33 또는 L12 개선된 리더가 EFS 또는 huPGK 프로모터 카세트에 사용되었을 때와 유사하게 수행하였다. 중요하게도, 새로운 변이체에 대한 'ON' 수준(TRAP이 없는 경우)은 중첩되지 않는 tbs/Kozak 변이체와 유사했으며, 이는 새로운 변이체 내의 Kozak 서열이 효율적인 번역을 지시하는 데 효과적임을 나타낸다.
- [0625] 데이터는 모든 tbs-Kozak 변이체가 유사한 수준의 'ON' 전이유전자 발현이 가능함을 나타내며, 이는 Kozak 서열이 번역 개시의 효율적인 수준을 지시한다는 것을 보여준다. 변이체 1은 원래 리포터 구성과 비교할 때 TRAP에 의해 제대로 억제되지 않았다(TRAP의 존재 하에 GFP 발현 수준이 10배 더 높음). 그러나, 놀랍게도 변이체 0, 2

및 3은 원래 구성에 비해 더 낮은 수준으로 억제되었다.

패널 I	최적의 3'tbs-Kozak 접합
Original	[kagnn]acaGCCACCATG
Variant '0'	[kaGCC][GAGAT]G
Variant '1'	[kagGC][GAGCA]TG
Variant '2'	[kagnG][GAGCC]ATG
Variant '3'	[kagnn][GAGAC]CATG

표 II: 최적의 3'tbs-Kozak 접합 서열

[패널 I] 상류 tbs 서열의 3' 말단과 중첩되도록 설계된 변이체 Kozak 서열. 확장된 Kozak 서열은 주요 전이유전자 ATG 개시 코돈이 tbs의 3' 말단 KAGNN 반복부의 9개 뉴클레오타이드 하류에 배치되도록, 즉 중첩이 없도록 배치하였다. 주요 전이유전자 ATG 개시 코돈을 상류 tbs에 더 가깝게 위치시키기 위해, 3' 말단 tbs 반복부의 컨센서스 KAGNN 반복부가 유지되는 동시에 효율적인 확장된 Kozak 컨센서스 서열(본원에서 GNNRVVATG로 정의됨)이 유지되도록 4개의 변이체를 설계하였다. KAGNN 반복부 서열은 대괄호 안에 있고, Kozak 서열은 굵은 대문자로 표시하였다.

실시예 8: 개선된 중첩되는 tbs-Kozak 변이체를 전장의 인트론 함유 EF1a 프로모터에 통합.

이전 실시예는 중첩되지 않는 tbs/Kozak 변이체와 비교하여 중첩되는 tbs-Kozak 변이체가 억제를 개선했음을 나타내었으며, 이는 EFS 프로모터(내포된 인트론을 제거함으로써 절단된 EF1a 프로모터)의 맥락에서 입증되었다. tbs-Kozak 변이체가 전장 EF1a 프로모터(즉, 인트론의 스플라이싱-아웃 후) 내에서 유사하게 '수행하였는'지 평가하기 위해, 3개의 tbs-Kozak 변이체(0, 2 및 3)를 EF1a 프로모터를 포함하는 GFP 리포터 카세트에 클로닝하였다(도 16A). 스플라이싱 후, 5'UTR은 엑손1(즉, L33 개선된 리더), 엑손2의 첫 번째 뉴클레오타이드를 포함하는 짧은 12 nt 서열, 그 다음 tbs-Kozak 변이체 서열(서열번호 106)을 포함한다. 이들 GFP 리포터 플라스미드를 tbs를 함유하지 않는 리포터와 함께 현탁액, 무혈청 HEK293T 세포 +/- TRAP-발현 플라스미드에 개별적으로 공동-형질감염시켰고, GFP 발현을 형질감염 2일 후에 유세포 분석에 의해 측정하였다. GFP 발현 점수(%GFP 양성 세포 × MFI)를 유세포 분석 데이터로부터 생성하고 플롯팅하였다(도 16B). 또한, 이러한 tbs-함유 GFP 발현 카세트를 HIV-1 기반 렌티바이러스 벡터 계통 플라스미드(여기서 MSD 및 crSD는 불활성화됨, 아래 참조)에 클로닝하고, 현탁액, 무혈청 HEK293T 세포에서 유사한 실험을 수행하여, GFP 발현 점수를 측정하였다(도 16C). 데이터는 실제로 중첩되는 tbs-Kozak 변이체가 사용된 중첩되지 않는 tbs/Kozak 변이체와 비교하여 개선된 전이유전자 억제를 나타낼 수 있었음을 보여준다.

실시예 9: tbs의 3' 말단 KAGNN 반복에 의한 코어 Kozak 서열의 점점 더 많은 폐색은 TRAP에 의한 점진적으로 더 큰 전이유전자 억제를 초래한다.

실시예 7에서, 제한된 수의 중첩되는 tbs-Kozak 변이체를 생성하고 인트론을 함유하지 않은 프로모터 EFS 및 huPGK의 맥락에서 테스트하였다. 그런 다음 이러한 변이체를 실시예 5에서 인트론을 함유하는 전체 EF1a 프로모터에서 테스트했는데, 이는 이 억제하기 어려운 프로모터가 중첩되는 tbs-Kozak 변이체를 사용하여 억제될 수 있음을 보여준다.

tbs의 3' 말단 KAGNN 반복부 내에 코어 Kozak 서열을 '숨김'으로써 개선된 TRAP 억제의 원칙을 추가로 예시하기 위해, 새로운 변이체 패널을 설계하였다(표 IV 참조). 이들은 하기 세 개의 '중첩 그룹'으로 암호화되는 모든 가능한 변이체를 기초로 하였다; KAGatg(KAGNN 컨센서스가 코어 Kozak의 ATG와 가능한 한 많이 중첩되는 경우, 즉 코어 Kozak의 ATG의 처음 두 뉴클레오타이드가 겹침), KAGNatg(KAGNN 컨센서스가 코어 Kozak의 ATG의 첫 번째 뉴클레오타이드와 겹치는 경우) 및 KAGNNatg(KAGNN 컨센서스가 코어 Kozak의 ATG와 겹치지 않는 경우). (실시예 3 및 5의 초기 변이체 tbs-kzkV0, tbs-kzkV1 및 tbs-kzkV2는 이러한 정의된 그룹에 속한다). 'GT' 디뉴클레오타이드를 생성한 이 그룹 내의 모든 변이체는 생성/평가되지 않았는데 그 이유는 이들이 잠복 스플라이스 공여자 부위를 생성할 수 있는 바람직하지 않은 가능성 때문이다.

표 III: 추가의 예시를 위해 생성된 중첩되는 tbs-Kozak 변이체

KAGatg 또는 KAGNatg 또는 KAGNNatg의 컨센서스와 관련된 '중첩 그룹'을 나타내는 가능한 모든 변이체를 생성하였으며, 단, 원치 않는 (잠복) 스플라이스 공여자 부위를 생성할 수 있는 'GT' 디뉴클레오타이드에 이르게 하는 것은 제외하였다. 처음 10개 KAGNN 반복부는 명확성을 위해 여기에서 컨센서스로 표시하였지만 주로 서열 번호 8의 처음 48개 뉴클레오타이드였다. 3' 말단 tbs KAGNN은 각 변이체에 암호화된 대로 표시하였다(이탤릭체 및 대

팔호); 코어 Kozak 컨센서스는 굵게 표시하였으며 더 광범위하고 확장된 Kozak 컨센서스의 일부로 제시된 모든 뉴클레오타이드는 밑줄로 표시하였다.

중첩	그룹	tbs-Kozak 변이체 ID	중첩되는 tbs 3'KAGNN-Kozak 변이체
N/A	중첩되지 않는 tbs-Kozak		[KAGNN]x10-[KAGNN]ACAGCCACCATG
KAGatg	tbskzkV0.G		[KAGNN]x10-[GAGAT]G
	tbskzkV0.T		[KAGNN]x10-[TAGAT]G
KAGNatg	tbskzkV1.0		[KAGNN]x10-[GAGCA]TG
	tbskzkV1.1		[KAGNN]x10-[GAGAA]TG
	tbskzkV1.2		[KAGNN]x10-[GAGGA]TG
	tbskzkV1.3		[KAGNN]x10-[TAGAA]TG
	tbskzkV1.4		[KAGNN]x10-[TAGCA]TG
	tbskzkV1.5		[KAGNN]x10-[TAGGA]TG
KAGNNatg	tbskzkV2.0		[KAGNN]x10-[GA GCC]ATG
	tbskzkV2.1		[KAGNN]x10-[GA GAA]ATG
	tbskzkV2.2		[KAGNN]x10-[GA GAG]ATG
	tbskzkV2.3		[KAGNN]x10-[GA GAC]ATG
	tbskzkV2.4		[KAGNN]x10-[GA GAT]ATG
	tbskzkV2.5		[KAGNN]x10-[GA GCA]ATG
	tbskzkV2.6		[KAGNN]x10-[GA GCG]ATG
	tbskzkV2.7		[KAGNN]x10-[GA GCT]ATG
	tbskzkV2.8		[KAGNN]x10-[GA GGA]ATG
	tbskzkV2.9		[KAGNN]x10-[GA GGG]ATG
	tbskzkV2.10		[KAGNN]x10-[GA GGC]ATG
	tbskzkV2.11		[KAGNN]x10-[TA GAA]ATG
	tbskzkV2.12		[KAGNN]x10-[TA GAG]ATG
	tbskzkV2.13		[KAGNN]x10-[TA GAC]ATG
	tbskzkV2.14		[KAGNN]x10-[TA GAT]ATG
	tbskzkV2.15		[KAGNN]x10-[TA GCA]ATG
	tbskzkV2.16		[KAGNN]x10-[TA GCG]ATG
	tbskzkV2.17		[KAGNN]x10-[TA GCC]ATG
	tbskzkV2.18		[KAGNN]x10-[TA GCT]ATG
	tbskzkV2.19		[KAGNN]x10-[TA GGA]ATG
	tbskzkV2.20		[KAGNN]x10-[TA GGG]ATG
	tbskzkV2.21		[KAGNN]x10-[TA GGC]ATG

[0636]

[0637]

일반적으로, 대부분의 변이체는 RVVATG의 선호되는 코어 Kozak 컨센서스를 따랐으며, 동시에 표시된 KAGNN tbs 컨센서스를 포함하는 것으로 제한하였다. 이들 변이체를 pEF1a-GFP 리포터 플라스미드에 클로닝하였고, 따라서 5'UTR은 (인트론의 스플라이싱 후) L33 리더(엑손 1)와 엑손 2의 짧은 12nt 서열을 포함하였으며, 이는 중첩되는 tbs-Kozak 변이체가 사용되지 않는 한 TRAP에 의해 덜 억제되는 것으로 실시예 5에서 이전에 밝혀졌다. 현탁액(무혈청) HEK293T 세포를 전형적으로 렌티바이러스 벡터(LV) 형질감염/생산을 반영하는 조건(예: 나트륨 부티레이트 유도 포함) 하에서 TRAP-발현 플라스미드가 있거나 없이 이러한 변이체로 개별적으로 형질감염시켰고, 전형적인 LV 수확 시간(형질감염 후 2일)에 유세포 분석을 수행하였다. +/- TRAP 조건에 대해 전역 GFP 발현 점수를 생성한 다음(%GFP × MFI; ArbU) 배수 억제 값을 생성하고 도 17A에 플롯팅하였다. 결과는 코어 Kozak 컨센서스 서열이 3' 말단 KAGNN tbs 반복부와 더 많이 중첩될수록, TRAP 억제 수준이 더 우수함을 입증한다. 각 중첩 그룹의 배수 차이(fold difference)에 대한 통계 분석(T-테스트)은 더 많은 코어 Kozak이 3' 말단 KAGNN tbs 반복부와 중첩됨에 따라 억제의 유의한 증가를 보여주며, 여기서 최고의 억제 점수는 개시 코돈의 3분의 2가 KAGNN 반복부의 일부를 형성하는 KAGatg 그룹에서 나온다. 모든 중첩 변이체는 중첩되지 않는 tbs 변이체와 비교하여 TRAP에 의해 통계적으로 더 큰 전이유전자 억제를 생성하였다.

[0638]

데이터를 도 17B에 추가로 계층화하였으며, 여기서 억제되지 않은 'ON' 전이유전자 수준을 최고에서 최저로 표시하였다. KAGatg 중첩 그룹의 두 변이체는 강조표시하였는데 TRAP 억제 측면에서 이 두 가지 가장 성능이 우수한 변이체 중에서 GAGatg 변이체(tbskzkV0.G)가 가장 높은 'ON' 수준을 제공했음을 보여주며, 이는 아마도 그것이 RVVATG의 코어 Kozak 컨센서스를 따르는 반면 TAGatg(tbskzkV0.T)는 그렇지 않기 때문일 것이다.

[0639]

이 모든 데이터는 중첩되는 tbs-Kozak 변이체가 중첩되지 않는 tbs 변이에 비해 TRAP 억제를 중재하는 데 더 효과적이며, 바람직하게는 벡터-형질도입된 표적 세포에서 우수한 'ON' 전이유전자 발현을 보장하기 위해 tbs-Kozak 중첩이 RVVATG의 코어 Kozak 컨센서스를 따른다는 일반 원칙을 뒷받침한다. 따라서, 이들 신규한 tbs-Kozak 변이체의 사용은 바이러스 벡터 생산 동안 보다 효율적인 전이유전자 억제를 가능하게 할 것이며, 상기

전이유전자 단백질 활성이 바이러스 벡터 역가 및/또는 활성에 유해한 경우 잠재적으로 바이러스 벡터 역가의 증가를 초래할 것이다.

[0640] 실시예 10: 인트론을 포함하는 공통 프로모터의 TRAP 매개 억제를 개선하기 위한 최적의 중첩되는 tbs-Kozak 변이체의 추가적 사용.

[0641] 실시예 8에서, 중첩되는 tbs-Kozak 변이체의 사용은 인트론을 함유하는 전장 EF1a 프로모터를 사용할 때 TRAP 매개 억제를 개선하는 것으로 나타났다. 유전자 치료를 위한 바이러스 벡터 계놈에 널리 사용되는 유사한 프로모터 - 예를 들어 CAG 프로모터 - 의 경우 내장된 엑손/인트론 서열의 존재는 TRAP 매개 억제의 정도가 '천연' 엑손 서열 내의 서열에 의해 영향을 받을 수 있음을 의미한다. TRAP 매개 억제를 개선하는 관점에서, 엑손 서열을 변경하는 것이 분명하거나 실현 가능하지 않을 수 있으며, 특히 이것이 스플라이싱 향상(예: 스플라이스 공여자 부위에 가까운 스플라이스 인핸서 요소)과 관련된 경우 그러하다. 널리 사용되는 CAG 프로모터는 CMV 인핸서 요소, 코어 닭 β -액틴 유전자 프로모터-엑손1-인트론 서열, 및 토끼 β -글로빈 유전자로부터의 스플라이스 수용체-엑손 서열을 포함한다. 본 발명의 다른 곳에서 놀랍게도 EF1a 프로모터로부터의 엑손 1(L33)이 모든 유형의 tbs 변이체의 상류에서 사용되어 TRAP 매개 억제를 개선할 수 있다는 것이 발견되었으며, 이는 안정적인 TRAP-tbs 복합체 형성을 지원하기 위한 우수한 서열 컨텍스트를 제공하는 것으로 추정되며, 효율적인 번역 억제를 가능하게 한다. 중첩되는 tbs-Kozak 변이체는 EF1a(인트론 함유) 및 인트론이 없는 여러 다른 프로모터 모두에서 TRAP 매개 억제를 돕는 것으로 나타났다.

[0642] 이 실시예에서(도 18A 참조), [a] CAG 프로모터의 '천연' 5'UTR 영역 내에 tbskzkV0.G 변이체(다른 실시예에서 변이체 '0'으로도 지칭됨)를 배치하고(서열번호 117), [b] 전체 '천연' 인트론 함유 5'UTR 영역을 tbskzkV0.G 변이체를 포함하는 EF1a 5'UTR-인트론 영역으로 교체함으로써(서열번호 118), 두 구성 모두를 CAG 프로모터로부터 TRAP 매개 억제를 개선하는 데 적용하였다. 상응하는 스플라이싱된 서열을 각각 서열번호 119 및 서열번호 120으로 나타내었다. 또한, 바이러스 벡터 계놈에서 전형적으로 인트론 없이 사용되는 프로모터 - 이 경우 CMV - 에 이중 인트론을 포함하는 인공 5'UTR이 추가될 수 있다는 것을 예시로 보여주었으며, 이의 발현은 TRAP에 의해 효율적으로 억제되는 것으로 이전에 나타났다. 구체적으로, tbskzkV0.G 변이체를 함유하는 EF1a 5'UTR-인트론 영역(서열번호 118)이 있는 5'UTR을 CMV 프로모터 컨텍스트에서 사용하였다.

[0643] 이러한 리포터 작제물(GFP 암호화)을 바이러스 벡터 생산 시나리오를 모델링하는 현탁액(무혈청) HEK293T 세포에서 'ON' 발현 수준과 TRAP 매개 억제 모두에 대해 평가하였다. 세포를 GFP 리포터 플라스미드 +/- pTRAP로 형질감염시키고, 형질감염 후 나트륨 부티레이트로 배양물을 유도하고(전형적인 바이러스 벡터 생산에 따라) 및 형질감염 후 ~2일(즉, 전형적인 바이러스 벡터 수확 지점)에서 GFP 발현에 대해 세포를 분석하였다. GFP 발현 점수(% GFP 양성 $\times$ MFI; ArbU)를 생성하고 플롯팅하였으며(도 18B), TRAP-억제 점수를 표시하였다. 이들 데이터는 전형적인 바이러스 벡터 생산 세포에서 TRAP 매개 발현이 tbskzkV0.G 변이체를 사용하는 CAG 프로모터로부터 3배에서 30배까지 개선될 수 있음을 나타낸다. 더욱이, 데이터는 서로 다른 프로모터로부터의 '천연' 5'UTR 영역 서열이 tbsKzkV0.G 변이체를 함유하는 인트론 함유 EF1a 5'UTR로 대체될 수 있고, 실질적으로 개선된 TRAP 매개 억제(30-40배 내지 >100배) 및 TRAP의 부재 시 높은 유전자 발현의 유지(즉, 바이러스 벡터-형질도입된 표적 세포에서의 발현을 모델링) 양자 모두에 이르게 함을 보여준다. 따라서, 새로운 EF1a-5'UTR-인트론-tbskzkV0.G 서열은 표적 세포에서 인트론에 의해 부여되는 알려진 이점(즉, 유전자 발현 증가)을 이중 프로모터에 제공하는 데 유용할 수 있으며, 동시에 바이러스 벡터 생산 동안 전이유전자 단백질의 효율적인 억제를 가능하게 하며, 상기 전이유전자 단백질 활성이 바이러스 벡터 역가에 유해한 경우 잠재적으로 바이러스 벡터 역가의 증가를 초래할 수 있다. 이것은 EF1a-5'UTR-인트론 서열을 다른 중첩되는 tbs-Kozak 서열과 함께 사용하는 경우에도 적용된다.

[0644] 실시예 11 - 주요 스플라이스 공여자 단독의 돌연변이 또는 인접한 잠복 스플라이스 공여자 부위의 추가적 돌연변이와의 조합이 비정상적인 스플라이스, 벡터 역가 및 변형된 U1 snRNA에 대한 반응에 미치는 영향의 평가.

[0645] 주요 스플라이스 공여자 부위 돌연변이 단독 또는 잠복 스플라이스 공여자(crSD, 이제 'crSD1'이라고도 함) 돌연변이(MSD-2KO에 존재함)와의 조합의 영향을 평가하기 위해, 'MSD-1KO'라는 또 다른 변이체를 클로닝하였다(도 20A). 이 변이체는 MSD 부위의 GT>CA 돌연변이만을 가지고 있었다. MSD-1KO 변이체 계놈을 표준 LV 계놈 및 MSD-2KO 계놈(모두 EF1a-GFP 전이유전자 카세트 포함)과 함께 사용하여 vRNA의 패키징 영역을 표적으로 하는 변형된 U1 snRNA(256U1)가 있거나 없이 현탁액(무혈청) HEK293T 세포의 일시적 형질감염에 의해 LV-GFP 조 수확 물질을 생성하였다. LV-GFP 역가를 도 20D에 표시하였으며, MSD 부위의 돌연변이만으로도 LV 역가를 감소시키기 에 충분하고, 이러한 역가가 MSD-2KO 벡터에 대해 관찰된 것과 동일한 수준으로 회복될 수 있음을 보여준다.

SL2 루프 내 MSD 영역으로부터의 비정상적인 스플라이싱의 생산 후 세포 분석을 스플라이싱되지 않은 생성물 또는 비정상적으로 스플라이싱된 생성물 모두의 검출을 허용하는 프라이머를 사용하여 RT-PCR에 의해 추출된 polyA-선택된 mRNA에 대해 수행하였다(도 20C; 프라이머의 위치는 도 23 참조). 데이터는 MSD 단독의 돌연변이는 바로 하류의 잠복 스플라이스 공여자(crSD[1])를 활성화한다는 것을 보여준다. 분석은 또한 SL2의 MSD 영역으로부터의 비정상적인 스플라이싱이 'MSD-2KO' 변이체에 존재하는 MSD와 인접한 잠복 스플라이스 공여자(crSD[1]) 둘 다의 돌연변이에 의해 완전히 제거되지만, 이것이 패키징 순서에서 더 하류인 SL4 루프 내에 존재하는 또 다른 마이너한 스플라이스 공여자 부위(crSD2)를 활성화시키는 것을 밝혀내었다(이는 이전 실시예에서도 분명하였다 - 도 4Bii 참조).

[0646] **실시예 12 - SL2의 주요 스플라이스 공여자 및 잠복 스플라이스 공여자 부위의 돌연변이와 조합된 패키징 서열의 SL4 루프의 잠복 스플라이스 공여자 돌연변이의 영향 평가.**

[0647] 실시예 11에서 MSD 및 crSD1 둘 다의 돌연변이가 패키징 신호의 SL2 영역으로부터의 비정상적인 스플라이싱을 완전히 제거했지만 이것이 SL4 루프에서 추가적인 잠복 스플라이스 공여자를 활성화시키는 것으로 나타났다(도 20A 참조; 'crSD2'). SL4의 crSD2 부위를 돌연변이시켜 활성을 제거할 수 있는지 평가하기 위해 - 그리고 패키징 서열에 대한 추가의 변형이 효율적인 패키징에 필요한 RNA의 접힘(folding)을 손상시킬 수 있다는 점을 염두에 두고 - MSD-3KO_1 및 MSD-3KO_2 변이체를 설계하였다. 이들은 SL2의 MSD-2KO 돌연변이와 crSD2 부위의 'GT' 디뉴클레오티드의 단일 뉴클레오티드 돌연변이를 포함하였다(도 20A 참조). MSD-3KO_1은 GT 디뉴클레오티드의 G를 'C'로 돌연변이시킨 반면, MSD-3KO_2 돌연변이는 GT 디뉴클레오티드의 T>C 돌연변이였다. T>C 돌연변이는 SL4 루프, 및 또한 더 넓은 패키징 서열의 3차 모델에서 더 나은 염기쌍을 예측한다는 점에 유의하라(Keane et al. Science. 2015 May 22; 348(6237): 917-921)). 이러한 변형을 함유하는 LV-EF1a-GFP 벡터를 실시예 11에 기재된 바와 같이 생산하였고, 생산 후 세포로부터의 벡터 RNA의 분석(도 21A) 및 생성된 벡터 수확물의 적정(도 21B)을 포함하였다. MSD-3KO_1 변이체의 비정상적인 스플라이싱은 절제된 것으로 보이지 않았고 RT-PCR 생성물의 엑손-엑손 접합부의 시퀀스에서 이 생성물이 10개 뉴클레오티드 하류에서 활성화된 또 다른 잠복 스플라이스 공여자 부위(crSD3)에서 유래된 것으로 밝혀졌다(도 21A의 라인 7 및 8 참조; crSD3의 서열에 대해서는 도 20A 참조). MSD-4KO_1 및 MSD-4KO_2의 작제 및 테스트는 crSD3로부터의 비정상적인 스플라이싱이 crSD2의 G>C 돌연변이에 의해 실제로 활성화된 것임을 입증하였는데, 왜냐하면 이것이 crSD3의 돌연변이로 폐지되었기 때문이다(도 21A; 라인 11 내지 14). 그러나, 놀랍게도 MSD-3KO_2 변이체의 경우 crSD2로부터의 비정상적인 스플라이싱이 폐지되었을 뿐만 아니라 crSD3 부위가 활성화되지 않았으며(도 21A; 라인 9 및 10), 이는 crSD3에 대한 잠복 스플라이싱 인핸서 또한 SL4 루프의 MSD-3KO_2의 단일 T>C 돌연변이에 의해 폐지되었음을 나타낸다.

[0648] 더 놀라운 것은 MSD-2KOm5 돌연변이가 crSD2 또는 crSD3로부터의 잠복 스플라이싱에 미치는 영향이었다. 이전 실시예에서 MSD-2KOm5 변이체는 다른 MSD-2KO 스플라이스 공여자 돌연변이체보다 덜 감쇠되었으며, 이것은 변형된 U1 snRNA에 의한 역가 회복에 있어 추가 이점으로 이어질 수 있는 것으로 나타났다. 이론에 얽매이고자 함이 없이, vRNA에 의한 U1 snRNA 동원이 안정성에 도움이 될 수 있다는 가설에 기초하여 MSD-2KOm5 변이체를 내인성 U1 snRNA와의 최대 어닐링이 (기능적 스플라이스 공여자 없이) 발생할 수 있도록 설계하였다(SL1 루프를 표적으로 하는 변형된 U1 snRNA의 사용에 대해 별도로 및 추가로). 도 20B는 스플라이스 공여자 부위를 포함하지 않는 돌연변이된 게놈에도 불구하고 표준, MSD-2KO 및 MSD-2KOm5 벡터 게놈 vRNA가 내인성 U1 snRNA와 어닐링될 것으로 예측되는 방법을 표시한다. 이론적으로 MSD-2KOm5 변이체는 MSD-2KO 변이체보다 더 많은 안정성(즉, 염기쌍에 참여하는 더 많은 수의 수소 결합)으로 내인성 U1 snRNA를 동원할 수 있으며, 아마도 표준 LV 게놈보다 더 큰 정도일 수 있지만, 결정적으로는 스플라이싱 이벤트로 이어지지 않는다. 추가적으로, MSD-2KOm5는 안정적인 SL2 루프가 형성될 수 있도록 설계하였으며, 이는 패키징 서열을 올바르게 접는 데 중요할 수 있기 때문이다. MSD-2KOm5(및 관련 MSD-3KO_2 및 MSD-4KO_2 변이체)에 대해 동일한 분석을 수행하였을 때, 놀랍게도 SL4 내의 crSD2 또는 crSD3로부터의 비정상적인 스플라이싱이 발생하지 않았다는 것이 발견되었다. 다시, 이론에 얽매이고자 함이 없이, SL2 루프에 대한 내인성 U1 snRNA 동원의 근접성은 SL4 루프의 잠복 스플라이스 부위의 인식을 방해할 수 있다고 제안된다.

[0649] 이 추가 연구에서 생산된 모든 스플라이스 공여자 돌연변이 벡터는 역가가 감소했지만(MSD-2KOm5가 가장 덜 감쇠됨), 모두 변형된 U1 snRNA의 트랜스로의 공급에 의해 실질적으로 부스트되었다(도 21B).

[0650] 종합하면, 이는 MSD 및 crSD1 부위 모두의 기능적 제거를 유도하는 마이너한 돌연변이가 패키징 서열로부터의 일반적인 비정상적인 스플라이싱의 강력한 감소를 유도하지만, 그것을 완전히 제거하기 위해서는 SL4 루프의 crSD2/3 부위의 보조 돌연변이가 필요할 수 있음을 나타낸다. MSD 및 crSD1 부위로부터의 비정상적인 스플라이싱의 기능적 절제를 유도하는 바람직한 변형은 MSD-2KOm5 변이체에 의해 설명된 바와 같이 보다 실질적인 변형

이며, 왜냐하면 이것은 하류 SL4 루프의 마이너한 잠복 스플라이스 공여자 부위의 활성화를 허용하지 않는 놀라운 효과를 또한 가지고, 이 변형을 포함하는 벡터는 이의 패키징 영역을 표적으로 하는 변형된 U1 snRNA의 존재 하에 가장 높은 역가로 만들어질 수 있기 때문이다.

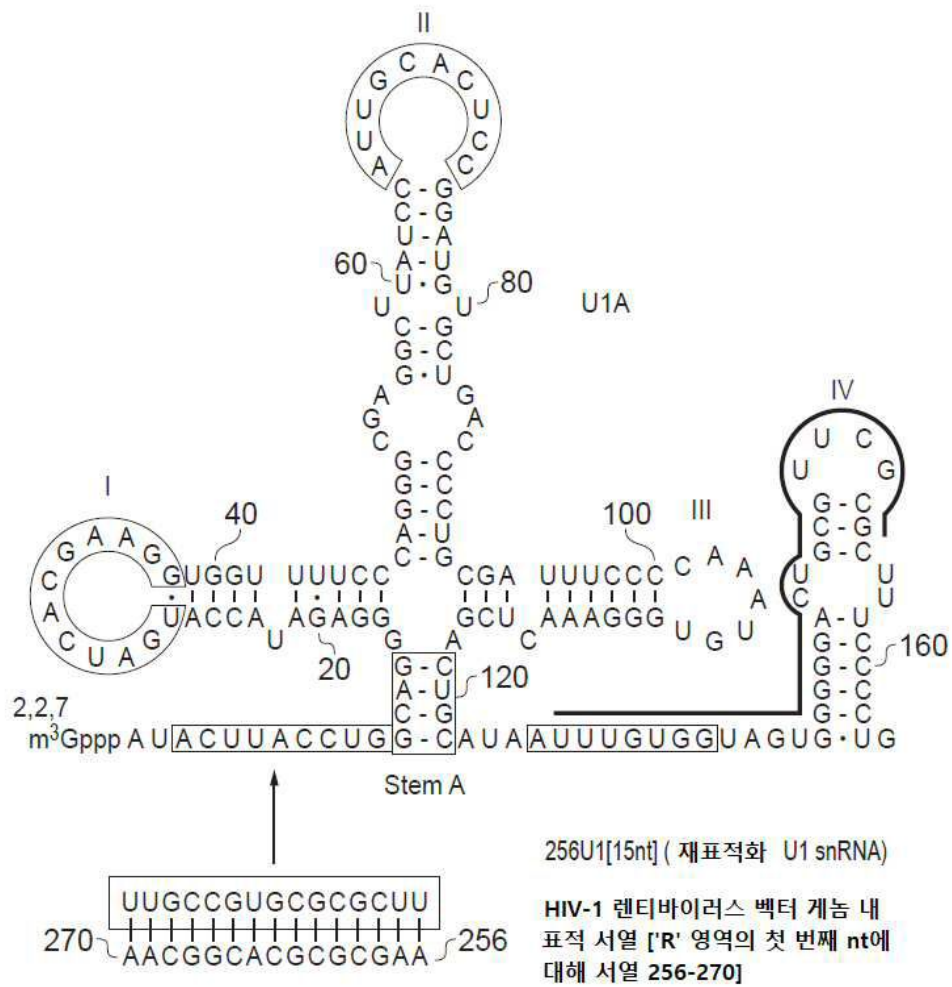
[0651] **실시예 13 - 스플라이스 공여자 돌연변이된 LV 게놈은 표적 세포의 상류 세포성 프로모터에 의한 전사 '리드-스루' 이벤트를 더 적게 발생시킨다.**

[0652] 표적 세포로의 렌티바이러스 벡터 통합은 반무작위이며, LV는 전사적으로 활성인 세포 유전자로의 통합을 선호하는 것으로 나타난다. 상류 세포성 프로모터로부터 통합된 LV로의 전사 리드-스루 또는 '리드-인'이 발생할 수 있고, LV 게놈 내의 서열과 동원/상호작용할 수 있다는 것이 다른 사람들에 의해 밝혀졌다(예를 들어 Moiani et al. J Clin Invest. 2012 May 1; 122(5): 1653-1666 참조). 본 발명의 경우 MSD-돌연변이된 LV의 추가적 이점은 표적 세포의 염색체 내 통합된 LV의 효과에 관한 개선된 환자 안전성일 것으로 예상된다. 야생형 HIV-1의 경우, 주요 스플라이스 공여자에 대한 내인성 U1 snRNA의 동원이 5'LTR 폴리아데닐화 신호의 역제를 초래한다는 것이 입증되었기 때문에, 리드-인 동안 표준 LV에서 발생하는 유사한 메커니즘은 U1 snRNA의 동원으로 인해 아마도 이를 악화시킬 수 있다. 따라서 LV 패키징 신호 내에서 MSD(및 잠복 SD)의 제거는 5'-SINLTR 암호화된 polyA 신호의 사용 증가로 인해 전사 리드-인 감소로 이어질 수 있다(도 22 참조).

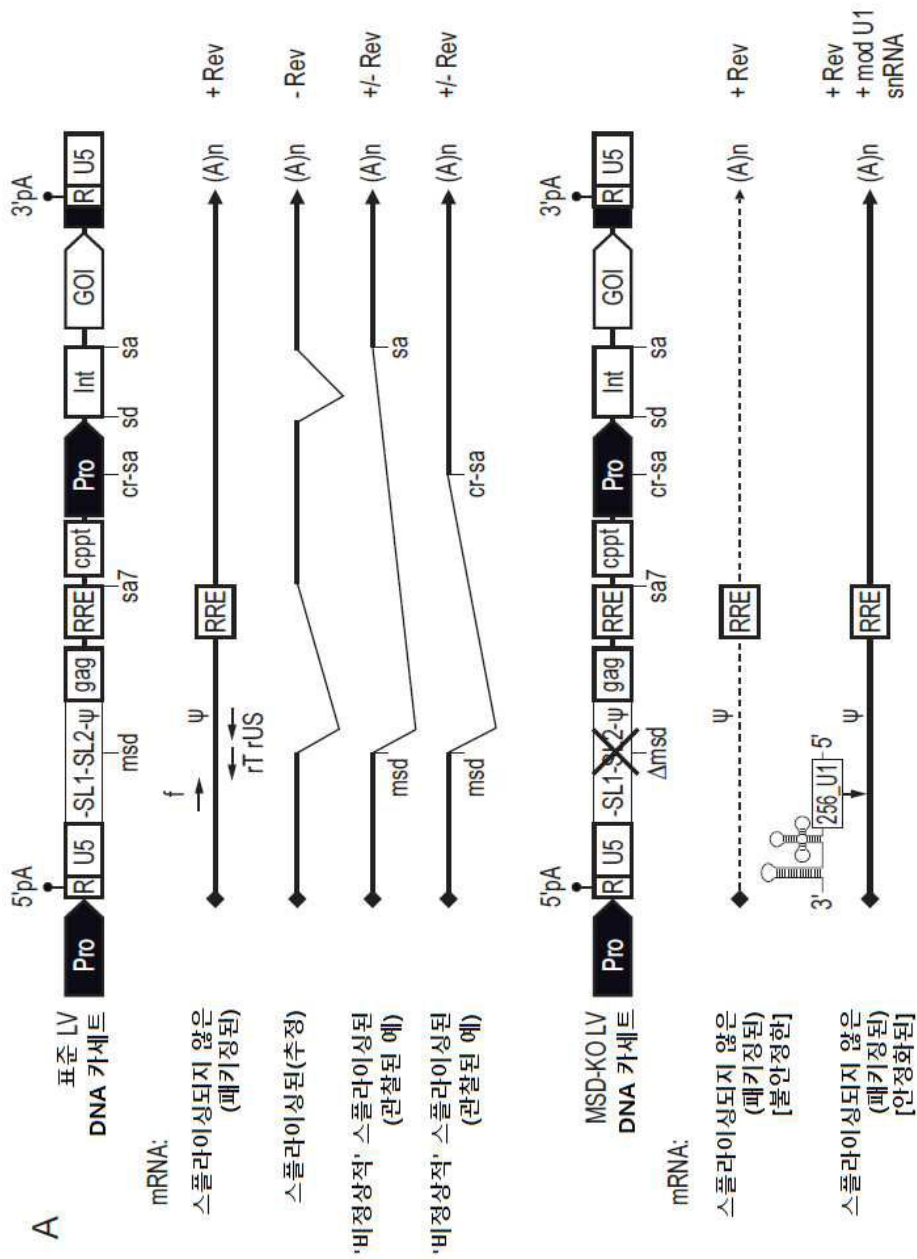
[0653] 이를 테스트하기 위해, 실시예 12에서 생성된 MSD-돌연변이된 LV 벡터 및 표준 LV 벡터(도 21B 참조)를 사용하여 HEK293T 세포 및 92BR 1차 세포(당나귀 섬유아세포)를 일치된 MOI에서 형질도입하였다(256U1의 존재하에서 생성한 단지 MSD 만-돌연변이된 LV 벡터 제제를 사용하였는데 이는 이들이 표준 벡터 제제와 유사한 역가를 가졌기 때문이었다). 형질도입된 세포 배양물을 10일 동안 계대하여 통합되지 않은 LV cDNA 및 이들로부터 생성된 잠재적 RNA를 모두 소실시켰고, 그런 다음 세포 RNA를 추출하고, DNase 처리하고, LV의 패키징 영역에 대한 프라이머를 사용하여 RT-qPCR 분석을 실시하였다(프라이머의 위치는 도 22 참조). 검출된 HIV 패키징 RNA의 카피를 GAPDH RNA 신호(RNA 로딩 대조군)로 정규화한 다음 통합된 LV DNA 카피 수에 대해 정규화하였다. 표준 LV 벡터(256U1 없이 생성됨)에 대해 검출된 HIV 패키징 RNA의 전체 정규화된 카피를 각 세포 유형에 대해 1로 설정하였고, 다른 모든 관련 정규화된 카피 수를 이것에 대해 상대적으로 설정하였다. 도 23은 이 분석의 결과를 보여주며, 실제로 LV 게놈 내에 돌연변이된 MSD 및 잠복 SD를 갖는 구성이 전사 리드-인 이벤트의 2배에서 5배 사이의 감소를 초래했음을 나타낸다. 이러한 모든 데이터는 MSD/crSD 돌연변이된 LV가 형질도입된 세포에서 표준 LV에 대해 더 나은 안전성 프로파일을 생성하고, LV 백본 동원 및/또는 LV 서열과의 상호 작용을 허용하는 경향이 적다는 것을 나타낸다.

도면

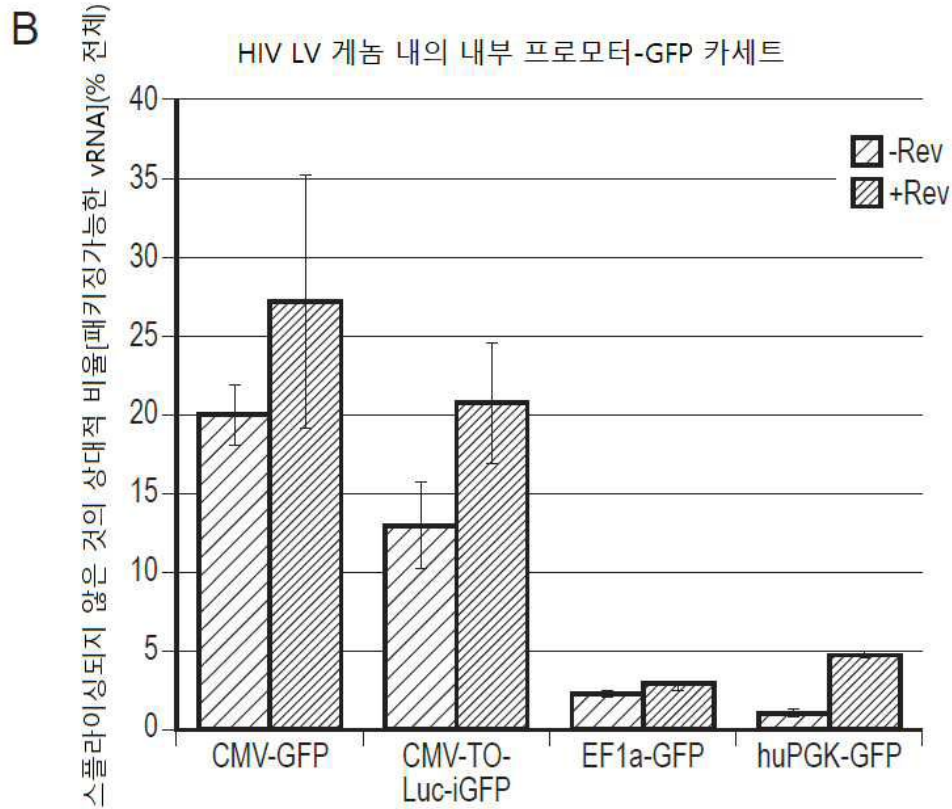
도면1



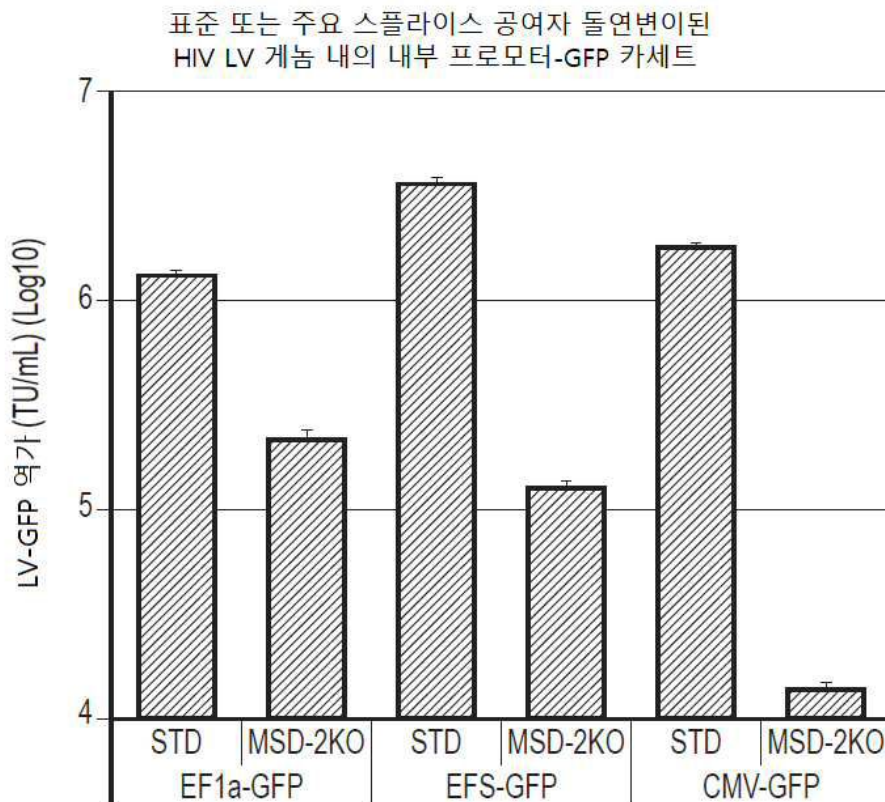
도면2a



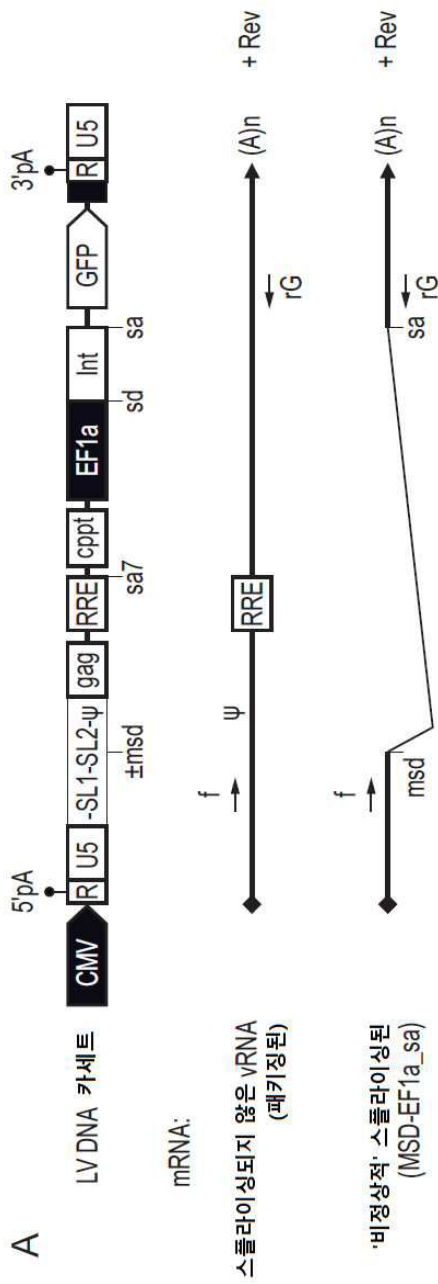
도면2b



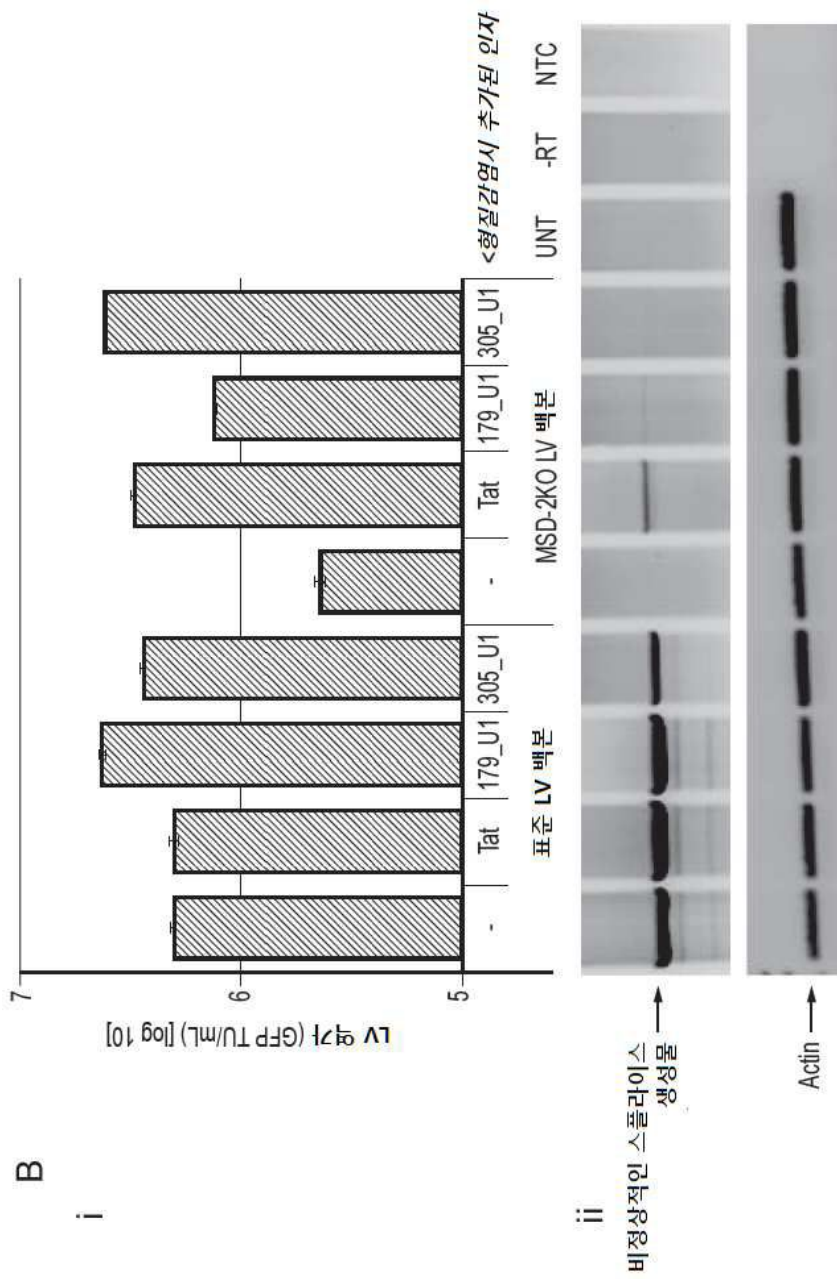
도면3



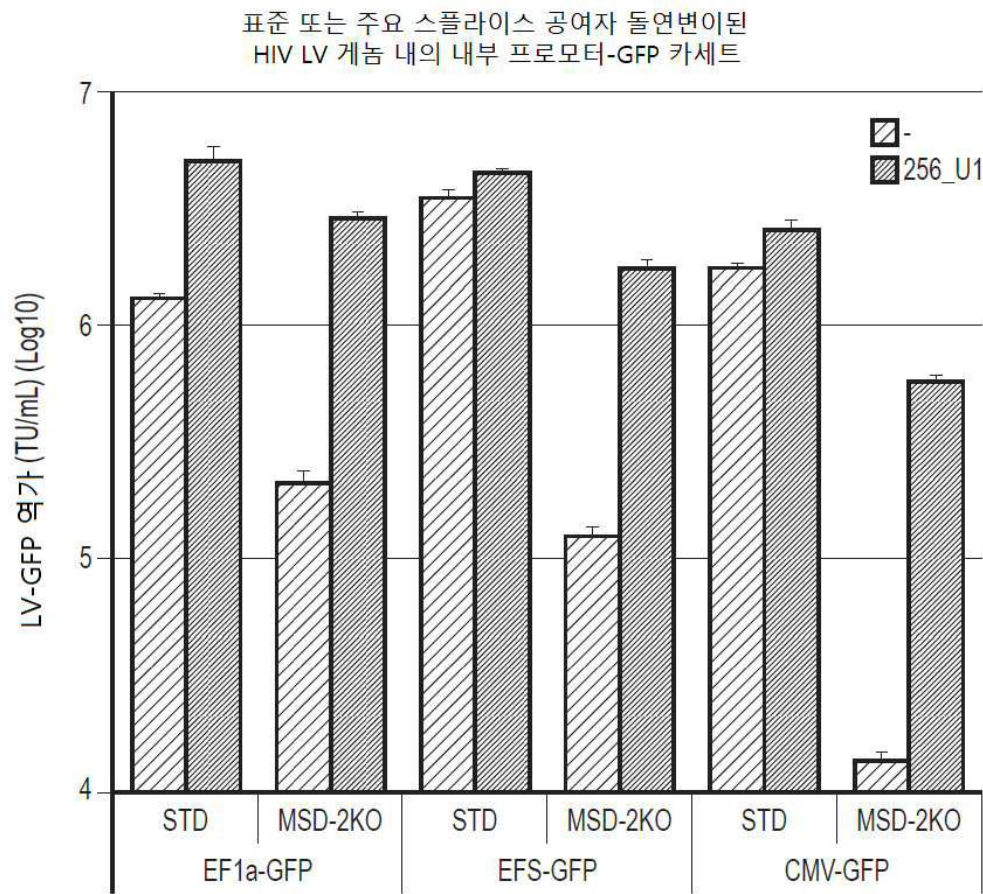
도면4a



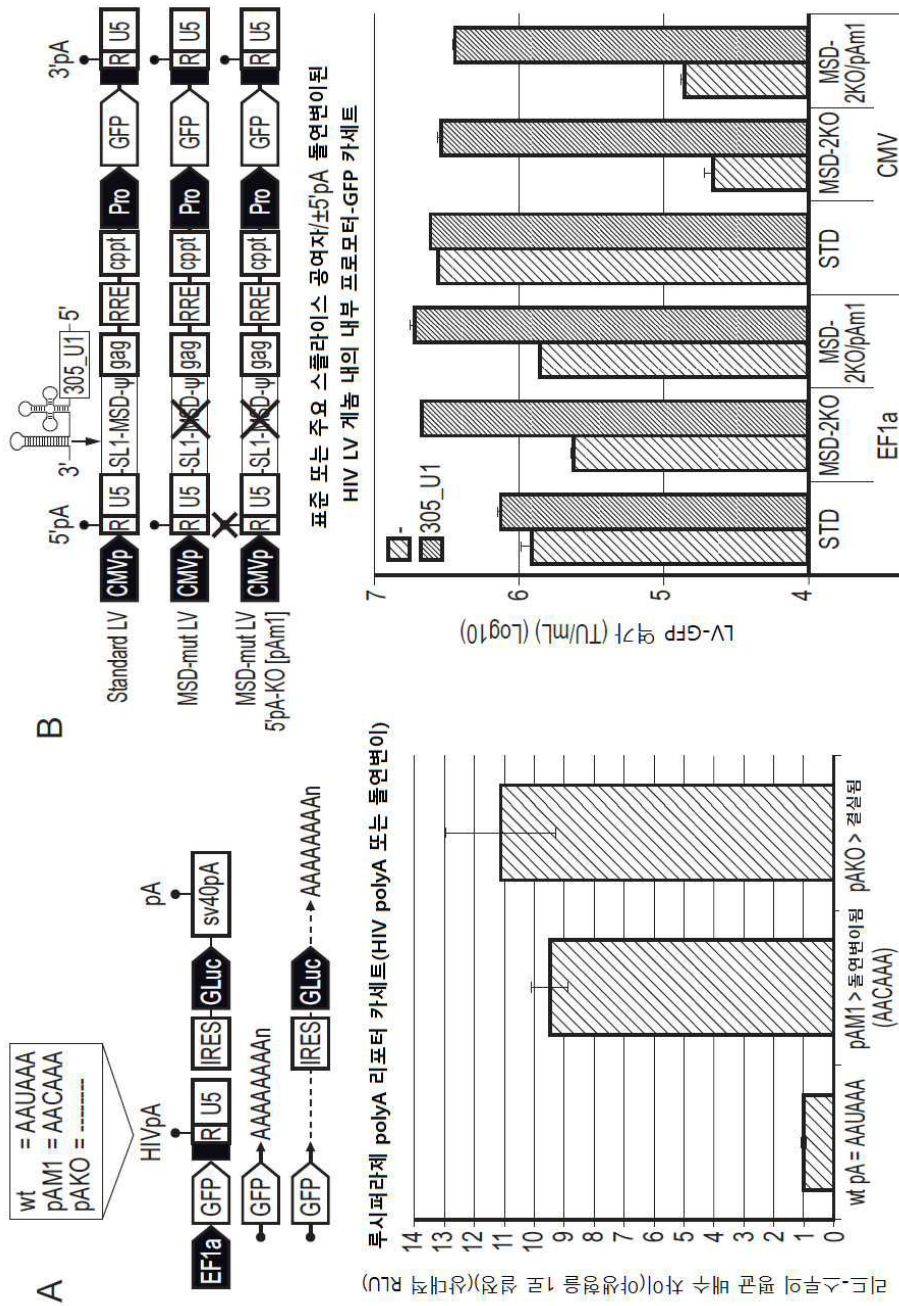
도면4b



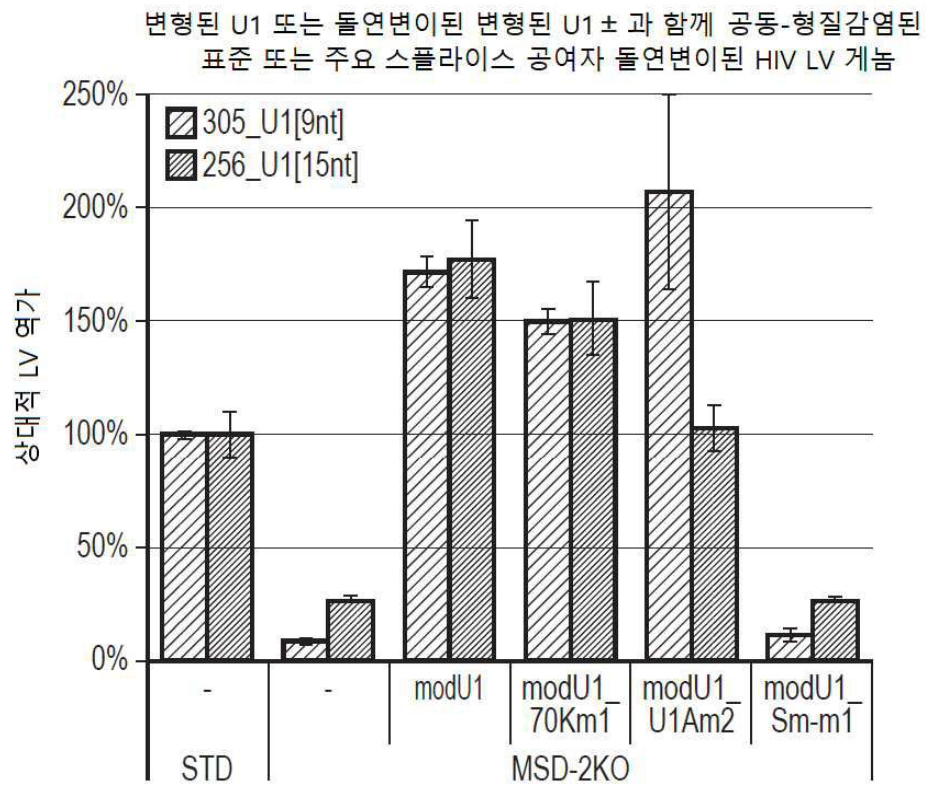
도면5



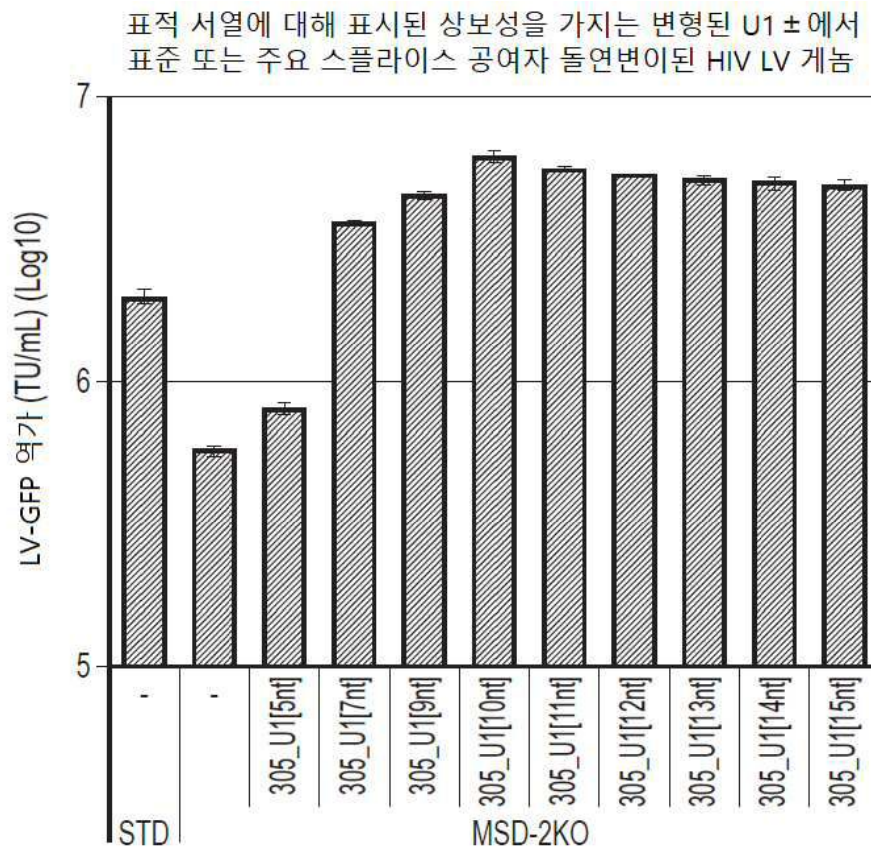
도면6



도면7

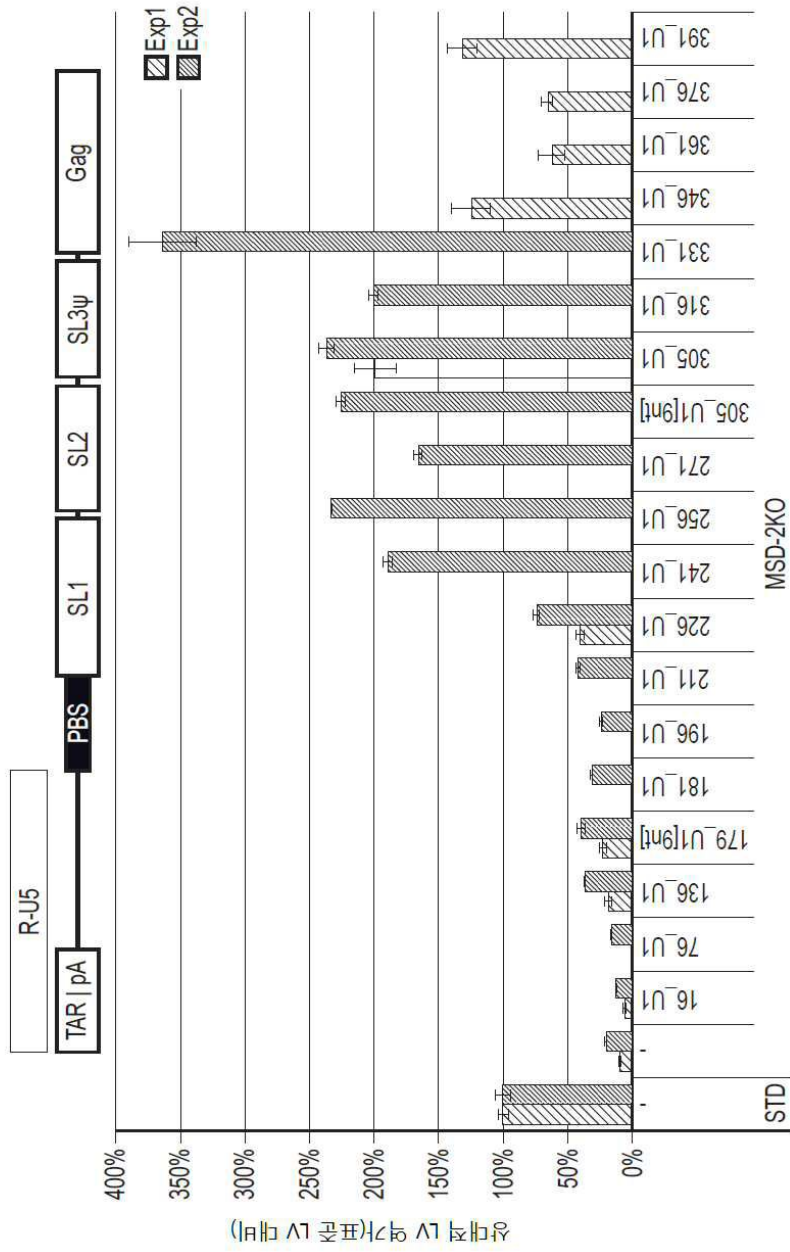


도면8

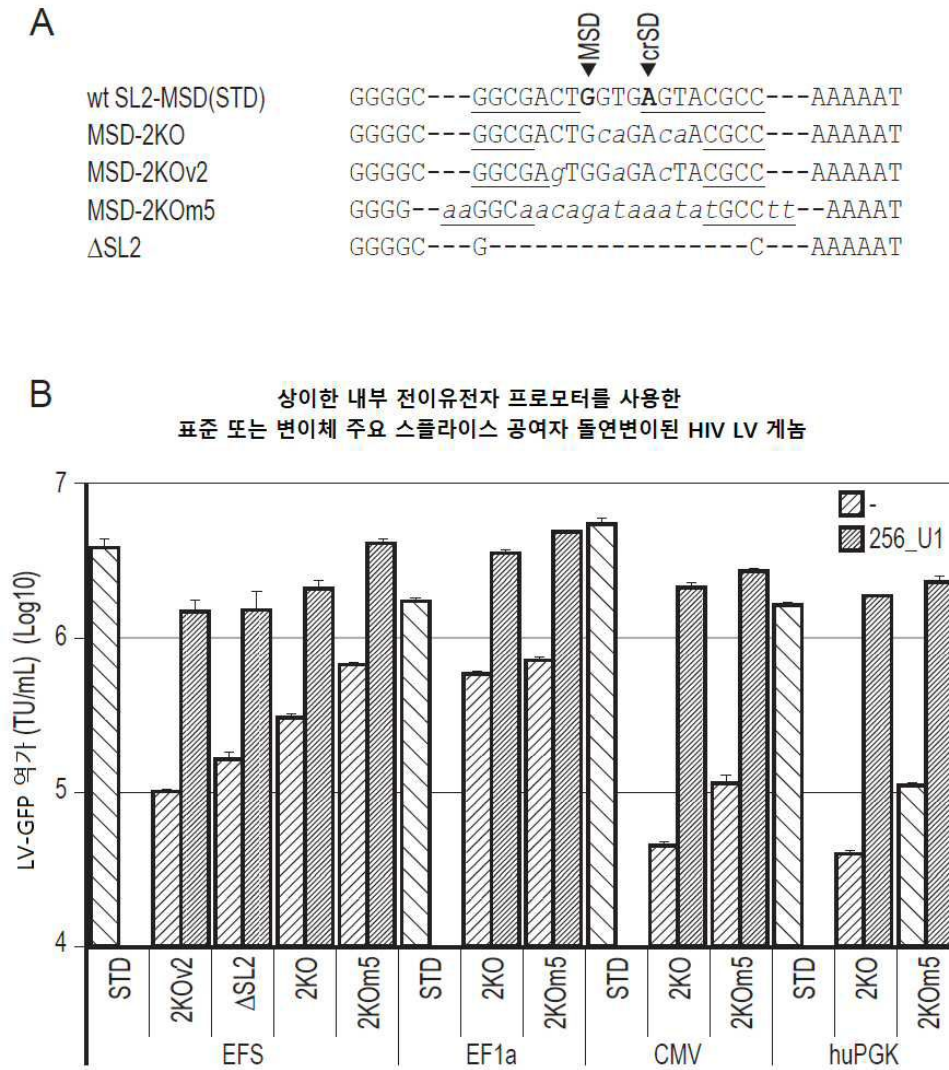


도면9

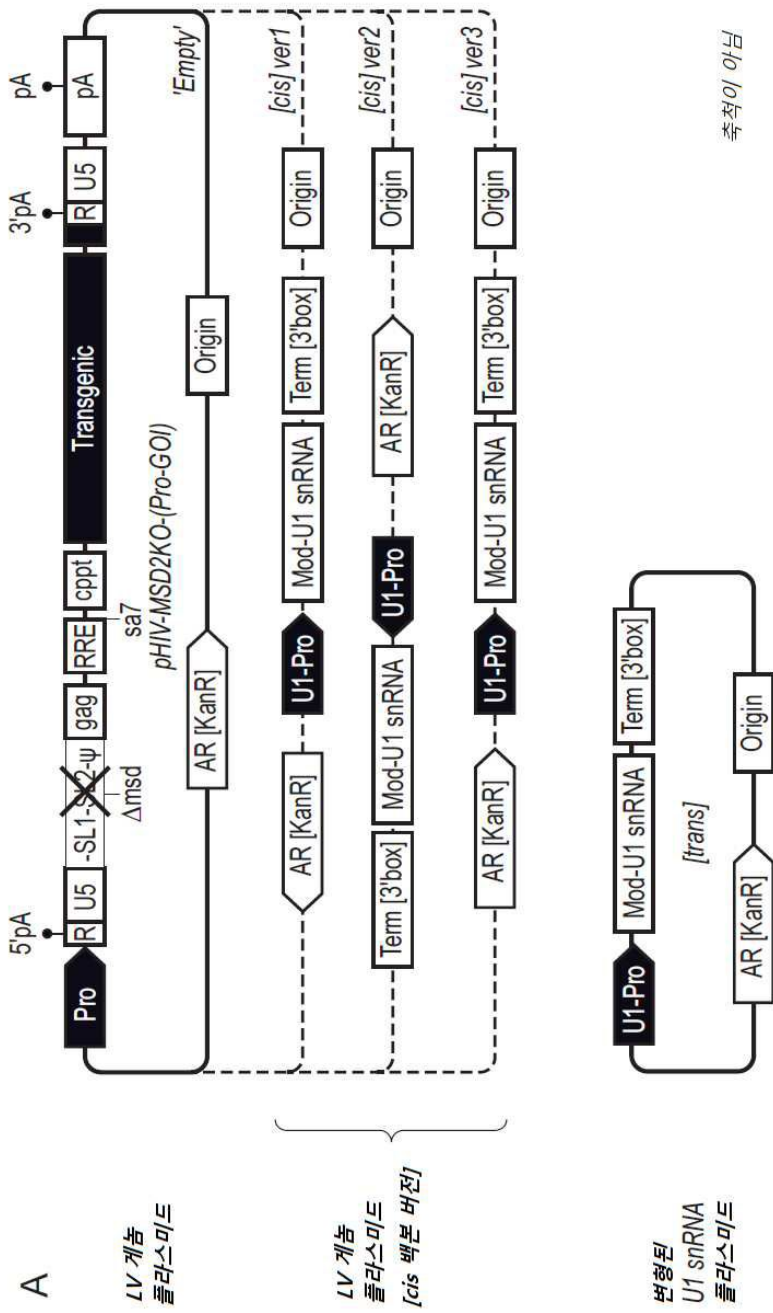
VRNA의 5' 패키징 영역을 표적으로 하는 변형된 U1 ± 과 함께 공동-형질감염된
표준 또는 주요 스플라이스 공여자 돌연변이된 HIV LV 계통



도면10



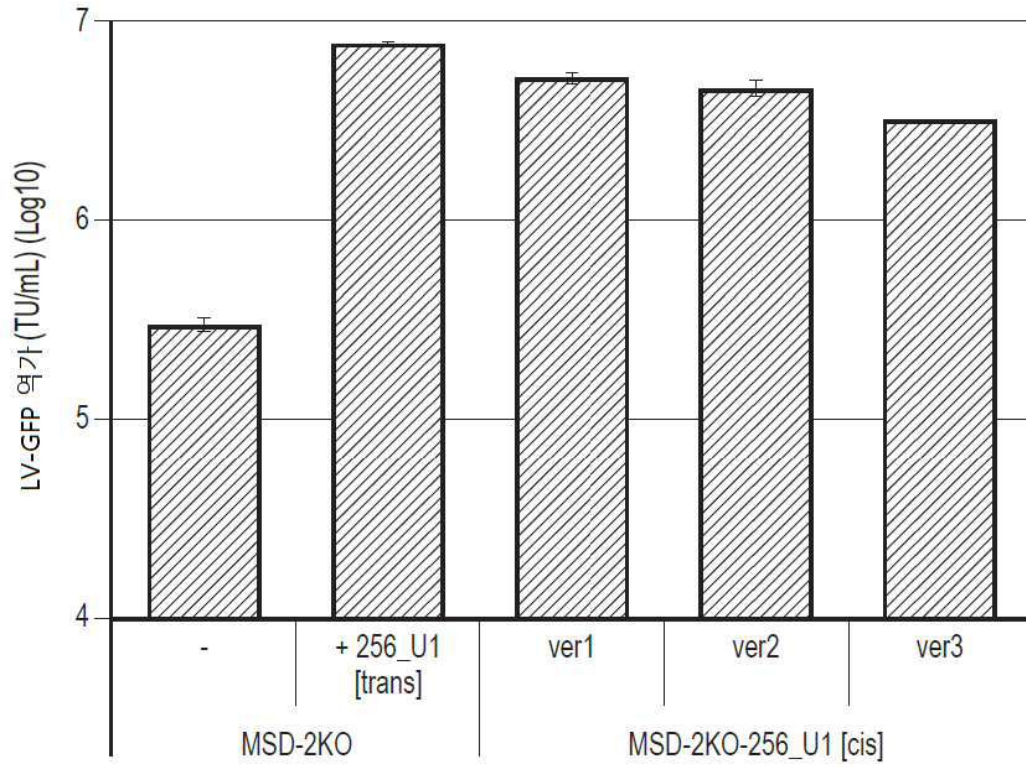
도면11a



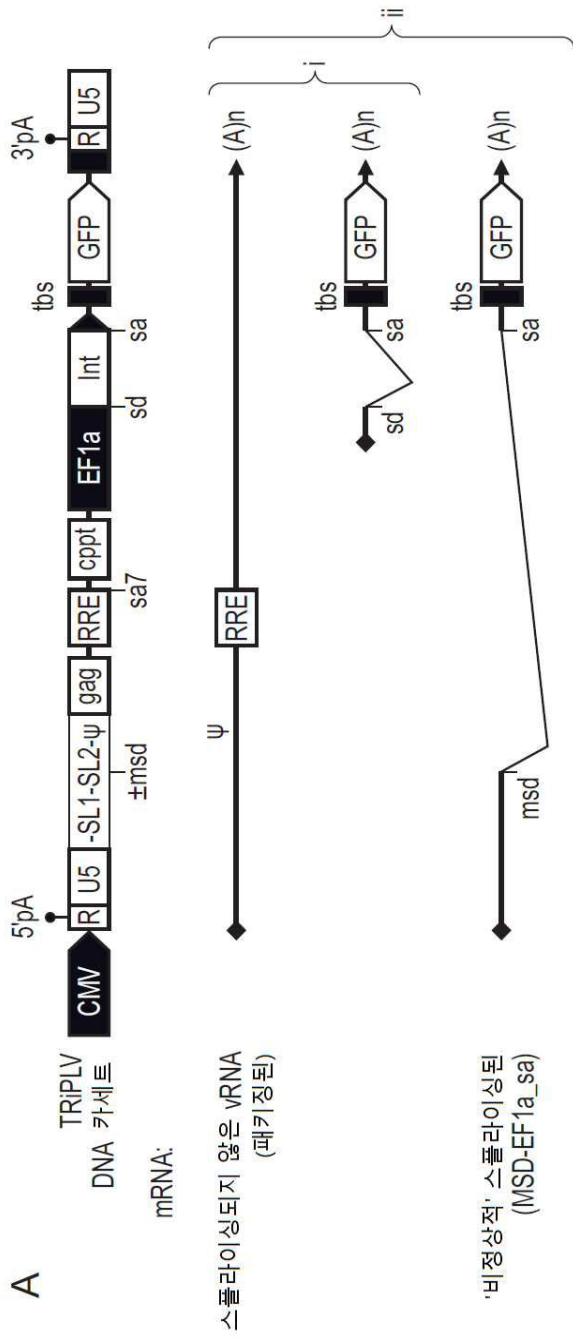
도면11b

B

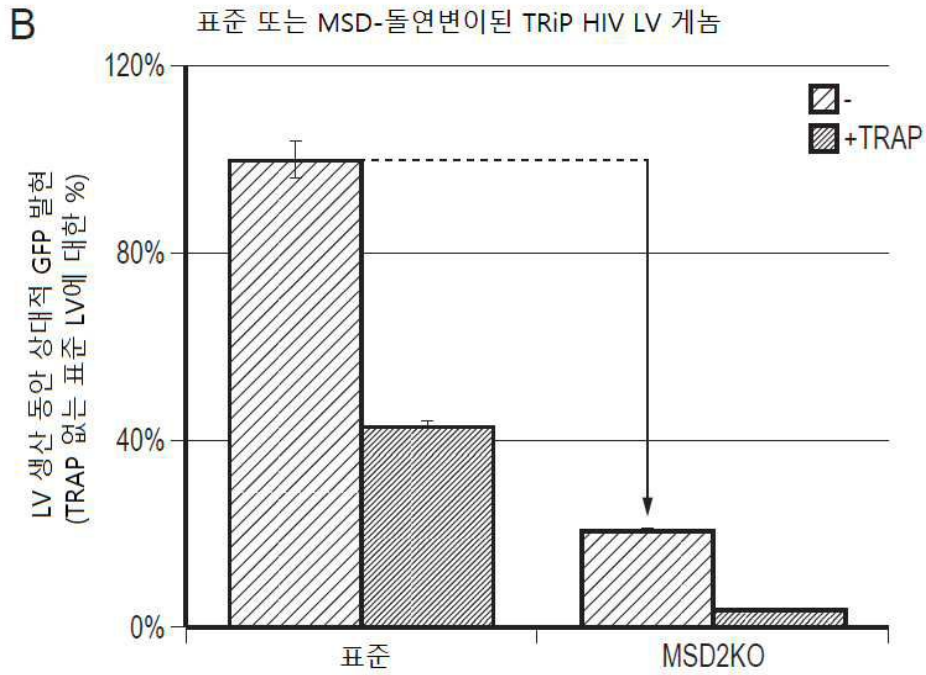
트랜스로 또는 시스로 변형된 U1 snRNA 카세트 ± 에서
주요 스플라이스 공여자 돌연변이된 HIV LV 계통



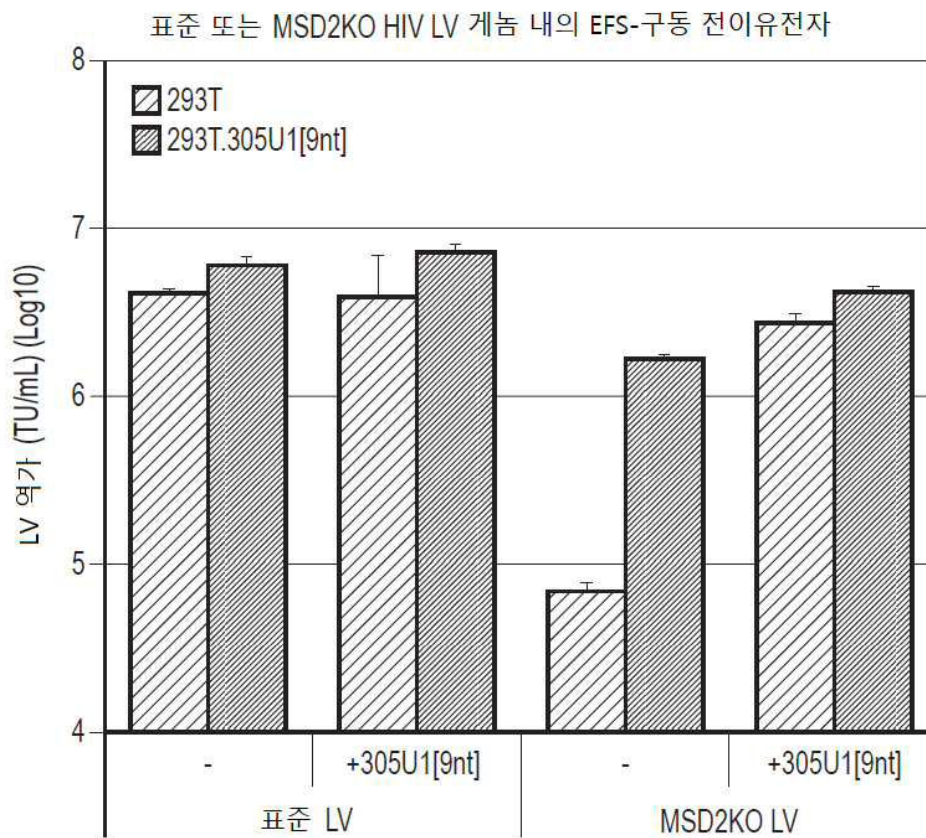
도면12a



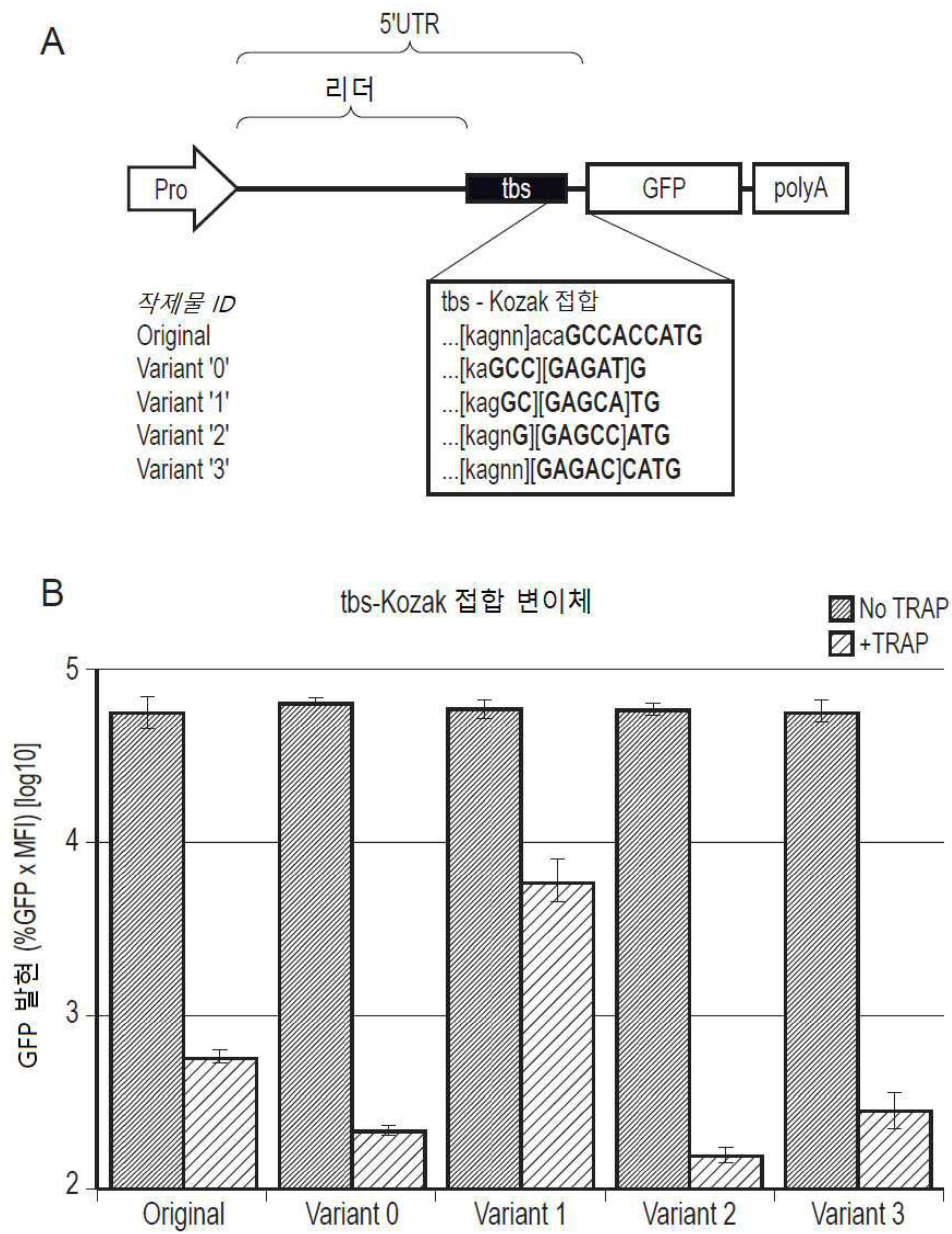
도면12b



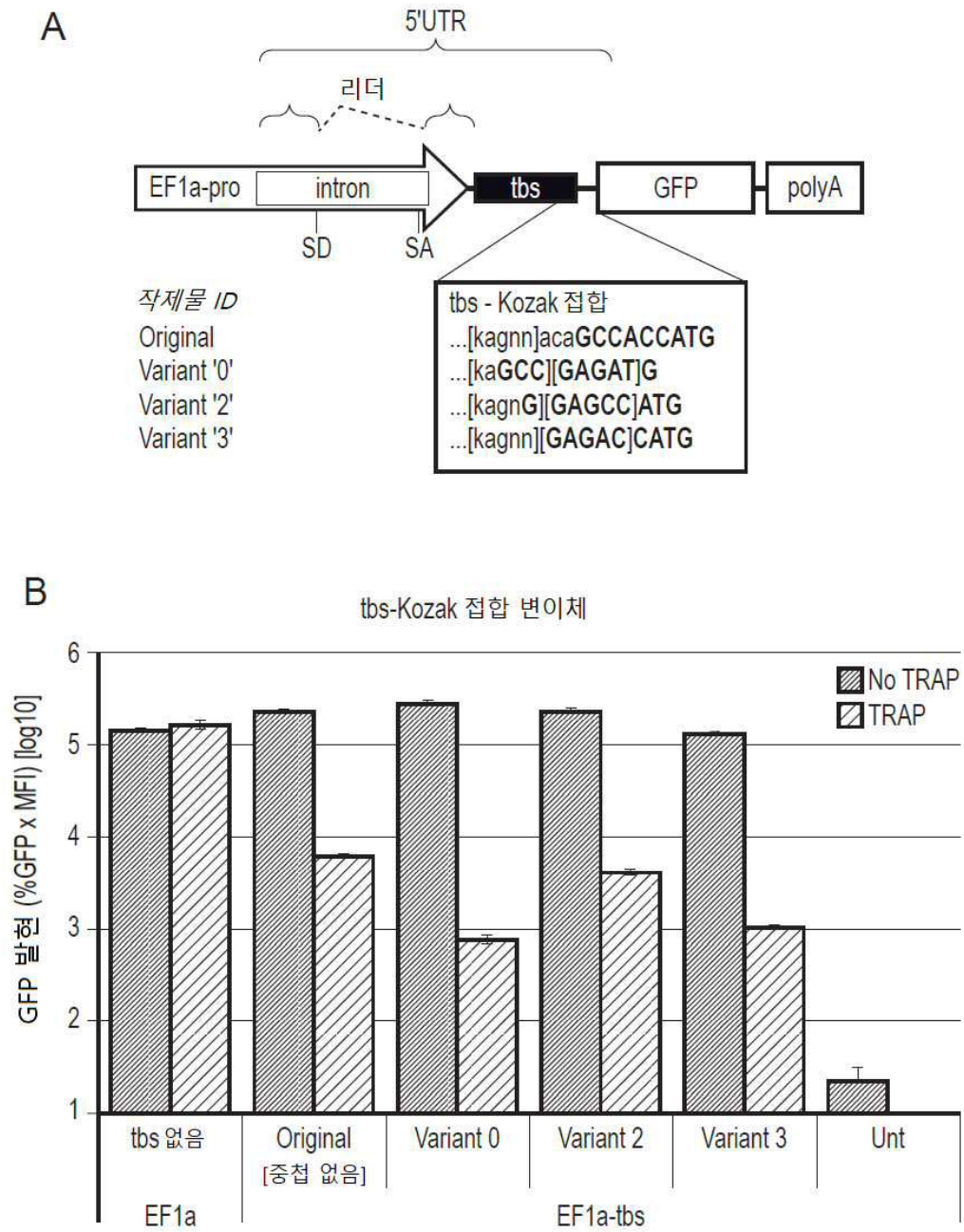
도면13



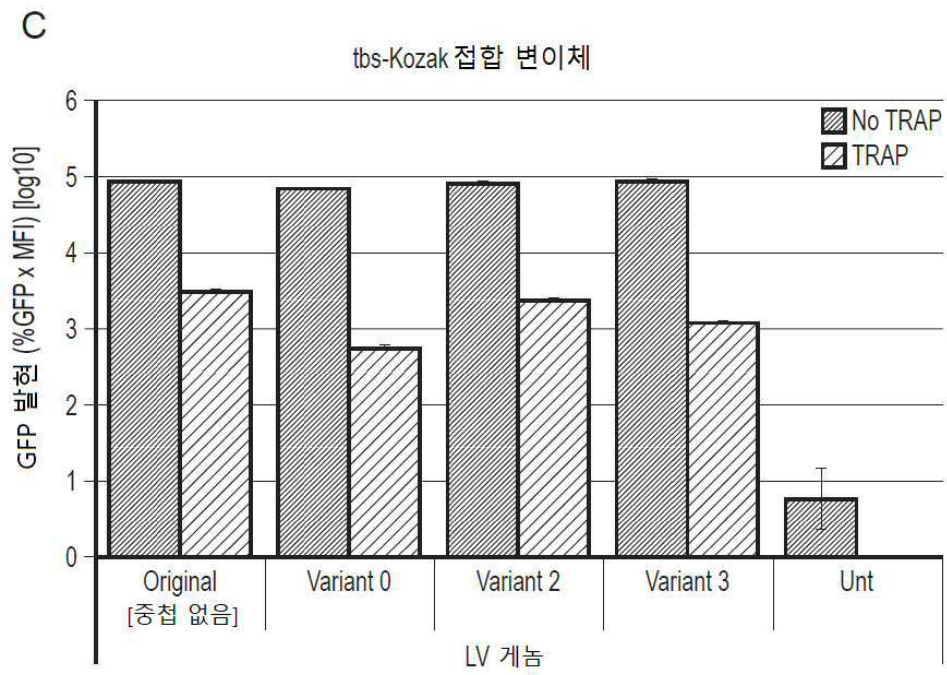
도면14



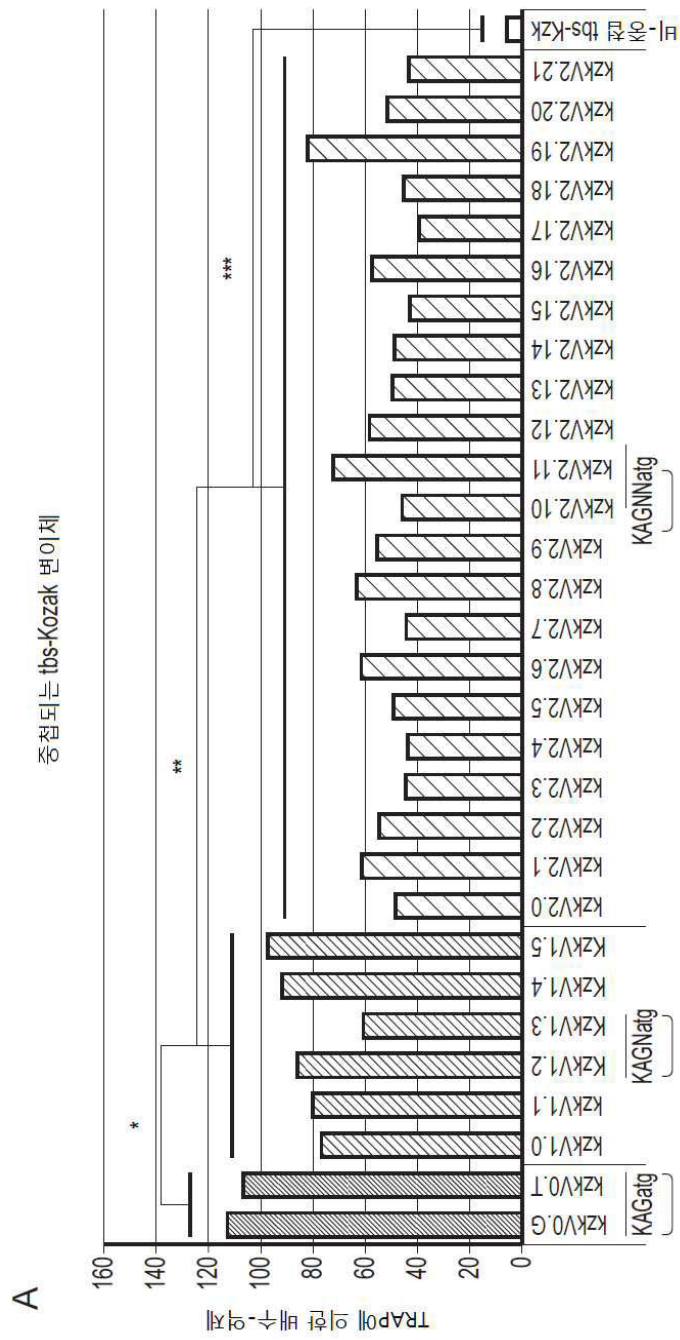
도면16a



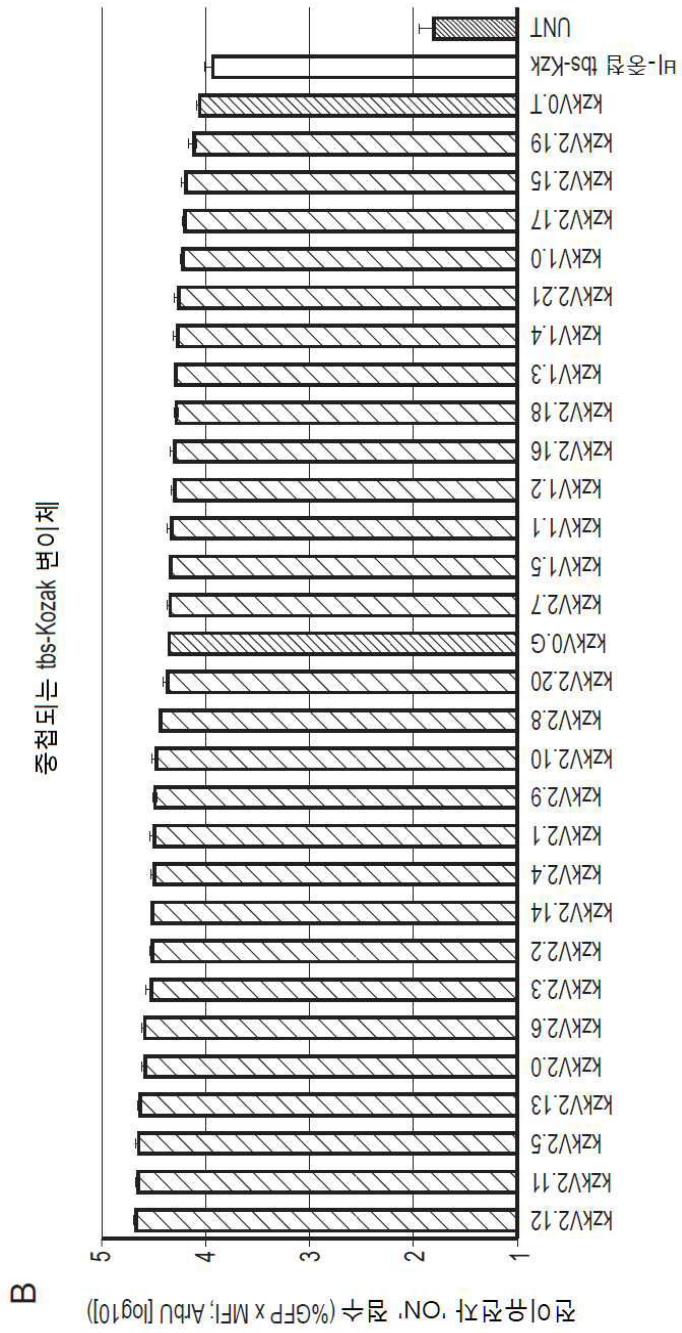
도면16b



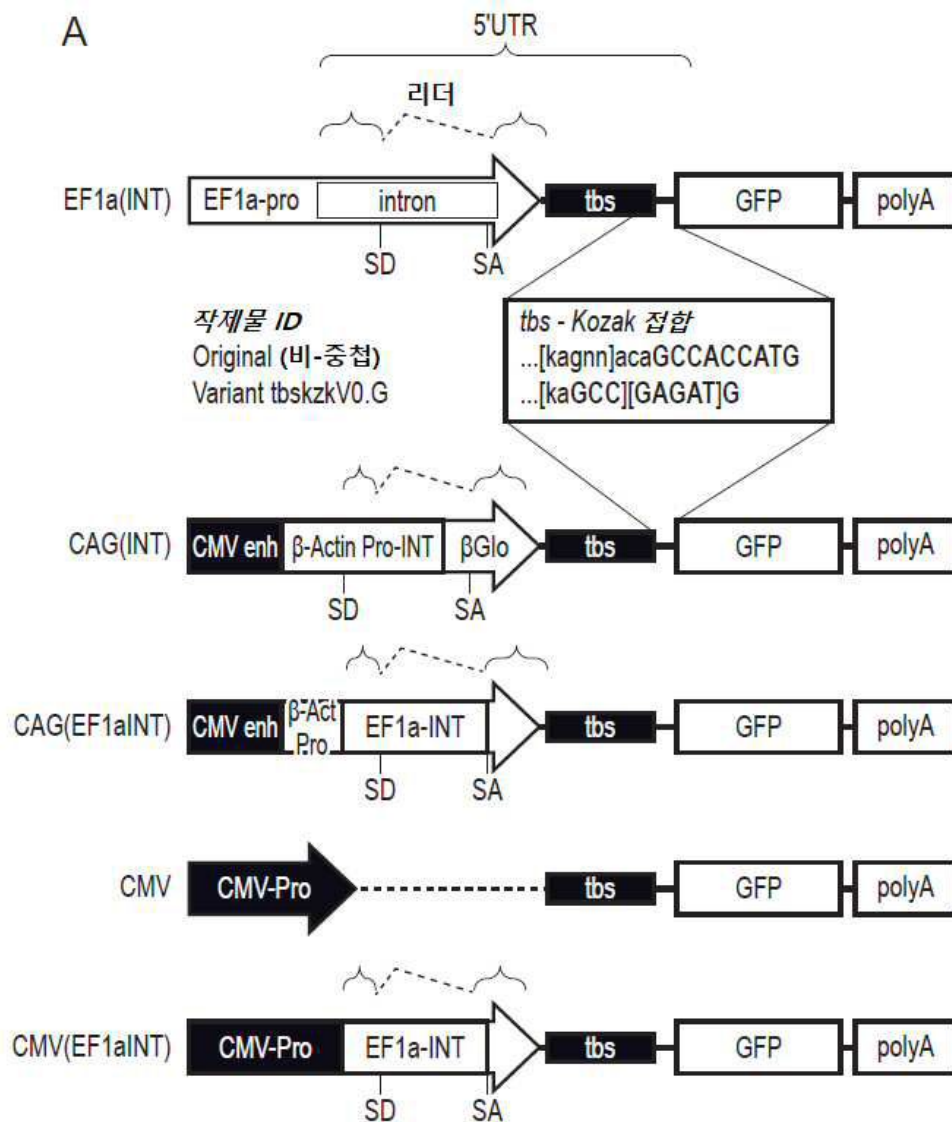
도면17a



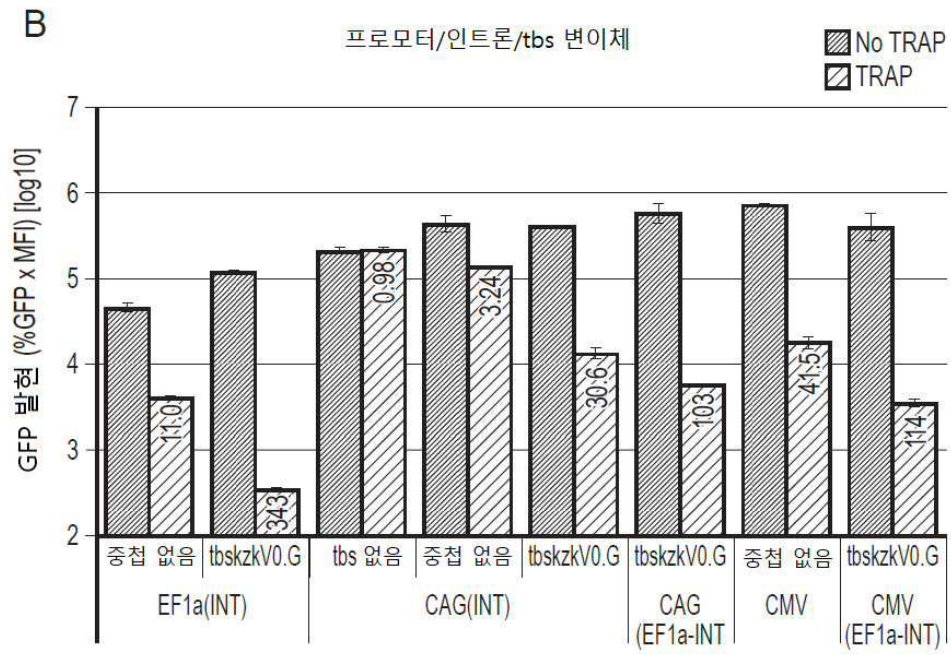
도면17b



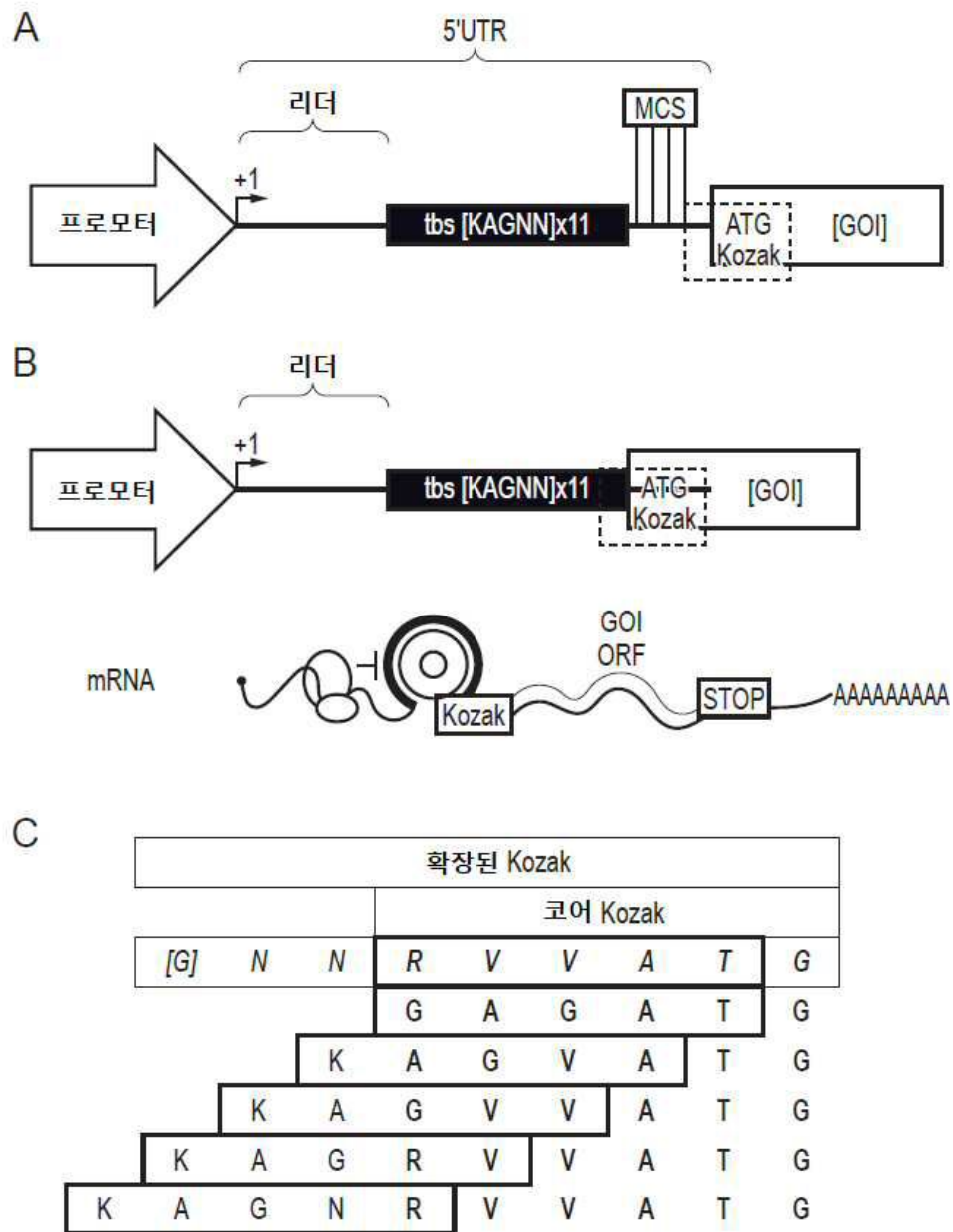
도면18a



도면18b



도면19



도면20a

A

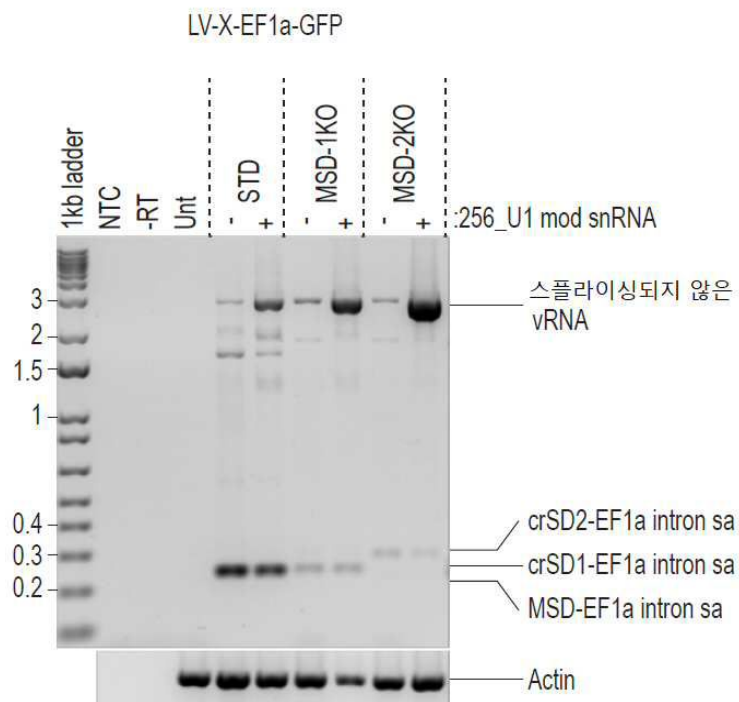
	SL2 루프	SL4 루프
wt SL2-MSD-SL4(STD)	5' GGC--GGCGACTGGT AG TACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGGT AG TACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-1KO	5' GGC--GGCGACTGcaGAGTACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGcaGAGTACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-2KO	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-3KO_1	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-3KO_2	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-4KO_1	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-4KO_2	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GGC--GGCGACTGcaGAcaACGCC--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-2KOm5	5' GG--aaGGCaaCagataaaatatGCCtt--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GG--aaGGCaaCagataaaatatGCCtt--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-3KOm5_2	5' GG--aaGGCaaCagataaaatatGCCtt--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GG--aaGGCaaCagataaaatatGCCtt--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG
MSD-4KOm5_2	5' GG--aaGGCaaCagataaaatatGCCtt--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG	5' GG--aaGGCaaCagataaaatatGCCtt--AAAAAT--/-GAGATGGGTGCAGAGCGTCAG

B

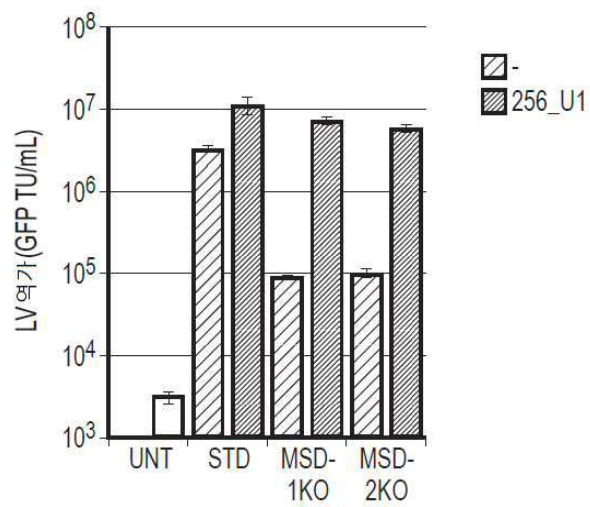
Endog. U1 snRNA	3' guccaucaua-5'
wt SL2-MSD(STD)	5' g'gcgacuggug a guacgcc...3'
Endog. U1 snRNA	3' guccaucaua-5'
MSD-2KO	5' g'gcgacugcaga a acgcc...3'
Endog. U1 snRNA	3' guccaucaua-5'
MSD-2KOm5	5' g'gcacagauaaauaugcc...3'

도면20b

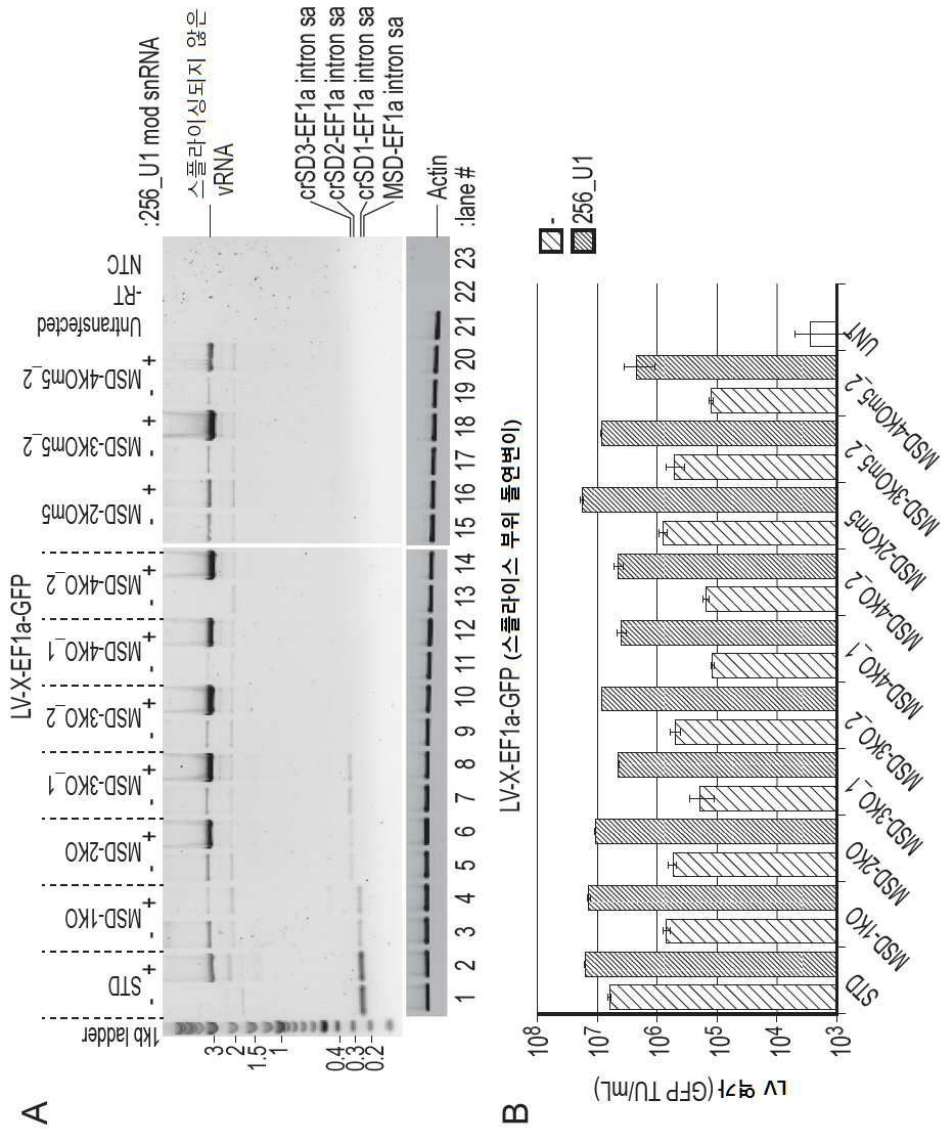
C



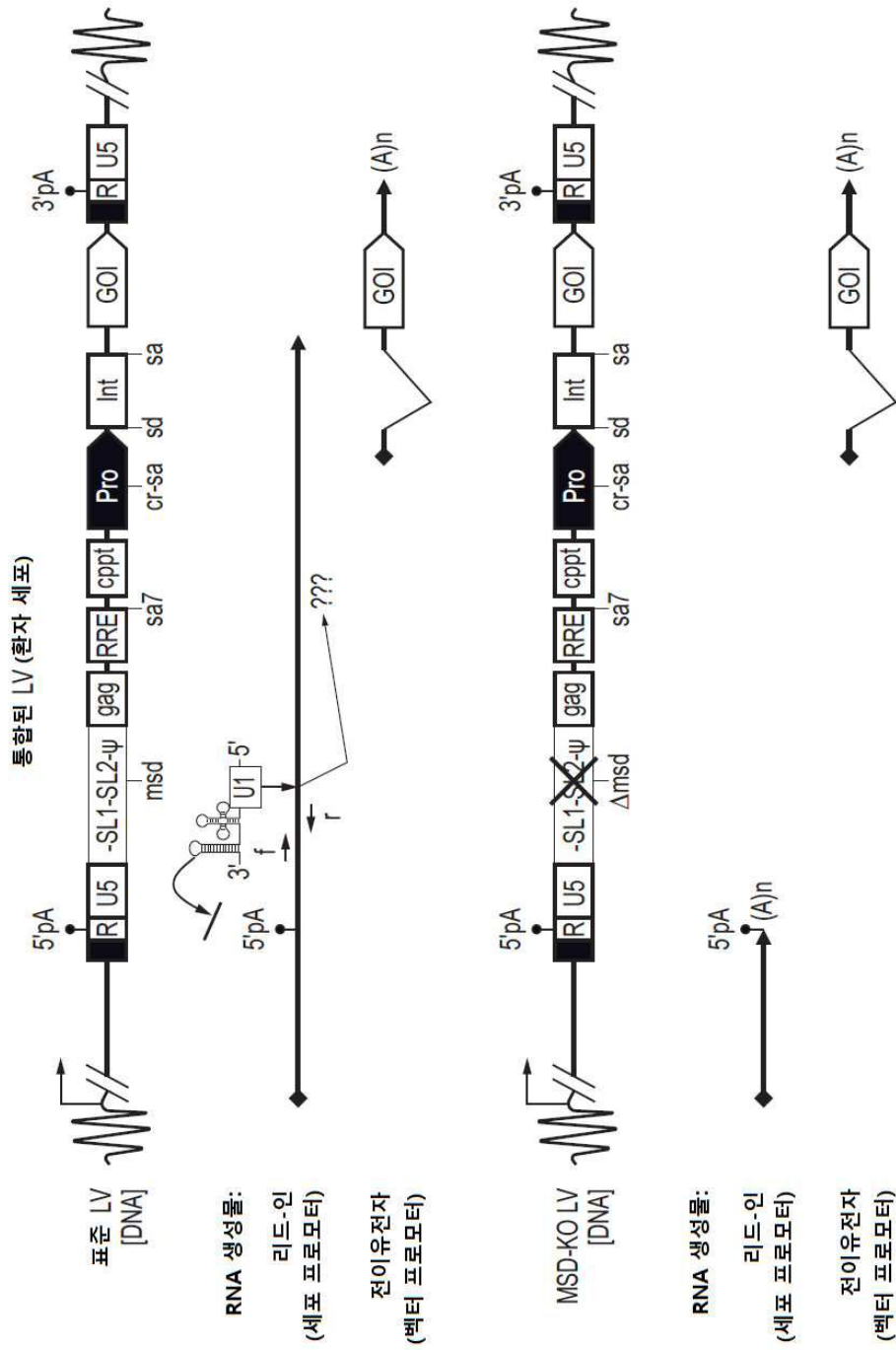
D



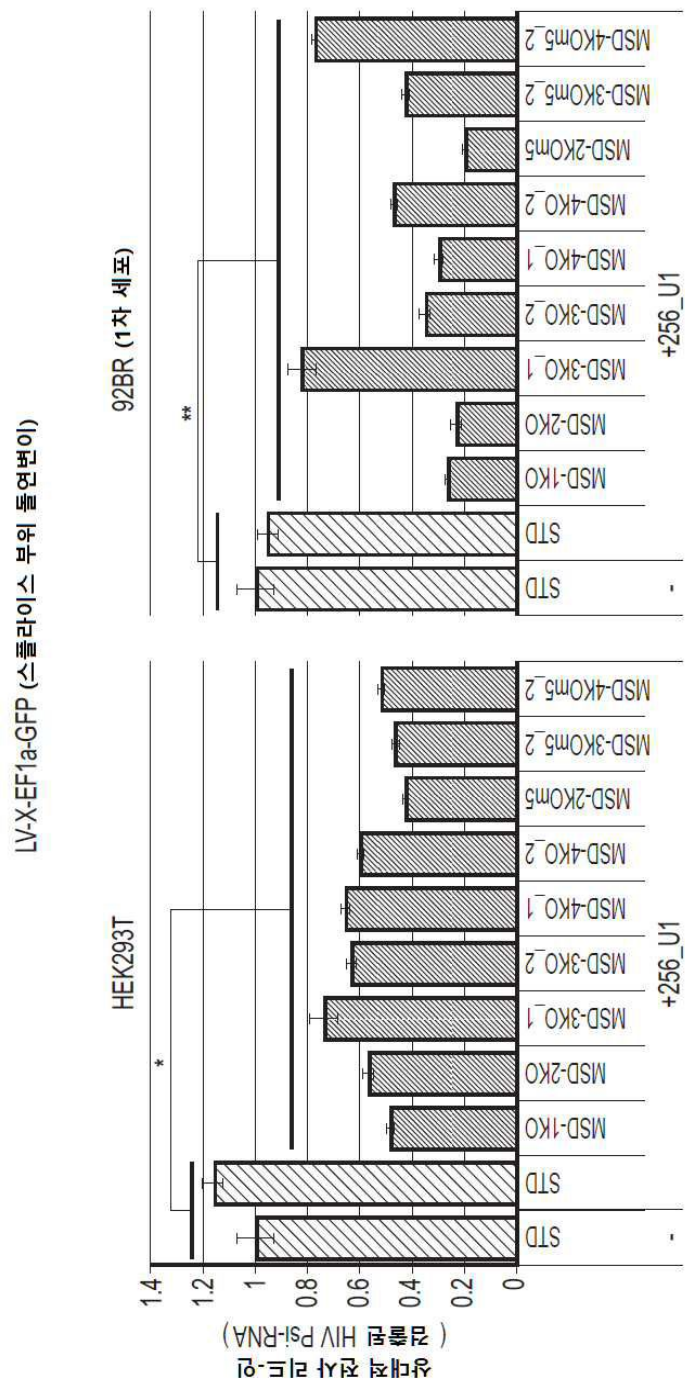
도면21



도면22



도면23



서열 목록

SEQUENCE LISTING

<110> Oxford BioMedica (UK) Limited

<120> PRODUCTION OF LENTIVIRAL VECTORS

<130> P118969PCT

<140> PCT/GB2021/050247

<141> 2021-02-04
 <150> GB 2001996.4
 <151> 2020-02-13
 <150> GB 2017820.8
 <151> 2020-11-11
 <160> 242
 <170> PatentIn version 3.5
 <210> 1
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> splice donor region
 <400> 1
 ggggcggcga ctggtgagta cgccaaaaat 30
 <210> 2
 <211> 30
 <212>
 > DNA
 <213> Artificial Sequence
 <220><223> mutated splice donor region, MSD-2KO
 <400> 2
 ggggcggcga ctgcagacaa cgccaaaaat 30
 <210> 3
 <211> 8
 <212> DNA
 <213> Artificial Sequence
 <220><223> splice donor site consensus sequence
 <400> 3
 tggtragt 8
 <210> 4
 <211> 5
 <212> DNA
 <213> Artificial Sequence
 <220><223> major splice donor consensus sequence
 <400> 4

ctggt	5
<210> 5	
<211> 6	
<212> DNA	
<213> Artificial Sequence	
<220><223> splice donor region	
<400> 5	
cagaca	6
<210> 6	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> mutated splice donor region	
<400> 6	
ggcgactgca gacaacgcc	19
<210> 7	
<211> 9	
<212> DNA	
<213> Artificial Sequence	
<220><223> mutated splice donor region	
<400> 7	
gtggagact	9
<210> 8	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> mutated splice donor region	
<400> 8	
ggcgagtgga gactacgcc	19
<210> 9	
<211> 19	
<212> DNA	
<213> Artificial Sequence	

<220><223> splice donor region

<400> 9

ggcgactggt gagtacgcc 19

<210> 10

<211> 5

<212> DNA

<213> Artificial Sequence

<220><223> cryptic splice donor site consensus sequence

<400> 10

tgagt 5

<210> 11

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> mutated splice donor region, MSD-2K0v2

<400> 11

ggggcggcga gtggagacta cgccaaaaat 30

<210> 12

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> mutated splice donor region, MSD-2K0m5

<400> 12

ggggaaggca acagataaat atgccttaa at 32

<210> 13

<211> 7

<212> DNA

<213> Artificial Sequence

<220><223> major splice donor and cryptic splice donor region core sequence

<400> 13

gtgagta 7

<210> 14

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> mutated splice donor region

<400> 14

aaggcaacag ataaatatgc ctt 23

<210> 15

<211> 642

<212> DNA

<213> Artificial Sequence

<220><223> 256U1 snRNA sequence

<400>

> 15

taaggaccag ctcttttggg agagaacaga cgcaggggcg ggagggaaaa agggagaggc 60

agacgtcact tccccttggc ggctctggca gcagattggt cggttgagtgc gcagaaaggc 120

agacggggac tgggcaaggc actgtcgggtg acatcacgga cagggcgact tctatgtaga 180

tgaggcagcg cagaggctgc tgcttcgcca cttgctgctt caccacgaag gagttcccgt 240

gccctgggag cgggttcagg accgctgac ggaagtgaga atcccagctg tgtgtcaggg 300

ctggaaaggg ctcgggagtgc cgcggggcaa gtgaccgtgt gtgtaaagag tgaggcgtat 360

gaggctgtgt cggggcagag gcccaagatc tcatttgccg tgcgcgcttg caggggagat 420

accatgatca cgaaggtggt tttcccaggc cgaggcttat ccattgcact ccggatgtgc 480

tgacccttgc gatttcccca aatgtgggaa actcgactgc ataatttgtg gtagtggggg 540

actgcgttcg cgctttcccc tggtttcaaa agtagactgt acgctaaggg tcatatcttt 600

ttttgttttg gtttgtgtct tggttggcgt cttaaagtgt aa 642

<210> 16

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> HIV-1 target sequence

<400> 16

gaccagatct gagcc 15

<210> 17

<211>

> 17

<212> DNA

<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 17	
atggctcaga tctggtc	17
<210> 18	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 18	
tgggagctct ctggc	15
<210> 19	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 19	
atgccagaga gctccca	17
<210> 20	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 20	
taaagcttgc cttga	15
<210> 21	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 21	
attcaaggca agcttta	17
<210> 22	
<211> 15	

<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 22	
tagagatccc tcaga	15
<210> 23	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 23	
attctgaggg atctcta	17
<210> 24	
<211> 9	
<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 24	
gcagtggcg	9
<210> 25	
<211> 11	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400>	
> 25	
atgccactg c	11
<210> 26	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 26	
agtggcgccc gaaca	15

<210> 27
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220><223> U1 snRNA target-annealing sequence
 <400> 27
 attgttcggg cgccact 17
 <210> 28
 <211> 15
 <212> DNA
 <213> Artificial Sequence
 <220><223> HIV-1 target sequence
 <400> 28
 gggacttgaa agcga 15
 <210> 29
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220><223> U1 snRNA target-annealing sequence
 <400> 29
 attcgctttc aagtccc 17
 <210> 30
 <211> 15
 <212> DNA
 <213> Artificial Sequence
 <220><223> HIV-1 target sequence
 <400> 30
 aagggaacc agagg 15
 <210> 31
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220><223> U1 snRNA target-annealing sequence

<400> 31
atcctctggg ttcctt 17
<210> 32
<211> 15
<212> DNA
<213> Artificial Sequence
<220><223> HIV-1 target sequence
<400> 32
ggactcggct tgctg 15
<210> 33
<211> 17
<212> DNA
<213> Artificial Sequence
<220><223> U1 snRNA target-annealing sequence
<400>
> 33
atcagcaagc cgagtcc 17
<210> 34
<211> 15
<212> DNA
<213> Artificial Sequence
<220><223> HIV-1 target sequence
<400> 34
aagcgcgcac ggcaa 15
<210> 35
<211> 17
<212> DNA
<213> Artificial Sequence
<220><223> U1 snRNA target-annealing sequence
<400> 35
attgcccgtg cgcgctt 17
<210> 36
<211> 15
<212> DNA

<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 36	
gaggcgaggg gcggc	15
<210> 37	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 37	
atgccgcccc tcgcctc	17
<210> 38	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 38	
gactggtgag tacgc	15
<210> 39	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 39	
atgcgtactc accagtc	17
<210> 40	
<211> 11	
<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 40	
aattttgact a	11
<210> 41	
<211> 11	

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400>

> 41

atgtcaaaat t 11

<210> 42

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> HIV-1 target sequence

<400> 42

aattttgact agcgg 15

<210> 43

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400> 43

atccgctagt caaaatt 17

<210> 44

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> HIV-1 target sequence

<400> 44

gcggaggcta gaagg 15

<210> 45

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400> 45

atccttctag cctccgc 17

<210> 46
 <211> 15
 <212> DNA
 <213> Artificial Sequence
 <220><223> HIV-1 target sequence
 <400> 46
 agagagatgg gtgcg 15

<210> 47
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220><223> U1 snRNA target-annealing sequence
 <400> 47
 atgcaccca tctctct 17

<210> 48
 <211> 15
 <212> DNA
 <213> Artificial Sequence
 <220><223> HIV-1 target sequence
 <400> 48
 agagcgtcgg tatta 15

<210> 49
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220><223> U1 snRNA target-annealing sequence
 <400>
 > 49
 attaatactg acgtctt 17

<210> 50
 <211> 15
 <212> DNA
 <213> Artificial Sequence
 <220><223> HIV-1 target sequence

<400> 50
agcgggggag aatta 15

<210> 51
<211> 17
<212> DNA
<213> Artificial Sequence
<220><223> U1 snRNA target-annealing sequence
<400> 51
attaattctc ccccgct 17

<210> 52
<211> 15
<212> DNA

<213> Artificial Sequence
<220><223> HIV-1 target sequence
<400> 52
gatcgcatg ggaaa 15

<210> 53
<211> 17
<212> DNA
<213> Artificial Sequence
<220><223> U1 snRNA target-annealing sequence
<400> 53
attttcccat cgcgatc 17

<210> 54
<211> 15
<212> DNA
<213> Artificial Sequence
<220><223> HIV-1 target sequence
<400> 54
aaattcggtt aaggc 15

<210> 55
<211> 17
<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400> 55

atgccttaac cgaattt 17

<210> 56

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> HIV-1 target sequence

<400> 56

gatcttcaga cctgg 15

<210> 57

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400>

> 57

atccaggtct gaagatc 17

<210> 58

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> HIV-1 target sequence

<400> 58

ttacacaagc ttaat 15

<210> 59

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400> 59

atattaagct tgtgtaa 17

<210> 60

<211> 15

<212> DNA	
<213> Artificial Sequence	
<220><223> HIV-1 target sequence	
<400> 60	
tagtagacat aatag	15
<210> 61	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 61	
atctattatg tctacta	17
<210> 62	
<211> 9	
<212> DNA	
<213> Artificial Sequence	
<220><223> Control U1 snRNA target sequence	
<400> 62	
ctacaggaa	9
<210> 63	
<211> 11	
<212> DNA	
<213> Artificial Sequence	
<220><223> U1 snRNA target-annealing sequence	
<400> 63	
atttcctgta g	11
<210> 64	
<211> 9	
<212> DNA	
<213> Artificial Sequence	
<220><223> Control U1 snRNA target sequence	
<400> 64	
tcacatgtg	9

<210> 65

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400> 65

atcacagatg a 11

<210> 66

<211> 153

<212> DNA

<213> Artificial Sequence

<220><223> modified U1 snRNA sequence (main U1 snRNA sequence [clover leaf]
(nt 410-562) of the U1_256 sequence)

<400> 66

gcaggggaga taccatgac acgaaggtgg ttttcccagg gcgaggctta tccattgcac 60

tccggatgtg ctgaccctg cgatttccc aaatgtggga aactcgactg cataatttgt 120

ggtagtgggg gactgcgttc gcgtttccc ctg 153

<210> 67

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> HIV-1 target sequence

<400> 67

agctctctcg acgca 15

<210> 68

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> U1 snRNA target-annealing sequence

<400> 68

attgcgtcga gagagct 17

<210> 69

<211> 7

<212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-Kozak junction variant
 <400> 69

gagaatg 7

<210> 70
 <211> 7
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-Kozak junction variant
 <400> 70

gagcatg 7

<210> 71
 <211> 7
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-Kozak junction variant
 <400> 71

gaggatg 7

<210> 72
 <211> 7
 <212> DNA
 <213>

> Artificial Sequence
 <220><223> 3' tbs-Kozak junction variant
 <400> 72

tagaatg 7

<210> 73
 <211> 7
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-Kozak junction variant
 <400> 73

tagcatg 7

<210> 74

<211> 7

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 74

taggatg

7

<210> 75

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 75

gagaaatg

8

<210> 76

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 76

gagacatg

8

<210> 77

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 77

gagagatg

8

<210> 78

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 78
gagcaatg 8

<210> 79
<211> 8
<212> DNA
<213> Artificial Sequence
<220><223> 3' tbs-Kozak junction variant
<400> 79
gagccatg 8
<210> 80
<211> 8
<212> DNA
<213>
> Artificial Sequence
<220><223> 3' tbs-Kozak junction variant
<400> 80
gagcgatg 8
<210> 81
<211> 8
<212> DNA
<213> Artificial Sequence
<220><223> 3' tbs-Kozak junction variant
<400> 81
gaggaatg 8
<210> 82
<211> 8
<212> DNA
<213> Artificial Sequence
<220><223> 3' tbs-Kozak junction variant
<400> 82
gaggcatg 8

<210> 83
<211> 8
<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 83

gagggatg 8

<210> 84

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 84

tagaaatg 8

<210> 85

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 85

tagacatg 8

<210> 86

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 86

tagagatg 8

<210> 87

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 87

tagcaatg 8

<210> 88

<211> 8

<212> DNA

<213>

> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 88

tagccatg 8

<210> 89

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 89

tagcgatg 8

<210> 90

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 90

taggaatg 8

<210> 91

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 91

taggcatg 8

<210> 92

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 92

tagggatg 8

<210> 93

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> L33 improved leader derived from a splicing event

<400> 93

ctttttcgca acgggtttgc cgccagaaca caggtgtcgt gaaaa

45

<210> 94

<211> 75

<212> PRT

<213> Bacillus subtilis

<400> 94

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Thr Ala Tyr Gly Glu Met Lys Ser Glu Lys Lys

65 70 75

<210> 95

<211> 86

<212> PRT

<213> Aminomonas paucivorans

<400> 95

Met Lys Glu Gly Glu Glu Ala Lys Thr Ser Val Leu Ser Asp Tyr Val

1 5 10 15

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

20 25 30

Gly Gln Glu Thr Lys Phe Ala His Thr Glu Lys Leu Asp Asp Gly Glu

35 40 45
Val Trp Ile Ala Gln Phe Thr Glu His Thr Ser Ala Ile Lys Val Arg
50 55 60
Gly Ala Ser Glu Ile His Thr Lys His Gly Met Leu Phe Ser Gly Arg
65 70 75 80
Gly Arg Asn Glu Lys Gly

85

<210> 96

<211> 81

<212> PRT

<213> Desulfotomaculum hydrothermale

<400> 96

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

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

<210> 97

<211> 74

<212> PRT

<213> Bacillus stearothermophilus

<400> 97

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

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

35 40 45

His Thr Ser Ala Ile Lys Val Arg Gly Lys Ala Tyr Ile Gln Thr Arg

50 55 60

His Gly Val Ile Glu Ser Glu Gly Lys Lys

65 70

<210> 98

<211> 74

<212> PRT

<213> Bacillus stearothermophilus

<400> 98

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

1 5 10 15

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

20 25 30

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

35 40 45

His Thr Ser Ala Ile Lys Val Arg Gly Lys Ala Tyr Ile Gln Thr Arg

50 55 60

His Gly Val Ile Glu Asn Glu Gly Lys Lys

65 70

<210> 99

<211> 76

<212> PRT

<213> Bacillus halodurans

<400> 99

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

1 5 10 15

Asn Gly Val Asn Val Phe Gly Met Thr Arg Gly Thr Asp Thr Arg Phe

20 25 30

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

35 40 45

Thr Glu His Thr Ser Ala Val Lys Ile Arg Gly Lys Ala Ile Ile Gln
50 55 60

Thr Ser Tyr Gly Thr Leu Asp Thr Glu Lys Asp Glu

65 70 75

<210> 100

<211> 80

<212> PRT

<213> Carboxydotherrmus hydrogenoformans

<400> 100

Met Val Cys Asp Asn Phe Ala Phe Ser Ser Ala Ile Asn Ala Glu Tyr

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Arg Gly Lys Ala Glu Ile Tyr Thr Lys His Gly Val Ile Lys Asn Glu

65 70 75 80

<210> 101

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> L33 improved leader

<400> 101

ctttttcgca acgggtttgc cgccagaaca cag

33

<210> 102

<211> 12

<212> DNA

<213> Artificial Sequence

<220><223> L12 improved leader

<400> 102

ctttttcgca ac	12
<210> 103	
<211> 9	
<212> DNA	
<213> Artificial Sequence	
<220><223> extended Kozak consensus sequence	
<220><221> misc_feature	
<222> (2)..(3)	
<223> n is a, c, g, or t	
<400> 103	
gnnrvvatg	9
<210> 104	
<211> 6	
<212> DNA	
<213> Artificial Sequence	
<220><223> core Kozak consensus sequence	
<400>	
> 104	
rvvatg	6
<210> 105	
<211> 32	
<212> DNA	
<213> Artificial Sequence	
<220><223> EMCV M loops	
<400> 105	
cgtggttttc ctttgaaaaa cacgatgata cc	32
<210> 106	
<211> 56	
<212> DNA	
<213> Artificial Sequence	
<220><223> optimal (overlapping) tbs-Kozak	
<400> 106	
gagtttagcg gaggtagagaa gagcggagcc gaggcctagca gagacgagcc gagatg	56
<210> 107	
<211> 65	

<212> DNA

<213>

> Artificial Sequence

<220><223> optimal (overlapping) tbs-MCS-Kozak

<400> 107

gagtttagcg gaggaggag gagcggagcc gaggctagca gagacgagaa gagctctaga 60

ccatg 65

<210> 108

<211> 6

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 108

tagatg 6

<210> 109

<211> 8

<212> DNA

<213> Artificial Sequence

<220>

<223> 3' tbs-Kozak junction variant

<400> 109

gagatatg 8

<210> 110

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 110

gagctatg 8

<210> 111

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 111

tagatatg 8

<210> 112

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 112

tagctatg 8

<210> 113

<211> 6

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak consensus sequence

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<400> 113

kagnng 6

<210> 114

<211> 6

<212> DNA

<

213> Artificial Sequence

<220><223> 3' tbs-Kozak consensus sequence

<400> 114

kagatg 6

<210> 115

<211> 7

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak consensus sequence

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<400> 115
kagnntg 7

<210> 116
<211> 7
<212> DNA
<213> Artificial Sequence
<220><223>
> 3' tbs-Kozak consensus sequence
<220><221> misc_feature
<222> (4)..(4)
<223> n is a, c, g, or t
<400> 116
kagnatg 7

<210> 117
<211> 1082
<212> DNA
<213> Artificial Sequence
<220><223> Chicken beta-Actin / Rabbit beta-globin chimeric 5'UTR-intron
with tbs-kzkV0.G variant
<400> 117

cggcggggcg gaacgttgcc ttgccccgt gccccgtcc gcgccgcctc gcgccgcccg 60
ccccggctct gactgaccgc gttactccca caggtgagcg ggcgggacgg cccttctccc 120

tccgggctgt aattagcgt tggtttaatg acggctcgtt tcttttctgt ggctgcgtga 180
aagccttaaa gggtctccgg agggcctttg tgcggggggg agcggctcgg ggggtgcgtg 240
cgtgtgtgtg tgcgtgggga gcgccgcgtg cggccccgcgc tgcccgccgg ctgtgagcgc 300
tgcgggcgcg gcgcggggct ttgtgcgtc cgcgtgtgcg cgaggggagc gcgggcccgg 360
ggcggtgccc cgcggtgcgg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt 420
gcgtgggggg gtgagcaggg ggtgtgggcg cggcggtcgg gctgtaacct cccctggca 480
ccccctccc cgagttgctg agcacggccc ggcttcgggt gcggggctcc gtgcggggcg 540

tggcgcgggg ctgccgtgc cgggcggggg gtggcggcag gtgggggtgc cgggcggggc 600
ggggccgcct cgggcggggg agggctcggg ggaggggcgc ggcggccccg gagcgcggc 660
ggctgtcgag gcgcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagaggcg 720
cagggacttc ctttgtccca aatctggcgg agccgaaatc tgggaggcgc cgccgcaccc 780

cctctagcgg gcgcgggcga agcgggtgcgg cgccggcagg aaggaaatgg gcggggaggg 840

ccttcgtgcg tcgccgcgcc gccgtcccct tctccatctc cagcctcggg gctgccgcag 900

ggggacggct gccttcgggg gggacggggc agggcggggt tcggcttctg gcgtgtgacc 960

ggcggcttta gagcctctgc taacatgtt catgccttct tcttttctt acagctcctg 1020

ggcaaagagt ttagcggagt ggagaagagc ggagccgagc ctagcagaga cgagccgaga 1080

tg 1082

<210> 118

<211> 1040

<212> DNA

<213> Artificial Sequence

<220><223> EF1a 5'UTR-intron with tbskzkV0.G variant

<400> 118

ctttttcgca acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg 60

cctggcctct ttacgggtta tggcccttgc gtgccttgaa ttacttccac ctggctgcag 120

tacgtgattc ttgatccga gcttcgggtt ggaagtgggt gggagagtgc gtggccttgc 180

gcttaaggag ccccttcgcc tcgtgcttga gttgtggcct ggccctggcg ctggggccgc 240

cgctgcgaa tctgggtgca ccttcgcgcc tgtctcgtg ctttcgataa gtctctagcc 300

atttaaaatt ttgatgacc tgctgcgacg ctttttttct ggcaagatag tcttgtaaat 360

gcgggccaag atcagcacac tggatatttcg gtttttgggg ccgcgggcgg cgacggggcc 420

cgtgcgtccc agcgcatatg ttccgagagg cggggcctgc gagcgcgcc accgagaatc 480

ggacgggggt agtctcaagc tgccggcct gctctggtgc ctggcctgc gccgccgtgt 540

atgccccgc cctgggcggc aaggttgcc cggtcgccac cagttgcgtg agcggaaaga 600

tggccgcttc ccggccctgc tgcaggagc acaaaatgga ggacgcggcg ctccggagag 660

cgggcgggtg agtcacccac acaaaggaaa agggccttcc cgtcctcagc cgtcgttca 720

tgtgactcca cggagtaccg ggcccgctcc aggcacctcg attagtctc cagcttttgg 780

agtacgtcgt ctttaggttg gggggagggg ttttatgcga tggagtttcc ccacactgag 840

tgggtggaga ctgaagttag gccagcttgg cacttgatgt aattctcctt ggaatttgcc 900

ctttttgagt ttgatcttg gttcattctc aagcctcaga cagtggttca aagttttttt 960

cttcatttc aggtgtcgtg aaaagagttt agcggagtgg agaagagcgg agccgagcct 1020

agcagagacg agccgagatg 1040

<210> 119

<211> 161

<212> DNA

<213> Artificial Sequence

<220><223> spliced sequence corresponding to SEQ ID NO: 117

<400> 119

cggcgggagg gaaagtggc ttgccccgt gccccgtcc gcgcgcctc gcgcgcccc	60
ccccggtct gactgaccg gttactcca cagctcctgg gaaaagagtt tagcggagt	120
gagaagagcg gagccgagcc tagcagagac gagccgagat g	161

<210> 120

<211> 101

<212> DNA

<213> Artificial Sequence

<220><223> spliced sequence corresponding to SEQ ID NO: 118

<400> 120

ctttttcgca acgggtttgc cgccagaaca caggtgtcgt gaaaagagtt tagcggagt	60
gagaagagcg gagccgagcc tagcagagac gagccgagat g	101

<210> 121

<211> 1026

<212> DNA

<213> Artificial Sequence

<220><223> Chicken beta-Actin / Rabbit beta-globin chimeric 5'UTR-intron

<400> 121

cggcgggagg gaaagtggc ttgccccgt gccccgtcc gcgcgcctc gcgcgcccc	60
ccccggtct gactgaccg gttactcca caggtgagcg ggccggacgg cctttctcc	120
tccgggtgt aattagcgt tggtttaatg acggctcgtt tcttttctgt ggctgcgtga	180
aagccttaaa gggctccggg agggcctttg tgcggggggg agcggctcgg ggggtgcgtg	240
cgtgtgtgtg tgcgtgggga gcgcgcgtg cggccccgcg tgcggggcgg ctgtgagcg	300
tgcgggcgcg gcgcggggct ttgtgcgtc cgcgtgtgcg cgaggggagc gcgggcccgg	360
ggcgggtgcc cgcgggtcgg gggggctgcg aggggaacaa aggctgcgtg cgggggtgtg	420
gcgtgggggg gtgagcaggg ggtgtgggcg cggcggctcg gctgtaacct cccctggca	480
ccccctccc cgagttgctg agcacggccc ggcttcgggt gcggggctcc gtgcggggcg	540
tggcgcgggg ctgcccgtgc cgggcggggg gtggcggcag gtgggggtgc cgggcggggc	600

gggggccgcct cggggccgggg agggctcggg ggagggggcgc ggcggccccc gagcgccggc 660
 ggctgtcag ggcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagaggcg 720
 cagggacttc ctttgtccca aatctggcgg agccgaaatc tgggaggcgc cgcgcaccc 780
 cctctagcgg gcgcgggcga agcgggtgcgg cgcgggcagg aaggaaatgg gcggggaggg 840
 cttctgtgcg tcgcccgcgc gccgtcccct tctccatctc cagcctcggg gctgccgcag 900

 ggggacggct gccttcgggg gggacggggc agggcgggggt tcggcttctg gcgtgtgacc 960
 ggcggttta gagcctctgc taaccatgtt catgccttct tcttttct acagctcctg 1020
 ggcaaa 1026
 <210> 122
 <211> 984
 <212> DNA
 <213> Artificial Sequence
 <220><223> EF1a 5'UTR-intron
 <400> 122

 ctttttcga acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg tcccccggg 60
 cctggcctct ttacgggtta tggcccttgc gtgccttgaa ttacttcac ctggctgcag 120
 tacgtgattc ttgatccga gcttcgggtt ggaagtgggt gggagagtgc gtggccttgc 180

 gcttaaggag ccccttcgcc tcgtgcttga gttgtggcct ggcctgggcg ctggggccgc 240
 cgctgcgaa tctggtggca ctttcgcgc tgtctcgtg ctttcgataa gtctctagcc 300
 atttaaaatt ttgatgacc tgctgcgacg cttttttct ggcaagatag tcttgtaaat 360
 gcgggccaa atcagcacac tggatatttc gtttttgggg ccgcgggcgg cgacggggcc 420
 cgtgcgtccc agcgcacatg ttcggcgagg cggggcctgc gagcgcgcc accgagaatc 480
 ggacgggggt agtctcaagc tgcccggcct gctctggtgc ctggcctcgc gccgccgtgt 540
 atcgccccgc cctgggcggc aaggctggcc cggtcggcac cagttgcgtg agcggaaaga 600

 tggccgcttc cgggccttgc tgcagggagc aaaaaatgga ggacgcggcg ctcgggagag 660
 cgggcgggtg agtcacccac acaaaggaaa agggcctttc cgtcctcagc cgtcgttca 720
 tgtactcca cggagtaccg ggcgcgtcc aggcacctcg attagtctc cagcttttgg 780
 agtacgtcgt ctttaggttg gggggagggg ttttatgcga tggagtttcc ccacactgag 840
 tgggtggaga ctgaagttag gccagcttgg cacttgatgt aattctcctt ggaatttgcc 900
 ctttttgagt ttggatcttg gttcattctc aagcctcaga cagtggttca aagttttttt 960
 cttccatttc aggtgtcgtg aaaa 984

<210> 123

<211> 1030

<212> DNA

<213> Artificial Sequence

<220><223> Chicken beta-Actin / Rabbit beta-globin chimeric 5'UTR-intron-tbs
consensus

<220><221> misc_feature

<222> (1027)..(1030)

<223> kagn(2-3) may be repeated 10-11 times

<220><221> misc_feature

<222> (1030)..(1030)

<223> n may be repeated 2-3 times

<400> 123

cggcgggaggc gaacgttgcc ttcccccgt gccccgtcc gcgccgcctc gcgccggccg 60

ccccggctct gactgaccgc gttactccca caggtgagcg ggccgggacgg cccttctccc 120

tccgggctgt aattagcgt tggtttaatg acggctcgtt tcttttctgt ggctgcgtga 180

aagccttaaa gggtccggg agggcctttg tgcggggggg agcggtcgg gggtgcgtg 240

cgtgtgtgtg tgcgtgggga gcgccgcgtg cggcccgccg tgcccgccgg ctgtgagcgc 300

tgcgggcgcg gcgcggggct ttgtgcgtc cgctgtgctg cgaggggagc gcggggccggg 360

ggcggtgccc cgcggtcggg gggggctgcg aggggaacaa aggttgcgtg cgggggtgtgt 420

gcgtgggggg gtgagcaggg ggtgtgggcg cggcggtcgg gctgtaacct cccctggca 480

ccccctccc cgagttgctg agcacggccc ggcttcgggt gcggggctcc gtgcggggcg 540

tggcgcgggg ctgccgtgc cgggcggggg gtggcggcag gtgggggtgc cgggcggggc 600

ggggccgcct cgggcggggg agggctcggg ggaggggcgc ggccggcccc gagcgccggc 660

ggctgtcgag gcgcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagaggcg 720

cagggaactt ctttgtccca aatctggcgg agccgaaatc tgggaggcgc cgccgcaccc 780

cctctagcgg gcgcgggcga agcggtgcgg cgcgggcagg aaggaaatgg gcggggaggg 840

cttctgtgct tgcggcgccc gccgtccct tctccatctc cagcctcggg gctgccgcag 900

ggggacggct gccttcgggg gggacggggc agggcggggt tcggcttctg gcgtgtgacc 960

ggcggtttta gagcctctgc taacctgtt catgccttct tctttttcct acagtcctg 1020

ggcaaaksgn 1030

<210> 124

<211> 988

<212> DNA

<213> Artificial Sequence

<220><223> EF1a 5'UTR-intron-tbs consensus

<220><221> misc_feature

<222> (985)..(988)

<223> kagn(2-3) may be repeated 10-11 times

<220><221> misc_feature

<222> (988)..(988)

<223> n may be repeated 2-3 times

<400> 124

ctttttcgca acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg ttccccgagg 60

cctggcctct ttacgggtta tggcccttgc gtgccttgaa ttacttcac ctggctgcag 120

tacgtgattc ttgatccga gcttcgggtt ggaagtgggt gggagagttc gtggccttgc 180

gcttaaggag ccccttcgcc tcgtgcttga gttgtggcct ggcctgggag ctggggccgc 240

cgcgtgcgaa tctggtggca ccttcgcgcc tgcctcctg ctttcgataa gtctctagcc 300

atttaaaatt ttgatgacc tgctcgcagc ctttttttct ggcaagatag tcttgtaaat 360

gcgggccaag atcagcacac tggatatttcg gtttttgggg ccgcgggagg cgacggggcc 420

cgtgcgtccc agcgcacatg ttcggcgagg cggggcctgc gagcgcggcc accgagaatc 480

ggacgggggt agtctcaagc tgcccggcct gctctggtgc ctggcctcgc gccgccgtgt 540

atcgccccgc cctgggaggc aaggctggcc cggtcggcac cagtgcgtg agcggaaaga 600

tggccgcttc ccggccctgc tgcagggagc acaaaatgga ggacgcggcg ctggggagag 660

cgggcgggtg agtcacccac acaaaggaaa agggcctttc cgtcctcagc cgtcgttca 720

tgtgactcca cggagtaccg ggcccgctcc aggcacctcg attagttctc cagcttttgg 780

agtacgtcgt ctttaggttg gggggagggg ttttatgcga tggagtttcc ccacactgag 840

tgggtggaga ctgaagttag gccagcttgg cacttgatgt aattctcctt ggaatttgcc 900

ctttttgagt ttggatcttg gttcattctc aagcctcaga cagtggttca aagttttttt 960

cttccatttc aggtgtcgtg aaaakagn 988

<210> 125

<211> 55

<212> RNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]11

<400> 125

gaguuuagcg gaguggagaa gagcggagcc gagccuagca gagacgagug gagcu 55

<210> 126

<211> 55

<212> RNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]11

<400> 126

gaguuuagcg gaguggagaa gagcggagcc gagccuagca gagacgagaa gagcu 55

<210> 127

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]6

<400> 127

uaguuuaguu uaguuuaguu uaguuuaguu 30

<210> 128

<211> 30

<212> RNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]6

<400> 128

uaguuuaguu gaguuuaguu gaguuuaguu 30

<210> 129

<211> 32

<212> RNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]6

<400> 129

gaguuuagagu ugaguugagu uugaguugag uu 32

<210> 130

<211> 32

<212> RNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]6

<400> 130

uaguuugagu uuaguugagu uuuaguugag uu 32

<210> 131

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]11

<400> 131

gagtttagcg gagtggagaa gagcggagcc ggcctagca gagacgagtg gagct 55

<210> 132

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]10

<400> 132

gagtttagcg gagtggagaa gagcggagcc ggcctagca gagacgagaa 50

<210> 133

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]9

<400> 133

gagtttagcg gagtggagaa gagcggagcc ggcctagca gagac 45

<210> 134

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> TRAP binding site variant - [KAGNN]8

<400> 134

gagtttagcg gagtggagaa gagcggagcc ggcctagca 40

<210> 135
 <211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><223> TRAP binding site variant - [KAGNN]7
 <400> 135
 gagtttagcg gaggagagaa gagcggagcc gagcc 35
 <210> 136
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> TRAP binding site variant - [KAGNN]6
 <400> 136
 gagtttagcg gaggagagaa gagcggagcc 30
 <210> 137
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> TRAP binding site variant - [KAGNNN]3[KAGNN]8
 <400> 137
 gagtttagcg gaggagagaa agagacggag ccgagaccta gcagagacga gaagagct 58
 <210> 138
 <211> 56
 <212> DNA
 <213> Artificial Sequence
 <220><223> TRAP binding site variant - [KAGNNN]1[KAGNN]10
 <400> 138
 gagtttagcg gaggagagaa gagacggagc cgagcctagc agagacgaga agagct 56
 <210> 139
 <211> 55
 <212> DNA
 <213> Artificial Sequence
 <220><223> TRAP binding site variant - [KAGNN]11

<400> 139
gagtttagcg gaggggagaa gagcggagcc gaggcctagca gagacgagaa gagct 55
<210> 140
<211> 41
<212> DNA
<213> Artificial Sequence
<220><223> TRAP binding site variant - [KAGNNN]1[KAGNN]7
<400> 140
gagtttagcg gaggggagaa agagcggagc cgagcctagc a 41
<210> 141
<211> 36
<212> DNA
<213> Artificial Sequence
<220><223> TRAP binding site variant - [KAGNNN]1[KAGNN]6
<400> 141
gagtttagcg gaggggagaa agagcggagc cgagcc 36
<210> 142
<211> 6
<212> DNA
<213> Artificial Sequence
<220><223> 3' tbs-Kozak junction variant
<400> 142
gagatg 6
<210> 143
<211> 7
<212> DNA
<213> Artificial Sequence
<220><223> 3' tbs-Kozak junction variant
<400> 143
kagvatg 7
<210> 144
<211> 8
<212> DNA
<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400>

> 144

kagvvatg 8

<210> 145

<211> 9

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<400> 145

kagrsvatg 9

<210> 146

<211> 10

<212> DNA

<213> Artificial Sequence

<220><223> 3' tbs-Kozak junction variant

<220><221> misc_feature

<222> (4)..(4)

<223> n is a, c, g, or t

<400> 146

kagnrvvatg 10

<210> 147

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> optimal 3' tbs-Kozak junction variant

<400> 147

kagccgagat g 11

<210> 148

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> optimal 3' tbs-Kozak junction variant

<220><221> misc_feature

<222> (4)..(4)
 <223> n is a, c, g, or t
 <400> 148
 kagnggagcc atg 13
 <210> 149
 <211> 14

 <212> DNA
 <213> Artificial Sequence
 <220><223> optimal 3' tbs-Kozak junction variant
 <220><221> misc_feature
 <222> (4)..(5)
 <223> n is a, c, g, or t
 <400> 149
 kagnngagac catg 14
 <210> 150
 <211> 12
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-Kozak junction variant
 <400> 150
 kaggcgagca tg 12
 <210> 151
 <211> 15
 <212> DNA
 <213> Artificial Sequence

 <220><223> spacer variant
 <400> 151
 atagcagaga cggt 15
 <210> 152
 <211> 10
 <212> DNA
 <213> Artificial Sequence
 <220><223> spacer variant

<400> 152	
atagcagaga	10
<210> 153	
<211> 5	
<212> DNA	
<213> Artificial Sequence	
<220><223> spacer variant	
<400> 153	
atagc	5
<210> 154	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> spacer variant	
<400> 154	
atatcagaga cggctagcgt atacca	26
<210> 155	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> spacer variant	
<400> 155	
atatcagaga cggct	15
<210> 156	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220><223> spacer variant	
<400> 156	
agagacggct	10
<210> 157	
<211> 5	
<212> DNA	

<213> Artificial Sequence	
<220><223> spacer variant	
<400> 157	
tacca	5
<210> 158	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 158	
gagctctaga vvatg	15
<210> 159	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 159	
gagctcgtcg acvatg	16
<210> 160	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 160	
gagctcgaat tcgaavvatg	20
<210> 161	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 161	
gagctctaga cgtcgacvat g	21
<210> 162	
<211> 22	

<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 162	
gagctctaga attcgaavva tg	22
<210> 163	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 163	
gagctctaga tatcgatrvv atg	23
<210> 164	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 164	
kagactagta cttaaagcttr vvatg	25
<210> 165	
<211> 15	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 165	
gagctctaga ccatg	15
<210> 166	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> 3' tbs-MCS-Kozak variant	
<400> 166	
gagctcgteg accatg	16

<210> 167
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-MCS-Kozak variant
 <400> 167
 gagctcgaat tcgaacctg 20
 <210> 168
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-MCS-Kozak variant
 <400> 168
 gagctctaga cgtcgacct g 21

 <210> 169
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-MCS-Kozak variant
 <400> 169
 gagctctaga attcgaacca tg 22
 <210> 170
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-MCS-Kozak variant
 <400> 170
 gagctctaga tatcgatacc atg 23
 <210> 171
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> 3' tbs-MCS-Kozak variant
 <400> 171

kagactagta cttaaagctta ccatg

25

<210> 172

<211> 74

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 1 containing L33 Improved

leader, optimal (overlapping) tbs ([KAGNN]8)-Kozak junction

<400> 172

ctttttcgca acgggtttgc cgccagaaca caggagtta gcggagtga gaagagcgga 60

gccgagccga gatg 74

<210> 173

<211> 89

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 2 containing L33 Improved

leader, optimal (overlapping) tbs ([KAGNN]11)-Kozak junction

<400> 173

ctttttcgca acgggtttgc cgccagaaca caggagtta gcggagtga gaagagcgga 60

gccgagccta gcagagacga gccgagatg 89

<210> 174

<211> 68

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 3 containing L12 Improved

leader, optimal (overlapping) tbs ([KAGNN]11)-Kozak junction

<400> 174

ctttttcgca acgagtttag cggagtggag aagagcggag ccgagcctag cagagacgag 60

ccgagatg 68

<210> 175

<211> 101

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 4 for intron-containing
5'UTRs, resulting in a spliced leader comprising L33, optimal
(overlapping) tbs ([KAGNN]11)-Kozak junction

<400> 175

ctttttcgca acgggtttgc cgccagaaca caggtgtcgt gaaaagagtt tagcggagtg 60
gagaagagcg gagccgagcc tagcagagac gagccgagat g 101

<210>

> 176

<211> 56

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 5 containing improved spacer,
optimal (overlapping) tbs ([KAGNN]8)-Kozak junction

<400> 176

atagcagaga cggtgagtt tagcggagtg gagaagagcg gagccgagcc gagatg 56

<210> 177

<211> 98

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 6 containing L33 Improved
leader tbs ([KAGNN]11)-MCS-Kozak

<400> 177

ctttttcgca acgggtttgc cgccagaaca caggagtta gcgagtgga gaagagcgga 60

gccgagccta gcagagacga gaagagctct agaccatg 98

<210> 178

<211> 80

<212> DNA

<213> Artificial Sequence

<220><223> illustrative nucleic acid sequence 7 containing improved spacer,
tbs ([KAGNN]11)-MCS-Kozak

<400> 178

atagcagaga cggtgagtt tagcggagtg gagaagagcg gagccgagcc tagcagagac 60

gagaagagct ctagaccatg 80

<210> 179

<211> 12	
<212> DNA	
<213> Artificial Sequence	
<220><223> tbs-Kozak region	
<400> 179	
acagccacca tg	12
<210> 180	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> HpaI variant tbs-Kozak region	
<400> 180	
gagttaacgc caccatg	17
<210> 181	
<211> 13	
<212> DNA	
<213> Artificial Sequence	
<220><223> consensus Kozak sequence (DNA)	
<400> 181	
gccgccrcca tgg	13
<210> 182	
<211> 13	
<212> RNA	
<213> Artificial Sequence	
<220><223> consensus Kozak sequence (RNA)	
<400> 182	
gccgccrcca ugg	13
<210> 183	
<211> 10	
<212> RNA	
<213> Artificial Sequence	
<220><223> extended Kozak consensus sequence	
<220><221> misc_feature	

<222> (2)..(3)

<223> n is a, c, g, or u

<400> 183

gnnrvvvaugg

10

<210> 184

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> original 3' tbs-Kozak junction sequence

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<400> 184

kagnnacagc caccatg

17

<210> 185

<211> 67

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222>

> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature

 <222> (49)..(50)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (54)..(55)
 <223> n is a, c, g, or t
 <400> 185
 kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnacagc 60
 caccatg 67
 <210> 186
 <211> 56
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222> (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222>
 > (9)..(10)
 <223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 186

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gagatg

56

<210> 187

<211> 56

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 187

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn tagatg

56

<210> 188

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 188

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn gagcatg

57

<210> 189
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222> (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature

 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)

<223> n is a, c, g, or t

<400> 189

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn gagaatg 57

<210> 190

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223>

> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)
 <223> n is a, c, g, or t
 <400> 190
 kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gaggatg 57
 <210> 191
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222>
 > (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 191

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn tagaatg

57

<210> 192

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 192

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn tagcatg

57

<210> 193

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 193

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn taggatg

57

<210> 194

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221>

> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 194

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gagccatg

58

<210> 195

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 195

kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn gagaaatg

58

<210> 196

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223>

> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 196

kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn gagagatg

58

<210> 197

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222>

> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 197

kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn gagacatg

58

<210> 198
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222> (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)

 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)

 <223> n is a, c, g, or t
 <220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 198

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn gagatatg

58

<210> 199

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

```

<222> (44)..(45)
<223> n is a, c, g, or t
<220><221> misc_feature
<222> (49)..(50)
<223> n is a, c, g, or t

<400> 199
kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gagcaatg      58
<210> 200
<211> 58
<212> DNA
<213> Artificial Sequence
<220><223> overlapping tbs 3'KAGNN-Kozak variant
<220><221> misc_feature
<222> (4)..(5)
<223> n is a, c, g, or t
<220><221> misc_feature
<222> (9)..(10)
<223> n is a, c, g, or t
<220><221> misc_feature
<222> (14)..(15)
<223> n is a, c, g, or t
<220><221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t
<220><221>
> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t
<220><221> misc_feature
<222> (29)..(30)
<223> n is a, c, g, or t
<220><221> misc_feature
<222> (34)..(35)
<223> n is a, c, g, or t

```

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 200

kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn gagcgatg

58

<210> 201

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 201

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gagctatg

58

<210> 202

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223>

> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature

<222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)
 <223> n is a, c, g, or t
 <400> 202

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gaggaatg 58

<210> 203
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222>
 > (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 203

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gagggatg

58

<210> 204

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 204

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn gaggcag

58

<210> 205

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature
 <222> (9)..(10)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t

<220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)
 <223> n is a, c, g, or t

<400> 205
 kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn tagaaatg
 <210> 206
 <211> 58
 <212> DNA
 <213> Artificial Sequence

58

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221>

> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 206

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn tagagatg

58

<210> 207
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222> (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature

 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)

<223> n is a, c, g, or t

<400> 207

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn tagacatg 58

<210> 208

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223>

> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)
 <223> n is a, c, g, or t
 <400> 208
 kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn tagatatg 58
 <210> 209
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><223> overlapping tbs 3'KAGNN-Kozak variant
 <220><221> misc_feature
 <222>
 > (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 209

kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn tagcaatg 58

<210> 210

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 210

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn tagcgatg

58

<210> 211

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 211

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn tagccatg 58

<210> 212

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221

> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 212

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn tagctatg

58

<210> 213

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 213

kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn kaggnkaggn taggaatg

58

<210> 214

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223>

> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 214

kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn tagggatg

58

<210> 215

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> overlapping tbs 3'KAGNN-Kozak variant

<220><221> misc_feature

<222>

> (4)..(5)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (9)..(10)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (14)..(15)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (24)..(25)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (29)..(30)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (34)..(35)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (39)..(40)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (44)..(45)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (49)..(50)

<223> n is a, c, g, or t

<400> 215

kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn kaggnkagnn taggcatg

58

<210> 216
 <211> 164
 <212> RNA
 <213> Artificial Sequence
 <220><223> U1 snRNA molecule
 <400> 216
 auacuuaccu ggcaggggag auaccaugau cacgaaggug guuuucccag ggcgaggcuu 60
 auccaugca cuccggaugu gcugaccccu gcgauuuccc caaauugugg aaacucuacu 120

 gcauaauuug ugguaguggg ggacugcguu cgcgcuuucc ccug 164
 <210> 217
 <211> 13
 <212> DNA
 <213> Artificial Sequence
 <220><223> mutated splice donor (MSD) region, deltaSL2
 <400> 217
 ggggcgcaaa aat 13
 <210> 218
 <211> 55
 <212> DNA
 <213> Artificial Sequence
 <220><223> tbs consensus sequence
 <220><221> misc_feature
 <222> (4)..(5)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (9)..(10)

 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (14)..(15)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (19)..(20)
 <223> n is a, c, g, or t

<220><221> misc_feature
 <222> (24)..(25)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (29)..(30)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (34)..(35)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (39)..(40)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (44)..(45)

 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (49)..(50)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (54)..(55)
 <223> n is a, c, g, or t
 <400> 218
 kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnnkagnn kagnn 55
 <210> 219
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> wild type SL2-MSD-SL4(STD) sequence including SL2 loop region
 <400> 219
 ggcggcgcact ggtgagtacg ccaaaaat 28
 <210> 220

 <211> 22
 <212> DNA

<213> Artificial Sequence

<220><223> wild type SL2-MSD-SL4(STD) sequence including SL4 loop region

<400> 220

gagatgggtg cgagagcgtc ag 22

<210> 221

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> MSD-1KO sequence including SL2 loop region

<400> 221

ggcggcgact gcagagtacg ccaaaaaat 28

<210> 222

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-1KO sequence including SL4 loop region

<400> 222

gagatgggtg cgagagcgtc ag 22

<210> 223

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> MSD-2KO sequence including SL2 loop region

<400> 223

ggcggcgact gcagacaacg ccaaaaaat 28

<210> 224

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-2KO sequence including SL4 loop region

<400> 224

gagatgggtg cgagagcgtc ag 22

<210> 225

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> MSD-3KO_1 sequence including SL2 loop region

<400> 225

ggcggcgact gcagacaacg ccaaaaat 28

<210> 226

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-3KO_1 sequence including SL4 loop region

<400> 226

gagatggctg cgagagcgtc ag 22

<210> 227

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> MSD-3KO_2 sequence including SL2 loop region

<400> 227

ggcggcgact gcagacaacg ccaaaaat 28

<210> 228

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-3KO_2 sequence including SL4 loop region

<400> 228

gagatgggcg cgagagcgtc ag 22

<210> 229

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> MSD-4KO_1 sequence including SL2 loop region

<400> 229

ggcggcgact gcagacaacg ccaaaaat 28

<210> 230

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-4KO_1 sequence including SL4 loop region

<400> 230

gagatggctg cgagagcatc ag 22

<210> 231

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> MSD-4KO_2 sequence including SL2 loop region

<400> 231

ggcggcgact gcagacaacg ccaaaaat 28

<210> 232

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-4KO_2 sequence including SL4 loop region

<400> 232

gagatgggcg cgagagcatc ag 22

<210> 233

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> MSD-2KOm5 sequence including SL2 loop region

<400> 233

ggaaggcaac agataaatat gccttaaaat 30

<210> 234

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-2K0m5 sequence including SL4 loop region

<400> 234

gagatgggtg cgagagcgtc ag

22

<210> 235

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> MSD-3K0m5_2 sequence including SL2 loop region

<400> 235

ggaaggcaac agataaatat gccttaaaat

30

<210> 236

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-3K0m5_2 sequence including SL4 loop region

<400> 236

gagatgggcg cgagagcgtc ag

22

<210> 237

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223>

MSD-4K0m5_2 sequence including SL2 loop region

<400> 237

ggaaggcaac agataaatat gccttaaaat

30

<210> 238

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> MSD-4K0m5_2 sequence including SL4 loop region

<400> 238

gagatgggcg cgagagcatc ag

22

<210> 239

<211> 11

<212> RNA
 <213> Artificial Sequence
 <220><223> endogenous U1 snRNA
 <400> 239
 auacuuaccu g 11

<210> 240
 <211> 19
 <212> RNA
 <213> Artificial Sequence
 <220><223> wild type major splice donor site sequence, wt SL2-MSD(STD)
 <400> 240
 ggcgacuggu gaguacgcc 19

<210> 241
 <211> 19
 <212> RNA
 <213> Artificial Sequence
 <220><223> variant major splice donor site, MSD-2KO
 <400> 241
 ggcgacugca gacaacgcc 19

<210> 242
 <211> 19
 <212> RNA
 <213> Artificial Sequence
 <220>
 ><223> variant major splice donor site, MSD-2K0m5
 <400> 242
 ggcaacagau aaauaugcc 19