(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0156462
(43) 공개일자 2022년11월25일

(51) 국제특허분류(Int. Cl.)
C12N 15/85 (2006.01) *C07K 14/435* (2006.01)
C12N 15/10 (2017.01) *C12N 15/113* (2010.01)
C12N 9/22 (2006.01) *C12Q 1/6897* (2018.01)
G01N 33/50 (2017.01) *G01N 33/68* (2006.01)
(52) CPC특허분류
C12N 15/85 (2013.01)
C07K 14/43595 (2013.01)
(21) 출원번호 10-2022-0060420
(22) 출원일자 2022년05월17일
심사청구일자 없음
(30) 우선권주장
1020210064300 2021년05월18일 대한민국(KR)

(71) 출원인
한국생명공학연구원
대전광역시 유성구 과학로 125 (어은동)
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
(72) 발명자
김선욱
대전광역시 유성구 과학로 125(어은동)
박영호
대전광역시 유성구 과학로 125(어은동)
(뒷면에 계속)
(74) 대리인
제일특허법인(유)

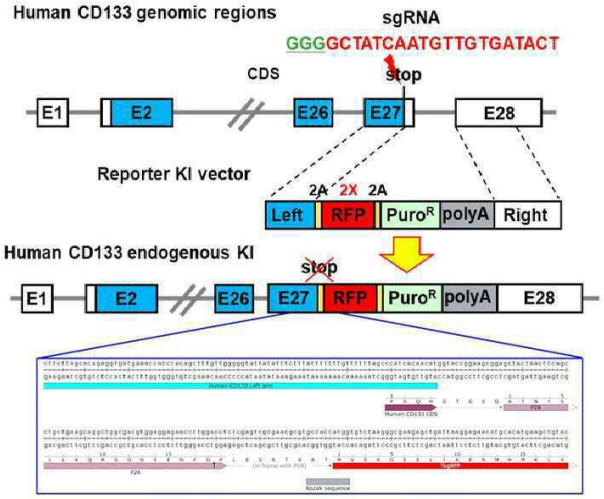
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **CD133 넉인 형광 리포터 벡터 시스템, 그 형질전환된 세포 및 이를 이용한 암줄기세포 검출방법**

(57) 요 약

본 발명은 암줄기세포 검출을 위한 CD133 넉인 형광 리포터 벡터 시스템, 이것으로 세포를 형질전환하는 방법, 상기 형질전환된 세포 및 이를 이용한 암줄기세포 확인 방법에 관한 것이다. 본 발명에 따른 형질전환된 세포는 CD133 내생 유전자의 종결 코돈을 결여시키고, 절단성 펩타이드 뒤에 형광단백질인 RFP(Red fluorescence protein)를 삽입하여 CD133이 존재하면 RFP를 발현하므로, 이를 통해 암줄기세포의 존재 유무를 실시간으로 모니터링할 수 있다. 나아가, 본 발명에 따른 형질전환된 세포는 다양한 항암 치료 약물의 스크리닝 및 약물의 내성을 확인하는데 이용될 수 있다. 따라서, 본 발명은 암 질환의 예방 또는 치료 개발에 유용하게 사용될 수 있다.

대 표 도 - 도1



(52) CPC특허분류
C12N 15/102 (2013.01)
C12N 15/113 (2013.01)
C12N 9/22 (2013.01)
C12Q 1/6897 (2018.05)
G01N 33/5044 (2013.01)
G01N 33/68 (2013.01)
C12N 2310/20 (2017.05)

(72) 발명자
송인성
서울특별시 송파구 올림픽로43길 88(풍납동, 서울아산병원 울산대학교 의과대학 의생명과학교실)
이동길
대전광역시 유성구 과학로 125(어은동)
정예나
대전광역시 유성구 과학로 125(어은동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711125186
과제번호	KGM4252021
부처명	과학기술정보통신부
과제관리(전문)기관명	한국생명공학연구원
연구사업명	한국생명공학연구원연구운영비지원(R&D)(주요사업비)
연구과제명	미니돼지 자원 활용 범용/맞춤형 인공혈액 개발사업
기 여 율	40/100
과제수행기관명	한국생명공학연구원
연구기간	2020.01.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1415168583
과제번호	20009707
부처명	산업통상자원부
과제관리(전문)기관명	한국산업기술평가관리원
연구사업명	바이오산업기술개발(R&D)
연구과제명	헬스케어 산업소재 적용 맞춤형 세포주 개발 및 제작 서비스 플랫폼
기 여 율	30/100
과제수행기관명	한국생명공학연구원
연구기간	2020.04.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711142232
과제번호	2019R1A2C1090362
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	난치성 폐암치료를 위한 암줄기세포 신규조절인자 기능 규명 및 임상유의성 검증
기 여 율	30/100
과제수행기관명	울산대학교
연구기간	2021.03.01 ~ 2022.02.28

명 세 서

청구범위

청구항 1

CD133 표적 부위 및 CD133의 발현 확인용 마커 단백질을 코딩하는 유전자를 포함하는 제1 벡터 및

상기 마커 단백질을 코딩하는 유전자를 CRISPR 상동-유도 수선(homology-directed repair) 방법으로 세포의 크로모좀에 삽입시키기 위한 sgRNA 및 Cas 단백질을 코딩하는 유전자를 포함하는 제2 벡터를 포함하는 암줄기세포 검출용 벡터 세트.

청구항 2

제1항에 있어서,

상기 마커 단백질은 녹색 형광 단백질(GFP), 황색 형광 단백질(YFP), 청록색 형광 단백질(CFP) 또는 적색 형광 단백질(RFP)인, 암줄기세포 검출용 벡터 세트.

청구항 3

제1항에 있어서,

상기 제1 벡터의 CD133 표적 부위는 5'-상동성 암(homology arm) 및 3'-상동성 암을 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 4

제3항에 있어서,

상기 5'-상동성 암은 CD133 유전자의 종결코돈의 업스트림(up-stream) 쪽으로 500 내지 2500 bp 이내에 위치하는 폴리뉴클레오타이드이고, 상기 3'-상동성 암은 CD133 유전자의 종결코돈의 다운스트림(down-stream) 쪽으로 1000 내지 3000 bp 이내에 위치하는 폴리뉴클레오타이드인 것인, 암줄기세포 검출용 벡터 세트.

청구항 5

제3항에 있어서,

상기 5'-상동성 암은 서열번호 43으로 표시되는 염기서열을 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 6

제3항에 있어서,

상기 3'-상동성 암은 서열번호 44로 표시되는 염기서열을 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 7

제1항에 있어서,

상기 sgRNA는 서열번호 45로 표시되는 염기서열을 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 8

제1항에 있어서,

상기 Cas 단백질은 Cas9, Cpf1, c2c1, C2c2, Cas13, c2c3, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e(CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a, Cas8a1, Cas8a2, Cas8b, Cas8c, Csn1, Csx12, Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1(CasA), Cse2(CasB), Cse3(CasE), Cse4(CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 또는 Cu1966로 구성된 군으로부터 선

택되는 어느 하나인 것인, 암줄기세포 검출용 벡터 세트.

청구항 9

제1항에 있어서,

상기 제1 벡터는 항생제 저항성 유전자를 추가로 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 10

제1항에 있어서,

상기 제1 벡터는 서열번호 41로 표시되는 염기서열을 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 11

제1항에 있어서,

상기 제2 벡터는 서열번호 38로 표시되는 염기서열을 포함하는 것인, 암줄기세포 검출용 벡터 세트.

청구항 12

제1항의 암줄기세포 검출용 벡터 세트로 세포를 형질전환하는 단계를 포함하는, 형질전환 세포 제조방법.

청구항 13

마커 단백질이 CD133의 내생 유전자 영역인 크로모좀에 삽입되어 CD133의 발현 정도를 확인할 수 있는 형질전환 세포.

청구항 14

제1항의 암줄기세포 검출용 벡터 세트로 형질전환시킨 세포 또는 제13항의 형질전환 세포의 마커 단백질의 발현 수준을 측정하는 단계를 포함하는 암줄기세포 확인 방법.

청구항 15

제12항의 방법으로 제조된 형질전환 세포 또는 제13항의 형질전환 세포에 후보 약물을 처리하여 마커 단백질의 발현 변화 수준을 측정하는 단계를 포함하는, 항암 후보 약물 스크리닝 또는 평가 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 암줄기세포 검출을 위한 CD133 넉인 형광 리포터 벡터 시스템, 이것으로 세포를 형질전환하는 방법, 상기 형질전환된 세포 및 이를 이용한 암줄기세포 검출방법에 관한 것이다.

배 경 기 술

[0002] 인체의 각 조직 내에 줄기세포가 존재하여 조직을 유지시켜 주듯이, 암조직 내에도 1% 내지 2% 정도의 암줄기세포(cancer stem cells, CSCs)가 존재한다. 암줄기세포는 정상적인 줄기세포처럼 자가복제능력(self-renewal)이 있으며, 다른 세포로 분화할 수 있는 능력(differentiation)을 갖고 있다. 이와 관련한 암줄기세포 이론(cancer stem cell theory)은 종양세포들을 수술, 항암화학요법 및 방사선 치료로 완전히 제거하더라도 암줄기세포가 남아 있으면 종양은 다시 재발하게 된다는 것으로, 암을 완전하게 치료하기 위해서는 암줄기세포를 완벽하게 제거해야 한다.

[0003] 이러한 암줄기세포의 분리 및 분석방법은 표준화되었으며, 많은 연구에서 암줄기세포를 암세포(Cancer cells, CCs)와 구분할 수 있는 특이적인 세포 표면 항원(CD13, CD44, CD105, CD133 등)을 발굴하고 있다. 또한, 암줄기세포에서 비정상적으로 활성화되어 있는 많은 신호전달경로가 밝혀지고 있으며, 약물내성의 원인으로 보고되고 있다.

[0004] 종양조직에서 암줄기세포의 분리 및 분석은 CD44, CD90, CD133 등과 같은 세포 표면 항원에 대한 항체를 이용하여 유세포분석(FACS, fluorescence activated cell sorting)이나 자기활성세포분석(MACS, magnetic-activated cell sorting)을 통해 수행될 수 있다(Hyun-Joo Lee *et al.*, *J Gastric Cancer* 2010;10(3):99-105). 하지만, 이러한 방법을 사용할 경우, 암줄기세포의 변화 유무는 확인할 수 있지만 실시간으로 어떻게 변화하는지는 알 수 없다. 따라서, 암줄기세포의 변화 유무 확인과 함께 실시간으로 모니터링할 수 있는 방법에 대한 연구가 필요한 실정이다.

선행기술문헌

비특허문헌

[0005] (비특허문헌 0001) Hyun-Joo Lee et al., J Gastric Cancer 2010;10(3):99-105

발명의 내용

해결하려는 과제

[0006] 이에, 본 발명자들은 CRISPR/Cas9 및 상동-유도 수선(HDR, homology-directed repair) 시스템을 이용하여 암줄기세포 특이적인 세포 표면 항원인 CD133의 내생 유전자(endogenous gene)의 종결 코돈을 결여시킨 후, 형광단백질을 코딩하는 유전자를 포함하는 넉인(knock-in) 리포터 벡터를 암세포에 도입시켜 암줄기세포의 존재뿐만 아니라 이를 실시간으로 모니터링 할 수 있음을 확인함으로써 본 발명을 완성하였다.

과제의 해결 수단

[0007] 상기 과제를 해결하기 위하여, 본 발명의 일 측면은, CD133 표적 부위 및 CD133의 발현 확인용 마커 단백질을 코딩하는 유전자를 포함하는 제1 벡터 및 상기 마커 단백질을 코딩하는 유전자를 CRISPR 상동-유도 수선(homology-directed repair) 방법으로 세포의 크로모좀에 삽입시키기 위한 sgRNA 및 Cas 단백질을 코딩하는 유전자를 포함하는 제2 벡터를 포함하는 암줄기세포 검출용 벡터 세트를 제공한다.

[0008] 본 발명의 다른 측면은, 상기 암줄기세포 검출용 벡터 세트로 세포를 형질전환하는 단계를 포함하는 형질전환 세포를 제조하는 방법을 제공한다.

[0009] 본 발명의 또 다른 측면은, 마커 단백질이 CD133의 내생 유전자 영역인 크로모좀에 삽입되어 CD133의 발현 정도를 확인할 수 있는 형질전환 세포를 제공한다.

[0010] 본 발명의 또 다른 측면은, 암줄기세포 검출용 벡터 세트로 형질전환시킨 세포 또는 상기 형질전환 세포의 마커 단백질의 발현 수준을 측정하는 단계를 포함하는 암줄기세포를 확인하는 방법을 제공한다.

[0011] 본 발명의 또 다른 측면은, 암줄기세포 검출용 벡터 세트로 형질전환시킨 세포 또는 상기 형질전환 세포에 후보 약물을 처리하여 마커 단백질의 발현 변화 수준을 측정하는 단계를 포함하는 항암 후보 약물 스크리닝 또는 평가하는 방법을 제공한다.

발명의 효과

[0012] 본 발명에 따른 형질전환된 세포는 CD133 내생 유전자의 종결 코돈을 결여시키고, 절단성 펩타이드 뒤에 형광단백질을 삽입하여 CD133의 발현 정도를 반영하는 형광단백질의 발현 수준을 확인 가능하므로, 이를 통해 암줄기세포의 존재 여부를 실시간으로 모니터링할 수 있다. 나아가, 본 발명에 따른 형질전환된 세포는 다양한 항암치료 약물의 스크리닝 및 약물의 내성을 확인하는데 이용될 수 있다. 따라서, 본 발명은 암 질환의 예방 또는 치료 개발에 유용하게 사용될 수 있다.

도면의 간단한 설명

[0013] 도 1은 CD133 표적 sgRNA-Cas9 벡터와 HDR(homology-directed repair)를 통해 CD133 내생 유전자에 삽입되는 CD133 넉인 형광 리포터 벡터 시스템을 도식화한 것이다.

도 2는 CCTop-CRISPR-Cas9 target online predictor site를 이용한 CD133 표적 sgRNA의 위치와 오프 표적(off

target)의 가능성을 확인한 것이다.

도 3은 CD133 표적 sgRNA의 CD133 유전자상의 위치 및 CD133 표적 sgRNA를 기준으로 HDR에 필요한 LA(left arm) 및 RA(right arm) 영역을 표시한 것이다.

도 4는 CD133 표적 sgRNA-Cas9 벡터 제작과정을 도식화한 것이다.

도 5는 pCAG-SpCas9-GFP-U6-sgRNA 벡터의 개요도를 나타낸 것이다.

도 6은 pCAG-SpCas9-GFP-U6-CD133-sgRNA 벡터의 개요도를 나타낸 것이다.

도 7은 CCTop-CRISPR-Cas9 target online predictor site를 이용한 H11 표적 sgRNA의 위치와 오프 표적의 가능성을 확인한 것이다.

도 8은 H11 표적 sgRNA의 H11 locus상의 위치를 확인한 것이다.

도 9는 pCAG-SpCas9-GFP-U6-hH11-sgRNA 벡터의 개요도를 나타낸 것이다.

도 10은 Loxp-TP2A-TagRFP 2X-T2A-Puro-polyA-Loxp-Zsgreen in pCAGGS 벡터의 개요도를 나타낸 것이다.

도 11은 human CD133 Left arm 벡터의 개요도를 나타낸 것이다.

도 12a 내지 도 12d는 human CD133 Right arm 벡터(Human CD133 endogenous HDR 벡터)의 개요도를 나타낸 것이다.

도 13은 Loxp-TP2A-Tag RFP 2X-T2A-puro-polyA in pCAGGS 벡터의 개요도를 나타낸 것이다.

도 14는 Human CD133 partial promoter 포함 RMCE 2X RFP reporter 벡터의 개요도를 나타낸 것이다.

도 15는 HT-29 세포에 CD133 표적 sgRNA-Cas9 벡터와 CD133 내생 프로모터-RFP 벡터를 도입시킨 후, 암줄기세포의 구 형성 및 RFP 발현을 확인한 것이다.

도 16은 HT-29 세포, NCCIT 세포 및 SH-SY5Y 세포에 H11 표적 sgRNA 벡터 및 CD133 부분적 프로모터-RFP 벡터; 또는 CD133 표적 sgRNA-Cas9 벡터 및 CD133 내생 프로모터-RFP 벡터를 도입시킨 후, 암줄기세포의 구 형성 및 RFP 발현을 확인한 것이다.

도 17은 HT-29 CD133-RFP Knock-in 세포의 genomic DNA의 integration PCR 결과 및 분석 결과를 나타낸 것이다.

도 18은 HCT-116 CD133-RFP Knock-in 세포의 genomic DNA의 integration PCR 결과를 나타낸 것이다.

도 19는 HT-29 CD133-RFP Knock-in 세포의 genomic DNA의 Genotyping PCR 결과 및 분석 결과를 나타낸 것이다.

도 20(가)는 HT-29 CD133-RFP Knock-in 세포를 혈청이 포함된 배지 및 무혈청 배지에서 배양하였을 경우의 구 형성 정도를 관찰한 것이다. 도 20(나)는 다른 암세포주인 NCCIT CD133-RFP Knock-in, SNU-16 CD133-RFP Knock-in 세포를 무혈청 배지에서 배양하였을 경우 구 형성 정도를 관찰한 것이다.

도 21은 HCT-116 CD133-RFP Knock-in 세포를 혈청이 포함된 배지 및 무혈청 배지에서 배양하였을 경우의 구 형성 정도를 관찰한 것이다.

도 22는 CD133 microbead 및 FACS를 이용한 세포 분리 방법을 도식화한 것이다.

도 23은 HT-29 세포 및 HT-29 CD133-RFP Knock-in 세포를 각각 CD133 항체/자성, CD133 항체/유세포분석기 및 CD133-RFP/유세포분석기를 이용하여 분리 후 RT-qPCR을 통해 대장암 줄기세포 인자들 및 다분화능 인자들을 비교한 것이다.

도 24는 10일간 배양한 HCT-116 CD133-RFP Knock-in 세포의 구형형성 및 RFP 발현을 관찰한 것이다.

도 25는 RFP 형광 발현이 약한 HCT-116 CD133-RFP Knock-in 세포 그룹(FACS sorting, RFPlow) 및 RFP 형광 발현이 강한 HCT-116 CD133-RFP Knock-in 세포 그룹(FACS sorting, RFPhigh)에서의 상대적인 P2A-RFP mRNA 발현량을 비교한 것이다.

도 26은 RFP 형광 발현이 약한 HCT-116 CD133-RFP Knock-in 세포 그룹(FACS sorting, RFPlow) 및 RFP 형광 발현이 강한 HCT-116 CD133-RFP Knock-in 세포 그룹(FACS sorting, RFPhigh)에서의 대장암 줄기세포 마커, 전사

인자 마커 및 분화세포 마커의 mRNA 발현량을 나타낸 것이다.

도 27은 5-fluorouracil 및/또는 frugoside 처리 여부에 따른 HT-29 CD133-RFP Knock-in 세포주에서의 RFP 발현량을 형광현미경으로 관찰한 결과 및 REP 양성인 세포를 포함하는 콜로니 형성 차이를 soft-agar colony-forming assay로 측정한 결과를 각각 도시한다.

도 28은 5-fluorouracil 및/또는 frugoside 처리 여부에 따른 HT-29 CD133-RFP Knock-in 세포주에서의 RFP 발현량을 광학현미경 및 형광현미경으로 관찰한 결과 및 REP 양성인 세포를 포함하는 스피어 형성 차이를 sphere formation assay로 측정한 결과를 각각 도시한다.

발명을 실시하기 위한 구체적인 내용

[0014] 이하, 본 발명에 대하여 상세히 설명하도록 한다.

[0015] 본 발명의 일 측면은, CD133 표적 부위 및 CD133의 발현 확인용 마커 단백질을 코딩하는 유전자를 포함하는 제1 벡터, 및 상기 마커 단백질을 코딩하는 유전자를 CRISPR 상동-유도 수선(homology-directed repair) 방법으로 세포의 크로모좀에 삽입시키기 위한 sgRNA 및 Cas 단백질을 코딩하는 유전자를 포함하는 제2 벡터를 포함하는 암줄기세포 검출용 벡터 세트를 제공한다. 또한, 상기 세포는 동물에서 유래한 세포일 수 있으며, 바람직하게는 포유 동물, 특히 인간에서 유래한 인간 세포일 수 있다.

[0016] 상기 마커 단백질은 형광 단백질일 수 있으며, 구체적으로 녹색 형광 단백질(GFP; green fluorescent protein), 황색 형광 단백질(YFP; yellow fluorescence protein), 청록색 형광 단백질(CFP; cyan fluorescent protein) 또는 적색 형광 단백질(RFP; red fluorescent protein)일 수 있다. 특히, 상기 마커 단백질은 적색 형광 단백질일 수 있다.

[0017] 상기 마커 단백질을 코딩하는 유전자는 녹색 형광 단백질, 황색 형광 단백질, 청록색 형광 단백질 또는 적색 형광 단백질을 코딩하는 유전자일 수 있으며, 서열번호 47 내지 50 중 어느 하나로 표시되는 염기서열을 포함하는 것일 수 있다. 또한, 상기 마커 단백질을 코딩하는 유전자중 적어도 하나가 벡터에 포함될 수 있으며, 상기 유전자 복수개가 하나의 벡터에 포함될 수도 있다.

[0018] 상기 CD133 표적 부위는 *CD133* 유전자 내 종결 코돈(stop codon)의 업스트림(up-stream) 또는 다운스트림(down-stream) 쪽으로 각각 약 100 내지 약 4000 뉴클레오티드(bp)를 갖는 영역일 수 있다.

[0019] 상기 제1 벡터의 CD133 표적 부위는 5'-상동성 암(homology arm) 및 3'-상동성 암을 포함할 수 있다. 상기 5'-상동성 암은 CD133 유전자의 종결코돈의 업스트림(up-stream)쪽으로 500 내지 2500 bp 이내에 위치하는 폴리뉴클레오타이드이고, 상기 3'-상동성 암은 CD133 유전자의 종결코돈의 다운스트림(down-stream)쪽으로 1000 내지 3000 bp 이내에 위치하는 폴리뉴클레오타이드일 수 있다.

[0020] 구체적으로, CD133 표적 부위의 5'-상동성 암은 서열번호 43으로 표시되는 염기서열을 포함하는 것일 수 있다. 상기 5'-상동성 암은 서열번호 43으로 표시되는 염기서열과 적어도 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 것일 수 있다. 또한, CD133 표적 부위의 3'-상동성 암은 서열번호 44로 표시되는 염기서열을 포함하는 것일 수 있다. 상기 3'-상동성 암은 서열번호 44로 표시되는 염기서열과 적어도 90%, 95%, 96%, 97%, 98%, 또는 99% 동일한 것일 수 있다.

[0021] 상기 sgRNA는 두 개의 RNA, 즉, crRNA(CRISPR RNA) 및 tracrRNA(trans-activating crRNA)를 구성요소로 포함하는 이중 RNA(dual RNA); 또는 표적 서열과 전부 또는 일부 상보적인 서열을 포함하는 제1부위 및 크리스퍼 연관 단백질과 상호작용하는 서열을 포함하는 제2부위를 포함하는 형태를 말하나, 크리스퍼 연관 단백질이 표적 서열에서 활성을 가질 수 있는 형태라면 제한 없이 본 발명의 범위에 포함될 수 있다.

[0022] 구체적으로, 상기 sgRNA는 *CD133* 유전자에 상보적으로 결합할 수 있는 서열을 포함할 수 있다. 특히, 상기 sgRNA는 *CD133* 유전자의 종결 코돈(stop codon) 근처에서 20 bp 이내의 상보적으로 결합할 수 있는 서열을 포함할 수 있다. 상기 sgRNA는 *CD133*를 코딩하는 DNA에 상보적으로 결합 가능한 15개 내지 35개의 뉴클레오티드를 포함하는 올리고뉴클레오티드일 수 있다. 상기 sgRNA는 *CD133* 유전자의 표적 서열에 상보적으로 결합할 수 있는 15개 내지 35개, 20개 내지 25개, 또는 23개의 뉴클레오티드를 포함하는 올리고뉴클레오티드일 수 있다. 상기 sgRNA는 서열번호 45로 표시되는 염기서열을 포함하는 것일 수 있다.

[0023] 상기 Cas 단백질은 Cas9, Cpf1, c2c1, C2c2, Cas13, c2c3, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e(CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a, Cas8a1, Cas8a2, Cas8b, Cas8c, Csn1, Csx12, Cas10,

Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1(CasA), Cse2(CasB), Cse3(CasE), Cse4(CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 또는 Cul966로 구성된 군으로부터 선택되는 어느 하나일 수 있다. 구체적으로, 상기 Cas 단백질은 Cas9일 수 있다.

[0024] 이때, 상기 Cas 단백질 또는 유전자의 서열 정보는 NCBI(National Center for Biotechnology Information)의 GenBank와 같은 공지의 데이터 베이스에서 얻을 수 있다. 상기 Cas9를 코딩하는 유전자는 서열번호 46으로 표시되는 염기서열을 포함하는 것일 수 있다.

[0025] 상기 제1 벡터는 항생제 저항성 유전자를 추가로 포함할 수 있다. 상기 항생제 저항성 유전자는 페니실린, 스트렙토마이신, 네오마이신(Neomycin, G418), 퓨로마이신(Puromycin) 및 히그로마이신(Hygromycin)로 이루어진 군에서 선택되는 1종 이상의 항생제에 대한 저항성 유전자일 수 있다. 구체적으로 상기 항생제 저항성 유전자는 퓨로마이신에 대한 저항성 유전자일 수 있으며, 서열번호 51로 표시되는 염기서열을 포함하는 것일 수 있다.

[0026] 상기 제1 벡터는 서열번호 41로 표시되는 염기서열을 포함하는 것일 수 있다. 또한, 상기 제2 벡터는 서열번호 38로 표시되는 염기서열을 포함하는 것일 수 있다.

[0027] 본 발명의 다른 측면은, 상기 암줄기세포 검출용 벡터 세트로 세포를 형질전환하는 단계를 포함하는 형질전환 세포를 제조하는 방법을 제공한다. 이때, 상기 암줄기세포 검출용 벡터 세트는 상술한 바와 동일하다.

[0028] 본 발명의 또다른 측면은, 상기 암줄기세포 검출용 벡터 세트로 인간 세포를 형질전환하는 단계를 포함하는 형질전환 인간 세포를 제조하는 방법을 제공한다. 이때, 상기 암줄기세포 검출용 벡터 세트는 상술한 바와 동일하다.

[0029] 상기 벡터를 세포 내로 형질전환하는 방법은 당해 분야에서 공지된 방법을 이용할 수 있으며, 예를 들면, 일시적 형질감염(transient transfection), 미세주사, 형질도입(transduction), 세포융합, 칼슘 포스페이트 침전법, 리포좀 매개된 형질감염(liposome-mediated transfection), DEAE 덱스트란-매개된 형질감염(DEAE Dextran-mediated transfection), 폴리브렌-매개된 형질감염(polybrene-mediated transfection), 전기침공법(electroporation), 유전자 총(gene gun) 또는 세포 내로 핵산을 유입시키기 위한 다른 공지의 방법에 의해 세포 내로 도입할 수 있다(Wu et al., J. Bio. Chem., 267:963-967, 1992; Wu and Wu, J. Bio. Chem., 263:14621-14624, 1988). 하지만, 이에 제한하는 것은 아니다.

[0030] 상기 세포에서 벡터가 잘 도입되었는지 확인하는 방법은 예로서, 당업자에게 널리 공지된 분자 생물학적 검정, 예를 들면, 서던 및 노던 블로팅, RT-PCR 및 PCR; 생화학적 검정, 예로서, 면역학적 방법(가령, ELISAs 및 웨스턴 블롯)에 의한 특정 단백질/펩티드의 존재 또는 부재의 검출을 포함한다.

[0031] 상기 방법은 추가적으로, 마커 단백질을 발현하는 세포를 단리하는 단계를 포함할 수 있다. 또한, 상기 방법은 추가적으로, 마커 단백질을 발현하는 인간 세포를 단리하는 단계를 포함할 수 있다.

[0032] 상기 마커 단백질을 발현하는 세포를 단리하는 방법은 항생제를 처리하여 항생제에 저항성을 갖는 세포에 대해 1차 선별하는 과정을 통해 마커 단백질을 발현하는 세포를 단리할 수 있다. 또한, 상기 마커 단백질을 발현하는 세포를 단리하는 방법은 FACS(Fluorescence-activated Cell Sorting)를 이용해서 형광단백질을 발현하는 세포를 선별하는 과정을 통해 마커 단백질을 발현하는 세포를 단리할 수 있다. 나아가, 상기 마커 단백질을 발현하는 세포를 단리하는 방법은 항생제를 처리하여 항생제에 저항성을 갖는 세포에 대해 1차 선별한 후, FACS(Fluorescence-activated Cell Sorting)를 이용해서 형광단백질을 발현하는 세포를 선별하는 과정을 통해 마커 단백질을 발현하는 세포를 단리할 수 있다.

[0033] 또한, 상기 마커 단백질을 발현하는 인간 세포를 단리하는 방법은 항생제를 처리하여 항생제에 저항성을 갖는 인간 세포에 대해 1차 선별하는 과정을 통해 마커 단백질을 발현하는 인간 세포를 단리할 수 있다. 또한, 상기 마커 단백질을 발현하는 인간 세포를 단리하는 방법은 FACS(Fluorescence-activated Cell Sorting)를 이용해서 형광단백질을 발현하는 인간 세포를 선별하는 과정을 통해 마커 단백질을 발현하는 세포를 단리할 수 있다. 나아가, 상기 마커 단백질을 발현하는 인간 세포를 단리하는 방법은 항생제를 처리하여 항생제에 저항성을 갖는 인간 세포에 대해 1차 선별한 후, FACS(Fluorescence-activated Cell Sorting)를 이용해서 형광단백질을 발현하는 인간 세포를 선별하는 과정을 통해 마커 단백질을 발현하는 인간 세포를 단리할 수 있다.

[0034] 본 발명의 일 실시예에서는 CD133 표적 sgRNA-Cas9 벡터와 CD133 내생 프로모터-RFP 벡터를 각각 2 ㎍와 5 ㎍씩 넣어 형질감염시키고, 7일 후 세포가 90% 정도 배양되면 배양액 10 ㎖에 퓨로마이신(#A11138-03, gibco) 3 ㎕를

넣어서 퓨로마이신 선별을 진행하였으며, 14일 후 세포가 70% 내지 80% 정도 배양되었을 때, FACS(Fluorescence-activated Cell Sorting)를 이용해서 RFP만 발현하는 세포만 2차로 선별하였다.

[0035] 본 발명의 또 다른 측면은, 마커 단백질이 CD133의 내생 유전자 영역인 크로모좀에 삽입되어 CD133의 발현 정도를 확인할 수 있는 형질전환 세포를 제공한다. 이때, 상기 마커 단백질은 상술한 바와 동일하다. 상기 형질전환 세포는 상기 "형질전환 세포를 제조하는 방법"으로 제조될 수 있다.

[0036] 본 발명의 다른 측면은, 마커 단백질이 CD133의 내생 유전자 영역인 크로모좀에 삽입되어 CD133의 발현 정도를 확인할 수 있는 형질전환 인간 세포를 제공한다. 이때, 상기 마커 단백질은 상술한 바와 동일하다. 상기 형질전환 인간 세포는 상기 "형질전환 인간 세포를 제조하는 방법"으로 제조될 수 있다.

[0037] 본 발명의 또 다른 측면은, 암줄기세포 검출용 벡터 세트로 형질전환시킨 세포 또는 상기 마커 단백질이 CD133의 내생 유전자 영역인 크로모좀에 삽입되어 CD133의 발현 정도를 확인할 수 있는 형질전환 세포의 마커 단백질의 발현 수준을 측정하는 단계를 포함하는 암줄기세포를 확인하는 방법을 제공한다. 이때, 암줄기세포 검출용 벡터 세트 및 형질전환 세포는 상술한 바와 동일하다.

[0038] 본 발명의 다른 측면은, 암줄기세포 검출용 벡터 세트로 형질전환시킨 인간 세포 또는 상기 마커 단백질이 CD133의 내생 유전자 영역인 크로모좀에 삽입되어 CD133의 발현 정도를 확인할 수 있는 형질전환 인간 세포의 마커 단백질의 발현 수준을 측정하는 단계를 포함하는 암줄기세포를 확인하는 방법을 제공한다. 이때, 암줄기세포 검출용 벡터 세트 및 형질전환 인간 세포는 상술한 바와 동일하다.

[0039] 상기 마커 단백질의 발현 수준을 측정하는 방법은 당업자에게 널리 공지된 분자 생물학적 검정, 예를 들면, 서던 및 노던 블로팅, RT-PCR 및 PCR; 생화학적 검정, 예로서, 형광 현미경, 면역학적 방법(예를 들면, FACS, ELISA 및 웨스턴 블롯)을 이용하여 확인할 수 있다. 본 발명의 일 실시예에서는 형광 현미경을 이용하여 마커 단백질 즉, 형광 단백질의 발현 수준을 측정하였다.

[0040] 본 발명의 또 다른 측면은, 상기 형질전환 세포 또는 상기 형질전환 인간 세포에 후보 약물을 처리하여 마커 단백질의 발현 변화 수준을 측정하는 단계를 포함하는 항암 후보 약물 스크리닝 또는 평가하는 방법을 제공한다. 이때, 암줄기세포 검출용 벡터 세트, 형질전환 세포 및 형질전환 인간 세포는 상술한 바와 동일하다.

[0041] 상기 형질전환 세포 또는 형질전환 인간 세포에 후보 약물을 처리하여 RFP의 발현 변화 수준을 측정하였을 때, RFP의 발현 수준이 감소하는 경우 해당 후보 약물이 항암 효과가 있는 약물로 스크리닝할 수 있다.

[0042] 또한, 상기 형질전환 세포 또는 형질전환 인간 세포에 기존의 항암 치료 약물을 처리하여 RFP의 발현 변화 수준을 측정하였을 때, RFP의 발현 수준이 증가 또는 유의미한 변화가 없는 경우 해당 항암 치료 약물은 약물 내성이 생길 수 있다는 것을 확인할 수 있다.

[0043] 이하, 본 발명의 이해를 돕기 위하여 바람직한 실시예를 제시한다. 그러나 하기의 실시예는 본 발명을 보다 쉽게 이해하기 위하여 제공되는 것일 뿐, 하기 실시예에 의해 본 발명의 내용이 한정되는 것은 아니다.

[0044] **I. CD133 넉인 형광 리포터 벡터 시스템 제조**

[0045] 인간세포에서 CRISPR/Cas9 및 sgRNA을 이용한 HDR(homology-directed repair)를 통해 CD133의 발현을 실시간으로 모니터링할 수 있도록, CD133 내생 유전자(endogenous gene)의 종결코돈을 결여시키고 2A 펩타이드 뒤에 형광을 발현하는 RFP(Red fluorescence protein)를 삽입하였다(도 1). 상기 2A 펩타이드는 self-processing viral peptide이다.

[0046] **실시예 1. CD133 내생 유전자(endogenous gene)에 삽입하기 위한 CD133 표적 sgRNA-Cas9 벡터 제작**

[0047] Cas9 및 sgRNA를 발현하는 플라스미드 벡터를 구축하기 위해 NCBI Gene site(https://www.ncbi.nlm.nih.gov/gene/)에서 CD133 유전자의 염기서열을 검색하고, 종결 코돈(stop codon) 근처에서 20 bp의 sgRNA 서열을 설정하였으며, 이를 하기 표 1에 나타내었다.

표 1

[0048]

CD133	F	TCAGCTATCAATGTTGTGAT	서열번호 1
	R	ATCACAACATTGATAGCTGA	서열번호 2
H11	F	CAAAATAGGTTAGCCTTGCG	서열번호 3
	R	CGCAAGGCTAACCTATTTTG	서열번호 4

[0049] sgRNA는 sgRNA의 위치와 오프 표적(off target)의 가능성을 확인해주는 CCTop-CRISPR-Cas9 target online predictor site(https://crispr.cos.uni-heidelberg.de/)를 이용하여 선택하였으며, 마크로젠(macrogen)사의 duplex oligo 제작 서비스(https://dna.macrogen.com/)를 이용해서 제작하였다. 이때, CCTop-CRISPR-Cas9 target online predictor를 이용해서 오프 표적 수가 적으면서 유전자 가위의 자르는 효율이 높은 CD133 표적 sgRNA 염기서열을 선정하였다(도 2). CD133을 표적으로 하는 sgRNA 염기서열의 CD133 유전자상의 위치 및 CD133 표적 sgRNA를 기준으로 HDR(homology-directed repair)에 필요한 LA(left arm) 및 RA(right arm) 영역을 선택하였다(도 3).

[0050] 그 후, 각각의 20 bp의 sgRNA를 인서트(insert)로, pCAG-SpCas9-GFP-U6-sgRNA plasmid(#79144, Addgene)를 BbsI(#R0539, NEB) 제한효소로 절단해서 백본 벡터(backbone vector)로 사용하였다. 이때, 인서트와 백본 벡터를 16℃ 온도에서 1시간 동안 라이게이션(ligation)한 후, HIT Competent Cells-DH5alpha High 108(#RH618, RBC)을 사용해서 LB broth 플레이트에 형질전환하여, pCAG-SpCas9-GFP-U6-CD133 gRNA 벡터(서열번호 38)를 제조하였다(도 4 내지 도 6).

[0051] **실시예 2. CD133 부분적 프로모터-RFP(partial promoter) 벡터를 H11 로커스(locus)에 삽입하기 위한 H11 표적 sgRNA 벡터 제작**

[0052] CD133 부분적 프로모터-RFP 벡터를 H11 로커스(세이프 하버(safe harbor) 중 하나)에 삽입시키기 위해서 CD133 표적 sgRNA를 선정하였고(도 2), 상기 실시예 1과 동일한 방법으로 H11 표적 sgRNA를 제작하였다. 이때, CCTop-CRISPR-Cas9 target online predictor를 이용해서 오프 표적 수가 적으면서 유전자 가위의 자르는 효율이 높은 H11 표적 sgRNA 염기서열을 선정하였다(도 7). 또한, H11 표적 sgRNA 염기서열 위치를 확인하였으며, 이를 도 8에 나타내었다.

[0053] 그 후, 실시예 1과 동일하게 각각의 20 bp의 sgRNA를 인서트로 pCAG-SpCas9-GFP-U6-sgRNA plasmid를 BbsI(#R0539, NEB) 제한효소로 절단해서 백본 벡터로 사용하였다. 이때, 인서트와 백본 벡터를 16℃ 온도에서 1시간 동안 라이게이션한 후, HIT Competent Cells-DH5alpha High 108(#RH618, RBC)을 사용해서 LB broth 플레이트에 형질전환하여 pCAG-SpCas9-GFP-U6 H11 gRNA 벡터(서열번호 39)를 제조하였다(도 9).

[0054] **실시예 3. CD133 내생 프로모터(endogenous promoter)-RFP 벡터 제작**

[0055] CD133 내생 프로모터-RFP 영역을 얻기 위해서 NCBI Gene site(https://www.ncbi.nlm.nih.gov/gene/)에서 CD133 genomic DNA의 염기서열을 찾아서 종결 코돈 앞, 뒤에서 LA와 RA를 설정하였으며, 이를 하기 표 2에 나타내었다.

표 2

[0056]

CD133	Homo sapiens prominin 1(PROM1), RefSeqGene on chromosome 4 NCBI Reference Sequence: NG_011696.2			
	Endogenous Left arm	F	ATCAAGTGTATCATATGTTCAGTGTCCCCTCCAAAACTC	서열번호 5
		R	GTAGCTCCGCTTCCGGTACCATGTTGTGATGGGCTAAAAAACAA	서열번호 6
	Endogenous Right arm	F	AGCCCAAGCTAGATCGAATTCTAGCTGATGTTGAAACTGCTTGAGG	서열번호 7
		R	GGCGGCCGCTCTAGCCTCGAGGCACACTCCACACATGGCA	서열번호 8
	Partial	F	GATTCATTAATGCAGCTGTGCAGGGGCTGCAGAGGGGAAAT	서열번호 9
		R	TCTCCAGATCCATCGATCGCTGCCGCCGCCCTTCCCTTA	서열번호 10

[0057] 그 후, CD133의 LA 및 RA를 클로닝하기 위해서 HEK293 세포(#CRL-1573™, ATCC)에서 추출한 genomic DNA를 KOD OneTM PCR Master Mix(#KMM-201, TOYOBO)를 이용하여 하기 표 3의 조건으로 PCR하여 CD133 부분적 프로모터-RFP의 PCR 산물을 얻었다.

표 3

[0058]

온도	시간	
98℃	2분 - 초기 변성	
98℃	10초	35 사이클 (1kbp 이하: 10초, 1-2kbp: 20초, 2-3kbp: 30초)
62℃	10초	
68℃	30초	
68℃	5분 - 최종 신장	

[0059] 얻어진 LA의 PCR 산물부터 순차적으로 Loxp-TP2A-TagRFP 2X-T2A-Puro-polyA-Loxp-Zsgreen in pCAGGS(도 10, 서열번호 40)를 백본 벡터에 삽입한 후, RA의 PCR 산물을 서브 클로닝하여 제작하였다.

[0060] 구체적으로, 상기 Loxp-TP2A-TagRFP 2X-T2A-Puro-polyA-Loxp-Zsgreen in pCAGGS 백본 벡터는 제한 효소를 포함한 TP2A-TagRFP 2X-T2A-Puro-polyA 염기서열을 합성하여 pCAG-loxP-STOP-loxP-ZsGreen(#51269, Addgene) 벡터의 HindIII-HF®(#R3104L, NEB), BamHI-HF®(#R3136L, NEB) 제한효소로 절단한 후, In-Fusion® HD Cloning Kit(#639649, TAKARA) 및 HIT Competent Cells-DH5alpha High 108(#RH618, RBC)을 이용하여 LB broth 플레이트에 형질전환하여 CD133 HDR을 위한 백본 벡터를 제작하였다. CD133 LA는 상기 Loxp-TP2A-TagRFP 2X-T2A-Puro-polyA-Loxp-Zsgreen in pCAGGS를 NdeI(#R0111, NEB), KpnI-HF®(#R3142, NEB) 제한효소로 절단한 후, In-Fusion® HD Cloning Kit(#639649, TAKARA) 및 HIT Competent Cells-DH5alpha High 108(#RH618, RBC)을 이용하여 LB broth 플레이트에 형질전환하여 CD133 LA가 삽입된 서브 클로닝된 벡터를 제작하였다. 이때, In-Fusion® HD Cloning Kit 는 제조사의 매뉴얼에 따라 수행하였다(도 11).

[0061] CD133 RA는 상기 CD133 left arm-P2A-Tag RFP 2X-T2A-puro-polyA-loxp-Zsgreen vector를 EcoRI-HF®(#R3101, NEB), PspX I(#R0656, NEB) 제한효소로 절단한 후, In-Fusion® HD Cloning Kit(#639649, TAKARA) 및 HIT Competent Cells-DH5alpha High 108(#RH618, RBC)을 이용하여 LB broth 플레이트에 형질전환하여 CD133 LA와 RA가 모두 삽입된 CD133 내생 프로모터-RFP(Human CD133 endogenous HDR) 벡터(도 12a, 서열번호 41)를 제작하였다.

[0062] 또한, 상기 방법과 동일하게 CD133 내생 프로모터-YFP(Human CD133 endogenous HDR) 벡터(도 12b, 서열번호 54) 및 CD133 내생 프로모터-CFP(Human CD133 endogenous HDR) 벡터(도 12c, 서열번호 55) 및 CD133 내생 프로모터-EGFP(Human CD133 endogenous HDR) 벡터(도 12d, 서열번호 56)를 제조하였다.

[0063] **실시예 4. CD133 부분적 프로모터-RFP(partial promoter)-RFP 벡터 제작**

[0064] CD133 부분적 프로모터 영역을 얻기 위해 Genecopoeia site(https://www.genecopoeia.com/product/search/detail.php?prt=22&cid=&key=HPRM36287&type=promoter&choose=HPRM36287)에서 상업적으로 판매하는 프로모터 클론 중 TSS(transcription start site)영역을 포함하는 가장 긴 프로모터 클론(#Cat. HPRM36287) 영역(서열번호 52)을 HEK293 세포(#CRL-1573™, ATCC)에서 추출한 genomic DNA에서 KOD OneTM PCR Master Mix(#KMM-201, TOYOBO)를 이용하여 상기 표 3의 조건으로 PCR하여 CD133 부분적 프로모터의 PCR 산물을 얻었다.

[0065] PCR 산물과 Pvu II-HF®(#R3151, NEB), Cla I(#R0197, NEB) 제한효소로 절단해서 미리 준비한 Loxp-TP2A-Tag RFP 2X-T2A-puro-polyA in pCAGGS(도 13, 서열번호 42)를 백본 벡터로 사용하였다. 상기 Loxp-TP2A-Tag RFP 2X-T2A-puro-polyA in pCAGGS 백본 벡터는 Loxp-TP2A-Tag RFP 2X-T2A-puro-polyA-Loxp-Zsgreen in pCAGGS를 백본 벡터로 하여 BamHI(#R3136L, NEB)과 HindIII(#R3104L, NEB)제한 효소를 처리하여 불필요한 Zsgreen-bGH polyA를 제거하고 비특이적 염기서열로 이부분을 채웠으며, H11 로커스(safe harbor) 타겟 염기서열을 포함한 영역을 합성한 후 Zsgreen-bGH polyA가 제거된 백본 벡터의 PciI(#R0655L, NEB)과 PspXI(#R0656L, NEB)제한효소로 자른 후 infusion 방법을 통해 크로모좀의 H11 로커스에 삽입되면서 부분적인 CD133 프로모터의 역할을 볼 수 있는 벡터로 제작하였다.

[0066] 그 후, In-Fusion® HD Cloning Kit(#639649, TAKARA) 및 HIT Competent Cells-DH5alpha High 108(#RH618, RBC)을 이용하여 LB broth 플레이트에 형질전환하여 CD133 부분적 프로모터-RFP(Human CD133 promoter in RMCE 2X RFP reporter) 벡터를 제작하였다(도 14, 서열번호 53). 이때, In-Fusion® HD Cloning Kit는 제조사의 매뉴얼에 따라 수행하였다.

[0067] **실시예 5. CD133 표적 sgRNA-Cas9 벡터 및 CD133 내생 프로모터-RFP 벡터를 이용한 암세포 형질전환 및 확인**

[0068] 실시예 1 및 실시예 3에서 제조한 CD133 표적 sgRNA-Cas9 벡터 및 CD133 내생 프로모터-RFP 벡터 또는 실시예 2 및 실시예 3에서 제조한 H11 표적 sgRNA 벡터 및 CD133 부분적 프로모터-RFP 벡터를 이용해 암세포를 형질전환시킨 후 RFP가 발현되는 것을 확인하였다.

[0069] 먼저, 인간 대장암 세포주인 HT-29 세포(ATCC®, #HIB-38™), 태생성 암종 세포주인 NCCIT 세포(ATCC®, #CRL-2073™), 인간 위암 세포주인 SNU-16 세포(ATCC®, CRL-5974™) 및 신경아세포종 세포주인 SH-SY5Y 세포(ATCC®, #CRL-2266™)는 ATCC에서 구입하여 배양하였다. 또한, 인간 대장암 세포주인 HCT-116 세포(ATCC®, CCL-247EMT™)을 ATCC에서 구입하여 배양하였다. 이때, HT-29 세포 및 HCT-116 세포 배양 조건은 5% CO_2, 37℃ 인큐베이터에서 10% FBS(#16000, ThermoFisher) 및 1% 페니실린/스트렙토마이신(#15140, ThermoFisher)을 넣은 Dulbecco의 Eagle 's media(DMEM, #11995, ThermoFisher)을 사용해서 배양하였다.

[0070] 또한, NCCIT 세포 배양 조건은 5% CO_2, 37℃ 인큐베이터에서 10% FBS(#16000, ThermoFisher) 및 1% 페니실린/스트렙토마이신(#15140, ThermoFisher)을 넣은 RPMI-1640(#11875, ThermoFisher)을 사용해서 배양하였다. SH-SY5Y 세포 배양 조건은 5% CO_2, 37℃ 인큐베이터에서 10% FBS(#16000, ThermoFisher) 및 1% 페니실린/스트렙토마이신(#15140, ThermoFisher)을 넣은 DMEM/F12(#11320, ThermoFisher)을 사용해서 배양하였다.

[0071] 그 후, HT-29 세포, SH-SY5Y 세포 및 SNU-16 세포를 100 mm 접시에 70% 내지 80% 정도 세포가 배양되었을 때, 4D-Nucleofector™ 장비를 이용하여 각 세포의 조건(HT-29, SF kit, FF-137 program; SH-SY5Y, SF kit, CA-137 program; NCCIT, SF kit, EH-100 program; SNU16, SF kit, DS-150 program)에 따라 제조사의 매뉴얼에 따라 형질감염 시켰다. 형질감염시 CD133 표적 sgRNA-Cas9 벡터와 CD133 내생 프로모터-RFP 벡터를 각각 2 ㎍와 5 ㎍씩 넣어주었다. 7일 후 세포가 90% 정도 배양되면 배양액 10 ㎖에 퓨로마이신(#A11138-03, gibco) 3 ㎕를 넣어서 퓨로마이신 선별을 진행했다. 퓨로마이신 선별 14일 후 세포가 70% 내지 80% 정도 배양되었을 때, FACS(Fluorescence-activated Cell Sorting)를 이용해서 RFP만 발현하는 세포만 2차로 선별하였다.

[0072] 또한, 60mm 접시에 60% 내지 70% 정도 HCT-116 세포가 배양되었을 때, 4D-Nucleofector™ 장비 및 lipofectamine 3000(#L3000001, ThermoFisher)을 이용하여 제조사의 매뉴얼에 따라 HCT-116 세포를 형질감염시켰다. 형질감염시 CD133 표적 sgRNA-Cas9 벡터와 CD133 내생 프로모터-RFP 벡터를 각각 2 ㎍와 5 ㎍씩 넣어주었다. 형질전환 이후 세포가 90% 정도 배양되면 배양액 10 ㎖에 퓨로마이신(#A11138-03, gibco) 2 ㎕를 넣어서 퓨로마이신 선별을 진행하였다. 퓨로마이신 선별 14일 후 세포가 70% 내지 80% 정도 배양되었을 때, FACS를 이용해서 RFP만 발현하는 세포만 2차로 선별하였다.

[0073] 그 결과, 도 15에 나타난 바와 같이, HT-29 세포에 CD133 표적 sgRNA-Cas9 벡터와 CD133 내생 프로모터-RFP 벡터를 도입시킨 후 RFP가 발현되는 것을 확인하였다. 또한, 퓨로마이신으로 선별한 후 14일이 지나자 암줄기세포의 대표적인 특징인 구를 형성하는 것(sphere formation)이 관찰되었다. 형광현미경을 통해 구를 확인한 결과, 암줄기세포는 RFP가 강하게 발현을 하지만, 주변의 암세포는 형광이 거의 보이지 않는 것을 확인하였다.

[0074] 또한, HT-29, NCCIT 및 SH-SY5Y 암세포 모두 구형을 형성하려고 하는 부분에서 RFP의 형광이 증가하는 것을 확인하였고, CD133 부분적 프로모터-RFP 벡터를 사용한 경우보다 CD133 내생 프로모터-RFP 벡터를 사용한 쪽에서 형광 발현이 더 잘되는 것을 확인하였다(도 16). 나아가, SNU-16 암세포 모두 구형을 형성하려고 하는 부분에서 RFP의 형광이 증가하는 것을 확인하였고, CD133 부분적 프로모터-RFP 벡터를 사용한 경우보다 CD133 내생 프로모터-RFP 벡터를 사용한 쪽에서 형광 발현이 더 잘되는 것을 확인하였다(데이터 미도시). 이를 통해, CD133 표적 sgRNA-Cas9 벡터 및 CD133 내생 프로모터-RFP 벡터를 사용한 리포터 유전자의 발현이 CD133 모니터링에 효과적임을 확인하였다. 다시 말해, CD133 유전자 뒤에 RFP가 잘 삽입되었으며, 암줄기세포의 마커로서 제 기능을 하고 있는 것을 확인하였다.

[0075] **실시예 6. 암세포 형질전환 확인**

[0076] 실시예 5에서 삽입한 벡터가 정확히 삽입 되었는지 확인하기 위해서, HT-29 CD133-RFP Knock-in 세포에서 genomic DNA를 추출하여 integration PCR 및 genotyping PCR을 진행하였고, HCT-116 CD133-RFP Knock-in 세포

에서 genomic DNA를 추출하여 integration PCR을 진행하였다.

[0077] integration PCR은 제조사의 매뉴얼에 따라 DNeasy Blood & Tissue kit(#69506, Qiagen)를 사용해서 세포에서 genomic DNA를 추출하였다. HT-29 CD133-RFP Knock-in 세포에서 추출한 genomic DNA는 KOD OneTM PCR Master Mix(#KMM-201, TOYOBO)를 이용해서 상기 표 3의 조건으로 integration PCR을 수행하였고, HCT-116 CD133-RFP Knock-in 세포에서 추출한 genomic DNA는 ExPrime Taq Premix(#G-5000, Genetbio)를 이용해서 95℃ 5분 초기 변성, 95℃ 30초, 62℃ 30초, 72℃ 1분 35cycle, 72℃ 5분 최종 신장 3단계로 integration PCR을 수행하였다. 이때, integration PCR에 사용한 프라이머를 표 4에 나타내었다.

표 4

[0078]

CD133	F	GGGCCAACTTCTCTCTGTTG	서열번호 11
	R	ACTCGAGAGGTCCAGGGTTC	서열번호 12

[0079] HT-29 CD133-RFP Knock-in 세포 관련 integration PCR를 진행한 샘플은 Agarose, LE, Analytical Grade(#V3125, Promega)와 0.5X TAE(50X TAE, #TR2002-100-00, Biosesang 3차수를 이용해 0.5X로 희석)를 이용해 제작한 0.8% 아가로스 겔에 120V로 40분 동안 전기영동하였다. 또한, HCT-116 CD133-RFP Knock-in 세포 관련 integration PCR를 진행한 샘플은 Agarose, LE, Analytical Grade(#V3125, Promega)와 0.5X TAE(50X TAE, #TR2002-100-00, Biosesang 3차수를 이용해 0.5X로 희석)를 이용해 제작한 0.8% 아가로스 겔에 90V로 60분 동안 전기영동하였다. 그 후, 샘플의 표적 사이즈의 밴드는 Geneclean® turbo kit(#111102-600, MP)를 사용해 DNA를 추출하였다. 추출된 DNA는 마크로젠사의 sequencing service(https://dna.macrogen.com/)를 통해 분석하였다.

[0080] 그 결과, HT-29 CD133-RFP Knock-in 세포의 CD133 유전자의 종결 코돈이 제거되고 RFP가 삽입되었으며, 프레임 쉬프트(frame shift)도 일어나지 않았음이 확인되었으며(도 17), HCT-116 CD133-RFP Knock-in 세포의 CD133 유전자의 종결 코돈이 제거되고 RFP가 삽입된 점이 확인되었다(도 18).

[0081] 또한, HT-29 CD133-RFP Knock-in 세포를 대상으로 한 genotyping PCR은 제조사의 매뉴얼에 따라 DNeasy Blood & Tissue kit(#69506, Qiagen)를 사용해서 세포에서 genomic DNA를 추출하였다. 추출한 genomic DNA와 ExPrime Taq Premix(2X, 8-strip)(#G-6000, GeNet Bio)를 이용해서 하기 표 5의 조건으로 genotyping PCR을 진행했다.

표 5

[0082]

온도	**시간**	
98℃	10분 - 초기 변성	
95℃	30초	35 사이클
62℃	30초	
72℃	30초	
72℃	5분 - 최종 신장	

[0083] genotyping PCR을 진행한 샘플은 Agarose, LE, Analytical Grade(#V3125, Promega)와 0.5X TAE(50X TAE; TR2002-100-00, Biosesang 3차수를 이용해 0.5X로 희석)를 이용해 제작한 0.8% 아가조스 겔에 120V 조건으로 40분 동안 전기영동하였다. 그 후, 샘플의 밴드 크기를 비교해서 야생형(Wild-type), 헤테로(Hetero), 호모(Homo) 타입의 세포를 구분하였다. Genomic DNA의 증폭을 위한 PCR primer는 Primer3 소프트웨어(https://bioinfo.ut.ee/primer3-0.4.0/)를 사용해서 in silico 방법으로 설계하였으며, 하기 표 6에 나타내었다.

표 6

[0084]

hCD133 WT/Insert	F	GCTGTGTTGCTGCTGGTCTA	서열번호 13
	R1	GGTACAGAGGGTGGACTGGA	서열번호 14
	R2	CCAGCCAATTTCAAGAGAGC	서열번호 15

[0085] 그 결과, 도 19에 나타난 바와 같이, HT-29 CD133-RFP Knock-in 세포의 CD133 유전자의 종결 코돈이 제거되고

RFP가 삽입되었으며, 프레임 쉬프트(frame shift)도 일어나지 않았음을 확인하였다.

[0086] **실시예 7. 암줄기세포의 구형 형성 배양**

[0087] 상기 실시예 5에서 형질전환시킨 HT-29 CD133-RFP Knock-in 세포 및 HCT-116 CD133-RFP Knock-in 세포를 37℃, 가습된 5% CO_2 인큐베이터에서 10% FBS(#16000, Gibco) 및 1% 페니실린/스트렙토마이신(Gibco)이 첨가된 DMEM(Gibco)에서 배양하였다. 구를 형성하는 암줄기세포를 증가시키기 위해 일반 배지로 배양된 세포를 2×10^5개로 계대 배양하였고, 혈청이 첨가되지 않은 DMEM 배지에 EGF(20 ng/㎖), bFGF(10 ng/㎖), 인슐린(5 ㎍/㎖), 1X B27 supplement를 첨가한 무혈청 배지를 사용하여 배양하였다. HT-29 CD133-RFP Knock-in 세포를 일반 배지와 무혈청 배지에서 10일 동안 배양하면서 확인하였다. 또한, HCT-116 CD133-RFP Knock-in 세포를 일반 배지와 무혈청 배지에서 7일 동안 배양하면서 확인하였고, NCCIT CD133-RFP Knock-in 세포 및 SNU-16 CD133-RFP Knock-in 세포를 무혈청 배지를 사용하여 7일간 배양하였다.

[0088] 그 결과, 도 20(가)에 나타난 바와 같이, 일반배지에서 배양시킨 HT-29 CD133-RFP Knock-in 세포와 비교하여 무혈청 배지에서 배양시킨 HT-29 CD133-RFP Knock-in 세포의 구 형성 및 RFP 발현이 증가하는 것을 확인하였다.

[0089] 또한, 다른 암세포주에 형질전환 시킨 NCCIT CD133-RFP Knock-in 세포와 SNU-16 CD133-RFP Knock-in 세포에서 위의 무혈청 배지를 사용하여 7일간 배양하였을 때, 도 20(나)에 나타난 바와 같이, 구형형성 및 RFP 발현의 증가를 확인하였다.

[0090] 또한, 도 21에 나타난 바와 같이, 일반배지에서 배양시킨 HCT-116 CD133-RFP Knock-in 세포와 비교하여 무혈청 배지에서 배양시킨 HCT-116 CD133-RFP Knock-in 세포의 구 형성 및 RFP 발현이 증가하는 것을 확인하였다.

[0091] **실시예 8. Real-time qPCR을 이용한 CD133 발현 암줄기세포 분리능 비교**

[0092] **실시예 8-1. HT-29 CD133-RFP Knock-in 세포의 암줄기세포 분리능**

[0093] HT-29 세포 및 HT-29 CD133-RFP Knock-in 세포를 RFP 및 CD133 항체를 이용한 분리를 통해 암줄기세포 분리능을 비교하였다.

[0094] 구체적으로, RFP가 표지된 세포의 분리는 유세포분석기를 사용하였으며, 무혈청 배지로 배양한 HT-29 CD133-RFP Knock-in 세포를 0.25% Trypsin/EDTA를 통해 떼어내고 1700 rpm 조건에서 3분 동안 원심분리한 후 상층액을 제거하였다. 그 후, 1 ㎖ 내지 4 ㎖의 FACS sorting buffer를 사용하여 세포를 재현탁하고 유세포분석기를 통해 분리하여 RFP가 표지된 세포만을 분리하였다.

[0095] 또한, CD133 항체 및 유세포분석기를 이용한 CD133을 발현하는 세포의 분리는, 항-CD133-PE 항체(#130-080-801, Miltenyi biotech)을 사용하여 수행하였다. 먼저, 무혈청 배지에서 배양한 HT-29 세포를 0.25% Trypsin/EDTA를 사용해 떼어내고 1700 rpm 조건에서 3분 동안 원심분리하고 상층액을 제거하였다. 그 후, FACS sorting buffer 1 ㎖을 넣어 재현탁하고, 항-CD133-PE 항체를 20 ㎕/㎖의 농도로 넣어서 잘 섞어준 뒤 20분 동안 4도에서 반응시켰다. FACS sorting Buffer 2 mL을 추가로 넣어 섞어준 뒤, 1700 rpm, 3분 원심분리하고 상층액을 제거했다. 이후 유세포분석기를 통해 CD133을 발현하는 세포를 분리하였다.

[0096] 한편, CD133 항체 및 자성을 이용한 분리는 CD133 microbead(#130-097-049, Miltenyi biotech)를 사용하여 수행하였다. 구체적으로, 배양한 세포를 0.25% trypsin/EDTA를 사용해 떼어내고 300×g 조건에서 10분 동안 원심분리한 후 상층액을 제거하여 1×10^8개의 세포를 300 ㎕의 버퍼(pH7.2, 0.5% BSA, 2mM EDTA in PBS)를 사용해 재현탁하였다. 재현탁한 세포에 100 ㎕ FcR blocking reagent 및 100 ㎕ CD133 microbead를 추가하여 잘 섞어준 후, 2℃ 내지 8℃ 온도에서 30분 동안 반응시켰다. 그 후, 상기 버퍼 2 ㎖로 세척한 후 300×g 조건에서 10분 동안 원심분리한 후 상층액을 제거하여 500 ㎕ 버퍼로 재현탁하였다. MACS separator(#130-091-051, Miltenyi)에 LD column(#130-042-901, Miltenyi)을 설치하고, 버퍼 2 ㎖을 흘려주어 컬럼을 세척하였다. 재현탁한 세포를 컬럼에 로딩한 후 버퍼 1 ㎖씩 2회 흘려주었다. 이때, CD133을 발현하지 않는 세포는 분리된다. 그 후, 컬럼을 MACS separator로부터 분리하고 버퍼 3 ㎖을 넣고 플런저로 밀어내어 CD133을 발현하는 세포를 분리하였다(도 22).

[0097] 상기 분리된 세포들의 대장암 줄기세포 표지 인자(CD133, CD24, CD26, CD44, EpCAM, ALCAM, ALDH1A1)들을 확인하기 위해서 제조사의 매뉴얼에 따라 RNeasy plus mini kit(#74136, Qiagen)를 사용해서 세포에서 total RNA를 추출하였다. total RNA(1 ㎍)은 gDNA Eraser(# RR047A, Takara Bio Inc)가 있는 PrimeScript RT reagent kit를

사용해서 cDNA로 합성하였다.

[0098] RT-PCR을 이용해 유전자의 상대적 발현 수준을 확인하기 위해, TB Green Premix Ex TaqTM(#RR420A, TaKaRa) 및 Mx3000p qPCR System(Agilent Technologies)를 사용해 하기 표 7의 조건으로 수행 및 분석하였다.

표 7

[0099]

온도	시간	
95℃	10분 - 초기 변성	
95℃	20초	50사이클
60℃	20초	

[0100] 비교 분석을 위해 cDNA 발현 수준을 B2M으로 정량화한 후, fold-changes로 표현하였다. 이때, 야생형(Wild-type) 세포는 C△Ct(control delta Ct) 값에 사용하였다. S△Ct(Sample delta Ct) 값은 B2M의 Ct 값과 표적 유전자 사이의 차이로 계산하였다. 샘플과 대조군 사이의 상대적 유전자 발현 수준은 S△Ct-C△Ct(Sample delta Ct - control delta Ct) 공식을 사용하여 측정하였다. cDNA의 증폭을 위해 사용된 PCR 프라이머는 Primer3 소프트웨어(https://bioinfo.ut.ee/primer3-0.4.0/)를 사용해서 in silico 방법으로 설계하였으며 하기 표 8에 나타내었다.

표 8

[0101]

B2M	F	CTCCGTGGCCTTAGGTGTG	서열번호 16
	R	TCTCTGCTGGATGACGTGAG	서열번호 17
CD133	F	AAGCATTGGCATCTTCTATGG	서열번호 18
	R	TTCAAGAGAGTTCGCAAGTCC	서열번호 19
CD24	F	GATTTATTCCAGTGAAACAACAACTG	서열번호 20
	R	ACCACGAAGAGACTGGCTGT	서열번호 21
EpCAM	F	GCTCAGGAAGAATGTGTCTGTG	서열번호 22
	R	GCCAGCTTTGAGCAAATGAC	서열번호 23
CD26	F	CTCTGAACGCTCACTTCCGA	서열번호 24
	R	TTTGTTCAGCAGAACCACGG	서열번호 25
CD44	F	AGACTCCCATTCGACAACAGG	서열번호 26
	R	GGAACTGTCCTCTGGGCTTG	서열번호 27
ALCAM	F	ACTTGACGTACCTCAGAATCTCA	서열번호 28
	R	CATCGTCGTACTGCACACTTT	서열번호 29
ALDH1A1	F	GCACGCCAGACTTACCTGTC	서열번호 30
	R	CCTCCTCAGTTGCAGGATTAAAG	서열번호 31
c-MYC	F	GTCCTCGGATTCTCTGCTCTC	서열번호 32
	R	CATCTTCTTGTTCCTCCTCAGAGT	서열번호 33
OCT3/4	F	GCCCGAAAGAGAAAGCGAAC	서열번호 34
	R	AACCACACTCGGACCACATC	서열번호 35
NANOG	F	TGTGTTCTCTTCCACCCAGC	서열번호 36
	R	ACCAGGTCTTCACCTGTTTG	서열번호 37

[0102] HT-29 CD133-RFP Knock-in 세포주를 CD133 항체-microbead를 사용하여 CD133 발현 그룹으로 분리한 것(Miltenyi CD133+)과, 유세포분석기를 이용하여 HT-29 CD133-RFP Knock-in 세포에서 RFP 형광(FACS sorting, RFP+) 및 HT-29 세포에서 CD133 항체-PE(CD133 Ab+) 그룹으로 분리한 후 실시간 qPCR을 이용해 대장암 줄기세포 인자들(CD133, CD24, CD26, CD44, EpCAM, ALCAM, ALDHA1)과 CD133의 다운 스트림(downstream)으로 알려진 다분화능(c-MYC, OCT3/4, NANOG) 인자들을 비교한 결과, RFP 형광으로 분리한 그룹(FACS sorting, RFP+)이 CD133 항체로 분리한 그룹(Miltenyi CD133+ 및 CD133 Ab+)에 비해 대부분의 대장암 줄기세포 인자들 및 다분화능 인자들의 발현이 더 증가된 것을 확인하였다(도 23). 이를 통해, CD133 표적 sgRNA-Cas9 벡터 및 CD133 내생 프로모터-RFP 벡터를 사용한 CD133 넉인 형광 리포터 벡터 시스템이 기존의 시스템에 비해 암줄기세포에 대한 특이성이 우수한 것을 확인하였다.

[0103] **실시예 8-2. HCT-116 CD133-RFP Knock-in 세포의 암줄기세포 분리능**

[0104] 구체적으로, RFP가 표지된 세포의 분리는 FACS를 사용하였으며, 무혈청 배지로 10일간 배양한 HCT-116 CD133-RFP Knock-in 세포를 0.25% Trypsin/EDTA를 통해 떼어내고 1700 rpm 조건에서 3분 동안 원심분리한 후 상층액을 제거하였다. 그 후, 1 ㎖ 내지 4 ㎖의 FACS sorting buffer를 사용하여 세포를 재현탁하고 FACS를 통해 분리하여 RFP가 표지된 세포만을 분리하였다(도 24). 이때, RFP 형광 발현이 약한 그룹(FACS sorting, RFPlow) 및 RFP 형광 발현이 강한 그룹(FACS sorting, RFPhigh)으로 구분되었다. HCT-116 CD133-RFP Knock-in 세포를 대상으로 실시간 qPCR을 이용해 P2A-RFP(Knock-in 부분의 리포터 서열) mRNA 발현량을 비교한 결과, 각 그룹별로 분리가 제대로 되었음을 확인하였다(도 25).

[0105] 나아가, endogenous promoter를 사용한 CD133 리포터 시스템의 암줄기세포 분리 특이성 여부를 확인하기 위하여, RFPlow 및 RFPhigh 그룹을 대상으로, 대장암 줄기세포 인자들(CD133, CD24, CD26, CD44, EpCAM, ALCAM, ALDHA1)과 stemness 관련 전사인자로서 CD133의 downstream으로 알려진 다분화능 인자들(c-MYC, OCT3/4, SOX2, NANOG) 및 분화된 세포 관련 인자들(SLC26A3, CEACAM1, CD177, MUC2, TFF3)의 발현량을 비교분석하였다.

[0106] 상기 분리된 세포들의 대장암 줄기세포 표지 인자들(CD133, CD24, CD26, CD44, EpCAM, ALCAM, ALDH1A1), 전사인자이자 다분화능 인자들(c-MYC, OCT3/4, SOX2, NANOG) 및 분화된 세포 인자들(SLC26A3, CEACAM1, CD177, MUC2, TFF3)의 상대적 발현수준을 확인하기 위해서 제조사의 프로토콜에 따라 RNeasy plus mini kit(#74136, Qiagen)를 사용해서 세포에서 total RNA를 추출하였다. Total RNA(1㎍)은 gDNA Eraser(# RR047A, Takara Bio Inc)가 있는 PrimeScript RT reagent kit를 사용해서 cDNA로 합성하였다. 유전자의 상대적 발현 수준 분석을 위하여 실시간 PCR을 하였으며, TB Green Premix Ex TaqTM(#RR420A, TaKaRa) 및 Mx3000p qPCR System(Agilent Technologies)를 사용해 분석하였다. qRT-PCR을 위한 반응 매개 변수로서 95℃ 10분 초기 변성, 95℃ 20초, 60℃ 20초, 50cycle, melting curve 3단계로 수행하였다.

[0107] 비교 분석을 위해 cDNA 발현 수준을 B2M으로 정량화한 다음 fold-changes로 표현하였다. Wild-type 세포는 control delta Ct(C△Ct)값에 사용되었다. Sample delta Ct(S△Ct)값은 B2M의 Ct 값과 표적 유전자 사이의 차이로 계산하였다. Sample과 control 사이의 상대적 유전자 발현 수준은 Sample delta Ct - control delta Ct(S△Ct-C△Ct) 공식을 사용해서 측정하였다. cDNA의 증폭을 위한 PCR primer는 Primer3 소프트웨어(https://bioinfo.ut.ee/primer3-0.4.0/)를 사용해서 in silico 방법으로 설계하였다.

[0108] 또한, B2M, CD133, CD24, EpCAM, CD26, CD44, ALCAM, ALDHA1, c-MYC, OCT3/4 및 NANOG cDNA의 증폭을 위한 PCR primer는 상기 표 8의 프라이머를 사용하였고, SOX2, SLC26A3, CEACAM1, CD177, MUC2 및 TFF3 cDNA의 증폭을 위한 PCR primer는 하기 표 9의 프라이머를 사용하였다.

표 9

[0109]

SOX2	F	TACAGCATGATGCAGGACCA	서열번호 57
	R	CGAGCTGGTCATGGAGTTGTA	서열번호 58
SLC26A3	F	ACACATATCCGTGGGTCCGT	서열번호 59
	R	GTTGCATTGCGATCTGGGAC	서열번호 60
CEACAM1	F	GCTCTGATAGCAGTAGCCCTG	서열번호 61
	R	ATTGGAGTGGTCCTGAGCTGC	서열번호 62
CD177	F	CATCTCTCAGGAGGTGGGCT	서열번호 63
	R	CCCGATTTGTCTGGTGTGGT	서열번호 64
MUC2	F	GTCACCACGGAGGTTTCGTA	서열번호 65
	R	CCTTGGCCGAGTACATGACA	서열번호 66
TFF3	F	GTACGTGGGCCTGTCTGCAA	서열번호 67
	R	TGAAACACCAAGGCACTCCA	서열번호 68

[0110] RFP 형광이 강한 그룹(RFPhigh)이 RFP 형광이 약한 그룹(RFPlow)에 비해 대부분의 대장암 줄기세포 인자들 및 다분화능 인자들의 발현이 더 증가되었고 분화된 세포 관련 인자들은 감소되는 것으로 나타났다(도 26). 종합하면, endogenous promoter를 사용한 CD133 리포터 시스템이 다른 대장암 세포주인 HCT-116에서도 암줄기세포에 대한 특이성이 있음이 확인되었다.

[0111] **실시예 9. CD133 발현 암줄기세포의 항암 후보 약물 스크리닝 능력 평가**

[0112] 상기 HT-29 CD133-RFP Knock-in 세포주에 5-fluorouracil(5-FU) 항암제 및/또는 항산화 효소인

Sulfiredoxin(Srx) 억제제인 frugoside를 처리하였을 때 세포의 콜로니 형성이 감소하는지 여부를 확인하기 위하여 soft agar colony-forming assay를 실시하였고, 세포의 스피어(sphere) 형성 감소 여부를 확인하기 위하여 sphere-formation assay를 실시하였다.

[0113] 구체적으로, soft-agar colony forming assay를 위하여 세포는 0.3% agar를 함유한 세포 성장배지에 현탁하였다. 이후, 대조군으로서 성장배지, 50 μM의 5-FU를 추가로 함유한 성장배지, 1 μM의 frugoside를 추가로 함유한 성장배지 또는 50 μM의 5-FU 및 1 μM의 frugoside를 추가로 함유한 성장배지에 HT-29 CD133-RFP Knock-in 세포주를 0.6% agar 층에 플레이팅하였다. 15일간 플레이팅된 세포를 성장시킨 뒤 10분간 0.01% 크리스탈 바이올렛(Sigma)로 염색한 뒤 콜로니 수를 측정하였다. 또한, 형광 현미경으로 각 배지에서의 10^4 세포를 갖는 콜로니 수를 측정하였다.

[0114] 한편, sphere formation assay를 위하여 세포를 ultra-low-attachment 6 well plate에 각 웰당 10^4 세포를 접종하였다. 배지로서 4mg/ml 소혈청 알부민, 10ng/ml의 basic fibroblast growth factor, 20ng/ml epidermal growth factor, 100μg/ml apotransferrin 및 25μg/ml의 인슐린이 첨가된 무혈청 DMEM/F12 배지에서 1주일간 배양한 뒤 형성된 sphere 수를 계수하였다.

[0115] 분석결과, 5-FU와 frugoside가 동시에 처리된 HT-29 CD133-RFP-Knock-in 세포에서 형성된 스피어 및 콜로니 수가 5-FU 또는 frugoside가 단독 처리된 HT-29 CD133-RFP-Knock-in 세포에서의 스피어 및 콜로니 수보다 적게 나타났으며, RFP 형광도 감소되는 것으로 나타났다(도 27 및 도 28). 따라서, CD133 발현 암 줄기세포에 항암 후보 약물을 처리하여 RFP 발현 변화 수준이 감소하면 해당 후보 약물이 항암효과가 있음을 스크리닝하는데 사용될 수 있다는 점이 확인되었다.

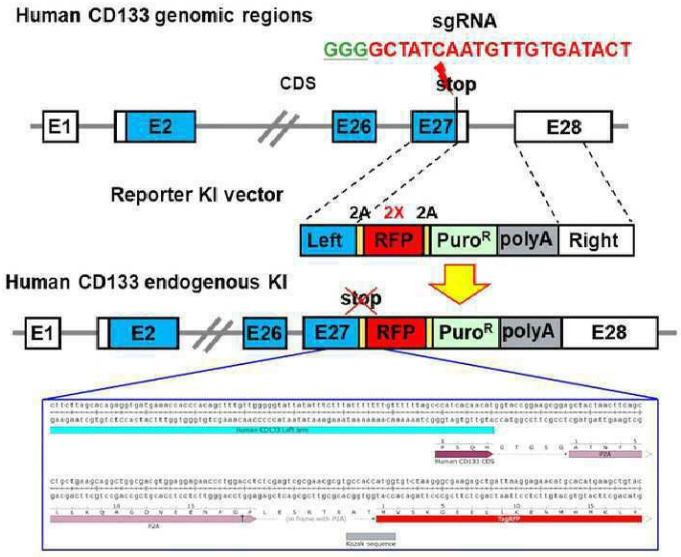
[0116] **실시예 10. CD133 발현 암줄기세포를 이용한 항암 후보 약물 스크리닝 또는 약물 내성 평가**

[0117] 본원발명의 벡터세트로 형질전환된 세포(예컨대, HT-29 CD133-RFP Knock-in 세포주)에 후보 약물을 처리하여 RFP의 발현 변화 수준을 측정하였을 때, RFP의 발현 수준이 감소하는 경우 해당 후보 약물이 항암 효과(특히, 암줄기세포에 대한 억제효과)가 있는 약물로 스크리닝할 수 있다.

[0118] 또한, 본원발명의 벡터세트로 형질전환된 세포에 기존의 항암 치료 약물을 처리하여 RFP의 발현 변화 수준을 측정하였을 때, RFP의 발현 수준이 증가 또는 유의미한 변화가 없는 경우 해당 항암 치료 약물은 약물 내성이 생길 수 있다는 것을 확인할 수 있다.

도면

도면1



도면2

Sequence: TCAGCTATCAATGTTGTGATGGG

Efficacy score by CRISPRater: 0.63 MEDIUM

Oligo pair with 5' extension: fwd: TAggTCAGCTATCAATGTTGTGAT rev: AAACATCACAACATTGATAGCTGA

Oligo pair with 5' substitution: fwd: TAggAGCTATCAATGTTGTGAT rev: AAACATCACAACATTGATAGCT

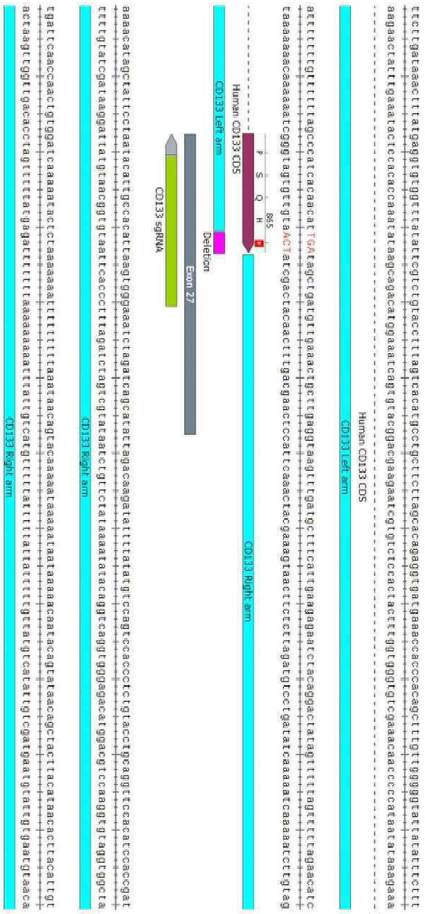
Legend for off-target site position: E = exonic; I = intronic; x = intergenic

Legend for the CRISPRater score: LOW efficacy (score<=0.58) MEDIUM efficacy (0.58<=score<=0.74) HIGH efficacy (score>0.74)

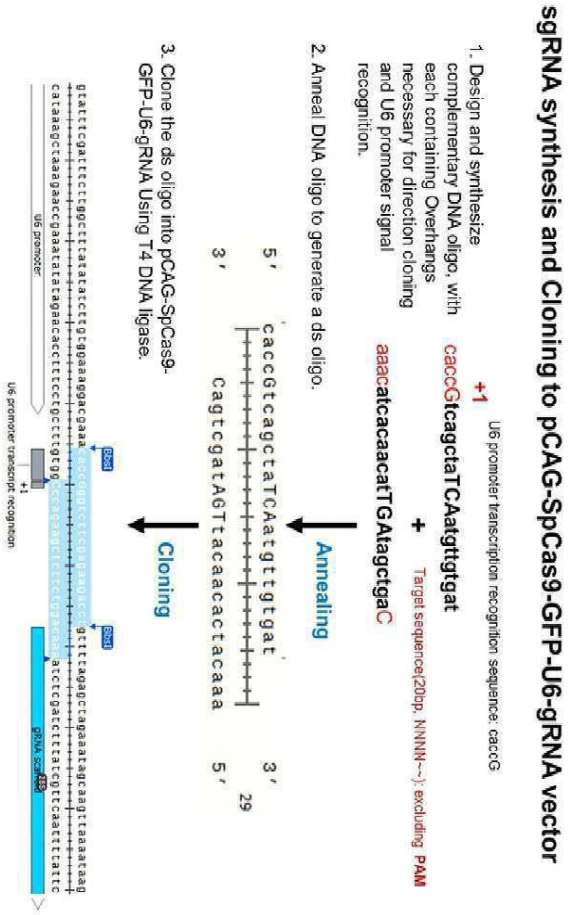
Top 20 offtarget sites out of 77 (including on target; for full list see xls file)

Coordinates	strand	MM	target_seq	PAM	distance		gene name	gene id
chr4:159710060-159710082	+	0	TCAGCTAT[CAATGTTGTGAT]	GGG	0	E	PROM1	ENSG00000007062
chr1:75991151-75991173	-	3	AAAGCTAA[CAATGTTGTGAT]	TGG	6242	I	RP11-550H2.1	ENSG00000225605
chr14:62295126-62295148	-	4	AGAGGTAT[CCATGTTGTGAT]	AGG	NA	-	NA	NA
chr6:25277426-25277448	+	4	TCATTTTT[CAGTGTTGTGAT]	AGG	16630	-	LRRC16A	ENSG00000079691
chr5:175200755-175200777	+	4	GCAGAAAT[CAAAGTTGTGAT]	TGG	0	E	ARL2BP05	ENSG00000248355
chr2:170251267-170251289	+	4	TCAGGACC[CAGTGTTGTGAT]	TGG	15131	I	MYO3B	ENSG00000071909
chr12:85516933-85516955	+	4	TCTTGATT[AATTGTTGTGAT]	GGG	50923	-	RP11-629O19.1	ENSG00000257855
chr21:33689226-33689248	-	4	TCAGTGAT[CCTTGTTGTGAT]	TGG	859	I	SYNJ1	ENSG00000159082
chrX:5455895-5455917	+	4	TCAACCAT[GAAAGTTGTGAT]	GGG	45644	-	SNORA20	ENSG00000201660
chr2:161038791-161038813	+	3	TCAGCTGT[GAATCTTGTGAT]	GGG	2950	-	AC142119.1	ENSG00000236989
chr13:51607745-51607767	+	4	TCAGAAAT[TACTGTTGTGAT]	AGG	8285	I	ATP5EJP1	ENSG00000224451
chr1:37169608-37169630	-	4	TCTGTCCT[CAATGCTGTGAT]	TGG	7974	-	MIR4255	ENSG00000264698
chr15:38606483-38606505	-	4	TCAACCAT[GAATCCTTGTGAT]	GGG	40908	-	RASGRP1	ENSG00000172575
chr3:178536330-178536352	-	4	TCTTCTAG[CAATGCTGTGAT]	AGG	0	E	KCNMB2	ENSG00000197584
chr18:72533459-72533481	-	4	ATAGCTAT[CCATGATGTGAT]	TGG	3199	-	CBLN2	ENSG00000141668
chrY:11289078-11289050	+	4	ACAGATAT[CATTCTTGTGAT]	TGG	16968	-	DUX4L16	ENSG00000258567
chr3:7586523-7586545	+	4	ACAGATAT[CATTCTTGTGAT]	TGG	41436	-	ROBO2	ENSG00000185008
chr21:9250025-9250047	-	4	ACAGATAT[CATTCTTGTGAT]	TGG	74966	-	KP-21264C1.2	ENSG00000278932
chr20:18907090-18907112	+	3	GCAGTTAT[CAATGTTTTGAT]	GGG	36	I	PPP1R16B	ENSG00000101445
chr4:94603868-94603890	-	4	TAATCTAT[CAGTGTGTGTGAT]	GGG	4169	I	PDLIM5	ENSG00000131110

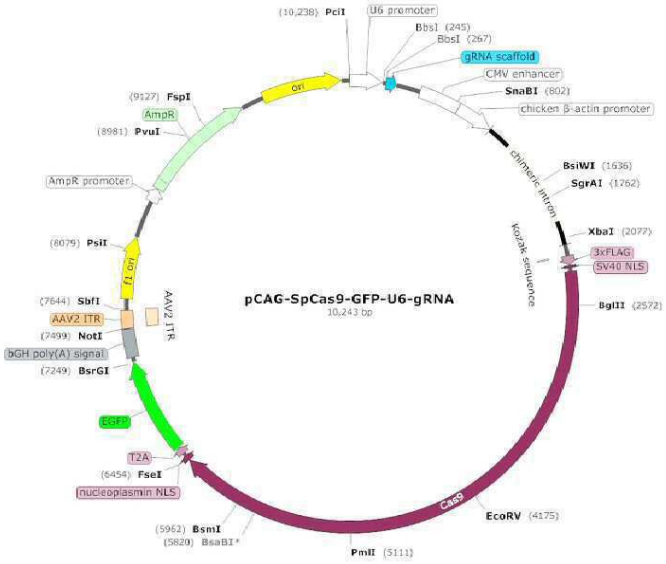
도면3


CD133 Right arm
CD133 sgRNA
CD133 Left arm
Deletion
Exon 27
Human CD133 CDS
P
S
Q
H
Y
865
CD133 Right arm
Human CD133 CDS
CD133 Left arm

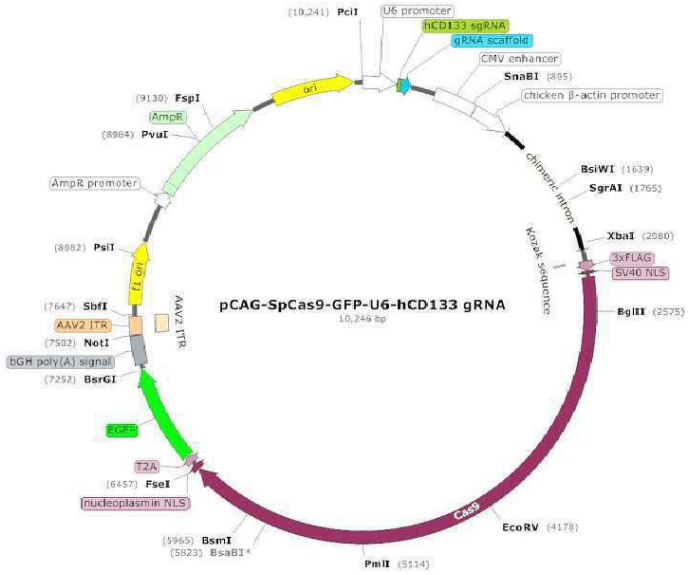
도면4


sgRNA synthesis and Cloning to pCAG-SpCas9-GFP-U6-gRNA vector
1. Design and synthesize complementary DNA oligo, with each containing Overhangs necessary for direction cloning and U6 promoter signal recognition.
+1
U6 promoter transcription recognition sequence: caccG
caccGtcagctaTCAatgttgtgat
+
Target sequence(20bp, NNNNN~): excluding PAM
aaacatcacaacatTGAtagctgaC
2. Anneal DNA oligo to generate a ds oligo.
Annealing
5′ caccGtcagctaTCAatgttgtgat 3′
3′ CagtcgatAGTtacaacactacaaa 5′
29
3. Clone the ds oligo into pCAG-SpCas9-GFP-U6-gRNA Using T4 DNA ligase.
Cloning
U6 promoter
U6 promoter transcript: recognition
+1
BbsI
BbsI
gRNA scaffold

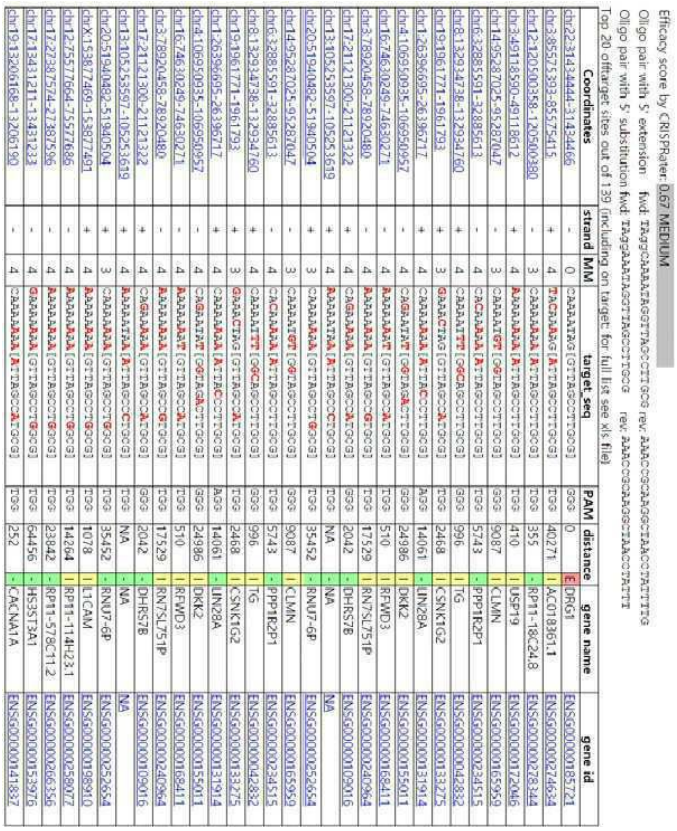
도면5


pCAG-SpCas9-GFP-U6-gRNA
10,243 bp
(10,238) PciI
U6 promoter
BbsI (245)
BbsI (267)
gRNA scaffold
CMV enhancer
SnaBI (802)
chicken β-actin promoter
chimeric intron
BsiWI (1636)
SgrAI (1762)
Kozak sequence
XbaI (2077)
3xFLAG
SV40 NLS
BglII (2572)
EcoRV (4175)
Cas9
PmlI (5111)
(5962) BsmI
(5820) BsaBI*
nucleoplasmin NLS
(6454) FseI
T2A
EGFP
(7249) BsrGI
bGH poly(A) signal
(7499) NotI
AAV2 ITR
(7644) SbfI
AAV2 ITR
U6 T
(8079) PsiI
AmpR promoter
(8981) PvuI
AmpR
(9127) FspI
Ori

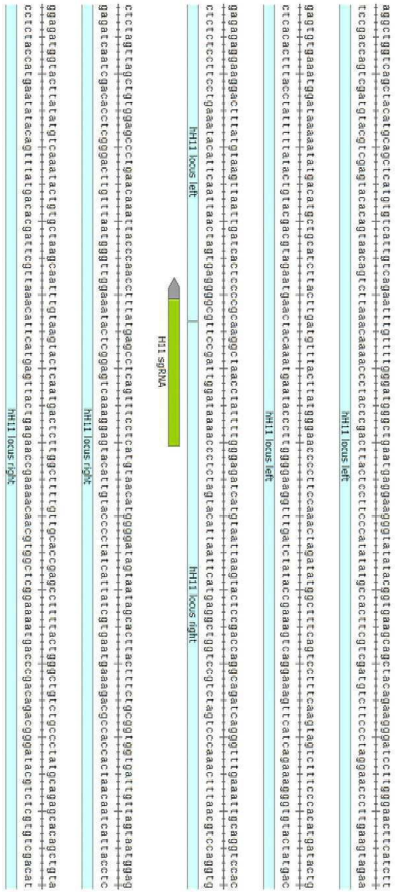
도면6


pCAG-SpCas9-GFP-U6-hCD133 gRNA
10,246 bp
U6 promoter
hCD133 sgRNA
gRNA scaffold
CMV enhancer
chicken β-actin promoter
chimeric intron
Kozak sequence
3xFLAG
SV40 NLS
ori
AmpR
AmpR promoter
ori f1
AAV2 ITR
bGH poly(A) signal
EGFP
T2A
nucleoplasmin NLS
Cas9
(10,241) PciI
SnaBI (885)
BsiWI (1639)
SgrAI (1765)
XbaI (2080)
BglII (2575)
EcoRV (4178)
PmlI (5114)
(5823) BsaBI*
(5965) BsmI
(6457) FseI
(7252) BsrGI
(7502) NotI
(7647) SbfI
(8082) PsiI
(8984) PvuI
(9130) FspI

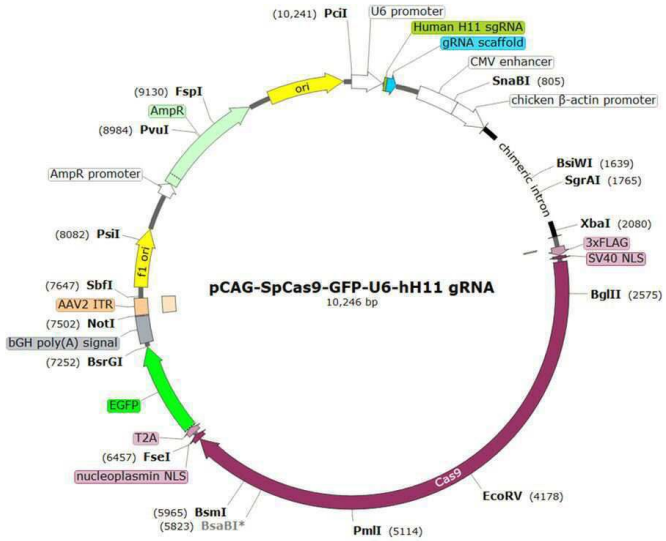
도면7


Efficacy score by CRISPRater: 0.67 MEDIUM
Oligo pair with 5' extension fwd: TAggCAAAATAGGTTAGCCTTGCG rev: AAACCGCAAGGCTAACCTATTTTG
Oligo pair with 5' substitution fwd: TAggAAAATAGGTTAGCCTTGCG rev: AAACCGCAAGGCTAACCTATTT
Top 20 offtarget sites out of 139 (including on target; for full list see xls file)
Coordinates
strand
MM
target seq
PAM
distance
gene name
gene id
chr22:31434444:31434466 - 0 CAAAATAGGTTAGCCTTGCGGGG 0 0 ERGI ENSG00000185721
chr3:85575393:85575415 + 4 TACCAAAGGTAATAGCCTTGCGG 0 40271 AC018361.1 ENSG00000274634
chr12:120500368:120500390 + 3 CAAAAAAGAAATAGCCTTGCGGG 0 355 RP11-18C24.8 ENSG00000228344
chr3:49118590:49118612 + 4 AAAAAAAGAAATAGCCTTGCGGG 0 410 USP19 ENSG00000172046
chr14:95287025:95287047 - 3 CAAAATGTTGATAGCCTTGCGGG 0 9087 CLMN ENSG00000165949
chr6:32885501:32885513 - 4 CACAAAAAATATAGCCTTGCGGG 0 5743 PPP1R2P1 ENSG00000234515
chr8:132934736:132934760 + 4 CAAAATTTTGAGAGCCTTGCGGG 0 996 TG ENSG00000042832
chr19:1961771:1961793 + 3 GAAACTAGGTATAGCCTTGCGGG 0 2468 CSNK1G2 ENSG00000133275
chr20:51940482:51940504 + 3 AAAAAAAAAAATAGCCTTGCAGG 0 NA NA NA
chr14:95287025:95287047 - 3 CAAAATGTTGATAGCCTTGCGGG 993 9087 CLMN ENSG00000165949
chr6:32885501:32885513 - 4 CACAAAAAATATAGCCTTGCGGG 993 5743 PPP1R2P1 ENSG00000234515
chr8:132934736:132934760 + 3 GAAACTAGGTATAGCCTTGCGGG 993 996 TG ENSG00000042832
chr19:1961771:1961793 + 4 CAAAATAGGTTAGCCTTGCGGGG 993 2468 CSNK1G2 ENSG00000133275
chr4:106950035:106950057 + 4 CAGAAATATGGGTAGCCTTGCGG 956 24986 DKK2 ENSG00000151011
chr16:74630249:74630271 - 4 AAAAAAAATGTTAGCCTTGCGGG 956 14061 LIN28A ENSG00000131914
chr3:78920458:78920480 - 4 AAAAAAAAAGTTAGCCTTGCGGG 956 17529 RN7SL751P ENSG00000240964
chr17:21121300:21123322 + 4 AAAAAAAAAGTTAGCCTTGCGGG 956 510 RFWD3 ENSG00000158411
chr13:105253597:105253619 + 4 AAAAAAAAAGTTAGCCTTGCGGG 956 2042 OIP5 ENSG00000109016
chr20:51940482:51940504 + 3 CAAAAAAAAAATAGCCTTGCAGG 956 35452 RNU7-6P NA NA
chrX:153877469:153877491 + 4 AAAAAAAAAGTTAGCCTTGCGGG 956 1078 L1CAM ENSG00000198910
chr12:75577664:75577686 - 4 AAAAAAAAAGTTAGCCTTGCGGG 956 14264 RP11-114H23.1 ENSG00000258072
chr12:27238724:27238746 - 4 AAAAAAAAAGTTAGCCTTGCGGG 956 23842 RP11-578C11.2 ENSG00000266356
chr17:13431211:13431233 - 4 GAAAAAAAAGTTAGCCTTGCGGG 956 64456 HS3ST3A1 ENSG00000153976
chr19:13206168:13206190 - 4 CAAAAAAAATTTAGCCTTGCAGG 956 252 CACNA1A ENSG00000141837

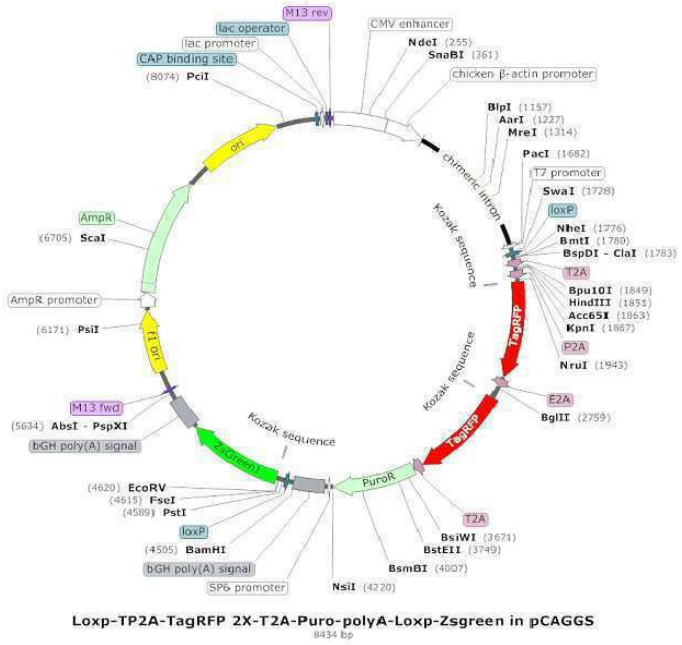
도면8



도면9


pCAG-SpCas9-GFP-U6-hH11 gRNA
10,246 bp
(10,241) PciI
U6 promoter
Human H11 sgRNA
gRNA scaffold
CMV enhancer
SnaBI (805)
chicken β-actin promoter
chimeric intron
BsiWI (1639)
SgrAI (1765)
XbaI (2080)
3xFLAG
SV40 NLS
BglII (2575)
EcoRV (4178)
PmlI (5114)
(5965) BsmI
(5823) BsaBI*
Cas9
nucleoplasmin NLS
(6457) FseI
T2A
EGFP
(7252) BsrGI
bGH poly(A) signal
(7502) NotI
AAV2 ITR
(7647) SbfI
f1 ori
(8082) PsiI
AmpR promoter
(8984) PvuI
AmpR
(9130) FspI
ori

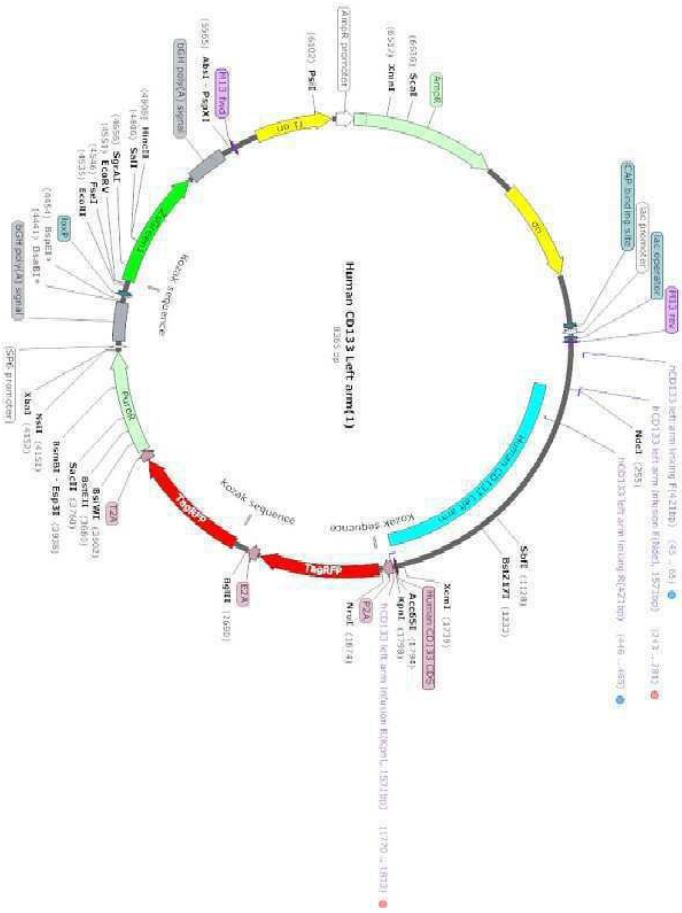
도면10



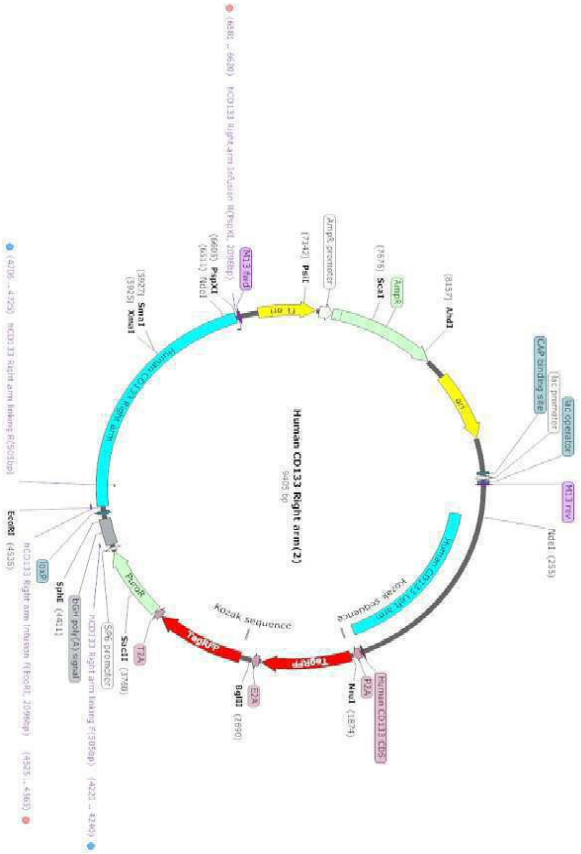
Loxp-TP2A-TagRFP 2X-T2A-Puro-polyA-Loxp-Zsgreen in pCAGGS

8434 bp

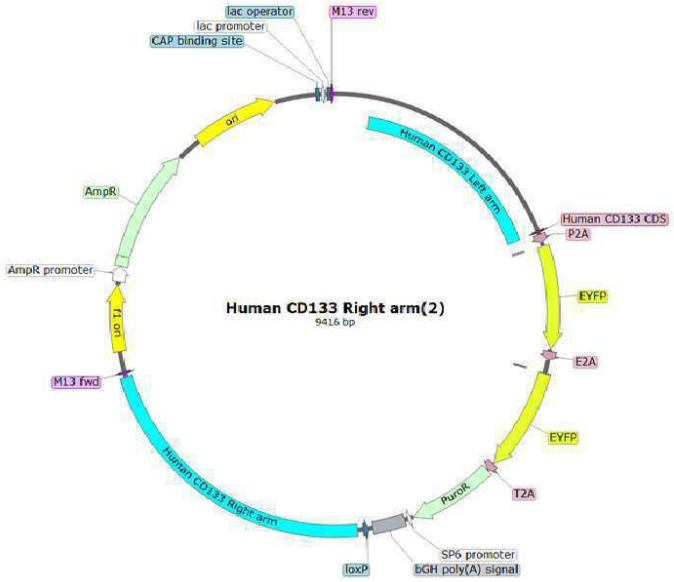
도면11



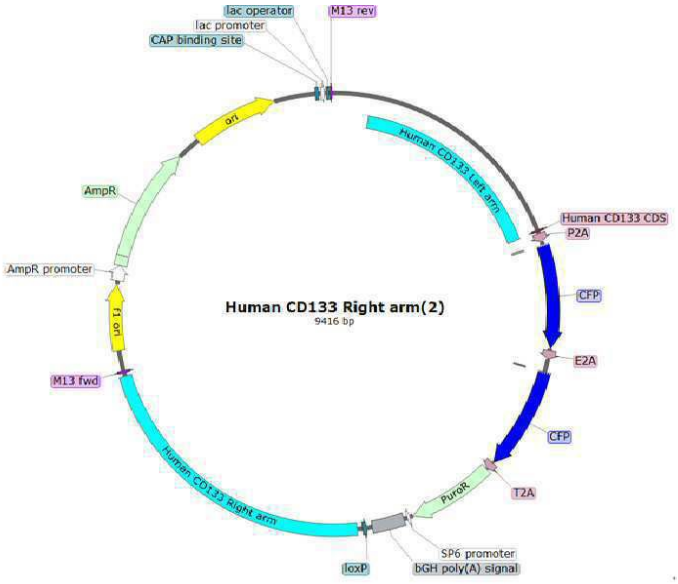
도면12a


Human CD133 Right arm(2)
9416 bp
M13 fwd
PspXI
PstI
AmpR promoter
ScaI
AmpR
AhdI
ori
CAP binding site
lac operator
lac promoter
M13 rev
PacI
PuroR
Kozak sequence
T2A
SrfI
E2A
PgiII
NcoI
P2A
Human CD133 CDS
SphI
EcoRI poly(A) signal
SP6 promoter
XmaI
SmaI

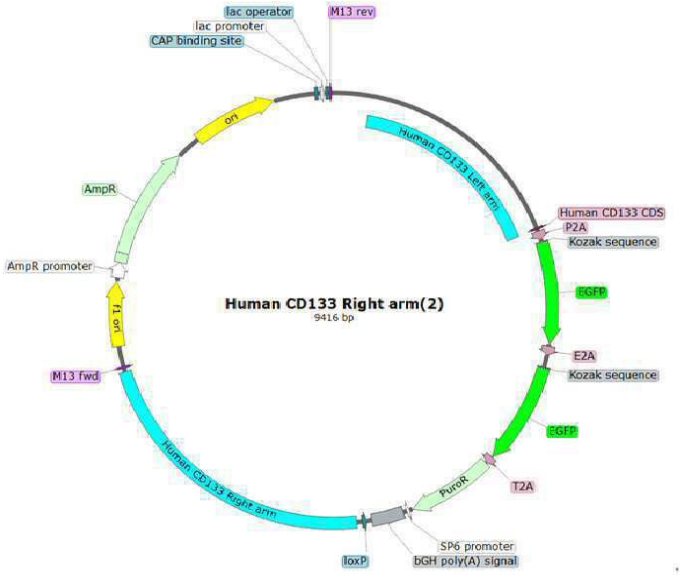
도면12b


lac operator
lac promoter
CAP binding site
M13 rev
ori
Human CD133 Left arm
AmpR
Human CD133 CDS
P2A
AmpR promoter
EYFP
Human CD133 Right arm(2)
9416 bp
ori
E2A
M13 fwd
EYFP
PuroR
T2A
Human CD133 Right arm
SP6 promoter
loxP
bGH poly(A) signal

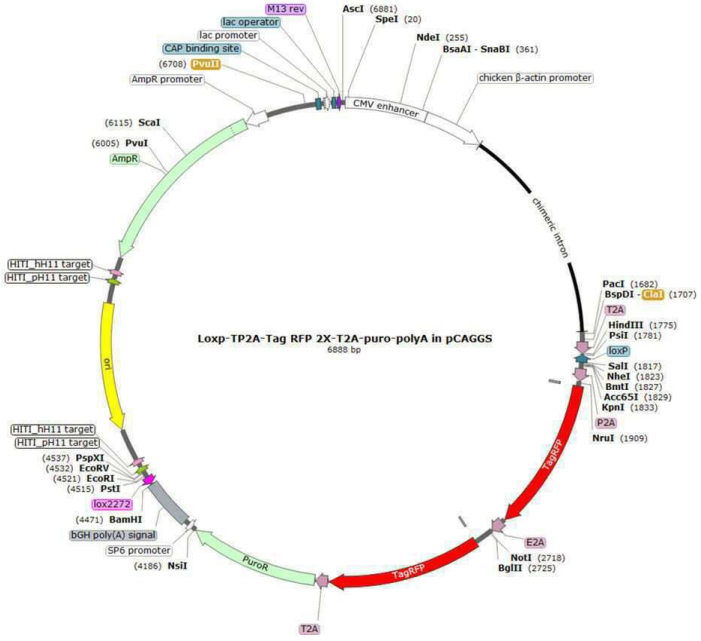
도면12c


lac operator
M13 rev
lac promoter
CAP binding site
ori
Human CD133 left arm
Human CD133 CDS
P2A
AmpR
CFP
AmpR promoter
Human CD133 Right arm(2)
9416 bp
E2A
f1 ori
CFP
M13 fwd
T2A
PuroR
Human CD133 Right arm
SP6 promoter
loxP
bGH poly(A) signal

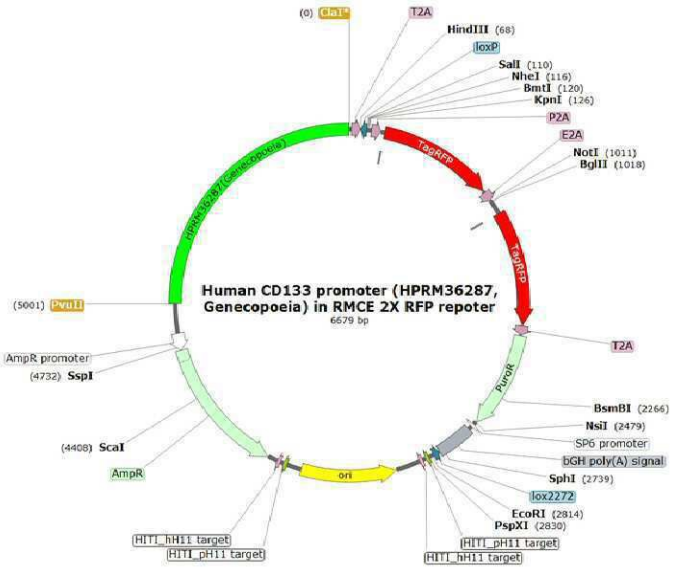
도면12d


lac operator
M13 rev
lac promoter
CAP binding site
ori
Human CD133 left arm
Human CD133 CDS
P2A
Kozak sequence
AmpR
AmpR promoter
EGFP
Human CD133 Right arm(2)
9416 bp
E2A
f1 ori
Kozak sequence
M13 fwd
EGFP
T2A
PuroR
Human CD133 Right arm
SP6 promoter
loxP
bGH poly(A) signal

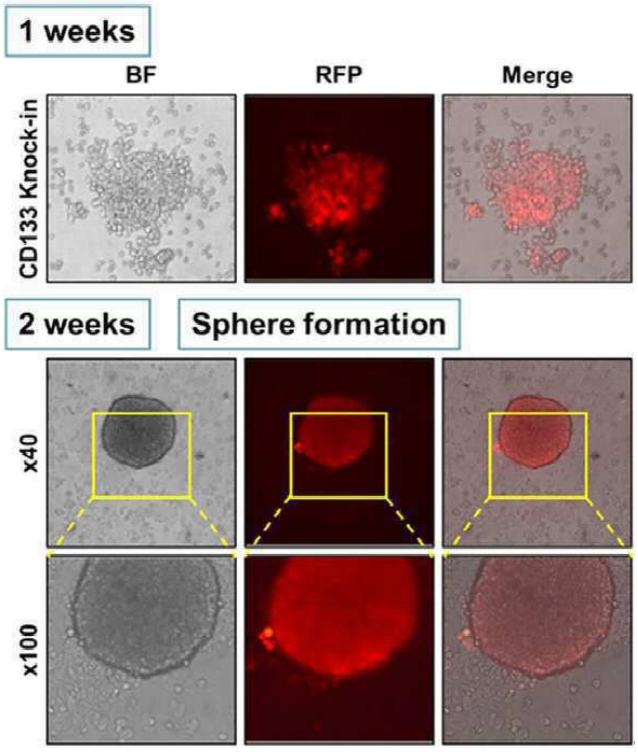
도면13


Loxp-TP2A-Tag RFP 2X-T2A-puro-polyA in pCAGGS
6888 bp
M13 rev
lac operator
lac promoter
CAP binding site
(6708) PvuII
AmpR promoter
AscI (6881)
SpeI (20)
NdeI (255)
BsaAI - SnaBI (361)
chicken β-actin promoter
CMV enhancer
chimeric intron
PacI (1682)
BspDI - ClaI (1707)
T2A
HindIII (1775)
PsiI (1781)
loxP
SalI (1817)
NheI (1823)
BmtI (1827)
Acc65I (1829)
KpnI (1833)
P2A
NruI (1909)
TagRFP
E2A
NotI (2718)
BglII (2725)
TagRFP
T2A
PuroR
(4186) NsiI
SP6 promoter
bGH poly(A) signal
(4471) BamHI
lox2272
(4515) PstI
(4521) EcoRI
(4532) EcoRV
(4537) PspXI
HITI_pH11 target
HITI_hH11 target
ori
(6005) PvuI
AmpR
(6115) ScaI
HITI_pH11 target
HITI_hH11 target

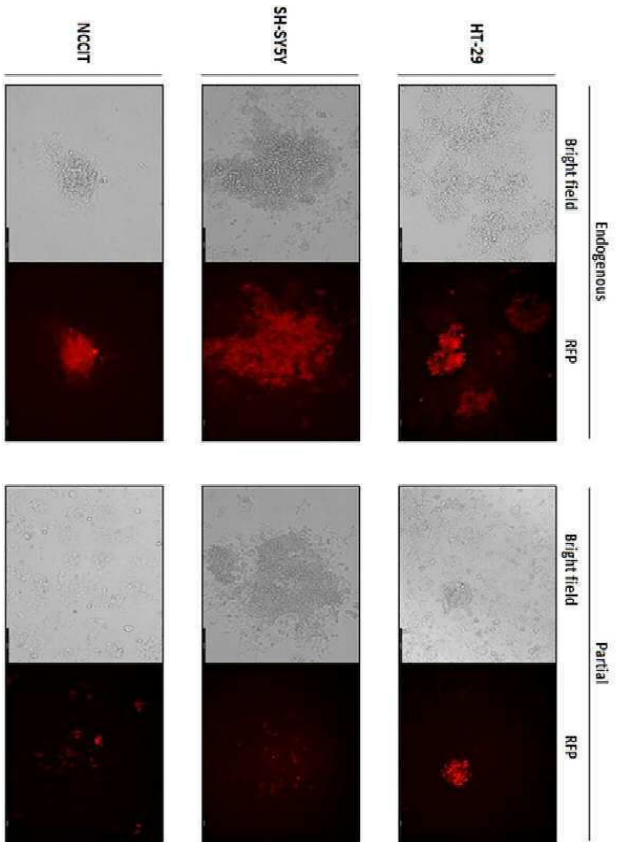
도면14


Human CD133 promoter (HPRM36287, Genecopoeia) in RMCE 2X RFP repoter
6679 bp
(0) ClaI*
T2A
HindIII (68)
loxP
SalI (110)
NheI (116)
BmtI (120)
KpnI (126)
P2A
E2A
NotI (1011)
BglII (1018)
TagRFP
TagRFP
T2A
PuroR
BsmBI (2266)
NsiI (2479)
SP6 promoter
bGH poly(A) signal
SphI (2739)
lox2272
EcoRI (2814)
PspXI (2830)
HITI_pH11 target
HITI_hH11 target
ori
HITI_hH11 target
HITI_pH11 target
(4408) ScaI
AmpR
AmpR promoter
(4732) SspI
(5001) PvuII

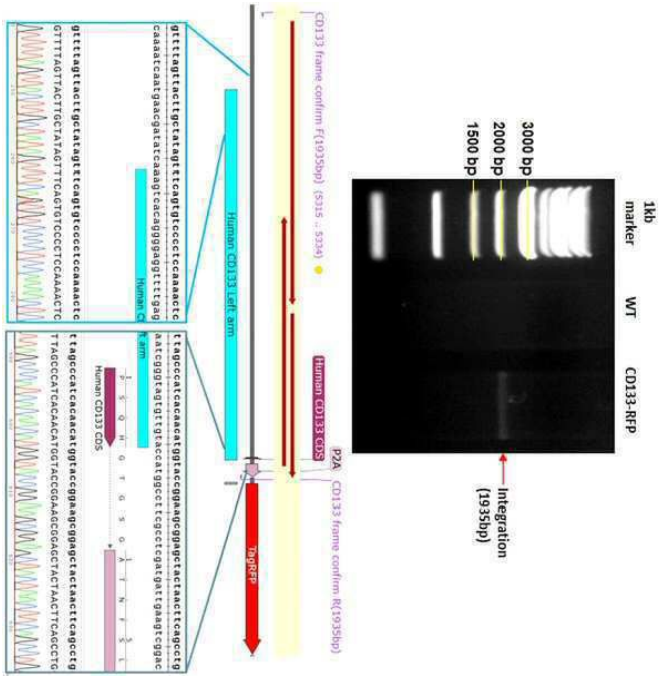
도면15


1 weeks
BF
RFP
Merge
CD133 Knock-in
2 weeks
Sphere formation
x40
x100

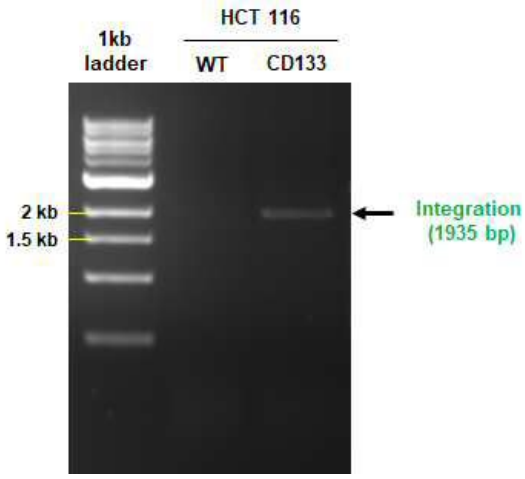
도면16


NCCIT
SH-SY5Y
HT-29
Bright field
RFP
Endogenous
Bright field
RFP
Partial

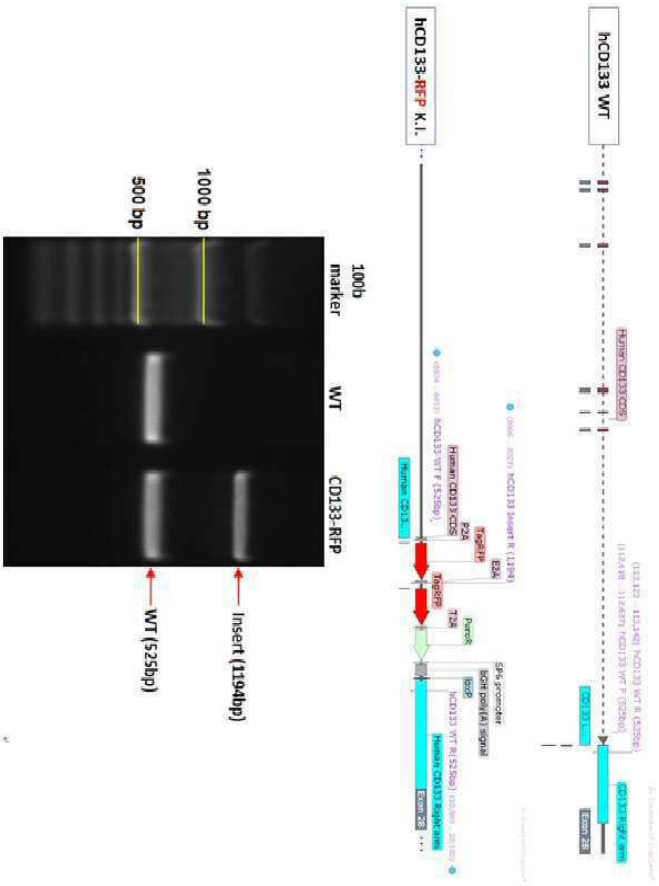
도면17


Human CD133 Left arm
Human CD133 CDS
P2A
TagRFP
CD133 frame confirm R(1935bp)
3000 bp
2000 bp
1500 bp
1kb marker
WT
CD133-RFP
Integration (1935bp)

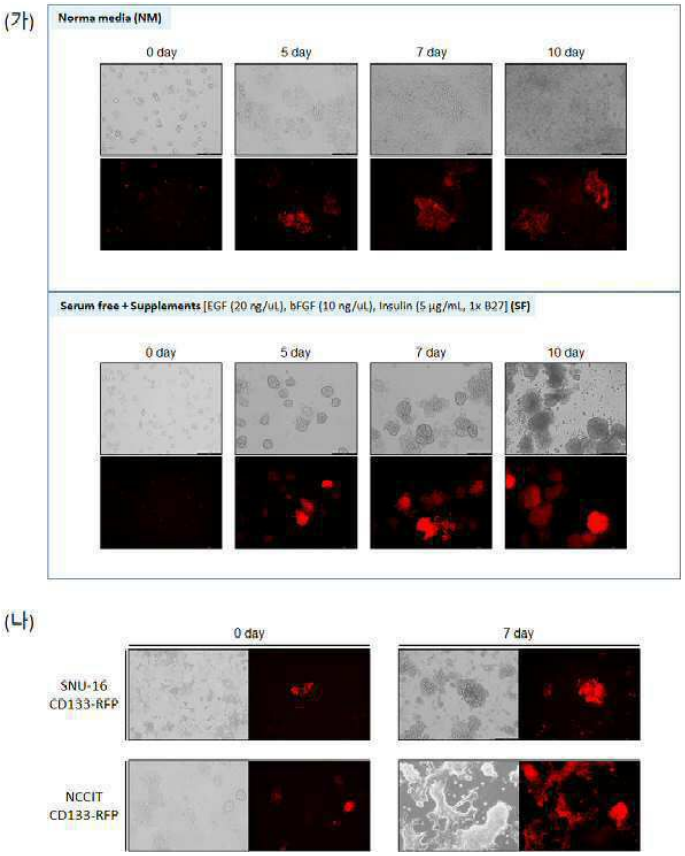
도면18


HCT 116
1kb ladder
WT
CD133
2 kb
1.5 kb
Integration (1935 bp)

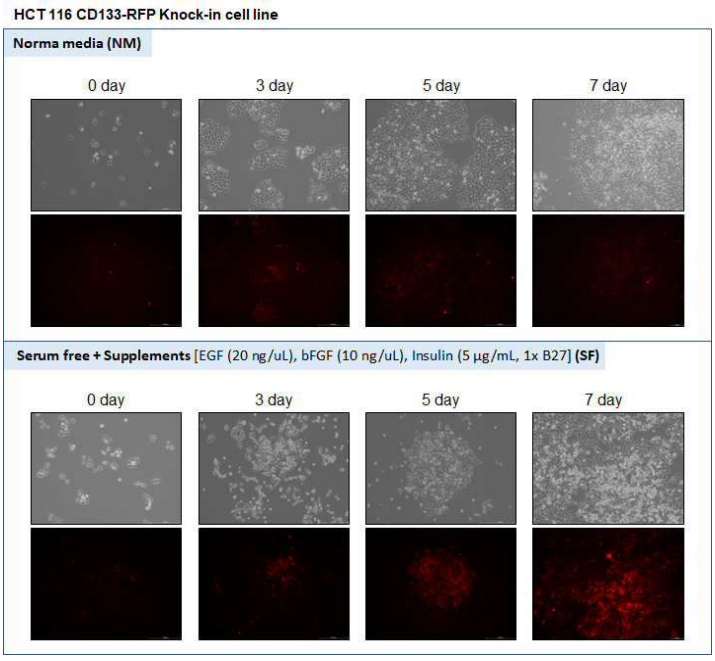
도면19


hCD133 WT
hCD133-RFP K.I.
100b marker
WT
CD133-RFP
1000 bp
500 bp
WT (525bp)
Insert (1194bp)

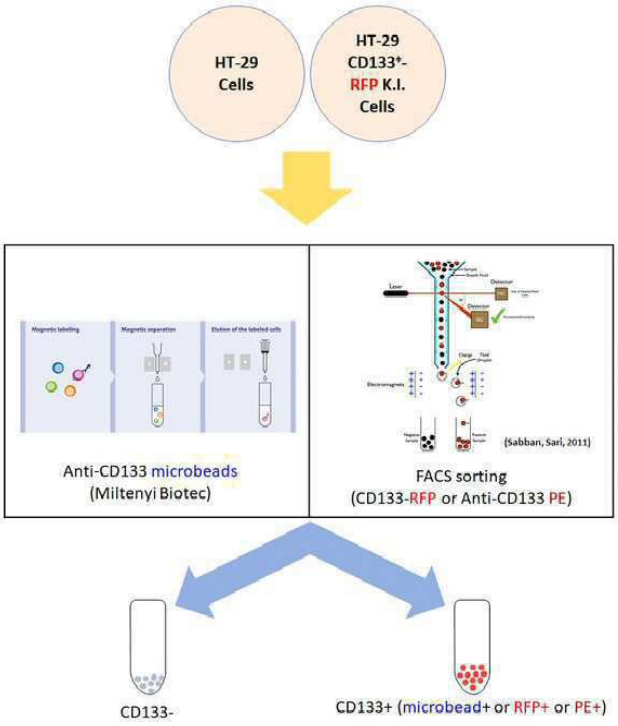
도면20


(가)
Norma media (NM)
0 day
5 day
7 day
10 day
Serum free + Supplements [EGF (20 ng/uL), bFGF (10 ng/uL), Insulin (5 μg/mL, 1x B27] (SF)
0 day
5 day
7 day
10 day
(나)
0 day
7 day
SNU-16
CD133-RFP
NCCIT
CD133-RFP

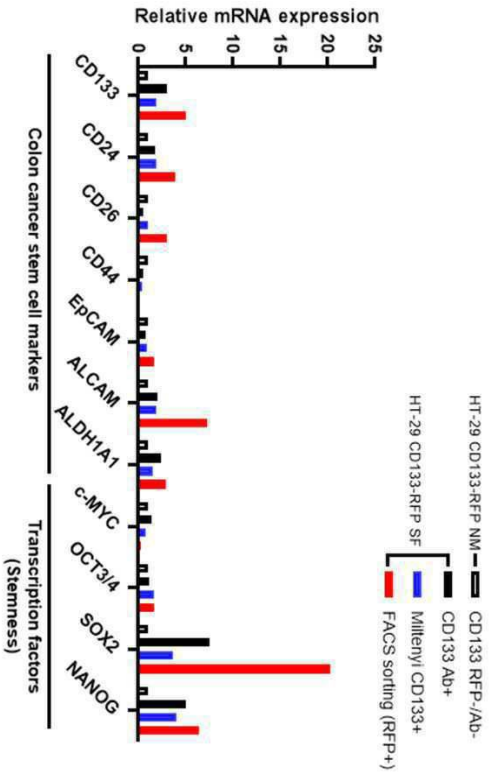
도면21


HCT 116 CD133-RFP Knock-in cell line
Norma media (NM)
0 day
3 day
5 day
7 day
Serum free + Supplements [EGF (20 ng/uL), bFGF (10 ng/uL), Insulin (5 μg/mL, 1x B27] (SF)
0 day
3 day
5 day
7 day

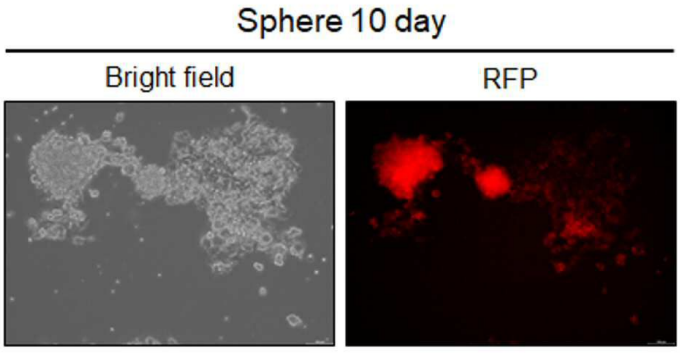
도면22


HT-29
Cells
HT-29
CD133+-
RFP K.I.
Cells
Anti-CD133 microbeads
(Miltenyi Biotec)
(Sabbari, Sari, 2011)
FACS sorting
(CD133-RFP or Anti-CD133 PE)
CD133-
CD133+ (microbead+ or RFP+ or PE+)

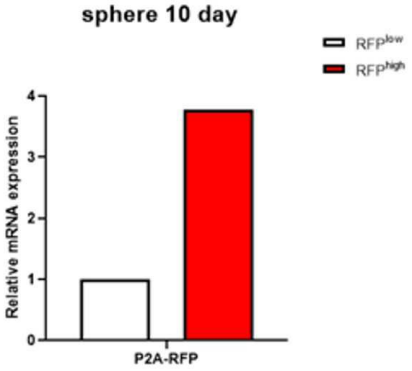
도면23


Relative mRNA expression
0
5
10
15
20
25
Colon cancer stem cell markers
Transcription factors
(Stemness)
CD133
CD24
CD26
CD44
EpCAM
ALCAM
ALDH1A1
c-MYC
OCT3/4
SOX2
NANOG
HT-29 CD133-RFP SF
HT-29 CD133-RFP NM
FACS sorting (RFP+)
Miltenyi CD133+
CD133 Ab+
CD133 RFP-/Ab-

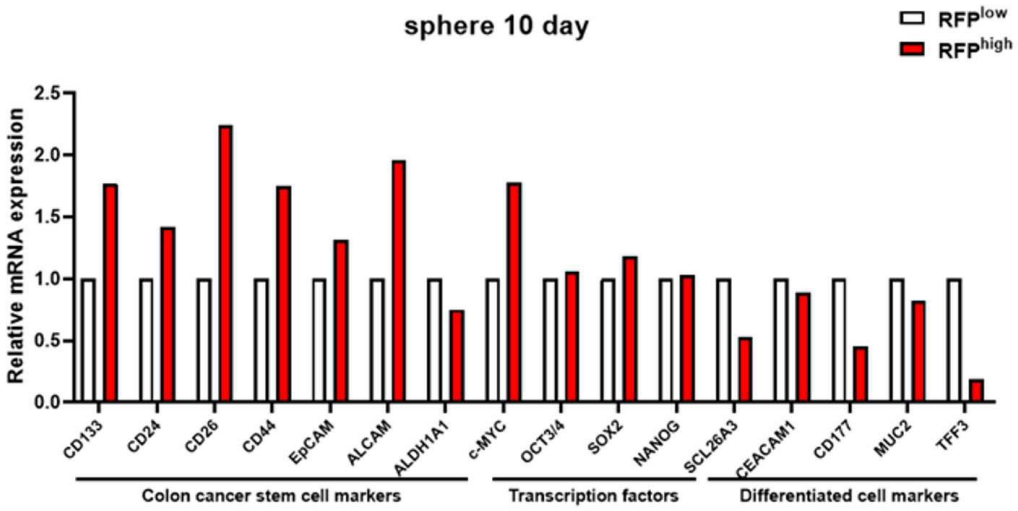
도면24



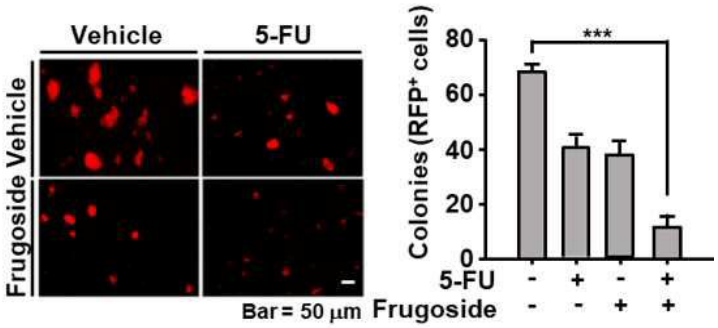
도면25



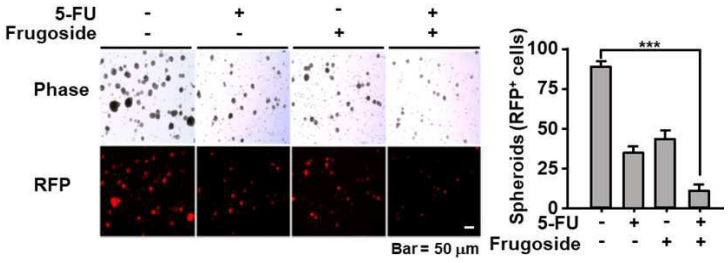
도면26



도면27



도면28



서 열 목 록

<110> Korea Research Institute of Bioscience and Biotechnology

University of Ulsan Foundation For Industry Cooperation

<120> CD133 KNOCK-IN FLUORESCENT REPORTER VECTOR SYSTEM, CELLS TRANSFORMED THEREWITH, AND CANCER STEM CELL DETECTION METHOD USING THE SAME

<130> FPD/202205-0002

<150> KR 2021-0064300

<151> 2021-05-18

<160> 68

<170> KoPatentIn 3.0

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for sgRNA_CD133_F

<400> 1

tcagctatca atgttgtgat 20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for sgRNA_CD133_R

<400> 2

atcacaacat tgatagctga 20

<210> 3

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for sgRNA_H11_F

<400> 3

caaaataggt tagccttgcg 20

<210> 4

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for sgRNA_H11_R

<400> 4

cgcaaggcta acctattttg 20

<210> 5

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for genomic DNA_CD133_Endogenous Left arm_F

<400> 5

atcaagtgta tcatatgttc agtgtcccct ccaaaactc 39

<210> 6

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for genomic DNA_CD133_Endogenous Left arm_R

<400> 6

gtagctccgc ttccggtacc atgttgtgat gggctaaaaa acaa 44

<210> 7

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for genomic DNA_CD133_Endogenous Right arm_F

<400> 7

agcccaagct agatcgaatt ctagctgatg ttgaaactgc ttgagg 46

<210> 8

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for genomic DNA_CD133_Endogenous Right arm_R

<400> 8

ggcggccgct ctagcctcga ggcacactcc acacatggca 40

<210> 9

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for genomic DNA_CD133_Partial_F

<400> 9

gattcattaa tgcagctgtg caggggctgc agaggggaaa t 41

<210> 10

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for genomic DNA_CD133_Partial_R

<400> 10

tctccagatc catcgatcgc tgccgccgcc cttccctta 39

<210> 11

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD133_F

<400> 11

gggccaactt ctctctgttg 20

<210> 12

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD133_R

<400> 12

actcgagagg tccagggttc 20

<210> 13

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_hCD133 WT/Insert_F

<400> 13

gctgtgttgc tgctggtcta 20

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_hCD133 WT/Insert_R1

<400> 14

ggtacagagg gtggactgga 20

<210> 15

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_hCD133 WT/Insert_R2

<400> 15

ccagccaatt tcaagagagc 20

<210> 16

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_B2M_F

<400> 16

ctccgtggcc ttaggtgtg 19

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_B2M_R

<400> 17

tctctgctgg atgacgtgag 20

<210> 18

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD133_F

<400> 18

aagcattggc atcttctatg g 21

<210> 19

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD133_R

<400> 19

ttcaagagag ttcgcaagtc c 21

<210> 20

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD24_F

<400> 20

gatttattcc agtgaaacaa caactg 26

<210> 21

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD24_R

<400> 21

accacgaaga gactggctgt 20

<210> 22

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_EpCAM_F

<400> 22

gctcaggaag aatgtgtctg tg 22

<210> 23

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_EpCAM_R

<400> 23

gccagctttg agcaaatgac 20

<210> 24

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD26_F

<400> 24

ctctgaacgc tcacttccga 20

<210> 25

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD26_R

<400> 25

tttgttcagc agaaccacgg 20

<210> 26

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD44_F

<400> 26

agactcccat tcgacaacag g 21

<210> 27

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_CD44_R

<400> 27

ggaactgtcc tctgggcttg 20

<210> 28

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_ALCAM_F

<400> 28

acttgacgta cctcagaatc tca 23

<210> 29

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_ALCAM_R

<400> 29

catcgtcgta ctgcacactt t 21

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_ALDH1A1_F

<400> 30

gcacgccaga cttacctgtc 20

<210> 31

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_ALDH1A1_R

<400> 31

cctcctcagt tgcaggatta aag 23

<210> 32

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_c-MYC_F

<400> 32

gtcctcggat tctctgctct c 21

<210> 33

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_c-MYC_R

<400> 33

catcttcttg ttcctcctca gagt 24

<210> 34

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_OCT3/4_F

<400> 34

gcccgaaaga gaaagcgaac 20

<210> 35

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_OCT3/4_R

<400> 35

aaccacactc ggaccacatc 20

<210> 36

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_NANOG_F

<400> 36

tgtgttctct tccacccagc 20

<210> 37

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for primer_NANOG_R

<400> 37

accaggtctt cacctgtttg 20

<210> 38

<211> 10246

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for pCAG-SpCas9-GFP-U6 CD133 gRNA vector

<400> 38

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60

ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120

aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180

atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

cgaaacaccg tcagctatca atgttgtgat gttttagagc tagaaatagc aagttaaaat 300

aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttgttttaga 360

gctagaaata gcaagttaaa ataaggctag tccgttttta gcgcgtgcgc caattctgca 420

gacaaatggc tctagtaacg gccgccagtg tgctggaatt caggctagtt attaatagta 480

atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac 540

ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac 600

gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt 660

acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagta cgccccctat 720

tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttatggga 780

ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg tcgaggtgag 840

ccccacgttc tgcttcactc tccccatctc ccccccctcc ccacccccaa ttttgtattt 900

atttatttttt taattatttt gtgcagcgat gggggcgggg ggggggggggg ggcgcgcgcc 960

aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg cggcggcagc 1020

caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc ggcggcggcc 1080

ctataaaaag cgaagcgcgc ggcgggcggg gagtcgctgc gacgctgcct tcgccccgtg 1140

ccccgctccg ccgccgcctc gcgccgcccg ccccggctct gactgaccgc gttactccca 1200

caggtgagcg ggcgggacgg cccttctcct ccgggctgta attagcgctt ggtttaatga 1260

cggcttgttt cttttctgtg gctgcgtgaa agccttgagg ggctccggga gggcccttg 1320

tgcgggggga gcggctcggg gggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 1380

ggctccgcgc tgcccggcgg ctgtgagcgc tgcgggcgcg gcgcggggct ttgtgcgctc 1440

cgcagtgtgc gcgaggggag cgcggccggg ggcggtgccc cgcggtgcgg gggggctgc 1500

gaggggaaca aaggctgcgt gcggggtgtg tgcgtggggg ggtgagcagg gggtgtgggc 1560

gcgtcggtcg ggctgcaacc ccccctgcac ccccctcccc gagttgctga gcacggcccg 1620

gcttcgggtg cggggctccg tacggggcgt ggcgcggggc tcgccgtgcc gggcgggggg 1680

tggcggcagg tgggggtgcc gggcggggcg gggccgcctc gggccgggga gggctcgggg 1740

gaggggcgcg gcggcccccg gagcgccggc ggctgtcgag gcgcggcgag ccgcagccat 1800

tgccttttat ggtaatcgtg cgagagggcg cagggacttc ctttgtccca aatctgtgcg 1860

gagccgaaat ctgggaggcg ccgccgcacc ccctctagcg ggcgcggggc gaagcggtgc 1920

ggcgccggca ggaaggaaat gggcggggag ggccttcgtg cgtcgccgcg ccgccgtccc 1980

cttctccctc tccagcctcg gggctgtccg cgggggggacg gctgccttcg gggggggacgg 2040

ggcagggcgg ggttcggctt ctggcgtgtg accggcggct ctagagcctc tgctaaccat 2100

gttcatgcct tcttcttttt cctacagctc ctgggcaacg tgctggttat tgtgctgtct 2160

catcattttg gcaaagaatt cctgaattct gcagatccgg tgccaccatg gactataagg 2220

accacgacgg agactacaag gatcatgata ttgattacaa agacgatgac gataagatgg 2280

ccccaaagaa gaagcggaag gtcggtatcc acggagtccc agcagccgac aagaagtaca 2340
gcatcggcct ggacatcggc accaactctg tgggctgggc cgtgatcacc gacgagtaca 2400
aggtgcccag caagaaattc aaggtgctgg gcaacaccga ccggcacagc atcaagaaga 2460
acctgatcgg agccctgctg ttcgacagcg gcgaaacagc cgaggccacc cggctgaaga 2520
gaaccgccag aagaagatac accagacgga agaaccggat ctgctatctg caagagatct 2580
tcagcaacga gatggccaag gtggacgaca gcttcttcca cagactggaa gagtccttcc 2640
tggtggaaga ggataagaag cacgagcggc accccatctt cggcaacatc gtggacgagg 2700

tggcctacca cgagaagtac cccaccatct accacctgag aaagaaactg gtggacagca 2760
ccgacaaggc cgacctgcgg ctgatctatc tggccctggc ccacatgatc aagttccggg 2820
gccacttcct gatcgagggc gacctgaacc ccgacaacag cgacgtggac aagctgttca 2880
tccagctggt gcagacctac aaccagctgt tcgaggaaaa ccccatcaac gccagcggcg 2940
tggacgccaa ggccatcctg tctgccagac tgagcaagag cagacggctg gaaaatctga 3000
tcgcccagct gcccggcgag aagaagaatg gcctgttcgg aaacctgatt gccctgagcc 3060
tgggcctgac ccccaacttc aagagcaact tcgacctggc cgaggatgcc aaactgcagc 3120

tgagcaagga cacctacgac gacgacctgg acaacctgct ggcccagatc ggcgaccagt 3180
acgccgacct gtttctggcc gccaagaacc tgtccgacgc catcctgctg agcgacatcc 3240
tgagagtgaa caccgagatc accaaggccc ccctgagcgc ctctatgatc aagagatacg 3300
acgagcacca ccaggacctg accctgctga aagctctcgt gcggcagcag ctgcctgaga 3360
agtacaaaga gattttcttc gaccagagca agaacggcta cgccggctac attgacggcg 3420
gagccagcca ggaagagttc tacaagttca tcaagcccat cctggaaaag atggacggca 3480
ccgaggaact gctcgtgaag ctgaacagag aggacctgct gcggaagcag cggaccttcg 3540

acaacggcag catcccccac cagatccacc tgggagagct gcacgccatt ctgcggcggc 3600
aggaagattt ttacccattc ctgaaggaca accgggaaaa gatcgagaag atcctgacct 3660
tccgcatccc ctactacgtg ggccctctgg ccaggggaaa cagcagattc gcctggatga 3720
ccagaaagag cgaggaaacc atcacccct ggaacttcga ggaagtggtg gacaagggcg 3780
cttccgccca gagcttcatc gagcggatga ccaacttcga taagaacctg cccaacgaga 3840
aggtgctgcc caagcacagc ctgctgtacg agtacttcac cgtgtataac gagctgacca 3900
aagtgaaata cgtgaccgag ggaatgagaa agcccgcctt cctgagcggc gagcagaaaa 3960

aggccatcgt ggacctgctg ttcaagacca accggaaagt gaccgtgaag cagctgaaag 4020

aggactactt caagaaaatc gagtgcttcg actccgtgga aatctccggc gtggaagatc 4080

ggttcaacgc ctccctgggc acataccacg atctgctgaa aattatcaag gacaaggact 4140

tcctggacaa tgaggaaaac gaggacattc tggaagatat cgtgctgacc ctgacactgt 4200

ttgaggacag agagatgatc gaggaacggc tgaaaaccta tgcccacctg ttcgacgaca 4260

aagtgatgaa gcagctgaag cggcggagat acaccggctg gggcaggctg agccggaagc 4320

tgatcaacgg catccgggac aagcagtccg gcaagacaat cctggatttc ctgaagtccg 4380

acggcttcgc caacagaaac ttcatgcagc tgatccacga cgacagcctg acctttaaag 4440

aggacatcca gaaagcccag gtgtccggcc agggcgatag cctgcacgag cacattgcca 4500

atctggccgg cagccccgcc attaagaagg gcatcctgca gacagtgaag gtggtggacg 4560

agctcgtgaa agtgatgggc cggcacaagc ccgagaacat cgtgatcgaa atggccagag 4620

agaaccagac cacccagaag ggacagaaga acagccgcga gagaatgaag cggatcgaag 4680

agggcatcaa agagctgggc agccagatcc tgaaagaaca ccccgtggaa aacacccagc 4740

tgcagaacga gaagctgtac ctgtactacc tgcagaatgg gcgggatatg tacgtggacc 4800

aggaactgga catcaaccgg ctgtccgact acgatgtgga ccatatcgtg cctcagagct 4860

ttctgaagga cgactccatc gacaacaagg tgctgaccag aagcgacaag aaccggggca 4920

agagcgacaa cgtgccctcc gaagaggtcg tgaagaagat gaagaactac tggcggcagc 4980

tgctgaacgc caagctgatt acccagagaa agttcgacaa tctgaccaag gccgagagag 5040

gcggcctgag cgaactggat aaggccggct tcatcaagag acagctggtg gaaacccggc 5100

agatcacaaa gcacgtggca cagatcctgg actcccggat gaacactaag tacgacgaga 5160

atgacaagct gatccgggaa gtgaaagtga tcaccctgaa gtccaagctg gtgtccgatt 5220

tccggaagga tttccagttt tacaaagtgc gcgagatcaa caactaccac cacgcccacg 5280

acgcctacct gaacgccgtc gtgggaaccg ccctgatcaa aaagtaccct aagctggaaa 5340

gcgagttcgt gtacggcgac tacaaggtgt acgacgtgcg gaagatgatc gccaagagcg 5400

agcaggaaat cggcaaggct accgccaagt acttcttcta cagcaacatc atgaactttt 5460

tcaagaccga gattaccctg gccaacggcg agatccggaa gcggcctctg atcgagacaa 5520

acggcgaaac cggggagatc gtgtgggata agggccggga ttttgccacc gtgcggaaag 5580

tgctgagcat gccccaagtg aatatcgtga aaaagaccga ggtgcagaca ggcggcttca 5640

gcaaagagtc tatcctgccc aagaggaaca gcgataagct gatcgccaga aagaaggact 5700

gggaccctaa gaagtacggc ggcttcgaca gccccaccgt ggcctattct gtgctggtgg 5760

tggccaaagt ggaaaagggc aagtccaaga aactgaagag tgtgaaagag ctgctgggga 5820

tcaccatcat ggaaagaagc agcttcgaga agaatcccat cgactttctg gaagccaagg 5880

gctacaaaga agtgaaaaag gacctgatca tcaagctgcc taagtactcc ctgttcgagc 5940

tggaaaacgg ccggaagaga atgctggcct ctgccggcga actgcagaag ggaaacgaac 6000

tggccctgcc ctccaaatat gtgaacttcc tgtacctggc cagccactat gagaagctga 6060

agggctcccc cgaggataat gagcagaaac agctgtttgt ggaacagcac aagcactacc 6120

tggacgagat catcgagcag atcagcgagt tctccaagag agtgatcctg gccgacgcta 6180

atctggacaa agtgctgtcc gcctacaaca agcaccggga taagcccatc agagagcagg 6240

ccgagaatat catccacctg tttaccctga ccaatctggg agcccctgcc gccttcaagt 6300

actttgacac caccatcgac cggaagaggt acaccagcac caaagaggtg ctggacgcca 6360

ccctgatcca ccagagcatc accggcctgt acgagacacg gatcgacctg tctcagctgg 6420

gaggcgacaa aaggccggcg gccacgaaaa aggccggcca ggcaaaaaag aaaaaggaat 6480

tcggcagtgg agagggcaga ggaagtctgc taacatgcgg tgacgtcgag gagaatcctg 6540

gcccagtgag caagggcgag gagctgttca ccggggtggt gcccatcctg gtcgagctgg 6600

acggcgacgt aaacggccac aagttcagcg tgtccggcga gggcgagggc gatgccacct 6660

acggcaagct gaccctgaag ttcatctgca ccaccggcaa gctgcccgtg ccctggccca 6720

ccctcgtgac caccctgacc tacggcgtgc agtgcttcag ccgctacccc gaccacatga 6780

agcagcacga cttcttcaag tccgccatgc ccgaaggcta cgtccaggag cgcaccatct 6840

tcttcaagga cgacggcaac tacaagaccc gcgccgaggt gaagttcgag ggcgacaccc 6900

tggtgaaccg catcgagctg aagggcatcg acttcaagga ggacggcaac atcctggggc 6960

acaagctgga gtacaactac aacagccaca acgtctatat catggccgac aagcagaaga 7020

acggcatcaa ggtgaacttc aagatccgcc acaacatcga ggacggcagc gtgcagctcg 7080

ccgaccacta ccagcagaac accccccatcg gcgacggccc cgtgctgctg cccgacaacc 7140

actacctgag cacccagtcc gccctgagca aagaccccaa cgagaagcgc gatcacatgg 7200

tcctgctgga gttcgtgacc gccgccggga tcactctcgg catggacgag ctgtacaagg 7260

aattctaact agagctcgct gatcagcctc gactgtgcct tctagttgcc agccatctgt 7320

tgtttgcccc tcccccgtgc cttccttgac cctggaaggt gccactccca ctgtcctttc 7380

ctaataaaat gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg 7440

tggggtgggg caggacagca agggggagga ttgggaagag aatagcaggc atgctgggga 7500

gcggccgcag gaacccctag tgatggagtt ggccactccc tctctgcgcg ctcgctcgct 7560

cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg cggcctcagt 7620

gagcgagcga gcgcgcagct gcctgcaggg gcgcctgatg cggtattttc tccttacgca 7680

tctgtgcggt atttcacacc gcatacgtca aagcaaccat agtacgcgcc ctgtagcggc 7740

gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc 7800
ctagcgcccg ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc 7860
cgtcaagctc taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc 7920
gaccccaaaa aacttgattt gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg 7980
gtttttcgcc ctttgacgtt ggagtccacg ttctttaata gtggactctt gttccaaact 8040
ggaacaacac tcaaccctat ctcgggctat tcttttgatt tataagggat tttgccgatt 8100
tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa 8160

atattaacgt ttacaatttt atggtgcact ctcagtacaa tctgctctga tgccgcatag 8220
ttaagccagc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc 8280
ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatgtg tcagaggttt 8340
tcaccgtcat caccgaaacg cgcgagacga aagggcctcg tgatacgcct atttttatag 8400
gttaatgtca tgataataat ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg 8460
cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga 8520
caataaccct gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat 8580

ttccgtgtcg cccttattcc ctttttgcg gcattttgcc ttcctgtttt tgctcaccca 8640
gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc 8700
gaactggatc tcaacagcgg taagatcctt gagagttttc gccccgaaga acgttttcca 8760
atgatgagca cttttaaagt tctgctatgt ggcgcggtat tatcccgtat tgacgccggg 8820
caagagcaac tcggtcgccg catacactat tctcagaatg acttggttga gtactcacca 8880
gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata 8940
accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag 9000

ctaaccgctt ttttgcacaa catgggggat catgtaactc gccttgatcg ttgggaaccg 9060
gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca 9120
acaacgttgc gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta 9180
atagactgga tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct 9240
ggctggttta ttgctgataa atctggagcc ggtgagcgtg gaagccgcgg tatcattgca 9300
gcactggggc cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag 9360
gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat 9420

tggtaactgt cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt 9480

taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa 9540

cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 9600

gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg 9660

gtggtttgtt tgccggatca agagctacca actcttttc cgaaggtaac tggcttcagc 9720

agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag 9780

aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc 9840

agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 9900

cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac 9960

accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga 10020

aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt 10080

ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag 10140

cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat ggaaaaacgc cagcaacgcg 10200

gccttttttac ggttcctggc cttttgctgg ccttttgctc acatgt 10246

<210> 39

<211> 9463

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for pCAG-SpCas9-GFP-U6 H11 gRNA vector

<400> 39

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60

ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120

aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180

atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

cgaaacaccg caaaataggt tagccttgcg gttttagagc tagaaatagc aagttaaaat 300

aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttgttttaga 360

gctagaaata gcaagttaaa ataaggctag tccgttttta gcgcgtgcgc caattctgca 420

gacaaatggc tctagtaacg gccgccagtg tgctggaatt caggctagtt attaatagta 480

atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac 540

ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac 600

gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt 660

acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagta cgccccctat 720

tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttatggga 780

ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg tcgaggtgag 840

ccccacgttc tgcttcactc tccccatctc ccccccctcc ccacccccaa ttttgtattt 900

atttatttttt taattatttt gtgcagcgat gggggcgggg ggggggggggg ggcgcgcgcc 960

aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg cggcggcagc 1020

caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc ggcggcggcc 1080

ctataaaaag cgaagcgcgc ggcgggcggg gagtcgctgc gacgctgcct tcgccccgtg 1140

ccccgctccg ccgccgcctc gcgccgcccg ccccggctct gactgaccgc gttactccca 1200

caggtgagcg ggcgggacgg cccttctcct ccgggctgta attagcgctt ggtttaatga 1260

cggcttgttt cttttctgtg gctgcgtgaa agccttgagg ggctccggga gggccctttg 1320

tgcgggggga gcggctcggg gggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 1380

ggctccgcgc tgcccggcgg ctgtgagcgc tgcgggcgcg gcgcggggct ttgtgcgctc 1440

cgcagtgtgc gcgaggggag cgcggccggg ggcggtgccc cgcggtgcgg gggggctgc 1500

gaggggaaca aaggctgcgt gcggggtgtg tgcgtggggg ggtgagcagg gggtgtgggc 1560

gcgtcggtcg ggctgcaacc ccccctgcac ccccctcccc gagttgctga gcacggcccg 1620

gcttcgggtg cggggctccg tacggggcgt ggcgcggggc tcgccgtgcc gggcgggggg 1680

tggcggcagg tgggggtgcc gggcggggcg gggccgcctc gggccgggga gggctcgggg 1740

gaggggcgcg gcggcccccg gagcgccggc ggctgtcgag gcgcggcgag ccgcagccat 1800

tgccttttat ggtaatcgtg cgagagggcg cagggacttc ctttgtccca aatctgtgcg 1860

gagccgaaat ctgggaggcg ccgccgcacc ccctctagcg ggcgcggggc gaagcggtgc 1920

ggcgccggca ggaaggaaat gggcggggag ggccttcgtg cgtcgccgcg ccgccgtccc 1980

cttctccctc tccagcctcg gggctgtccg cgggggggacg gctgccttcg gggggacgg 2040

ggcagggcgg ggttcggctt ctggcgtgtg accggcggct ctagagcctc tgctaaccat 2100

gttcatgcct tcttcttttt cctacagctc ctgggcaacg tgctggttat tgtgctgtct 2160

catcattttg gcaaagaatt cctgaattct gcagatccgg tgccaccatg gactataagg 2220

accacgacgg agactacaag gatcatgata ttgattacaa agacgatgac gataagatgg 2280

ccccaaagaa gaagcggaag gtcggtatcc acggagtccc agcagccgac aagaagtaca 2340

gcatcggcct ggacatcggc accaactctg tgggctgggc cgtgatcacc gacgagtaca 2400

aggtgcccag caagaaattc aaggtgctgg gcaacaccga ccggcacagc atcaagaaga 2460

acctgatcgg agccctgctg ttcgacagcg gcgaaacagc cgaggccacc cggctgaaga 2520

gaaccgccag aagaagatac accagacgga agaaccggat ctgctatctg caagagatct 2580

tcagcaacga gatggccaag gtggacgaca gcttcttcca cagactggaa gagtccttcc 2640

tggtggaaga ggataagaag cacgagcggc accccatctt cggcaacatc gtggacgagg 2700

tggcctacca cgagaagtac cccaccatct accacctgag aaagaaactg gtggacagca 2760

ccgacaaggc cgacctgcgg ctgatctatc tggccctggc ccacatgatc aagttccggg 2820

gccacttcct gatcgagggc gacctgaacc ccgacaacag cgacgtggac aagctgttca 2880

tccagctggt gcagacctac aaccagctgt tcgaggaaaa ccccatcaac gccagcggcg 2940

tggacgccaa ggccatcctg tctgccagac tgagcaagag cagacggctg gaaaatctga 3000

tcgcccagct gcccggcgag aagaagaatg gcctgttcgg aaacctgatt gccctgagcc 3060

tgggcctgac ccccaacttc aagagcaact tcgacctggc cgaggatgcc aaactgcagc 3120

tgagcaagga cacctacgac gacgacctgg acaacctgct ggcccagatc ggcgaccagt 3180

acgccgacct gtttctggcc gccaagaacc tgtccgacgc catcctgctg agcgacatcc 3240

tgagagtgaa caccgagatc accaaggccc ccctgagcgc ctctatgatc aagagatacg 3300

acgagcacca ccaggacctg accctgctga aagctctcgt gcggcagcag ctgcctgaga 3360

agtacaaaga gattttcttc gaccagagca agaacggcta cgccggctac attgacggcg 3420

gagccagcca ggaagagttc tacaagttca tcaagcccat cctggaaaag atggacggca 3480

ccgaggaact gctcgtgaag ctgaacagag aggacctgct gcggaagcag cggaccttcg 3540

acaacggcag catcccccac cagatccacc tgggagagct gcacgccatt ctgcggcggc 3600

aggaagattt ttacccattc ctgaaggaca accgggaaaa gatcgagaag atcctgacct 3660

tccgcatccc ctactacgtg ggccctctgg ccaggggaaa cagcagattc gcctggatga 3720

ccagaaagag cgaggaaacc atcacccccc ggaacttcga ggaagtggtg gacaagggcg 3780

cttccgccca gagcttcatc gagcggatga ccaacttcga taagaacctg cccaacgaga 3840

aggtgctgcc caagcacagc ctgctgtacg agtacttcac cgtgtataac gagctgacca 3900

aagtgaaata cgtgaccgag ggaatgagaa agcccgcctt cctgagcggc gagcagaaaa 3960

aggccatcgt ggacctgctg ttcaagacca accggaaagt gaccgtgaag cagctgaaag 4020

aggactactt caagaaaatc gagtgcttcg actccgtgga aatctccggc gtggaagatc 4080

ggttcaacgc ctccctgggc acataccacg atctgctgaa aattatcaag gacaaggact 4140

tcctggacaa tgaggaaaac gaggacattc tggaagatat cgtgctgacc ctgacactgt 4200

ttgaggacag agagatgatc gaggaacggc tgaaaaccta tgcccacctg ttcgacgaca 4260

aagtgatgaa gcagctgaag cggcggagat acaccggctg gggcaggctg agccggaagc 4320

tgatcaacgg catccgggac aagcagtccg gcaagacaat cctggatttc ctgaagtccg 4380

acggcttcgc caacagaaac ttcatgcagc tgatccacga cgacagcctg acctttaaag 4440

aggacatcca gaaagcccag gtgtccggcc agggcgatag cctgcacgag cacattgcca 4500

atctggccgg cagccccgcc attaagaagg gcatcctgca gacagtgaag gtggtggacg 4560

agctcgtgaa agtgatgggc cggcacaagc ccgagaacat cgtgatcgaa atggccagag 4620

agaaccagac cacccagaag ggacagaaga acagccgcga gagaatgaag cggatcgaag 4680

agggcatcaa agagctgggc agccagatcc tgaaagaaca ccccgtggaa aacacccagc 4740

tgcagaacga gaagctgtac ctgtactacc tgcagaatgg gcgggatatg tacgtggacc 4800

aggaactgga catcaaccgg ctgtccgact acgatgtgga ccatatcgtg cctcagagct 4860

ttctgaagga cgactccatc gacaacaagg tgctgaccag aagcgacaag aaccggggca 4920

agagcgacaa cgtgccctcc gaagaggtcg tgaagaagat gaagaactac tggcggcagc 4980

tgctgaacgc caagctgatt acccagagaa agttcgacaa tctgaccaag gccgagagag 5040

gcggcctgag cgaactggat aaggccggct tcatcaagag acagctggtg gaaacccggc 5100

agatcacaaa gcacgtggca cagatcctgg actcccggat gaacactaag tacgacgaga 5160

atgacaagct gatccgggaa gtgaaagtga tcaccctgaa gtccaagctg gtgtccgatt 5220

tccggaagga tttccagttt tacaaagtgc gcgagatcaa caactaccac cacgcccacg 5280

acgcctacct gaacgccgtc gtgggaaccg ccctgatcaa aaagtaccct aagctggaaa 5340

gcgagttcgt gtacggcgac tacaaggtgt acgacgtgcg gaagatgatc gccaagagcg 5400

agcaggaaat cggcaaggct accgccaagt acttcttcta cagcaacatc atgaactttt 5460

tcaagaccga gattaccctg gccaacggcg agatccggaa gcggcctctg atcgagacaa 5520

acggcgaaac cggggagatc gtgtgggata agggccggga ttttgccacc gtgcggaaag 5580

tgctgagcat gccccaagtg aatatcgtga aaaagaccga ggtgcagaca ggcggcttca 5640

gcaaagagtc tatcctgccc aagaggaaca gcgataagct gatcgccaga aagaaggact 5700

gggaccctaa gaagtacggc ggcttcgaca gccccaccgt ggcctattct gtgctggtgg 5760

tggccaaagt ggaaaagggc aagtccaaga aactgaagag tgtgaaagag ctgctgggga 5820

tcaccatcat ggaaagaagc agcttcgaga agaatcccat cgactttctg gaagccaagg 5880

gctacaaaga agtgaaaaag gacctgatca tcaagctgcc taagtactcc ctgttcgagc 5940

tggaaaacgg ccggaagaga atgctggcct ctgccggcga actgcagaag ggaaacgaac 6000

tggccctgcc ctccaaatat gtgaacttcc tgtacctggc cagccactat gagaagctga 6060

agggctcccc cgaggataat gagcagaaac agctgtttgt ggaacagcac aagcactacc 6120

tggacgagat catcgagcag atcagcgagt tctccaagag agtgatcctg gccgacgcta 6180

atctggacaa agtgctgtcc gcctacaaca agcaccggga taagcccatc agagagcagg 6240

ccgagaatat catccacctg tttaccctga ccaatctggg agcccctgcc gccttcaagt 6300

actttgacac caccatcgac cggaagaggt acaccagcac caaagaggtg ctggacgcca 6360

ccctgatcca ccagagcatc accggcctgt acgagacacg gatcgacctg tctcagctgg 6420

gaggcgacaa aaggccggcg gccacgaaaa aggccggcca ggcaaaaaag aaaaagtaag 6480

aattcctaga gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt 6540

ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta 6600

ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg 6660

ggtggggcag gacagcaagg gggaggattg ggaagagaat agcaggcatg ctggggagcg 6720

gccgcaggaa cccctagtga tggagttggc cactccctct ctgcgcgctc gctcgctcac 6780

tgaggccggg cgaccaaagg tcgcccgacg cccgggcttt gcccgggcgg cctcagtgag 6840

cgagcgagcg cgcagctgcc tgcaggggcg cctgatgcgg tattttctcc ttacgcatct 6900

gtgcggtatt tcacaccgca tacgtcaaag caaccatagt acgcgccctg tagcggcgca 6960

ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta 7020

gcgcccgctc ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt 7080

caagctctaa atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac 7140

cccaaaaaac ttgatttggg tgatggttca cgtagtgggc catcgccctg atagacggtt 7200

tttcgccctt tgacgttgga gtccacgttc tttaatagtg gactcttgtt ccaaactgga 7260

acaacactca accctatctc gggctattct tttgatttat aagggatttt gccgatttcg 7320

gcctattggt taaaaaatga gctgatttaa caaaaattta acgcgaattt taacaaaata 7380

ttaacgttta caattttatg gtgcactctc agtacaatct gctctgatgc cgcatagtta 7440

agccagcccc gacacccgcc aacacccgct gacgcgccct gacgggcttg tctgctcccg 7500

gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca 7560

ccgtcatcac cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttataggtt 7620

aatgtcatga taataatggt ttcttagacg tcaggtggca cttttcgggg aaatgtgcgc 7680

ggaacccCta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 7740

taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc 7800

cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa 7860

acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa 7920

ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 7980

atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa 8040

gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc 8100

acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 8160

atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 8220

accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag 8280

ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca 8340

acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca acaattaata 8400

gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc 8460

tggtttattg ctgataaatc tggagccggt gagcgtggaa gccgcggtat cattgcagca 8520

ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca 8580

actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 8640

taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 8700

tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt 8760

gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 8820

cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg 8880

gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga 8940

gcgcagatac caaatactgt ccttctagtg tagccgtagt taggccacca cttcaagaac 9000

tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 9060

ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag 9120

cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 9180

gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag 9240

gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca 9300

gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 9360

cgatttttgt gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc 9420

tttttacggt tcctggcctt ttgctggcct tttgctcaca tgt 9463

<210> 40

<211> 8434

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Loxp-TP2A-TagRFP
2x-T2A-Puro-polyA-Loxp-Zsgreen in pCAGGS HDR repoter vector

<400> 40

gattcgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 60

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 120

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 180

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 240

acatcaagtg tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc 300

cgcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta 360

cgtattagtc atcgctatta ccatggtcga ggtgagcccc acgttctgct tcactctccc 420

catctccccc ccctccccac ccccaatttt gtatttattt atttttaat tattttgtgc 480

agcgatgggg gcgggggggg gggggggggcg cgcgccaggc ggggcggggc ggggcgaggg 540

gcggggcggg gcgaggcgga gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa 600

gtttcctttt atggcgaggc ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg 660

ggcgggagtc gctgcgttgc cttcgccccg tgccccgctc cgcgccgcct cgcgccgccc 720

gccccggctc tgactgaccg cgttactccc acaggtgagc gggcgggacg gcccttctcc 780

tccgggctgt aattagcgct tggtttaatg acggctcgtt tcttttctgt ggctgcgtga 840

aagccttaaa gggctccggg agggcccttt gtgcgggggg gagcggctcg gggggtgcgt 900

gcgtgtgtgt gtgcgtgggg agcgccgcgt gcggcccgcg ctgcccggcg gctgtgagcg 960

ctgcgggcgc ggcgcggggc tttgtgcgct ccgcgtgtgc gcgaggggag cgcggccggg 1020

ggcggtgccc cgcggtgcgg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt 1080

gcgtgggggg gtgagcaggg ggtgtgggcg cggcggtcgg gctgtaaccc ccccctgcac 1140

ccccctcccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tgcggggcgt 1200

ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg 1260

gggccgcctc gggccgggga gggctcgggg gaggggcgcg gcggccccgg agcgccggcg 1320

gctgtcgagg cgcggcgagc cgcagccatt gccttttatg gtaatcgtgc gagagggcgc 1380

agggacttcc tttgtcccaa atctggcgga gccgaaatct gggaggcgcc gccgcacccc 1440

ctctagcggg cgcgggcgaa gcggtgcggc gccggcagga aggaaatggg cggggagggc 1500

cttcgtgcgt cgccgcgccg ccgtcccctt ctccatctcc agcctcgggg ctgccgcagg 1560

gggacggctg ccttcggggg ggacggggca gggcggggtt cggcttctgg cgtgtgaccg 1620

gcggctctag agcctctgct aaccatgttc atgccttctt ctttttccta cagatcctta 1680

attaataata cgactcacta tagggggtcg acccgccacc aatcatttaa atataacttc 1740

gtataatgta tgctatacga agttattcgg aattcgctag catcgatgga tctggagaag 1800

gacgcggcag cctactgact tgcggagatg tcgaagagaa ccctggccct aagcttcagc 1860

tgggtaccgg aagcggagct actaacttca gcctgctgaa gcaggctggc gacgtggagg 1920

agaaccctgg acctctcgag tcgcgaacgc gtgccaccat ggtgtctaag ggcgaagagc 1980

tgattaagga gaacatgcac atgaagctgt acatggaggg caccgtgaac aaccaccact 2040

tcaagtgcac atccgagggc gaaggcaagc cctacgaggg cacccagacc atgagaatca 2100

aggtggtcga gggcggccct ctccccttcg ccttcgacat cctggctacc agcttcatgt 2160

acggcagcag aaccttcatc aaccacaccc agggcatccc cgacttcttt aagcagtcct 2220

tccctgaggg cttcacatgg gagagagtca ccacatacga ggacgggggc gtgctgaccg 2280

ctacccagga caccagcctc caggacggct gcctcatcta caacgtcaag atcagagggg 2340

tgaacttccc atccaacggc cctgtgatgc agaagaaaac actcggctgg gaggccaaca 2400

ccgagatgct gtaccccgct gacggcggcc tggaaggcag aagcgacatg gccctgaagc 2460

tcgtgggcgg gggccacctg atctgcaact tcaagaccac atacagatcc aagaaacccg 2520

ctaagaacct caagatgccc ggcgtctact atgtggacca cagactggaa agaatcaagg 2580

aggccgacaa agagacctac gtcgagcagc acgaggtggc tgtggccaga tactgcgacc 2640

tccctagcaa actggggcac aaacttaata cgcgtggctc gggccagtgt actaattatg 2700

ctctcttgaa attggctgga gatgttgaga gcaacccagg tcccgagctc gcggccgcag 2760

atctttttcc ctctgccaaa aattatgggg acatcatgaa gccccttgag catctgactt 2820

ctggcgagct tgccaccatg gtgtctaagg gcgaagagct gattaaggag aacatgcaca 2880

tgaagctgta catggagggc accgtgaaca accaccactt caagtgcaca tccgagggcg 2940

aaggcaagcc ctacgagggc acccagacca tgagaatcaa ggtggtcgag ggcggccctc 3000

tcccttcgc cttcgacatc ctggctacca gcttcatgta cggcagcaga accttcatca 3060

accacaccca gggcatcccc gacttcttta agcagtcctt ccctgagggc ttcacatggg 3120

agagagtcac cacatacgag gacgggggcg tgctgaccgc tacccaggac accagcctcc 3180

aggacggctg cctcatctac aacgtcaaga tcagaggggt gaacttccca tccaacggcc 3240

ctgtgatgca gaagaaaaca ctcggctggg aggccaacac cgagatgctg taccccgctg 3300

acggcggcct ggaaggcaga agcgacatgg ccctgaagct cgtgggcggg ggccacctga 3360

tctgcaactt caagaccaca tacagatcca agaaacccgc taagaacctc aagatgcccg 3420

gcgtctacta tgtggaccac agactggaaa gaatcaagga ggccgacaaa gagacctacg 3480

tcgagcagca cgaggtggct gtggccagat actgcgacct ccctagcaaa ctggggcaca 3540

aacttaatgg ctccgagggc agaggaagcc ttctaacatg cggtgacgtg gaggagaatc 3600

ccggcccttc cgggatgacc gagtacaagc ccacggtgcg cctcgccacc cgcgacgacg 3660

tccccagggc cgtacgcacc ctcgccgccg cgttcgccga ctaccccgcc acgcgccaca 3720

ccgtcgatcc agaccgccac atcgagcggg tcaccgagct gcaagaactc ttcctcacgc 3780

gcgtcgggct cgacatcggc aaggtgtggg tcgcggacga cggcgccgcg gtggcggtct 3840

ggaccacgcc ggagagcgtc gaagcggggg cggtgttcgc cgagatcggc ccgcgcatgg 3900

ccgagttgag cggttcccgg ctggccgcgc agcaacagat ggaaggtctc ctggcgccgc 3960

accggcccaa ggagcccgcg tggttcctgg ccaccgtcgg cgtctcgccc gaccaccagg 4020

gcaagggtct gggcagcgcc gtcgtgctcc ccggagtgga ggcggccgag cgcgccgggg 4080

tgcccgcctt cctggagacc tccgcgcccc gcaacctccc cttctacgag cggctcggct 4140

tcaccgtcac cgccgacgtc gaggtgcccg aaggaccgcg cacctggtgc atgacccgca 4200

agcccggtgc ctgacatgca tctagagggc cctattctat agtgtcacct aaatgctaga 4260

gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc 4320

cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag 4380

gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag 4440

gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 4500

atggggatcc ccatcaagct gatccggaac ccttaatata acttcgtata atgtatgcta 4560

tacgaagtta ttaggtccct cgacctgcag cccaagctag atcgaattcg gccggccgat 4620

atcggcgcgc cgccaccatg gcccagtcca agcacggcct gaccaaggag atgaccatga 4680

agtaccgcat ggagggctgc gtggacggcc acaagttcgt gatcaccggc gagggcatcg 4740

gctacccctt caagggcaag caggccatca acctgtgcgt ggtggagggc ggcccccttgc 4800

ccttcgccga ggacatcttg tccgccgcct tcatgtacgg caaccgcgtg ttcaccgagt 4860

acccccagga catcgtcgac tacttcaaga actcctgccc cgccggctac acctgggacc 4920

gctccttcct gttcgaggac ggcgccgtgt gcatctgcaa cgccgacatc accgtgagcg 4980

tggaggagaa ctgcatgtac cacgagtcca agttctacgg cgtgaacttc cccgccgacg 5040

gccccgtgat gaagaagatg accgacaact gggagccctc ctgcgagaag atcatccccg 5100

tgcccaagca gggcatcttg aagggcgacg tgagcatgta cctgctgctg aaggacggtg 5160

gccgcttgcg ctgccagttc gacaccgtgt acaaggccaa gtccgtgccc cgcaagatgc 5220

ccgactggca cttcatccag cacaagctga cccgcgagga ccgcagcgac gccaagaacc 5280

agaagtggca cctgaccgag cacgccatcg cctccggctc cgccttgccc tgagtcagag 5340

ctcgctgatc agcctcgact gtgccttcta gttgccagcc atctgttgtt tgcccctccc 5400
ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt cctttcctaa taaaatgagg 5460
aaattgcatc gcattgtctg agtaggtgtc attctattct gggggggtggg gtggggcagg 5520
acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg gtgggctcta 5580
tggcttctga ggcggaaaga accagctggg gctcgagatc cactagttct agcctcgagg 5640
ctagagcggc cgccactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt 5700
acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa tagcgaagag 5760

gcccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg ggacgcgccc 5820
tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt 5880
gccagcgccc tagcgcccgc tcctttcgct ttcttcccct cctttctcgc cacgttcgcc 5940
ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta 6000
cggcacctcg accccaaaaa acttgattag ggtgatggtt cacgtagtgg gccatcgccc 6060
tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg 6120
ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt 6180

ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat 6240
tttaacaaaa tattaacgct tacaatttag gtggcacttt tcggggaaat gtgcgcggaa 6300
cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 6360
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg 6420
tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc 6480
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 6540
atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga 6600

gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc 6660
aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtactca ccagtcacag 6720
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc ataaccatga 6780
gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg 6840
ctttttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa ccggagctga 6900
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt 6960
tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 7020

ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt 7080

ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 7140

ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 7200

tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac 7260

tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 7320

aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt 7380

tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt 7440

ttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt 7500

gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 7560

agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 7620

tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg 7680

ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggt 7740

cgggctgaac gggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac 7800

tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg 7860

acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgagggag cttccagggg 7920

gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat 7980

ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt 8040

tacggttcct ggccttttgc tggccttttg ctcacatgtt ctttcctgcg ttatcccctg 8100

attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa 8160

cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc 8220

ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggttt cccgactgga 8280

aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccagg 8340

ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc 8400

acacaggaaa cagctatgac catgaggcgc gccg 8434

<210> 41

<211> 9404

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Human CD133 endogenous HDR repoter vector (RFP)

<400> 41

gattcgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 60

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 120

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 180

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 240

acatcaagtg tatcatatgt tcagtgtccc ctccaaaact catgttgaga tttaattgcc 300

attgtgacta tattaacagg cgagaccttg gacaggtggt taagtcatga gagttctgcc 360

ttcataaatg gattaatgcc attatcatgg gagtgggctc ctgataaaag gaggaagctc 420

gaccccattc cctgtctgtc tcttgcactc atttgccctt ccattatggg atgatgcgtc 480

atgaaagccc tcaccagatg ccagcaccag gttcttggac ttcgtggcct tcagaactgt 540

gaaaaacata ttttttttctt tgtaagttac ccagtctctg atattctgtt atagcagcag 600

aaaacatact aagacactgc ttaaattggc ttgttttgct catttcctgc ttctgaggtc 660

tccagagttc cctctcttcc aggataggat ggatatactg tctggaggag aggcaatgct 720

aggcaggagg agaagaacca gtgatcctgt tagtgtgccc agcatccctc tgctcaataa 780

ggcaaatgta gatgttcttt ggaaatgatt gccagagcct acagtccttc ataatgttca 840

gctccaggtc ctggtgaggt ggacagtaac aagagattga gccagcccct gcttgcaccc 900

ttttccaggg gccctttgag ccctttcgat aaagagtttg tttctttacc gaaagcataa 960

aaggtagtgc aagaagaggc acacagaacc aagcatgtgc actttctaaa gtgaaaaact 1020

ttagaaggaa tgtcttgcct ggcagatgcc ctgatgtcct ccgtgcatgt cctcccccaa 1080

cctcatgtgt ctgatgcagc actgcagtga ggcacggtgg atcctgcagg ttcagagggg 1140

ttgtgatcca cctgaagtca cagtcgggca gggtgtgcca gagttgaaac tgccgacctt 1200

ccagatggag ttcactagtc tccccctggc gtataccttt ttggttgtca aggccattat 1260

cacagagtca ctttagagga cactagggct actgttgatg gttcacccag agagccactt 1320

agaggccgtg tgcactgcct gatcttacac ccaggaattt gtagttgtaa caaggtccct 1380

gcggaacttc catttgcttc tgtcttctgt cctagaagtg gactttagaa agactcttgt 1440

ggtagacatg tgtactgctg tgttgctgct ggtctaatgt cttcaacacg tattcttagt 1500

gactcttaac acttttaacc ttagattgtg aatattcaca attcttactt gtcatagtct 1560

gtttatcctc tttatcttct atgctgtttt taaaagcaat acaaacagaa gctagagtaa 1620

ccagtgtttt gtctgacaca ttcttcttga taaactttat gaggtgtggt ttatattcgt 1680

ctgtaccttt agtcacatgc ctgcttctta gcacagaggt gatgaaacca cccacagctt 1740

tgttgggggt attatatttc tttatttttt tgtttttttag cccatcacaa catggtaccg 1800

gaagcggagc tactaacttc agcctgctga agcaggctgg cgacgtggag gagaaccctg 1860

gacctctcga gtcgcgaacg cgtgccacca tggtgtctaa gggcgaagag ctgattaagg 1920

agaacatgca catgaagctg tacatggagg gcaccgtgaa caaccaccac ttcaagtgca 1980

catccgaggg cgaaggcaag ccctacgagg gcacccagac catgagaatc aaggtggtcg 2040

agggcggccc tctccccttc gccttcgaca tcctggctac cagcttcatg tacggcagca 2100

gaaccttcat caaccacacc cagggcatcc ccgacttctt taagcagtcc ttccctgagg 2160

gcttcacatg ggagagagtc accacatacg aggacggggg cgtgctgacc gctacccagg 2220

acaccagcct ccaggacggc tgcctcatct acaacgtcaa gatcagaggg gtgaacttcc 2280

catccaacgg ccctgtgatg cagaagaaaa cactcggctg ggaggccaac accgagatgc 2340

tgtaccccgc tgacggcggc ctggaaggca gaagcgacat ggccctgaag ctcgtgggcg 2400

ggggccacct gatctgcaac ttcaagacca catacagatc caagaaaccc gctaagaacc 2460

tcaagatgcc cggcgtctac tatgtggacc acagactgga aagaatcaag gaggccgaca 2520

aagagaccta cgtcgagcag cacgaggtgg ctgtggccag atactgcgac ctccctagca 2580

aactggggca caaacttaat acgcgtggct cgggccagtg tactaattat gctctcttga 2640

aattggctgg agatgttgag agcaacccag gtcccgagct cgcggccgca gatctttttc 2700

cctctgccaa aaattatggg gacatcatga agccccttga gcatctgact tctggcgagc 2760

ttgccaccat ggtgtctaag ggcgaagagc tgattaagga gaacatgcac atgaagctgt 2820

acatggaggg caccgtgaac aaccaccact tcaagtgcac atccgagggc gaaggcaagc 2880

cctacgaggg cacccagacc atgagaatca aggtggtcga gggcggccct ctcccttcg 2940

ccttcgacat cctggctacc agcttcatgt acggcagcag aaccttcatc aaccacaccc 3000

agggcatccc cgacttcttt aagcagtcct tccctgaggg cttcacatgg gagagagtca 3060

ccacatacga ggacgggggc gtgctgaccg ctacccagga caccagcctc caggacggct 3120

gcctcatcta caacgtcaag atcagagggg tgaacttccc atccaacggc cctgtgatgc 3180

agaagaaaac actcggctgg gaggccaaca ccgagatgct gtaccccgct gacggcggcc 3240

tggaaggcag aagcgacatg gccctgaagc tcgtgggcgg gggccacctg atctgcaact 3300

tcaagaccac atacagatcc aagaaacccg ctaagaacct caagatgccc ggcgtctact 3360

atgtggacca cagactggaa agaatcaagg aggccgacaa agagacctac gtcgagcagc 3420

acgaggtggc tgtggccaga tactgcgacc tccctagcaa actggggcac aaacttaatg 3480

gctccgaggg cagaggaagc cttctaacat gcggtgacgt ggaggagaat cccggccctt 3540

ccgggatgac cgagtacaag cccacggtgc gcctcgccac ccgcgacgac gtccccaggg 3600

ccgtacgcac cctcgccgcc gcgttcgccg actaccccgc cacgcgccac accgtcgatc 3660

cagaccgcca catcgagcgg gtcaccgagc tgcaagaact cttcctcacg cgcgtcgggc 3720

tcgacatcgg caaggtgtgg gtcgcggacg acggcgccgc ggtggcggtc tggaccacgc 3780

cggagagcgt cgaagcgggg gcggtgttcg ccgagatcgg cccgcgcatg gccgagttga 3840

gcggttcccg gctggccgcg cagcaacaga tggaaggtct cctggcgccg caccggccca 3900

aggagcccgc gtggttcctg gccaccgtcg gcgtctcgcc cgaccaccag ggcaagggtc 3960

tgggcagcgc cgtcgtgctc cccggagtgg aggcggccga gcgcgccggg gtgcccgcct 4020

tcctggagac ctccgcgccc cgcaacctcc ccttctacga gcggctcggc ttcaccgtca 4080

ccgccgacgt cgaggtgccc gaaggaccgc gcacctggtg catgacccgc aagcccggtg 4140

cctgacatgc atctagaggg ccctattcta tagtgtcacc taaatgctag agctcgctga 4200

tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgtgcct 4260

tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca 4320

tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtggggca ggacagcaag 4380

ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtgggctc tatggggatc 4440

cccatcaagc tgatccggaa cccttaatat aacttcgtat aatgtatgct atacgaagtt 4500

attaggtccc tcgacctgca gcccaagcta gatcgaattc tagctgatgt tgaaactgct 4560

tgaggtaagt ttgatgcttt cattgaagag aatctacagg actatagttt ttagttttta 4620

gaacatcaaa acatagctat tcctaataca ttgccacatt aagtgggaaa tctagatcag 4680

tatattagac aagatatttt atatgtccag tccaccctct gtacctgcag gttccacatc 4740

caccgattga ttcaaccaac tgtggatcaa aaatactcta aaaaaatttt ttttaaataa 4800

cagtacaaaa ataaaaataa taaaaacaat acagtataac agctacttac ataacactta 4860

cattgtatta ggtattttga gtgatctaga gatgatatac agtatgtagg aggatgtgtg 4920

tagcggttct ggaaccaatc ccacatggat ccagagggac gactgcatgt atacatacga 4980

tacagtgacc ctctctgtat atatatgcat tgctatataa atgtgtgtgt atttatatat 5040

atatatacac acacacagtg agtatgtggg tgtatattaa aaaaataaaa tttttttact 5100

taaaaagata agattttta agtaaaaaaa taagggctgg tgcggtggct cacgcctgta 5160

atcccagcac tatgggaggc cgaggcgggc agatcacttg aggtcaggag ctcaaaacca 5220

gcctggccaa catggtgaaa ccctgtctct actaaaaaaa gttagccggg tgtggtgggg 5280

ggcccctgta atcctgcctg taatctcagc tactcaggag gccaaggcag gagaatcgct 5340

tcaacctggg aggcagaggt tgcagtgagc cgagatcaca ccactgcact ccagcgtggg 5400

caacagagtg agactccatc acaaaaaaaa aaaaaaagaa aagaaaagaa agtgtttttt 5460

tagagtaaaa caataagcca ggcatggtgg tacctgcctg tagtcccagc tactggggtg 5520

ggaggctgtg gtagaaggat tgcttgagcc caggagttca aagatgcagt gaaccatgat 5580

tatacctgtg aatagccaac acactcatgc actccaactt ggaaaatgta gtgaaaccct 5640
atctctaaac aaacaaacaa acaaacaaaa atatagaaat ctgtgtacca accacgtacc 5700
ttgagtatta tggacttaaa atttttttctg gctgggcatg gtagctcatg cctgtaatcc 5760
cagcactttg ggaagctgag gtgggcagat cacttggggt taggagttcg agaccagcct 5820
ggccaacatg gagaaacccc atctctacta aaaaatataa aaattagcca ggcatggtgg 5880
caggtgcctg tagtcccagc tacttgggag gctgaggctt aaacccggga gacggaggtt 5940
gcagtgagct gagatcgcgc cattgcactc cagcctgggc gacagagtga gactccctct 6000

caaaaaaata aataaataaa ataagataaa taaacaaaaa ccacaaagtt tttttctata 6060
ttgctatatt ttaacagatt gctttgcaac aaacatattg ctgtgcctct tgaacgagtt 6120
tatattttcc ttagttctta ggctcctggt ggggaccatt ttaggattat gcattcctgg 6180
ttttctcttt attaattcaa aagcaatttc taaaattgtg tttaactttc atttcagcat 6240
caggatactc aaagtggaaa ggatcacaga tttttggtag tttctgggtc tacaaggact 6300
ttccaaatcc aggagcaacg ccagtggcaa cgtagtgact caggcgggca ccaaggcaac 6360
ggcaccattg gtctctgggt agtgctttaa gaatgaacac aatcacgtta tagtccatgg 6420

tccatcacta ttcaaggatg actccctccc ttcctgtcta tttttgtttt ttacttttttt 6480
acactgagtt tctatttaga cactacaaca tatggggtgt ttgttcccat tggatgcatt 6540
tctatcaaaa ctctatcaaa tgtgatggct agattctaac atattgccat gtgtggagtg 6600
tgcctcgagg ctagagcggc cgccactggc cgtcgtttta caacgtcgtg actgggaaaa 6660
ccctggcgtt acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa 6720
tagcgaagag gcccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg 6780
ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 6840

cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttcccttt cctttctcgc 6900
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 6960
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatggtt cacgtagtgg 7020
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag 7080
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt 7140
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt 7200
taacgcgaat tttaacaaaa tattaacgct tacaatttag gtggcacttt tcggggaaat 7260

gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg 7320

agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa 7380

catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac 7440

ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac 7500

atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt 7560

ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc 7620

gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtactca 7680

ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc 7740

ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag 7800

gagctaaccg ctttttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa 7860

ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg 7920

gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa 7980

ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg 8040

gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt 8100

gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 8160

caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag 8220

cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat 8280

ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 8340

taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct 8400

tgagatcctt ttttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca 8460

gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc 8520

agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc 8580

aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct 8640

gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag 8700

gcgcagcggt cgggctgaac gggggttcg tgcacacagc ccagcttgga gcgaacgacc 8760

tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg 8820

agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgagggag 8880

cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt 8940

gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac 9000

gcggcctttt tacggttcct ggccttttgc tggccttttg ctcacatgtt ctttcctgcg 9060

ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc 9120

cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcccaata 9180

cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggttt 9240

cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag 9300

gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga 9360

taacaatttc acacaggaaa cagctatgac catgaggcgc gccg 9404

<210> 42

<211> 6888

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Loxp-TP2A-Tag RFP 2X-T2A-puro-polyA in pCAGGS

<400> 42

gattcgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 60

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 120

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 180

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 240

acatcaagtg tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc 300

cgcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta 360

cgtattagtc atcgctatta ccatggtcga ggtgagcccc acgttctgct tcactctccc 420

catctccccc ccctccccac ccccaatttt gtatttattt atttttaat tattttgtgc 480

agcgatgggg gcgggggggg gggggggcg cgcgccaggc ggggcggggc ggggcgaggg 540

gcggggcggg gcgaggcgga gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa 600

gtttcctttt atggcgaggc ggcggcggcg gcggccctat aaaaagcgaa gcgcgcggcg 660

ggcgggagtc gctgcgttgc cttcgccccg tgccccgctc cgcgccgcct cgcgccgccc 720

gccccggctc tgactgaccg cgttactccc acaggtgagc gggcgggacg gcccttctcc 780

tccgggctgt aattagcgct tggtttaatg acggctcgtt tcttttctgt ggctgcgtga 840

aagccttaaa gggctccggg agggcccttt gtgcgggggg gagcggctcg gggggtgcgt 900

gcgtgtgtgt gtgcgtgggg agcgccgcgt gcggcccgcg ctgcccggcg gctgtgagcg 960

ctgcgggcgc ggcgcggggc tttgtgcgct ccgcgtgtgc gcgaggggag cgcggccggg 1020

ggcggtgccc cgcggtgcgg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt 1080

gcgtgggggg gtgagcaggg ggtgtgggcg cggcggtcgg gctgtaaccc ccccctgcac 1140

ccccctcccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tgcggggcgt 1200

ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg 1260

gggccgcctc gggccgggga gggctcgggg gaggggcgcg gcggccccgg agcgccggcg 1320

gctgtcgagg cgcggcgagc cgcagccatt gccttttatg gtaatcgtgc gagagggcgc 1380

agggacttcc tttgtcccaa atctggcgga gccgaaatct gggaggcgcc gccgcacccc 1440

ctctagcggg cgcgggcgaa gcggtgcggc gccggcagga aggaaatggg cggggagggc 1500

cttcgtgcgt cgccgcgccg ccgtcccctt ctccatctcc agcctcgggg ctgccgcagg 1560

gggacggctg ccttcggggg ggacggggca gggcggggtt cggcttctgg cgtgtgaccg 1620

gcggctctag agcctctgct aaccatgttc atgccttctt ctttttccta cagatcctta 1680

attaatcgcc gccgcgttcg ccgacatcga tggatctgga gaaggacgcg gcagcctact 1740

gacttgcgga gatgtcgaag agaaccctgg ccctaagctt ataacttcgt ataatgtatg 1800

ctatacgaag ttattggtcg acgctagcgg taccggaagc ggagctacta acttcagcct 1860

gctgaagcag gctggcgacg tggaggagaa ccctggacct ctcgagtcgc gaacgcgtgc 1920

caccatggtg tctaagggcg aagagctgat taaggagaac atgcacatga agctgtacat 1980

ggagggcacc gtgaacaacc accacttcaa gtgcacatcc gagggcgaag gcaagcccta 2040

cgagggcacc cagaccatga gaatcaaggt ggtcgagggc ggccctctcc ccttcgcctt 2100

cgacatcctg gctaccagct tcatgtacgg cagcagaacc ttcatcaacc acacccaggg 2160

catccccgac ttctttaagc agtccttccc tgagggcttc acatgggaga gagtcaccac 2220

atacgaggac gggggcgtgc tgaccgctac ccaggacacc agcctccagg acggctgcct 2280

catctacaac gtcaagatca gaggggtgaa cttcccatcc aacggccctg tgatgcagaa 2340

gaaaacactc ggctgggagg ccaacaccga gatgctgtac cccgctgacg gcggcctgga 2400

aggcagaagc gacatggccc tgaagctcgt gggcgggggc cacctgatct gcaacttcaa 2460

gaccacatac agatccaaga aacccgctaa gaacctcaag atgcccggcg tctactatgt 2520

ggaccacaga ctggaaagaa tcaaggaggc cgacaaagag acctacgtcg agcagcacga 2580

ggtggctgtg gccagatact gcgacctccc tagcaaactg gggcacaaac ttaatacgcg 2640

tggctcgggc cagtgtacta attatgctct cttgaaattg gctggagatg ttgagagcaa 2700

cccaggtccc gagctcgcgg ccgcagatct ttttccctct gccaaaaatt atggggacat 2760

catgaagccc cttgagcatc tgacttctgg cgagcttgcc accatggtgt ctaagggcga 2820

agagctgatt aaggagaaca tgcacatgaa gctgtacatg gagggcaccg tgaacaacca 2880

ccacttcaag tgcacatccg agggcgaagg caagccctac gagggcaccc agaccatgag 2940

aatcaaggtg gtcgagggcg gccctctccc cttcgccttc gacatcctgg ctaccagctt 3000

catgtacggc agcagaacct tcatcaacca cacccagggc atccccgact tctttaagca 3060

gtccttccct gagggcttca catgggagag agtcaccaca tacgaggacg ggggcgtgct 3120

gaccgctacc caggacacca gcctccagga cggctgcctc atctacaacg tcaagatcag 3180

aggggtgaac ttcccatcca acggccctgt gatgcagaag aaaacactcg gctgggaggc 3240

caacaccgag atgctgtacc ccgctgacgg cggcctggaa ggcagaagcg acatggccct 3300

gaagctcgtg ggcgggggcc acctgatctg caacttcaag accacataca gatccaagaa 3360

acccgctaag aacctcaaga tgcccggcgt ctactatgtg gaccacagac tggaaagaat 3420

caaggaggcc gacaaagaga cctacgtcga gcagcacgag gtggctgtgg ccagatactg 3480

cgacctccct agcaaactgg ggcacaaact taatggctcc gagggcagag gaagccttct 3540

aacatgcggt gacgtggagg agaatcccgg cccttccggg atgaccgagt acaagcccac 3600

ggtgcgcctc gccacccgcg acgacgtccc cagggccgta cgcaccctcg ccgccgcgtt 3660

cgccgactac cccgccacgc gccacaccgt cgatccagac cgccacatcg agcgggtcac 3720

cgagctgcaa gaactcttcc tcacgcgcgt cgggctcgac atcggcaagg tgtgggtcgc 3780

ggacgacggc gccgcggtgg cggtctggac cacgccggag agcgtcgaag cgggggcggt 3840

gttcgccgag atcggcccgc gcatggccga gttgagcggt tcccggctgg ccgcgcagca 3900

acagatggaa ggtctcctgg cgccgcaccg gcccaaggag cccgcgtggt tcctggccac 3960

cgtcggcgtc tcgcccgacc accagggcaa gggtctgggc agcgccgtcg tgctccccgg 4020

agtggaggcg gccgagcgcg ccggggtgcc cgccttcctg gagacctccg cgccccgcaa 4080

cctccccttc tacgagcggc tcggcttcac cgtcaccgcc gacgtcgagg tgcccgaagg 4140

accgcgcacc tggtgcatga cccgcaagcc cggtgcctga catgcatcta gagggcccta 4200

ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg ccttctagtt 4260

gccagccatc tgttgtttgc ccctcccccg tgccttcctt gaccctggaa ggtgccactc 4320

ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 4380

ctattctggg gggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 4440

ggcatgctgg ggatgcggtg ggctctatgg ggatccataa cttcgtataa agtatcctat 4500

acgaagttat ctgcagcata gaattcagcg atatcactcg agggttcctg gaagtttaga 4560

tcagggctca ctgactcgct gcgccaaaat aggttagcct tgcggggtcg gtcgttcggc 4620

tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 4680

ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 4740

ccgcgttgct ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac 4800

gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg 4860

gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 4920
ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 4980
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 5040
gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 5100
tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 5160
tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc 5220
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 5280

ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 5340
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 5400
gttaagggat tttggtcatg agattatcaa aaaggagttc ctggaagttt agatcagggt 5460
cttcacctag atccttttac aaaataggtt agccttgcgg ggaattaaaa atgaagtttt 5520
aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 5580
gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc 5640
gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg 5700

cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 5760
gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 5820
gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca 5880
ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga 5940
tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 6000
ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 6060
cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 6120

accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 6180
cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 6240
tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact 6300
cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa 6360
acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc 6420
atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 6480
tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgaca tgttctttcc 6540

tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag ctgataccgc 6600

tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc 6660

aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag 6720

gtttcccgac tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt agctcactca 6780

ttaggcaccc caggctttac actttatgct tccggctcgt atgttgtgtg gaattgtgag 6840

cggataacaa tttcacacag gaaacagcta tgaccatgag gcgcgccg 6888

<210> 43

<211> 1534

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for 5-homology arm

<400> 43

ttcagtgtcc cctccaaaac tcatgttgag atttaattgc cattgtgact atattaacag 60

gcgagacctt ggacaggtgg ttaagtcatg agagttctgc cttcataaat ggattaatgc 120

cattatcatg ggagtgggct cctgataaaa ggaggaagct cgaccccatt ccctgtctgt 180

ctcttgcact catttgccct tccattatgg gatgatgcgt catgaaagcc ctcaccagat 240

gccagcacca ggttcttgga cttcgtggcc ttcagaactg tgaaaaacat atttttttct 300

ttgtaagtta cccagtctct gatattctgt tatagcagca gaaaacatac taagacactg 360

cttaaattgg cttgttttgc tcatttcctg cttctgaggt ctccagagtt ccctctcttc 420

caggatagga tggatatact gtctggagga gaggcaatgc taggcaggag gagaagaacc 480

agtgatcctg ttagtgtgcc cagcatccct ctgctcaata aggcaaatgt agatgttctt 540

tggaaatgat tgccagagcc tacagtcctt cataatgttc agctccaggt cctggtgagg 600

tggacagtaa caagagattg agccagcccc tgcttgcacc cttttccagg ggccctttga 660

gccctttcga taaagagttt gtttctttac cgaaagcata aaaggtagtg caagaagagg 720

cacacagaac caagcatgtg cactttctaa agtgaaaaac tttagaagga atgtcttgcc 780

tggcagatgc cctgatgtcc tccgtgcatg tcctccccca acctcatgtg tctgatgcag 840

cactgcagtg aggcacggtg gatcctgcag gttcagaggg gttgtgatcc acctgaagtc 900

acagtcgggc agggtgtgcc agagttgaaa ctgccgacct tccagatgga gttcactagt 960

ctccccctgg cgtatacctt tttggttgtc aaggccatta tcacagagtc actttagagg 1020

acactagggc tactgttgat ggttcaccca gagagccact tagaggccgt gtgcactgcc 1080

tgatcttaca cccaggaatt tgtagttgta acaaggtccc tgcggaactt ccatttgctt 1140

ctgtcttctg tcctagaagt ggactttaga aagactcttg tggtagacat gtgtactgct 1200

gtgttgctgc tggtctaatg tcttcaacac gtattcttag tgactcttaa cacttttaac 1260

cttagattgt gaatattcac aattcttact tgtcatagtc tgtttatcct ctttatcttc 1320

tatgctgttt ttaaaagcaa tacaaacaga agctagagta accagtgttt tgtctgacac 1380

attcttcttg ataaacttta tgaggtgtgg tttatattcg tctgtacctt tagtcacatg 1440

cctgcttctt agcacagagg tgatgaaacc acccacagct ttgttggggg tattatattt 1500

ctttattttt ttgtttttta gcccatcaca acat 1534

<210> 44

<211> 2063

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for 3-homology arm

<400> 44

tagctgatgt tgaaactgct tgaggtaagt ttgatgcttt cattgaagag aatctacagg 60

actatagttt ttagttttta gaacatcaaa acatagctat tcctaataca ttgccacatt 120

aagtgggaaa tctagatcag tatattagac aagatatttt atatgtccag tccaccctct 180

gtacctgcag gttccacatc caccgattga ttcaaccaac tgtggatcaa aaatactcta 240

aaaaaatttt tttttaaata acagtacaaa aataaaaata ataaaaacaa tacagtataa 300

cagctactta cataacactt acattgtatt aggtattttg agtgatctag agatgatata 360

cagtatgtag gaggatgtgt gtagcggttc tggaaccaat cccacatgga tccagaggga 420

cgactgcatg tatacatacg atacagtgac cctctctgta tatatatgca ttgctatata 480

aatgtgtgtg tatttatata tatatataca cacacacagt gagtatgtgg gtgtatatta 540

aaaaaataaa atttttttac ttaaaaagat aagatttttt aagtaaaaaa ataagggctg 600

gtgcggtggc tcacgcctgt aatcccagca ctatgggagg ccgaggcggg cagatcactt 660

gaggtcagga gctcaaaacc agcctggcca acatggtgaa accctgtctc tactaaaaaa 720

agttagccgg gtgtggtggg gggcccctgt aatcctgcct gtaatctcag ctactcagga 780

ggccaaggca ggagaatcgc ttcaacctgg gaggcagagg ttgcagtgag ccgagatcac 840

accactgcac tccagcgtgg gcaacagagt gagactccat cacaaaaaaa aaaaaaaaga 900

aaagaaaaga aagtgttttt ttagagtaaa acaataagcc aggcatggtg gtacctgcct 960

gtagtcccag ctactggggt gggaggctgt ggtagaagga ttgcttgagc ccaggagttc 1020

aaagatgcag tgaaccatga ttatacctgt gaatagccaa cacactcatg cactccaact 1080

tggaaaatgt agtgaaaccc tatctctaaa caaacaaaca aacaaacaaa aatatagaaa 1140

tctgtgtacc aaccacgtac cttgagtatt atggacttaa aatttttttct ggctgggcat 1200

ggtagctcat gcctgtaatc ccagcacttt gggaagctga ggtgggcaga tcacttgggg 1260

ttaggagttc gagaccagcc tggccaacat ggagaaaccc catctctact aaaaaatata 1320

aaaattagcc aggcatggtg gcaggtgcct gtagtcccag ctacttggga ggctgaggct 1380

taaacccggg agacggaggt tgcagtgagc tgagatcgcg ccattgcact ccagcctggg 1440

cgacagagtg agactccctc tcaaaaaaat aaataaataa aataagataa ataaacaaaa 1500

accacaaagt ttttttctat attgctatat tttaacagat tgctttgcaa caaacatatt 1560

gctgtgcctc ttgaacgagt ttatattttc cttagttctt aggctcctgg tggggaccat 1620

tttaggatta tgcattcctg gttttctctt tattaattca aaagcaattt ctaaaattgt 1680

gtttaacttt catttcagca tcaggatact caaagtggaa aggatcacag atttttggta 1740

gtttctgggt ctacaaggac tttccaaatc caggagcaac gccagtggca acgtagtgac 1800

tcaggcgggc accaaggcaa cggcaccatt ggtctctggg tagtgcttta agaatgaaca 1860

caatcacgtt atagtccatg gtccatcact attcaaggat gactccctcc cttcctgtct 1920

atttttgttt tttacttttt tacactgagt ttctatttag acactacaac atatggggtg 1980

tttgttccca ttggatgcat ttctatcaaa actctatcaa atgtgatggc tagattctaa 2040

catattgcca tgtgtggagt gtg 2063

<210> 45

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for CD133 sgRNA

<400> 45

tcagctatca atgttgtgat 20

<210> 46

<211> 4101

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for SpCas9

<400> 46

gacaagaagt acagcatcgg cctggacatc ggcaccaact ctgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc cagcaagaaa ttcaaggtgc tgggcaacac cgaccggcac 120

agcatcaaga agaacctgat cggagccctg ctgttcgaca gcggcgaaac agccgaggcc 180

acccggctga agagaaccgc cagaagaaga tacaccagac ggaagaaccg gatctgctat 240

ctgcaagaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccacagactg 300

gaagagtcct tcctggtgga agaggataag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gagaaagaaa 420

ctggtggaca gcaccgacaa ggccgacctg cggctgatct atctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa cagcgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga aaaccccatc 600

aacgccagcg gcgtggacgc caaggccatc ctgtctgcca gactgagcaa gagcagacgg 660

ctggaaaatc tgatcgccca gctgcccggc gagaagaaga atggcctgtt cggaaacctg 720

attgccctga gcctgggcct gaccccaac ttcaagagca acttcgacct ggccgaggat 780

gccaaactgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgtttctg gccgccaaga acctgtccga cgccatcctg 900

ctgagcgaca tcctgagagt gaacaccgag atcaccaagg ccccccctgag cgcctctatg 960

atcaagagat acgacgagca ccaccaggac ctgaccctgc tgaaagctct cgtgcggcag 1020

cagctgcctg agaagtacaa agagattttc ttcgaccaga gcaagaacgg ctacgccggc 1080

tacattgacg gcggagccag ccaggaagag ttctacaagt tcatcaagcc catcctggaa 1140

aagatggacg gcaccgagga actgctcgtg aagctgaaca gagaggacct gctgcggaag 1200

cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggaga gctgcacgcc 1260

attctgcggc ggcaggaaga tttttaccca ttcctgaagg acaaccggga aaagatcgag 1320

aagatcctga ccttccgcat cccctactac gtgggccctc tggccagggg aaacagcaga 1380

ttcgcctgga tgaccagaaa gagcgaggaa accatcaccc cctggaactt cgaggaagtg 1440

gtggacaagg gcgcttccgc ccagagcttc atcgagcgga tgaccaactt cgataagaac 1500

ctgcccaacg agaaggtgct gcccaagcac agcctgctgt acgagtactt caccgtgtat 1560

aacgagctga ccaaagtgaa atacgtgacc gagggaatga gaaagcccgc cttcctgagc 1620

ggcgagcaga aaaaggccat cgtggacctg ctgttcaaga ccaaccggaa agtgaccgtg 1680

aagcagctga aagaggacta cttcaagaaa atcgagtgct tcgactccgt ggaaatctcc 1740

ggcgtggaag atcggttcaa cgcctccctg ggcacatacc acgatctgct gaaaattatc 1800

aaggacaagg acttcctgga caatgaggaa aacgaggaca ttctggaaga tatcgtgctg 1860

accctgacac tgtttgagga cagagagatg atcgaggaac ggctgaaaac ctatgcccac 1920

ctgttcgacg acaaagtgat gaagcagctg aagcggcgga gatacaccgg ctggggcagg 1980

ctgagccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac aatcctggat 2040

ttcctgaagt ccgacggctt cgccaacaga aacttcatgc agctgatcca cgacgacagc 2100

ctgaccttta aagaggacat ccagaaagcc caggtgtccg gccagggcga tagcctgcac 2160

gagcacattg ccaatctggc cggcagcccc gccattaaga agggcatcct gcagacagtg 2220

aaggtggtgg acgagctcgt gaaagtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gaaatggcca gagagaacca gaccacccag aagggacaga agaacagccg cgagagaatg 2340

aagcggatcg aagagggcat caaagagctg ggcagccaga tcctgaaaga acaccccgtg 2400

gaaaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa tgggcgggat 2460

atgtacgtgg accaggaact ggacatcaac cggctgtccg actacgatgt ggaccatatc 2520

gtgcctcaga gctttctgaa ggacgactcc atcgacaaca aggtgctgac cagaagcgac 2580

aagaaccggg gcaagagcga caacgtgccc tccgaagagg tcgtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg attacccaga gaaagttcga caatctgacc 2700

aaggccgaga gaggcggcct gagcgaactg gataaggccg gcttcatcaa gagacagctg 2760

gtggaaaccc ggcagatcac aaagcacgtg gcacagatcc tggactcccg gatgaacact 2820

aagtacgacg agaatgacaa gctgatccgg gaagtgaaag tgatcaccct gaagtccaag 2880

ctggtgtccg atttccggaa ggatttccag ttttacaaag tgcgcgagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtcgtgggaa ccgccctgat caaaaagtac 3000

cctaagctgg aaagcgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaaga gcgagcagga aatcggcaag gctaccgcca agtacttctt ctacagcaac 3120

atcatgaact ttttcaagac cgagattacc ctggccaacg gcgagatccg gaagcggcct 3180

ctgatcgaga caaacggcga aaccggggag atcgtgtggg ataagggccg ggattttgcc 3240

accgtgcgga aagtgctgag catgccccaa gtgaatatcg tgaaaaagac cgaggtgcag 3300

acaggcggct tcagcaaaga gtctatcctg cccaagagga acagcgataa gctgatcgcc 3360

agaaagaagg actgggaccc taagaagtac ggcggcttcg acagccccac cgtggcctat 3420

tctgtgctgg tggtggccaa agtggaaaag ggcaagtcca agaaactgaa gagtgtgaaa 3480

gagctgctgg ggatcaccat catggaaaga agcagcttcg agaagaatcc catcgacttt 3540

ctggaagcca agggctacaa agaagtgaaa aaggacctga tcatcaagct gcctaagtac 3600

tccctgttcg agctggaaaa cggccggaag agaatgctgg cctctgccgg cgaactgcag 3660

aagggaaacg aactggccct gccctccaaa tatgtgaact tcctgtacct ggccagccac 3720

tatgagaagc tgaagggctc ccccgaggat aatgagcaga aacagctgtt tgtggaacag 3780

cacaagcact acctggacga gatcatcgag cagatcagcg agttctccaa gagagtgatc 3840

ctggccgacg ctaatctgga caaagtgctg tccgcctaca acaagcaccg ggataagccc 3900

atcagagagc aggccgagaa tatcatccac ctgtttaccc tgaccaatct gggagccccc 3960

gccgccttca agtactttga caccaccatc gaccggaaga ggtacaccag caccaaagag 4020

gtgctggacg ccaccctgat ccaccagagc atcaccggcc tgtacgagac acggatcgac 4080

ctgtctcagc tgggaggcga c 4101

<210> 47

<211> 711

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for TagRFP

<400> 47

atggtgtcta agggcgaaga gctgattaag gagaacatgc acatgaagct gtacatggag 60

ggcaccgtga acaaccacca cttcaagtgc acatccgagg gcgaaggcaa gccctacgag 120

ggcacccaga ccatgagaat caaggtggtc gagggcggcc ctctcccctt cgccttcgac 180

atcctggcta ccagcttcat gtacggcagc agaaccttca tcaaccacac ccagggcatc 240

cccgacttct ttaagcagtc cttccctgag ggcttcacat gggagagagt caccacatac 300

gaggacgggg gcgtgctgac cgctacccag gacaccagcc tccaggacgg ctgcctcatc 360

tacaacgtca agatcagagg ggtgaacttc ccatccaacg gccctgtgat gcagaagaaa 420

acactcggct gggaggccaa caccgagatg ctgtaccccg ctgacggcgg cctggaaggc 480

agaagcgaca tggccctgaa gctcgtgggc gggggccacc tgatctgcaa cttcaagacc 540

acatacagat ccaagaaacc cgctaagaac ctcaagatgc ccggcgtcta ctatgtggac 600

cacagactgg aaagaatcaa ggaggccgac aaagagacct acgtcgagca gcacgaggtg 660

gctgtggcca gatactgcga cctccctagc aaactggggc acaaacttaa t 711

<210> 48

<211> 717

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for TagYFP

<400> 48

atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60

ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120

ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180

ctcgtgacca ccttcggcta cggcctgcag tgcttcgccc gctaccccga ccacatgaag 240

cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300

ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360

gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420

aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480

ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540

gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600

tacctgagct accagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660

ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaag 717

<210> 49

<211> 717

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for TagCFP

<400> 49

atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60

ggcgacgtaa acggccacag gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120

ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180

ctcgtgacca ccctgacctg gggcgtgcag tgcttcagcc gctaccccga ccacatgaag 240

cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg taccatcttc 300

ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360

gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420

aagctggagt acaactacat cagccacaac gtctatatca ccgccgacaa gcagaagaac 480

ggcatcaagg cccacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540

gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600

tacctgagca cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660

ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaag 717

<210> 50

<211> 717

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for TagEGFP

<400> 50

atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60

ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120

ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180

ctcgtgacca ccctgaccta cggcgtgcag tgcttcagcc gctaccccga ccacatgaag 240

cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300

ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360

gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420

aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480

ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540

gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600

tacctgagca cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660

ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaag 717

<210> 51

<211> 600

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for PuroR

<400> 51

atgaccgagt acaagcccac ggtgcgcctc gccacccgcg acgacgtccc cagggccgta 60

cgcaccctcg ccgccgcgtt cgccgactac cccgccacgc gccacaccgt cgatccagac 120

cgccacatcg agcgggtcac cgagctgcaa gaactcttcc tcacgcgcgt cgggctcgac 180

atcggcaagg tgtgggtcgc ggacgacggc gccgcggtgg cggtctggac cacgccggag 240

agcgtcgaag cgggggcggt gttcgccgag atcggcccgc gcatggccga gttgagcggt 300

tcccggctgg ccgcgcagca acagatggaa ggtctcctgg cgccgcaccg gcccaaggag 360

cccgcgtggt tcctggccac cgtcggcgtc tcgcccgacc accagggcaa gggtctgggc 420

agcgccgtcg tgctccccgg agtggaggcg gccgagcgcg ccggggtgcc cgccttcctg 480

gagacctccg cgccccgcaa cctccccttc tacgagcggc tcggcttcac cgtcaccgcc 540

gacgtcgagg tgcccgaagg accgcgcacc tggtgcatga cccgcaagcc cggtgcctga 600

600

<210> 52

<211> 1673

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for CD133 partial promoter

HPRM36287(Genecopoeia)

<400> 52

tgcaggggct gcagagggga aatctgccct gattagaaga gatttgtttt ccacccatat 60

tagtttgatc tcacattgct ataaagaact acctgagact gggtttcaga ggctgtaccg 120

taggcatggc cagagaggct tcaggaaact tacaattatg gtggaagaag ggaaagcaag 180

cacatcttca catggcagaa ggagagacag agcgagcaaa ggggaagtgc cacacacttt 240

taaatcatca gatatcctga gaactcactc actatcatga gaacagcaag ggagaaatcc 300

gccctcataa tccaatcacc tcccactagg tccctctccc agtgaattat aattcaatat 360

gagatttgga tggggacaca gagccaaacc atctaacctt gccatgtcct actgtacttt 420

ttggaagttt tctgaactaa catttcttgt ctagaaatat acacactcaa gatcaatcta 480

actattaaaa acaaaaaatg cttattttcc agtacctgtt ttatgttaga ttctggggta 540

aaaagaaaaa ggtactgact taatgttcct gtctatcatt catttgtagc ttgtgcatcc 600

atcctttcat ctgtctatcc atatcacaaa catgtattaa acacctgctg tatgctagac 660

accatcctga aggtgcattt cttccctaag tctttttttt gcctgccgga gagtacagta 720

ttctcaattc aagataagaa gtggtgaata agagtgaagt gcaaattggt gaggacactt 780

agacaccgag tgattacttc tgacaggttt ccaccactca gctaagtagc aaggtgaaga 840

ctccgaggtt gagttgtaac tcacagagcg ggaagaccaa taggcagtga gaaaagaagt 900

tgcagtggct tcacgttata gcagatgcct gtgtgtacac atgggagcga atatggcttt 960

atgctgtttt tcaaccgctt cagccactag cattggtact gctgcacact aaggatccaa 1020

atgattgttt ttttgttttt tgttttgttt tgttttgttt tgttttgttt taaattgcat 1080

tgggattagg caacagaagg gtctaatgcg gccgggatga gacaggagag tttttaggag 1140

ggtagctgca ttctaagtaa gggactctgc ggggggaaaa gaggcgcaag cgttgcaaga 1200

agggagtgca gggggttgag caggcacctc tacaggaaat ggatgctgtc caggtgctgg 1260

tgggcgcccc agggctacgt ggcgaagcag ctcagccggt ccaatcagag tgcgtccagg 1320

gctcgggttt cgcgatcttt aagtgactga ggcagatccc cacgcggcac ctggccatgc 1380

tctcagctct cccgccgcgg tgagtatgtt taaggaatcc tttccattac ggcggcccca 1440

tacctaggtc cccgtccggg acagaggaag ccgcaacagg tccccccggg cacccgggcc 1500

tttctcctgc ctcccgccac gtccgagggt ccggccgcag cgccgcctga gcccctccgc 1560

ggccggcagt gggaggcggg ctctccgaaa gccgtcgcgg tggtcccaga agccgggtca 1620

taaataattc acgagccagg gtctggcgag ctaagggaag ggcggcggca gcg 1673

<210> 53

<211> 6679

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Human CD133 promoter in RMCE 2X RFP reporter

<400> 53

cgatggatct ggagaaggac gcggcagcct actgacttgc ggagatgtcg aagagaaccc 60

tggccctaag cttataactt cgtataatgt atgctatacg aagttattgg tcgacgctag 120

cggtaccgga agcggagcta ctaacttcag cctgctgaag caggctggcg acgtggagga 180

gaaccctgga cctctcgagt cgcgaacgcg tgccaccatg gtgtctaagg gcgaagagct 240

gattaaggag aacatgcaca tgaagctgta catggagggc accgtgaaca accaccactt 300

caagtgcaca tccgagggcg aaggcaagcc ctacgagggc acccagacca tgagaatcaa 360

ggtggtcgag ggcggccctc tcccttcgc cttcgacatc ctggctacca gcttcatgta 420

cggcagcaga accttcatca accacaccca gggcatcccc gacttcttta agcagtcctt 480

ccctgagggc ttcacatggg agagagtcac cacatacgag gacgggggcg tgctgaccgc 540

tacccaggac accagcctcc aggacggctg cctcatctac aacgtcaaga tcagaggggt 600

gaacttccca tccaacggcc ctgtgatgca gaagaaaaca ctcggctggg aggccaacac 660

cgagatgctg taccccgctg acggcggcct ggaaggcaga agcgacatgg ccctgaagct 720

cgtgggcggg ggccacctga tctgcaactt caagaccaca tacagatcca agaaacccgc 780

taagaacctc aagatgcccg gcgtctacta tgtggaccac agactggaaa gaatcaagga 840

ggccgacaaa gagacctacg tcgagcagca cgaggtggct gtggccagat actgcgacct 900

ccctagcaaa ctggggcaca aacttaatac gcgtggctcg ggccagtgta ctaattatgc 960

tctcttgaaa ttggctggag atgttgagag caacccaggt cccgagctcg cggccgcaga 1020

tcttttccc tctgccaaaa attatgggga catcatgaag ccccttgagc atctgacttc 1080

tggcgagctt gccaccatgg tgtctaaggg cgaagagctg attaaggaga acatgcacat 1140

gaagctgtac atggagggca ccgtgaacaa ccaccacttc aagtgcacat ccgagggcga 1200

aggcaagccc tacgagggca cccagaccat gagaatcaag gtggtcgagg gcggccctct 1260

ccccttcgcc ttcgacatcc tggctaccag cttcatgtac ggcagcagaa ccttcatcaa 1320

ccacacccag ggcatccccg acttctttaa gcagtccttc cctgagggct tcacatggga 1380

gagagtcacc acatacgagg acgggggcgt gctgaccgct acccaggaca ccagcctcca 1440

ggacggctgc ctcatctaca acgtcaagat cagaggggtg aacttcccat ccaacggccc 1500

tgtgatgcag aagaaaacac tcggctggga ggccaacacc gagatgctgt accccgctga 1560

cggcggcctg gaaggcagaa gcgacatggc cctgaagctc gtgggcgggg gccacctgat 1620

ctgcaacttc aagaccacat acagatccaa gaaacccgct aagaacctca agatgcccgg 1680

cgtctactat gtggaccaca gactggaaag aatcaaggag gccgacaaag agacctacgt 1740

cgagcagcac gaggtggctg tggccagata ctgcgacctc cctagcaaac tggggcacaa 1800

acttaatggc tccgagggca gaggaagcct tctaacatgc ggtgacgtgg aggagaatcc 1860

cggcccttcc gggatgaccg agtacaagcc cacggtgcgc ctcgccaccc gcgacgacgt 1920

ccccagggcc gtacgcaccc tcgccgccgc gttcgccgac taccccgcca cgcgccacac 1980

cgtcgatcca gaccgccaca tcgagcgggt caccgagctg caagaactct tcctcacgcg 2040

cgtcgggctc gacatcggca aggtgtgggt cgcggacgac ggcgccgcgg tggcggtctg 2100

gaccacgccg gagagcgtcg aagcgggggc ggtgttcgcc gagatcggcc cgcgcatggc 2160

cgagttgagc ggttcccggc tggccgcgca gcaacagatg gaaggtctcc tggcgccgca 2220

ccggcccaag gagcccgcgt ggttcctggc caccgtcggc gtctcgcccg accaccaggg 2280

caagggtctg ggcagcgccg tcgtgctccc cggagtggag gcggccgagc gcgccggggt 2340

gcccgccttc ctggagacct ccgcgccccg caacctcccc ttctacgagc ggctcggctt 2400

caccgtcacc gccgacgtcg aggtgcccga aggaccgcgc acctggtgca tgacccgcaa 2460

gcccggtgcc tgacatgcat ctagagggcc ctattctata gtgtcaccta aatgctagag 2520

ctcgctgatc agcctcgact gtgccttcta gttgccagcc atctgttgtt tgcccctccc 2580

ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt cctttcctaa taaaatgagg 2640

aaattgcatc gcattgtctg agtaggtgtc attctattct gggggggtggg gtggggcagg 2700

acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg gtgggctcta 2760

tggggatcca taacttcgta taaagtatcc tatacgaagt tatctgcagc atagaattca 2820

gcgatatcac tcgagggttc ctggaagttt agatcagggc tcactgactc gctgcgccaa 2880

aataggttag ccttgcgggg tcggtcgttc ggctgcggcg agcggtatca gctcactcaa 2940

aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa 3000

aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc 3060

tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 3120

caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc 3180

cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 3240

ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct 3300

gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 3360

agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 3420

gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct 3480

acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa 3540

gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt 3600

gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 3660

cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat 3720

caaaaaggag ttcctggaag tttagatcag ggtcttcacc tagatccttt tacaaaatag 3780

gttagccttg cggggaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa 3840

acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta 3900

tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc 3960

ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat 4020

ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta 4080

tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt 4140

aatagtttgc gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt 4200

ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg 4260

ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc 4320

gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc 4380

gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg 4440

cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga 4500

actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta 4560

ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct 4620

tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag 4680

ggaataaggg cgacacggaa atgttgaata ctcatactct tccttttca atattattga 4740

agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 4800

aaacaaatag gggttccgcg acatgttctt tcctgcgtta tcccctgatt ctgtggataa 4860

ccgtattacc gcctttgagt gagctgatac cgctcgccgc agccgaacga ccgagcgcag 4920

cgagtcagtg agcgaggaag cggaagagcg cccaatacgc aaaccgcctc tccccgcgcg 4980
ttggccgatt cattaatgca gctgtgcagg ggctgcagag gggaaatctg ccctgattag 5040
aagagatttg ttttccaccc atattagttt gatctcacat tgctataaag aactacctga 5100
gactgggttt cagaggctgt accgtaggca tggccagaga ggcttcagga aacttacaat 5160
tatggtggaa gaagggaaag caagcacatc ttcacatggc agaaggagag acagagcgag 5220
caaaggggaa gtgccacaca cttttaaatc atcagatatc ctgagaactc actcactatc 5280
atgagaacag caagggagaa atccgccctc ataatccaat cacctcccac taggtccctc 5340

tcccagtgaa ttataattca atatgagatt tggatgggga cacagagcca aaccatctaa 5400
ccttgccatg tcctactgta ctttttggaa gttttctgaa ctaacatttc ttgtctagaa 5460
atatacacac tcaagatcaa tctaactatt aaaaacaaaa aatgcttatt ttccagtacc 5520
tgttttatgt tagattctgg ggtaaaaaga aaaaggtact gacttaatgt tcctgtctat 5580
cattcatttg tagcttgtgc atccatcctt tcatctgtct atccatatca caaacatgta 5640
ttaaacacct gctgtatgct agacaccatc ctgaaggtgc atttcttccc taagtctttt 5700
ttttgcctgc cggagagtac agtattctca attcaagata agaagtggtg aataagagtg 5760

aagtgcaaat tggtgaggac acttagacac cgagtgatta cttctgacag gtttccacca 5820
ctcagctaag tagcaaggtg aagactccga ggttgagttg taactcacag agcgggaaga 5880
ccaataggca gtgagaaaag aagttgcagt ggcttcacgt tatagcagat gcctgtgtgt 5940
acacatggga gcgaatatgg ctttatgctg tttttcaacc gcttcagcca ctagcattgg 6000
tactgctgca cactaaggat ccaaatgatt gtttttttgt tttttgtttt gttttgtttt 6060
gttttgtttt gttttaaatt gcattgggat taggcaacag aagggtctaa tgcggccggg 6120
atgagacagg agagttttta ggagggtagc tgcattctaa gtaagggact ctgcgggggg 6180

aaaagaggcg caagcgttgc aagaagggag tgcagggggt tgagcaggca cctctacagg 6240
aaatggatgc tgtccaggtg ctggtgggcg ccccagggct acgtggcgaa gcagctcagc 6300
cggtccaatc agagtgcgtc cagggctcgg gtttcgcgat ctttaagtga ctgaggcaga 6360
tccccacgcg gcacctggcc atgctctcag ctctcccgcc gcggtgagta tgtttaagga 6420
atcctttcca ttacggcggc cccataccta ggtccccgtc cgggacagag gaagccgcaa 6480
caggtccccc cgggcacccg ggcctttctc ctgcctcccg ccacgtccga gggtccggcc 6540
gcagcgccgc ctgagcccct ccgcggccgg cagtgggagg cgggctctcc gaaagccgtc 6600

gcggtggtcc cagaagccgg gtcataaata attcacgagc cagggtctgg cgagctaagg 6660

gaagggcggc ggcagcgat 6679

<210> 54

<211> 9416

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Human CD133 endogenous HDR repoter vector (YFP)

<400> 54

gattcgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 60

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 120

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 180

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 240

acatcaagtg tatcatatgt tcagtgtccc ctccaaaact catgttgaga tttaattgcc 300

attgtgacta tattaacagg cgagaccttg gacaggtggt taagtcatga gagttctgcc 360

ttcataaatg gattaatgcc attatcatgg gagtgggctc ctgataaaag gaggaagctc 420

gaccccattc cctgtctgtc tcttgcactc atttgccctt ccattatggg atgatgcgtc 480

atgaaagccc tcaccagatg ccagcaccag gttcttggac ttcgtggcct tcagaactgt 540

gaaaaacata ttttttttctt tgtaagttac ccagtctctg atattctgtt atagcagcag 600

aaaacatact aagacactgc ttaaattggc ttgttttgct catttcctgc ttctgaggtc 660

tccagagttc cctctcttcc aggataggat ggatatactg tctggaggag aggcaatgct 720

aggcaggagg agaagaacca gtgatcctgt tagtgtgccc agcatccctc tgctcaataa 780

ggcaaatgta gatgttcttt ggaaatgatt gccagagcct acagtccttc ataatgttca 840

gctccaggtc ctggtgaggt ggacagtaac aagagattga gccagcccct gcttgcaccc 900

ttttccaggg gccctttgag ccctttcgat aaagagtttg tttctttacc gaaagcataa 960

aaggtagtgc aagaagaggc acacagaacc aagcatgtgc actttctaaa gtgaaaaact 1020

ttagaaggaa tgtcttgcct ggcagatgcc ctgatgtcct ccgtgcatgt cctcccccaa 1080

cctcatgtgt ctgatgcagc actgcagtga ggcacggtgg atcctgcagg ttcagagggg 1140

ttgtgatcca cctgaagtca cagtcgggca gggtgtgcca gagttgaaac tgccgacctt 1200

ccagatggag ttcactagtc tccccctggc gtataccttt ttggttgtca aggccattat 1260

cacagagtca ctttagagga cactagggct actgttgatg gttcacccag agagccactt 1320

agaggccgtg tgcactgcct gatcttacac ccaggaattt gtagttgtaa caaggtccct 1380

gcggaacttc catttgcttc tgtcttctgt cctagaagtg gactttagaa agactcttgt 1440

ggtagacatg tgtactgctg tgttgctgct ggtctaatgt cttcaacacg tattcttagt 1500
gactcttaac acttttaacc ttagattgtg aatattcaca attcttactt gtcatagtct 1560
gtttatcctc tttatcttct atgctgtttt taaaagcaat acaaacagaa gctagagtaa 1620
ccagtgtttt gtctgacaca ttcttcttga taaactttat gaggtgtggt ttatattcgt 1680
ctgtaccttt agtcacatgc ctgcttctta gcacagaggt gatgaaacca cccacagctt 1740
tgttgggggt attatatttc tttatttttt tgtttttag cccatcacaa catggtaccg 1800
gaagcggagc tactaacttc agcctgctga agcaggctgg cgacgtggag gagaaccctg 1860

gacctctcga gtcgcgaacg cgtgccacca tggtgagcaa gggcgaggag ctgttcaccg 1920
gggtggtgcc catcctggtc gagctggacg gcgacgtaaa cggccacaag ttcagcgtgt 1980
ccggcgaggg cgagggcgat gccacctacg gcaagctgac cctgaagttc atctgcacca 2040
ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac cttcggctac ggcctgcagt 2100
gcttcgcccg ctaccccgac cacatgaagc agcacgactt cttcaagtcc gccatgcccg 2160
aaggctacgt ccaggagcgc accatcttct tcaaggacga cggcaactac aagacccgcg 2220
ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat cgagctgaag ggcatcgact 2280

tcaaggagga cggcaacatc ctggggcaca agctggagta caactacaac agccacaacg 2340
tctatatcat ggccgacaag cagaagaacg gcatcaaggt gaacttcaag atccgccaca 2400
acatcgagga cggcagcgtg cagctcgccg accactacca gcagaacacc cccatcggcg 2460
acggccccgt gctgctgccc gacaaccact acctgagcta ccagtccgcc ctgagcaaag 2520
accccaacga gaagcgcgat cacatggtcc tgctggagtt cgtgaccgcc gccgggatca 2580
ctctcggcat ggacgagctg tacaagacgc gtggctcggg ccagtgtact aattatgctc 2640
tcttgaaatt ggctggagat gttgagagca acccaggtcc cgagctcgcg gccgcagatc 2700

tttttccctc tgccaaaaat tatggggaca tcatgaagcc ccttgagcat ctgacttctg 2760
gcgagcttgc caccatggtg agcaagggcg aggagctgtt caccggggtg gtgcccatcc 2820
tggtcgagct ggacggcgac gtaaacggcc acaagttcag cgtgtccggc gagggcgagg 2880
gcgatgccac ctacggcaag ctgaccctga agttcatctg caccaccggc aagctgcccg 2940
tgccctggcc caccctcgtg accaccttcg gctacggcct gcagtgcttc gcccgctacc 3000
ccgaccacat gaagcagcac gacttcttca agtccgccat gcccgaaggc tacgtccagg 3060
agcgcaccat cttcttcaag gacgacggca actacaagac ccgcgccgag gtgaagttcg 3120

agggcgacac cctggtgaac cgcatcgagc tgaagggcat cgacttcaag gaggacggca 3180

acatcctggg gcacaagctg gagtacaact acaacagcca caacgtctat atcatggccg 3240

acaagcagaa gaacggcatc aaggtgaact tcaagatccg ccacaacatc gaggacggca 3300

gcgtgcagct cgccgaccac taccagcaga acaccccccat cggcgacggc cccgtgctgc 3360

tgcccgacaa ccactacctg agctaccagt ccgccctgag caaagacccc aacgagaagc 3420

gcgatcacat ggtcctgctg gagttcgtga ccgccgccgg gatcactctc ggcatggacg 3480

agctgtacaa gggctccgag ggcagaggaa gccttctaac atgcggtgac gtggaggaga 3540

atcccggccc ttccgggatg accgagtaca agcccacggt gcgcctcgcc acccgcgacg 3600

acgtccccag ggccgtacgc accctcgccg ccgcgttcgc cgactacccc gccacgcgcc 3660

acaccgtcga tccagaccgc cacatcgagc gggtcaccga gctgcaagaa ctcttcctca 3720

cgcgcgtcgg gctcgacatc ggcaaggtgt gggtcgcgga cgacggcgcc gcggtggcgg 3780

tctggaccac gccggagagc gtcgaagcgg gggcggtgtt cgccgagatc ggcccgcgca 3840

tggccgagtt gagcggttcc cggctggccg cgcagcaaca gatggaaggt ctcctggcgc 3900

cgcaccggcc caaggagccc gcgtggttcc tggccaccgt cggcgtctcg cccgaccacc 3960

agggcaaggg tctgggcagc gccgtcgtgc tccccggagt ggaggcggcc gagcgcgccg 4020

gggtgcccgc cttcctggag acctccgcgc cccgcaacct ccccttctac gagcggctcg 4080

gcttcaccgt caccgccgac gtcgaggtgc ccgaaggacc gcgcacctgg tgcatgaccc 4140

gcaagcccgg tgcctgacat gcatctagag ggccctattc tatagtgtca cctaaatgct 4200

agagctcgct gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc 4260

tcccccgtgc cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat 4320

gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 4380

caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc 4440

tctatgggga tccccatcaa gctgatccgg aacccttaat ataacttcgt ataatgtatg 4500

ctatacgaag ttattaggtc cctcgacctg cagcccaagc tagatcgaat tctagctgat 4560

gttgaaactg cttgaggtaa gtttgatgct ttcattgaag agaatctaca ggactatagt 4620

ttttagtttt tagaacatca aaacatagct attcctaata cattgccaca ttaagtggga 4680

aatctagatc agtatattag acaagatatt ttatatgtcc agtccaccct ctgtacctgc 4740

aggttccaca tccaccgatt gattcaacca actgtggatc aaaaatactc taaaaaaatt 4800

ttttttaaat aacagtacaa aaataaaaat aataaaaaca atacagtata acagctactt 4860

acataacact tacattgtat taggtatttt gagtgatcta gagatgatat acagtatgta 4920

ggaggatgtg tgtagcggtt ctggaaccaa tcccacatgg atccagaggg acgactgcat 4980

gtatacatac gatacagtga ccctctctgt atatatatgc attgctatat aaatgtgtgt 5040

gtatttatat atatatatac acacacacag tgagtatgtg ggtgtatatt aaaaaaataa 5100

aatttttta cttaaaaaga taagattttt taagtaaaaa aataagggct ggtgcggtgg 5160

ctcacgcctg taatcccagc actatgggag gccgaggcgg gcagatcact tgaggtcagg 5220

agctcaaaac cagcctggcc aacatggtga aaccctgtct ctactaaaaa aagttagccg 5280

ggtgtggtgg ggggcccctg taatcctgcc tgtaatctca gctactcagg aggccaaggc 5340

aggagaatcg cttcaacctg ggaggcagag gttgcagtga gccgagatca caccactgca 5400

ctccagcgtg ggcaacagag tgagactcca tcacaaaaaa aaaaaaaaag aaaagaaaag 5460

aaagtgtttt tttagagtaa aacaataagc caggcatggt ggtacctgcc tgtagtccca 5520

gctactgggg tgggaggctg tggtagaagg attgcttgag cccaggagtt caaagatgca 5580

gtgaaccatg attatacctg tgaatagcca acacactcat gcactccaac ttggaaaatg 5640

tagtgaaacc ctatctctaa acaaacaaac aaacaaacaa aaatatagaa atctgtgtac 5700

caaccacgta ccttgagtat tatggactta aaatttttc tggctgggca tggtagctca 5760

tgcctgtaat cccagcactt tgggaagctg aggtgggcag atcacttggg gttaggagtt 5820

cgagaccagc ctggccaaca tggagaaacc ccatctctac taaaaaatat aaaaattagc 5880

caggcatggt ggcaggtgcc tgtagtccca gctacttggg aggctgaggc ttaaacccgg 5940

gagacggagg ttgcagtgag ctgagatcgc gccattgcac tccagcctgg gcgacagagt 6000

gagactccct ctcaaaaaaa taaataaata aaataagata aataaacaaa aaccacaaag 6060

ttttttcta tattgctata ttttaacaga ttgctttgca acaaacatat tgctgtgcct 6120

cttgaacgag tttatatttt ccttagttct taggctcctg gtggggacca ttttaggatt 6180

atgcattcct ggttttctct ttattaattc aaaagcaatt tctaaaattg tgtttaactt 6240

tcatttcagc atcaggatac tcaaagtgga aaggatcaca gatttttggt agtttctggg 6300

tctacaagga ctttccaaat ccaggagcaa cgccagtggc aacgtagtga ctcaggcggg 6360

caccaaggca acggcaccat tggtctctgg gtagtgcttt aagaatgaac acaatcacgt 6420

tatagtccat ggtccatcac tattcaagga tgactccctc ccttcctgtc tatttttgtt 6480

ttttactttt ttacactgag tttctattta gacactacaa catatggggt gtttgttccc 6540

attggatgca tttctatcaa aactctatca aatgtgatgg ctagattcta acatattgcc 6600

atgtgtggag tgtgcctcga ggctagagcg gccgccactg gccgtcgttt tacaacgtcg 6660

tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccctttcgc 6720

cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 6780

gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc gcggcgggtg tggtggttac 6840

gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg ctttcttccc 6900

ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt 6960
agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt agggtgatgg 7020
ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac 7080
gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta tctcggtcta 7140
ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat 7200
ttaacaaaaa tttaacgcga attttaacaa aatattaacg cttacaattt aggtggcact 7260
tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca ttcaaatatg 7320

tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt 7380
atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct 7440
gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca 7500
cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 7560
gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc 7620
cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg 7680
gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta 7740

tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct gacaacgatc 7800
ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt 7860
gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg 7920
cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct 7980
tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 8040
tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct 8100
cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 8160

acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc 8220
tcactgatta agcattggta actgtcagac caagtttact catatatact ttagattgat 8280
ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga taatctcatg 8340
accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc 8400
aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa 8460
ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag 8520
gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta 8580

ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta 8640

ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag 8700

ttaccggata aggcgcagcg gtcgggctga cgggggggtt cgtgcacaca gcccagcttg 8760

gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg 8820

cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag 8880

cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc 8940

cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa 9000

aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg 9060

ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct 9120

gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa 9180

gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 9240

cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 9300

ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 9360

attgtgagcg gataacaatt tcacacagga aacagctatg accatgaggc gcgccg 9416

<210> 55

<211> 9416

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Human CD133 endogenous HDR repoter vector (CFP)

<400> 55

gattcgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 60

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 120

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 180

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 240

acatcaagtg tatcatatgt tcagtgtccc ctccaaaact catgttgaga tttaattgcc 300

attgtgacta tattaacagg cgagaccttg gacaggtggt taagtcatga gagttctgcc 360

ttcataaatg gattaatgcc attatcatgg gagtgggctc ctgataaaag gaggaagctc 420

gaccccattc cctgtctgtc tcttgcactc atttgccctt ccattatggg atgatgcgtc 480

atgaaagccc tcaccagatg ccagcaccag gttcttggac ttcgtggcct tcagaactgt 540

gaaaaacata ttttttttctt tgtaagttac ccagtctctg atattctgtt atagcagcag 600

aaaacatact aagacactgc ttaaattggc ttgttttgct catttcctgc ttctgaggtc 660

tccagagttc cctctcttcc aggataggat ggatatactg tctggaggag aggcaatgct 720

aggcaggagg agaagaacca gtgatcctgt tagtgtgccc agcatccctc tgctcaataa 780

ggcaaatgta gatgttcttt ggaaatgatt gccagagcct acagtccttc ataatgttca 840

gctccaggtc ctggtgaggt ggacagtaac aagagattga gccagcccct gcttgcaccc 900

ttttccaggg gccctttgag ccctttcgat aaagagtttg tttctttacc gaaagcataa 960

aaggtagtgc aagaagaggc acacagaacc aagcatgtgc actttctaaa gtgaaaaact 1020

ttagaaggaa tgtcttgcct ggcagatgcc ctgatgtcct ccgtgcatgt cctcccccaa 1080

cctcatgtgt ctgatgcagc actgcagtga ggcacggtgg atcctgcagg ttcagagggg 1140

ttgtgatcca cctgaagtca cagtcgggca gggtgtgcca gagttgaaac tgccgacctt 1200

ccagatggag ttcactagtc tccccctggc gtataccttt ttggttgtca aggccattat 1260

cacagagtca ctttagagga cactagggct actgttgatg gttcacccag agagccactt 1320

agaggccgtg tgcactgcct gatcttacac ccaggaattt gtagttgtaa caaggtccct 1380

gcggaacttc catttgcttc tgtcttctgt cctagaagtg gactttagaa agactcttgt 1440

ggtagacatg tgtactgctg tgttgctgct ggtctaatgt cttcaacacg tattcttagt 1500

gactcttaac acttttaacc ttagattgtg aatattcaca attcttactt gtcatagtct 1560

gtttatcctc tttatcttct atgctgtttt taaaagcaat acaaacagaa gctagagtaa 1620

ccagtgtttt gtctgacaca ttcttcttga taaactttat gaggtgtggt ttatattcgt 1680

ctgtaccttt agtcacatgc ctgcttctta gcacagaggt gatgaaacca cccacagctt 1740

tgttgggggt attatatttc tttatttttt tgtttttag cccatcacaa catggtaccg 1800

gaagcggagc tactaacttc agcctgctga agcaggctgg cgacgtggag gagaaccctg 1860

gacctctcga gtcgcgaacg cgtgccacca tggtgagcaa gggcgaggag ctgttcaccg 1920

gggtggtgcc catcctggtc gagctggacg gcgacgtaaa cggccacagg ttcagcgtgt 1980

ccggcgaggg cgagggcgat gccacctacg gcaagctgac cctgaagttc atctgcacca 2040

ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac cctgacctgg ggcgtgcagt 2100

gcttcagccg ctaccccgac cacatgaagc agcacgactt cttcaagtcc gccatgcccg 2160

aaggctacgt ccaggagcgt accatcttct tcaaggacga cggcaactac aagacccgcg 2220

ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat cgagctgaag ggcatcgact 2280

tcaaggagga cggcaacatc ctggggcaca agctggagta caactacatc agccacaacg 2340

tctatatcac cgccgacaag cagaagaacg gcatcaaggc ccacttcaag atccgccaca 2400

acatcgagga cggcagcgtg cagctcgccg accactacca gcagaacacc cccatcggcg 2460

acggccccgt gctgctgccc gacaaccact acctgagcac ccagtccgcc ctgagcaaag 2520

accccaacga gaagcgcgat cacatggtcc tgctggagtt cgtgaccgcc gccgggatca 2580

ctctcggcat ggacgagctg tacaagacgc gtggctcggg ccagtgtact aattatgctc 2640

tcttgaaatt ggctggagat gttgagagca acccaggtcc cgagctcgcg gccgcagatc 2700

tttttccctc tgccaaaaat tatggggaca tcatgaagcc ccttgagcat ctgacttctg 2760

gcgagcttgc caccatggtg agcaagggcg aggagctgtt caccggggtg gtgcccatcc 2820

tggtcgagct ggacggcgac gtaaacggcc acaggttcag cgtgtccggc gagggcgagg 2880

gcgatgccac ctacggcaag ctgaccctga agttcatctg caccaccggc aagctgcccg 2940

tgccctggcc caccctcgtg accaccctga cctggggcgt gcagtgcttc agccgctacc 3000

ccgaccacat gaagcagcac gacttcttca agtccgccat gcccgaaggc tacgtccagg 3060

agcgtaccat cttcttcaag gacgacggca actacaagac ccgcgccgag gtgaagttcg 3120

agggcgacac cctggtgaac cgcatcgagc tgaagggcat cgacttcaag gaggacggca 3180

acatcctggg gcacaagctg gagtacaact acatcagcca caacgtctat atcaccgccg 3240

acaagcagaa gaacggcatc aaggcccact tcaagatccg ccacaacatc gaggacggca 3300

gcgtgcagct cgccgaccac taccagcaga acaccccat cggcgacggc cccgtgctgc 3360

tgcccgacaa ccactacctg agcacccagt ccgccctgag caaagacccc aacgagaagc 3420

gcgatcacat ggtcctgctg gagttcgtga ccgccgccgg gatcactctc ggcatggacg 3480

agctgtacaa gggctccgag ggcagaggaa gccttctaac atgcggtgac gtggaggaga 3540

atcccggccc ttccgggatg accgagtaca agcccacggt gcgcctcgcc acccgcgacg 3600

acgtccccag ggccgtacgc accctcgccg ccgcgttcgc cgactacccc gccacgcgcc 3660

acaccgtcga tccagaccgc cacatcgagc gggtcaccga gctgcaagaa ctcttcctca 3720

cgcgcgtcgg gctcgacatc ggcaaggtgt gggtcgcgga cgacggcgcc gcggtggcgg 3780

tctggaccac gccggagagc gtcgaagcgg gggcggtgtt cgccgagatc ggcccgcgca 3840

tggccgagtt gagcggttcc cggctggccg cgcagcaaca gatggaaggt ctcctggcgc 3900

cgcaccggcc caaggagccc gcgtggttcc tggccaccgt cggcgtctcg cccgaccacc 3960

agggcaaggg tctgggcagc gccgtcgtgc tccccggagt ggaggcggcc gagcgcgccg 4020

gggtgcccgc cttcctggag acctccgcgc cccgcaacct ccccttctac gagcggctcg 4080

gcttcaccgt caccgccgac gtcgaggtgc ccgaaggacc gcgcacctgg tgcatgaccc 4140

gcaagcccgg tgcctgacat gcatctagag ggccctattc tatagtgtca cctaaatgct 4200

agagctcgct gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc 4260

tcccccgtgc cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat 4320

gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 4380

caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc 4440

tctatgggga tccccatcaa gctgatccgg aacccttaat ataacttcgt ataatgtatg 4500

ctatacgaag ttattaggtc cctcgacctg cagcccaagc tagatcgaat tctagctgat 4560

gttgaaactg cttgaggtaa gtttgatgct ttcattgaag agaatctaca ggactatagt 4620

ttttagtttt tagaacatca aaacatagct attcctaata cattgccaca ttaagtggga 4680

aatctagatc agtatattag acaagatatt ttatatgtcc agtccaccct ctgtacctgc 4740

aggttccaca tccaccgatt gattcaacca actgtggatc aaaaatactc taaaaaaatt 4800

ttttttaaat aacagtacaa aaataaaaat aataaaaaca atacagtata acagctactt 4860

acataacact tacattgtat taggtatttt gagtgatcta gagatgatat acagtatgta 4920

ggaggatgtg tgtagcggtt ctggaaccaa tcccacatgg atccagaggg acgactgcat 4980

gtatacatac gatacagtga ccctctctgt atatatatgc attgctatat aaatgtgtgt 5040

gtatttatat atatatatac acacacacag tgagtatgtg ggtgtatatt aaaaaaataa 5100

aatttttta cttaaaaaga taagattttt taagtaaaaa aataagggct ggtgcggtgg 5160

ctcacgcctg taatcccagc actatgggag gccgaggcgg gcagatcact tgaggtcagg 5220

agctcaaaac cagcctggcc aacatggtga aaccctgtct ctactaaaaa aagttagccg 5280

ggtgtggtgg ggggcccctg taatcctgcc tgtaatctca gctactcagg aggccaaggc 5340

aggagaatcg cttcaacctg ggaggcagag gttgcagtga gccgagatca caccactgca 5400

ctccagcgtg ggcaacagag tgagactcca tcacaaaaaa aaaaaaaaag aaaagaaaag 5460

aaagtgtttt tttagagtaa aacaataagc caggcatggt ggtacctgcc tgtagtccca 5520

gctactgggg tgggaggctg tggtagaagg attgcttgag cccaggagtt caaagatgca 5580

gtgaaccatg attatacctg tgaatagcca acacactcat gcactccaac ttggaaaatg 5640

tagtgaaacc ctatctctaa acaaacaaac aaacaaacaa aaatatagaa atctgtgtac 5700

caaccacgta ccttgagtat tatggactta aaatttttc tggctgggca tggtagctca 5760

tgcctgtaat cccagcactt tgggaagctg aggtgggcag atcacttggg gttaggagtt 5820

cgagaccagc ctggccaaca tggagaaacc ccatctctac taaaaaatat aaaaattagc 5880

caggcatggt ggcaggtgcc tgtagtccca gctacttggg aggctgaggc ttaaacccgg 5940

gagacggagg ttgcagtgag ctgagatcgc gccattgcac tccagcctgg gcgacagagt 6000

gagactccct ctcaaaaaaa taaataaata aaataagata aataaacaaa aaccacaaag 6060

ttttttttcta tattgctata ttttaacaga ttgctttgca acaaacatat tgctgtgcct 6120

cttgaacgag tttatatttt ccttagttct taggctcctg gtggggacca ttttaggatt 6180

atgcattcct ggttttctct ttattaattc aaaagcaatt tctaaaattg tgtttaactt 6240
tcatttcagc atcaggatac tcaaagtgga aaggatcaca gatttttggt agtttctggg 6300
tctacaagga ctttccaaat ccaggagcaa cgccagtggc aacgtagtga ctcaggcggg 6360
caccaaggca acggcaccat tggtctctgg gtagtgcttt aagaatgaac acaatcacgt 6420
tatagtccat ggtccatcac tattcaagga tgactccctc ccttcctgtc tatttttgtt 6480
ttttactttt ttacactgag tttctattta gacactacaa catatggggt gtttgttccc 6540
attggatgca tttctatcaa aactctatca aatgtgatgg ctagattcta acatattgcc 6600

atgtgtggag tgtgcctcga ggctagagcg gccgccactg gccgtcgttt tacaacgtcg 6660
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccttcgc 6720
cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 6780
gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc gcggcgggtg tggtggttac 6840
gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg ctttcttccc 6900
ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt 6960
agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt agggtgatgg 7020

ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac 7080
gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta tctcggtcta 7140
ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat 7200
ttaacaaaaa tttaacgcga attttaacaa aatattaacg cttacaattt aggtggcact 7260
tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca ttcaaatatg 7320
tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt 7380
atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct 7440

gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca 7500
cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 7560
gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc 7620
cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg 7680
gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta 7740
tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct gacaacgatc 7800
ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt 7860

gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg 7920

cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct 7980

tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 8040

tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct 8100

cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 8160

acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc 8220

tcactgatta agcattggta actgtcagac caagtttact catatatact ttagattgat 8280

ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttga taatctcatg 8340

accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc 8400

aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa 8460

ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag 8520

gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta 8580

ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta 8640

ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag 8700

ttaccggata aggcgcagcg gtcgggctga cgggggggtt cgtgcacaca gcccagcttg 8760

gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg 8820

cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag 8880

cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc 8940

cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa 9000

aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg 9060

ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct 9120

gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa 9180

gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 9240

cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 9300

ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 9360

attgtgagcg gataacaatt tcacacagga aacagctatg accatgaggc gcgccg 9416

<210> 56

<211> 9416

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide seqeunce for Human CD133 endogenous HDR repoter vector

(EGFP)

<400> 56

gattcgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 60

tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 120

gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 180

agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 240

acatcaagtg tatcatatgt tcagtgtccc ctccaaaact catgttgaga tttaattgcc 300

attgtgacta tattaacagg cgagaccttg gacaggtggt taagtcatga gagttctgcc 360

ttcataaatg gattaatgcc attatcatgg gagtgggctc ctgataaaag gaggaagctc 420

gaccccattc cctgtctgtc tcttgcactc atttgccctt ccattatggg atgatgcgtc 480

atgaaagccc tcaccagatg ccagcaccag gttcttggac ttcgtggcct tcagaactgt 540

gaaaaacata ttttttttctt tgtaagttac ccagtctctg atattctgtt atagcagcag 600

aaaacatact aagacactgc ttaaattggc ttgttttgct catttcctgc ttctgaggtc 660

tccagagttc cctctcttcc aggataggat ggatatactg tctggaggag aggcaatgct 720

aggcaggagg agaagaacca gtgatcctgt tagtgtgccc agcatccctc tgctcaataa 780

ggcaaatgta gatgttcttt ggaaatgatt gccagagcct acagtccttc ataatgttca 840

gctccaggtc ctggtgaggt ggacagtaac aagagattga gccagcccct gcttgcaccc 900

ttttccaggg gccctttgag ccctttcgat aaagagtttg tttctttacc gaaagcataa 960

aaggtagtgc aagaagaggc acacagaacc aagcatgtgc actttctaaa gtgaaaaact 1020

ttagaaggaa tgtcttgcct ggcagatgcc ctgatgtcct ccgtgcatgt cctcccccaa 1080

cctcatgtgt ctgatgcagc actgcagtga ggcacggtgg atcctgcagg ttcagagggg 1140

ttgtgatcca cctgaagtca cagtcgggca gggtgtgcca gagttgaaac tgccgacctt 1200

ccagatggag ttcactagtc tccccctggc gtataccttt ttggttgtca aggccattat 1260

cacagagtca ctttagagga cactagggct actgttgatg gttcacccag agagccactt 1320

agaggccgtg tgcactgcct gatcttacac ccaggaattt gtagttgtaa caaggtccct 1380

gcggaacttc catttgcttc tgtcttctgt cctagaagtg gactttagaa agactcttgt 1440

ggtagacatg tgtactgctg tgttgctgct ggtctaatgt cttcaacacg tattcttagt 1500

gactcttaac acttttaacc ttagattgtg aatattcaca attcttactt gtcatagtct 1560

gtttatcctc tttatcttct atgctgtttt taaaagcaat acaaacagaa gctagagtaa 1620

ccagtgtttt gtctgacaca ttcttcttga taaactttat gaggtgtggt ttatattcgt 1680

ctgtaccttt agtcacatgc ctgcttctta gcacagaggt gatgaaacca cccacagctt 1740

tgttgggggt attatatttc tttatttttt tgtttttag cccatcacaa catggtaccg 1800

gaagcggagc tactaacttc agcctgctga agcaggctgg cgacgtggag gagaaccctg 1860

gacctctcga gtcgcgaacg cgtgccacca tggtgagcaa gggcgaggag ctgttcaccg 1920

gggtggtgcc catcctggtc gagctggacg gcgacgtaaa cggccacaag ttcagcgtgt 1980

ccggcgaggg cgagggcgat gccacctacg gcaagctgac cctgaagttc atctgcacca 2040

ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac cctgacctac ggcgtgcagt 2100

gcttcagccg ctaccccgac cacatgaagc agcacgactt cttcaagtcc gccatgcccg 2160

aaggctacgt ccaggagcgc accatcttct tcaaggacga cggcaactac aagacccgcg 2220

ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat cgagctgaag ggcatcgact 2280

tcaaggagga cggcaacatc ctggggcaca agctggagta caactacaac agccacaacg 2340

tctatatcat ggccgacaag cagaagaacg gcatcaaggt gaacttcaag atccgccaca 2400

acatcgagga cggcagcgtg cagctcgccg accactacca gcagaacacc cccatcggcg 2460

acggccccgt gctgctgccc gacaaccact acctgagcac ccagtccgcc ctgagcaaag 2520

accccaacga gaagcgcgat cacatggtcc tgctggagtt cgtgaccgcc gccgggatca 2580

ctctcggcat ggacgagctg tacaagacgc gtggctcggg ccagtgtact aattatgctc 2640

tcttgaaatt ggctggagat gttgagagca acccaggtcc cgagctcgcg gccgcagatc 2700

tttttccctc tgccaaaaat tatggggaca tcatgaagcc ccttgagcat ctgacttctg 2760

gcgagcttgc caccatggtg agcaagggcg aggagctgtt caccggggtg gtgcccatcc 2820

tggtcgagct ggacggcgac gtaaacggcc acaagttcag cgtgtccggc gagggcgagg 2880

gcgatgccac ctacggcaag ctgaccctga agttcatctg caccaccggc aagctgcccg 2940

tgccctggcc caccctcgtg accaccctga cctacggcgt gcagtgcttc agccgctacc 3000

ccgaccacat gaagcagcac gacttcttca agtccgccat gcccgaaggc tacgtccagg 3060

agcgcaccat cttcttcaag gacgacggca actacaagac ccgcgccgag gtgaagttcg 3120

agggcgacac cctggtgaac cgcatcgagc tgaagggcat cgacttcaag gaggacggca 3180

acatcctggg gcacaagctg gagtacaact acaacagcca caacgtctat atcatggccg 3240

acaagcagaa gaacggcatc aaggtgaact tcaagatccg ccacaacatc gaggacggca 3300

gcgtgcagct cgccgaccac taccagcaga acaccccat cggcgacggc cccgtgctgc 3360

tgcccgacaa ccactacctg agcacccagt ccgccctgag caaagacccc aacgagaagc 3420

gcgatcacat ggtcctgctg gagttcgtga ccgccgccgg gatcactctc ggcatggacg 3480

agctgtacaa gggctccgag ggcagaggaa gccttctaac atgcggtgac gtggaggaga 3540

atcccggccc ttccgggatg accgagtaca agcccacggt gcgcctcgcc acccgcgacg 3600

acgtccccag ggccgtacgc accctcgccg ccgcgttcgc cgactacccc gccacgcgcc 3660

acaccgtcga tccagaccgc cacatcgagc gggtcaccga gctgcaagaa ctcttcctca 3720

cgcgcgtcgg gctcgacatc ggcaaggtgt gggtcgcgga cgacggcgcc gcggtggcgg 3780

tctggaccac gccggagagc gtcgaagcgg gggcggtgtt cgccgagatc ggcccgcgca 3840

tggccgagtt gagcggttcc cggctggccg cgcagcaaca gatggaaggt ctcctggcgc 3900

cgcaccggcc caaggagccc gcgtggttcc tggccaccgt cggcgtctcg cccgaccacc 3960

agggcaaggg tctgggcagc gccgtcgtgc tccccggagt ggaggcggcc gagcgcgccg 4020

gggtgcccgc cttcctggag acctccgcgc cccgcaacct ccccttctac gagcggctcg 4080

gcttcaccgt caccgccgac gtcgaggtgc ccgaaggacc gcgcacctgg tgcatgaccc 4140

gcaagcccgg tgcctgacat gcatctagag ggccctattc tatagtgtca cctaaatgct 4200

agagctcgct gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc 4260

tcccccgtgc cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat 4320

gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg 4380

caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc 4440

tctatgggga tccccatcaa gctgatccgg aacccttaat ataacttcgt ataatgtatg 4500

ctatacgaag ttattaggtc cctcgacctg cagcccaagc tagatcgaat tctagctgat 4560

gttgaaactg cttgaggtaa gtttgatgct ttcattgaag agaatctaca ggactatagt 4620

ttttagtttt tagaacatca aaacatagct attcctaata cattgccaca ttaagtggga 4680

aatctagatc agtatattag acaagatatt ttatatgtcc agtccaccct ctgtacctgc 4740

aggttccaca tccaccgatt gattcaacca actgtggatc aaaaatactc taaaaaaatt 4800

ttttttaaat aacagtacaa aaataaaaat aataaaaaca atacagtata acagctactt 4860

acataacact tacattgtat taggtatttt gagtgatcta gagatgatat acagtatgta 4920

ggaggatgtg tgtagcggtt ctggaaccaa tcccacatgg atccagaggg acgactgcat 4980

gtatacatac gatacagtga ccctctctgt atatatatgc attgctatat aaatgtgtgt 5040

gtatttatat atatatatac acacacacag tgagtatgtg ggtgtatatt aaaaaaataa 5100

aatttttta cttaaaaaga taagattttt taagtaaaaa aataagggct ggtgcggtgg 5160

ctcacgcctg taatcccagc actatgggag gccgaggcgg gcagatcact tgaggtcagg 5220

agctcaaaac cagcctggcc aacatggtga aaccctgtct ctactaaaaa aagttagccg 5280

ggtgtggtgg ggggcccctg taatcctgcc tgtaatctca gctactcagg aggccaaggc 5340

aggagaatcg cttcaacctg ggaggcagag gttgcagtga gccgagatca caccactgca 5400

ctccagcgtg ggcaacagag tgagactcca tcacaaaaaa aaaaaaaaag aaaagaaaag 5460

aaagtgtttt tttagagtaa aacaataagc caggcatggt ggtacctgcc tgtagtccca 5520
gctactgggg tgggaggctg tggtagaagg attgcttgag cccaggagtt caaagatgca 5580
gtgaaccatg attatacctg tgaatagcca acacactcat gcactccaac ttggaaaatg 5640
tagtgaaacc ctatctctaa acaaacaaac aaacaaacaa aaatatagaa atctgtgtac 5700
caaccacgta ccttgagtat tatggactta aaatttttc tggctgggca tggtagctca 5760
tgcctgtaat cccagcactt tgggaagctg aggtgggcag atcacttggg gttaggagtt 5820
cgagaccagc ctggccaaca tggagaaacc ccatctctac taaaaaatat aaaaattagc 5880

caggcatggt ggcaggtgcc tgtagtccca gctacttggg aggctgaggc ttaaacccgg 5940
gagacggagg ttgcagtgag ctgagatcgc gccattgcac tccagcctgg gcgacagagt 6000
gagactccct ctcaaaaaaa taaataaata aaataagata aataaacaaa aaccacaaag 6060
tttttttcta tattgctata ttttaacaga ttgctttgca acaaacatat tgctgtgcct 6120
cttgaacgag tttatatttt ccttagttct taggctcctg gtggggacca ttttaggatt 6180
atgcattcct ggttttctct ttattaattc aaaagcaatt tctaaaattg tgtttaactt 6240
tcatttcagc atcaggatac tcaaagtgga aaggatcaca gatttttggt agtttctggg 6300

tctacaagga ctttccaaat ccaggagcaa cgccagtggc aacgtagtga ctcaggcggg 6360
caccaaggca acggcaccat tggtctctgg gtagtgcttt aagaatgaac acaatcacgt 6420
tatagtccat ggtccatcac tattcaagga tgactccctc ccttcctgtc tatttttgtt 6480
ttttactttt ttacactgag tttctattta gacactacaa catatggggt gtttgttccc 6540
attggatgca tttctatcaa aactctatca aatgtgatgg ctagattcta acatattgcc 6600
atgtgtggag tgtgcctcga ggctagagcg gccgccactg gccgtcgttt tacaacgtcg 6660
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccttcgc 6720

cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 6780
gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc gcggcgggtg tggtggttac 6840
gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg ctttcttccc 6900
ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt 6960
agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt agggtgatgg 7020
ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac 7080
gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta tctcggtcta 7140

ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat 7200

ttaacaaaaa tttaacgcga attttaacaa aatattaacg cttacaattt aggtggcact 7260

tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca ttcaaatatg 7320

tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt 7380

atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct 7440

gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca 7500

cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 7560

gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc 7620

cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg 7680

gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta 7740

tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct gacaacgatc 7800

ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt 7860

gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg 7920

cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct 7980

tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 8040

tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct 8100

cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 8160

acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc 8220

tcactgatta agcattggta actgtcagac caagtttact catatatact ttagattgat 8280

ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttga taatctcatg 8340

accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc 8400

aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa 8460

ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag 8520

gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta 8580

ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta 8640

ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag 8700

ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca gcccagcttg 8760

gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg 8820

cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag 8880

cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc 8940

cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa 9000

aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg 9060

ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct 9120

gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa 9180

gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 9240

cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 9300

ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 9360

attgtgagcg gataacaatt tcacacagga aacagctatg accatgaggc gcgccg 9416

<210> 57

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer SOX2_F

<400> 57

tacagcatga tgcaggacca 20

<210> 58

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer SOX2_R

<400> 58

cgagctggtc atggagttgt a 21

<210> 59

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer SLC26A3_F

<400> 59

acacatatcc gtgggtccgt 20

<210> 60

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer SLC26A3_R

<400> 60

gttgcattgc gatctgggac 20

<210> 61

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer CEACAM1_F

<400> 61

gctctgatag cagtagccct g 21

<210> 62

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer CEACAM1_R

<400> 62

attggagtgg tcctgagctg c 21

<210> 63

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer CD177_F

<400> 63

catctctcag gaggtgggct 20

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer CD177_R

<400> 64

cccgatttgt ctggtgtggt 20

<210> 65

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer MUC2_F

<400> 65

gtcaccacgg aggtttcgta 20

<210> 66

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer MUC2_R

<400> 66

ccttggccga gtacatgaca 20

<210> 67

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer TFF3_F

<400> 67

gtacgtgggc ctgtctgcaa 20

<210> 68

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence for primer TFF3_R

<400> 68

tgaaacacca aggcactcca 20