



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0134894  
(43) 공개일자 2018년12월19일

(51) 국제특허분류(Int. Cl.)  
C12N 15/90 (2006.01) C12N 15/65 (2006.01)  
(52) CPC특허분류  
C12N 15/902 (2013.01)  
C07K 16/00 (2013.01)  
(21) 출원번호 10-2018-7029371  
(22) 출원일자(국제) 2017년04월20일  
심사청구일자 없음  
(85) 번역문제출일자 2018년10월11일  
(86) 국제출원번호 PCT/US2017/028552  
(87) 국제공개번호 WO 2017/184831  
국제공개일자 2017년10월26일  
(30) 우선권주장  
62/325,385 2016년04월20일 미국(US)

(71) 출원인  
리제너론 파마슈티칼스 인코포레이티드  
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리  
버 로드 777  
(72) 발명자  
밥 로버트  
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리  
버 로드 777 리제너론 파마슈티칼스 인코포레이티  
드 내  
부라코프 다리야  
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리  
버 로드 777 리제너론 파마슈티칼스 인코포레이티  
드 내  
(뒷면에 계속)  
(74) 대리인  
특허법인와이에스장

전체 청구항 수 : 총 57 항

(54) 발명의 명칭 **발현 강화 유전자좌의 사용에 기초하여 항체를 만들기 위한 조성물 및 방법**

(57) 요약

본 발명은 진핵 세포에서 재조합 단백질을 부위 특이적으로 통합하고 및 발현시키는 것에 관한 것이다. 특히, 본 발명은 발현 강화 유전자좌를 사용하여 진핵 세포, 특히 차이나이즈 햄스터(*Cricetulus griseus*) 세포주에서 이중 특이적 항체를 포함하는 항체의 발현을 개선하기 위한 조성물 및 방법을 포함한다.

(52) CPC특허분류

**C07K 16/468** (2013.01)

**C12N 15/65** (2013.01)

**C07K 2317/31** (2013.01)

**C07K 2317/526** (2013.01)

**C12N 2510/02** (2013.01)

(72) 발명자

**첸 강**

미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버  
로드 777 리제너론 파마슈티칼스 인코포레이티드  
내

**팬들 제임스 피.**

미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버  
로드 777 리제너론 파마슈티칼스 인코포레이티드  
내

## 명세서

### 청구범위

#### 청구항 1

발현 강화 유전자와 내의 특정 부위에 통합된 외인성 핵산 서열을 포함하는 세포로서, 상기 외인성 핵산 서열은 이중 특이적 항원 결합 단백질을 암호화하는, 세포.

#### 청구항 2

제1항에 있어서, 상기 외인성 핵산 서열은 제1 LCF를 암호화하는 뉴클레오타이드 서열을 포함하는 제1 외인성 핵산, 제1 HCF를 암호화하는 뉴클레오타이드 서열을 포함하는 제2 외인성 핵산, 및 제2 HCF를 암호화하는 뉴클레오타이드 서열을 포함하는 제3 외인성 핵산을 포함하는, 세포.

#### 청구항 3

제2항에 있어서, 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 CH2 제1 CH3 도메인을 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 CH2 및 제2 CH3 도메인을 암호화하는, 세포.

#### 청구항 4

제3항에 있어서, 상기 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산 위치가 상이한, 세포.

#### 청구항 5

제3항에 있어서, 상기 제1 및 제2 CH3 도메인은 인간 IgG의 CH3 도메인이고, 상기 2개의 CH3 도메인 중 하나는 적어도 하나의 아미노산 위치가 변형되어 미변형 CH3 도메인과는 상이한 단백질 A 결합 특성을 가지는, 세포.

#### 청구항 6

제3항에 있어서, 상기 제1 및 제2 CH 도메인을 암호화하는 상기 뉴클레오타이드 서열은 상기 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 서로 상이한, 세포.

#### 청구항 7

제2항에 있어서, 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 VH를 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 VH를 암호화하는, 세포.

#### 청구항 8

제2항에 있어서, 상기 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 VL을 암호화하는, 세포.

#### 청구항 9

제1항에 있어서, 상기 외인성 핵산은 제2 LCF를 암호화하는 뉴클레오타이드 서열을 포함하는 추가 외인성 핵산을 추가로 포함하는, 세포.

#### 청구항 10

제9항에 있어서, 상기 제2 LCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 VL을 암호화하는, 세포.

#### 청구항 11

제2항에 있어서, 상기 제1 외인성 핵산은 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열에 작동 가능하게 연결된 제1 프로모터를 추가로 포함하고, 상기 제2 외인성 핵산은 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열에 작동 가능하게 연결된 제2 프로모터를 추가로 포함하며, 상기 제3 외인성 핵산은 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열에 작동 가능하게 연결된 제3 프로모터를 포함하되, 상기 제1, 제2 및 제3 프로모터는 동일하거나 상이한, 세포.

#### 청구항 12

제2항에 있어서, 상기 제1 외인성 핵산은 상기 제2 및 제3 외인성 핵산에 대해 상대적으로 상류에 위치하는, 세포.

#### 청구항 13

제12항에 있어서, 상기 제1 외인성 핵산에 대해 상대적으로 5'에 위치한 제1 재조합 효소 인식 부위(RRS), 및 상기 제2 및 제3 외인성 핵산 둘 모두에 대해 상대적으로 3'에 위치한 제2 재조합 효소 인식 부위(RRS)를 추가로 포함하되, 상기 제1 및 제2 RRS는 상이한, 세포.

#### 청구항 14

제12항에 있어서, 상기 제1 외인성 핵산에 대해 상대적으로 3'에 위치하고 상기 제2 및 제3 외인성 핵산에 둘 모두에 대해 상대적으로 5'에 위치하는 제3 RRS를 추가로 포함하되, 상기 제3 RRS는 상기 제1 및 제2 RRS와는 상이한, 세포.

#### 청구항 15

제13항에 있어서, 선별 표지 유전자를 포함하는 제4 외인성 핵산을 추가로 포함하는 세포.

#### 청구항 16

제15항에 있어서, 상기 제4 외인성 핵산은 제3 RRS를 추가로 포함하고, 상기 제1 외인성 핵산에 대해 3'에 위치하는, 세포.

#### 청구항 17

제2항에 있어서, 상기 제2 외인성 핵산은 상기 제1 및 제3 외인성 핵산에 대해 상대적으로 상류에 위치하는, 세포.

#### 청구항 18

제16항에 있어서, 상기 유전자좌에서 상기 외인성 핵산의 순서는 5'에서 3'까지, 상기 제1 외인성 핵산, 상기 제4 외인성 핵산, 상기 제2 외인성 핵산, 및 상기 제3 외인성 핵산이고, 상기 제2 외인성 핵산은 인간 IgG의 변형된 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함하고, 상기 제3 외인성 핵산은 상기 인간 IgG의 원시 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함하는, 세포.

#### 청구항 19

제16항에 있어서, 상기 유전자좌에서 상기 외인성 핵산의 순서는 5'에서 3'까지, 상기 제1 외인성 핵산, 상기 제2 외인성 핵산, 상기 제4 외인성 핵산, 및 상기 제3 외인성 핵산이고, 상기 제2 외인성 핵산은 인간 IgG의 원시 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함하고, 상기 제3 외인성 핵산은 상기 인간 IgG의 변형된 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함하는, 세포.

#### 청구항 20

제18항 또는 제19항에 있어서, 상기 제1, 제2 및 제3 프로모터는 동일하고, 상기 선별 표지 유전자가 작동 가능하게 연결되는 상기 프로모터와는 상이한, 세포.

#### 청구항 21

제1항에 있어서, 상기 이중 특이적 항원 결합 단백질은 T 세포 항원 및 종양 세포 항원에 특이적으로 결합하는, 세포.

#### 청구항 22

제1항에 있어서, 상기 발현 강화 유전자좌는 서열번호 1과 적어도 90% 동일한 뉴클레오타이드 서열을 포함하는 유전자좌, 및 서열번호 2와 적어도 90% 동일한 뉴클레오타이드 서열을 포함하는 유전자좌로 이루어진 군으로부터 선

택되는, 세포.

#### 청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 상기 세포는 CHO 세포인, 세포.

#### 청구항 24

백터 세트로서,

5'에서 3'까지, 제1 RRS, 제1 LCF를 암호화하는 뉴클레오타이드 서열을 포함하는 제1 핵산, 및 제3 RRS를 포함하는 제1 백터;

5'에서 3'까지, 상기 제3 RRS, 제1 HCF를 암호화하는 뉴클레오타이드 서열을 포함하는 제2 핵산, 및 제2 RRS를 포함하는 제2 백터를 포함하되,

상기 제1 또는 제2 핵산은 제2 HCF를 암호화하는 뉴클레오타이드 서열을 추가로 포함하고,

상기 제1 및 제2 HCF 및 상기 제1 LCF는 이중 특이적 항원 결합 단백질을 암호화하는, 백터 세트.

#### 청구항 25

제24항에 있어서, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 상기 제1 핵산에 포함되는, 백터 세트.

#### 청구항 26

제25항에 있어서, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 상기 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열의 상류 또는 하류에 위치하는, 백터 세트.

#### 청구항 27

제24항에 있어서, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 상기 제2 핵산에 포함되는, 백터 세트.

#### 청구항 28

제24항에 있어서, 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화하는, 백터 세트.

#### 청구항 29

제28항에 있어서, 상기 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산 위치가 상이한, 백터 세트.

#### 청구항 30

제28항에 있어서, 상기 제1 및 제2 CH3 도메인은 인간 IgG의 CH3 도메인이고, 상기 2개의 CH3 도메인 중 하나는 적어도 하나의 아미노산 위치가 변형되어 상기 미변형 CH3 도메인과는 상이한 단백질 A 결합 특성을 가지는, 백터 세트.

#### 청구항 31

제28항에 있어서, 상기 제1 및 제2 CH 도메인을 암호화하는 상기 뉴클레오타이드 서열은 상기 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 서로 상이한, 백터 세트.

#### 청구항 32

제24항에 있어서, 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 VH를 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 중쇄 VH를 암호화하는, 백터 세트.

#### 청구항 33

제24항에 있어서, 상기 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 VL을 암호화하는, 백터 세트.

#### 청구항 34

제24항에 있어서, 제1 경쇄 LCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 프로모터에 작동 가능하게 연결되고, 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 프로모터에 작동 가능하게 연결되며, 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제3 프로모터에 작동 가능하게 연결되는, 벡터 세트.

#### 청구항 35

제34항에 있어서, 상기 제1, 제2 및 제3 프로모터는 동일하거나 상이한, 벡터 세트.

#### 청구항 36

제24항에 있어서, 상기 제1 벡터에서의 상기 제1 핵산은 상기 제3 RRS에 대해 5'에 위치한 제2 프로모터를 추가로 포함하고, 상기 제2 벡터에서의 상기 제2 핵산은 상기 제3 RRS에 대해 3'에 위치한 선별 표지 유전자를 추가로 포함하는, 벡터 세트.

#### 청구항 37

제24항에 있어서, 상기 제1 핵산은 선별 표지 유전자의 5' 부분에 대해 5'에 위치한 상기 제1 벡터에서 제2 프로모터를 추가로 포함하고, 상기 제2 벡터 내의 상기 선별 표지의 3' 부분은 상기 선별 표지 유전자의 5' 부분의 상동 아암에 대해 3'에 존재하는, 벡터 세트.

#### 청구항 38

제24항에 있어서,

상기 제1 벡터는 5'에서 3'까지, 상기 제1 RRS, 상기 제1 핵산, 및 상기 제3 RRS를 포함하고;

상기 제2 벡터는 5'에서 3'까지, 상기 제3 RRS, 및 (제1 중쇄 또는 이의 단편을 암호화하는 상기 뉴클레오타이드 서열 및 제2 중쇄 또는 이의 단편을 암호화하는 상기 뉴클레오타이드 서열을 포함하는) 상기 제2 핵산을 포함하는, 벡터 세트.

#### 청구항 39

제24항에 있어서,

상기 제1 벡터는 5'에서 3'까지, 상기 제1 RRS, (제1 LCF를 암호화하는 상기 뉴클레오타이드 서열 및 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열을 포함하는) 상기 제1 핵산, 및 제3 RRS를 포함하고;

상기 제2 벡터는 5'에서 3'까지, 제3 RRS, 및 (제1 HCF를 암호화하는 뉴클레오타이드 서열을 포함하는) 제2 핵산을 포함하는, 벡터 세트.

#### 청구항 40

제38항 또는 제39항에 있어서, 상기 제1 핵산은 상기 제1 벡터에서의 상기 제3 RRS에 대해 5'에 위치한 선별 표지 유전자의 5' 부분을 추가로 포함하고, 상기 제2 핵산은 상기 제2 벡터에서의 상기 제3 RRS에 대해 3'에 위치한 상기 선별 표지 유전자의 나머지 3' 부분을 추가로 포함하며, 선택적으로, 상기 제1 벡터에서의 상기 제3 RRS는 상기 선별 표지 유전자 인트론의 5' 부분 내에 존재하고, 상기 제2 벡터에서의 상기 제3 RRS는 상기 선별 표지 유전자 인트론의 3' 부분 내에 존재하는, 벡터세트.

#### 청구항 41

제24항에 있어서, 하나 이상의 RRS 및 제2 LCF를 암호화하는 뉴클레오타이드 서열을 포함하는 제3 벡터를 추가로 포함하는 벡터 세트.

#### 청구항 42

제24항에 있어서, 상기 RRS를 인식하는 하나 이상의 제조합 효소를 암호화하는 제3 벡터를 추가로 포함하는 벡터 세트.

#### 청구항 43

세포의 발현 강화 유전자좌 내로의 통합을 위한 5' 상동 아암 및 3' 상동 아암이 측면에 위치한 이중 특이적 항

원 결합 단백질을 암호화하는 외인성 핵산 서열을 포함하는 벡터.

#### 청구항 44

세포 및 벡터 세트를 포함하는 시스템으로서,

상기 세포는 상기 세포 계놈의 발현 강화 유전자좌 내에 통합된, 제1 RRS, 제1 외인성 핵산, 제3 RRS, 제2 외인성 핵산, 및 제2 RRS를 5'에서 3'까지 포함하되, 상기 3개의 RRS는 서로 상이하고;

상기 벡터 세트는

5'에서 3'까지, 상기 제1 RRS, 제1 LCF를 암호화하는 뉴클레오티드 서열을 포함하는 제1 핵산, 및 상기 제3 RRS를 포함하는 제1 벡터, 및

상기 제3 RRS, 제1 HCF를 암호화하는 뉴클레오티드 서열을 포함하는 제2 핵산, 및 제2 RRS를 포함하는 제2 벡터를 포함하며;

상기 제1 핵산 또는 상기 제2 핵산 중 하나는 제2 HCF를 암호화하는 뉴클레오티드 서열을 추가로 포함하고;

상기 세포 내로 상기 벡터가 도입될 때, 상기 벡터에서의 상기 제1 및 제2 핵산은 상기 제1, 제2 및 제3 RRS에 의해 매개된 재조합을 통해 상기 발현 강화 유전자좌 내에 통합되는, 시스템.

#### 청구항 45

제44항에 있어서, 상기 제1 외인성 핵산은 제1 선별 표지 유전자를 포함하고, 상기 제2 외인성 핵산은 제2 선별 표지 유전자를 포함하며, 상기 제1 및 제2 선별 표지 유전자는 상이한, 시스템.

#### 청구항 46

제44항에 있어서,

상기 제1 벡터는 5'에서 3'까지, 상기 제1 RRS, 상기 제1 LCF를 포함하는 상기 제1 핵산, 및 상기 제3 RRS를 포함하고;

상기 제2 벡터는 5'에서 3'까지, 상기 제3 RRS, (상기 제1 HCF를 암호화하는 상기 뉴클레오티드 서열 및 상기 제2 HCF를 암호화하는 상기 뉴클레오티드 서열 둘 다를 포함하는) 상기 제2 핵산, 및 상기 제2 RRS를 포함하는, 시스템.

#### 청구항 47

제44항에 있어서,

상기 제1 벡터는 5'에서 3'까지, 상기 제1 RRS, (상기 제1 LCF를 암호화하는 상기 뉴클레오티드 서열 및 상기 제2 HCF를 암호화하는 상기 뉴클레오티드 서열을 포함하는) 제1 핵산, 및 제3 RRS를 포함하고;

상기 제2 벡터는 5'에서 3'까지, 상기 제3 RRS, (상기 제1 HCF를 암호화하는 상기 뉴클레오티드 서열을 포함하는) 상기 제2 핵산, 및 상기 제2 RRS를 포함하는, 시스템.

#### 청구항 48

제46항 또는 제47항에 있어서, 상기 제1 벡터에서의 상기 제1 핵산은 상기 제3 RRS에 대해 5'에 위치한 선별 표지 유전자의 5' 부분을 추가로 포함하고; 상기 제2 벡터에서의 상기 제2 핵산은 상기 제3 RRS에 대해 3'에 위치한 선별 표지 유전자의 나머지 3' 부분을 추가로 포함하는, 시스템.

#### 청구항 49

제48항에 있어서, 상기 제1 벡터에서의 상기 제3 RRS는 상기 선별 표지 유전자 인트론의 5' 부분 내에 존재하고; 상기 제2 벡터에서의 제3 RRS는 상기 선별 표지 유전자 인트론의 3' 부분 내에 존재하는, 시스템.

#### 청구항 50

제48항에 있어서, 상기 LCF를 암호화하는 상기 뉴클레오티드 서열은 제1 프로모터에 작동 가능하게 연결되고, 상기 제1 HCF를 암호화하는 상기 뉴클레오티드 서열은 제2 프로모터에 작동 가능하게 연결되며, 상기 제2 HCF를

암호화하는 상기 뉴클레오타이드 서열은 제3 프로모터에 작동 가능하게 연결되되, 상기 제1, 제2 및 제3 프로모터는 동일하거나 상이하고, 및/또는 상기 프로모터들은 상기 선별 표지 유전자가 작동 가능하게 연결되는 상기 프로모터와는 동일하거나 상이한, 시스템.

#### 청구항 51

제44항에 있어서, 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화하는, 시스템.

#### 청구항 52

제51항에 있어서, 상기 제1 및 제2 CH3 도메인 중 하나는 인간 IgG의 CH3 도메인이고, 다른 하나는 상기 인간 IgG의 변형된 CH3 도메인으로서 적어도 하나의 아미노산 위치의 변형을 포함하는, 시스템.

#### 청구항 53

제52항에 있어서, 상기 변형된 CH3 도메인을 암호화하는 상기 뉴클레오타이드 서열은 상기 제1 벡터 내에 있는, 시스템.

#### 청구항 54

제52항에 있어서, 상기 변형된 CH3 도메인을 암호화하는 상기 뉴클레오타이드 서열은 상기 제2 벡터 내에 있고 상기 미변형된 CH3 도메인을 암호화하는 상기 뉴클레오타이드 서열의 상류에 있는, 시스템.

#### 청구항 55

- (i) 제44항 내지 제54항 중 어느 한 항의 시스템을 제공하는 단계;
- (ii) 상기 벡터를 상기 세포 내로 형질감염에 의해 동시에 도입하는 단계; 및
- (iii) 상기 벡터에서의 상기 제1 및 제2 핵산이 상기 제1, 제2 및 제3 RRS에 의해 매개된 재조합을 통해 상기 세포의 상기 발현 강화 유전자와 내에 통합된 형질감염 세포를 선별하는 단계를 포함하는, 방법.

#### 청구항 56

제55항에 있어서,

- (iv) 상기 선별된 형질감염 세포에서 상기 제1 LCF, 상기 제1 HCF, 및 상기 제2 HCF를 발현시키는 단계; 및
- (v) 상기 선별된 형질감염 세포로부터 상기 제1 LCF, 상기 제1 HCF 및 상기 제2 HCF를 포함하는 상기 이중 특이적 항원 결합 단백질을 수득하는 단계를 추가로 포함하는, 방법.

#### 청구항 57

이중 특이적 항원 결합 단백질을 만드는 방법으로서,

- (i) 제1항 내지 제23항 중 어느 한 항의 상기 세포를 제공하는 단계;
- (ii) 상기 외인성 핵산 서열로부터 상기 이중 특이적 항원 결합 단백질을 발현시키는 단계; 및
- (iii) 상기 세포로부터 상기 이중 특이적 항원 결합 단백질을 수득하는 단계를 포함하는, 방법.

## 발명의 설명

## 기술 분야

### [0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 그 내용 전체가 참조로서 본원에 통합되는, 2016년 4월 20일에 출원된 미국 가출원 번호 제 62/325,385호의 우선권의 이익을 주장한다.



[0003] 기술분야

[0004] 본 개시는 진핵 세포에서의 재조합 단백질의 부위 특이적 통합 및 발현에 관한 것이다. 특히, 본 개시는 발현 강화 유전자좌를 사용함으로써 진핵 세포, 특히 차이니즈 햄스터(*Cricetulus griseus*) 세포주에서 이중 특이적 항체와 같은 항원 결합 단백질의 발현을 강화시키기 위한 조성물 및 방법에 관한 것이다.

배경 기술

[0005] 세포 발현 시스템의 목적은, 연구용이거나 치료용이거나를 불문하고, 특정 단백질의 제조를 위한 신뢰할 수 있고 효율적인 공급원을 제공하는 것이다. 포유류 세포에서의 재조합 단백질을 발현시키는 것은 치료용 단백질 제조를 위한 바람직한 방법인데, 이는, 예를 들어, 재조합 단백질을 적절하게 번역한 후 변형하는 포유류 발현 시스템의 능력에 기인한다.

[0006] 다양한 발현 시스템을 이용할 수 있음에도 불구하고, 재조합 단백질의 발현을 위한 효율적인 유전자 전달 및 통합 유전자의 안정성에 대한 어려움은 여전히 존재한다. 표적 전이 유전자의 장기적인 발현을 위해서는, 세포주의 표현형이 변하지 않도록 세포 유전자의 파괴를 최소화하는 것을 고려할 수 있다.

[0007] 안정적인 세포주를 조작하여 발현을 위한 다중 유전자, 예컨대 다중 특이적 항체에서와 같이 다중 항체 사슬을 수용하는 것은 특히 어렵다. 통합된 유전자의 발현 수준은 매우 다양할 수 있다. 유전자를 추가로 통합하면 국부적인 유전적 환경(즉, 자리 효과(position effect))으로 인해 발현의 다양성과 불안정성이 더 커질 수 있다. 다중 특이적 항원 결합 단백질의 생산을 위한 발현 시스템은 종종 특정 다중 결합 포맷으로 쌍을 이루도록 의도된 둘 이상의 상이한 면역 글로불린 사슬의 발현을 필요로 하고, 종종 원하는 이중 이량체(heterodimer) 또는 다량체(multimer) 조합보다는 동중 이량체의 생산에 유리하게 작용할 수 있다. 따라서, 당업계에는 개선된 포유류 발현 시스템에 대한 요구가 있다.

발명의 내용

[0008] 일 양태에서, 특정 부위에서 발현 강화 유전자좌 내에 통합된 외인성 핵산 서열을 포함하는 세포가 제공되며, 외인성 핵산 서열은 이중 특이적 항원 결합 단백질을 암호화한다.

[0009] 일부 구현예에서, 외인성 핵산 서열은 제1 경쇄 단편(LCF)을 암호화하는 뉴클레오타이드 서열을 함유하는 제1 외인성 핵산, 제1 중쇄 단편(HCF)을 암호화하는 뉴클레오타이드 서열을 함유하는 제2 외인성 핵산, 및 제2 HCF(또는 HCF\*로도 지칭되며 제2 HCF는 제1 HCF와 상이함)를 암호화하는 뉴클레오타이드 서열을 함유하는 제3 외인성 핵산을 포함하되, 제1 및 제2 HCF 및 제1 LCF는 이중 특이적 항원 결합 단백질을 형성한다. 특정 구현예에서, 제1 및 제2 HCF 및 제1 LCF는 적어도 2개의 가변 영역 및 이중 특이적 항원 결합 단백질의 2개의 CH3 불변 도메인을 함유한다. 일부 구현예에서, 2개의 가변 영역은 상이하다. 일부 구현예에서, 2개의 CH3 도메인은 상이하다. 일부 구현예에서, 각각의 외인성 핵산 서열은 특정 부위에서 발현 강화 유전자좌 내에 동시에 통합된다.

[0010] 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 불변 영역으로부터의 아미노산을 암호화하고 (예를 들어, CH1, CH2, 힌지 또는 CH3 도메인 중 하나 이상을 암호화하고), 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 불변 영역으로부터의 아미노산을 암호화한다. 제1 불변 영역으로부터의 아미노산은 제2 불변 영역으로부터의 아미노산과 동일하거나 상이할 수 있다. 특이적 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화하되, 제1 및 제2 CH3 도메인은 동일하거나 상이할 수 있다. 일부 구현예에서, 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산 위치가 상이하고 (예를 들어, 2개의 CH3 도메인 중 하나는 인간 IgG CH3 도메인이고, 다른 하나는 변형된 인간 IgG CH3 도메인임), 2개의 CH3 도메인은 상이한 단백질 A 결합 특성을 갖는다. 다른 구현예에서, 제1 및 제2 CH3 도메인을 암호화하는 뉴클레오타이드 서열은 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 서로 상이하다.

[0011] 다른 특이적 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 중쇄 가변(VH) 영역을 암호화하고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 VH 영역을 암호화하되, 제1 및 제2 중쇄는 동일하거나 상이한 VH 영역을 가질 수 있다. 또 다른 구현예에서, 제1 및 제2 VH는 동일하거나 상이한 불변 영역에 연결될 수 있다.

[0012] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오타이드 서열은 제1 경쇄 가변(VL) 영역을 암호화한다.

[0013] 일부 구현예에서, 외인성 핵산 서열은 제2 경쇄 가변(VL) 영역과 같은 제2 LCF를 암호화하는 뉴클레오타이드 서열

을 포함하는 추가 외인성 핵산을 함유한다. 일부 구현예에서, 제2 VL 영역을 암호화하는 뉴클레오타이드 서열은 제2 경쇄 불변 영역도 암호화한다.

- [0014] 다수의 외인성 핵산의 유전자좌에서의 상대 위치는 변할 수 있다. 일부 구현예에서, LCF 암호화 핵산은 2개의 HCF 암호화 핵산 모두에 대해 상대적으로 상류 또는 하류에 위치한다.
- [0015] 일부 구현예에서, HCF 암호화 서열 또는 LCF 암호화 서열 각각은 전사 조절 서열에 독립적으로 및 작동 가능하게 연결된다. 특이적 구현예에서, 제1 외인성 핵산은 제1 LCF를 암호화하는 뉴클레오타이드 서열에 작동 가능하게 연결된 제1 프로모터를 추가로 포함하고, 제2 외인성 핵산은 제1 HCF를 암호화하는 뉴클레오타이드 서열에 작동 가능하게 연결된 제2 프로모터를 추가로 포함하며, 제3 외인성 핵산은 제2 HCF를 암호화하는 뉴클레오타이드 서열에 작동 가능하게 연결된 제3 프로모터를 포함하되, 제1, 제2 및 제3 프로모터는 동일하거나 상이하고/상이하거나 상기 프로모터들은 제4 외인성 핵산이 작동 가능하게 연결되는 제4 프로모터와 동일하거나 상이하다. 일부 구현예에서, 제1, 제2 및 제3프로모터는 동일하다.
- [0016] 일부 구현예에서, 외인성 핵산 서열은 통합 부위에서 재조합 효소 인식 부위, 예를 들어, 제1 외인성 핵산에 대해 상대적으로 5'에 위치한 제1 재조합 효소 인식 부위(RRS), 및 제2 및 제3 외인성 핵산 둘 모두에 대해 상대적으로 3'에 위치한 제2 재조합 효소 인식 부위(RRS)를 추가로 포함하되, 제1 및 제2 RRS는 상이하다. 일부 구현예에서, 제3 RRS가 또한 포함되어 제1 외인성 핵산에 대해 상대적으로 3'에 및 제2 및 제3 외인성 핵산에 둘 모두에 대해 상대적으로 5'에 위치하며, 제3 RRS는 제1 및 제2 RRS와 상이하다.
- [0017] 일부 구현예에서, 외인성 핵산 서열은 선별 표지 유전자를 함유하는 제4 외인성 핵산을 포함할 수 있다. 특이적 구현예에서, 제4 외인성 핵산은 제1 외인성 핵산에 대해 상대적으로 3'에 위치한다. 특정 구현예에서, 제4 외인성 핵산은 분할된 유전자로서 통합된다. 다른 구현예에서, 제4 외인성 핵산 또는 선별 표지는 제4 외인성 핵산에 작동 가능하게 연결되는 제4 프로모터의 3'인 제3 RRS의 3'에 위치한다. 일부 구현예에서, 선별 표지 유전자는 삽입된 (선택적으로, 선별 표지 유전자의 인트론 내에 삽입된) 제3 RRS를 포함하되, 제3 RRS는 제1 및 제2 RRS와 상이하다.
- [0018] 특정 구현예에서, 유전자좌에서 외인성 핵산의 순서는 5'에서 3'까지 (LCF를 암호화하는) 제1 외인성 핵산, (선별 표지를 암호화하는) 제4 외인성 핵산, (제1 HCF를 암호화하는) 제2 외인성 핵산, 및 (제2 HCF를 암호화하는) 제3 외인성 핵산일 수 있는데, 일부 특정 구현예에서, 제2 외인성 핵산은 인간 IgG의 변형된 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 함유하고, 제3 외인성 핵산은 인간 IgG의 원시 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함한다.
- [0019] 특정 구현예에서, 유전자좌에서 외인성 핵산의 순서는 5'에서 3'까지 (LCF를 암호화하는) 제1 외인성 핵산, (제1 HCF를 암호화하는) 제2 외인성 핵산, (선별 표지를 암호화하는) 제4 외인성 핵산, 및 (제2 HCF를 암호화하는) 제3 외인성 핵산일 수 있으며, 제2 외인성 핵산은 인간 IgG의 원시 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함하고, 제3 외인성 핵산은 인간 IgG의 변형된 CH3 도메인을 암호화하는 뉴클레오타이드 서열을 포함한다.
- [0020] 일부 구현예에서, HCF 또는 LCF 코딩 서열에 연결된 프로모터는 동일하며, 선별 표지 유전자가 작동 가능하게 연결되는 프로모터와는 상이하다.
- [0021] 일부 구현예에서, 이중 특이적 항원 결합 단백질은 T 세포 항원 및 종양 세포 항원에 특이적으로 결합한다. 다른 적합한 이중 항원 특이성도 제공된다.
- [0022] 일부 구현예에서, 발현 강화 유전자좌는 서열번호 1과 적어도 90% 동일한 뉴클레오타이드 서열을 포함하는 유전자좌, 또는 서열번호 2와 적어도 90% 동일한 뉴클레오타이드 서열을 포함하는 유전자좌로부터 선택된다.
- [0023] 다양한 구현예에서, 세포는 CHO 세포이다.
- [0024] 또 다른 양태에서, 다수의 외인성 핵산의 부위 특이적 통합을 위해 설계된 벡터가 제공된다.
- [0025] 일부 구현예에서, 본 개시는 벡터 세트를 제공하는데, 벡터 세트는: (5'에서 3'까지) 제1 RRS, 제1 HCF를 암호화하는 뉴클레오타이드 서열을 함유하는 제1 핵산, 및 제3 RRS를 함유하는 제1 벡터; 및 (5'에서 3'까지) 제3 RRS, 제1 VH 영역을 암호화하는 뉴클레오타이드 서열을 함유하는 제2 핵산, 및 제2 RRS를 함유하는 제2 벡터를 포함하되, 제1 또는 제2 핵산 중 하나는 제2 HCF를 암호화하는 뉴클레오타이드를 추가로 포함하고, 제1 및 제2 HCF, 및 제1 LCF는 이중 특이적 항원 결합 단백질을 형성한다.

- [0026] 일부 구현예에서, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제1 핵산에 포함되고, 선택적으로는, 제1 LCF를 암호화하는 뉴클레오타이드 서열의 하류에 위치한다. 다른 구현예에서, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 핵산에 포함된다.
- [0027] 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 키메라 불변 영역을 암호화하고 (예를 들어, 임의의 동형으로부터의 CH1, 힌지 CH2, 또는 CH3 도메인, 또는 이의 단편 중 하나 이상을 암호화하고), 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 키메라 불변 영역을 암호화한다. 키메라 불변 영역의 실시예는 2014년 8월 7일에 공개되고 본원에 참조로서 통합된 PCT 국제 공보 제WO 2014/121087 A1호에 기술되어 있다. 제1 불변 영역으로부터의 아미노산은 제2 키메라 불변 영역과 동일하거나 상이할 수 있다. 특이적 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화하고, 제1 및 제2 CH3 도메인은 동일하거나 상이할 수 있다. 일부 구현예에서, 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산 위치가 상이하고 (예를 들어, 2개의 CH3 도메인 중 하나는 인간 IgG CH3 도메인이고, 다른 하나는 변형된 인간 IgG CH3 도메인임), 2개의 CH3 도메인은 상이한 단백질 A 결합 특성을 갖는다. 다른 구현예에서, 제1 및 제2 CH3 도메인을 암호화하는 뉴클레오타이드 서열은 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 서로 상이하다.
- [0028] 다른 특이적 구현예에서, 제1 VH 영역을 암호화하는 뉴클레오타이드 서열은 제1 중쇄를 암호화하고, 제2 VH 영역을 암호화하는 뉴클레오타이드 서열은 제2 중쇄를 암호화하고, 제1 및 제2 중쇄는 동일하거나 상이한 불변 영역을 가질 수 있다.
- [0029] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오타이드 서열은 제1 경쇄 가변 영역(VL)을 암호화한다.
- [0030] 일부 구현예에서, LCF 암호화 서열 또는 HCF 암호화 서열 각각은 프로모터와 같은 전사 조절 서열에 독립적으로 연결된다. 특이적 구현예에서, 제1 HCF 코딩 서열에 연결된 프로모터 및 제2 HCF 코딩 서열에 연결된 프로모터는 동일하다. 특이적 구현예에서, LCF(들) 및 HCF(들)에 연결된 프로모터는 모두 동일하다.
- [0031] 일부 구현예에서, 제1 벡터에서 제1 핵산은 제1 벡터의 제3 RRS에 대해 5'에 위치한 선별 표지 유전자의 5' 부분을 추가로 함유하고, 제2 핵산은 제2 벡터에서 제3 RRS에 대해 3'에 위치한 선별 표지 유전자의 나머지 3' 부분을 추가로 함유한다 - 즉, 선별 표지 유전자는 2개의 벡터로 분할된다. 다른 구현예에서, 선별 표지 및 선별 표지가 작동 가능하게 연결되는 프로모터는 2개의 벡터 사이에서 분할된다. 즉, 프로모터와 선별 표지 유전자는 상이한 벡터 상에 위치한다. 특정 구현예에서, 표지 유전자에 작동 가능하게 연결된 프로모터는 제3 RRS에 대해 제1 벡터 5'에 위치하고, 표지 유전자는 제2 벡터에서 제3 RRS에 대해 3'에 위치하고, 제2 핵산에 작동 가능하게 연결된 제2 프로모터 및 제3 핵산에 작동 가능하게 연결된 제3 프로모터의 5'에 위치한다. 일부 구현예에서, 제1 벡터에서의 제1 RRS는 선별 표지 유전자 인트론의 5' 부분 내에 존재하고, 제2 벡터에서의 제3 RRS는 선별 표지 유전자 인트론의 3' 부분 내에 존재한다.
- [0032] 특이적 구현예에서, 제1 벡터는 5'에서 3'까지, 제1 RRS, 제1 핵산, 및 제3 RRS를 포함하고, 제2 벡터는 5'에서 3'까지, 제3 RRS, 제2 핵산(제2 핵산은 제1 HCF를 암호화하는 뉴클레오타이드 서열 및 제2 HCF를 암호화하는 뉴클레오타이드 서열을 함유함), 및 제2 RRS를 포함한다. 다른 특이적 구현예에서, 제1 벡터는 5'에서 3'까지, 제1 RRS, 제1 핵산(제1 핵산은 제1 HCF를 암호화하는 뉴클레오타이드 서열 및 제2 HCF 영역을 암호화하는 뉴클레오타이드 서열을 포함함), 및 제3 RRS를 포함하고, 제2 벡터는 5'에서 3'까지, 제3 RRS, 제2 핵산(제2 핵산은 제1 HCF 영역을 암호화하는 뉴클레오타이드 서열을 포함함), 및 제2 RRS를 포함한다. 이들 특이적 구현예 중 어느 하나에서, 제1 핵산은 제1 벡터에서의 제3 RRS에 대해 5'에 위치한 선별 표지 유전자의 5' 부분을 추가로 포함할 수 있고, 제2 핵산은 제2 벡터에서 제3 RRS에 대해 3'에 위치한 선별 표지 유전자의 나머지 3' 부분을 추가로 포함하며, 선택적으로, 제1 벡터에서의 제3 RRS는 선별 표지 유전자 인트론의 5' 부분 내에 존재하고, 제2 벡터에서의 제3 RRS는 선별 표지 유전자 인트론의 3' 부분 내에 존재한다.
- [0033] 일부 구현예에서, 벡터 세트는 추가적인 벡터 또는 벡터들, 예를 들어, 하나 이상의 RRS 및 제2 LCF를 암호화하는 뉴클레오타이드 서열을 함유하는 벡터, 또는 RRS를 인식하는 하나 이상의 제조함 효소를 암호화하는 벡터를 포함할 수 있다.
- [0034] 다른 구현예에서, 본 개시는 상동 아암에 기초한 상동성 제조함을 통해 다수의 외인성 핵산의 부위 특이적 통합을 달성하도록 설계된 벡터를 제공한다. 일부 구현예에서, 벡터는 세포의 발현 강화 유전자와 내로의 통합을 위한 5' 상동 아암 및 3' 상동 아암이 측면에 위치한 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산 서열을 함유한다.

- [0035] 추가적인 양태에서, 본 개시는, 세포 및 하나 이상의 박테의 조합을 포함하는 시스템으로서, 함께 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산이 발현 강화 유전자와 내에 통합된 세포를 만드는데 사용할 수 있는 시스템을 제공한다.
- [0036] 특정 구현예에서, 세포 및 박터 세트를 포함하는 시스템이 제공되는데, 세포는, 그 게놈의 발현 강화 유전자와 내에 통합된 RRS 세트로서, 서로 상이하고 박터 세트 내의 관심 유전자와의 재조합 교환을 위해 하나 이상의 외인성 핵산, 예컨대 선별 표지 사이에서 이격된 RRS 세트를 함유하며, 박터 세트 내의 RRS는 세포 내의 RRS와 동일한 배치를 포함한다.
- [0037] 일부 구현예에서, 세포 및 박터 세트를 포함하는 시스템이 제공되는데, 세포는, 그 게놈의 발현 강화 유전자와 내에 통합된 제1 RRS, 제1 외인성 핵산, 제2 RRS, 제2 외인성 핵산, 및 제3 RRS를 5'에서 3'까지 함유하되, 3개의 RRS는 서로 상이하고; 박터 세트는 5'에서 3'까지 제1 RRS, 제1 면역 글로불린 사슬 또는 그의 단편을 암호화하는 뉴클레오티드 서열, 및 제2 RRS를 함유하는 제1 박터, 및 제2 RRS, 제2 면역 글로불린 사슬 또는 그의 단편을 암호화하는 뉴클레오티드 서열을 함유하는 제2 핵산, 및 제3 RRS를 함유하는 제2 박터를 포함하며; 제1 핵산 또는 제2 핵산 중 어느 하나는 제3 면역 글로불린 사슬 또는 그의 단편을 암호화하는 뉴클레오티드 서열을 추가로 포함한다. 세포 내로 박터가 도입될 때, 박터 내의 제1 및 제2 핵산은 제1, 제2 및 제3 RRS에 의해 매개된 재조합을 통해 발현 강화 유전자와 내에 통합된다.
- [0038] 일부 구현예에서, 세포 내의 제1 외인성 핵산은 제1 선별 표지 유전자를 포함하고, 세포 내의 제2 외인성 핵산은 제2 선별 표지 유전자를 포함하되, 제1 및 제2 선별 표지 유전자는 상이하다. 선별 표지는 세포 내에서 통합된 외인성 핵산과 교환된다.
- [0039] 일부 구현예에서, 제1 박터는 5'에서 3'까지, 제1 RRS, 제1 LCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 핵산, 및 제3 RRS를 포함하고; 제2 박터는 5'에서 3'까지, 제3 RRS, 제2 핵산(제2 핵산은 제1 HCF를 암호화하는 뉴클레오티드 서열 및 제2 HC를 암호화하는 뉴클레오티드 서열 모두를 함유함), 및 제2 RRS를 포함한다. 다른 구현예에서, 제1 박터는 5'에서 3'까지, 제1 RRS, 제1 LCF를 암호화하는 뉴클레오티드 서열 및 제2 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 핵산, 및 제3 RRS를 포함하고; 제2 박터는 5'에서 3'까지, 제3 RRS, 제1 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제2 핵산, 및 제2 RRS를 함유한다.
- [0040] 일부 구현예에서, 제1 박터는 5'에서 3'까지, 제1 RRS, 제1 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 핵산, 및 제3 RRS를 포함하고; 제2 박터는 5'에서 3'까지, 제3 RRS, 제2 핵산(제2 핵산은 제1 LCF를 암호화하는 뉴클레오티드 서열 및 제2 HCF를 암호화하는 뉴클레오티드 서열 모두를 함유함), 및 제2 RRS를 포함한다. 다른 구현예에서, 제1 박터는 5'에서 3'까지, 제1 RRS, 제1 HCF를 암호화하는 뉴클레오티드 서열 및 제2 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 핵산, 및 제3 RRS를 포함하고; 제2 박터는 5'에서 3'까지, 제3 RRS, 제1 LCF를 암호화하는 뉴클레오티드 서열을 함유하는 제2 핵산, 및 제2 RRS를 함유한다. 이들 구현예 중 어느 하나에서, 제1 박터 내의 제1 핵산은 제3 RRS에 대해 5'에 위치한 프로모터를 추가로 포함할 수 있고, 제2 박터 내의 제2 핵산은 제3 RRS에 대해 3'에 위치하고 프로모터가 작동 가능하게 연결될 선별 표지 유전자를 추가로 포함한다. 이들 특이적 구현예 중 어느 하나에서, 제1 박터 내의 제1 핵산은 제3 RRS에 대해 5'에 위치한 선별 표지 유전자의 5' 부분을 추가로 포함할 수 있고, 제1 박터 내의 제2 핵산은 제3 RRS에 대해 3'에 위치한 선별 표지 유전자의 나머지 3' 부분을 추가로 포함하며; 선택적으로는, 제1 박터 내의 제3 RRS는 선별 표지 유전자 인트론의 5' 부분 내에 존재하고; 제2 박터 내의 제3 RRS는 선별 표지 유전자 인트론의 3' 부분 내에 존재한다.
- [0041] 일부 구현예에서, LCF를 암호화하는 뉴클레오티드 서열은 제1 프로모터에 작동 가능하게 연결되고, 제1 HCF를 암호화하는 뉴클레오티드 서열은 제2 프로모터에 작동 가능하게 연결되며, 제2 HCF를 암호화하는 뉴클레오티드 서열은 제3 프로모터에 작동 가능하게 연결되되, 제1, 제2 및 제3 프로모터는 동일하고, 선별 표지 유전자(박터들 중 하나에 존재하는 경우)가 작동 가능하게 연결되는 프로모터와는 상이하다.
- [0042] 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오티드 서열은 제1 CH3 도메인을 암호화하고, 제2 HCF를 암호화하는 뉴클레오티드 서열은 제2 CH3 도메인을 암호화하되, 제1 및 제2 CH3 도메인은 동일하거나 상이할 수 있다. 일부 구현예에서, 2개의 CH3 도메인 중 하나는 인간 IgG의 원시 CH3 도메인이고, 다른 하나의 CH3 도메인은 인간 IgG의 변형된 CH3 도메인이다. 특이적 구현예에서, 변형된 CH3 도메인을 암호화하는 뉴클레오티드 서열은 제1 박터(즉, 제1 LCF를 암호화하는 박터) 내에, 선택적으로는, 제1 LCF를 암호화하는 뉴클레오티드 서열의 하류에 존재한다. 다른 특이적 구현예에서, 변형된 CH3 도메인을 암호화하는 뉴클레오티드 서열은 제2 박터 내에 있



고, 미변형된 CH3 도메인을 암호화하는 뉴클레오티드 서열의 상류에 있다.

[0043] 또 다른 양태에서, 본 개시는 이중 특이적 항원 결합 단백질의 제작 방법을 제공한다.

[0044] 일부 구현예에서, 상기 방법은: RRS를 갖는 세포, 및 이중 특이적 항원 결합 단백질을 함께 암호화하는 다수의 외인성 핵산 및 세포 내의 RRS와 일치하는 RRS를 함유하는 벡터 세트를 포함하는, 본원에 기술된 시스템을 제공하는 단계; 벡터를 형질 감염에 의해 세포 내로 도입하는 단계; 벡터 내의 외인성 핵산이 RRS에 의해 매개된 재조합을 통해 세포의 발현 강화 유전자와 내에 통합된, 형질 감염 세포를 선별하는 단계; 변형된 세포 내의 핵산에 의해 암호화된 폴리펩티드를 발현시키는 단계; 및 형질 감염 세포로부터 이중 특이적 항원 결합 단백질을 수득하는 단계를 포함한다.

[0045] 일부 구현예에서, 상기 방법은, 발현 강화 유전자와 내에 통합된 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산 서열을 세포가 함유하는 단계; 외인성 핵산 서열로부터 이중 특이적 항원 결합 단백질을 발현시키는 단계; 및 세포로부터 이중 특이적 항원 결합 단백질을 수득하는 단계를 포함할 수 있다.

### 도면의 간단한 설명

[0046] **도 1.** 발현 강화 유전자와 내에서의 통합을 위한 예시적인 이중 특이적 클로닝 전략. 경쇄("LC") 벡터(예: 공통 경쇄), 및 HC\*와 같은 이중 중쇄("HC") 벡터("\*"이 있으면 2개의 HC가 상이함을 나타내는데, 예를 들어, HC\*는 CH3 도메인 내에 변형을 함유하고/하거나 코돈 변형된 것임)는 관심 항체의 가변 영역을 적합한 벡터 내에 클로닝 함으로써 만들어진다. LC 벡터의 3' RRS 부위와 이중 HC 벡터의 5' RRS 부위는 동일하고 하이그로마인신 내성 유전자의 분할된 인트론에 포함되는데, 상기 하이그로마인신 내성 유전자는 인트론을 조합하고 절단하도록 조작되어 재조합체의 효율적인 선별을 위해 하이그로마인신 내성 유전자에 의해 암호화된 단백질을 발현시킨다. 화살표는 프로모터를 나타낸다.

**도 2.** 발현 강화 유전자와 내에서의 통합을 위한 예시적인 이중 특이적 클로닝 전략. 5' RRS(RRS1)를 갖는 공용 경쇄(예를 들어, WO 2013022782에 기술된 바와 같은 인간화 공용 경쇄(Humanized Universal Light Chain, ULC) VelocImmune® 마우스 유래의 공용 경쇄를 참조함)를 이용하면, 공용 경쇄에 대한 발현 카세트를 함유하는 선재 플라스미드에 제3 RRS(RRS3)가 측면에 위치한 하나의 중쇄(HC\*)를 삽입함으로써 새로운 이중 특이적 항체를 효율적으로 작제할 수 있다. 제2 중쇄(HC)는 RRS2 및 RRS3 부위를 갖는 제2 플라스미드 내에 클로닝된다.

**도 3.** 발현 강화 유전자와 내에서의 통합을 위한 예시적인 이중 특이적 클로닝 전략. 이중 특이적 항체의 3개의 뚜렷한 항체 사슬(AbC1, AbC2 및 AbC3)이 개별적인 벡터 내로 먼저 클로닝된다. AbC1 및 AbC3 벡터는 각각 항체 발현 카세트의 측면에 위치하는 RRS 부위를 갖는다. AbC2에 대한 발현 카세트는 AbC2 플라스미드로부터 절단된 뒤, AbC3 발현 플라스미드 내로 서브 클로닝되어, 5'에서 3'까지 RRS3 부위, AbC2 발현 카세트, AbC3 발현 카세트, 및 RRS2 부위를 함유하는 플라스미드를 생성한다. 이 플라스미드는 발현 강화 유전자좌에서 RRS1 및 RRS2를 보유하는 숙주 세포 내에 AbC1 플라스미드 및 재조합 효소와 함께 도입된다. 이중 특이적 항체 발현 세포주는 재조합 효소 매개 카세트 교환 이후 단리된다.

**도 4.** 발현 강화 유전자와 내에서의 통합을 위한 예시적인 이중 특이적 클로닝 전략. 이중 특이적 항체의 3개의 뚜렷한 항체 사슬(AbC1, AbC2 및 AbC3)이 개별적인 벡터 내로 먼저 클로닝된다. AbC1 및 AbC3 벡터는 각각 항체 발현 카세트의 측면에 위치하는 RRS 부위를 갖는다. AbC2에 대한 발현 카세트는 AbC2 플라스미드로부터 절단된 뒤, AbC1 발현 플라스미드 내로 서브 클로닝되어, 5'에서 3'까지 RRS1 부위, AbC1 발현 카세트, AbC2 발현 카세트, 및 RRS3 부위를 함유하는 플라스미드를 생성한다. 이 플라스미드는 발현 강화 유전자좌에서 RRS1 및 RRS2를 보유하는 숙주 세포 내에 AbC3 플라스미드 및 재조합 효소와 함께 도입된다. 이중 특이적 항체 발현 세포주는 재조합 효소 매개 카세트 교환 이후 단리된다.

**도 5.** 하나의 게놈 부위(EESYR<sup>®</sup>)에서 통합된 발현 카세트로부터 이중 특이적 항체의 발현. 재조합 효소 매개 카세트 교환에 의해 EESYR<sup>®</sup> 유전자좌에서 CHO 세포주 RSX4189-1, RSX4187-1, RSX4191-1, RSX4188-1을 생성하였다. EESYR<sup>®</sup> 유전자좌에서의 이중 특이적 Ab의 3개의 뚜렷한 항체 사슬((AbC1, AbC2, 및 AbC3)에 대한 발현 카세트의 배열은 좌측에 도시되어 있다. 4일 동안 셰이커 플라스크 배양액의 소모 배지에서 각각의 이중 특이적 항체의 역가를 HPLC에 의해 측정하였고, 이는 우측에 막대 그래프에 도시되어 있다.

### 발명을 실시하기 위한 구체적인 내용

[0047] 정의

[0048] 용어 "항체"는 본원에서 사용되는 바와 같이, 4개의 폴리펩티드 사슬, 즉 이황화 결합에 의해 상호 연결된 2개의 중(H)쇄 및 2개의 경(L)쇄를 포함하는 면역 글로불린 분자를 포함한다. 각각의 중쇄는 중쇄 가변 영역(본원에서 HCVR 또는 VH로 약칭됨)과 중쇄 불변 영역을 포함한다. 중쇄 불변 영역은 3개의 도메인 CH1, CH2 및 CH3을 포함한다. 각각의 경쇄는 경쇄 가변 영역(본원에서 LCVR 또는 VL로 약칭됨)과 경쇄 불변 영역을 포함한다. 경쇄 불변 영역은 하나의 도메인, CL을 포함한다. VH와 VL 영역은 추가적으로, 더 보존적이면서 프레임워크 영역(FR)인 영역들이 중간에 끼어 있는, 상보성 결정 영역(CDR)이라 지칭되는 초가변성 영역으로 다시 나누어질 수 있다. 각각의 VH 및 VL은 아미노 말단으로부터 카르복시 말단 쪽으로 다음 순서로 배열된 3개의 CDR과 4개의 FR을 포함한다: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 (중쇄 CDR은 HCDR1, HCDR2 및 HCDR3로서 약칭될 수 있고; 경쇄 CDR은 LCDR1, LCDR2 및 LCDR3로서 약칭될 수 있음).

[0049] "항원 결합 단백질"이라는 문구는, 적어도 하나의 CDR을 가지며 항원을 선택적으로 인식할 수 있는, 즉 항원을 적어도 마이크로몰 범위에 있는  $K_D$ 와 결합시킬 수 있는 단백질을 포함한다. 치료용 항원 결합 단백질(예: 치료 항체)는 나노몰 또는 피코몰 범위에 있는  $K_D$ 를 빈번하게 필요로 한다. 일반적으로, 항원 결합 단백질은 2개 이상의 CDR, 예를 들어 2, 3, 4, 5 또는 6개의 CDR을 포함한다. 예시적인 항원 결합 단백질은 항체, 항체의 중쇄 및 경쇄 가변 영역을 함유하는 폴리펩티드와 같은 항체의 항원 결합 단편(예: Fab 단편, F(ab')<sub>2</sub> 단편), 및 항체의 중쇄 및 경쇄 가변 영역을 함유하고 중쇄 및/또는 경쇄의 불변 영역(예컨대, 하나 이상의 불변 영역, 즉, CL, CH1, CH2, 및 CH3 도메인 중 하나 이상)으로부터의 추가 아미노산을 함유하는 단백질을 포함한다.

[0050] "이중 특이적 항원 결합 단백질"이라는 문구는 2개의 상이한 분자(예: 항원) 또는 동일한 분자(예: 동일한 항원) 상에서 2개 이상의 에피토프에 선택적으로 결합할 수 있거나, 상기 2개 이상의 에피토프에 대해 상이한 특이성을 갖는 항원 결합 단백질을 포함한다. 이러한 단백질의 항원 결합 부분, 또는 단편 항원 결합(Fab) 부분은 특정 항원에 대한 특이성을 나타내고, 일반적으로 면역 글로불린의 중쇄 가변 영역 및 경쇄 가변 영역으로 구성된다. 일부 환경에서, 중쇄 가변 영역 및 경쇄 가변 영역은 동족 쌍이 아닐 수 있는데, 달리 말해, 상이한 결합 특이성을 가질 수 있다.

[0051] 예시적 특이적 항원 결합 단백질은 "이중 특이적 항체"이며, 2개 이상의 에피토프에 선택적으로 결합할 수 있는 항체를 포함한다. 이중 특이적 항체는 일반적으로 2개의 상이한 중쇄를 포함하며, 각각의 중쇄는 2개의 상이한 분자(예, 항원) 또는 동일한 분자(예, 동일한 항원) 상에서 상이한 에피토프에 특이적으로 결합한다. 이중 특이적 항체가 2개의 상이한 에피토프(제1 에피토프 및 제2 에피토프)에 선택적으로 결합할 수 있는 경우, 제1 에피토프에 대한 제1 중쇄의 가변 영역의 친화성은 일반적으로 제2 에피토프에 대한 제1 중쇄의 가변 영역의 친화성보다 적어도 10배 내지 100배, 또는 1,000배 내지 10,000배 낮아지거나 그 반대가 될 것이다. 이중 특이적 항체와 같은 이중 특이적 항원 결합 단백질은 동일한 항원의 상이한 에피토프를 인식하는 중쇄의 가변 영역을 포함할 수 있다. 일반적인 이중 특이적 항체는 2개의 중쇄를 갖는데, 이들 각각은 (N 말단에서 C 말단까지) CH1 도메인, 힌지, CH2 도메인 및 CH3 도메인이 뒤따르는 3개의 중쇄 CDR, 및 면역 글로불린 경쇄를 갖는데, 상기 면역 글로불린 경쇄는 항체 결합 특이성을 부여하지 않지만 각각의 중쇄와 결합할 수 있거나, 각각의 중쇄와 결합하면서 중쇄 항원 결합 영역에 의해 결합된 에피토프 중 하나 이상과 결합할 수 있거나, 각각의 중쇄와 결합할 수 있고 중쇄 중 하나 또는 둘 모두를 하나 또는 둘 모두의 에피토프에 결합시킬 수 있다. 일 구현예에서, Fc 도메인은 적어도 CH2 및 CH3을 포함한다. Fc 도메인은 힌지, CH2 도메인 및 CH3 도메인을 포함할 수 있다.

[0052] 구현된 하나의 이중 특이적 포맷은 제1 중쇄(HC), 변형된 CH3을 갖는 제2 중쇄(HC\*), 및 공통 경쇄(LC)(동일한 경쇄의 2개의 카피)를 포함한다. 또 다른 구현예에는 제1 중쇄(HC), 공통 LC 및 (제2 HC가 ScFv의 N 말단에 융합된) HC-ScFv 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 HC, 동족 LC, 및 (제2 HC가 ScFv의 N 말단에 융합된) HC-ScFv 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 중쇄(HC), LC 및 Fc 도메인이 포함된다. 또 다른 구현예에는 제1 HC, LC, 및 (Fc가 ScFv의 C 말단에 융합된) ScFv-Fc 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 HC, 공통 LC, 및 (Fc가 ScFv의 N 말단에 융합된) Fc-ScFv 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 HC, LC, 및 (제2 HC가 ScFv의 C 말단에 융합된) ScFv-HC가 포함된다.

[0053] 특정 구현예에서, 하나의 중쇄(HC)는 원시 서열 또는 "야생형" 서열일 수 있고, 제2 중쇄는 Fc 도메인에서 변형될 수 있다. 다른 구현예에서, 하나의 중쇄(HC)는 원시 서열 또는 "야생형" 서열일 수 있고, 제2 중쇄는 코돈 변형될 수 있다.

- [0054] 용어 "세포"는, 재조합 핵산 서열을 발현하는데 적합하고, 외인성 핵산을 안정적으로 통합시키고 발현을 강화할 수 있는 유전자좌를 갖는 임의의 세포를 포함한다. 세포는 비인간 동물 세포와 같은 포유류 세포, 인간 세포, 또는 예를 들어 하이브리도마 또는 쿼드로마와 같은 세포 융합체를 포함한다. 일부 구현예에서, 세포는 인간, 원숭이, 유인원, 햄스터, 랫트, 또는 마우스 세포이다. 일부 구현예에서, 세포는 다음의 세포로부터 선택된 포유류 세포이다: CHO(예를 들어, CHOK1, DXB-11 CHO, Veggie-CHO), COS(예를 들어, COS-7), 망막 세포, Vero, CV1, 신장(예를 들어, HEK293, 293 EBNA, MSR 293, MDCK, HaK, BHK), HeLa, HepG2, WI38, MRC 5, Colo205, HB 8065, HL-60(예를 들어, BHK21), Jurkat, Daudi, A431(상피), CV-1, U937, 3T3, L 세포, C127세포, SP2/0, NS-0, MMT 060562, Sertoli 세포, BRL 3A 세포, HT1080 세포, 골수종 세포, 종양 세포, 및 전술된 세포로부터 유래한 세포주. 일부 구현예에서, 세포는 하나 이상의 바이러스 유전자, 예를 들어, 바이러스 유전자를 발현하는 망막 세포(예를 들어, PER.C6<sup>TM</sup> 세포)를 포함한다.
- [0055] "세포 밀도"는, 예를 들어, mL 당 총 (생존 및 사멸) 세포 수로서의 시료의 부피 당 세포 수를 지칭한다. 세포 수는 수동으로 또는 자동으로, 예컨대 유동 세포 계측기로 계수할 수 있다. 자동화된 세포 계수기는 예를 들어 표준 트립판 블루 흡수 기술을 사용하여 생존 세포 또는 사멸 세포 또는 생존/사멸 세포의 수를 계수하기에 적합하도록 구성되었다. "생존 세포 밀도" 또는 "생존 세포 농도"라는 구문은 시료의 부피 당 생존 세포의 수를 지칭한다("생존 세포 계수"로서도 지칭됨). 세포 밀도를 결정하기 위해 임의의 수의 잘 알려진 수동 또는 자동화 기술이 사용될 수 있다. 커패시턴스 또는 광학 밀도가 부피 당 세포 수와 상관된, 배양물의 온라인 바이오매스 측정치가 측정될 수 있다. 생산 배양물과 같은 세포 배양물에서의 최종 세포 밀도는 출발 세포주에 따라, 예를 들어, 약 1.0 내지  $10 \times 10^6$  세포/mL의 범위 내에서 달라진다. 일부 구현예에서, 최종 세포 밀도는 생산 세포 배양물로부터 관심 단백질을 수확하기 전에 1.0 내지  $10 \times 10^6$  세포/mL에 도달한다. 다른 구현예에서, 최종 세포 밀도는 적어도  $5.0 \times 10^6$  세포/mL, 적어도  $6 \times 10^6$  세포/mL, 적어도  $7 \times 10^6$  세포/mL, 적어도  $8 \times 10^6$  세포/mL, 적어도  $9 \times 10^6$  세포/mL, 또는 적어도  $10 \times 10^6$  세포/mL에 도달한다.
- [0056] 용어 "코돈 변형"은 코돈에 의해 암호화된 아미노산을 변화시키지 않고 단백질 코딩 서열이 하나 이상의 뉴클레오타이드, 즉 하나 이상의 코돈에서 변형되어 뉴클레오타이드 서열의 코돈 변형된 버전이 생성되는 것을 의미한다. 뉴클레오타이드 서열의 코돈 변형은 핵산에 기초한 검정(예를 들어, 특히 하이브리드화에 기초한 검정, PCR 등)에서 뉴클레오타이드 서열을 그의 코돈 변형 버전과 구별하기에 편리한 기초를 제공할 수 있다. 일부 실시예에서, 뉴클레오타이드 서열의 코돈은 숙주 세포에서 암호화된 단백질의 개선되거나 최적화된 발현을 제공하도록 당업계에 잘 알려진 코돈 최적화 기술을 사용하여 변형된다(Gustafsson, C. 외의 2004, *Trends in Biotechnology*, 22:346-353; Chung, B.K.-S. 외의 2013, *Journal of Biotechnology*, 167:326-333; Gustafsson, C. 외의 2012, *Protein Expr Purif*, 83(1): 37-46). 이러한 기술을 사용하는 서열 설계 소프트웨어 툴 또한 당업계에 잘 알려져 있는데, 이에는, Codon Optimizer (Fuglsang A. 2003, *Protein Expr Purif*, 31:247-249), Gene Designer (Villalobos A, 외의 2006, *BMC Bioinforma*, 7:285), 및 OPTIMIZER (Puigbo P, 외의 2007, *Nucleic Acids Research*, 35:W126-W131)가 특히 포함되지만 이들로 한정되지 않는다.
- [0057] 문구 "상보성 결정 영역", 또는 용어 "CDR"은 보통 (즉, 야생형 동물에서) 면역 글로불린 분자(예: 항체 또는 T 세포 수용체)의 경쇄 또는 중쇄 가변 영역의 2개의 프레임워크 영역 사이에 나타나는, 유기체의 면역 글로불린 유전자의 핵산 서열에 의해 암호화된 아미노산 서열을 포함한다. CDR은 예를 들어, 생식계열 서열 또는 재배열되거나 재배열되지 않은 서열에 의해 암호화될 수 있으며, 예를 들어, 미처리 B 세포 또는 성숙한 B 세포 또는 T 세포에 의해 암호화될 수 있다. 일부 경우에(예: CDR3의 경우), CDR은, (예컨대, 재배열되지 않은 핵산 서열에서는) 서로 인접하지 않지만 B 세포 핵산 서열 내에서는 인접하는 2개 이상의 서열(예: 생식계열 서열)에 의해, 예를 들어 서열의 접합 또는 연결(예를 들어, 중쇄 CDR3을 형성하기 위한 V-D-J 재조합)의 결과로서 암호화될 수 있다.
- [0058] 용어 "발현 강화 유전자좌"는 세포의 게놈 내의 유전자좌로서, 적합한 유전자 또는 작제물이 서열/서열들 내에 또는 이에 가깝게 외인성으로 첨가(즉, 통합)되거나, 서열/서열들에 "작동 가능하게 연결"될 때, 게놈 내의 다른 영역 또는 서열과 비교하여 더 높은 레벨의 발현을 나타내는 서열/서열들을 함유하는 유전자좌를 지칭한다.
- [0059] "강화된"이라는 용어가 강화된 발현을 설명하도록 사용되는 경우, 예를 들어, 동일 발현 작제물의 단일 카피의 무작위 인티그란트(integrants) 풀과 비교하여, 외인성 서열을 게놈 내에 무작위 통합하거나 상이한 유전자좌에서 통합하는 것에 의해 일반적으로 관찰되는 것에 비해 발현이 적어도 1.5배 내지 적어도 3배 강화된 것을 포함

한다. 실질적으로 동일한 조건에서, 본 발명의 서열의 부재에서 동일한 유전자의 발현 수준에 비해, 예컨대 동일한 종의 계놈 내의 또 다른 유전자좌에서 통합시키는 경우에 비해, 본 발명의 서열을 사용할 때 수 배의 발현 강화가 관찰되었다. 강화된 재조합 효율은 (예를 들어, 재조합 효소-인식 부위("RRS")를 이용하여) 재조합할 수 있는 유전자좌의 능력 강화를 포함한다. 강화는 재조합 효소 인식 부위 등을 사용하지 않은 무작위 재조합(일반적으로 0.1%임)에 대한 재조합의 효율을 지칭한다. 바람직한 강화된 재조합 효율은 무작위에 비해 약 10배이거나 약 1%이다. 명시되지 않는 한, 청구된 발명은 특정한 재조합 효율로 제한되지 않는다. 강화된 발현 유전자좌는 일반적으로 숙주 세포에 의한 관심 단백질의 높은 생산성을 지지한다. 따라서, 강화된 발현은 단순히 배양된 세포의 높은 카피 수에 의해 높은 역가를 달성하기 보다는, 세포 당 많은 관심 단백질을 생산하는 것(상승된 그램 당 단백질 역가)을 포함한다. 특정 생산성 Qp(pg/세포/일, 즉 pcd)는 지속 가능한 생산성의 척도로 간주된다. 5 pcd보다 크거나, 10 pcd보다 크거나, 15 pcd보다 크거나, 20 pcd보다 크거나, 25 pcd보다 크거나, 심지어 30 pcd보다 큰 Qp를 나타내는 재조합 숙주 세포가 바람직하다. 발현 강화 유전자좌, 또는 "핫스팟(hotspot)" 내에 삽입된 관심 유전자를 갖는 숙주 세포는 높은 특정 생산성을 나타낸다.

[0060] 문구 "외인성으로 추가된 유전자", "외인성으로 추가된 핵산", 또는 간단히 "외인성 핵산"이 관심 유전자좌와 관련하여 사용되는 경우, 이러한 문구는 유전자좌가 자연에서 발견될 때 관심 유전자좌 내에 존재하지 않는 임의의 DNA 서열 또는 유전자를 지칭한다. 예를 들어, CHO 유전자좌(예컨대, 서열번호 1 또는 서열번호 2의 서열을 포함하는 유전자좌) 내의 "외인성 핵산"은 자연에서 특정 CHO 유전자좌 내에서 발견되지 않은 햄스터 유전자(즉, 햄스터 계놈 내의 또 다른 유전자좌의 햄스터 유전자), 임의의 다른 종의 유전자(예컨대, 인간 유전자), 키메라 유전자(예컨대, 인간/마우스), 또는 관심 CHO 유전자좌 내에 존재하는 것으로 자연에서 발견되지 않은 임의의 다른 유전자일 수 있다.

[0061] "중쇄" 또는 "면역 글로불린 중쇄"라는 문구는 임의의 유기체의 면역 글로불린 중쇄 불변 영역 서열을 포함하며, 달리 명시되지 않는 한 중쇄 가변 도메인을 포함한다. 중쇄 가변 도메인은 달리 명시되지 않는 한, 3개의 중쇄 CDR 및 4개의 FR 영역을 포함한다. 일반적인 중쇄는, (N 말단에서 C 말단까지) 가변 도메인에 이어서, CH1 도메인, 힌지, CH2 도메인 및 CH3 도메인을 갖는다. 용어 "중쇄의 단편" 또는 "중쇄 단편"(본원에서는 "HCF"로도 지칭됨)은 중쇄의 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90 또는 100개 이상의 아미노산으로 이루어진 펩티드를 포함하고, 하나 이상의 CDR, 하나 이상의 FR과 결합된 하나 이상의 CDR, CH1, 힌지, CH2 또는 CH3 중 하나 이상, 가변 영역, 불변 영역, 불변 영역의 단편(예: CH1, CH2, CH3) 또는 이들의 조합을 포함할 수 있다. 예시적인 HCF는 VH를 포함하고, Fc 영역의 전체 또는 일부를 포함한다. 문구 "HCF를 암호화하는 뉴클레오타이드 서열"은 HCF로 이루어지는 폴리펩티드를 암호화하는 뉴클레오타이드 서열 및 HCF를 함유하는 폴리펩티드(예를 들어, 특정 HCF 외에 추가 아미노산을 함유할 수 있는 폴리펩티드)를 암호화하는 뉴클레오타이드 서열을 포함한다. 예를 들어, HCF를 암호화하는 뉴클레오타이드 서열은, 특히, VH로 이루어지는 폴리펩티드, CH3에 연결된 VH로 이루어지는 폴리펩티드, 전체 중쇄로 이루어지는 폴리펩티드를 암호화하는 뉴클레오타이드 서열을 포함한다.

[0062] 핵산 서열의 맥락에서 "상동 서열"은 참조 핵산 서열과 실질적으로 상동인 서열을 지칭한다. 일부 구현예에서, 2개의 서열의 대응하는 뉴클레오타이드의 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상이 잔기의 관련 구간에 걸쳐 동일한 경우 2개의 서열은 실질적으로 상동인 것으로 간주된다. 일부 구현예에서, 관련 구간은 완전한 (즉, 전장) 서열이다.

[0063] "경쇄"라는 문구는 임의의 유기체의 면역 글로불린 경쇄 불변 영역 서열을 포함하며, 달리 명시되지 않는 한 인간 카파(kappa) 및 람다(lambda) 경쇄를 포함한다. 경쇄 가변 (VL) 도메인은 달리 명시되지 않는 한, 일반적으로 3개의 경쇄 CDR 및 4개의 프레임워크(FR) 영역을 포함한다. 일반적으로, 전장 경쇄는 아미노 말단에서 카복시 말단까지, FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4를 포함하는 VL 도메인, 및 경쇄 불변 도메인을 포함한다. 본 발명과 함께 사용될 수 있는 경쇄는, 예를 들어, 이중 특이적 항체에 의해 선택적으로 결합된 제1 에피토프 또는 제2 에피토프에 선택적으로 결합하지 않는 것들을 포함한다. 적절한 경쇄는 항체의 항원 결합 영역에 의해 결합된 하나 또는 둘 모두의 에피토프에 결합할 수 있거나 이의 결합에 기여할 수 있는 것들을 포함한다. 용어 "경쇄의 단편" 또는 "경쇄 단편"(또는 "LCF")는 경쇄의 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90 또는 100개 이상의 아미노산으로 이루어진 펩티드를 포함하고, 하나 이상의 CDR, 하나 이상의 FR과 결합된 하나 이상의 CDR, 가변 영역, 불변 영역, 불변 영역의 단편 또는 이들의 조합을 포함할 수 있다. 예시적인 LCF는 VL을 포함하고 경쇄 불변 영역("CL")의 전체 또는 일부를 포함한다. 문구 "LCF를 암호화하는 뉴클레오타이드 서열"은 LCF로 이루어지는 폴리펩티드를 암호화하는 뉴클레오타이드 서열 및 LCF를 함유하는 폴리펩티드(예를 들어, 특정 LCF 외에 추가 아미노산을 함유할 수 있는 폴리펩티드)를 암호화하는 뉴클레오타이드 서열을 포함한다. 예를 들어, LCF



를 암호화하는 뉴클레오타이드 서열은, 특히, VL로 이루어지거나 전체 경색로 이루어지는 폴리펩티드를 암호화하는 뉴클레오타이드 서열을 포함한다.

- [0064] 문구 "작동 가능하게 연결된"은 연결된 분자가 의도한대로 기능하는 방식으로 핵산 또는 단백질이 연결된 것을 지칭한다. DNA 영역은 이들이 서로 기능적으로 연관되어 있을 때 작동 가능하게 연결된다. 예를 들어, 프로모터는 상기 프로모터가 상기 서열의 전사에 참여할 수 있으면 암호화 서열에 작동 가능하게 연결되고; 리보솜-결합 부위는 번역이 가능하도록 배치되는 경우 암호화 서열에 작동 가능하게 연결된다. 일반적으로, 작동 가능하게 연결된다는 것은 인접성을 포함하나, 반드시 필요하지는 않다. 분비 리더(leaders)와 같은 서열의 경우, 접근 및 해독 틀(reading frame)에서 적절한 배치는 일반적인 특징이다. 관심 유전자좌의 발현 강화 서열은 관심 유전자(GOI)와 기능적으로 연관되는 경우, 예를 들어, 발현 강화 서열이 존재함으로써 GOI의 발현이 강화되는 경우, GOI에 작동 가능하게 연결된다.
- [0065] 관심 유전자좌, 예컨대 서열번호 1 또는 서열번호 2, 또는 이들의 단편을 기술할 때, "동일성 백분율"은 연속적인 상동성의 영역을 따라 인용된 동일성을 표시하는 상동 서열을 포함한다는 의미이지만, 비교된 서열에 상동성이 없는 갭, 결실 또는 삽입의 존재는 동일성 백분율의 계산에 고려되지 않는다.
- [0066] 본원에서 사용된 바와 같이, 예컨대 중 동족체(homolog)를 갖는 서열번호 1 또는 이의 단편 사이의 "동일성 백분율" 결정은 중 동족체가 비교될 상동 서열을 정렬에서 갖지 않는 경우 서열 비교를 포함하지 않을 것이다(즉, 경우에 따라, 서열번호 1 또는 이의 단편은 그 지점에 삽입을 갖거나, 중 동족체는 갭 또는 결실을 갖는다). 따라서, "동일성 백분율"은 갭, 결실 및 삽입에 대한 불이익을 포함하지 않는다.
- [0067] "인식 부위" 또는 "인식 서열"은 DNA 골격에 결합하고 DNA 골격의 부위 특이적인 절단을 유도하기 위해 뉴클레아제 또는 다른 효소에 의해 인식된 특정 DNA 서열이다. 엔도뉴클레아제는 DNA 분자 내에서 DNA를 절단한다. 인식 부위는 당업계에서 인식 표적 부위로도 지칭된다.
- [0068] "재조합 효소 인식 부위(또는 "RRS")"는 재조합 효소, 예컨대 Cre 재조합 효소(Cre) 또는 플립파아제(flp)에 의해 인식되는 특정 DNA 서열이다. 부위 특이적인 재조합 효소의 하나 이상의 표적 인식 서열이 유기체의 게놈 내에 전략적으로 배치되는 경우, 부위 특이적인 재조합 효소는 결실, 도치 및 전좌를 비롯한 DNA 재배열을 수행할 수 있다. 하나의 예시에서, Cre는 8-bp 스페이스에 의해 분리된 2개의 13-bp 역위 반복으로 이루어진 이의 DNA 표적 인식 부위 *loxP*에서 재조합 사건을 특이적으로 매개한다. 예컨대, DNA의 재조합-매개 교환을 용이하게 하기 위해 하나 이상의 재조합 효소 인식 부위가 사용될 수 있다. 재조합 효소 인식 부위, 예컨대, *lox* 부위의 변이체 또는 돌연변이가 또한 사용될 수 있다(Araki, N. 등, 2002, *Nucleic Acids Research*, 30:19, e103).
- [0069] "재조합 효소 매개 카세트 교환" 또는 "RMCE"는 게놈 표적 카세트를 공여자 카세트로 정확하게 치환하는 공정과 관련이 있다. 이러한 공정을 수행하기 위해 전형적으로 제공되는 분자 조성물은 (1) 특정 재조합 효소에 특이적인 인식 표적 부위에 의한 5' 및 3' 둘 다의 측부에 위치한 게놈 표적 카세트, (2) 인식 표적 부위의 매칭에 의해 측부에 위치한 공여자 카세트, 및 (3) 부위-특이적인 재조합 효소를 포함한다. 재조합 효소 단백질이 당업계에 널리 공지되어 있으며(Turan, S. 및 Bode J., 2011, *FASEB J.*, 25, pp. 4088-4107), 뉴클레오타이드의 수득 또는 손실 없이 특정 인식 표적 부위(DNA 서열) 내에서 DNA의 정확한 절단을 가능케 한다. 통상적인 재조합 효소/부위 조합은 비제한적으로 Cre/*lox* 및 Flp/*frrt*를 포함한다. 상업적으로 이용 가능한 키트도 R4-*attP* 부위를 함유하는 벡터 및 RMCE를 위해 phiC31 인테그라아제(integrase)를 암호화하는 벡터를 제공한다. (미국 공개 출원 제US20130004946 또한 예를 들어 참조한다.)
- [0070] "부위 특이적 통합" 또는 "표적화된 삽입"은 유전자 또는 핵산 서열을 게놈상의 특정 위치에 직접 삽입 또는 통합을 유도하도록, 즉, DNA를 인접한 폴리뉴클레오타이드 사슬 내의 2개의 뉴클레오타이드 사이의 특정 부위에 유도하도록 사용되는 유전자 표적화 방법을 지칭한다. 부위 특이적 통합 또는 표적화된 삽입은, 각각의 유전자가 자체 조절 요소(예컨대, 프로모터, 증강자 및/또는 전사적 말단 서열 등)를 갖는 다중 유전자와 같은 다중 발현 단위 또는 카세트를 포함하는 특정 핵산에 대해 수행될 수도 있다. "삽입" 및 "통합"은 상호교환적으로 사용된다. 유전자 또는 핵산 서열(예컨대, 발현 카세트를 포함하는 핵산 서열)의 삽입은 이용되는 유전자 편집 기법에 따라 하나 이상의 핵산의 치환, 또는 결실을 야기할 수 있다(또는 이를 위해 조작될 수 있다).
- [0071] "안정적 통합"은 숙주 세포의 게놈 내에 통합된 외인성 핵산이 세포 배양물 내에서 연속된 기간 동안, 예를 들어, 적어도 7일, 적어도 10일, 적어도 15일, 적어도 20일, 적어도 25일, 적어도 30일, 적어도 35일, 적어도 40일, 적어도 45일, 적어도 50일, 적어도 55일 또는 적어도 60일 이상 통합상태를 유지하는 것을 의미한다. 대규모 제조 및 정제를 위해 이중 특이적 항원 결합 단백질을 만드는 것이 어려운 과제라는 것을 이해해야 한다. 안

정성 및 복제 가능성은 임의의 생체분자의 재현성, 특히 치료학적으로 사용되는 생체분자의 재현성에 필수적이다. 본 개시의 방법에 의해 만들어진, 이중 특이적 항체를 발현하는 안정적 클론은 치료 생체분자를 생성하기 위한 일관되고 재현 가능한 방법을 제공한다.

[0072] **일반적인 설명**

[0073] 본 개시는 숙주 세포에서 발현 강화 유전자좌를 사용하여 숙주 세포, 특히 차이니즈 햄스터(*Cricetulus griseus*) 세포주에서 다중 폴리펩티드의 발현을 개선하는 조성물 및 방법을 제공한다. 보다 구체적으로, 본 개시는, 이중 특이적 항원 결합 단백질을 함께 암호화하는 다수의 외인성 핵산을 CHO 세포와 같은 숙주 세포에서 발현 강화 유전자좌 내의 특정 부위 내에 통합하도록 설계된 조성물 및 방법을 제공한다. 특히, 본 개시는 발현 강화 유전자좌 내의 특정 부위에 통합된 다수의 외인성 핵산을 함유하는 세포를 제공하며, 상기 다수의 외인성 핵산은 함께 이중 특이적 항원 결합 단백질을 암호화한다. 본 개시는 다수의 외인성 핵산을 발현 강화 유전자좌 내에 부위 특이적으로 통합하도록 설계된 핵산 벡터를 추가로 제공한다. 본 개시는, 발현 강화 유전자좌 내에 벡터로부터의 다수의 외인성 핵산을 부위 특이적으로 통합하기 위해 2개 이상의 재조합 효소 인식 부위(RRS)를 함유하는 숙주 세포, 및 일치하는 RRS와 다수의 외인성 핵산을 함유하는 벡터 세트를 포함하는 시스템을 추가적으로 제공한다. 또한, 본 개시는 본원에 개시된 세포, 벡터 및 시스템을 사용하여 이중 특이적 항원 결합 단백질을 만드는 방법을 제공한다.

[0074] **발현 강화 유전자좌 내의 특정 부위에 통합된 다수의 외인성 핵산을 갖는 세포**

[0075] 일 양태에서, 본 개시는 발현 강화 유전자좌 내의 특정 부위에 통합된 외인성 핵산 서열을 함유하는 세포를 제공하며, 상기 외인성 핵산 서열은 이중 특이적 항원 결합 단백질을 암호화한다.

[0076] 본원에 제공된 세포는 높은 역가 및/또는 높은 특정 생산성(pg/세포/일)으로 이중 특이적 항원 결합 단백질(예: 이중 특이적 항체)을 생산할 수 있다. 일부 구현예에서, 세포는 적어도 5 mg/L, 10 mg/L, 15 mg/L, 20 mg/L, 25 mg/L, 30 mg/L, 35 mg/L, 40 mg/L, 45 mg/L, 또는 50 mg/L 이상의 역가로 이중 특이적 항원 결합 단백질을 생산한다. 일부 구현예에서, 세포는 적어도 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 또는 50% 이상의 이중 특이적 항원 결합 단백질 역가 대 총 항원 결합 단백질 역가의 비율로 이중 특이적 항원 결합 단백질을 생산한다. 일부 구현예에서, 이중 특이적 항원 결합 단백질을 생산하는 세포는 일당 세포당 생산된 총 항원 결합 단백질(pg 단위)에 기초하여 결정된 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 피코그램/세포/일(pcd) 이상의 특정 생산성을 갖는다.

[0077] 발현 강화 유전자좌 내의 특정 부위에 통합된 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산을 포함하는 숙주 세포는 생산 배양물 중에서, 예를 들어 1 내지  $10 \times 10^6$  세포/mL의 높은 세포 밀도를 나타낸다. 다른 구현예에서, 이중 특이적 항원 결합 단백질을 암호화하는 숙주 세포는 적어도  $5 \times 10^6$  세포/mL,  $6 \times 10^6$  세포/mL,  $7 \times 10^6$  세포/mL,  $8 \times 10^6$  세포/mL,  $9 \times 10^6$  세포/mL, 또는  $10 \times 10^6$  세포/mL의 최종 세포 밀도를 갖는다.

[0078] 일부 구현예에서, 이중 특이적 항원 결합 단백질은 상이한 항원 특이성을 갖는 2개의 HC 단편("HCF") 및 2개의 LCF를 함유한다. 2개의 VL 영역이 사용되는 경우, 이들은 동일하거나 상이할 수 있다. 특이적 구현예에서, 2개의 VL 영역은 공통 경쇄와 같이 동일하다.

[0079] 일부 구현예에서, 2개의 HCF 각각은 CH1, CH2 또는 CH3와 같은 중쇄 불변 영역으로부터의 아미노산을 포함한다. 특이적 구현예에서, 2개의 HCF 각각은 CH3 도메인을 포함한다. 특정 구현예에서, 2개의 HCF 각각은 불변 영역, 즉 완전한 불변 영역을 포함한다.

[0080] 일부 구현예에서, 2개의 HCF 각각은 VH를 포함하고, 2개의 VH는 동일하거나 상이할 수 있다.

[0081] 일부 구현예에서, 이중 특이적 항원 결합 단백질은 2개의 중쇄(즉, 2개의 완전한 중쇄)를 포함한다.

[0082] 일부 구현예에서, 2개의 LCF 각각은 VL을 포함한다. 특이적 구현예에서, 각각의 LCF는 경쇄 불변 영역으로부터의 아미노산을 포함하는 아미노산 서열에 작동 가능하게 연결되는 VL 영역으로 구성된다. 특이적 구현예에서, 각각의 VL 영역은 CL 영역에 작동 가능하게 연결된다. 즉, 이중 특이적 항원 결합 단백질이 경쇄(즉, 완전한 경쇄)를 포함한다.

[0083] 일부 구현예에서, 발현 강화 유전자좌 내에 통합된 외인성 핵산 서열은, 제1 LCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 외인성 핵산, 제1 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제2 외인성 핵산, 및 제2

HCF를 암호화하는 뉴클레오타이드 서열을 함유하는 제3 외인성 핵산을 포함한다.

- [0084] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오타이드 서열은 경쇄 가변(VL) 영역 서열을 암호화할 수 있다. 특이적 구현예에서, 제1 VL 영역을 암호화하는 뉴클레오타이드 서열은 제1 경쇄를 암호화한다.
- [0085] 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 중쇄 불변 영역으로부터의 아미노산을 암호화하고 (예를 들어, CH1, 힌지, CH2, 또는 CH3 중 하나 이상을 암호화하고), 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 중쇄 불변 영역으로부터의 아미노산을 암호화한다. 제1 중쇄 불변 영역으로부터의 아미노산은 제2 중쇄 불변 영역으로부터의 아미노산과 동일하거나 상이할 수 있다. 예를 들어, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화하며, 제1 및 제2 CH3 도메인은 동일하거나, 이중 특이적 항원 결합 단백질에 대해 아래 본원에 기술된 바와 같이, 하나 이상의 아미노산 위치가 상이할 수 있다.
- [0086] 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 VH를 암호화하고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 VH를 암호화한다.
- [0087] 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 중쇄를 암호화하고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 중쇄를 암호화한다. 제1 및 제2 중쇄는 동일한 불변 영역을 가질 수 있거나, 하나 이상의 아미노산이 상이할 수 있다. 상이한 중쇄 불변 도메인(예: 상이한 CH3 도메인)을 갖는 이중 특이적 항원 결합 단백질의 다양한 예가 아래 본원에 더 기술된다. 암호화된 아미노산 서열과는 독립적으로, 2개의 중쇄 불변 영역으로부터의 아미노산을 암호화하는 뉴클레오타이드 서열은 2개의 뉴클레오타이드 서열 중 하나가 코돈 변형될 수 있다는 점에서 상이할 수 있는데, 이는 핵산에 기초한 검출 검정에서 2개의 뉴클레오타이드 서열을 구별하기 위한 편리한 기초를 제공한다.
- [0088] 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 프로모터를 함유하는 전사 조절 서열에 독립적으로 및 작동 가능하게 연결된다. "독립적으로"는, 각각의 코딩 서열이 프로모터와 같은 별도의 전사 조절 서열에 작동 가능하게 연결되어, 코딩 서열의 전사가 별도로 조절 및 제어된다는 것을 의미한다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들은 동일하다. 일부 구현예에서, 2개의 HCF를 함유하는 폴리펩티드의 전사를 유도하는 프로모터들뿐만 아니라 LCF를 함유하는 폴리펩티드의 전사를 유도하는 프로모터(예: CMV 프로모터)도 모두 동일하다. 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 유도 프로모터 또는 억제 프로모터에 독립적으로 및 작동 가능하게 연결된다. 유도 프로모터 또는 억제 프로모터는, 예를 들어, 생산 단계(유가 배양)에서만 생산이 이루어지도록 하고, 성장 단계(시드 트레인 배양) 동안에는 생산이 이루어지지 않도록 한다. 각각의 유전자 산물의 생산(발현)의 미세 조절은 상이한 프로모터에 의해 달성될 수 있다.
- [0089] 하나의 이러한 실시예에서, 세포가 테트라시클린 억제 단백질(TetR)을 발현하도록 먼저 조작되고, 각각의 HCF 및 LCF 코딩 뉴클레오타이드 서열을 TetR에 의해 활성이 조절되는 프로모터의 전사 조절 하에 놓인다. 2개의 탠덤 TetR 작동자(TetO)는 CMV 프로모터의 바로 하류에 위치한다. 일부 구현예에서, 각각의 HCF 및/또는 LCF 코딩 뉴클레오타이드 서열은 적어도 하나의 TetR 작동자(TetO) 또는 Arc 작동자(ArcO)의 상류에서 프로모터에 독립적으로 및 작동 가능하게 연결된다. 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 CMV/TetO 또는 CMV/ArcO 하이브리드 프로모터에 독립적으로 및 작동 가능하게 연결된다. 적절한 추가 프로모터는 아래 본원에 기술된다.
- [0090] 유전자좌 내에서의 다수의 외인성 핵산의 상대 위치는 변할 수 있다. 임의의 이론에 구속되고자 하는 것은 아니지만, 2개의 HCF 함유 폴리펩티드의 균형된 (즉, 비슷한) 발현 레벨을 달성하는 것은 중요한 것으로 여겨진다. 일부 구현예에서, LCF 암호화 핵산은 2개의 HCF 암호화 핵산 모두에 대해 상대적으로 상류에 위치한다. LCF 함유 폴리펩티드 및 2개의 HCF 함유 폴리펩티드의 발현을 유도하기 위한 3개의 프로모터가 동일한 경우, 적합한 배열은, 5'에서 3'까지, LCF를 암호화하는 뉴클레오타이드 서열, 제1 HCF를 암호화하는 뉴클레오타이드 서열, 뉴클레오타이드 서열(예: 선별 표지 유전자)에 작동 가능하게 연결된 추가적인 상이한 프로모터, 및 제2 HCF를 암호화하는 뉴클레오타이드 서열을 포함할 수 있다. 다른 적합한 배열은, 5'에서 3'까지, LCF를 암호화하는 뉴클레오타이드 서열, 뉴클레오타이드 서열(예: 선별 표지 유전자)에 작동 가능하게 연결된 추가적인 상이한 프로모터, 제1 HCF를 암호화하는 뉴클레오타이드 서열, 및 제2 HCF를 암호화하는 뉴클레오타이드 서열을 포함할 수 있다. HCF를 암호화하는 뉴클레오타이드 서열이 불변 영역 서열을 암호화하는 경우, 상류에 위치한 뉴클레오타이드 서열은 불변 영역 서열의 변형된 버전(예: 변형된 CH3)을 암호화할 수 있거나, 상류에 위치한 뉴클레오타이드 서열은 불변 영역 서열의 변형된 버전을 암호화하고 다른 하나는 불변 영역 서열의 미변형 버전을 암호화할 수 있다.

- [0091] 일부 구현예에서, 세포는 유전자와 내에 역시 통합된 하나 이상의 RRS를 추가로 함유한다. 일부 구현예에서, 세포는 서로 상이하고 외인성 핵산 서열의 측면에 위치하는 제1 및 제2 RRS를 포함하며, 외인성 핵산 서열은 차례로 제1 LCF 암호화 핵산, 제1 HCF 암호화 핵산, 및 제2 HCF 암호화 핵산을 함유한다. 특이적 구현예에서, LCF 암호화 핵산은 2개의 HCF 암호화 핵산 모두에 대해 상대적으로 상류에 위치하고, 세포는, 제1 LCF 암호화 핵산에 대해 상대적으로 3'에 위치하고 HCF 암호화 외인성 핵산 중 하나 또는 둘 모두에 대해 상대적으로 5'에 위치하는 제3 RRS를 포함하되, 제3 RRS는 제1 및 제2 RRS와 상이하다. 제3 RRS는, HCF 또는 LCF 암호화 서열 중 임의의 2개 사이에 배치될 수 있는 유전자 인트론에 포함되도록 조작될 수 있다.
- [0092] **이중 특이적 항원 결합 단백질**
- [0093] 본 개시에 기술된 세포, 벡터 및 시스템에서의 클로닝 및 생산에 적합한, 이중 특이적 항체와 같은 이중 특이적 항원 결합 단백질은 이중 특이적 항원 결합 단백질의 임의의 특정 포맷으로 한정되지 않는다.
- [0094] 다양한 구현예에서, 이중 특이적 항원 결합 단백질은 각각 항원 결합 모이어티 (예: VH 영역) 및 CH3 도메인을 함유하는 2개의 폴리펩티드를 포함하되, 2개의 폴리펩티드의 항원 결합 모이어티는 상이한 항원 특이성을 가지고, 2개의 CH3 도메인은 CH3 도메인 중 하나가 적어도 하나의 아미노산 위치가 변형되어 2개의 폴리펩티드 사이에서 분별 단백질 A 결합 특성을 생성한다는 점에서 서로에 대한 이중 이량체이다. 미국 특허 제8,586,713호에 기술된 이중 특이적 항체를 예를 들어 참조한다. 이러한 방식으로, 분별 단백질 A 단리법을 사용하여 동종 이량체로부터 이중 이량체 이중 특이적 항원 결합 단백질을 쉽게 단리할 수 있다.
- [0095] 일부 구현예에서, 이중 특이적 항원 결합 단백질은, 상이한 항원 특이성을 가지고 CH3 도메인에서 적어도 하나의 아미노산 위치가 상이한 2개의 중쇄를 포함하여, 2개의 중쇄 사이의 분별 단백질 A 결합 특성을 생성한다.
- [0096] 일부 구현예에서, 2개의 폴리펩티드는 인간 IgG의 CH3 도메인을 함유하되, 2개의 폴리펩티드 중 하나는 IgG1, IgG2 및 IgG4로부터 선택된 인간 IgG의 CH3 도메인을 함유하고, 2개의 폴리펩티드 중 다른 하나는 IgG1, IgG2 및 IgG4로부터 선택된 인간 IgG의 변형된 CH3를 함유하며, 상기 변형은 단백질 A에 대한 변형된 CH3 영역의 결합을 감소시키거나 제거한다. 특정 구현예에서, 2개의 폴리펩티드 중 하나는 인간 IgG1의 CH3 도메인을 함유하고, 2개의 폴리펩티드 중 다른 하나는 인간 IgG1의 변형된 CH3 도메인을 함유하되, 상기 변형은 IMGT 엑손 넘버링 시스템에서의 (i) 95R 및 (ii) 95R과 96F로 이루어지는 군으로부터 선택된다. 다른 특정 구현예에서, 변형된 CH3 도메인은 IMGT 엑손 넘버링 시스템에서의 16E, 18M, 44S, 52N, 57M, 및 82I로 이루어지는 군으로부터 선택된 1개 내지 5개의 추가 변형을 포함한다.
- [0097] 다른 다양한 구현예에서, 2개의 폴리펩티드는 마우스 IgG의 CH3 도메인을 함유하되, 2개의 폴리펩티드 중 하나는 변형되지 않은 마우스 IgG의 CH3 도메인을 함유하고, 2개의 폴리펩티드 중 다른 하나는 마우스 IgG의 변형된 CH3 도메인을 함유하며, 상기 변형은 단백질 A에 대한 변형된 CH3 영역의 결합을 감소시키거나 제거한다. 다양한 구현예에서, 마우스 IgG CH3 영역이 변형되어, 다음으로 이루어지는 군으로부터 선택된 특정 위치(EU 넘버링)에 특정 아미노산을 포함한다: 252T, 254T, 및 256T; 252T, 254T, 256T, 및 258K; 247P, 252T, 254T, 256T, 및 258K; 435R 및 436F; 252T, 254T, 256T, 435R, 및 436F; 252T, 254T, 256T, 258K, 435R, 및 436F; 247P, 252T, 254T, 256T, 258K, 435R, 및 436F; 및 435R. 특정 구현예에서, 다음으로 이루어지는 군으로부터 선택된 특정 변형 군이 만들어진: M252T, S254T, S256T; M252T, S254T, S256T, I258K; I247P, M252T, S254T, S256T, I258K; H435R, H436F; M252T, S254T, S256T, H435R, H436F; M252T, S254T, S256T, I258K, H435R, H436F; I247P, M252T, S254T, S256T, I258K, H435R, H436F; 및 H435R.
- [0098] 다양한 구현예에서, 이중 특이적 항원 결합 단백질은 마우스 및 랫트 단클론 항체 또는 항원 결합 단백질의 하이브리드, 예를 들어 마우스 IgG2 및 랫트 IgG2b의 하이브리드이다. 이들 구현예에 따르면, 이중 특이적 항체는 하나의 중쇄 쌍/경쇄 쌍 각각을 포함하는 2개의 항체 이중 이량체로 이루어지며, 이들의 Fc 부분을 통해 결합한다. 원하는 이중 이량체는 2개의 부모 항체 동종 이량체 및 이중 특이적 이중 이량체의 혼합물로부터 용이하게 정제될 수 있는데, 이는 단백질 A에 대한 이중 특이적 항체의 결합 속성이 부모 항체의 결합 속성과 다르고, 따라서 랫트 IgG2b는 단백질 A에 결합하지 않는 반면 마우스 IgG2a는 단백질 A에 결합하기 때문이다. 결과적으로, 마우스-랫트 이중 이량체는 단백질 A에 결합하지만, 마우스 IgG2a 동종 이량체보다 더 높은 pH에서는 용출되는데, 이는 이중 특이적 이중 이량체의 선택적 정제를 가능하게 한다.
- [0099] 다른 다양한 구현예에서, 이중 특이적 항원 결합 단백질은 당업계에서 "노브-인투-홀(knobs-into-holes)"로 지칭되는 포맷이다(예: 미국 특허 제7,183,076호 참조). 이들 구현예에서, 2개의 항체의 Fc 부분은 조작되어, 이



중 하나는 돌출 "노브"가 되고 다른 하나는 상보성 "홀"을 제공한다. 동일한 세포에서 생산될 때, 중쇄는 조작된 "노브"를 조작된 "홀"과 결합시킴으로써 동종 이량체 보다는 이종 이량체를 우선적으로 형성하는 것으로 알려져 있다.

- [0100] 또 다른 구현예에서, 제1 중쇄 및 제2 중쇄는 2개의 중쇄 사이의 상호 작용이 가능하도록 CH3 도메인에서 하나 이상의 아미노산 변형을 포함한다. CH3-CH3 계면 아미노산 잔기는 하전된 아미노산과 대체되어 정전기적으로 부적합한 동종 이량체를 형성할 수 있다. (PCT 공보 제W02009089004호, 및 유럽 공보 제EP1870459호를 예를 들어 참조한다.)
- [0101] 다른 구현예에서, 이종 이량체의 우선 형성을 촉진하기 위해 제1 중쇄는 동형 IgA의 CH3 도메인을 포함하고, 제2 중쇄는 IgG의 CH3 도메인을 포함한다(또는 그 반대이다). (PCT 공보 제W02007110205호를 예를 들어 참조한다.)
- [0102] 다른 구현예에서, Fab-아암 교환(PCT 공보 제W02008119353호; PCT 공보 제W02011131746호), 꼬인 나선 도메인 상호 작용(PCT 공보 제W02011034605호) 또는 류신 지퍼 펩티드(leucine zipper peptides)(Kostelny, 외의 *J. Immunol.* 1992, 148(5):1547-1553)와 같은 이종 이량체의 형성을 촉진시키는 조작 방법에 의해 다양한 포맷이 면역 글로불린 사슬과 통합될 수 있다.
- [0103] 이종 특이적 항원 결합 단백질 생성하는 데 사용될 수 있는 면역 글로불린 중쇄 가변 영역은 당업계에 공지된 임의의 방법을 사용해 생성할 수 있다. 예를 들어, 제1 중쇄는 제1 항원으로 면역화된 제1 동물의 성숙한 B 세포의 계능으로부터 유래된 핵산에 의해 암호화되는 가변 영역을 포함하고, 제1 중쇄는 제1 항원을 특이적으로 인식하며; 제2 중쇄는 제2 항원으로 면역화된 제2 동물의 성숙한 B 세포로부터 유래된 핵산에 의해 암호화되는 가변 영역을 포함하고, 제2 중쇄는 제2 항원을 특이적으로 인식한다. 면역 글로불린 중쇄 가변 영역 서열은 파지 디스플레이와 같은, 당업계에 공지된 임의의 다른 방법에 의해서도 획득될 수 있다. 다른 실시예에서, 중쇄 가변 영역을 암호화하는 핵산에는 당업계에서 기술되었거나 달리 이용 가능한 항체의 핵산이 포함된다. 일부 구현예에서, 2개의 중쇄 코딩 서열 중 하나는 핵산 기반의 검정에서 2개의 코딩 서열을 구별하기 위한 편리한 기초를 제공하기 위해 코돈 변형된 것이다.
- [0104] 2개의 상이한 에피토프(또는 2개의 상이한 항원)를 인식하는 2개의 중쇄를 포함하는 이종 특이적 항체는 이들이 동일한 경쇄(즉, 동일한 가변 및 불변 도메인을 갖는 경쇄)와 쌍을 이룰 수 있는 곳에서 더 쉽게 분리된다. 예를 들어, 미국 특허 제8,586,713호 및 본원에 기술된 기술분야에서 설명된 바와 같이, 중쇄 가변 도메인과 그의 표적 항원 간의 선택도 및/또는 친화성을 방해하지 않거나 실질적으로 방해하지 않으면서 특이성이 상이한 2개의 중쇄와 쌍을 이룰 수 있는 경쇄를 생성하는 다양한 방법이 당업계에 공지되어 있다.
- [0105] 이종 특이적 항원 결합 단백질은 이종 항원 특이성 및 관련된 유용한 용도를 가질 수 있다.
- [0106] 일부 실시예에서, 종양 항원 및 T 세포 항원에 대한 결합 특이성을 갖는 이종 특이적 항원 결합 단백질은 세포 상의 항원(예: CD20)을 표적화하고, T 세포 상의 항원(예: CD3과 같은 T 세포 수용체)를 또한 표적화하도록 만들어질 수 있다. 이러한 방식으로, 이종 특이적 항원 결합 단백질은 환자의 T 세포뿐만 아니라 환자에서의 관심 세포(예: CD20 결합을 통한 림프종 환자에서의 B 세포)를 표적화한다. 다양한 구현예에서, 이종 특이적 항원 결합 단백질은 T 세포 수용체에 결합할 때, 예컨대 CD3에 결합할 때 T 세포를 활성화시켜 T 세포 활성화를 특정한 선별 종양 세포에 결합시키도록 설계된다.
- [0107] 하나의 모이어티가 CD3에 결합하고 다른 하나의 모이어티가 표적 항원에 결합하는 이종 특이적 항원 결합 단백질의 맥락에서, 표적 항원은 종양 관련 항원일 수 있다. 특이적인 종양 관련 항원의 비한정적인 예에는, 예를 들어, AFP, ALK, BAGE 단백질, BIRC5 (서바이빈), BIRC7,  $\beta$ -카테닌, bcr-abl, BRCA1, BCMA, BORIS, CA9, 탄산무수화효소 IX, 카스파제-8, CALR, CCR5, CD19, CD20(MS4A1), CD22, CD30, CD40, CDK4, CEA, CLEC-12, CTLA4, 사이클린-B1, CYP1B1, EGFR, EGFRvIII, ErbB2/Her2, ErbB3, ErbB4, ETV6-AML, EpCAM, EphA2, Fra-1, FOLR1, GAGE 단백질(예: GAGE-1, -2), GD2, GD3, GloboH, 글리피칸-3, GM3, gp100, Her2, HLA/B-raf, HLA/k-ras, HLA/MAGE-A3, hTERT, LMP2, MAGE 단백질(예: MAGE-1, -2, -3, -4, -6, 및 -12), MART-1, 메소텔린, ML-IAP, Muc1, Muc2, Muc3, Muc4, Muc5, Muc16 (CA-125), MUM1, NA17, NY-BR1, NY-BR62, NY-BR85, NY-ESO1, OX40, p15, p53, PAP, PAX3, PAX5, PCTA-1, PLAC1, PRLR, PRAME, PSMA (FOLH1), RAGE 단백질, Ras, RGS5, Rho, SART-1, SART-3, Steap-1, Steap-2, TAG-72, TGF- $\beta$ , TMPRSS2, 톰슨-누벨 항원(Tn), TRP-1, TRP-2, 티로시나제, 및 유로플라킨-3이 포함된다.
- [0108] 일부 구현예에서, 이종 특이적 항원 결합 단백질은 (본원에 참조로서 통합된 미국 특허 출원 공개 번호 제

US2014/0088295A1호 및 제US20150266966A1에 기술된 바와 같은) 항CD3 x 항CD20 이중 특이적 항체, 항CD3 x 항Mucin 16 이중 특이적 항체(예: 항CD3 x 항Muc16 이중 특이적 항체), 및 항CD3 x 항전립선 특이적 막 항원 이중 특이적 항체(예: 항CD3 x 항-PSMA 이중 특이적 항체)를 포함한다. 다른 구현예에서, 이중 특이적 항원 결합 단백질은 CD3에 결합하는 하나의 모이어티를 포함한다. 예시적인 항CD3 항체 모이어티는 미국 특허 출원 공개 번호 제US2014/0088295A1호 및 제US20150266966A1호, 및 2017년 3월 30일 공개되고 그 전체가 본원에 참조로서 통합된 국제 공보 제WO 2017/053856호에 기술되어 있다. 또 다른 구현예에서, 이중 특이적 항원 결합 단백질은 CD3에 결합하는 하나의 모이어티 및 BCMA, CD19, CD20, CD28, CLEC-12, Her2, HLA 단백질, MAGE 단백질, Muc16, PSMA, 또는 Steap-2에 결합하는 하나의 모이어티를 포함한다.

[0109] 하나의 모이어티가 T 세포 수용체에 결합하고(예컨대 CD3에 결합하고) 다른 하나가 표적 항원에 결합하는 이중 특이적 항원 결합 단백질의 맥락에서, 표적 항원은 감염성 질병 관련 항원일 수 있다. 감염성 질환 관련 항원의 비한정적인 예에는, 예를 들어, 바이러스 입자의 표면에서 발견되거나 바이러스에 감염된 세포 상에서 우선적으로 발견되는 항원이 포함되며, 상기 바이러스는 HIV, 간염 (A, B 또는 C), 헤르페스 바이러스 (예: HSV-1, HSV-2, CMV, HAV-6, VZV, 엡스타인 바 바이러스), 아데노 바이러스, 인플루엔자 바이러스, 플라비 바이러스, 에코 바이러스, 코감기 바이러스, 콕사키 바이러스(coxsackie virus), 코로나 바이러스, 호흡기 세포융합 바이러스, 이차선염 바이러스, 로타 바이러스, 홍역 바이러스, 풍진 바이러스, 파보 바이러스, 우두 바이러스, HTLV, 뎅기 바이러스, 유두종 바이러스, 연체 동물 바이러스, 폴리오 바이러스, 광견병 바이러스, JC 바이러스 및 아르보 바이러스성 뇌염 바이러스로 이루어지는 군으로부터 선택된다. 대안적으로, 표적 항원은 박테리아의 표면에서 발견되거나, 박테리아에 감염된 세포 상에서 우선적으로 발견되는 항원일 수 있으며, 상기 박테리아는 클라미디아(chlamydia), 리케차(rickettsia), 마이코 박테리아(mycobacteria), 포도상 구균, 연쇄상 구균, 폐렴 구균, 수막 구균, 임균, 클렙시엘라(klebsiella), 프로테우스(proteus), 세라티아(serratia), 슈도모나스(pseudomonas), 레지오넬라(legionella), 디프테리아(diphtheria), 살모넬라균(salmonella), 간균(bacilli), 콜레라(cholera), 파상풍(tetanus), 보툴리누스 중독(botulism), 탄저병(anthrax), 역병(plague), 랩토스피라(leptospira), 및 라임 병균(Lyme disease bacteria)로 이루어지는 군으로부터 선택된다. 특정 구현예에서, 표적 항원은 곰팡이의 표면에서 발견되거나 곰팡이에 감염된 세포 상에서 우선적으로 발견되는 항원이며, 상기 곰팡이는 칸디다(알비칸스, 크루세이, 글라브라타, 트리파칼리스 등), 크립토코커스 네오포르만스(Cryptococcus neoformans), 아스페리길루스(푸미가투스, 니제르 등), 털곰팡이목(털곰팡이속, 활털곰팡이속, 거미줄 곰팡이속 등), 스포로트릭스 첸키아이(Sporothrix schenckii), 블라스토미세스 더마티티디스(Blastomyces dermatitidis), 파라코시디오이디즈 브라질리엔시스(Paracoccidioides brasiliensis), 콕시디오이데스 이미티스(Coccidioides immitis), 및 히스토플라스마속(Histoplasma capsulatum)으로 이루어지는 군으로부터 선택된다. 특정 구현예에서, 표적 항체는 기생충의 표면 상에서 발견되거나 기생충에 감염된 세포 상에서 우선적으로 발견되는 항원이며, 상기 기생충은 이질 아메바(Entamoeba histolytica), 대장 발란티듐(Balantidium coli), 파울러 자 유아메바(Naegleria fowleri), 가시 아메바종(Acanthamoeba sp.), 람블 편모충(Giardia lamblia), 크립토스포리디움종(Cryptosporidium sp.), 주폐포자충(Pneumocystis carinii), 삼일열원충(Plasmodium vivax), 바베시아 미코티(Babesia microti), 트리파노소마 브루세이(Trypanosoma brucei), 트리파노소마 크루지(Trypanosoma cruzi), 리슈만 편모충(Leishmania donovani), 톡소포자충(Toxoplasma gondii), 쥐모양 선충(Nippostrongylus brasiliensis), 붉은 여우 조충(Taenia crassiceps), 및 말레이 사상충(Brugia malayi)으로 이루어지는 군으로부터 선택된다. 특정한 병원균 관련 항원의 비한정적 예에는, 예를 들어, HIV gp120, HIV CD4, B형 간염 당단백질 L, B형 간염 당단백질 M, B형 간염 당단백질 S, C형 간염 E1, C형 간염 E2, 간세포 특이적 단백질, 단순 포진 바이러스 gB(herpes simplex virus gB), 시토크로 바이러스 gB(cytomegalovirus gB), 및 HTLV 외피 단백질(HTLV envelope protein)이 포함된다.

[0110] 동일한 세포의 표면 상에서 각각이 결합 파트너를 향하는 (즉, 각각이 상이한 표적을 향하는) 2개의 결합 모이어티를 포함하는 이중 특이적 결합 단백질도 만들 수 있다. 이러한 디자인은 동일한 세포의 표면 상에서 표적 모두를 발현하는 특정 세포 또는 세포 유형을 표적으로 하기에 특히 적합하다. 표적이 다른 세포 상에서 개별적으로 나타날 수 있다 하더라도, 이들 결합 단백질의 결합 모이어티는 각각의 결합 모이어티가 비교적 낮은 (예를 들어, 낮은 나노몰 또는 가령, 500, 600, 700, 800 나노몰과 같은 100 나노몰 KD보다 큰 높은 나노몰) 친화성으로 그의 표적에 결합하도록 선택된다. 이러한 방식으로, 장기간의 표적 결합은 2개의 표적이 동일한 세포 상에서 근접한 상황에서만 선호된다.

[0111] 각각 동일한 표적의 상이한 에피토프에서 동일한 표적에 결합하는 2개의 결합 모이어티를 포함하는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인은 결합 단백질로 표적을 성공적으로 차단하는 확률을 최대화하는데 특히 적합하다. 예를 들어, 막관통 경로 또는 세포 표면 수용체의 다수의 세포 외 루프는 동일한 이중 특이적

결합 분자에 의해 표적화될 수 있다.

[0112] 면역 신호 전달의 음성 조절체를 모으고 활성화시켜 면역을 억제시키는 2개의 결합 모이어티를 포함하는 이중 특이적 결합 단백질을 만들 수 있다. 표적이 동일한 세포 상에 있는 경우 cis의 억제제가 달성될 수 있고, 표적이 상이한 세포 상에 있는 경우 trans의 억제제가 달성될 수 있다. 예를 들어, cis의 억제는 항IgGRIIb 결합 모이어티 및 항FelD1 결합 모이어티를 갖는 이중 특이적 결합 단백질로 달성될 수 있으므로, IgGRIIb는 FelD1에 대한 면역 반응을 하향 조절하기 위해 FelD1의 존재 하에서만 모인다. 예를 들어, trans의 억제는 항BTLA 결합 모이어티 및 조직 특이적 관심 항원에 특이적으로 결합하는 결합 모이어티를 갖는 이중 특이적 결합 단백질로 달성될 수 있으므로, 억제성 BTLA 분자의 클러스터링은 자기 면역 질병을 잠재적으로 다루는 선택된 표적 조직에서만 발생한다.

[0113] 다중 성분 수용체를 활성화시키는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인에서, 수용체의 2개의 성분을 향하는 2개의 결합 모이어티는 수용체에 결합하고, 가교 결합하며, 수용체로부터의 신호 전달을 활성화한다. (결합이 수용체를 가교 결합시키는 경우) 이는, 예를 들어, IFNAR1에 결합하는 결합 모이어티 및 IFNAR2에 결합하는 결합 모이어티를 갖는 이중 특이적 결합 단백질을 사용해 수행될 수 있다. 이러한 이중 특이적 결합 단백질은 인터페론 치료에 대한 대안을 제공할 수 있다.

[0114] 반투과성 배리어, 예를 들어, 혈액-뇌 배리어를 가로질러 결합 모이어티를 운반하는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인에서, 하나의 결합 모이어티는 특정 선택성 배리어를 통과할 수 있는 표적에 결합하고, 다른 하나의 모이어티는 치료 활성을 갖는 분자를 표적화하되, 치료 활성을 갖는 표적 분자는 배리어를 정상적으로 통과할 수 없다. 이러한 종류의 이중 특이적 결합 단백질은 치료제가 달리 도달할 수 없는 조직에 치료제를 전달하는데 유용하다. 일부 실시예는 장 또는 폐 내에 치료제를 운반하기 위해 pIGR 수용체를 표적화하거나, 혈액-뇌 배리어를 가로질러 치료제를 운반하기 위해 트랜스페린 수용체를 표적화하는 것을 포함한다.

[0115] 특이적 세포 또는 세포 유형 내에 결합 모이어티를 운반하는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인에서, 하나의 결합 모이어티는 세포 내로 쉽게 내부화되는 세포 표면 단백질(예: 수용체)을 표적화한다. 다른 하나의 모이어티는 치료 효과를 일으키는 세포간 단백질을 표적화한다.

[0116] 대식 면역 세포의 표면 수용체 및 감염성 병원균의 표면 분자(예: 효모 또는 박테리아)에 결합하여 대식 면역 세포의 근위에 감염성 병원균을 운반함으로써 병원균의 식균 작용을 용이하게 하는 이중 특이적 결합 단백질. 이러한 디자인의 예는 CD64 또는 CD89 분자 및 병원균을 또한 표적화하는 이중 특이적 항체일 것이다.

[0117] 항체 가변 영역을 하나의 결합 모이어티로서 가지고 비-Ig 모이어티를 제2 결합 모이어티로서 갖는 이중 특이적 결합 단백질. 항체 가변 영역은 표적화를 달성하는 반면, 비-Ig 모이어티는 Fc에 연결된 작동자(effector) 또는 독소(toxin)이다. 이러한 방식으로, 리간드(예: 작동자 또는 독소)가 항체 가변 영역에 의해 결합된 표적에 전달된다.

[0118] 임의의 2개의 단백질 모이어티가 Fc의 맥락에서 서로의 근위에 전달되도록 Ig 영역(예: CH2 및 CH3 영역을 함유하는 Ig 서열)에 각각 결합된 2개의 모이어티를 가지는 이중 특이적 결합 단백질. 이러한 디자인의 예는 트랩, 예를 들어 동종 또는 이종 이량체 트랩 분자가 포함된다.

#### [0119] **발현 강화 유전자좌**

[0120] 본 발명에서의 용도에 적합한 발현 강화 유전자좌는, 예를 들어, 서열번호 1에 대해 실질적 상동성을 갖는 뉴클레오타이드 서열을 갖는, 미국 특허 제8,389,239호에 기술된 바와 같은, 유전자좌("EESYR<sup>®</sup> 유전자좌"로서도 본원에 지칭됨), 서열번호 2 또는 서열번호 3에 대해 실질적 상동성을 갖는 뉴클레오타이드 서열을 포함하는, 미국 출원 일련 번호 제14/919,300호에 기술된 바와 같은, 유전자좌("YARS 유전자좌"로서도 본원에 지칭됨), 및 당업계에서 문서화된 (예: US 20150167020A1, 및 미국 특허 제6,800,457호) 다른 발현 강화 유전자좌 및 서열을 포함한다.

[0121] 일부 구현예에서, 본 발명에 사용된 발현 강화 유전자좌는, 서열번호 1에 대해 실질적 상동성을 갖는 뉴클레오타이드 서열을 포함하는 유전자좌, 서열번호 2 또는 서열번호 3에 대해 실질적 상동성을 갖는 뉴클레오타이드 서열을 포함하는 유전자좌로부터 선택된다. 이들 유전자좌는 서열에 대한 작동 가능한 연결 내에 (즉, 서열 내 또는 서열에 인접한 범위 내에) 통합된 유전자의 발현 강화를 제공할 뿐 아니라, 게놈 내의 다른 서열과 비교하여 더 큰 제조 효율 및 강화된 통합 안정성을 나타낸다.

- [0122] 서열번호 1, 서열번호 2 및 서열번호 3은 CHO 세포로부터 동정되었다. (예컨대 인간 또는 마우스와 같은) 다른 포유류 종은 동정된 발현 강화 영역에 대해 제한된 상동성을 갖는 것으로 확인되었지만, 상동 서열은 중국 햄스터(*Cricetulus griseus*,)의 다른 조직 유형 또는 다른 상동의 종으로부터 유래된 세포주에서 발견될 수 있고, 당업계에 널리 공지된 기법에 의해 단리될 수 있다. 예컨대, 중간 혼성화 또는 PCR-기반 기법에 의해 다른 상동 서열을 동정할 수 있다. 또한, 당업계에 널리 공지된 부위 지향적 또는 무작위 돌연변이 유발 기술에 의해 서열번호 1, 서열번호 2 또는 서열번호 3에 제시된 뉴클레오타이드 서열을 변경할 수 있다. 이어서, 생성된 서열 변이체는 발현 강화 활성에 대해 시험될 수 있다. 발현 강화 활성을 갖는 서열번호 1, 서열번호 2 또는 서열번호 3과 핵산 동일성이 적어도 약 90% 동일한 DNA는 통상적인 실험에 의해 단리되며, 발현 강화 활성을 나타낼 것으로 기대된다.
- [0123] 통합 부위, 하나 이상의 외인성 핵산이 삽입된 부위 또는 뉴클레오타이드 위치는 발현 강화 서열(예컨대, 서열번호 1, 서열번호 2 또는 서열번호 3) 중 어느 하나의 내에 있거나 이에 인접한 임의의 위치일 수 있다. 관심 유전자와 내 또는 이에 인접한 특정 염색체 위치가 통합된 외인성 유전자의 안정한 통합 및 효율적인 전사를 지원 하는지 여부는 당업계에 널리 공지된 표준 절차에 따라, 예를 들어 미국 특허 제8,389,239호 및 미국 특허 출원 일련번호 제14,919,300호에 기술된 바와 같이 결정될 수 있다.
- [0124] 본원에서 고려되는 통합 부위는 발현 강화 서열 내, 또는 서열에 근접한 범위 내, 예를 들어, 염색체 DNA 상의 발현 강화 서열의 위치에 대해 약 1 kb 미만, 500 염기쌍(bp), 250 bp, 100 bp, 50 bp, 25 bp, 10 bp, 또는 약 5 bp 미만의 상류(5') 또는 하류(3')에 위치한다. 또 다른 일부 구현예에서, 사용된 통합 부위는 염색체 DNA 상의 발현 강화 서열의 위치에 대해 약 1000, 2500, 또는 5000 이상의 염기쌍 상류(5') 또는 하류(3')에 위치한다.
- [0125] 큰 게놈 영역, 예컨대 지지체/기질 부착 영역이 염색체 DNA의 효율적인 복제 및 전사를 위해 사용되는 것으로 당업계에서 이해된다. 지지체-부착 영역(SAR) 또는 기질-결합 또는 기질 부착 영역(MAR)으로도 일컬어지는 것으로 공지된 지지체/기질 부착 영역(S/MAR)은 핵 기질이 부착되는 진핵 생물 게놈 DNA 영역이다. 임의의 하나의 이론에 구속됨이 없이, S/MAR은 전형적으로 비-암호화 영역에 도표화(mapping)되고, 주어진 전사 영역(예컨대, 염색질 도메인)을 이의 인접으로부터 분리하고, 전사를 가능하게 하는 인자, 예컨대 DNAses 또는 중합효소 인식 부위를 위한 기구 및/또는 결합을 위한 플랫폼이 또한 제공된다. 몇몇 S/MAR은 길이가 약 14 내지 20 kb인 것으로 특성분석되었다(Klar 등, 2005, *Gene* 364:79-89). 이와 같이, 발현 강화 유전자좌에서의 (예컨대, 서열번호 1 또는 서열번호 2, 또는 서열번호 3 내에서 또는 그 근처에서의) 유전자 통합이 강화된 발현을 제공할 것으로 기대된다. 일부 구현예에서, 발현 강화 유전자좌 내의 특이적 부위에서 통합된 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산 서열을 포함하는 숙주 세포는 높은 특이적 생산성을 나타낸다. 다른 구현예에서, 이중 특이적 항원 결합 단백질을 암호화하는 숙주 세포는 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 또는 30 피코그램/세포/일(pcd)의 특이적 생산성을 갖는다.
- [0126] 일부 구현예에서, 통합 부위는 서열번호 1의 뉴클레오타이드 서열을 포함하는 유전자좌 내에 있다. 특이적 구현예에서, 통합 부위는 서열번호 1의 뉴클레오타이드 서열 내, 또는 이에 대한 근접한 범위 내에 있다. 특정 구현예에서, 통합 부위는 10-13,515; 20-12,020; 1,020-11,020; 2,020-10,020; 3,020-9,020; 4,020-8,020; 5,020-7,020; 6,020-6,920; 6,120-6,820; 6,220-6,720; 6,320-6,620; 6,420-6,520; 6,460-6,500; 6,470-6,490; 및 6,475-6,485로 넘버링 된 뉴클레오타이드의 걸침 위치들로부터 선택된 서열번호 1 내의 위치에 있다. 다른 구현예에서, 통합 부위는 서열번호 1의 뉴클레오타이드 5,000-7,400, 5,000-6,500, 6,400-7,400 및 서열번호 1의 뉴클레오타이드 6,400-6,500으로 이루어지는 군으로부터 선택되는 서열 내에 있다. 특이적 구현예에서, 서열번호 1의 뉴클레오타이드 6471 내지 6473의 "act" 삼중항 전, 후, 또는 내의 통합 부위.
- [0127] 일부 구현예에서, 통합 부위는 서열번호 2 또는 서열번호 3의 뉴클레오타이드 서열을 포함하는 유전자좌 내에 있다. 특이적 구현예에서, 통합 부위는 서열번호 2의 뉴클레오타이드 서열 내, 또는 이에 대한 근접한 범위 내에 있다. 특이적 구현예에서, 통합 부위는 서열번호 3의 뉴클레오타이드 서열 내, 또는 이에 대한 근접한 범위 내에 있다. 일부 구현예에서, 통합 부위는 서열번호 3의 뉴클레오타이드 1990-1991, 1991-1992, 1992-1993, 1993-1994, 1995-1996, 1996-1997, 1997-1998, 1999-2000, 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2008, 2008-2009, 2009-2010, 2010-2011, 2011-2012, 2012-2013, 2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018, 2018-2019, 2019-2020, 2020-2021 또는 2021-2022 내에 있다. 특이적 구현예에서, 통합은 서열번호 3의 뉴클레오타이드 2001~2022에 있거나 그 내에 있다. 일부 구현예에서, 외인성 핵산은 서열번호 3의 뉴클레오타이드 2001~2002 또는 뉴클레오타이드 2021~2022 뉴클레오타이드에서 또는 그 안에 삽



입되고, 삽입의 결과 서열번호 3의 뉴클레오티드 2002~2021이 결실된다.

[0128] **발현 강화 유전자좌 내로의 부위 특이적 통합**

[0129] 발현 강화 유전자좌 내에, 즉 본원에 기술된 바와 같은 발현 강화 유전자좌 내의 하나의 특정 부위 내에 다수의 외인성 핵산을 부위 특이적 방식으로 통합시키는 것은, 예를 들어, 상동 재조합에 의한 방식 및 당업계에서 기술된 재조합 효소 매개 카세트 교환에 의한 방식을 비롯한 여러 가지 방식으로 달성될 수 있다 (미국 특허 제 8,389,239호 및 본원에 기술된 종래 기술을 예를 들어 참조한다).

[0130] 일부 구현예에서, 다수의 외인성 핵산을 함유하는 핵산 서열 또는 관심 유전자의 통합에 편리한 발현 강화 유전자좌 내에 적어도 2개, 즉 2개 이상의 상이한 재조합 효소 인식 서열(RRS)을 함유하는 세포가 제공된다. 이러한 세포는, 아래 본원 및 당업계에 기술된 바와 같은 상동 재조합을 포함하는 다양한 수단, 예를 들어 미국 특허 제 8,389,239호 및 그에 개시된 기술에 의해 2개 이상의 RRS를 함유하는 외인성 핵산 서열을 원하는 유전자좌 내에 도입함으로써 수득될 수 있다.

[0131] 특이적 구현예에서, 다수의 외인성 핵산의 통합에 편리한 발현 강화 유전자좌 내에 3개 이상의 재조합 효소 인식 서열(RRS)을 함유하는 세포가 제공된다. 특정 구현예에서, 예를 들어 2개의 별개 외인성 핵산의 통합을 매개할 수 있는 발현 강화 유전자좌 내에 3개의 상이한 재조합 효소 인식 서열(RRS)을 함유하는 세포로서, 게놈 내의 5' RRS 및 중간 RRS는 통합 대상 제1 외인성 핵산의 측면에 위치하는 5' RRS 및 3' RRS와 일치하고, 게놈 내의 중간 RRS 및 3' RRS는 통합 대상 제2 외인성 핵산의 측면에 위치하는 5' RRS 및 3' RRS와 일치하는 세포가 제공된다.

[0132] 적합한 RRS는 *LoxP*, *Lox511*, *Lox5171*, *Lox2272*, *Lox2372*, *Loxm2*, *Lox-FAS*, *Lox71*, *Lox66* 및 이들의 돌연변이를 포함하는 군으로부터 선택될 수 있으며, 부위 특이적 재조합 효소는 Cre 재조합 효소이거나 이의 유도체가 재조합 효소 매개 카세트 교환(RMCE)을 달성하는 데 사용된다. 다른 실시예에서, 적합한 RRS는 FRT, F3, F5, FRT 돌연변이-10, FRT 돌연변이+10 및 이들의 돌연변이를 포함하는 군으로부터 선택될 수 있으며, 이러한 시나리오에서는, 부위 특이적 재조합 효소 Flp 재조합 효소 또는 이의 유도체가 RMCE를 달성하는 데 사용된다. 또 다른 실시예에서, RRS는 attB, attP 및 이들의 돌연변이를 포함하는 군으로부터 선택될 수 있으며, 이 경우에는, 부위 특이적 재조합 효소 phiC31 인테그라아제 또는 이의 유도체가 RMCE를 달성하는 데 사용된다.

[0133] 다른 구현예에서, 원시 세포가 상동 재조합 기술에 의해 변형되어 다수의 외인성 핵산을 함유하는 핵산 서열을 발현 강화 유전자좌 내의 특이적 부위 내에 통합시킨다.

[0134] 상동 재조합의 경우, 상동성 폴리뉴클레오티드 분자(즉, 상동성 아암)가 정렬되어 이들의 서열 구간을 교환한다. 전이유전자가 상동성 게놈 서열에 의해 측부에 위치할 경우, 이러한 교환 동안 전이 유전자가 도입될 수 있다. 하나의 실시예에서, 재조합 효소 인식 부위는 상동 재조합에 의해 통합 부위에서 숙주 세포 게놈 내에 도입될 수 있다. 다른 실시예에서, 다수의 외인성 핵산, 예를 들어 이중 특이적 항원 결합 단백질질을 함께 암호화하는 다수의 핵산으로서, 표적 유전자좌에서의 서열과 상동인 서열("상동성 아암")이 측면에 위치하는 핵산을 함유하는 핵산 서열이 숙주 게놈 내에 삽입된다.

[0135] 진핵 세포에서의 상동 재조합은 상기 통합 부위에서 염색체 DNA에 절단을 도입함으로써 촉진될 수 있다. 이는 통합 특정 부위에 특정 뉴클레아제를 표적화하여 수행될 수 있다. 상기 표적 유전자좌에서 DNA 서열을 인식하는 DNA-결합 단백질은 당해 기술 분야에 알려져 있다. 유전자 표적화 벡터는 또한 상동 재조합을 용이하게 하는데 사용된다.

[0136] 상동 재조합을 달성하기 위한 유전자 표적화 벡터의 작제 및 뉴클레아제의 선택은 본 발명이 속하는 당업자의 기술 범위 내에 있다. 일부 실시예에서, 모듈 구조를 가지며 개별 징크 핑거 도메인(zinc finger domains)을 함유하는 징크 핑거 뉴클레아제(ZFN)는 표적 서열(예를 들어, 표적화 통합 부위) 내 특정 3-뉴클레오티드 서열을 인식한다. 일부 구현예에서, 다수의 표적 서열을 표적화하는 개별 징크 핑거 도메인의 조합과 함께 ZFN이 사용될 수 있다. 전사 촉진제 유사(TAL) 이펙터 뉴클레아제(TALEN) 또한 부위 특이적 게놈 편집을 위해 사용될 수 있다. TAL 이펙터 단백질 DNA-결합 도메인은 일반적으로 FokI과 같은 제한 뉴클레아제의 비-특이적 절단 도메인과 조합하여 사용된다. 일부 구현예에서, TAL 이펙터 단백질 DNA-결합 도메인 및 제한 뉴클레아제 절단 도메인을 포함하는 융합 단백질은 본 발명의 유전자좌 내 표적 서열에서 DNA를 인식하고 절단하는데 사용된다(Boch J 등, 2009 *Science* 326:1509-1512). RNA 유전자가위(RNA-guided endonucleases; RGEN)는 세균 적응성 면역 기구로부터 개발된 프로그램 가능한 게놈 조작 툴이다. 이 시스템-간헐적으로 반복되는 회문 구조 염기 서열 집합체(the clustered regularly interspaced short palindromic repeats; CRISPR)/CRISPR-관련(Cas) 면역-에

서 두 개의 RNA로 복합체를 형성할 때 상기 단백질 Cas9는 서열-특이적 엔도뉴클레아제를 형성하며 이들 RNA 중 하나는 표적 선별을 유도한다. RGEN은 구성성분(Cas9 및 tracrRNA) 및 표적-특이적 CRISPR RNA(crRNA)로 이루어진다. DNA 표적 절단 효율 및 절단 부위의 위치 모두는 프로토스페이스 인접 모티프(protospacer adjacent motif; PAM)의 위치, 표적 인식에 대한 추가 요건에 따라 다양하다(원고 M113.539726로서 2014년 3월 14일자 온라인 공개된 Chen, H. 등의 *J. Biol. Chem.*). 특정 표적 유전자좌에 대해 고유한 서열(예컨대, 서열번호 1, 서열번호 2, 또는 서열번호 3)은, 16~17 염기쌍 일치로 잠재적인 표적 외 부위를 밝혀낼 수 있는 CHO 계통에 이들 서열 중 많은 서열을 정렬함으로써 동정될 수 있다.

[0137] 일부 구현예에서, 5' 및 3' 상동 아암이 측면에 위치하는 관심 핵산(예: 하나 이상의 선별 표지 유전자의 측면에 선택적으로 위치하는 하나 이상의 RRS를 함유하는 핵산, 또는 함께 이중 특이적 항원 결합 단백질을 암호화하는 다수의 외인성 핵산을 함유하는 핵산)을 가지는 표적화 벡터가 하나 이상의 추가 벡터 또는 mRNA를 갖는 세포 내에 도입된다. 하나의 구현예에서, 하나 이상의 추가 벡터 또는 mRNA는 징크 핑거 뉴클레아제(ZFN), ZFN 이량체, 전사 촉진제 유사 이펙터 뉴클레아제(TALEN), TAL 이펙터 도메인 융합 단백질 및 RNA-유도된 DNA 엔도뉴클레아제를 포함하되 이들로 한정되지 않는 부위 특이적 뉴클레아제를 암호화하는 뉴클레오티드 서열을 함유한다. 특정 구현예에서, 하나 이상의 벡터 또는 mRNA는 가이드 RNA, tracrRNA 및 Cas 효소를 암호화하는 뉴클레오티드 서열을 포함하는 제1 벡터, 및 공여자(외인성) 뉴클레오티드 서열을 포함하는 제2 벡터를 포함한다. 이러한 공여자 서열은 관심 유전자를 암호화하는 뉴클레오티드 서열, 또는 인식 서열, 또는 표적화된 삽입을 위한 임의의 하나의 이러한 외인성 요소를 포함하는 유전자 카세트를 함유한다. mRNA가 사용되는 경우, mRNA는 당업자에게 공지된 통상의 형질감염 방법에 의해 세포 내로 형질감염될 수 있으며, 효소, 예컨대 전이효소 또는 엔도뉴클레아제를 암호화할 수 있다. 세포 내로 도입되는 mRNA는 일시적일 수 있고 계통 내로 통합되지 않지만, mRNA는 통합이 발생하기 위해 필수적이거나 유익한 외인성 핵산을 운반할 수 있다. 일부 경우에, mRNA는 부속 폴리뉴클레오티드의 부작용이 장기간 지속되는 임의의 위험을 없애기 위해 선택되는데, 이 경우, 핵산의 원하는 통합을 달성하기 위해서는 단기간의 발현만이 요구된다.

#### [0138] 부위 특이적 통합을 위한 벡터

[0139] 부위 특이적 통합을 통해 외인성 핵산을 발현 강화 유전자좌 내에 도입하기 위한 핵산 벡터가 본원에 제공된다. 적합한 벡터는, RMCE를 통한 통합을 위해 RRS가 측면에 위치한 외인성 핵산 서열을 함유하도록 설계된 벡터, 및 상동 재조합을 통한 통합을 위해 상동성 아암이 측면에 위치한 외인성 관심 핵산을 함유하도록 설계된 벡터를 포함한다.

[0140] 다양한 구현예에서, 벡터는 RMCE를 통한 부위 특이적 통합을 달성하기 위해 제공된다. 일부 구현예에서, 벡터는 표적 유전자좌 내에 다수의 핵산의 동시 통합을 달성하도록 설계된다. 순차 통합과 대조적으로, 동시 통합은 이중 특이적 항원 결합 단백질을 생산하는 바람직한 클론의 효율적이고 신속한 단리를 가능하게 한다.

[0141] 일부 구현예에서, 벡터 세트가 제공되는데, 벡터 세트는 각각 하나 이상의 핵산의 측면에 위치하는 적어도 2개의 RRS를 함유하는 2개 이상의 벡터를 포함하되, 벡터 세트에서의 상기 핵산은 함께 이중 특이적 항원 결합 단백질을 암호화한다.

[0142] 일 구현예에서, 벡터 세트는, (5'에서 3'까지) 제1 RRS, 제1 LCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 핵산, 및 제3 RRS를 함유하는 제1 벡터; 및 (5'에서 3'까지) 제3 RRS, 제1 HCF를 암호화하는 뉴클레오티드 서열을 포함하는 제2 핵산, 및 제2 RRS를 함유하는 제2 벡터를 포함하되; 제1 또는 제2 핵산 중 하나는 제2 HCF를 암호화하는 뉴클레오티드를 추가로 포함하고; 제1 및 제2 HCF, 및 제1 LCF는 이중 특이적 항원 결합 단백질의 영역(예: 가변 영역)을 암호화한다. 일부 구현예에서, 제2 HCF를 암호화하는 뉴클레오티드 서열은 제1 벡터 상의 제1 핵산(즉, 벡터 상의 제1 LCF 및 제2 HCF)에 포함되고, 예를 들어 제1 LCF를 암호화하는 뉴클레오티드 서열의 하류에 선택적으로 배치되며; 다른 구현예에서, 제2 HCF를 암호화하는 뉴클레오티드 서열은 제2 벡터 상의 제2 핵산(하나의 벡터 상의 제1 HCF 및 제2 HCF)에 포함된다.

[0143] HCF를 암호화하는 뉴클레오티드 서열은 아미노산, 예를 들어 불변 영역으로부터의 아미노산 또는 도메인(들)을 암호화하거나, 전체 불변 영역을 암호화할 수 있다. 특이적 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오티드 서열은 하나 이상의 불변 영역, 예컨대, CL, CH1, 힌지 CH2, CH3 또는 이들의 조합을 암호화할 수 있다. 특정 구현예에서, HCF를 암호화하는 뉴클레오티드 서열은 CH3 도메인을 암호화할 수 있다. 예를 들어, 제1 HCF를 암호화하는 뉴클레오티드 서열은 제1 CH3 도메인을 암호화할 수 있고, 제2 HCF를 암호화하는 뉴클레오티드 서열은 제2 CH3 도메인을 암호화할 수 있다. 제1 및 제2 CH3 도메인은 동일하거나, 적어도 하나의 아미노산이 상이하다. CH3 도메인 또는 불변 영역에서의 차이(예: 상이한 단백질 A 결합 특성, 정전 스티어링(electrostatic

steering), 또는 "노브 앤 홀(knob-and-hole)" 포맷으로 나타나는 차이)는 본원에 기술된 이중 특이적 항원 결합 단백질에 대해 임의의 포맷을 취할 수 있다. 임의의 아미노산 서열 차이와 별개로, 2개의 HCF 코딩 뉴클레오타이드 서열은 하나 이상의 뉴클레오타이드 서열이 코돈 최적화되었다는 것이 또한 상이하다.

[0144] 일부 구현예에서, 각각의 HCF 코딩 뉴클레오타이드 서열은 가령 프로모터를 함유하는 전사 조절 서열에 독립적으로 및 작동 가능하게 연결된다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들은 동일하다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들뿐만 아니라 LCF 함유 폴리펩티드의 전사를 유도하는 프로모터(예: CMV 프로모터)도 모두 동일하다. 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 유도 프로모터 또는 억제 프로모터에 독립적으로 및 작동 가능하게 연결된다. 유도 프로모터 또는 억제 프로모터는, 예를 들어, 생산 단계(유가 배양)에서만 생산이 이루어지도록 하고, 성장 단계(시드 트레인 배양) 동안에는 생산이 이루어지지 않도록 한다. 유도 프로모터 또는 억제 프로모터는 또한 하나 이상의 관심 유전자의 차별적 발현이 이루어지도록 한다. 일부 구현예에서, 각각의 HCF 및/또는 LCF 코딩 뉴클레오타이드 서열은 적어도 하나의 TetR 작동자(TetO) 또는 Arc 작동자(ArcO)의 상류에서 프로모터에 독립적으로 및 작동 가능하게 연결된다. 다른 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 CMV/TetO 또는 CMV/ArcO 하이브리드 프로모터에 독립적으로 및 작동 가능하게 연결된다. 하이브리드 프로모터(조절 융합 단백질로도 지칭됨)의 실시예는 2003년 12월 11일에 공개된 국제 공보 제W003101189A1호에서 (본원에 참조로서 통합됨) 확인할 수 있다.

[0145] 일부 구현예에서, 제1 벡터에서의 제1 핵산은 제1 벡터에서의 제3 RRS에 대해 5'에 위치한 선별 표지 유전자의 5' 부분을 추가로 포함하고, 제2 핵산은 제2 벡터에서의 제3 RRS에 대해 3'에 위치한 선별 표지 유전자의 나머지 3' 부분을 추가로 포함한다. 이들 구현예에서, 제1, 제2 및 제3 RRS는 제1 및 제2 핵산의 부위 특이적 통합을 매개하는데, 이는 편리한 선별을 위해, 적절히 및 동시 통합된 클론에서의 선별 표지 유전자의 5' 부분과 3' 부분을 결합시킨다. 특정 구현예에서, 제1 벡터에서의 제3 RRS는 선별 표지 유전자 인트론의 5' 부분 내에 있도록 설계되고; 제2 벡터에서의 제3 RRS는 선별 표지 유전자 인트론의 3' 부분 내에 있도록 설계된다. 또 다른 구현예에서, 제1 벡터에서의 제3 RRS는 제3 RRS가 작동 가능하게 연결되는 (다른 하나의 벡터로부터는 분리됨) 선별 표지 유전자와 프로모터 사이에 있도록 설계되고, 제1 벡터에서의 제3 RRS는 프로모터의 3'에 있도록 설계되며, 제2 벡터에서의 제3 RRS는 선별 표지 유전자의 5'에 있도록 설계된다.

[0146] 전술한 벡터 세트는 3개 이상의 벡터를 포함할 수 있다. 예를 들어, 전술한 2개의 벡터에 추가하여, 세트는 제2 LCF를 암호화하는 뉴클레오타이드 서열의 측면에 위치하는 적어도 2개의 RRS를 포함하는 제3 벡터를 포함할 수 있다. 벡터 세트는 RRS를 인식하는 하나 이상의 재조합 효소를 암호화하는 벡터를 포함할 수도 있다.

[0147] 다른 구현예에서, 벡터는 상동 재조합을 통해 부위 특이적 통합을 달성하도록 제공된다. 일부 실시예에서, 숙주 게놈 내에 통합될 폴리뉴클레오타이드 서열은 이중 특이적 항원 결합 단백질을 암호화하는 핵산의 후속 통합을 위해 바람직한 유전자좌에 통합된 하나 이상의 RRS를 갖는 세포를 생성하기 위한 DNA 서열, 예컨대 RRS 또는 하나 이상의 선별 표지 유전자의 측면에 위치하는 다수의 RRS일 수 있다. 다른 실시예에서, 숙주 게놈 내에 통합될 폴리뉴클레오타이드 서열은 이중 특이적 항원 결합 단백질을 함께 암호화하는 다수의 핵산을 포함한다. 예를 들어, 폴리뉴클레오타이드 서열은 이중 특이적 항체의 2개의 상이한 중쇄 및 공통 경쇄를 암호화하는 핵산을 포함한다. 일부 구현예에서, 이중 특이적 항원 결합 단백질을 함께 암호화하는 다수의 핵산은 각각 독립하여 (즉, 별도로) 조절 서열(예: 프로모터, 증강자, 전사 종결 서열, 또는 이의 조합)에 작동 가능하게 연결된다. 즉, 다수의 핵산 각각에 대한 조절 서열(예: 프로모터)은 동일하거나 상이할 수 있는 (즉, 동일하거나 상이한 뉴클레오타이드 서열을 함유하는) 별도의 것이다. 다수의 핵산 중 하나의 핵산이 다수의 코딩 서열을 포함하는 경우, 폴리펩티드의 N 말단 부분에 대해 코딩하는 각각의 코딩 서열 또는 각각의 뉴클레오타이드 서열은 자체 조절 서열(예: 프로모터)에 독립적으로 및 작동 가능하게 연결된다.

[0148] 발현 강화 유전자좌 내의 서열과 상동인 서열을 선별하고, 선별된 서열을 표적 벡터에 상동 아암으로서 포함시키는 것은 당업자의 기술 범위 내에 있다. 일부 구현예에서, 벡터 또는 작제물은 제1 상동 아암 및 제2 상동 아암을 포함하며, 결합된 제1 및 제2 상동 아암은 유전자좌 내에서 내인성 서열을 치환하는 표적화 서열을 포함한다. 또 다른 구현예에서, 제1 및 제2 상동 아암은 유전자좌 내의 내인성 서열 내에 통합되거나 삽입되는 표적화 서열을 포함한다. 일부 구현예에서, 상동 아암은 서열번호 1, 서열번호 2, 또는 서열번호 3에 존재하는 뉴클레오타이드 서열과 상동인 뉴클레오타이드 서열을 포함한다. 특이적 구현예에서, 벡터는 서열번호 3의 뉴클레오타이드 1001~2001에 상응하는 뉴클레오타이드 서열을 갖는 5' 상동 아암, 및 서열번호 3의 뉴클레오타이드 2022~3022에 상응하는 뉴클레오타이드를 갖는 3' 상동 아암을 함유한다. 상동 아암, 예컨대 제1 상동 아암(5' 상동 아암으로도 불림) 및 제2 상동 아암(3' 상동 아암으로도 불림)은 유전자좌 내에서의 표적화 서열에 상동이다. 상동 아암은,



적어도 1 kb, 또는 적어도 약 2 kb, 또는 적어도 약 3 kb, 또는 적어도 약 4 kb, 또는 적어도 5 kb, 또는 적어도 약 10 kb를 포함하는 유전자와 내에서의 영역 또는 표적화 서열을 5'에서 3'까지 확장시킬 수 있다. 다른 구현예에서, 제1 및 제2 상동 아암으로부터 선택된 표적화된 서열의 총 뉴클레오티드 수는 1 kb 이상, 또는 약 2 kb 이상, 또는 약 3 kb 이상, 또는 약 4 kb 이상, 또는 5 kb 이상, 또는 약 10 kb 이상이다. 일부 예에서, 5' 상동 아암과 3' 상동 아암(표적화 서열과 상동임) 사이의 거리는 적어도 5 bp, 10 bp, 20 bp, 30 bp, 40 bp, 50 bp, 60 bp, 70 bp, 80 bp, 90 bp, 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 또는 적어도 1 kb, 또는 적어도 약 2 kb, 또는 적어도 약 3 kb, 또는 적어도 약 4 kb, 또는 적어도 5 kb, 또는 적어도 약 10 kb를 포함한다. 서열번호 3의 뉴클레오티드 1001~2001 및 2022~3022가 5' 및 3' 상동 아암으로서 선택된 예에서, 2개의 상동성 아암 사이의 거리는 20 뉴클레오티드(서열번호 3의 뉴클레오티드 2002~2021에 상응함)일 수 있으며; 이러한 상동 아암은 서열번호 3을 포함하는 유전자와 내에서의 (예를 들어, 서열번호 3의 뉴클레오티드 1990~2021 또는 2002~2021 내에서의) 외인성 핵산 서열의 통합과, 서열번호 3의 뉴클레오티드 2002~2021의 동시 결실을 매개할 수 있다.

[0149] 발현 강화 유전자와 내로의 부위 특이적 통합을 위한 외인성 핵산을 도입하기 위한, 본원에 개시된 벡터는, 외인성 관심 핵산 및 암호화된 폴리펩티드의 발현을 유도하고, 외인성 관심 핵산이 성공적으로 통합된 세포의 선별하고 동정하기 위한 추가 유전자 및 서열을 포함할 수 있다. 이러한 추가 서열은, 예를 들어, 아래 본원에도 기술된 전사 및 번역 조절 서열, 선별 표지 유전자 등을 포함한다.

#### [0150] 조절 서열

[0151] 발현 강화 유전자와 내에 외인성 핵산을 부위 특이적 방식으로 도입하기 위한, 본원에 개시된 벡터 및 부위 특이적 통합의 결과로서 획득된 세포는 외인성 관심 핵산 및 암호화된 폴리펩티드의 발현을 유도하기 위한 조절 서열을 포함할 수 있다. 조절 서열은 전사 프로모터, 증강자, 적합한 mRNA 리보솜 결합 부위를 암호화하는 서열, 및 전사 및 번역의 종결을 제어하는 서열을 포함한다. 전사 및 번역 조절 서열은 바이러스 공급원에 의해 제공될 수 있다. 예를 들어, 공통적으로 사용된 프로모터 및 증강자는 폴리오마(polyoma), 아데노바이러스(adenovirus) 2, 시미안 바이러스 40 (SV40), 마우스 또는 인간 사이토메갈로바이러스 (CMV), CMV 극초기 (CMV-IE) 또는 CMV 주요 IE (CMV-MIE) 프로모터뿐만 아니라 RSV, SV40 후기 프로모터, SL3-3, MMTV, 유비퀴틴 (Ubi), 유비퀴틴 C (Ubc) 및 HIV LTR 프로모터와 같은 바이러스로부터 유래한다. 이러한 제어 서열이 선택된 숙주 세포와 양립할 수 있다면, 바이러스 게놈 프로모터, 제어 및/또는 신호 서열이 발현을 유도하는 데 사용될 수 있다. 관심 단백질이 발현될 세포 유형에 따라, 비-바이러스 세포 프로모터(예컨대, b-글로빈 및 EF-1a 프로모터)가 사용될 수도 있다. SV40 바이러스 게놈, 예를 들어 SV40 기원, 초기 및 후기 프로모터, 증강자, 스플라이스 및 폴리아데닐화 부위로부터 유래된 DNA 서열을 사용해서 외인성 DNA 서열의 발현에 유용한 다른 유전 요소를 제공할 수 있다. 초기 및 후기 프로모터는 둘 다 SV40 바이러스 복제 기점을 또한 포함하는 단편으로서 SV40 바이러스에서 쉽게 획득되기 때문에 특히 유용하다(Fiers 등의 *Nature* 273:113, 1978). 더 작거나 더 큰 SV40 단편이 사용될 수도 있다. 통상적으로, Hind III 부위로부터 SV40 복제 기점에 위치한 BglI 부위를 향해 연장되는 약 250 bp의 서열이 포함된다. (예를 들어, 화학적 화합물, 보조 인자, 조절 단백질에 의해 유도된) 유도성 프로모터는 항원 결합 단백질의 생산이 성장 단계(시드 트레인 배양) 동안이 아닌 생산 단계(유가 배양)에서만 일어나게 하는데 사용되거나 이에 특히 유용하다. 유도성 프로모터의 예에는 알코올 탈수소 효소 I 유전자 프로모터, 테트라시클린 반응성 프로모터 시스템, 글루코코르티코이드 수용체 프로모터, 에스트로젠 수용체 프로모터, 엑디손 수용체 프로모터, 메탈로티오네인 기반 프로모터 및 T7 폴리머라제 기반 프로모터가 포함된다. 바이시스트로닉 벡터(bicistronic vector)를 통해 다수의 전사물을 발현시키는 데 적합한 서열이 전술되었으며(Kim S. K. 및 Wold B. J., *Cell* 42:129, 1985), 본 발명에 사용될 수 있다. 단백질의 멀티시스트론(multicistronic) 발현을 위한 적합한 전략의 예에는 2A 펩티드의 사용(Szymczak 등, *Expert Opin Biol Ther* 5: 627-638 (2005)) 및 내부 리보솜 개시 부위("IRES")가 포함되며, 둘 모두는 당업계에 널리 공지되어 있다. 미국 특허 제4,634,665호(Axel 등) 및 미국 특허 제4,656,134호(Ringold 등)에 기재된 것들과 같은 다른 유형의 발현 벡터도 유용할 것이다.

#### [0152] 선별 표지

[0153] 외인성 핵산을 발현 강화 유전자와 내에 부위 특이적 방식으로 도입하기 위한, 본원에 개시된 벡터, 및 부위 특이적 통합의 결과로서 획득된 세포는 하나 이상의 선별 표지 유전자를 포함할 수 있다.

[0154] 일부 구현예에서, 선별 표지 유전자는, 예를 들어, Kaufman, R. J. (1988) *Meth Enzymology* 185:537의 표 1에 기술된 것들과 같은 약물 내성, DHFR-MTX 내성을 참조하며, DHFR-MTX 내성, P-당단백질 및 다중 약물 내성(MDR)

- 다양한 친유성 세포독성 제제(예: 아드리마이신, 콜히친, 빈크리스틴), 및 아데노신 탈아미노화 효소(ADA)-Xyl-A 또는 아데노신 및 2'-디옥시코포르마이신을 포함한다. 다른 우성 선별 표지는 미생물에 의해 유도된 항생제 내성 유전자, 예컨대 네오마이신, 카나마이신 또는 하이그로마이신 내성을 포함한다. 포유류 숙주에 대해 여러 가지 적절한 선별 시스템이 존재한다(전술된 Sambrook, pgs 16.9-16.15). 2개의 우성 선별 표지를 사용하는 동시-형질감염 프로토콜이 또한 기술되어 있다(Okayama 및 Berg, *Mol. Cell Biol* 5:1136, 1985).

[0155] 다른 구현예에서, 선별 표지 유전자는, 경우에 따라 성공적으로 삽입 및/또는 치환되었거나 성공적으로 삽입 및/또는 대체되지 않은 유전자 카세트의 인식을 위한 검출 가능한 신호를 제공하거나 생성할 수 있는 폴리펩티드를 암호화한다. 적합한 예에는 형광 표지 또는 단백질, 특히 검출 가능한 신호를 생성하는 화학 반응을 촉매하는 효소가 포함된다. 디스코소마 산호(DsRed), 녹색 형광 단백질(GFP), 향상된 녹색 형광 단백질(eGFP), 시안 형광 단백질(CFP), 향상된 시안 형광 단백질(eCFP), 황색 형광 단백질(YFP), 향상된 황색 형광 단백질(eYFP) 및 원-적색 형광 단백질(예컨대, mKate, mKate2, mPlum, mRaspberry 또는 E2- crimson)을 포함하되 이들로 한정되지 않는 형광 표지의 예가 당업계에 널리 공지되어 있다. 예를 들어, Nagai, T., 등의 2002 *Nature Biotechnology* 20:87-90; Heim, R. 등의 23 February 1995 *Nature* 373:663-664; 및 Strack, R.L. 등의 2009 *Biochemistry* 48:8279-81을 또한 참조한다.

#### [0156] 이중 특이적 항원 결합 단백질 제작 시스템

[0157] 추가적인 양태에서, 본 개시는, 세포 및 하나 이상의 박테리의 조합을 포함하는 시스템으로서, 함께 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산이 발현 강화 유전자와 내에 통합된 세포를 만드는데 사용할 수 있는 시스템을 제공한다. 상기 시스템은, 예를 들어, 키트의 형태로 제공될 수 있다.

[0158] 일부 구현예에서, 시스템은, 효율적인 박테리가 구성될 수 있고 RMCE를 통해 다수의 외인성 핵산이 발현 강화 유전자와 내에 동시에 통합될 수 있도록 설계된다. 동시 통합은 원하는 클론이 신속하게 단리될 수 있게 하는데, 안정한 세포주의 생성을 위해서는 발현 강화 유전자좌를 사용하는 것도 중요하다.

[0159] 일부 구현예에서, RMCE를 통해 다수의 외인성 핵산을 통합하도록 설계된 전술한 박테리 세트, 및 발현 강화 유전자와 내의 특정 부위에서 통합되고 박테리 세트의 RRS와 일치하는 RRS를 함유하는 세포 중 임의의 하나를 포함하는 시스템이 제공된다. 예를 들어, 시스템은 세포 및 박테리 세트를 포함하되, 세포는 5'에서 3'까지 발현 강화 유전자와 내에 통합된 제1 RRS, 제1 외인성 핵산, 제2 RRS, 제2 외인성 핵산, 및 제3 RRS를 포함하되, 상기 3개의 RRS는 서로 상이하고; 박테리 세트는 5'에서 3'까지 제1 RRS, 제1 LCF(예: 제1 VL)를 암호화하는 뉴클레오타이드 서열을 포함하는 제1 핵산, 및 제2 RRS를 포함하는 제1 박테리, 및 제2 RRS, 제1 HCF(예: 제1 VH)를 암호화하는 뉴클레오타이드 서열을 포함하는 제2 핵산, 및 제3 RRS를 포함하는 제2 박테리를 포함하되, 상기 제1 핵산 또는 제2 핵산 중 하나는 제1 HCF(예: 제2 VH)를 암호화하는 뉴클레오타이드 서열을 추가로 포함한다. 세포 내로 박테리가 도입될 때, 박테리에서의 제1 및 제2 핵산은 제1, 제2 및 제3 RRS에 의해 매개된 재조합을 통해 발현 강화 유전자와 내에 통합된다. 박테리로부터 유전자좌와 내에 적절히 통합된 핵산을 갖는 형질감염체의 스크리닝을 용이하게 하기 위해, 시스템의 세포에서의 제1 외인성 핵산은 제1 선별 표지 유전자를 포함할 수 있고, 세포에서의 제2 외인성 핵산은 제2 선별 표지 유전자를 포함할 수 있으며, 제1 및 제2 선별 표지 유전자는 서로 상이하고 박테리에 의해 제공된 임의의 선별 표지 유전자와도 상이하며, 특이적 구현예에서, 제1 및 제2 선별 표지 유전자는 (음성 선별을 가능하게 하는) 형광 단백질을 암호화하고, 박테리 상의 제1 및 제2 핵산은 분할 포맷의 추가 선별 표지 유전자를 제공하여 양성 선별을 가능하게 한다. 의도된 재조합으로 클론을 단리하는 효율은 제한될 수 있지만(약 1%), 음성 선별만으로도 신속한 클론 단리가 가능할 수 있다. 양성 선별(새로운 형광제, 또는 약물 내성, 또는 항생제)과 결합된 음성 선별은 양성 클론의 단리 효율을 (약 80%까지) 상당히 개선할 수 있다.

[0160] 시스템은 추가적인 성분, 시약, 또는 정보, 예를 들어, 시스템의 박테리(들)를 형질감염에 의해 시스템의 세포 내에 도입하기 위한 프로토콜을 포함할 수 있다. 비제한적인 형질감염 방법은 화학-기반 형질감염 방법을 포함하고, 이는 리포좀; 나노 입자; 칼슘 포스페이트(Graham 등 (1973) *Virology* 52 (2): 456-67, Bacchetti 등 (1977) *Proc Natl Acad Sci USA* 74 (4): 1590-4 및 Kriegler, M (1991) *Transfer and Expression: A Laboratory Manual*. New York: W. H. Freeman and Company. pp. 96-97); 덴드리머; 또는 양이온 중합체, 예컨대 DEAE-텍스트란 또는 폴리에틸렌이민을 포함한다. 비화학적 방법은 전기천공법; 초음파천공법; 및 광학 형질감염을 포함한다. 입자-기반 형질감염은 유전자 총의 자석 보조 형질감염을 포함한다(Bertram, J. (2006) *Current Pharmaceutical Biotechnology* 7, 277-28). 바이러스 방법이 또한 형질감염에 사용될 수 있다. mRNA 전달은 트랜스메신저(TransMessenger) 및 트랜스아이티(TransIT®)를 사용하는 방법을 포함한다(Bire 등 *BMC Biotechnology* 2013, 13:75). 이중 DNA를 세포 내로 도입하는데 통상적으로 사용되는 방법 중 하나는, 예컨대

Wigler 등에 의해 기술된 바(Proc. Natl. Acad. Sci. USA 77:3567, 1980)와 같은 칼슘 포스페이트 침전이다.

박테리아 원형질체와 포유류 세포의 폴리에틸렌-유도된 융합(Schaffner 등, (1980) *Proc. Natl. Acad. Sci. USA* 77:2163)은 이중 DNA를 도입하는 또 다른 유용한 방법이다. 전기천공법이 또한 DNA를 숙주 세포의 세포질 내로 직접 도입하기 위해 사용될 수 있는데, 이는 예컨대 Potter 등(Proc. Natl. Acad. Sci. USA 81:7161, 1988) 또는 Shigekawa 등(BioTechniques 6:742, 1988)에 기술된 바와 같다. 이중 DNA를 포유류 세포로 도입하는데 유용한 어떤 시약은 리포펙틴(Lipofectin<sup>®</sup>) 시약 및 리포펙타민(Lipofectamine<sup>®</sup>) 시약(Gibco BRL, Gaithersburg, Md.)을 포함한다. 이들 시약은 모두 지질-핵산 복합체(또는 리포솜)를 형성하는데 사용되는 상업적으로 이용가능한 시약으로, 배양된 세포에 적용시, 핵산을 상기 세포 내로 쉽게 흡수시킨다.

[0161] **이중 특이적 항원 결합 단백질 제작 방법**

[0162] 본 개시는 이중 특이적 항원 결합 단백질의 제작 방법을 또한 제공한다. 본 방법을 사용함으로써, 높은 역가 및/또는 높은 특정 생산성(pg/세포/일)으로 이중 특이적 항원 결합 단백질(예: 이중 특이적 항체)을 생산할 수 있다. 일부 구현예에서, 이중 특이적 항원 결합 단백질은 적어도 5 mg/L, 10 mg/L, 15 mg/L, 20 mg/L, 25 mg/L, 30 mg/L, 35 mg/L, 40 mg/L, 45 mg/L, 또는 50 mg/L 이상의 역가로 생산된다. 일부 구현예에서, 이중 특이적 항원 결합 단백질은 적어도 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40% 또는 50% 이상의 이중 특이적 항원 결합 단백질 역가 대 총 항원 결합 단백질 역가의 비율로 생산된다. 일부 구현예에서, 이중 특이적 항원 결합 단백질은 일당 세포당 생산된 총 항원 결합 단백질(pg 단위)에 기초하여 결정된 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 피코그램/세포/일 이상의 특정 생산성으로 생산된다.

[0163] 일 구현예에서, 상기 방법은 본원에 개시된 시스템을 사용하고, 시스템에서의 벡터를 시스템의 세포 내에 형질감염에 의해 도입한다. 외인성 핵산이 RMCE를 통해 표적 발현 강화된 유전자좌 내에 적절하게 통합된 형질감염된 세포를 스크리닝하고 동정할 수 있다. 2개의 HCF 함유 폴리펩티드 및 적어도 하나의 LCF 함유 폴리펩티드는 통합된 핵산으로부터 발현될 수 있고, 3개의 폴리펩티드를 함유하는 이중 특이적 항원 결합 단백질은 동정된 형질감염 세포로부터 수득될 수 있고, 공지된 방법을 사용해 정제될 수 있다.

[0164] 일부 구현예에서, 방법은 (i) 세포 및 벡터 세트를 포함하는 시스템을 제공하는 단계로서, 상기 세포는 5'에서 3'까지, 그 게놈의 발현 강화 유전자좌 내에 통합된 제1 RRS, 제1 외인성 핵산, 제2 RRS, 제2 외인성 핵산, 및 제3 RRS를 포함하되, 상기 3개의 RRS는 서로 상이하고; 벡터 세트는 5'에서 3'까지, 제1 RRS, 제1 LCF(예: 제1 VL)를 암호화하는 뉴클레오티드 서열을 포함하는 제1 핵산, 및 제2 RRS를 포함하는 제1 벡터, 및 제2 RRS, 제1 HCF(예: 제1 VH)를 암호화하는 뉴클레오티드 서열을 포함하는 제2 핵산, 및 제3 RRS를 포함하는 제2 벡터를 포함하되, 상기 제1 핵산 또는 제2 핵산 중 하나는 제1 HCF(예: 제2 VH)를 암호화하는 뉴클레오티드 서열을 추가로 포함하는, 단계; (ii) 벡터를 세포 내에 동시 도입하는 단계; 및 (iii) 벡터에서의 제1 및 제2 핵산이 제1, 제2 및 제3 RRS에 의해 매개된 재조합을 통해 발현 강화 유전자좌 내에 동시 통합된 형질전환 세포에 대해 스크리닝하는 단계를 포함한다.

[0165] 상기 방법의 특이적 구현예에서, 벡터로부터 유전자좌 내에 적절히 통합된 핵산을 갖는 형질전환체에 대한 스크리닝을 용이하게 하기 위해, 시스템의 세포에서의 제1 외인성 핵산은 제1 선별 표지 유전자를 포함할 수 있고, 세포에서의 제2 외인성 핵산은 제2 선별 표지 유전자를 포함할 수 있으며(제1 및 제2 선별 표지 유전자는 서로 상이함); 벡터 상의 제1 및 제2 핵산은, 추가 선별 표지 유전자를 암호화하는 완전한 서열이 동시 통합 이후에 제공되는 분할 포맷으로 추가 선별 표지를 함께 암호화한다. 제1 및 제2 선별 표지에 대한 선별(음성 선별) 및 추가 선별 표지에 대한 선별(양성 선별)을 위한 형질전환체에 대한 스크리닝이 수행될 수 있다.

[0166] 또 다른 구현예에서, 상기 방법은 세포의 발현 강화 유전자좌 내의 특정 부위에서 통합된 외인성 핵산 서열을 갖는 세포를 단순히 사용하는데, 상기 외인성 핵산 서열은 이중 특이적 항원 결합 단백질을 암호화하고 이중 특이적 항원 결합 단백질을 세포로부터 발현시킨다. 클로닝된 발현 카세트는 특이적 통합 부위 내에서 인접한다.

[0167] 본 명세서는 다음의 실시예들에 의해 추가로 예시되며, 이들 실시예는 어떤 방식으로든 제한하는 것으로 간주되지 않아야 한다. 모든 인용된 참조 문헌(문헌 참조, 발행된 특허 및 본 출원 전반에 걸쳐 인용된 공개된 특허 출원을 포함함)의 내용은 본원에 참조로서 명백하게 포함된다.

[0168] **실시예**

[0169] **실시예 1: 이중 특이적 항체 발현 플라스미드의 클로닝:**

[0170] 이중 특이적 항체의 중쇄 및 경쇄 성분은 당업계에서 공지된 방법을 사용해 하이브리도마 세포, B 세포, 또는 재

조합 항체 유전자 라이브러리부터 클로닝될 수 있다. 예를 들어, 항체는 5' RACE PCR, 또는 리더 펩티드, 프레임워크 1 서열, 프레임워크 4 서열, 또는 불변 영역 서열에 대한 프라이머를 사용하는 PCR에 의해 하이브리도마 또는 B 세포로부터 클로닝될 수 있다. 대안적으로, 항체 발현 세포에서의 항체 유전자 또는 mRNA는 차세대 시퀀싱에 의해 서열화될 수 있고, 이어서 생물정보학(bioinformatics)을 통해 확인될 수 있다. 합성 DNA 기술에 의해 항체 단백질을 서열화하고 상응하는 항체 유전자를 클로닝하는 것이 또한 가능하다. 효모 또는 파지 라이브러리와 같은 재조합 항체 라이브러리 또한 항체 유전자의 공급원이다.

[0171] CHO 발현 세포주 RSX4189-1, RSX4187-1, RSX4191-1, 및 RSX4188-1은 각각 3개의 구별되는 폴리펩티드로 이루어진 이중 특이적 항체를 생산한다: AbC1, AbC2, 및 AbC3 (도 5). RSX4189-1의 작제용 플라스미드를 생성하기 위해, AbC1 유전자에 대해 3'에 있던 Mfe I로 AbC1 플라스미드를 소화시켜 선형화하였다. Mfe I 소화에 의해 AbC2 플라스미드로부터 절제한 AbC2 발현 카세트를 선형화된 AbC1 플라스미드의 Mfe I 부위에 결합시켰다. 결합 생성물을 DH10B 대장균(*E. coli*)으로 형질전환시켰다. 형질전환 및 암피실린 함유 LB 배지에서 성장시킨 후, 개별 대장균 콜로니를 AbC1 및 AbC2 유전자를 함유하는 원하는 플라스미드의 보유 여부에 대해 분석하였다. AbC3 플라스미드 및 AbC1-AbC2 이중 발현 플라스미드에 대한 maxi-prep DNA의 서열은 생어 시퀀싱(Sanger sequencing)에 의해 확인하였다. Cre 발현 플라스미드 pRG858과 함께 이들 2개의 플라스미드를 리포펙타민을 사용해 EESYR<sup>®</sup> 유전자좌에서 RRS1 및 RRS3 부위를 보유하는 EESYR<sup>®</sup> 숙주 세포 내로 형질감염시켰다. 형질감염 세포를 12일 동안 항생제로 선별하였고, 이어서 재조합 세포를 RSX4189-1의 풀(pool)로 모았다.

[0172] RSX4187-1 작제용 플라스미드를 생성하기 위해, Mlu I 및 Nhe I 부위가 측면에 위치한 AbC3 발현 카세트를 AbC2 유전자에 대해 3'에 있는 AbC2 플라스미드 내의 Mlu I 및 Spe I 부위 내에 클로닝하였다. 조합된 AbC2-AbC3 플라스미드, AbC1 플라스미드, 및 Cre 플라스미드 pRG858을 리포펙타민을 사용해 EESYR<sup>®</sup> 숙주 세포 내에 공동 형질감염시켰다. RMCE를 거친 세포를 RSX4187-1의 풀로 모았다.

[0173] RSX4191-1 작제용 플라스미드를 생성하기 위해, AbC3 발현 카세트를 AbC1 유전자에 대해 3'에 있는 AbC1 플라스미드 내의 Mfe I 부위 내에 클로닝하였다. 조합된 AbC1-AbC3 플라스미드, AbC2 플라스미드, 및 Cre 플라스미드 pRG858을 리포펙타민을 사용해 EESYR<sup>®</sup> 숙주 세포 내에 공동 형질감염시켰다. RMCE를 거친 세포를 RSX4191-1의 풀로 모았다.

[0174] RSX4188-1 작제용 플라스미드를 생성하기 위해, Mlu I 및 Nhe I 부위가 측면에 위치한 AbC2 발현 카세트를 AbC3 유전자에 대해 3'에 있는 AbC3 플라스미드 내의 Mlu I 및 Spe I 부위 내에 클로닝하였다. 조합된 AbC3-AbC2 플라스미드, AbC1 플라스미드, 및 Cre 플라스미드 pRG858을 리포펙타민을 사용해 EESYR<sup>®</sup> 숙주 세포 내에 공동 형질감염시켰다. RMCE를 거친 세포를 RSX4188-1의 풀로 모았다.

## [0175] 실시예 2: EESYR<sup>®</sup> 유전자좌로부터 이중 특이적 항체의 발현

[0176] 이중 특이적 항체 발현 세포주 RSX4189-1, RSX4187-1, RSX4191-1, 및 RSX4188-1을 무혈청 배지에서의 현탁 배양하였다. 이들의 이중 특이적 항체 발현 레벨을 정량화하기 위해, 배양물의 세포 수를 Guava 유량 세포 계측기 상에서 계수하였고, 1 ml의 배지당 2백만 세포를 함유하는 새로운 셰이커 플라스크 배양을 개시하였다. 4일 후에, 세포를 제거하기 위해 소모 배지를 원심분리한 후 수확하였다. 이중 특이적 항체 역가는 이중 특이적 항체에 특이적인 단백질 A HPLC 검정을 사용해 확인하였다. RSX4189-1, RSX4187-1, RSX4191-1, 및 RSX4188-1로부터 발현된 이중 특이적 항체 단백질의 역가는 각각 37.8 mg/L, 40.5 mg/L, 48.3 mg/L, 및 21.8 mg/L이었다. 이들 세포주로부터 발현된 모든 항체 단백질(이중 특이적 항체 단백질 및 단일 특이적 항체 단백질을 포함함)의 총 역가, 및 이중 특이적 항체 단백질 역가 대 총 항체 단백질 역가의 비율은 아래 표에 도시되어 있다.

AbPID/RSX	이중 특이적 역가 (mg/L)	총 역가 (mg/L)	% 이중 특이적
RSX4187-1	40.5	211.7	19.10%
RSX4188-1	21.8	218.3	10.00%
RSX4189-1	37.8	199.5	19.00%
RSX4191-1	48.3	172.9	27.90%

[0177]



[0178]	서열 목록	
[0179]	서열번호 1	
[0180]	13515 염기	
[0181]	DNA	
[0182]	<i>Cricetulus griseus</i>	
[0183]	tctagaaaca aaaccaaaaa tattaagtca ggcttggctt caggtgctgg ggtggagtgc	60
[0184]	tgacaaaaat acacaaattc ctggctttct aaggcttttt cggggattca ggtattgggt	120
[0185]	gatggtagaa taaaaatctg aaacataggt gatgtatctg ccatactgca tgggtgtgta	180
[0186]	tgtgtgtgta tgttgtctg tgttgtgtcc cagacagaaa taccatgaag gaaaaaaca	240
[0187]	cttcaaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag	300
[0188]	cacatgccag aggaggtgag gactgaacca ttcaacacaa gtggtgaata gtcctgcaga	360
[0189]	cacagagagg gccagaagca ctcagaactc caggggttca ggagtgggtc tctggaggct	420
[0190]	tctgcccttg gaggttcctg aggaggaggc ttccatattg aaaatgtagt tagtggccgt	480
[0191]	ttccattagt acagtgacta gagagagctg agggaccact ggactgaggc ctatagctc	540
[0192]	agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaaggaaa cagcaggtgt	600
[0193]	gaggggtcag gggcaagtta gtgggagagg tcttccagat gaagtagcag gaacggagac	660
[0194]	gcactggatg gcccacttg tcaaccagca aaagcttggga tcttgttcta agaggccagg	720
[0195]	gacatgacaa gggatgatctc gggtttttaa aggcctttgtg ttacctaalc acttctatta	780
[0196]	gtcagatact ttgtaacaca aatgagtact tggcctgtat tttagaaact tctgggatcc	840
[0197]	tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctcccagcca ggccctggac	900
[0198]	ccgggtccat tcatgcaaat actcaggac agattcttca ctaggtactg atgagctgtc	960
[0199]	ttggatgcaa atgtggcctc ttcattttac tacaagtcac catgagtcag gaggtgctgt	1020
[0200]	ttgcacagtg tgactaagtg atggagtgtt gactgcagcc attcccggcc ccagcttgtg	1080
[0201]	agagagatcc ttttaaatg aaagtaagct caaagttacc acgaagccac acatgtataa	1140
[0202]	actgtgtgaa taatctgtgc acatacacia accatgtgaa taatctgtgt acatgtataa	1200
[0203]	actgtgtgaa taatctgtgt gcagcctttc cttacctact acctccagt gatcaggttt	1260
[0204]	ggactgcctg tgtgctactg gacctgaat gtccccaccg ctgtcccctg tcttttacga	1320
[0205]	ttctgacatt ttttaataat tcagcggctt cccctctgct ctgtgcctag ctataccttg	1380
[0206]	gtactctgca ttttggtttc tgtgacattt ctctgtgact ctgtacatt ctcagatgac	1440
[0207]	atgtgacaca gaagggtgtc cctctggaga catgtgatgt cctgtcatt agtggaatca	1500
[0208]	gatgccccca aactgttgtc cagtgttttg gaaagtgaca cgtgaaggag gatcaggaaa	1560
[0209]	agaggggtgg aaatcaagat gtgtctgagt atctcatgtc cctgagtggt ccaggtgct	1620
[0210]	gacttcactc cccaagtga gggaggccat ggtgagtaca cacacctcac acatactata	1680
[0211]	tccaacacac acacacacac acacacacac acgcacgcac gcacgcacgc acgcacacat	1740
[0212]	gcacacacac gaactacatt tcacaaacca catacgcata ttacacccca aacgtatcac	1800
[0213]	ctatacatat cacacatata caccctcca cacatcacac acataccaca cccacacaca	1860



[0214]	gcacacacat acataggcac acattcacac accacacata tacattttgtg tatgcataca	1920
[0215]	tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtiacga cacatgcata	1980
[0216]	cacacacata ggcacacatt gagcacacac atacatttgt gtacgcacac tacatagaca	2040
[0217]	tatatgcatt tgtatatgca cacatgcatg cacacataca taggcacaca tagagcacac	2100
[0218]	acatacatTT gtgtatgcac acatgcacac accaatcaca tgggaagact caggTtcttc	2160
[0219]	actaaggTtc acatgaactt agcagTtccT ggttatctcg tgaacttgg aagattgctg	2220
[0220]	tggagaagag gaagcgTtgg cttgagccct ggcagcaatt aaccccgccc agaagaagta	2280
[0221]	ggttttaaaa tgagagggtc tcaatgtgga acccgaggg cgccagTtca gagaagagac	2340
[0222]	ctaccaagc caactgagag caaaggcaga gggatgaacc tgggatgtag ttTgaacctc	2400
[0223]	tgtaccagct gggcttcatg ctattttgtt atatctttat taaatattct tttagtttta	2460
[0224]	tgtgcgtgaa taccttgcTt gcataaatgt atgggcactg tatgtgttct tggTgccggt	2520
[0225]	ggaggccagg agagggcagT gatcctccgg agctggcgTt tgagacagTt gtgaccaca	2580
[0226]	gtgtggggTc tgggaactgg gtcttagtgt tccgaagtg cagctggggc tcttaacctc	2640
[0227]	tgagccatcc ctccagcttc aagaaactta tttcttagg acatggggga agggatccag	2700
[0228]	ggctttaggc ttgtttgttc agcaataact ctttctgtgt attttgaatt ttattttatt	2760
[0229]	ttactttttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaat	2820
[0230]	cctctgtct cagtctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca	2880
[0231]	ctttagagtc cccagccctt ctggacactt gttccaagta taatatatat atatatatat	2940
[0232]	atatatatat atatatatat atatatTtgT tgtgtgtgtt tgtgtgtgta tgagacactt	3000
[0233]	gctctaaggT tatcataat atccttgatt tgcTtttaat ttatttttta attaaaaatg	3060
[0234]	attagctaca tgcacctgt atgcgtctgt atcatctata tatccttccT tccttctctc	3120
[0235]	tccttctctc ttcttcttct caccccaag catctatttt caaatccttg tgccgaggag	3180
[0236]	atgccaagag tctcgTtggg ggagatggTg agggggcgat acaggggaag agcaggagga	3240
[0237]	aagggggaca gactggTgtg ggtctttgga gagctcagga gaatagcagc gatcttccct	3300
[0238]	gtccctggTg tcacctctta cagccaacac catTTtgTgg cctggcagaa gagTtgtcaa	3360
[0239]	gctggTgcga ggtctgccac acaaccccaa tctggcccca agaaaaggca cctgtgtgtg	3420
[0240]	actctggggt taaaggcgct gcctggTcgt ctccagctgg acttgaaact cccgtttaat	3480
[0241]	aaagagTtct gcaaaataat acccgagag tcacagtgc aggttcccgT gctttcctga	3540
[0242]	agcgccaggc acgggttccc taggaaatgg ggccttgctt gccaagctcc cacggcttgc	3600
[0243]	cctgcaaacg gcctgaatga tctggcactc tgcgttgcca ctgggatgaa atggaaaaaa	3660
[0244]	gaaaaagaag aagtgtctct ggaagcgggc gcgtcacac aaaccgcaa cgatttgtta	3720
[0245]	aacactctcc attgagaatc tggagtgcgg ttgccctcta ctggggagct gaagacagct	3780
[0246]	agtgggggcg gggggaggac cgtgctagca tccttcacg gtgctcgctg gctgtggtgc	3840
[0247]	atgccgggaa ccgaaacgcg gaactaaagt caagtcttgc tttggtggaa ctgacaatca	3900
[0248]	acgaaatcac ttcgattgtt ttctcttttt tactggaatt cttggatttg atagatgggg	3960
[0249]	gaggatcaga gggggagggg aggggcgggg agacggaggg aggaggggag gaggggagga	4020

[0250]	ggggaggagg ggaggagggg aagggatgga ggaaaatact aacttttcta attcaacatg	4080
[0251]	acaaagattc ggagaaagtg caccgctagt gaccgggagg aggaatgccc tattgggcat	4140
[0252]	tatatccct gtcttcta ggaatcaaac tcttggttcc agcaccaagg attctgagcc	4200
[0253]	tatcctattc aagacagtaa ctacagccca cacggaagag gctatacaac tgaagaaata	4260
[0254]	aaattttcac tttatttcat ttctgtgact gcatgttcac atgtagagag ccacctgtgt	4320
[0255]	ctaggggctg atgtgctggg cagtagagtt ctgagcccg taaactggaac aacccagaac	4380
[0256]	tcccaccaca gtiagagctt gctgagagag ggaggccctt ggtgagattt ctttgtgtat	4440
[0257]	ttatttagag acagggcttc atactgtagt ccaagctagc ctccagctca cagaaattct	4500
[0258]	cctgttccgg tttccaaagt actggagtta tgagtgtgtg ttaattgaac gctaagaatt	4560
[0259]	tgctgattga agaaaacctc aagtgggttt ggctaattcc cacgaccca gaggctgagg	4620
[0260]	caggaggaat gagagaattc aaggtttgcc agagccacag ggtgagctca atgtggagac	4680
[0261]	tgtgaggggtg agctcaatgt ggagactgtg aggggtgagct caatgtggag actgtgaggg	4740
[0262]	tgagctcaat gtggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca	4800
[0263]	atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggggtgtgt	4860
[0264]	tgagtggtag agcagttagt tgatttgaca tgcttgaggt ctcccgttcc atctgtggcc	4920
[0265]	ctgcaacagg aaggaggagg ggaagggggg gaacgagaga gaggaaagag agacagaagc	4980
[0266]	taagataggg aatgagagag gaaggaagaa acgggaagaa attcagactc cttcctgagt	5040
[0267]	tccgccaacg cctagtgaac tcctgtgcac accctaaggt ggcctttgtg tggcactggc	5100
[0268]	ttgggtggtc gggaaaggca ttttcagctt gttgcagaac tgccacagta gcatgctggg	5160
[0269]	tccgtgaaag tttctgccc ttaacaagaa gtcttacta cttgtgacct caccagtga	5220
[0270]	aatttcttta attgtctcct ggtgttcttg gttttgcatt tttgtttcta aggatacatt	5280
[0271]	cctgggtgat gtcatgaagt ccccaaagac acagtggggc tgtgttggat tgggaaagat	5340
[0272]	gatttatctg ggggtgtcaa aggaaaagaa gggaaacagg cacttgggaa aatgtcctcc	5400
[0273]	cgccccccg aattttggct tggcaaccgt ggtggaggag caagaaacac gtggacgttt	5460
[0274]	gaggaggcat ggggtcctag gaggacagga agcagaagga gagagctggg ctgacagcct	5520
[0275]	gcaggcattg cacagtttca gaaggagatt acagcatgac tgagttttta gggatccaac	5580
[0276]	agggacctgg gtagagattc tgtgggctct gaggcaactt gacctcagcc agatggtatt	5640
[0277]	tgaataacct gctcttagag ggaaaacaga catagcaaac agagccacgt ttagtgatga	5700
[0278]	aactctcact ttgcctgagt catgtgcggc catgccagg ggtcaggctg aactcaact	5760
[0279]	caaaaacaag tgagaaattg aagacaatcc gtggtggcag ctactggaag ggccaccaca	5820
[0280]	tccccagaaa gattggagct gctaaaaagc cattttgtat aggcacagtt atcttgaatg	5880
[0281]	catggagcag agattacgga aaatcgaga atgttaatga ggcaacattc gatttgagtc	5940
[0282]	attcagtgtg ggaaccagc acgcttccat ccctaaaag gaacatcttg ctctcagtc	6000
[0283]	aatggaaat aaaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat	6060
[0284]	tttcacacgc acagtggata atttcatgt tggagtttat ttgtgctaaa aggcagaaaa	6120
[0285]	gggtaaaaag cacatcttaa gatttatgag gttctacgaa taaaaataat gttacttaca	6180

[0286]	gctattcctt aattagtagc cccttccacc tgtggtaatt tcctgagata gtcagtgggg	6240
[0287]	aaaagatctc tcttctctt ctttctccc cccccctct ctcctcctt cctcctctc	6300
[0288]	ctcctctctc tccctcctc ccccttctt tcttctttg ctccttctc tctgctctt	6360
[0289]	tctcccttct tcttctatt attctaagta gcttttaaca gcacaccaat tacctgtgta	6420
[0290]	taacgggaaa acacaggctc aagcagctta gagaagattg atctgtgttc actagcgtgc	6480
[0291]	aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttctt atttggcacg	6540
[0292]	gcacttagga agtggaaacat gcctaattca ctggtttga ccaccttct ctataatgga	6600
[0293]	ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag gcctgttaaa	6660
[0294]	tgggtactctt atttgcaaag aaggctgtaa ctgttagctt taaaagcctc tctcaagaa	6720
[0295]	agaagggaga aaggatatgg ctagacata ctaatagact taaccactgt gaaaagcctt	6780
[0296]	agtatgaatc agatagaacc tatttttaac tcagtttga aaaaaataat ctttatattt	6840
[0297]	atttgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gaaccacatg	6900
[0298]	tagcagggtgc tggaggaggc cagaagaggg caccagatct cctggaactg acaccacaca	6960
[0299]	tggttatgag ctgcctgatg tgggtgctgg gaactgaact ctcgtgttct gcaagagcag	7020
[0300]	caactgttct cttaactgat gagccatctc tccagccccc cccataattt taattgttca	7080
[0301]	tttttagtaa tttattcat aatcaattat cacagtataa aacaatgatt ttatatatat	7140
[0302]	catatacata tcaaggatga cagtgagggg gatatgtgtg tgtgtgtgtg tgtgtgtgtg	7200
[0303]	tgtgtgtgtg tgtgttatt gtgtgtgtgc tttttaagaa ggtgccatag tcaactgcatt	7260
[0304]	tctctgaagg atttcaaagg aatgagacat gtctgtctgc caggaaacct atcttctct	7320
[0305]	ttgggaatct gacccaaatg aggtattctg aggaactgaa tgaagagctc aagtagcagt	7380
[0306]	gtcttaaac caaatgtgt gtctagagaa agtcaacgtc atcagtgagc tgaggagaga	7440
[0307]	tttactgagc ggaagacaag cgctctttga ttttaagtgc tcgaacagtc acggctgtgg	7500
[0308]	agtggagcct gtgtcaggt ctgaggcagt ctttgctagc cagctgtgat gagcagtga	7560
[0309]	gaaagggtgg agatggaggc aggggtgggag cagggtctatg gttcagacta ggtatcgtga	7620
[0310]	gcacaccagc tggttgactt gtggctctgt ggtcaggcgt tgtaaacgcc ctcagggtca	7680
[0311]	ggcagtcaca ttgcttgaag ctgaatgggt gaggcaacac agagagtga aagaaggcaa	7740
[0312]	agtaccacct ctccccgac ccaggctact tctgggttat agctgagact ccggacagca	7800
[0313]	tgcaaccagc tggtttagagc ttcagggaaa acttgatgtc tgcatgttgc tatgaaatgt	7860
[0314]	gattcggtag atctggagaa aatttataat gctggctcag tcaagcactg aacaaaggta	7920
[0315]	ccttggcttt gggagctaca tgacattgac ttgtaggcag acttttttt tctgcccgc	7980
[0316]	caattcccag ataaccaata tggaggctca atattaatta taaatgctcg gctgatagct	8040
[0317]	caggcttgtt actagctaac tcttccaaact taaatgaacc catttctatt atctacattc	8100
[0318]	tgccacgtga cttaccttg tacttctgt ttcctctct tgtctgactc tgcccttctg	8160
[0319]	cttcccagag tcttagtct ggttctctg cctaacctta tctgcccag ctgctgacca	8220
[0320]	agcatttata attaatatta agtctcccag tgagactctc atccaggag gacttgggtg	8280
[0321]	ctccccctc ctcatgcca tccgtgtctt cctcttccct cgttcccc tctcttctt	8340

[0322]	gctcttcctc ctccaccct cctttcatag tattgatggc aagggtgttc tagaatggag	8400
[0323]	gagtgcccat aggcattgcaa agaaaccagt taggatgtct tgtgaggggt tgtaatcata	8460
[0324]	agcgatggac acaattcaag ccacagagtg aagacggaag gatgcactgt gctctagagc	8520
[0325]	aacttctggg gcagaatcac aggggtgagtt tctgacttga gggcgaagag gccacgagga	8580
[0326]	agggagttag tttgtctgag ctagaagcta cgccccacct cttggtagca gacctgcccc	8640
[0327]	caagcatgct ttgttaatca tgtgggatct gattttcctc taaatctatg ttcaactctt	8700
[0328]	aagaaaatgt gaatttcac attaaaattt agatatacgt cttttggtgg ggggggtgta	8760
[0329]	aaaaatcctc aagaataagg atttctgggg gccggagaga tggctcagag gttaagagaa	8820
[0330]	ctggttgctc ttctagacat tctgagttca attcccagca accacatggt ggctcacaac	8880
[0331]	catctgtaat gcgacctggt gccatcttct gacatgcatg gatacatgca ggcagaaagc	8940
[0332]	tgtatacata gtaaatgtat aaatcttttt ttaaaaagag tatggattct gccgggtgtt	9000
[0333]	ggtggcgcac gcctttaatc ccagcactct ggaggcagag gcaggtggat ctctgtgagt	9060
[0334]	tcgagaccag ccigtctat aagagctagt tccaggacag cctccaaagc cacagagaaa	9120
[0335]	ccctgtctcg aaaaaccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaga gtatggattc	9180
[0336]	taagaaagcc gtaacagctg gagctgtgta cggagttcag cgtggtaact gaagaacaga	9240
[0337]	cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt gttgttttga	9300
[0338]	gaccggctct tatgtctctc aggtggcct caaactgtg atcttcccgc ctctacctct	9360
[0339]	caagtcctgg gactacttgg ctcataaaac agttttgtc gggtccctg aagttatggt	9420
[0340]	tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggct tgaatcccag	9480
[0341]	acaatgactg catctcagga cagttgggaa gaggacaatg gcagaaggac ttagaaaaga	9540
[0342]	tagactggag ggtggaaaag cagcaggaac agagaaacaa aacaggaagc ttgctatcca	9600
[0343]	gggccactct ggagtcctgt ggcaagatgg aagcgggcta ggggaataca tttgtgctac	9660
[0344]	tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat caatgttgaa	9720
[0345]	ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg cctagagaaa	9780
[0346]	ggaatggccc ctgcagaata gacagagtga atgggtccct ttatcatttg ctaaagtga	9840
[0347]	ggagaaataa acatccttcc atagagtttc aggtaaatga accccacagt tcatctgtgc	9900
[0348]	cgtgggtggag gcctggccaa cagttaaaaa gattagacac ggacaaagtc tgaaggaaac	9960
[0349]	acctcgaata ggaagaggag agccacctca ttctgtaact ttctcaagg ggaagatgtt	10020
[0350]	ccaagagtgg gaataaatgg tcaaaggggg gatttttaat taggaaaacg atttctgta	10080
[0351]	tcacttgtga aactggaggt tgatttgggg cataggacaa tagatttgat gctttgcaaa	10140
[0352]	aagctgtttc aaagcagaga aatggaatag agacaattat gtagcgagga gggagggtgg	10200
[0353]	ggcgaagatg gagacagaga agtggaagct gactttaggg aagaggaaca tagaccacag	10260
[0354]	ggggggggcg gggggcaggg gcggggggcg gggctcaaag gaggcagtgg gaacgttgct	10320
[0355]	agtgttcgca gcgtaagcgt gaatgtgcaa gcgtctttgt ggtgtgtgac caggagtagc	10380
[0356]	gtggctggct tgtgtgtgc ttgtaatccc agtctttgag gtttccacac tgttccacag	10440
[0357]	tgggtgtgat ttccctcgg agagcatgag ggctctgctt tccccacatc ctccccagcg	10500

[0358]	ttcgttggta ttgttttcca agatgttagt gggtagagaca aagcctctct gttagattgc	10560
[0359]	ctttaacagg tgacaaaaaa agctcaacca ggagacattt ttgccttctt ggaaggtaat	10620
[0360]	gctcccatgt agagcaatgg gacctatctc taaggtgagg ctactcttgc agtttgcacc	10680
[0361]	cagctcttct gatgcaggaa ggaagttagt gggcaagcaa gactgtttgc ttccttgcgat	10740
[0362]	ggacacattc tgcacacaaa ggctcaggag gggagaaggc tgtttgatgt ttagcactca	10800
[0363]	ggaagggccc tgatgcatct gtgattagct gtctccatct gtggagcaga cacggactaa	10860
[0364]	ctaaaaacca gtgtttttaa attgtcaagc ctttaagggtg aggaaattga cttattgtgc	10920
[0365]	tgggccatac gtagagcaag tgctctgcat tgggccaacc cccggctctg gtttctaggc	10980
[0366]	accagaatgg cctagaacta actcacaatc ctcccatcc aggtctcagg tgctagaatg	11040
[0367]	aaccactata ccagcctgcc tgccctgccia cctgccttcc taaattttaa atcatgggga	11100
[0368]	gtaggggaga atacacttat cttagttagg gtttctattg ctgtgaagag acaccatgag	11160
[0369]	catggcaact cttataaagg aaaacattta gttgggtggc agtttcagag gttttagtag	11220
[0370]	attgtcatca tggctgggaa catgatggca tgcagacaga catggtgctg gagaaaggga	11280
[0371]	tgagagtcct acatcttgca ggcaacagga cctcagctga gacactggct ggtaccctga	11340
[0372]	gcataggaaa cctcacagcc caccctcaca gtgacatatt tccttcaaca aagccatacc	11400
[0373]	tcctaatagt gccactccct atgagatgac agggccaatt acattcaaac tgctataaca	11460
[0374]	ctttaaagta ttttattttt attattgtaa attatgtatg tagctgggtg gtggcagccg	11520
[0375]	aggtgcacgc ctttaatccc agcacttggg aggcagaggc agatggatct ctgtgagttc	11580
[0376]	aagaccagcc tggctataaa gagctagtig caaggaagga tatacaaaga acagttctag	11640
[0377]	gatagccttc aaagccacag agaagtgtg tcttgaaaac caaaaattgt gctgggacct	11700
[0378]	gtctctgctt tggttgcttc ccactccccc agagctggac tcttgggtcaa cactgaatca	11760
[0379]	gctgcaaaat aaactcctgg attcctctct tgtaacagga gcccgaaagtc aggcgcccac	11820
[0380]	ttgtcttctc gcaggattgc catagacttt tctgtgtgac ccaccattcc agactgaagt	11880
[0381]	agagatggca gtggcagaga ctgggaaggc tgcaacgaaa acaggaagtt attgcaccct	11940
[0382]	gggaatagtc tggaaatgaa gcttcaaac ttgcttcatg ttcagttgta cacagactca	12000
[0383]	ctcccagggt gactcacacg tgtaaatatt cctgactatg tctgcaactgc ttttatctga	12060
[0384]	tgcttccttc ccaaaatgcc aagtgtacaa ggtgagggaa tcacccttgg attcagagcc	12120
[0385]	cagggtcgtc ctcttaacc tggacttgtc tttctccggc agcctctgac acccctcccc	12180
[0386]	ccattttctc tatcagaagg tctgagcaga gttggggcac gctcatgtcc tgatacactc	12240
[0387]	cttgtcttcc tgaagatcta acttctgacc cagaaagatg gctaagggtg tgaagtgttt	12300
[0388]	gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggatg tggcagcatg	12360
[0389]	taccggaaat ccagaaactg gggaggtaga gacggatgag tgcccggggc tagctggctg	12420
[0390]	ctcagccagc ctagctgaat tgccaaattc caactcctat tgaaaaacct ttaccaaaca	12480
[0391]	aacaaacaaa caaataataa caacaacaac aacaacaaac taccctatc aaggtgggcg	12540
[0392]	gtctttggct cttgaggaat gactcaccca aacccaaagc ttgccacagc tgttctctgg	12600
[0393]	cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac ttcctgggct	12660

[0394]	gggctgtgtg ctctaggcca ccaggaactt tctgtcttg ctctctgtct ggcacagcca	12720
[0395]	gagcaccagc acccagcagg tgcacacacc tcctccgtg cttcttgagc aaacacaggt	12780
[0396]	gccttggtct gtctattgaa cgggagtaag ttcttgacaga tgtatgcatg gaaacaacat	12840
[0397]	tgtcctgggtt ttattttctac tgttgtgata aaaaccgggg aactccagga agcagctgag	12900
[0398]	gcagaggcaa atgcaaggaa tgctgcctcc tagcttgctc cccatggctt gccgggcctg	12960
[0399]	ctttctgcaa gccctttctc ccccatggc atgctgaca tgaacagcgt ttgaaatgct	13020
[0400]	ctcaaagtgc actttcaaag aaggcttctc tgatcttgct aactaaatca gaccatgttt	13080
[0401]	caccgtgcat tatctttctg ctgtctgtct gtctgtctgt ctgtctatct gtctatcatc	13140
[0402]	tatcaatcat ctatctatct atcttctatt tatctaccta tcattcaatc atctatcttc	13200
[0403]	taactagtta tcatttatct atttgtttac ttactttttt tatttgagac agtatttctc	13260
[0404]	tgagtgcagc ctttggtgtg cctggaacct attctgtaac caggctgtcc tcaaactcac	13320
[0405]	agagatccaa ctgcctctgc ctctctgggtg ctgggggttaa agacgtgcac caccaacgcc	13380
[0406]	ccgctctatc atctatttat gtacttatta ttcagtcatt atctatcctc taactatcca	13440
[0407]	tcatctgtct atccatcatc tatctatcta tctatctatc tatctatcta tctatcatcc	13500
[0408]	atctataatc aattg	13515
[0409]	서열번호 2	
[0410]	14931 염기	
[0411]	DNA	
[0412]	Cricetulus griseus	
[0413]	기타_특징 (2176)..(2239)	
[0414]	n은 a, c, g, t이거나 뉴클레오티드가 없음	
[0415]	catgtacact tatgcaagta tgatatggcc caacacagta ttttacacca atttttatct	60
[0416]	ataaaatata catgtacatc aaaatatatt attaataata acatcattat tctttctttc	120
[0417]	caagtaataa acacatacac tgaaattttg gttcttgtgg ataattttaa tgaaacagga	180
[0418]	aatgcaaatt tatcttagca tgtttacttc actttctttg catagataac cagtaatcac	240
[0419]	attgatggat catgtagtga aatgtatttt taggtatcta aggaattttg gcttcgtttt	300
[0420]	gtgcttggtg aactgaatt ctattcctaa caacagtgtg taaggattct gtctgatttc	360
[0421]	ttttaccagt atttgtccat ttgcattttc ttattatttc atggctgctg ttctagaaag	420
[0422]	tggaaggtag tgtgtcaagt ctgtttaaca tgtttccctg atgatcagtg tettaacacc	480
[0423]	tctctgagta catgttggcc aatgtcgttt ctagacccat ctattcttgc ttgacttate	540
[0424]	ctggtacatg cctgccaaga aatttctcct caccctttct gtctcttcac tgatttactt	600
[0425]	gatgtgtgga ttacacattg atcatatgga aatagaagat acaattttct ttattcacag	660
[0426]	tttggaagac tttaactctc atagatcatc attatttttt gctactgttc cctatgctat	720
[0427]	ggtgaaattt ccatttgaat aattgcttaa acaattaaca agaaagaatc tatttttact	780
[0428]	tgcaataact tccatttcag aacatttact acactgttac tataatccaa aactagtttt	840
[0429]	atataatcatg tgagaaatga ctaattcata atttggccat gacatttttt tcagaaacag	900

[0430]	aaaaagtgc caatacatc acaatgctat aaatattaag acttcagcaa attaaatatt	960
[0431]	tattcatgat atcacataaa attcatttat tatgttttat ttaaatgtgt ttttaaaca	1020
[0432]	gtggatcac taaatattaa gttagatgtg tttatgtgct taatgaattt atattttaga	1080
[0433]	atgttataag ttgtatatag tcaaataatgt aataaatttt attttttagg tctttctcat	1140
[0434]	taaggatatt taattttggg tcccttttcc agagtgactc tagctcatga tgagttgaca	1200
[0435]	taaaaactaa acagtacaaa atgtacattg cattcagtat tgcacttgat ctttgactg	1260
[0436]	aagtttgagt cagttcatc atttagtact tgggaagtac attaagctaa ctttcattgc	1320
[0437]	tctggcaaaa tgctcgataa gataagagtc tattgtggaa agccatggca gcaggaaagt	1380
[0438]	aagactgctg atgatgttta atccatagtc aagacgcaga aggagatgaa tgctggtatc	1440
[0439]	caacattttt tgctgttcat tttctctaga accctagtcc ataaagatgt atgacttgca	1500
[0440]	ttcaaatgc gtcccttca gttgttcaac ttttctgtaa atatccttc aggcattgtc	1560
[0441]	agaagattgt ttcgcaataa cttctcaatc cattcaagtt gatagtgcag attaatcact	1620
[0442]	gcagaataaa agcctgtaac ttggctcacg tgccaaggaa tatgcacact cctgacacat	1680
[0443]	caataagtaa atcaaatgt agcttttgcc tttacattg ccagacttat gtaatgttct	1740
[0444]	gcacgttctt cctccatcac tttttattct aatggtgttt ccttgacatt gaatcacgt	1800
[0445]	gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta	1860
[0446]	gatttactat tttacttaga attttttata attgagagaa tataatattt tcacagttaa	1920
[0447]	ctatctgctg taaatagagg attttaaaaa aaatctctat aactttttt tacaacacac	1980
[0448]	agtaaaatta agtataaatt taataaagtc actatgttga tttcaaagtg tgctacgcc	2040
[0449]	acggtgtgca cgcaggtgta gcagaagatg ccactaaggt gggctaaggc c gatgggtg	2100
[0450]	gggtctgcgc tccctggaga tgagccccag gcggttcctt ggcaatcagc tgcatcatg	2160
[0451]	atgccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	2220
[0452]	nnnnnnnnnn nnnnnnnnnc tgggtgactt tatggaaaga atttgataga tttcatgatg	2280
[0453]	tagaagaatt ttattaggct tattttacag gagactaaga ccttgggacc taaagatata	2340
[0454]	tgggtcctga gaatcaggaa atgggtagag acgtggttga tggatgaga cagattttag	2400
[0455]	agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa	2460
[0456]	caattatgct gtctttttc tttctgttgt atgatctgat gatgtagccc ccttgccaag	2520
[0457]	ttccctgac ccccttgcca agttccctga ttgtaacagt atataagcat tgcttgagag	2580
[0458]	catattcaac tacattgagt gtgtctgtct gtcatttcct cgccgattcc tgatttctcc	2640
[0459]	ttgagccttt tcccttgctc tccctcggtc ggtgggtctcc acgagaggcg gtccgtggca	2700
[0460]	aaagtgtata aatgttctaa aacatttgaa ctctaaaaca tgcaaatga aaaattaaaa	2760
[0461]	taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaacaata ctatataata	2820
[0462]	ttttgttatt agaattcaaa atcacattag ttggatttaa tttgaacatt gcattctttc	2880
[0463]	aataataatt tcaataaaaa agtttcccc atgatatag aaaataataa catatgtatc	2940
[0464]	tatctattta tttaactaca catatatagc atttgtttca actaaaataa atgaatgagc	3000
[0465]	aaagcaccta agtaattggt gtctattata tttatgaagc caatagtttc aaataaatta	3060

[0466]	tcatgcataa ggaggtattg caaatgttaa accttttttg aaacagatat tcccagttac	3120
[0467]	agaaattata atttctaate ttctctataa gtagaatgat gataattaat ataggccatt	3180
[0468]	tgtaaataat gttcagatta aaatattctc tatttcacta gagaagaatg atattaaatg	3240
[0469]	tattatattt tatttcccat ttgttttgca ccactattct atatccctca gcagttttaa	3300
[0470]	tttgtttcac catatgtgtg tgtgtttgta tcttaaatat ggccactaaa ttagaataat	3360
[0471]	ttaatataaa tctttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct	3420
[0472]	tttaattgtc caagaatact ttttcttcta ttttaggact gatcagaccc aggactaata	3480
[0473]	ttttatatgt actaattcta tgtacaaaaa tatgttatta tctcatgaat tctgtctcaa	3540
[0474]	tattgaggta ataaaaatag tccatcatga actttaaaat taaaataatg attaattaat	3600
[0475]	ttttattcat atttgtttg tatgaatggg tatacatcac atgtgtgcct ggtgactgtg	3660
[0476]	aatgtcagga gaaggtatga aagccactgg aattggaata agagataata tttagatgt	3720
[0477]	tatgtgggtg ctgagaatta gacgcaagcc atcttcaaga atagccagca tactatacca	3780
[0478]	ctgagtaate cattcatccc tcaataatta tctttgtaga cagtaaatat atttctaaac	3840
[0479]	tataaatgac cagaaaaatt aatgtattat taatgaagac attcatctca tgtgacacac	3900
[0480]	ttcacctgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca	3960
[0481]	cttactattt ctaaagctgt ccaatggggg ccagtcacca gtcagcacc agtgagataa	4020
[0482]	tccatgaatg catttatate ttaggaaaaa ttcttatcta tgtagtattt agaacatttt	4080
[0483]	catgtgaggg gataaacaag gaagcacaga tgctttctga tagaaacttt ctctttaatt	4140
[0484]	catctagaaa aaaaaaacct ctgaggaaaa tctctcttgc tctcctcca atgctctatt	4200
[0485]	cagcatcttc tcctacttta attctagatc tttttctcta tgcctccttg ctgctgcct	4260
[0486]	gctggtctg ctctatgcct ccccatgtca cttttctttg ctatctcacc gttaccttct	4320
[0487]	ctgcctcact ctctgccttc ttctctgctt ctacatggc caggctctgg acaattatag	4380
[0488]	ttatatgtta cattctcata acacatgata tgtcacatag ttctctcag gctagggata	4440
[0489]	tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagttg gtgatggttc	4500
[0490]	ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtacatg aagactgaat	4560
[0491]	tttcttcaaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgtat	4620
[0492]	ttgtcttttag aggcaaccac atattgtctg cagggcctgc ttggaatttg ctctgaagtt	4680
[0493]	agcttgtttg tgtaaaagga agaatcctat atcagcctga gaaatgtaaa atatcctagc	4740
[0494]	atttcaagtc atcaaaaata tatggagagt ataatcatc ctctcgacta ttcatagtca	4800
[0495]	tatttgtgtc caccaaglat aaaacacact accaaagggc tgtggaaaaa atcgccataa	4860
[0496]	ctgttcttat tagggaggca tagcagtggg acctgaggaa gttacagcaa caaccagtca	4920
[0497]	tccagtcaat aaccccatgg ctttgccact tggaggtacc caataatgtt tggctttgcc	4980
[0498]	gagtaggact ccaacaaatt cagagggtca atttttaaat gctggttgtc actgctgaac	5040
[0499]	agtccatttg cctctgcat aattccacaa tggaaaagctt tttaactga ttgccaatca	5100
[0500]	ttaaacagcc tactcagcat aaacaggtat gatattatc tgcattttgt tacattacta	5160
[0501]	gatgaattcc tatttcttcc tacaatagtg gaactgaaaa aagatacaca atcatactac	5220



[0502]	ccctctacta atcttatgac ttatatcatt tcaattttca gaccataatg caaactattg	5280
[0503]	accaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa	5340
[0504]	ggcggactta ttttgtttta tttcttagca tgcatagcaa tacatgattt gaggtttata	5400
[0505]	taataaaggg acaataaatc ttcaagaaac ttacccttac tgaattaaaa tattaagaa	5460
[0506]	ggtcacacat ttactcaaat atattagact actgggcaaa tagacatgaa aagtagagtt	5520
[0507]	aatattgagg taggccttct gtgaaatgtc taaggaaatt atgtttcata cagtgtgtaa	5580
[0508]	ccaagtggga atcatalcag aaagcagtca aaagcttata ttacaagtaa cagatgcttg	5640
[0509]	gttatatgac ctcccagagc ttgactgtct atacacaaaa agtgggtgta ataaaactgt	5700
[0510]	aatttgggct atgttttttt aaatggcttc accaacaatga aaggaaggga atgagcatgt	5760
[0511]	catggatgct tagagattat gcttcagca agaagaattg agctttggct cttattacag	5820
[0512]	aaacatgaca aggtgtgagt tttatttatt agaaattata taatatttta agctggggac	5880
[0513]	taaaaatfff attgaaacaa acaggcaagg gataggcatg tactagaagc aaaaatagga	5940
[0514]	tgtcaatgct gtaatgttat ttttggacc aaaatagtat ttcctataga aatgacaatg	6000
[0515]	atcttaggtt attattcttc ataaagatga caagttcaca agatatecta gttcattaaa	6060
[0516]	atcgttttag tcatttaata gagtgtgtg atagattaca caaaggaaag cacttacgat	6120
[0517]	gagaaataat gatatccaca attattttct taattcttag aaacattcta ttgttatatc	6180
[0518]	tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata	6240
[0519]	ttgtctatgg tgacatttca aagaacatgt gacgtacagt gtagcacaga taaagaacat	6300
[0520]	aactgcagct gaatcagtaa ctaaacttac atacattaaa tctgccatgt tggcaacagt	6360
[0521]	gtgtgacta ccaaaggatg tactaatgct cagcacactc ccctatgtca cctttgttc	6420
[0522]	atcattacat cataggctta ttttgtttgc ttttgaaac tagaccaagt cttttgtgtc	6480
[0523]	ttccaagca cagagctcat taatttacct catagacttg ttaaacttct tctggttcat	6540
[0524]	caattgaata gaaatactca ctactaatta tgtgagaccc tgccagtacc atagcacatg	6600
[0525]	gataatffff acataaaaca tgcatacaag taagattatt cagactgaac atgaatttta	6660
[0526]	gagaaatcag gaaggagtat atgggagtgg ttggagtgag actagagaaa tgtaattaaa	6720
[0527]	ctataatctc aatacaaaaga tctactaagc aaaaaacatg aaacattgtc attcaagtga	6780
[0528]	aacatcagtc ttcaaatgg aaagatattt ttactaggaa aatgtctggt agatggttat	6840
[0529]	tatctagaaa acacaaaaat tagaaaacgg taaactttta taaaaagaat aatacaatga	6900
[0530]	gactacatga aaagtcttta actaatgaaa caaatatctt gaaacttttt tcttaaaagt	6960
[0531]	ttaatatcaa taaccatcat ggaaattcaa attaaaacta tttacatatt acccctgaaa	7020
[0532]	taataactaa tacccaataa aaataatata aacaaaaaat ggcaatgcat gccatcatgg	7080
[0533]	atttgggaga gagaatgttc attgcagttc tgaatggata ctgggtgccac cacggtgaaa	7140
[0534]	atctctgtat aggtccttcc aaaagctgaa aatagacata tcacaagacc tgccacacat	7200
[0535]	ttttcaagca aatacccaaa ggactctacc tgactgcaga gacactttct cataaaatat	7260
[0536]	tattgttgat ctattcataa tatctggaaa atagaaacag ccaagatgcc catcaactga	7320
[0537]	ttaatagatg ataaaattat tgtacatttc agtghtaatat tattcagttt ttaagaaaa	7380

[0538]	tgaattatg taataagcat gtaaatggat atatcttgaa acaaccattc cccattatat	7440
[0539]	tacctaaca ttgaaagtc aaaatcatat gatcttttta gtggatctac taatcttttg	7500
[0540]	ctatatgtat ttatttgaac tacccatgga tgtgagataa ttggttaaca cagcacatgg	7560
[0541]	gagagcatgg gatcattcaa ggaagattag agagaatgca ttttttagga gataatggag	7620
[0542]	gagcaataga aaggattaaa tgaggttact gatgaaagt atggtttagag aaggcaatat	7680
[0543]	gaggagggat aactagcact tagggccttt tgaaaaagac atagagaaaa tactattgta	7740
[0544]	gaaacttcct ataattgggtg tatagttata tacaccaaag agctcagatg gagttaccct	7800
[0545]	ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat	7860
[0546]	agattgggaa gcatttactt tggcttacag ttctaacggg ataaaaattc atgatgaaag	7920
[0547]	aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc	7980
[0548]	ttgaagtgtg agaattgtac agagagaaca aactgcaaat gaccagaaaa tgcttttgga	8040
[0549]	tcagagccca taccctctcg actgacttct ccagaaattc tgaacaaata aaactcccca	8100
[0550]	aacagagcca taactgaagg tccagtgtct gagactacta ggggtatttc ttattcaaac	8160
[0551]	cactacaatg ggggtggggg agcaatctc caagtaggca ctacacacag acaataaaaa	8220
[0552]	actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtgag ctacatagag	8280
[0553]	cacaattatt gtatttaaat taccctttat gatcttaca aacttgacag taagatcata	8340
[0554]	ttgctaaaga aaccacatat ttgaatcagg gaacatggtg atatctagtt gttcttcaac	8400
[0555]	tggaaacttc atgctttctg cccagcattc atgttgctgg aaagagcaat gtacactacc	8460
[0556]	agtgtagaaa ttaaatcatc aatcttatca agatgtggat cctataagtt acaataaaaa	8520
[0557]	ttagcctgat aagatatccc caccagaaga atattcacat aaatgctatg ggagcaacaa	8580
[0558]	gctattttct aaattagctt taatcctatt ctacaagaga gaatccatat ctagaatagt	8640
[0559]	tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tcctaattat	8700
[0560]	attgttctac aaaaataaaa taactcctaa tgacttggtg ctgcagtaat aagtttagtat	8760
[0561]	gttgctcaac ttcacaaaga gaagttttgt cttacaataa atggcaatta aagcagcccc	8820
[0562]	acaagattta tatcataccg atctcctcat ggcctatgca tctagaagct aggaacaaaa	8880
[0563]	gaggacccta agagagacat acatggtecc cctggagaag ggggaagggg caagacctcc	8940
[0564]	aaagctaatt gggagcatgg gggaggggag agggagttag aagaaagaga aggggataaa	9000
[0565]	aggagggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag	9060
[0566]	caagaaaaga gataccatag tagagggagc ctgtgtatgt taaatagaaa actggcacta	9120
[0567]	gggaattgtc caaagatcca caaggtccaa ctaataatct aagcaatagt cgagaggcta	9180
[0568]	ccttaaaagc ctttctctga taatgagatt gatgactacc ttatatacca tcctagagcc	9240
[0569]	ttcatccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagttg	9300
[0570]	cagacaggga ggagtgatga gcaaagtcaa gaccaggctg gagaacaca cagaacagc	9360
[0571]	agacctgaaa aaaaatttgc acatggaccc cagactgata gctgggagtc cagcatagga	9420
[0572]	cttttctaga aaccctgaat gaggatatca gtttggaggt ctggttaatc tatggggaca	9480
[0573]	ctggtagtgg atcaatatct atccctagtt catgactgga atttgggtac ccattccaca	9540

[0574]	tggaggaatt ctctgtcagc ctagacacat gggggagggt ctaggtcctg ctccaaataa	9600
[0575]	tgtgttagac ttigaagaac tcccttgaga agactcacc tccctgggga gcagaaaggg	9660
[0576]	gatgggatga gggtttgtga gggacaggag aggaggggag ggtgaggga ctgggattga	9720
[0577]	caagtaaag atgcttgttt ctaattttaa tgaataaagg aaaagtaaaa gaagaaaaga	9780
[0578]	aaacaggcca aaagattata aaagacagag gtggtgggtg actataaaga aacactatta	9840
[0579]	tctaaataaa aatatgtcag aagcacacat gaacttatag tgtttatgaa agtatgtata	9900
[0580]	ataactacat aatctcaagc caagaaaaa atatcatctt tcagtgtatga aggtgatatt	9960
[0581]	atttctccca gaattaaagc caaagaccta atgaaagtaa ttatcttcaa aagggtgaaa	10020
[0582]	atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg	10080
[0583]	ctcaattgaa ttccattcaa tgttacagtt tagataaaca gttttagat aaactcaca	10140
[0584]	tgtatcatct ctttttattt ttgaccaa cagcttctca tctgttattc agaataattc	10200
[0585]	ctcgatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac	10260
[0586]	cacatacata ttigccattg ggaaacaaag tctaaatga tgttggtcac atcttctcta	10320
[0587]	ctagtctct ccccgctcca aagaaccttg gtatatgtgc ctcatcttac agagagagga	10380
[0588]	aagcaggaac tgagcatccc ttacttgcca tctcaaccc aaaatttgca tcattgctca	10440
[0589]	gctctgccct tctcataatga cagttacaag tcaaggcttc caaagtcct ctgtcatgtt	10500
[0590]	tgggtgtcaat agtttatata gatgacttca tgtcttcata tctaatgtct tatatagatt	10560
[0591]	aatattaaac aatgttattt ctctaaccac attttaaat aatttaaaaa tccattaatt	10620
[0592]	gltgtctataa aatgcagaca gagtgtgag acacaatata agcctgatga tctgaatttg	10680
[0593]	aaactcacac ccaccacatg gagaatcaac ttccaaaaat ttctctatta ctccacact	10740
[0594]	tacaccattg tacaacaca ataataatga acaaatgaa atgaaataaa aaattaagtc	10800
[0595]	tctgtaggta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgctttatgg	10860
[0596]	ttacacttta ccatcttcca ttaattataa ggacttcaat catggcagaa ctatgctgtt	10920
[0597]	attgtctcag tgtaacctaa ccagggtgtc cagatgttct taatgtggac acctaacta	10980
[0598]	tttgatattt gggttaagat ctttccctct ttcagaagaa acctcaggac agagggaatc	11040
[0599]	ttgtctttta attttgatg tgtagacttt ttccatttca aatafacatg aaacaagtga	11100
[0600]	tgaagaaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca	11160
[0601]	atattatcat ggaatcgct gttatacact gagtgtttgg caataaggga tttttagaag	11220
[0602]	aaggagtttt tattctcaac aggttcctta agtttagctc aaataaatct aagcaatcca	11280
[0603]	ctctagaatt aaatagtct ctaagggcac agctatgaat agagctcaat ttacatataa	11340
[0604]	aattttgttc accatttatg tcattccagt ttctattagt acaaggaaaa tacaaaaat	11400
[0605]	ttagatgtca atatcaagt aatagttcat ctctttttt aatatatac acctaaatca	11460
[0606]	ccattttctc agaaaaatct ggcctgaagt tctgtctgga acttcaacat gaaaaaatg	11520
[0607]	cacagcttgc tattataaat cctagttagt tttaagatt catgtctggt gtctgactca	11580
[0608]	gaggggccag aggctagaca aatatTTTTT gaatcttcat tgtgaagatt tttaagtatt	11640
[0609]	attttaatat aaataacaaa gatgatggat aatgtaactt tgtacagttc atagacgctg	11700

[0610]	aactactttg tgcttaaaat gttagtccccc tatcataaat gataggatgat aagtgtatgt	11760
[0611]	ttaatacttt cctctgagc tatattcatg tactagagaa ttatttttaa catgaaaaga	11820
[0612]	ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcaggtg aatgccagga	11880
[0613]	gcaacgtttt tcttctacag aggatgcttt gctgccaagc aacctggttg tgtggaaatg	11940
[0614]	ttcctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tattttcagg	12000
[0615]	ttagagcttg gtccttggag tattatcttt taccagaaaa ttcatagtat tctttcaata	12060
[0616]	actaacaact aaacttttcg ataaaaaaga attggaattt caatttttaa gcctgagtaa	12120
[0617]	aattcttgtg aatcaggata ttttatttta agtcttatct tttaaaaagt tattttattt	12180
[0618]	tttaaaaaat tataatafac ttccataatt tccctccttc acttttcttt acaaacactt	12240
[0619]	ctatagatca ccatgtgttt ttttttttac atttatggcc tctttctgtt cattgttatt	12300
[0620]	acatacaaat agtcttgctt atagaagaac accacaattt gttacctgat aacaaattat	12360
[0621]	caaccttaaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa	12420
[0622]	gatatatgtg tgtgcacata tatagataca catatatgta ggatttttaa ttttagattt	12480
[0623]	tagacatcaa aattatttat atgactgaga aactagacac tataaatgag cattcagtat	12540
[0624]	tcaacaccgt gattttagat attgtcacaa tgacagaaaa ttttcttata gaaaatttta	12600
[0625]	agttttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtatttt	12660
[0626]	tagtttataa taaaagttac atataattaa aatggttggc acaaaacaac atttgagcat	12720
[0627]	ttttctatt tactatcaag tagtatcatt ttgaaataat aatttgacta gtttcaaaaa	12780
[0628]	tgaaaacaaa atttaacta aatgcctaatt ctagcctgat aacattttta tgaatgaaat	12840
[0629]	tattcaatag ttttatcaat taggggcccc aaacttttcc taaaataaaa cttttaattt	12900
[0630]	ttttccattt ttatttaaat tagaaacaaa attgttttac atgtaaatca gagtttcctc	12960
[0631]	accctccctt tctccctgtc cctcactaac accctacttg tcccatacca tttctgctcc	13020
[0632]	ccaggaggagg tgaggccttc catggggaaa ctccagagtc tgtctatcct ttcggatagg	13080
[0633]	gcctaggccc tcacccattt gtctaggcta aggcctcaca agtttactcc tatgctagtg	13140
[0634]	ataagtactg atctactaca agagacacca tagatttcct aggccttctc actgacaccc	13200
[0635]	atgttcatgg ggtctggaac aatcatatgc tagtttcccta ggtatcagtc tggggaccat	13260
[0636]	gagctcccc ttgttcaggc caactgtttc tgtgggtttc accaccctgg tcttgactgc	13320
[0637]	tttgctcacc actcctccct ttctgttaact ggggttccagt acaattccgt gtttagctgt	13380
[0638]	gggtgtctac ttctactttc atcagcttct gggatggagc ctctaggata gcatacaatt	13440
[0639]	agtcatcacc tcattatcag ggaagggcatt ttaaagtagc ctctccattg ttgcttggat	13500
[0640]	tgttagttag tgtcatcttt gtagatctct ggacatttcc ctagtgccag atatctcttt	13560
[0641]	aaacctacaa gactacctt attatgggat ctcttttctt gctctcgtct attcttcag	13620
[0642]	acaaaatctt cctgctccct tatattttcc tctcccttcc tcttctcccc ttctcattct	13680
[0643]	cctagatcca tcttcccttc ccccatgctc ccaagagaga tgttgctcag gagatcttgt	13740
[0644]	tccttaacce ttttcttggg gatctgtctc tcttaggggt gtccttggtt ctagcttct	13800
[0645]	ctggaagtgt ggattgtaag ctggtaaatca tttgctccat gtctaaaatc catatatgag	13860

[0646]	tgatgtttgt ctttttgtga ctgggttacc tcaactcaaaa tggttttcttc catatgtctg	13920
[0647]	tggatttcaa tagcacaaa aacatacagt atcttggggc aacactaacc aaacaagtga	13980
[0648]	aagaccagta tagcaagaac tttgagttta aagaaagaaa ttaaagaaga taccagaaaa	14040
[0649]	tggaaagatc tcccatgctc tttgataggc agaatacaaca tagtaaaaat ggcaatcttg	14100
[0650]	ccaaaatcca tctacagact caatgcaatc cccattaaat accagcacac ttcttcacag	14160
[0651]	acctgaaaga ataatactta actttatatg gagaaacaaa agaccagga taggccaac	14220
[0652]	aaccctgtac aatgaaggca cttccagagg catcccatc cctgacttca agctctatta	14280
[0653]	tagagtaata atcctgaaa cagcttggtg atggcacaaa aatagacagg tagaccaatg	14340
[0654]	gaattgagtt gaaaaccctg atattaaccc acatatctat gaacacctga ctttgacaaa	14400
[0655]	gaagctaagg ttatacaatg taagaaagaa agcatcttca acaaatcgtg ctggcataac	14460
[0656]	tggatgctgg catgtagaag actgcagata gatccatgct taatgccatg cacaaaactt	14520
[0657]	aagtccaaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag	14580
[0658]	aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca	14640
[0659]	gtagcacaga caatcagatc aataatcaat aaatgggacc tctgaaact gagaagcttc	14700
[0660]	tgtaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat	14760
[0661]	tcaccaatcc tatatctgac agagggtgct tctctatttg caaagaacac aataagctag	14820
[0662]	tttttaaac accaattaat ccgattataa agttgggtag agaactaat aaagaattgt	14880
[0663]	taacagagca atctaacttg gcagaaagac acataagaaa gtgctcacca t	14931
[0664]	서열번호 3	
[0665]	4001 염기	
[0666]	DNA	
[0667]	Cricetulus griseus	
[0668]	ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacatttc agtgtaatat	60
[0669]	tattcagttt ttaagaaaa tgaaattatg taataagcat gtaaatggat atatcttgaa	120
[0670]	acaaccattc cccattatat tacctaaaca ttgaaagtcc aaaatcatat gatcttttta	180
[0671]	gtggatctac taatcttttg ctatatgtat ttatttgaac tacccatgga tgtgagataa	240
[0672]	ttggtaacaa cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca	300
[0673]	ttttttagga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagtg	360
[0674]	atggttagag aaggcaatat gaggagggat aactagcact tagggccttt tgaaaaagac	420
[0675]	atagagaaaa tactatttga gaaacttcct ataattggtg tatagttata tacaccaag	480
[0676]	agctcagatg gagttaccct ataattgaaa tattaactac tttttatcac tgtgataaaa	540
[0677]	catcctgaac agagcaacat agattgggaa gcattttact ttgcttacag ttctaacggg	600
[0678]	ataaaaattc atgatgaaag aatgaatatg tcagcaacaa gcagtagcaa tggcctgaga	660
[0679]	agcaggtgag agctcacatc ttgaagtgtg agaattgtgc agagagaaca aactgcaaat	720
[0680]	gaccagaaaa tgcttttggg tcagagccca taccctctg actgacttct ccagaaattc	780
[0681]	tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta	840

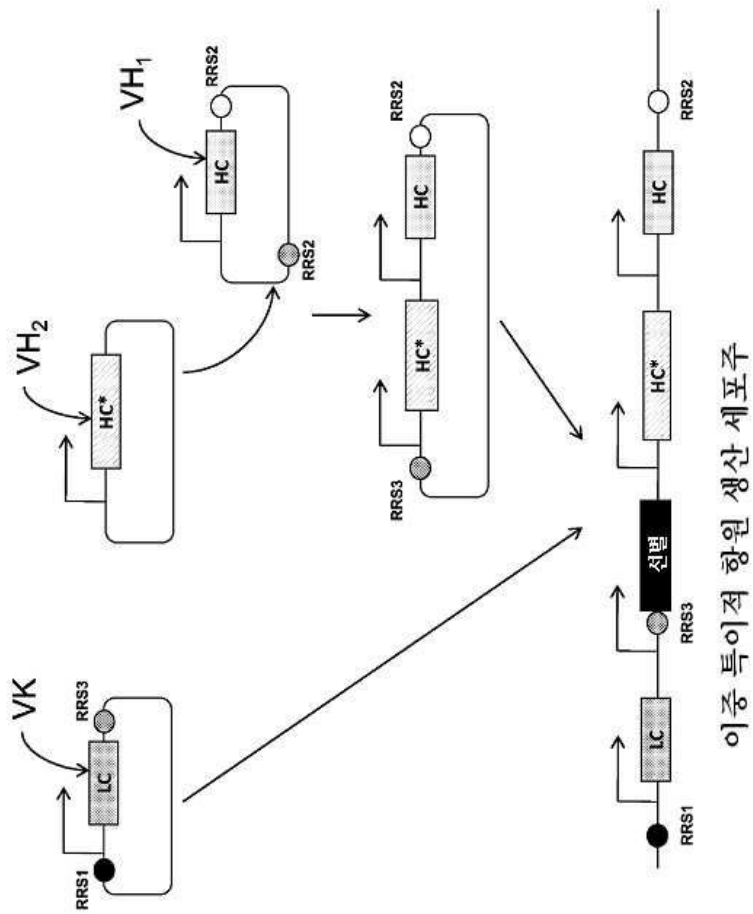


[0682]	ggggtatttc ttattcaaac cactacaatg gggtaggggg agcaatcctc caagtaggca	900
[0683]	ctacacacag acaataaaaa actctagtaa ctggaatgga ttgacttatt tgaattactt	960
[0684]	gccagtggag ctacatagag cacaattatt gtattttaat taccctttat gatcttacia	1020
[0685]	aacttgacag taagatcata ttgctaaaga aaccacatat ttgaatcagg gaacatgggtg	1080
[0686]	atatctagtt gttcttcaac tggaaacttc atgctttctg cccagcattc atgttgctgg	1140
[0687]	aaagagcaat gtacactacc agtgtagaaa ttaaatcatc aatcttatca agatgtggat	1200
[0688]	cctataagtt acaataaaaa ttagcctgat aagatatccc caccagaaga atattcacat	1260
[0689]	aatgctatg ggagcaacaa gctattttct aaattagctt taatcctatt ctacaagaga	1320
[0690]	gaatccatat ctagaatagt tatagggatc aagaacccat ggcttgattg gtcataggcc	1380
[0691]	caatgggaga tcctaattatt attgttctac aaaatgaaaa taactcctaa tgacttgttg	1440
[0692]	ctgcagtaat aagtttagat gttgctcaac tctcacaaga gaagttttgt cttacaataa	1500
[0693]	atggcaatta aagcagcccc acaagattta tatcataccg atctcctcat ggcctatgca	1560
[0694]	tctagaagct aggaaacaaa gaggacccta agagagacat acatgggtccc cctggagaag	1620
[0695]	gggaaggggg caagacctcc aaagctaatt gggagcatgg gggaggggag agggagttag	1680
[0696]	aagaaagaga aggggataaa aggagggaga ggaggacaag agagagaagg aagatctagt	1740
[0697]	caagagaaga tagaggagag caagaaaaga gataccatag tagagggagc cttgtatgtt	1800
[0698]	taaatagaaa actggcacta gggaattgtc caaagatcca caaggtccaa ctaataatct	1860
[0699]	aagcaatagt cgagaggcta ccttaaaagc ctttctctga taatgagatt gatgactacc	1920
[0700]	ttatatacca tcctagagcc ttcatccagt agctgatgga agcagaagca gacatctaca	1980
[0701]	gctaaacact gagctagttg cagacaggga ggagtgatga gcaaagtcaa gaccaggctg	2040
[0702]	gagaaacaca cagaaacagc agacctgaaa aaaatgttgc acatggacct cagactgata	2100
[0703]	gctgggagtc cagcatagga cttttctaga aaccctgaat gaggatatca gtttgagggt	2160
[0704]	ctggttaatc tatggggaca ctggtagtgg atcaatatit atccctagtt catgactgga	2220
[0705]	atttggttac ccattccaca tggaggaatt ctctgtcagc ctagacacat gggggagggt	2280
[0706]	ctaggtcctg ctccaaataa tgtgttagac ttggaagaac tcccttgaga agactcaccc	2340
[0707]	tccctgggga gcagaaaagg gatgggatga gggttggtga gggacaggag agggagggag	2400
[0708]	ggtgagggaa ctgggattga caagtaaatg atgcttgttt ctaattttaa tgaataaagg	2460
[0709]	aaaagtaaaa gaagaaaaga aaacaggcca aaagattata aaagacagag gtggtgggtg	2520
[0710]	actataaaga aacactatta tctaaataaa aacatgtcag aagcacacat gaacttatag	2580
[0711]	tgtttatgaa agtatgtata ataactacat aatctcaagc caagaaaaaa atatcatctt	2640
[0712]	tcagtgatga aggtgatitit atttctccca gaattaaagc caaagacctc atgaaagtaa	2700
[0713]	ttatcttcaa aaggttgaaa atacatactt tgcaatacac agatctgcct agaaatctca	2760
[0714]	tgttcacaat acacatgatg ctcaattgaa ttccattcaa tgttacagtt tagataaaca	2820
[0715]	gtttgtagat aaactcacia tgtatcatit ctttttatit tttagacaaa cagcttctca	2880
[0716]	tctgttatte agaataatc ctcgatggca ggatatccat cccaattggg ggaaggggag	2940
[0717]	aatttgaaga aaacctagac cacatacata tttagcattg ggaacaaaag tctaaaatga	3000

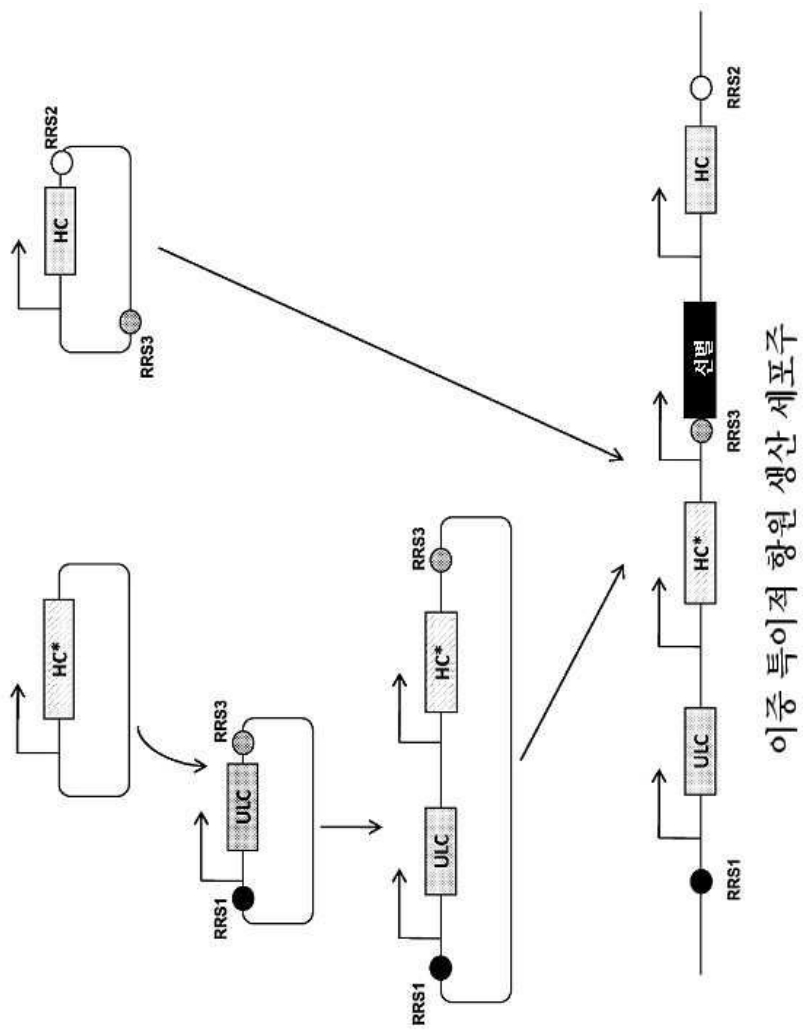
[0718]	tgtttgttcac atctttctcta ctatgtcctct ccccgtecca aagaaccttg gtatatgtgc	3060
[0719]	ctcatttttac agagagagga aagcaggaac tgagcatccc ttacttgcca tcctcaaccc	3120
[0720]	aaaatttgca tcattgtcga gctctgccct tctcatatga cagttacaag tcaaggcttc	3180
[0721]	caaagtcctt ctgtcatgtt tgggtgtcaat agtttataca gatgacttca tgtcttcata	3240
[0722]	tctaattgtct tatatagatt aatattaaac aatgttattt ctctaaccac attttaaatt	3300
[0723]	aatttaaaaa tcattaatt gtgtctataa aatgcagaca gagtgctgag acacaatata	3360
[0724]	agcctgatga tcigaatttg aaactcacac ccaccacatg gagaatcaac ttccaaaaat	3420
[0725]	tttctatta ctccacact tacaccattg tacaacaca ataataatga acaaaatgaa	3480
[0726]	atgaaataaa aaattaagtc tctgtaggta atgctactgt gcagcaaaag taaaaatggc	3540
[0727]	agcttaagct tgctttatgg ttacacttta ccatcttcca ttaattataa ggacttcaat	3600
[0728]	catggcagaa ctatgtctgtt attgtctcag tgtaacctaa ccagggtgtc cagatgttct	3660
[0729]	taatgtggac acctaaacta ttgatattt gggttaagat ctttccctct ttcagaagaa	3720
[0730]	acctcaggac agagggaatc ttgtctttta attttgagtc tgtagacttt ttccatttca	3780
[0731]	aatatacatg aaacaagtga tgaagaaaat taatcaaaag gtgggaattg caatgatatt	3840
[0732]	aggttcaata ttaagcttca atattatcat ggaatcgctt gttatacact gagtgtttgg	3900
[0733]	caataaggga tttttagaag aaggagtttt tattctcaac aggttcctta agtttagctc	3960
[0734]	aaataaatct aagcaatcca ctctagaatt aaatagtttc c	4001

도면

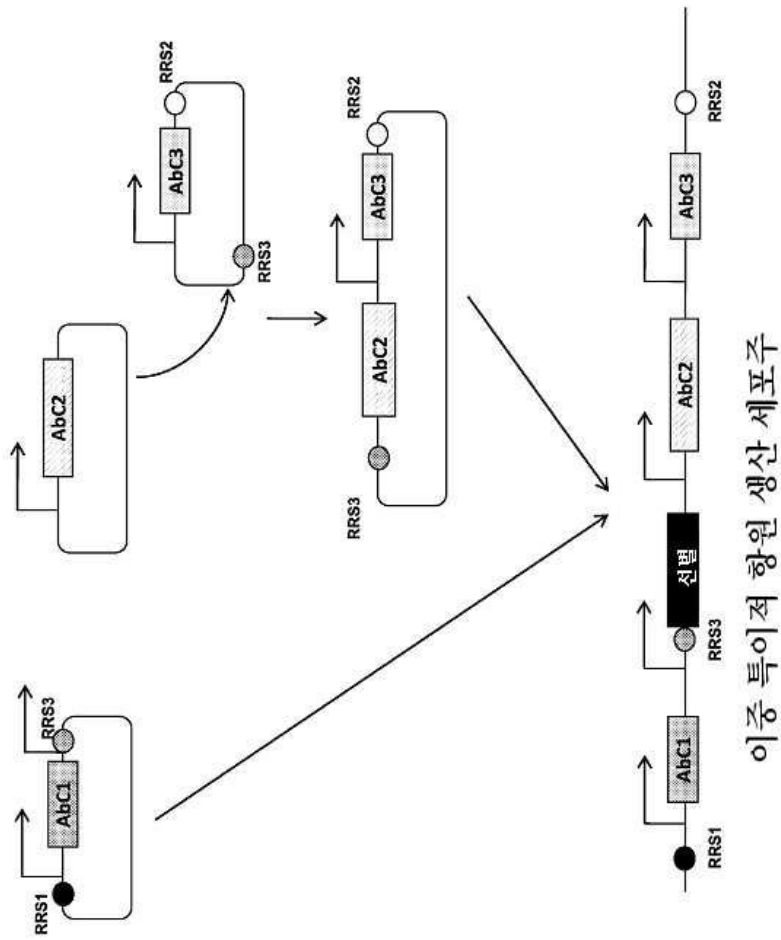
도면1



도면2

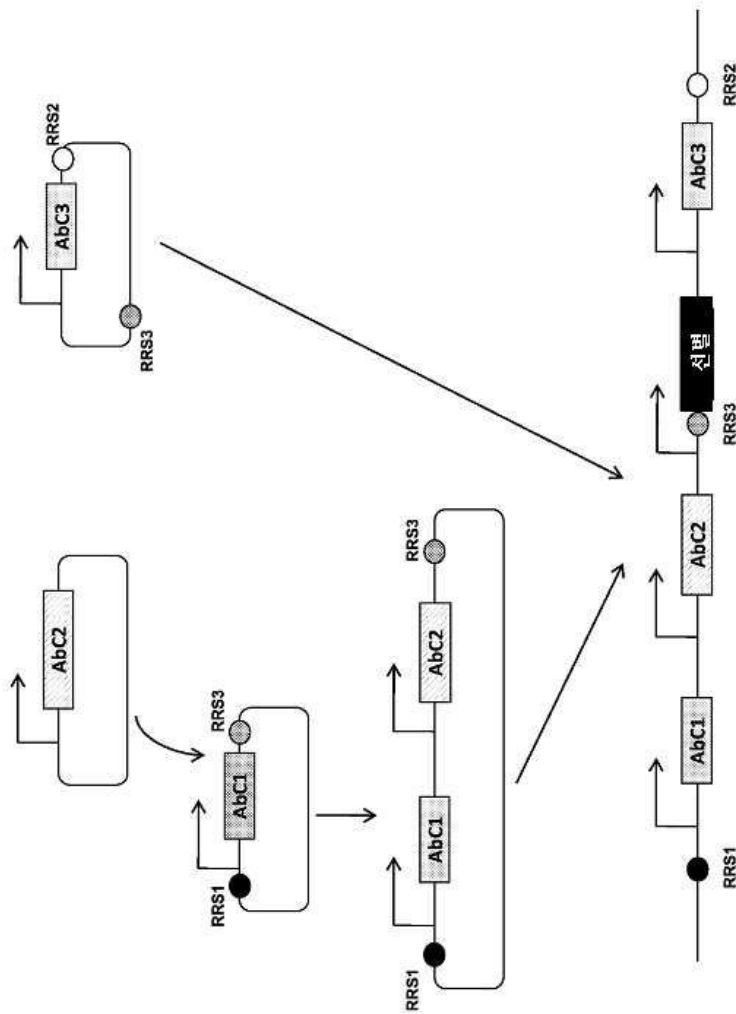


도면3



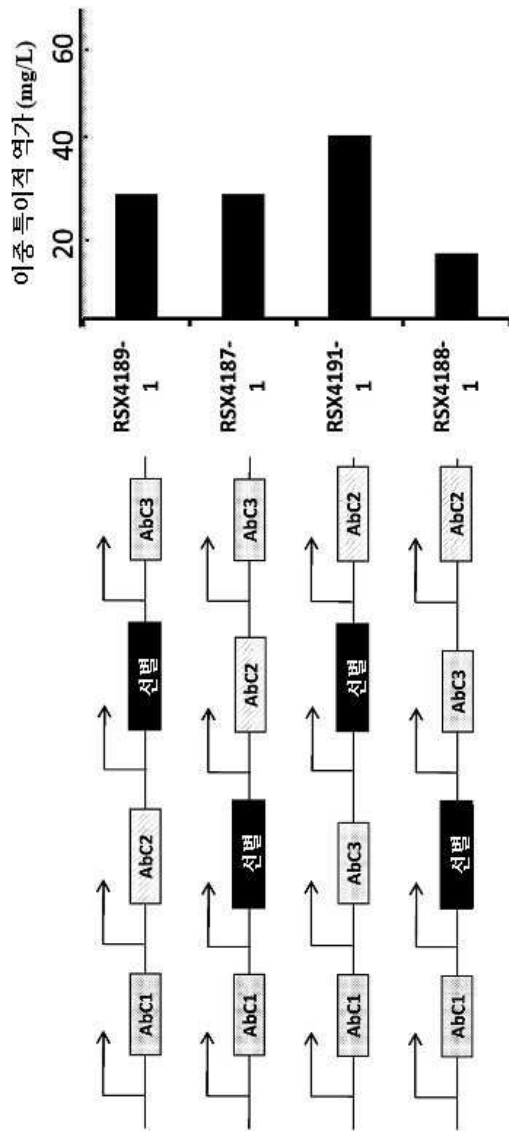


도면4



이중 특이적 항원 세포주

도면5



## 서열 목록

### SEQUENCE LISTING

- <110> Regeneron Pharmaceuticals, Inc.
- <120> COMPOSITIONS AND METHODS FOR MAKING ANTIBODIES BASED ON USE OF AN  
EXPRESSION-ENHANCING LOCUS
- <130> 32354PCT (T0047W001)
- <150> 62/325,385
- <151> 2016-04-20
- <160> 3
- <170> PatentIn version 3.5
- <210> 1

<211> 13515

<212> DNA

<213> *Cricetulus griseus*

<400> 1

tctagaaca aaacaaaaa tattaagtca ggcttgctt caggigctgg ggtggagtgc	60
tgacaaaaat acacaaattc ctggctttct aaggcttttt cggggattca ggtattgggt	120
gatggtagaa taaaaatctg aaacataggt gatgtatctg ccatactgca tgggtgtgta	180
tgtgtgtgta tgttgtctg tgttgtgtcc cagacagaaa taccatgaag gaaaaaaca	240
cttcaaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag	300
cacatgccag aggaggtgag gactgaacca ttcaacacaa gtggtgaata gtctgcaga	360
cacagagagg gccagaagca ctcagaactc caggggtca ggagtgggtc tctggaggct	420
tctgcccttg gaggttctctg aggaggaggc ttccatattg aaaatgtagt tagtgccgt	480
ttccattagt acagtacta gagagagctg agggaccact ggactgaggc ctatgctc	540
agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaaggaaa cagcaggtgt	600
gaggggtcag gggcaagtta gtgggagagg tcttcagat gaagtagcag gaacggagac	660
gcactggatg gcccacttg tcaaccagca aaagcttga tcttgttcta agaggccagg	720
gacatgacaa gggatgatctc ggtttttaaa aggcctttgtg ttacctaact acttctatta	780
gtcagatact ttgtaacaca aatgagtact tggcctgtat tttagaaact tctgggatcc	840
tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctccagcca ggccctggac	900
ccgggtccat tcatgcaaat actcaggac agattcttca ctaggtagt atgagctgtc	960
ttggatgcaa atgtggcctc ttcattttac tacaagtcac catgagtcag gaggtgctgt	1020
ttgcacagtg tgactaagtg atggagtgtt gactgcagcc attcccggcc ccagcttgtg	1080
agagagatcc ttttaaatg aaagtaagct caaagttacc acgaagccac acatgtataa	1140
actgtgtgaa taatctgtgc acatacacia accatgtgaa taatctgtgt acatgtataa	1200
actgtgtgaa taatctgtgt gcagcctttc cttacctact acctccagt gatcaggttt	1260
ggactgcctg tgtgtactg gacctgaat gtccccaccg ctgtccctg tcttttacga	1320
ttctgacatt ttttaataat tcagcggtt cccctctgct ctgtgcctag ctataccttg	1380
gtactctgca ttttggtttc tgtgacattt ctctgtgact ctgctacatt ctcatgac	1440
atgtgacaca gaaggtgttc cctctggaga catgtgatgt cctgtcatt agtggaatca	1500
gatgccccca aactgtgtc cagtgtttgg gaaagtgaca cgtgaaggag gatcaggaaa	1560
agaggggtgg aaatcaagat gtgtctgagt atctcatgtc cctgagtggc ccaggctgct	1620

gacttcactc ccccaagtga gggaggccat ggtgagtaca cacacctcac acatactata	1680
tccaacacac acacacacac acacacacac acgcacgcac gcacgcacgc acgcacacat	1740
gcacacacac gaactacatt tcacaaacca catacgcata ttacacccca aacgtatcac	1800
ctatacatat cacacatata cacccttcca cacatcacac acataccaca cccacacaca	1860
gcacacacat acataggcac acattcacac accacacata tacattttgtg tatgcataca	1920
tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtacgca cacatgcata	1980
cacacacata ggcacacatt gagcacacac atacatttgt gtacgcacac tacatagaca	2040
tatatgcatt tgtatatgca cacatgcatg cacacatata taggcacaca tagagcacac	2100
acatacatat gtgtatgcac acatgcacac accaatcaca tgggaagact caggttcttc	2160
actaaggttc acatgaactt agcagttcct ggttatctcg tgaaacttgg aagattgctg	2220
tggagaagag gaagcgttgg cttgagccct ggcagcaatt aaccccgccc agaagaagta	2280
ggttttaaaa tgagagggtc tcaatgtgga acccgaggg cgccagttca gagaagagac	2340
ctaccaagc caactgagag caaaggcaga gggatgaacc tgggatgtag tttagaacctc	2400
tgtaccagct gggttcatg ctattttgtt atatctttat taaatattct tttagtttta	2460
tgtgcgtgaa taccttgctt gcataaatgt atgggcactg tatgtgttct tgggtgccgt	2520
ggaggccagg agagggcatt gatcctccgg agctggcgtt tgagacagtt gtgaccaca	2580
gtgtggggtc tgggaactgg gtcttagtgt tccgaagtg cagctggggc tcttaacctc	2640
tgagccatcc ctccagcttc aagaaactta ttttcttagg acatggggga agggatccag	2700
ggcttttaggc ttgtttgttc agcaaatact ctttctgtgt attttgaatt ttattttatt	2760
ttactttttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaat	2820
cctctgtct cagctctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca	2880
cttttagagtc cccagccctt ctggacactt gttccaagta taatatatat atatatatat	2940
atatatatat atatatatat atatatgtg tgtgtgtgtt tgtgtgtgta tgagacactt	3000
gctctaaggg tatcatatat atccttgatt tgcctttaat ttatttttta attaaaaatg	3060
attagctaca tgcacctgt atgcgtctgt atcatctata tatccttctt tccttctctc	3120
tccttctctc tccttctctt caccccaag catctatctt caaatccttg tgccgaggag	3180
atgccaagag tctcgttggg ggagatggtg agggggcgat acagggaag agcaggagga	3240
aagggggaca gactggtgtg ggtcttttga gagctcagga gaatagcagc gatcttccct	3300
gtccctggtg tcacctctta cagccaacac cattttgtgg cctggcagaa gagttgtcaa	3360
gtggtgcga ggtctgccac acaaccccaa tctggcccca agaaaaggca cctgtgtgtg	3420

actctggggt taaaggcgt gcctggctgt ctccagctgg acttgaaact cccgtttaat 3480

aaagagtctt gcaaaataat acccgagag tcacagtgcc aggttcccgt gctttcctga 3540

agcgccaggc acgggttccc taggaaatgg ggccttgctt gccaaagtcc cacggcttgc 3600

cctgcaaacg gcctgaatga tctggcactc tgcgttgcca ctgggatgaa atggaaaaaa 3660

gaaaaagaag aagtgtctct ggaagcgggc gcgtcacac aaaccgcaa cgatttgtta 3720

aacactctcc attgagaatc tggagtgcgg ttgccctcta ctggggagct gaagacagct 3780

agtggggcgg gggggaggac cgtgctagca tccttcacg gtgctcgctg gctgtggtgc 3840

atgccgggaa ccgaaacgag gaactaaagt caagtcttgc tttggtggaa ctgacaatca 3900

acgaaatcac ttcgattgtt ttctctttt tactggaatt ctgggatttg atagatgggg 3960

gaggatcaga gggggagggg aggggcgggg agacggaggg aggaggggag gaggggagga 4020

ggggaggagg ggaggagggg aagggatgga ggaaaatact aacttttcta attcaacatg 4080

acaaagattc ggagaaagt caccgctagt gaccggaggg aggaatgccc tattgggcat 4140

tatattccct gtcgtctaata ggaatcaaac tcttggttcc agcaccaagg attctgagcc 4200

tatcctattc aagacagtaa ctacagccca cacggaagag gctatacaac tgaagaaata 4260

aaattttcac tttatttcat ttctgtgact gcatgttcac atgtagagag ccacctgtgt 4320

ctaggggctg atgtgctggg cagtagagtt ctgagcccg taactggaac aaccagaac 4380

tcccaccaca gtiagagctt gctgagagag ggaggccctt ggtgagattt ctttgtgtat 4440

ttatttagag acagggtctc atactgtagt ccaagctagc ctccagctca cagaaattct 4500

cctgttccgg ttccaaagt actggagtta tgagtgtgtg ttaattgaac gctaagaatt 4560

tgctgattga agaaaacctc aagtgggttt ggctaattcc cacgaccca gaggctgagg 4620

caggaggaat gagagaattc aaggtttgcc agagccacag ggtgagctca atgtggagac 4680

tgtgagggtg agctcaatgt ggagactgtg agggtagact caatgtggag actgtgaggg 4740

tgagctcaat gtggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca 4800

atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggggtgtgt 4860

tgagtggtag agcagttagt tgatttgaca tgcttgaggt ctcccgttcc atctgtggcc 4920

ctgcaacagg aaggagggga ggaagggggg gaacgagaga gaggaaagag agacagaagc 4980

taagataggg aatgagagag gaaggaagaa acgggaagaa attcagactc ctctctgagt 5040

tccccaacg cctagtgaac tcctgtgcac accctaaggt ggcctttgtg tggcactggc 5100

ttgggtggtc gggaaaggca ttttcagctt gttgcagaac tgccacagta gcatgctggg 5160



tccgtgaaag tttctgcccc ttaacaagaa gtctctacta cttgtgacct caccagtga	5220
aatttcttta attgtctcct ggtgttctgg gttttgcatt tttgtttcta aggatacatt	5280
cctgggtgat gtcatgaagt ccccaaagac acagtggggc tgtgttggat tgggaaagat	5340
gatttatctg ggggtgtcaaa aggaaaagaa gggaacagg cacttgggaa aatgtcctcc	5400
cggccacccg aattttggct tggcaaccgt ggtggaggag caagaaacac gtggacgttt	5460
gaggaggcat ggggtcctag gaggacagga agcagaagga gagagctggg ctgacagcct	5520
gcaggcattg cacagtttca gaaggagatt acagcatgac tgagttttta gggatccaac	5580
agggacctgg gtagagattc tgtgggctct gaggcaactt gacctcagcc agatggtatt	5640
tgaataacct gccttagag ggaaaacaga catagcaaac agagccacgt ttagtgatga	5700
aactctcact ttgcctgagt catgtgcggc catgccagg ggtcaggctg aactcaact	5760
caaaaacaag tgagaaattg aagacaatcc gtggtggcag ctactggaag ggccaccaca	5820
tccccagaaa gagtggagct gctaaaaagc catttgtgat aggcacagtt atcttgaatg	5880
catggagcag agattacgga aaaatcgaga atgttaatga ggcaacattc gaggtagtc	5940
attcagtgtg ggaaaccag acgcttccat ccctaaaag gaacatcttg ctctcagtca	6000
aaatggaaat aaaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat	6060
tttcacacgc acagtggata attttcatgt tggagtttat ttgtgctaaa aggcagaaaa	6120
gggtaaaaag cacatcttaa gatttatgag gtctacgaa taaaaataat gttacttaca	6180
gctattcctt aattagtacc ccttccacc tgtggtaatt tcctgagata gtcagtgggg	6240
aaaagatctc tctttctctt ctttctccc cteccctcct cteccctcct cctccctcc	6300
ctccctcctc tccctcctc ccccttctt tctttcttg ctecttctcc tctgctcct	6360
tctcccttct tcttctattt attctaagta gcttttaaca gcacaccaat tacctgtgta	6420
taacgggaaa acacaggtc aagcagctta gagaagattg atctgtgttc actagcgtgc	6480
aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttcctt atttggcacg	6540
gcacttagga agtggaaatc gcctaatact ctggtttgta ccaccttcc ctataatgga	6600
ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag gcctgttaaa	6660
tggtagctctt atttgcaaag aaggctgtaa ctgttagctt taaaagcctc tcctcaagaa	6720
agaagggaga aaggatatgg ctgacatat ctaatagact taaccactgt gaaaagcctt	6780
agtatgaatc agatagaacc tatttttaac tcagttttga aaaaaataat ctttatattt	6840
atttgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gaaccacatg	6900
tagcaggtgc tggaggaggc cagaagaggg caccagatct cctggaactg acaccacaca	6960
tggttatgag ctgcctgatg tgggtgctgg gaactgaact ctcgtgttct gcaagagcag	7020

caactgttct cttaaactgat gagccatctc tccagccccc ccataatct taattgttca	7080
ttttagtaaa ttttattcat aatcaattat cacagtataa aacaatgatt ttatatatat	7140
catatacata tcaaggatga cagttagggg gatatgtgtg tgtgtgtgtg tgtgtgtgtg	7200
tgtgtgtgtg tgtgttatct gtgtgtgtgc tttttaagaa ggtgccatag tcaactgcatt	7260
tctctgaagg atttcaaagg aatgagacat gtctgtctgc caggaaccct atcttctctt	7320
ttgggaatct gacccaaatg aggtattctg aggaactgaa tgaagagctc aagtagcagt	7380
gtcttaaac caaatgtgtc gtctagagaa agtcaacgtc atcagtgagc tgaggagaga	7440
tttactgagc ggaagacaag cgctctttga ttttaagtgc tcgaacagtc acggctgtgg	7500
agtggagcct gtgtcaggt ctgaggcagt ctttgctagc cagctgtgat gagcagtga	7560
gaaaggttg agatggagc aggggtgggag cagggtatg gttcagacta ggtatcgtga	7620
gcacaccagc tgggtgactt gtggtctgtg ggtcaggcgt tgtaaacgcc cttaggtca	7680
ggcagtcaca ttgcttgaag ctgaatgggt gaggcaacac agagagtga aagaaggcaa	7740
agtaccacct ctccccgac ccaggtcact tctgggttat agctgagact ccggacagca	7800
tgaaccagc tggttagagc ttcagggaag acttgatgtc tgcattgttc tatgaaatgt	7860
gattcggtag atctggagaa aatttataat gctggctcag tcaagcactg aacaaggta	7920
ccttggcttt gggagctaca tgacattgac ttgtaggcag actttttttt ttctgcccgc	7980
caattccag ataaccaata tggaggctca atattaatta taaatgctcg gctgatagct	8040
caggcttgtt actagctaac tctccaact taaatgaacc catttctatt atctacattc	8100
tgccacgtga ctttacctg tacttctgt tctctctct tgtctgactc tgcccttctg	8160
cttccagag tcttagtct ggttctctg cctaacctta tctgcccag ctgctgacca	8220
agcatttata attaatatta agtctcccag tgagactctc atccaggag gacttgggtg	8280
ctccccctc ctcatgcca tccgtgtctt cctcttccct cgttcccc tcttctct	8340
gtcttctct ctcaccct ctttcatag tattgatggc aagggtgttc tagaatggag	8400
gagtggccat aggcattgaa agaaaccagt taggatgtc tgtgaggggt tgtaatcata	8460
agcgatggac acaattcaag ccacagagt aagacggaag gatgcactgt gctctagagc	8520
aacttctggg gcagaatcac aggggtgagt tctgacttga gggcgaagag gccacagga	8580
aggagtgag tttgtctgag ctagaagcta cgccaccct cttgtagca gacctgcca	8640
caagcatgct ttgttaatca tgtgggatct gattttctc taaatctatg ttcaactctt	8700
aagaaaatgt gaattctcac attaaaattt agatatacgt cttttggtgg ggggggtgta	8760
aaaaatctc aagaatatgg atttctgggg gccggagaga tggtcagag gtttaagagaa	8820
ctggttgctc ttctagacat tctgagttca attccagca accacatggt ggctcacaac	8880

catctgtaat gcgacctggg gccatcttct gacatgcatg gatacatgca ggcagaaagc	8940
tgtatacata gtaaattgat aaatcttttt ttaaaaagag tatggattct gccgggtgtt	9000
ggtaggcgcac gcctttaatc ccagcactct ggaggcagag gcaggaggat ctctgtgagt	9060
tcgagaccag cciggtctat aagagctagt tccaggacag cctccaaagc cacagagaaa	9120
ccctgtctcg aaaaacaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaga gtatggattc	9180
taagaaagcc gtaacagctg gagctgtgta cggagttagc cgtggtacta gaagaacaga	9240
cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt gttgttttga	9300
gaccggctct tatgtctctc aggtggcct caaactgctg atcttcccgc ctctacctct	9360
caagtcctgg gactacttgg ctcataaaac agtttttgtc gggctccctg aagttatgg	9420
tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggct tgaatcccag	9480
acaatgactg catctcagga cagttgggaa gaggacaatg gcagaaggac ttagaaaaga	9540
tagactggag ggtggaaaag cagcaggaac agagaaacaa aacaggaagc ttgctatcca	9600
gggccactct ggagtcctgt ggcaagatgg aagcgggcta ggggaataca tttgtgctac	9660
tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat caatgttgaa	9720
ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg cctagagaaa	9780
ggaatggccc ctgcagaata gacagagtga atggtgccct ttatcatttg cttaaagtga	9840
ggagaaataa acatccttcc atagagtttc aggtaaatga accccacagt tcatctgtgc	9900
cgtggtggag gcctggccaa cagttaaaaa gattagacac ggacaaagtc tgaaggaaac	9960
acctcgaata ggaagaggag agccacctca ttctgtaact ttctcaagg ggaagatgtt	10020
ccaagagtgg gaataaatgg tcaaaggggg gatttttaat taggaaaacg atttctgtg	10080
tcacttgtga aactggaggt tgatttgggg cataggacaa tagatttgat gctttgcaa	10140
aagctgtttc aaagcagaga aatggaatag agacaattat gtagcgagga gggagggtgg	10200
ggcgaagatg gagacagaga agtggaagct gactttaggg aagaggaaca tagaccacag	10260
gggcggggcg gggggcaggg gcggggggcg gggctcaaag gaggcagtgg gaacgttgct	10320
agtgttcgca gcgtaagcgt gaatgtgcaa gcgtctttgt ggtgtgtgac caggagtagc	10380
gtggctggct tgtgtgtctc ttgtaatccc agtctttgag gtttccacac tgttccacag	10440
tgggtgtgat tttccctcgg agagcatgag ggctctgctt tccccacatc ctccccagcg	10500
ttcgttggtg ttgttttcca agatgttagt gggtagagaca aagcctctct gttgatttgc	10560
ctttaacagg tgacaaaaaa agctcaacca ggagacattt ttgccttctt ggaaggtaat	10620

gctcccatgt agagcaatgg gacccatctc taaggtgagg ctactcttgc agtttgcacc	10680
cagctcttct gatgcaggaa ggaagttggt gggcaagcaa gactgtttgc ttcttgcgat	10740
ggacacattc tgcacacaaa ggctcaggag gggagaaggc tgtttgatgt ttagcactca	10800
ggaaggcccc tgatgcatct gtgattagct gctccatct gtggagcaga cacggactaa	10860
ctaaaaacca gtgtttttaa attgtcaagc ctttaagggtg aggaaattga cttattgtgc	10920
tgggccatac gtagagcaag tgctctgcat tgggccaacc cccggctctg gtttctaggc	10980
accagaatgg cctagaacta actcacaatc ctccattcc aggtctcagg tgctagaatg	11040
aaccactata ccagcctgcc tgectgecta cctgccttcc taaattttaa atcatgggga	11100
gtaggggaga atacacttat cttagttagg gtttctattg ctgtgaagag acaccatgag	11160
catggcaact cttataaagg aaaacattta gttgggtggc agtttcagag gtttttagtac	11220
attgtcatca tggtcgggaa catgatggca tgcagacaga catggtgctg gagaaaggga	11280
tgagagtcc t acatcttga ggcaacagga cctcagctga gacactggct ggtaccctga	11340
gcataggaaa cctcacagcc caccctcaca gtgacatatt tccttcaaca aagccatacc	11400
tcctaatagt gccactccct atgagatgac agggccaatt acattcaaac tgctataaca	11460
ctttaaagta ttttattttt attattgtaa attatgtatg tagctgggtg gtggcagccg	11520
aggtgcacgc ctttaatccc agcacttggg aggcagaggc agatggatct ctgtgagttc	11580
aagaccagcc tggctataaa gagctagttg caaggaagga tatacaaaga acagttctag	11640
gatagccttc aaagccacag agaagtgtg tcttgaaaac caaaaattgt gctgggacct	11700
gtctctgctt tggttgcttc ccaactcccc agagctggac tcttggtcaa cactgaatca	11760
gttgcaaaat aaactcctgg attcctctct tgltaacagga gcccgaagtc aggcgcccac	11820
ttgtcttctc gcaggattgc catagacttt ttctgtgtgc ccaccattcc agactgaagt	11880
agagatggca gtggcagaga ctgggaaggc tgcaacgaaa acaggaagtt attgcaccct	11940
gggaatagtc tggaaatgaa gcttcaaac ttgcttcatg ttcagttgta cacagactca	12000
ctcccaggtt gactcacagc tgtaaatatt cctgactatg tctgactgc ttttatctga	12060
tgcttccttc ccaaaatgcc aagtgtacaa ggtgagggaa tcacccttgg attcagagcc	12120
cagggtcgtc ctcttaacc tggacttgtc ttctccggc agcctctgac accctcccc	12180
ccattttctc tatcagaagg tctgagcaga gttggggcac gctcatgtcc tgatacactc	12240
cttgtcttcc tgaagatcta acttctgacc cagaaagatg gctaaggtgg tgaagtgttt	12300
gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggatg tggcagcatg	12360
taccgaaat ccagaaactg gggaggtaga gacgatgag tgcccggggc tagctggctg	12420
ctcagccagc ctactgtaatt tgccaaattc caactcctat tgaaaaacct ttaccaaaca	12480

aacaaacaaa caaataataa caacaacaac aacaacaaac taccccatat aaggtgggcg 12540  
gctcttggct cttaggaat gactcacca aacccaaagc ttgccacagc ttttctctgg 12600  
cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac ttcttgggct 12660  
gggctgtgtg ctctaggcca ccaggaactt tctgtcttg ctctctgtct ggcacagcca 12720  
  
gagcaccagc acccagcagg tgcacacacc tccctccgtg cttcttgagc aaacacaggt 12780  
gccttggctt gtctattgaa cgggagtaag ttcttgcaga tgtatgcatg gaaacaacat 12840  
tgtcctgggtt ttatttctac tgttgtgata aaaaccgggg aactccagga agcagctgag 12900  
gcagaggcaa atgcaaggaa tctgcctcc tagcttctc cccatggctt gccgggcctg 12960  
ctttctgcaa gcccttctct ccccatggc atgctgaca tgaacagcgt ttgaaatgct 13020  
ctcaaatgtc actttcaaag aaggcttctc tgatcttct aactaaatca gaccatgttt 13080  
cacctgcat tatcttctct ctgtctgtct gtctgtctgt ctgtctatct gtctatcatc 13140  
  
tatcaatcat ctatctatct atcttctatt tatctaccta tcattcaatc atctatcttc 13200  
taactagtta tcatttattt atttgtttac ttactttttt tatttgagac agtatttctc 13260  
tgagtgcagc ccttggctgt cctggaaccc attctgtaac cagctgttcc tcaaactcac 13320  
agagatccaa ctgcctctgc ctctctgggt ctgggggtta agacgtgcac caccaacgcc 13380  
ccgtctatc atctatttat gtacttatta ttcagtcatt atctatctc taactatcca 13440  
tcctctgtct atccatcatc tatctatcta tctatctatc tatctatcta tctatcatc 13500  
atctataatc aattg 13515

<210> 2

<211> 14931

<212> DNA

<213> Cricetulus griseus

<220><221> misc\_feature

<222> (2716)..(2239)

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

<400> 2

catgtacact tatgcaagta tgatatggcc caacacagta ttttacacca atttttatct 60  
ataaaatata catgtacatc aaaatatatt attaataata acatcattat tctttctttc 120  
caagtaataa acacatacac tgaaattttg gttcttgtgg ataatttta tgaaacagga 180  
aatgcaaatt tatcttagca tgtttacttc actttctttg catagataac cagtaatcac 240  
attgatggat catgtagtga aatgtatatt taggtatcta aggaattttg gcttcgtttt 300



gtgcttgttg acactgaatt ctattcctaa caacagtgtg taaggattct gtctgatttc	360
ttttaccagt atttgtccat ttgcattttc tttattattc atggctgctg ttctagaaag	420
tggaaggtag tgtgtcaagt ctgtttaaca tgtttccctg atgatcagtg tcttaacacc	480
tctctgagta caigtgggcc aatgtcgttt ctagacccat ctattcttgc ttgacttate	540
ctggtacatg cctgccaaga aatttctcct catcctttct gtctcttcac tgatttactt	600
gatgtgtgga tttcacattg atcatatgga aatagaagat acaattttct ttattcacag	660
tttggaagac tticaatctc atagatcatc attatTTTTT gctactgttc cctatgctat	720
ggtgaaatTT ccatttgaat aattgcttaa acaattaaca agaaagaatc tatttttact	780
tgcaataact tccatttcag aacatttact aactgttac tataatccaaa aactagtTTT	840
atatatcatg tgagaaatga ctaattcata atttggccat gacatttttt tcagaaacag	900
aaaaagtgc caatacatc acaatgctat aaatatTAAG acttcagcaa attaaatatt	960
tattcatgat atcacataaa attcattttat tatgttttat ttaaatgtgt ttttaaaca	1020
gtggtatcac taaatattaa gttagatgtg tttatgtgct taatgaattt atattttaga	1080
atgttataag ttgtatatag tcaaatatgt aataaatTTT attttttagg tctttctcat	1140
taaggTattt taattttggg tcccttttcc agagtgactc tagctcatga tgagttgaca	1200
taaaaactaa acagtacaaa atgtacattg cattcagtat tgcacttgat ctttgcactg	1260
aagtttgagt cagttcatac atttagtact tgggaagtac attaaactaa ctttcattgc	1320
tctggcaaaa tgctcgataa gataagagtc tattgtggaa agccatggca gcaggaaagt	1380
aagactgctg atgatgttta atccatagtc aagacgcaga aggagatgaa tgctggtatc	1440
caacattttt tgctgttcat tttctctaga accctagtcc ataaagatgt atgacttgca	1500
ttcaaaatgc gtcccttca gttgttcaac ttttctgtaa atatcctttc aggcatgtct	1560
agaagattgt ttcgcaaata cttctcaatc cattcaagtt gatagtgcag attaatcact	1620
gcagaataaa agcctgtaac ttggctcacg tgccaaggaa tatgcacact cctgacacat	1680
caataagtaa atcaaagtgt agcttttgcc tttaacattg ccagacttat gtaatgttct	1740
gcacgttctt cctccatcac tttttattct aatgggtgtt ccttgacatt gaatcacgt	1800
gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta	1860
gatttactat tttacttaga attttttata attgagagaa tataatattt tcacagttaa	1920
ctatctgctg taaatagagg attttaaaaa aaatctctat aacttttttt tacaacacac	1980
agtaaaatta agttaaaatt taataaagtc actatgttga tttcaaagtg tgctacgcc	2040
acggtgtgta cgcagggtga gcagaagatg ccactaaggt gggttaaggc cgatgggttg	2100
gggtctgcgc tccttgaga tgagccccag gcggttcctt ggcaatcagc tgcatcatg	2160

atgcccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	2220
nnnnnnnnnn nnnnnnnnnn tgggtgactt tatggaaaga atttgataga ttcatgatg	2280
tagaagaatt ttattagct tattttacag gagactaaga ccctgggacc taaagatc	2340
tgggtcctga gaatcaggaa atgggtagag acgtgggtga tggtagaga cagattttag	2400
agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa	2460
caattatgct gttctttttc tttctgttgt atgatctgat gatgtagccc ccttgccaag	2520
ttccctgac ccccttgcca agttccctga ttgtaacagt atataagcat tgcttgagag	2580
catattcaac tacattgagt gtgtctgtct gtcatttctc cgccgattcc tgatttctcc	2640
ttgagccttt tcccttgttc tccctcggtc ggtgggtctcc acgagaggcg gtccgtggca	2700
aaagtgtata aatgttctaa aacatttgaa ctctaaaaca tgcaaatga aaaattaaaa	2760
taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaaacaata ctatataata	2820
ttttgttatt agaattcaaa atcacattag ttggatttaa tttgaacatt gcattctttc	2880
aataataatt tcaataaaaa aagtttccc atgatatag aaaataataa catatgtatc	2940
tatctattta tttaactaca catatatagc atttgtttca actaaaataa atgaatgagc	3000
aaagcaccta agtaattggt gtctattata tttatgaagc caatagtttc aaataaatta	3060
tcatgcataa ggaggtattg caaatgttaa accttttttg aaacagatat tcccagttac	3120
agaaattata atttctaate tttctataa gtagaatgat gataattaat ataggccatt	3180
tgtaataat gttcagatta aaatattctc tatttcacta gagaagaatg atattaatg	3240
tattatattt tatttcccat ttgtttgca ccactattct atatccctca gcagtttaaa	3300
ttgtttcac catatgtgtg tgtgtttgta tcttaaatat ggcaactaaa ttagaataat	3360
ttaatataaa tctttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct	3420
tttaattgtc caagaatact ttttcttcta ttttaggact gatcagaccc aggactaata	3480
ttttatatgt actaattcta tgtacaaaa tatgttatta tctcatgaat tctgtctcaa	3540
tattgaggta ataaaaatag tccatcatga actttaaaat taaaataatg attaattaat	3600
ttttattcat attttgtttg tatgaatggt tatacatcac atgtgtgcct ggtgactgtg	3660
aatgtcagga gaaggtafga aagccactgg aattggaata agagataata ttgagatgt	3720
tatgtgggtg ctgagaatta gacgcaagcc atcttcaaga atagccagca tactatacca	3780
ctgagtaac cattcatccc tcaataattia tctttgtaga cagtaaatat atttctaaac	3840
tataaatgac cagaaaaatt aatgtattat taatgaagac attcatctca tgtgacacac	3900
ttcacctgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca	3960
cttactatct ctaaagctgt ccaatggggg ccagtcacca gtcagcacc agtgagataa	4020

tccatgaatg catttatatc ttaggaaaaa ttcttatcta tgtagtattt agaacatttt	4080
catgtgaggg gataaacaag gaagcacaga tgctttctga tagaaacttt ctctttaatt	4140
catctagaaa aaaaaaacct ctcaggaaaa tctctcttgc tctcctccca atgctctatt	4200
cagcatcttc tcctactta attctagatc tttttctcta tgcctccttg ctgctgccct	4260
gctggctctg ctctatgcct ccccatgtca cttttctttg ctatctcacc gttaccttct	4320
ctgcctcact ctctgccttc ttctctgctt ctcacatggc caggtctctgg acaattatag	4380
ttatatgtta cattctcata acacatgata tgcacatag tttctctcag gctagggata	4440
tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagttg gtgatggttc	4500
ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtacatg aagactgaat	4560
ttttctcaaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgtat	4620
ttgtctttag aggcaaccac atattgtctg cagggcctgc tttgaatttg ctctgaagtt	4680
agcttgtttg tgtaaaagga agaatcctat atcagcctga gaaatgtaaa atatcctagc	4740
atttcaagtc atcaaaatta tatggagagt ataaatcatc cttctgacta ttcatagtca	4800
tatttgtgtc caccaagtat aaaacacact accaaagggc tgtggaaaaa atcgccataa	4860
ctgttcttat tagggaggca tagcagtggi acctgaggaa gttacagcaa caaccagtca	4920
tccagtcaat aaccccatgg ctttgccact tggaggfacc caataatgtt tggctttgcc	4980
gagtaggact ccaacaaatt cagaggggtca atttttaaat gctggttgtc actgctgaac	5040
agtccattg ccctctgcct aattccacaa tggaaagctt ttactactga ttgccaatca	5100
ttaaacagcc tactcagcat aaacaggtat gatattattc tgcattttgt tacattacta	5160
gatgaattcc tatttcttcc tacaatagtg gaactgaaaa aagatacaca atcactactac	5220
ccctctacta atcttatgac ttatatcatt tcaattttca gaccataatg caaactattg	5280
acaaaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa	5340
ggcggactta tttgtttta tttcttagca tgcatagcaa tacatgattt gaggtttata	5400
taataaaggg acaataaatc ttcaagaaac ttaccctac tgaattaaaa tattaagaa	5460
ggtcacacat ttactcaaat atattagact actgggcaaa tagacatgaa aagtagagtt	5520
aatattgagg taggccttct gtgaaatgtc taaggaaatt atgtttcata cagtgtgtaa	5580
ccaagtggga atcatatcag aaagcagtca aaagcttata ttacaagtaa cagatgcttg	5640
gttatatgac ctcccagagc ttgactgtct atacacaaaa agtgggtgtta ataaaactgt	5700
aatttgggct atgttttttt aaatggcttc accaacatga aaggaaggga atgagcatgt	5760

catggatgct tagagattat gcttccagca agaagaattg agctttggct cttattacag	5820
aaacatgaca aggtgtgagt tttattttatt agaaattata taatatTTTA agctggggac	5880
taaaaatttt attgaaacaa acaggcaagg gataggcatg tactagaagc aaaaatagga	5940
tgtcaatgct gtaatgttat tttttggacc aaaatagtat ttcctataga aatgacaatg	6000
atcttaggtt attattcttc ataaagatga caagttcaca agatatecta gttcattaaa	6060
atcgTTTTag tcatTTaata gagtGctgtg atagattaca caaaggaaag cacttacgat	6120
gagaaataat gatatecaca attattttct taattcttag aaacattcta ttgttatatc	6180
tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata	6240
ttgtctatgg tgacatttca aagaacatgt gacgtacagt gtagcacaga taaagaacat	6300
aactgcagct gaatcagtaa ctaaacttac atacattaaa tctgccatgt tggcaacagt	6360
gtgtgcacta ccaaaggatg tactaatgct cagcacactc ccctatgtca cctttgttc	6420
atcattacat cataggctca ttttgtttgc ttttgaaatc tagaccaagt cttttgtgtc	6480
ttccaagca cagagctcat taatttacct catagacttg ttaaacttct tctggttcat	6540
caattgaata gaaatactca ctactaatta tgtgagaccc tgccagtacc atagcacatg	6600
gataattttt acataaaaca tgcatacaag taagattatt cagactgaac atgaatttta	6660
gagaaatcag gaaggagtat atgggagtgg ttggagtgag actagagaaa tgtaattaaa	6720
ctataatctc aatacaaaaga tctactaagc aaaaaacatg aaacattgtc attcaagtga	6780
aacatcagtc ttcaaatgg aaagatatTT ttactaggaa aatgtctggt agatggttat	6840
tatctagaaa acacaaaaat tagaaaacgg taaactTTaa taaaaagaat aatacaatga	6900
gactacatga aaagtTctta actaatgaaa caaatatctt gaaactTTTT tcttaaaagt	6960
ttaatatcaa taaccatcat ggaaattcaa attaaaacta tttacatatt acccctgaaa	7020
taataactaa tacccaataa aaataatata aacaaaaaat ggcaatgcat gccatcatgg	7080
atttgggaga gagaatgttc attgcagttc tgaatggata ctggtgccac cacggtgaaa	7140
atctctgtat aggtccttcc aaaagctgaa aatagacata tcacaagacc tgccacacat	7200
ttttcaagca aatacccaaa ggactctacc tgactgcaga gacactttct cataaaatat	7260
tattgttgat ctattcataa tatctggaaa atagaaacag ccaagatgcc catcaactga	7320
ttaatagatg ataaaattat tgtacatttc agtgaatat tattcagttt ttaagaaaaa	7380
tgaaattatg taataagcat gtaaatggat atatcttgaa acaaccattc cccattatat	7440
tacctaaca ttgaaagtc aaaatcataT gatcttttta gtggatctac taatcttttg	7500
ctatatgtat tttattgaac taccatgga tgtgagataa ttggtacaa cagcacatgg	7560
gagagcatgg gatattcaa ggaagattag agagaatgca ttttttagga gataatggag	7620

gagcaataga aaggattaaa tgaggttact gatgaaagtg atggttagag aaggcaatat	7680
gaggagggat aactagcact tagggccttt tgaaaaagac atagagaaaa tactattgta	7740
gaaacttcct ataattggtg tatagttata tacaccaaag agctcagatg gagttaccct	7800
ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat	7860
agattgggaa gcattttactt tggcttacag ttctaacggg ataaaaattc atgatgaaag	7920
aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc	7980
ttgaagtgtg agaattgtagc agagagaaca aactgcaaat gaccagaaaa tgcttttggg	8040
tcagagccca taccctctg actgacttct ccagaaattc tgaacaaata aaactcccca	8100
aacagagcca taactgaagg tccagtgtct gagactacta ggggtatttc ttattcaaac	8160
cactacaatg gggtaggggg agcaatctc caagtaggca ctacacacag acaataaaaa	8220
actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtggag ctacatagag	8280
cacaattatt gtattttaat taccctttat gatcttaca aacttgacag taagatcata	8340
ttgctaaaga aaccacatat ttgaatcagg gaacatggtg atatctagtt gtcttcaac	8400
tggaaacttc atgctttctg ccagcattc atgttgctgg aaagagcaat gtacactacc	8460
agtgtagaaa ttaaatcatc aatcttatca agatgtggat cctataagtt acaataaaaa	8520
ttagcctgat aagatatccc caccagaaga atattcacat aaatgctatg ggagcaacaa	8580
gctattttct aaattagctt taatcctatt ctacaagaga gaatccatat ctagaatagt	8640
tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tcctaatatt	8700
attgttctac aaaatgaaaa taactcctaa tgacttggtg ctgcagtaat aagttagtat	8760
gttgctcaac tctcacaaga gaagttttgt cttacaataa atggcaatta aagcagcccc	8820
acaagattta tatcataccg atctctcat ggcctatgca tctagaagct aggaaacaaa	8880
gaggacccta agagagacat acatggtecc cctggagaag gggaaggggg caagacctec	8940
aaagctaatt gggagcatgg gggagggggag agggagttag aagaaagaga aggggataaa	9000
aggagggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag	9060
caagaaaaga gataccatag tagagggagc cttgtatgtt taaatagaaa actggcacta	9120
gggaattgtc caaagatcca caaggtccaa ctaataatct aagcaatagt cgagaggcta	9180
ccttaaaagc ctttctctga taatgagatt gatgactacc ttatatacca tcctagagcc	9240
ttcatccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagtgt	9300
cagacaggga ggagtgaiga gcaaagtcaa gaccaggctg gagaacaca cagaaacagc	9360
agacctgaaa aaaaattgtc acatggaccc cagactgata gctgggagtc cagcatagga	9420
cttttctaga aaccctgaat gaggatatca gtttgagggt ctggttaatc tatggggaca	9480

ctggtagtgg atcaatatatt atccctagtt catgactgga atttgggtac ccattccaca	9540
tggaggaatt ctctgtcagc ctagacacat gggggagggt ctaggtcctg ctccaaataa	9600
tgtgttagac tttgaagaac tcccttgaga agactcaccc tccctgggga gcagaaaggg	9660
gatgggatga gggtttggtga gggacaggag aggaggggag ggtgaggga ctgggattga	9720
caagtaaag atgcttgttt ctaattttaa tgaataaagg aaaagtaaaa gaagaaaaga	9780
aaacaggcca aaagattata aaagacagag gtggtgggtg actataaaga aacactatta	9840
tctaaataaa aatatgtcag aagcacacat gaacttatag tgtttatgaa agtatgtata	9900
ataactacat aatctcaagc caagaaaaaa atatcatctt tcagtgatga aggtgatatt	9960
atttctccca gaattaaagc caaagacctt atgaaagtaa ttatcttcaa aaggttgaag	10020
atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg	10080
ctcaattgaa ttccattcaa tgttacagtt tagataaaca gttttagat aaactcaca	10140
tgtatcatct ctttttattt ttgaccaa cagcttctca tctgttattc agaataattc	10200
ctcgatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac	10260
cacatacata ttgccattg ggaaacaaag tctaaatga tgtgtttcac atcttctcta	10320
ctagtcctct ccccgctcca aagaaccttg gtatatgtgc ctcatcttac agagagagga	10380
aagcaggaac tgagcatccc ttacttgcca tctcaaccc aaaatttgca tcattgctca	10440
gtcttgccct tctcatatga cagttacaag tcaaggcttc caaagtcct ctgtcatgtt	10500
tgggtgcaat agtttataca gatgacttca tgtcttcata tctaatgtct tatatagatt	10560
aatattaaac aatgttattt ctctaaccac attttaaat aatttaaaaa tccattaatt	10620
gtgtctataa aatgcagaca gagtgtgag acacaatata agcctgatga tctgaatttg	10680
aaactcacac ccaccacat gagaatcaac ttccaaaaat tttctatta ctccacact	10740
tacaccattg tacaaacaca ataataatga acaaatgaa atgaataaaa aaattaagtc	10800
tctgtaggta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgctttatgg	10860
ttacacttta ccattctcca ttaattataa ggacttcaat catggcagaa ctatgctgtt	10920
attgtctcag tgtaacctaa ccagggtgtc cagatgttct taatgtggac acctaaacta	10980
tttgatattt gggttaagat ctttcctct ttcagaagaa acctcaggac agagggaatc	11040
ttgtctttta attttagtc ttagacttt ttccatttca aatatacatg aaacaagtga	11100
tgaagaaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca	11160
atattatcat ggaatgcct gttatacact gagtgtttgg caataaggga tttttagaag	11220

aaggagtttt tattctcaac aggttcctta agtttagctc aaataaatct aagcaatcca	11280
ctctagaatt aaatagtctt ctaagggcac agctatgaat agagctcaat ttacatataa	11340
aattttgttc accatttatg tcattccagt ttctattagt acaaggaaaa tacaaaaat	11400
ttagatgtca atatcaagtg aatagtccat ctcccttttt aatatatctc acctaaatca	11460
ccattttctc agaaaaatct ggcctgaagt tctgtctgga acttcaacat gaaaaaatg	11520
cacagcttgc tattataaat cctagttagt tttaagatt catgtctggt gtctgactca	11580
gaggggccag aggctagaca aatatTTTTT gaatcttcat tgtgaagatt tttaatgatt	11640
attttaatat aaataacaaa gatgatggat aatgtaactt tgtacagttc atagacgctg	11700
aactactttg tgcttaaaat gttagtctcc tatcataaat gatagggtgat aagtgtatgt	11760
ttaatacttt ccctctgagc tataattcatg tactagagaa ttattttaaa catgaaaaga	11820
ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcaggtg aatgccagga	11880
gcaacgtttt tcttctacag aggatgcttt gctgccaagc aacctggttg tgtggaaatg	11940
ttcctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tattttcagg	12000
ttagagcttg gtccctggag tattatcttt taccagaaaa ttcatagtat tctttcaata	12060
actaacaact aaacttttcg ataaaaaaga attggaattt caattttaaa gcctgagtaa	12120
aattcttgtg aatcaggata ttttatttta agtcttatct tttaaaaagt tattttattt	12180
tttaaaaaat tataatatac ttccataatt tcctctcttc acttttcttt acaaacactt	12240
ctatagatca ccatgtgttt ttttttttac atttatggcc tctttctgtt cattgttatt	12300
acatacaaat agtcttgctt atagaagaac accacaattt gttacctgat aacaaattat	12360
caacccttaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa	12420
gatatatgtg tgtgcacata tatagataca catatatgta ggatttttaa ttttagattt	12480
tagacatcaa aattatttat atgactgaga aactagacac tataaatgag cattcagtat	12540
tcaacaccgt gattttagat attgtcacia tgacagaaaa ttttcttata gaaaatttta	12600
agttttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtatttt	12660
tagtttataa taaaaagtac atataattaa aatggttggc acaaaacaac atttgagcat	12720
ttttctatt tactatcaag tagtatcatt ttgaaataat aatttgacta gtttcaaaaa	12780
tgaaaacaaa atttaacta aatgcctaatt ctgcctgat aacattttta tgaatgaaat	12840
tattcaatag tgitatcaat taggggccca aaacttttcc taaaataaaa cttttaattt	12900
ttttccattt ttatttaaat tagaaacaaa attgttttac atgtaaatca gagtttcttc	12960
acctccctt tctccctgtc cctcactaac acctacttg tccatacca tttctgtctc	13020
ccaggaggagg tgaggccttc catggggaaa cttcagagtc tgtctatcct ttcggatagg	13080



gcctaggccc tcacccatit gtctaggcta aggcctcaca agtttactcc tatgctagt	13140
ataagtagtg atctactaca agagacacca tagatttctt aggccttcctc actgacaccc	13200
atgttcatgg ggtctggaac aatcatatgc tagtttctta ggtatcagtc tggggacccat	13260
gagctcccc ttgttcaggt caactgtttc tgtgggtttc accaccctgg tcttgactgc	13320
tttgctcctc actcctccct ttctgtaact gggttccagt acaattccgt gtttagctgt	13380
gggtgtctac ttctactttc atcagcttct gggtatggagc ctctaggata gcatacaatt	13440
agtcactcctc tcattatcag ggaagggcat ttaaagtagc ctctccattg ttgcttggat	13500
tgttagttgg tgcactcttt gtagatctct ggacatttcc ctagtccag atctctcttt	13560
aaacctacaa gactacctct attatgggtat ctcttttctt gctctcgtct attcttccag	13620
acaaaatctt cctgctccct tataatttcc tctccctcc tcttctccc ttctcattct	13680
cctagatcca tcttccctc ccccatgctc ccaagagaga tgttgctcag gagatcttgt	13740
tccttaaccc ttttcttggg gatctgtctc tcttagggtt gtccttgitt cctagcttct	13800
ctggaagtgt ggattgtaag ctggtaatca ttgtctccat gtctaaaatc catatatgag	13860
tgatgtttgt ctttttgtga ctgggttacc tcaactcaaaa tggtttcttc catatgtctg	13920
tggatttcaa tagcacaaac aacatacagc atcttggggc aacactaacc aaacaagtga	13980
aagaccagta tagcaagaac tttagattta aagaaagaaa ttaaagaaga taccagaaaa	14040
tggaaagatc tccatgctc tttagataggc agaatacaaa tagtaaaaat ggcaatcttg	14100
ccaaaatcca tctacagact caatgcaatc cccattaaat accagcacac ttcttcacag	14160
acctgaaaga ataacttta actttatatg gagaaacaaa agaccagga taggccaac	14220
aacctgtac aatgaaggca ctccagagg catcccatc cctgacttca agctctatta	14280
tagagtaata atctgaaaa cagcttggta atggcacaaa aatagacagg tagaccaatg	14340
gaattgagtt gaaaacctg atattaaccc acatatctat gaacacctga ctttgacaaa	14400
gaagctaagg ttatacaatg taagaaagaa agcatcttca acaaatcgtg ctggcataac	14460
tggatgctgg catgtagaag actgcagata gatccatgct taatgccatg cacaaaactt	14520
aagtccaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag	14580
aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca	14640
gtagcacaga caatcagatc aataatcaat aatgggacc tctgaaact gagaagcttc	14700
tgtaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat	14760
tcaccaatcc tatatctgac agagggtgc tctctatttg caaagaacac aataagctag	14820
tttttaaac accaattaat ccgattataa agttgggtag agaactaat aaagaattgt	14880
taacagagca atctaacttg gcagaaagac acataagaaa gtgctcacca t	14931

<210> 3

<211> 4001

<212> DNA

<213> *Cricetulus griseus*

<400> 3

```

ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacatttc agtgtaatat      60
tattcagttt ttaagaaaaa tgaaattatg taataagcat gtaaatggat atatcttgaa    120
acaaccattc cccattatat tacctaaca ttgaaagtcc aaaatcatat gatcttttta    180
gtggatctac taatcttttg ctatatgtat ttatttgaac taccatgga tgtgagataa    240
ttggttaaca cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca    300
tttttttaga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagtg    360
atggttagag aaggcaatat gaggagggat aactagcact tagggccttt tgaaaaagac    420

atagagaaaa tactattgta gaaacttcct ataattgggtg tatagttata tacaccaaag    480
agctcagatg gagttaccct ataattggaaa tattaactac tttttatcac tgtgataaaa    540
catcctgaac agagcaacat agattgggaa gcatttactt tggcttacag ttctaacggg    600
ataaaaattc atgatgaaag aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga    660
agcaggtgag agctcacatc ttgaagtgtg agaatgtagc agagagaaca aactgcaaat    720
gaccagaaaa tgcttttggg tcagagccca taccctctg actgacttct ccagaaattc    780
tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta    840

ggggtatttc ttattcaaac cactacaatg ggggtgggggg agcaatcctc caagtaggca    900
ctacacacag acaataaaaa actctagtaa ctggaatgga ttgacttatt tgaattactt    960
gccagtggag ctacatagag cacaattatt gtattttaa taccctttat gatcttaca    1020
aacttgacag taagatcata ttgctaaaga aaccacatat ttgaatcagg gaacatggtg    1080
atatctagtt gttcttcaac tggaaacttc atgctttctg ccagcattc atgttgctgg    1140
aaagagcaat gtacactacc agtgtagaaa ttaaattcatc aatcttatca agatgtggat    1200
cctataagtt acaataaaaa ttagcctgat aagatatccc caccagaaga atattcacat    1260

aaatgctatg ggagcaacaa gctattttct aaattagctt taatcctatt ctacaagaga    1320
gaatccatat ctagaatagt tatagggatc aagaacccat ggcttgattg gtcataggcc    1380
caatgggaga tcctaattatt attgttctac aaaatgaaaa taactcctaa tgacttgttg    1440
ctgcagtaat aagttagtat gttgtctaac tctcacaaga gaagttttgt cttacaataa    1500
atggcaatta aagcagcccc acaagattta tatcataccg atctcctcat ggcctatgca    1560

```

tctagaagct aggaaacaaa gaggacccta agagagacat acatgggtccc cctggagaag	1620
gggaaggggg caagacctcc aaagctaatt gggagcatgg gggaggggag agggagttag	1680
aagaaagaga aggggataaa aggagggaga ggaggacaag agagagaagg aagatctagt	1740
caagagaaga tagaggagag caagaaaaga gataccatag tagaggggagc cttgtatgtt	1800
taaatagaaa actggcacta gggaattgtc caaagatcca caaggtccaa ctaataatct	1860
aagcaatagt cgagaggcta ccttaaaagc ctttctctga taatgagatt gatgactacc	1920
ttatatacca tcttagagcc ttcatccagt agctgatgga agcagaagca gacatctaca	1980
gctaaacact gagctagtgt cagacaggga ggagtgtatga gcaaagtcaa gaccaggctg	2040
gagaaacaca cagaaacagc agacctgaaa aaaatgttgc acatggacct cagactgata	2100
gctgggagtc cagcatagga cttttctaga aacctgaat gaggatatca gtttggaggt	2160
ctggttaate tatggggaca ctggtagtgg atcaatatit atccctagtt catgactgga	2220
atttgggtac ccattccaca tggaggaatt ctctgtcagc ctacacacat gggggaggtt	2280
ctaggtcctg ctccaaataa tgtgttagac ttgaagaac tcccttgaga agactcaccc	2340
tccctgggga gcagaaaagg gatgggatga ggtttgttga gggacaggag aggaggggag	2400
ggtgagggaa ctgggattga caagttaatg atgcttgttt ctaattttaa tgaataaagg	2460
aaaagtaaaa gaagaaaaga aaacaggcca aaagattata aaagacagag gtggtgggtg	2520
actataaaga aacactatta tctaaataaa aacatgtcag aagcacacat gaacttatag	2580
tgtttatgaa agtatgtata ataactacat aatctcaagc caagaaaaaa atatcatctt	2640
tcagtgatga aggtgatttt atttctccca gaattaaagc caaagacctt atgaaagtaa	2700
ttatcttcaa aaggttgaaa atacatactt tgcaatacac agatctgcct agaaatctca	2760
tgttcacaat acacatgatg ctcaattgaa ttccattcaa tgttacagtt tagataaaca	2820
gtttgtagat aaactcaca tgtatcattt ctttttattt tttagacaaa cagcttctca	2880
tctgttattc agaataattc ctcatggca ggatatccat cccaattggg ggaaggggag	2940
aatttgaaga aaacctagac cacatacata ttgccattg ggaaacaaag tctaaaatga	3000
tgttgttcac atcttctcta ctatgcctct ccccgctcca aagaaccttg gtatatgtgc	3060
ctcattttac agagagagga aagcaggaac tgagcatccc ttacttgcca tctcaaccc	3120
aaaatttgca tcattgtcga gctctgccct tctcatatga cagttacaag tcaaggtctc	3180
caaagtcctt ctgtcatgtt tgggtgtcaat agtttataca gatgacttca tgtcttcata	3240
tctaattgtt tatatagatt aatattaaac aatgttattt ctctaaccac attttaaatt	3300

aatttaaaaa tccattaatt gtgtctataa aatgcagaca gagtgctgag acacaatata	3360
agcctgatga tctgaatttg aaactcacac ccaccacatg gagaatcaac ttccaaaaat	3420
tttcttatta ctccacact tacaccattg tacaacaca ataataatga acaaaatgaa	3480
atgaaataaa aaattaagtc tctgtaggia atgctactgt gcagcaaaag taaaaatggc	3540
agcttaagct tgctttatgg ttacacttta ccatcttcca ttaattataa ggacttcaat	3600
catggcagaa ctatgctgtt attgtctcag tgtaacctaa ccaggtgttc cagatgttct	3660
taatgtggac acctaaacta ttgatattt gggttaagat ctttcctct ttcagaagaa	3720
acctcaggac agaggggaatc ttgtctttta attttgagtc tgtagacttt ttccatttca	3780
aatatcatg aaacaagtga tgaagaaaat taatcaaaag gtgggaattg caatgatatt	3840
aggttcaata ttaagcttca atattatcat ggaatcgctt gttatacact gagtgtttgg	3900
caataaggga tttttagaag aaggagtttt tattctcaac aggttcctta agtttagctc	3960
aaataaatct aagcaatcca ctctagaatt aaatagtttc c	4001