



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2018년11월19일

(11) 등록번호 10-1920187

(24) 등록일자 2018년11월13일

(51) 국제특허분류(Int. Cl.)

*C12N 15/85* (2006.01)

(52) CPC특허분류

*C12N 15/85* (2013.01)*C12N 2800/107* (2013.01)

(21) 출원번호 10-2018-7005004(분할)

(22) 출원일자(국제) 2011년07월06일

심사청구일자 2018년02월20일

(85) 번역문제출일자 2018년02월20일

(65) 공개번호 10-2018-0021247

(43) 공개일자 2018년02월28일

(62) 원출원 특허 10-2013-7000305

원출원일자(국제) 2011년07월06일

심사청구일자 2016년03월09일

(86) 국제출원번호 PCT/JP2011/065916

(87) 국제공개번호 WO 2012/005378

국제공개일자 2012년01월12일

(30) 우선권주장

JP-P-2010-154782 2010년07월07일 일본(JP)

(56) 선행기술조사문헌

Nat Methods. 2007 Sep;4(9):747-53.

H.sapiens BAC clone RP11-115A14 GenBank:  
AC093770.4 (2005.05.13.)

(73) 특허권자

다이이찌 산쿄 가부시키가이샤

일본 도쿄도 츠오쿠 니혼바시 혼쵸 3-5-1

(72) 발명자

니시미야 다이스케

일본 도쿄도 에도가와구 기타카사이 1쵸메 16방  
13고 다이이찌 산쿄 가부시키가이샤 나이

이노우에 다츠야

일본 도쿄도 에도가와구 기타카사이 1쵸메 16방  
13고 다이이찌 산쿄 가부시키가이샤 나이

(74) 대리인

특허법인코리아나

전체 청구항 수 : 총 15 항

심사관 : 유성전

(54) 발명의 명칭 외래 유전자 발현 증강 활성을 갖는 DNA 요소

**(57) 요약**

본 발명에서는, 신규한 DNA 요소를 사용하여 포유류 세포에서의 외래 유전자의 높은 발현을 안정적으로 달성하기 위한 방법이 개시된다. 보다 구체적으로는, 외래 유전자 발현 단위체가 도입된 유전자자리 주변의 염색질 구조를 변화시켜 전사 활성화를 증강시키는 DNA 요소가 개시된다.

(52) CPC특허분류  
*C12N 2830/46* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

서열 목록에서 SEQ ID NO: 3로 표기되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.

#### 청구항 2

삭제

#### 청구항 3

제 1 항에 따른 폴리뉴클레오티드의 폴리뉴클레오티드 서열을 포함하는 외래 유전자 발현 벡터.

#### 청구항 4

제 3 항에 있어서, 외래 유전자에 의해 인코딩되는 단백질이 다량체 (multimeric) 단백질인 외래 유전자 발현 벡터.

#### 청구항 5

제 3 항에 있어서, 외래 유전자에 의해 인코딩되는 단백질이 이종-다량체 (hetero-multimeric) 단백질인 외래 유전자 발현 벡터.

#### 청구항 6

제 5 항에 있어서, 외래 유전자에 의해 인코딩되는 단백질이 항체 또는 이의 기능적 단편인 외래 유전자 발현 벡터.

#### 청구항 7

제 3 항에 따른 외래 유전자 발현 벡터가 도입된 형질전환 세포.

#### 청구항 8

제 7 항에 있어서, 포유류에서 유래한 배양 세포인 형질전환 세포.

#### 청구항 9

제 8 항에 있어서, 포유류에서 유래한 배양 세포가 COS-1 세포, 293 세포 및 CHO 세포로 이루어지는 군에서 선택되는 세포인 형질전환 세포.

#### 청구항 10

제 7 항에 있어서, 외래 유전자에 의해 인코딩되는 단백질이 다량체 단백질인 형질전환 세포.

#### 청구항 11

제 10 항에 있어서, 외래 유전자에 의해 인코딩되는 단백질이 이종-다량체 단백질인 형질전환 세포.

#### 청구항 12

제 11 항에 있어서, 외래 유전자에 의해 인코딩되는 단백질이 항체 또는 이의 기능적 단편인 형질전환 세포.

#### 청구항 13

제 7 항에 따른 형질전환 세포를 배양하고, 생성된 배양 산물로부터 외래 유전자에 의해 인코딩되는 단백질을 수득하는 것을 포함하는 것을 특징으로 하는, 단백질 제조 방법.

#### 청구항 14

제 1 항에 따른 폴리뉴클레오티드를 사용하는 것을 특징으로 하는, 외래 유전자 또는 외래 유전자 발현 벡터가 도입된 형질전환 세포에서의 외래 유전자 발현 증강 방법.

### 청구항 15

제 1 항에 있어서, 폴리뉴클레오티드가 형질전환 세포에서의 외래 유전자 발현 증강에 사용되는 폴리뉴클레오티드.

### 청구항 16

제 3 항에 따른 외래 유전자 발현 벡터를 사용하는 것을 특징으로 하는, 외래 유전자 또는 외래 유전자 발현 벡터가 도입된 형질전환 세포에서의 외래 유전자 발현 증강 방법.

## 발명의 설명

### 기술 분야

[0001]

본 발명은 DNA 요소를 갖는 외래 유전자 발현 벡터를 사용함으로써 외래 단백질 분비 능력이 증강된 형질전환된 포유류 숙주 세포 및 상기 숙주 세포를 사용하여 외래 단백질을 제조하는 방법에 관한 것이다.

### 배경 기술

[0002]

유전자 재조합 기술의 발달로 인해, 치료 단백질 및 항체 약물과 같은 단백질 의약품에 대한 상업적 시장은 빠르게 확장되고 있다. 특히, 항체 약물은 높은 특이성을 가지며, 인간 신체에 투여되더라도 이상 면역반응을 일으키지 않으므로, 이의 개발이 활발히 수행되고 있다.

[0003]

항체 약물로 대표되는 단백질 의약품이 제조되는 숙주 세포로서, 미생물, 효모, 곤충, 동물 또는 식물 세포, 트랜스제닉 (transgenic) 동물 또는 식물 세포 등이 사용될 수 있다. 단백질 의약품이 생물학적 활성 또는 면역원성을 갖도록 하기 위해서 접힘 (folding) 또는 글리코실화와 같은 번역-후 변형이 필수적이므로, 복잡한 번역-후 변형을 갖는 미생물로는 작업이 수행될 수 없거나, 상이한 글리칸 구조를 갖는 식물은 생물반응기로서 작동하는 숙주 세포로서 적합하지 않다. 인간의 것과 유사한 글리칸 구조를 갖고 안전하다는 것을 고려하여, 인간과 밀접히 관련된 종으로부터의 것인 CHO 세포와 같은 배양된 포유류 세포의 사용이 현재 표준이며, 이러한 세포를 사용하여 번역-후 변형이 수행될 수 있다.

[0004]

배양된 포유류 세포가 숙주 세포로서 사용되는 경우, 미생물 등과 비교하여 생장률이 낮고, 생산성이 낮으며, 비용이 높은 것 등의 문제점이 존재한다 (비특허 문헌 1 참조). 또한, 임상 시험에서 단백질 의약품을 사용하기 위해서는 다양한 제품을 투여하는 것이 필요하다. 그러므로, 이의 생산 능력의 결여는 또한 전세계적인 문제점이다. 따라서, 배양된 포유류 세포에서의 외래 유전자 생산성을 향상시키기 위해서, 프로모터, 인핸서, 약물 선별 마커, 유전자 증폭 및 배양 공학 기술 등의 많은 연구가 지금까지 수행되어 왔다. 그러나 현재 상황으로는, 유전자 발현을 균일하게 증가시킬 수 있는 시스템이 아직 구축되지 못하고 있다. 외래 단백질의 낮은 생산성의 원인 중 하나로서, "위치 효과" 가 고려된다 (비특허 문헌 2 참조). 외래 유전자가 숙주 세포에 도입되는 경우, 이는 숙주 염색체 게놈 내로 무작위하게 통합되고, 외래 유전자의 전사는 외래 유전자가 통합된 부위 주변의 DNA에 의해 크게 영향받는다. 위치 효과는 외래 유전자의 삽입 부위, 카피 수, 구조 등과 같은 인자들에 의해 영향받으나, 염색체 내의 삽입 부위를 제어하기에 매우 어렵다.

[0005]

상기 문제점을 해결하기 위해서, 조절 폴리뉴클레오티드 서열 (또한 DNA 요소로서 알려져 있음) 예컨대 유전자 자리 제어 부위 (LCR), 스캐폴드/매트릭스 부착 부위 (S/MAR), 인슬레이터, 유비퀴티스 (ubiquitous) 염색질 개방 요소 (UCOE) 및 항-리프레서 (STAR 요소) 가 최근에 확인되었다 (비특허 문헌 3 내지 6 참조). LCR 은 내재성 유전자 자리에서 염색질 구조를 개방하는데 필요하지 않다. 그러나, LCR 은 외래 유전자가 통합된 DNA 주변의 염색질 구조를 개방하고 외래 유전자 발현 단위체와 함께 사용되는 경우 광범위한 염색질을 리모델링하는 능력을 갖는 전사 조절 요소이며, AT-풍부 부위를 필요로 하는 것으로 여겨진다 (비특허 문헌 7 참조).

[0006]

LCR 로 대표되는 상기 언급한 DNA 요소는 종종 프로모터와 조합으로 사용되며, DNA 요소를 프로모터와 조합으로 사용하는 경우 외래 유전자의 발현 수준이 프로모터만 사용하는 경우에 비해 증가하는 것으로 알려져 있다. 그러나, 극소수 유형의 DNA 요소가 지금까지 보고되어 있고, 외래 유전자 발현의 증강에 기여하는 다양한 메커니즘은 서로 상이하다. 또한, DNA 요소 및 프로모터가 조합으로 사용되더라도, DNA 요소 및 프로모터의 제어 하에 충분량의 치료 단백질이 제조되지 않는다. 그러므로, 외래 단백질의 생산성을 증가시킬 수 있는

DNA 요소에 대한 충분한 지식을 얻었다고 할 수는 없다.

[0007] 따라서, 본 발명의 목적은 배양된 포유류 세포와 같은 숙주 세포에서의 외래 유전자 발현을 증강시키는데 있어서 높은 활성을 갖는 DNA 요소를 사용하여, 단백질 의약품에서 사용되는 외래 단백질의 제조를 증가시키는 방법을 제공하는 것이다.

#### 인용 목록

##### 비특허 문헌

[0010] NPL 1: Florian M. Wurm.(2004) Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol. 22(11): 1393-1398

[0011] NPL 2: Ted H. J. Kwaks and Arie P. Otte. (2006) Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. TRENDS in Biotechnol. 24(3): 137-142

[0012] NPL 3: Pierre-Alain Girod, Duc-QuangNguyen, et al. (2007) Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nat. Methods. 4(9) : 747-753

[0013] NPL 4: Adam C. Bell, Adam G. West, Gary Felsenfeld (2001) Insulators and Boundaries: Versatile Regulatory Elements in the Eukaryotic Genome, Science 291: 447-450

[0014] NPL 5: Steven Williams, Tracey Mustoe.et al. (2005) CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnol. 5(17) : 1-9

[0015] NPL 6: Arie P. Otte, Ted H. J. Kwaks. et al. (2007) Various Expression-AugmentingDNA Elements Benefit from STAR-Select, a Novel High Stringency Selection System for Protein Expression Biotechnol. Prog. 23: 801-807

[0016] NPL 7: Qiliang Li, Kenneth R. Peterson, Xiangdong Fang, and George Stamatoyannopoulos, (2002) Locus control regions, Blood 100(9): 3077-3086

#### 발명의 내용

##### 해결하려는 과제

[0017] 상기 기재한 바와 같이, 조절 폴리뉴클레오티드 서열인 DNA 요소의 유형이 아직 많지 않으며, 또한, 이들 중 극 소수의 DNA 요소만이 외래 유전자 발현을 증강시키는데 있어서 매우 유효하다. 본 발명의 목적은 외래 유전자 발현 단위체가 도입되는 유전자자리 주변의 염색질 구조에서의 변화 등에 수반되는 전사 활성화를 증강시키는 DNA 요소를 사용하여 포유류 세포에서의 높은 발현을 안정적으로 얻는 방법을 제공하는 것이다.

##### 과제의 해결 수단

[0018] 본 발명자들은 상기 문제점을 해결하기 위해 광범위한 연구를 수행하였고, 그 결과로, 배양된 포유류 세포에서 하나 이상의 특정 유형의 DNA 요소를 사용하여 발현되는 외래 단백질의 생산성 및 분비가 향상될 수 있다는 것을 발견하였으며, 그에 따라 발명을 완결하였다.

[0019] 즉, 본 발명은 하기의 발명을 포함한다:

[0020] (1) 서열 목록에서 SEQ ID NO: 1로 표기되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.

[0021] (2) 서열 목록에서 SEQ ID NO: 2로 표기되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.

[0022] (3) 서열 목록에서 SEQ ID NO: 3으로 표기되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.

[0023] (4) 서열 목록에서 SEQ ID NO: 4로 표기되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.

[0024] (5) 서열 목록에서 SEQ ID NO: 5로 표기되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.

[0025] (6) 서열 목록에서 SEQ ID NOS: 1 내지 5 중 어느 하나로 표기되는 폴리뉴클레오티드 서열의 3000 개 이상의 연속 뉴클레오티드를 포함하는 폴리뉴클레오티드.

- [0026] (7) 서열 목록에서 SEQ ID NOS: 1 내지 5 중 어느 하나로 표기되는 폴리뉴클레오티드 서열의 2000 개 이상의 연속 뉴클레오티드를 포함하는 폴리뉴클레오티드.
- [0027] (8) 서열 목록에서 SEQ ID NOS: 1 내지 5 중 어느 하나로 표기되는 폴리뉴클레오티드 서열의 1500 개 이상의 연속 뉴클레오티드를 포함하는 폴리뉴클레오티드.
- [0028] (9) (1) 내지 (8) 중 어느 하나에 따른 폴리뉴클레오티드의 폴리뉴클레오티드 서열에 대해 95% 이상의 상동성을 갖는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.
- [0029] (10) (1) 내지 (8) 중 어느 하나에 따른 폴리뉴클레오티드의 폴리뉴클레오티드 서열에 대해 99% 이상의 상동성을 갖는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.
- [0030] (11) (1) 내지 (10) 중 어느 하나에 따른 폴리뉴클레오티드의 폴리뉴클레오티드 서열 중 2 개 이상의 서열을 포함하는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드.
- [0031] (12) (1) 내지 (10) 중 어느 하나에 따른 폴리뉴클레오티드에서 선택되는 폴리뉴클레오티드 중 2 개 이상의 유형으로 이루어지는 폴리뉴클레오티드.
- [0032] (13) (1) 내지 (12) 중 어느 하나에 따른 폴리뉴클레오티드의 폴리뉴클레오티드 서열을 포함하는 외래 유전자 발현 벡터.
- [0033] (14) 외래 유전자에 의해 인코딩되는 단백질이 다량체 (multimeric) 단백질인, (13)에 따른 외래 유전자 발현 벡터.
- [0034] (15) 외래 유전자에 의해 인코딩되는 단백질이 이종-다량체 (hetero-multimeric) 단백질인, (14)에 따른 외래 유전자 발현 벡터.
- [0035] (16) 외래 유전자에 의해 인코딩되는 단백질이 항체 또는 이의 기능적 단편인, (15)에 따른 외래 유전자 발현 벡터.
- [0036] (17) (13) 내지 (16) 중 어느 하나에 따른 외래 유전자 발현 벡터가 도입된 형질전환 세포.
- [0037] (18) 포유류에서 유래한 배양 세포인, (17)에 따른 형질전환 세포.
- [0038] (19) 포유류에서 유래한 배양 세포가 COS-1 세포, 293 세포 및 CHO 세포로 이루어지는 군에서 선택되는 세포인, (18)에 따른 형질전환 세포.
- [0039] (20) 외래 유전자에 의해 인코딩되는 단백질이 다량체 단백질인, (17) 내지 (18) 중 어느 하나에 따른 형질전환 세포.
- [0040] (21) 외래 유전자에 의해 인코딩되는 단백질이 이종-다량체 단백질인, (20)에 따른 형질전환 세포.
- [0041] (22) 외래 유전자에 의해 인코딩되는 단백질이 항체 또는 이의 기능적 단편인, (21)에 따른 형질전환 세포.
- [0042] (23) (17) 내지 (22) 중 어느 하나에 따른 형질전환 세포를 배양하고, 생성된 배양 산물로부터 외래 유전자에 의해 인코딩되는 단백질을 수득하는 것을 포함하는 것을 특징으로 하는, 단백질 제조 방법.
- [0043] (24) (1) 내지 (12) 중 어느 하나에 따른 폴리뉴클레오티드 또는 (13) 내지 (16) 중 어느 하나에 따른 외래 유전자 발현 벡터를 사용하는 것을 특징으로 하는, 외래 유전자 또는 외래 유전자 발현 벡터가 도입된 형질전환 세포에서의 외래 유전자 발현 증강 방법.
- [0044] (25) 형질전환 세포에서의 외래 유전자 발현을 증강시키기 위한, (1) 내지 (12) 중 어느 하나에 따른 폴리뉴클레오티드의 용도.

### 발명의 효과

- [0045] 본 발명에 따라서, DNA 요소를 사용하여 외래 유전자 발현 벡터를 포유류 숙주 세포 내로 도입함으로써, 치료 단백질, 항체 등을 위한 외래 유전자의 발현이 유의하게 증강될 수 있다.

### 도면의 간단한 설명

- [도 1] 도 1 은 칩-온-칩 (ChIP-on-chip) 처리한 샘플이 항-아세틸화 히스톤 H3 항체로 특이적으로 염색질-면역

침전되었다는 것이 GAPDH 부위의 증폭에 의해 확인된 그래프를 나타낸다.

[도 2] 도 2 는 DNA 요소가 삽입된 SEAP 발현 벡터의 모식도이다.

[도 3] 도 3 은 DNA 요소가 없거나 DNA 요소 A2, A7, A18, B5 또는 C14 를 갖는 안정적으로 발현하는 CHO 세포주에서의, CMV 프로모터의 제어 하의 SEAP 발현을 나타내는 그래프이다. 발현 증강에 대한 DNA 요소 A2, A7, A18, B5 및 C14 의 효과를 확인하였다.

[도 4] 도 4 는 DNA 요소가 없거나 DNA 요소 A2 또는 A7 을 갖는 안정적으로 발현하는 CHO 세포주에서의, EF-1 $\alpha$  또는 SV40 프로모터의 제어 하의 SEAP 발현을 나타내는 2 개 그래프를 포함한다. 발현 증강에 대한 DNA 요소 A2 및 A7 의 효과를 확인하였다.

[도 5] 도 5 는 DNA 요소가 삽입된 항체 발현 (항체 유전자 X 중쇄 및 경쇄 동시 발현) 벡터의 모식도이다.

[도 6] 도 6 은 DNA 요소가 없거나 DNA 요소 A7 을 갖는 안정적으로 발현하는 CHO 세포주에서의, CMV 또는 EF-1 $\alpha$  프로모터의 제어 하의 항체 분비 수준 (ELISA 방법에 의해 측정) 을 나타내는 2 개 그래프를 포함한다. 발현 증강에 대한 DNA 요소 A7 의 효과를 확인하였다.

[도 7] 도 7 은 DNA 요소 A2 및 관련 서열의 서열 길이를 나타내는 표이다.

[도 8] 도 8 은 DNA 요소가 없거나 DNA 요소 A2 또는 관련 서열을 갖는 안정적으로 발현하는 CHO 세포주에서의 SEAP 발현을 나타내는 3 개 그래프를 포함한다. 발현 증강에 대한 DNA 요소 A2 및 관련 서열의 효과를 확인하였다.

[도 9] 도 9 는 DNA 요소 A7 및 관련 서열의 서열 길이를 나타내는 표이다.

[도 10] 도 10 은 DNA 요소가 없거나 DNA 요소 A7 또는 관련 서열을 갖는 안정적으로 발현하는 CHO 세포주에서의 SEAP 발현을 나타내는 3 개 그래프를 포함한다. 발현 증강에 대한 DNA 요소 A7 및 관련 서열의 효과를 확인하였다.

[도 11] 도 11 은 DNA 요소 A18 및 관련 서열의 서열 길이를 나타내는 표이다.

[도 12] 도 12 는 DNA 요소가 없거나 DNA 요소 A18 또는 관련 서열을 갖는 안정적으로 발현하는 CHO 세포주에서의 SEAP 발현을 나타내는 그래프이다. 발현 증강에 대한 DNA 요소 A18 및 관련 서열의 효과를 확인하였다.

[도 13] 도 13 은 DNA 요소 B5 및 관련 서열의 서열 길이를 나타내는 표이다.

[도 14] 도 14 는 DNA 요소가 없거나 DNA 요소 B5 또는 관련 서열을 갖는 안정적으로 발현하는 CHO 세포주에서의 SEAP 발현을 나타내는 그래프이다. 발현 증강에 대한 DNA 요소 B5 및 관련 서열의 효과를 확인하였다.

[도 15] 도 15 는 DNA 요소 C14 및 관련 서열의 서열 길이를 나타내는 표이다.

[도 16] 도 16 은 DNA 요소가 없거나 DNA 요소 C14 또는 관련 서열을 갖는 안정적으로 발현하는 CHO 세포주에서의 SEAP 발현을 나타내는 3 개 그래프를 포함한다. 발현 증강에 대한 DNA 요소 C14 및 관련 서열의 효과를 확인하였다.

[도 17] 도 17 은 DNA 요소가 없거나 DNA 요소 A2, A7, A18, B5 또는 C14 를 갖는 안정적으로 발현하는 HEK293 세포주에서의 SEAP 발현을 나타내는 그래프이다. HEK293 세포에서의 발현 증강에 대한 DNA 요소 A2, A7, A18, B5 및 C14 의 효과를 확인하였다.

[도 18] 도 18 은 DNA 요소 A2, A7 또는 A18 의 전체 길이 서열을 기준으로 출발점 및 종료점에서의 뉴클레오티드를 나타내는 도면이다.

[도 19] 도 19 는 DNA 요소 B5 또는 C14 의 전체 길이 서열을 기준으로 출발점 및 종료점에서의 뉴클레오티드를 나타내는 도면이다.

### 발명을 실시하기 위한 구체적인 내용

#### [0047] 구현예의 설명

이하, 실시예를 참조로 하여 본 발명을 구체적으로 기재할 것이다. 그러나, 이들 실시예는 본 발명의 기술적 범주를 제한하지 않는다. 본 발명의 실시예에서 사용되는 플라스미드, 제한 효소, DNA 변형 효소 등은 시판 제품이며 통상의 절차에 따라 사용될 수 있다. 또한, DNA 클로닝, 폴리뉴클레오티드 서열 결정, 숙주

세포의 형질전환, 형질전환된 숙주 세포의 배양, 수득한 배양액으로부터의 항체 단리, 항체 정제 등에 사용된 절차는 또한 당업자에게 잘 알려져 있으며 문헌에서 이용가능하다.

[0049] 본원에서 사용되는 바와 같은 용어 "유전자" 는 DNA 뿐 아니라 이의 mRNA, cDNA, 및 이의 RNA 를 포함한다.

[0050] 본원에서 사용되는 바와 같은 용어 "폴리뉴클레오티드" 는 핵산과 동일한 의미로 사용되며 또한 DNA, RNA, 프로브, 올리고뉴클레오티드 및 프라이머를 포함한다.

[0051] 본원에 사용되는 바와 같은 용어 "폴리펩티드" 및 "단백질" 은 구별 없이 사용된다.

[0052] 본원에서 사용되는 바와 같은 용어 "유전자 발현" 은 mRNA 가 유전자로부터 전사되는 현상 및/또는 단백질이 mRNA 로부터 번역되는 현상을 지칭한다.

[0053] 본원에서 사용되는 바와 같은 용어 "외래 유전자" 는 숙주 세포 내로 인위적으로 도입되는 유전자를 지칭한다.

[0054] 본원에서 사용되는 바와 같은 용어 "외래 단백질" 은 외래 유전자에 의해 인코딩된 단백질을 지칭한다.

[0055] 본원에서 사용되는 바와 같은 용어 "유전자 발현 단위체" 는 전사의 판독 프레임 (frame) 방향으로, 적어도 프로모터 부위, 외래 유전자 및 전사 터미네이터 부위 (폴리 (A) 부가 신호) 를 갖는 폴리뉴클레오티드를 지칭한다.

[0056] 본원에서 사용되는 바와 같은 용어 "외래 유전자 발현 증강 활성" 은 외래 유전자를 포함하는 유전자 발현 단위체 주변의 DNA 에 대한 전사 및 번역에 유리한 환경을 만들고 전사 및 번역 효율을 유의하게 향상시킴으로써 숙주 세포에서의 외래 단백질 제조를 증강시키는 활성을 지칭한다.

[0057] 본원에서 사용되는 바와 같은 용어 "DNA 요소" 는 폴리뉴클레오티드가 유전자 발현 단위체 부근에 위치하거나 유전자 발현 단위체를 포함하는 외래 유전자 발현 벡터 내에 위치하는 경우 외래 유전자 발현 증강 활성을 갖는 폴리뉴클레오티드를 지칭한다.

[0058] 본원에서 사용되는 바와 같은 용어 "항체의 기능적 단편" 은 Fab, F(ab')<sub>2</sub> 등을 포함하며 항원-결합 활성을 갖는 항체의 부분적 단편을 지칭한다. 그러나, 상기 용어는 상기 단편이 항원에 대한 결합 친화성을 갖는 것 이면, 이들 분자에 제한되지 않는다.

#### 1. 외래 유전자 발현 증강에 사용되는 DNA 요소

[0060] 실시예 1 에서 나타낸 바와 같이, 본 발명에 따른 DNA 요소는 아세틸화 히스톤 H3 과 게놈 DNA 사이의 상호작용을 사용하여 수득될 수 있다. 일반적으로, 히스톤 (H3 및 H4) 의 아세틸화는 전사 활성화와 연관되는 것으로 여겨지며, 2 개의 주요 가설이 지지되고 있다. 한 가설은, 히스톤 말단부 (tail) 가 아세틸화됨으로써 이것이 전기적으로 중화되어 DNA-히스톤 상호작용의 약화를 초래하는 방식으로 히스톤의 아세틸화가 뉴클레오솜 형태에서의 변화와 연관된다는 것이다 (Mellor J. (2006) Dynamic nucleosomes and gene transcription. Trends Genet. 22(6) : 320-329). 다른 가설은 히스톤의 아세틸화가 다양한 전사 인자의 리크루트먼트 (recruitment) 와 연관된다는 것이다 (Nakatani Y. (2001) Histone acetylases - versatile players. Genes Cells. 6 (2) : 79-86). 두 가설 중 어느 쪽의 가설에서나, 히스톤의 아세틸화가 전사 활성화와 연관되며, 항-아세틸화 히스톤 H3 항체를 사용하여 염색질 면역침전법 (ChIP) 을 수행함으로써, 아세틸화 히스톤 H3 과 상호작용하는 DNA 요소를 농축시킬 수 있다는 높은 가능성이 존재한다.

[0061] 본 발명에서, A2 는 외래 유전자 발현 증강에 사용되는 DNA 요소의 예이다. A2 는 인간 염색체 15 번의 80966429 ~ 80974878 부위에 위치하며 8450 bp 의 폴리뉴클레오티드 서열이고, AT 함량이 62.2% 이다. A2 의 폴리뉴클레오티드 서열은 서열 목록에서 SEQ ID NO: 1 로 표기된다.

[0062] A7, A18, B5 및 C14 는 유사한 DNA 요소의 예이다. A7 은 인간 염색체 11 번의 88992123 ~ 89000542 부위에 위치하며 8420 bp 의 폴리뉴클레오티드 서열이고, AT 함량이 64.52% 이다. A7 의 폴리뉴클레오티드 서열은 서열 목록에서 SEQ ID NO: 2 로 표기된다.

[0063] A18 은 인간 염색체 4 번의 111275976 ~ 111284450 부위에 위치하며 8475 bp 의 폴리뉴클레오티드 서열이고, AT 함량이 62.54% 이다. A18 의 폴리뉴클레오티드 서열은 서열 목록에서 SEQ ID NO: 3 으로 표기된다.

[0064] B5 는 인간 염색체 1 번의 143034684 ~ 143043084 부위에 위치하며 8401 bp 의 폴리뉴클레오티드 서열이고, AT 함량이 66.37% 이다. B5 의 폴리뉴클레오티드 서열은 서열 목록에서 SEQ ID NO: 4 로 표기된다.

- [0065] 마지막으로, C14 는 인간 염색체 11 번의 46089056 ~ 46097482 부위에 위치하며 8427 bp 의 폴리뉴클레오티드 서열이고, AT 함량이 63.81% 이다. C14 의 폴리뉴클레오티드 서열은 서열 목록에서 SEQ ID NO: 5 로 표기된다.
- [0066] 본 발명에서, DNA 요소의 외래 유전자 발현 증강 활성을 지표로서 SEAP 와 같은 리포터 유전자에 의해 인코딩된 단백질의 활성을 사용하여 측정될 수 있다. DNA 요소의 존재 하 리포터 단백질의 활성이 DNA 요소가 부재하는 경우에 비해 바람직하게는 2 배 이상, 보다 바람직하게는 4 배 이상, 보다 더 바람직하게는 5 배 이상 증가하는 경우, DNA 요소는 외래 유전자 발현 증강 활성을 갖는 것으로 측정될 수 있다. 심지어 활성이 2 배 이상 증가하는 경우에도, 이것이 세포 배양 규모 및 세포 배양 시간을 감소시켜, 그 결과로 수율을 증가시키고 세포 배양 비용을 감소시킬 수 있다는 것이 예측된다. 수율이 증가하는 경우, 의약으로서 사용되는 외래 단백질을 안정적으로 공급할 수 있다. 또한, 세포 배양 비용이 감소되는 경우, 의약으로서 사용되는 외래 단백질에 대한 비용이 감소하고, 외래 단백질을 투여받을 환자에 대한 재정상의 부담이 또한 감소한다.
- [0067] 본 발명에서, 상기 DNA 요소 중 임의 하나는 단독으로 사용될 수 있으며, DNA 요소 1 개 유형의 2 카페 이상이 사용될 수 있다. 대안적으로는, 상기 DNA 요소의 2 개 이상의 상이한 유형이 조합으로 사용될 수 있다.
- [0068] A2, A7, A18, B5 및 C14 는 본 발명에서 사용되는 DNA 요소의 바람직한 예이다.
- [0069] 본 발명에서 사용되는 DNA 요소는 SEQ ID NOS: 1 내지 5 로 표기되는 임의의 폴리뉴클레오티드 서열에 대해 80% 이상의 상동성을 갖는 폴리뉴클레오티드 서열을 포함하며 외래 유전자 발현 증강 활성을 갖는 폴리뉴클레오티드 서열일 수 있다. 80% 이상의 상동성은 바람직하게는 90% 이상의 상동성, 보다 바람직하게는 95% 이상의 상동성, 가장 바람직하게는 99% 이상의 상동성이다. 폴리뉴클레오티드 서열 상동성 탐색은 예를 들어, FASTA 또는 BLAST 와 같은 프로그램을 사용하여 일본 DNA 데이터뱅크 등에서 수행될 수 있다.
- [0070] 본 발명에서 사용되는 DNA 요소는 염격한 조건 하에 SEQ ID NOS: 1 내지 5 로 표기되는 폴리뉴클레오티드 서열로 이루어지는 군에서 선택되는 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드에 상보적인 폴리뉴클레오티드 서열로 이루어지는 폴리뉴클레오티드에 하이브리드화되며 외래 유전자 발현 증강 활성을 갖는 DNA 요소일 수 있다.
- [0071] 본원에서 사용되는 바와 같은 용어 "엄격한 조건" 은 소위 특이적인 하이브리드가 형성되지만 비 (non)-특이적인 하이브리드는 형성되지 않는 조건을 지칭한다. 예를 들어, SEQ ID NOS: 1 내지 5 로 표기되는 폴리뉴클레오티드 서열로 이루어지는 군에서 선택되는 폴리뉴클레오티드 서열에 대해 높은 상동성을 갖는 폴리뉴클레오티드 서열, 즉 80% 이상, 바람직하게는 90% 이상, 보다 바람직하게는 95% 이상, 가장 바람직하게는 99% 이상의 상동성을 갖는 폴리뉴클레오티드 서열로 이루어지는 핵산의 상보적 가닥이 하이브리드화하고, 낮은 상동성을 갖는 폴리뉴클레오티드 서열을 포함하는 핵산의 상보적 가닥이 하이브리드화하지 않는 조건이, 예시적인 염격한 조건이다. 보다 구체적으로, 나트륨 염의 농도가 15 내지 750 mM, 바람직하게는 50 내지 750 mM, 보다 바람직하게는 300 내지 750 mM 이고, 온도가 25 내지 70°C, 바람직하게는 50 내지 70°C, 보다 바람직하게는 55 내지 65°C 이고, 포름아미드 농도가 0 내지 50%, 바람직하게는 20 내지 50%, 보다 바람직하게는 35 내지 45% 인 조건을 예로 들 수 있다. 또한, 염격한 조건으로서, 나트륨 염의 농도가 일반적으로 15 내지 600 mM, 바람직하게는 50 내지 600 mM, 보다 바람직하게는 300 내지 600 mM 이고 온도가 50 내지 70°C, 바람직하게는 55 내지 70°C, 보다 바람직하게는 60 내지 65°C 인, 하이브리드화 후 필터 세척 조건을 예로 들 수 있다.
- [0072] 당업자는 Molecular Cloning (Sambrook, J. et al., Molecular Cloning: a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory Press, 10 Skyline Drive Plainview, NY (1989)) 등을 참조로 하여 이러한 상동 유전자를 용이하게 수득할 수 있다. 또한, 상기 언급한 폴리뉴클레오티드 서열의 상동성은 FASTA 탐색 또는 BLAST 탐색에 의해 동일한 방식으로 측정될 수 있다.
- [0073] 상기 언급한 폴리뉴클레오티드 서열 내로의 돌연변이 (결실, 치환 및/또는 삽입) 도입은 이러한 기술 분야에서 알려져 있는 방법, 예컨대 쿤켈 (Kunkel) 법 또는 캡이 있는 이중체 (gapped duplex)법에 의해, 또는 이러한 방법을 기초로 하여 수행될 수 있다. 예를 들어, 위치-지정 돌연변이유발 방법을 이용하는 돌연변이 도입 키트 (예를 들어, Mutant-K(TaKaRa Bio, Inc. 사제), Mutant-G(TaKaRa Bio, Inc. 사제), 또는 LA PCR 시험관내 돌연변이유발 시리즈 키트 (TaKaRa Bio, Inc. 사제)) 등이 사용될 수 있다. 이러한 돌연변이화된 폴리뉴클레오티드가 또한 본 발명의 DNA 요소로서 사용될 수 있다.
- [0074] 본 발명의 DNA 요소로서, 서열 목록에서 SEQ ID NOS: 1 내지 5 중 임의 하나로 표기되는 폴리뉴클레오티드 서열의 3000 개 이상 또는 2000 개 이상의 연속 뉴클레오티드를 포함하는 부분적 단편이 사용될 수 있다. 이러한

한 부분적 단편의 예는 다음을 포함한다: A2 의 부분적 단편인 A2-1 내지 A2-17; A7 의 부분적 단편인 A7-1 내지 A7-18; A18 의 부분적 단편인 A18-1 내지 A18-4; B5 의 부분적 단편인 B5-1 내지 B5-6; 및 C14 의 부분적 단편인 C14-1 내지 C14-14. 그러나, DNA 요소는 외래 유전자 발현 증강 활성을 갖는 것이면, 이들 부분적 단편에 제한되지 않는다.

[0075] 본 발명에서, 상기 부분적 단편 중 임의 하나는 단독으로 사용될 수 있으며, 부분적 단편 1 개 유형의 2 카피 이상이 또한 사용될 수 있다. 대안적으로, 부분적 단편의 2 개 이상의 상이한 유형이 조합으로 사용될 수 있다. 또한, 임의의 상기 언급한 DNA 요소의 전체 길이 서열 및 부분적 단편이 조합으로 사용될 수 있다.

상기 조합에서, 전체 길이 서열 및 부분적 단편은 동일한 DNA 요소 또는 상이한 DNA 요소에서 유래할 수 있다.

[0076] A2 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, A2-1 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 1 ~ 3000 의 폴리뉴클레오티드 서열에 상응하고; A2-2 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2801 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-3 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 5401 ~ 8450 의 폴리뉴클레오티드 서열에 상응하고; A2-4 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 2700 의 폴리뉴클레오티드 서열에 상응하고; A2-5 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 2200 의 폴리뉴클레오티드 서열에 상응하고; A2-6 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 3700 의 폴리뉴클레오티드 서열에 상응하고; A2-7 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 5000 의 폴리뉴클레오티드 서열에 상응하고; A2-8 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 4001 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-9 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 1 ~ 3700 의 폴리뉴클레오티드 서열에 상응하고; A2-10 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-11 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2801 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-12 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-13 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-14 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2801 ~ 8450 의 폴리뉴클레오티드 서열에 상응하고; A2-15 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 1 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-16 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-17 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 8450 의 폴리뉴클레오티드 서열에 상응한다.

[0077] A7 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, A7-1 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 601 ~ 3600 의 폴리뉴클레오티드 서열에 상응하고; A7-2 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3601 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-3 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 5401 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-4 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3401 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-5 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 4500 의 폴리뉴클레오티드 서열에 상응하고; A7-6 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 4401 ~ 7400 의 폴리뉴클레오티드 서열에 상응하고; A7-7 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 2401 ~ 5400 의 폴리뉴클레오티드 서열에 상응하고; A7-8 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1 ~ 3600 의 폴리뉴클레오티드 서열에 상응하고; A7-9 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 5400 의 폴리뉴클레오티드 서열에 상응하고; A7-10 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 2401 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-11 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3401 ~ 7400 의 폴리뉴클레오티드 서열에 상응하고; A7-12 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 4401 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-13 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1 ~ 5400 의 폴리뉴클레오티드 서열에 상응하고; A7-14 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-15 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 2401 ~ 7400 의 폴리뉴클레오티드 서열에 상응하고; A7-16 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3401 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-17 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-18 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 7400 의 폴리뉴클레오티드 서열에 상응한다.

[0078] A18 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, A18-1 은 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 1 ~ 5040 의 폴리뉴클레오티드 서열에 상응하고; A18-2 는 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 1001 ~ 6002 의 폴리뉴클레오티드 서열에 상응하고; A18-3 은 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 2001 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A18-4 는 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 3000 ~ 7000 의 폴리뉴클레오티드 서열에 상응한다.

- [0079] A2, A7 및 A18 의 각 단편의 출발점 및 종료점을 또한 도 18 에서 나타낸다.
- [0080] B5 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, B5-1 은 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 1 ~ 4001 의 폴리뉴클레오티드 서열에 상응하고; B5-2 는 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 1 ~ 3200 의 폴리뉴클레오티드 서열에 상응하고; B5-3 은 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 2491 ~ 5601 의 폴리뉴클레오티드 서열에 상응하고; B5-4 는 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 5373 ~ 8401 의 폴리뉴클레오티드 서열에 상응하고; B5-5 는 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 901 ~ 4001 의 폴리뉴클레오티드 서열에 상응하고; B5-6 은 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 4001 ~ 7000 의 폴리뉴클레오티드 서열에 상응한다.
- [0081] C14 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, C14-1 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 4015 의 폴리뉴클레오티드 서열에 상응하고; C14-2 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1987 ~ 5014 의 폴리뉴클레오티드 서열에 상응하고; C14-3 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 4020 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-4 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-5 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 6011 의 폴리뉴클레오티드 서열에 상응하고; C14-6 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 4939 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-7 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 5014 의 폴리뉴클레오티드 서열에 상응하고; C14-8 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 2994 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-9 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 4020 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-10 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1 ~ 5014 의 폴리뉴클레오티드 서열에 상응하고; C14-11 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1987 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-12 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 2994 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-13 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-14 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1987 ~ 8141 의 폴리뉴클레오티드 서열에 상응한다.
- [0082] B5 및 C14 의 각 단편의 출발점 및 종료점을 또한 도 19 에서 나타낸다.
- [0083] 2. 폴리뉴클레오티드의 취득
- [0084] 본 발명에서, 이후 기재될, 제조를 증가시킬 외래 단백질을 인코딩하는 외래 유전자를 포함하는 폴리뉴클레오티드는 하기 기재된 바와 같은 통상적인 절차에 의해 수득될 수 있다. 예를 들어, 이러한 폴리뉴클레오티드는 외래 유전자의 단편을 기준으로 하여 합성된 DNA 프로브를 사용하여 외래 유전자 발현 세포 또는 조직에서 유래한 cDNA 라이브러리를 스크리닝함으로써 단리될 수 있다. mRNA 는 이러한 기술 분야에서 통상 사용되는 방법에 의해 제조될 수 있다. 예를 들어, 세포 또는 조직을 구아닌 시약, 페놀 시약 등으로 처리하여 총 RNA 를 수득한 후, 올리고(dT) 셀룰로오스 컬럼 또는 폴리 U-세파로오스 컬럼 (담체로서 세파로오스 2B 함유) 등을 사용하는 친화성 컬럼 방법에 의해, 또는 배치 (batch) 방법에 의해 폴리 (A) + RNA (mRNA) 를 수득한다. 또한, 폴리 (A) + RNA 를 수크로오스 밀도-구배 원심분리 등에 의해 추가로 분해할 수 있다. 이후, 그에 따라 수득한 mRNA 를 주형으로 사용하고, 올리고 dT 프라이머 및 역방향 전사효소를 또한 사용하여 단일 가닥 cDNA 를 합성한다. 그에 따라 수득한 단일 가닥 cDNA 로부터, DNA 중합효소 I, DNA 리가아제, RNase H 등 을 사용하여 이중 가닥 cDNA 를 합성한다. 이에 따라 합성된 이중 가닥 cDNA 를 T4 DNA 중합효소를 사용하여 평활화 (blunted) 한 후, 어댑터 (예컨대 EcoRI 어댑터) 에 라이게이션하고, 인산화 등을 수행하고, 생성된 DNA 를 λgt11 과 같은 람다 파지 내로 혼입하여 생체내 (in vivo) 패키징을 달성함으로써, cDNA 라이브러리를 제조할 수 있다. 또한, 람다 파지 외의 플라스미드 벡터를 사용하여 cDNA 라이브러리를 제조할 수 있다. 그 후에, 표적 DNA 를 포함하는 클론 (양성 클론) 을 cDNA 라이브러리에서 선택할 수 있다.
- [0085] 외래 유전자 함유 단백질 또는 폴리뉴클레오티드 제조를 증가시키기 위해 사용되는 상기 언급한 DNA 요소가 통상적인 절차 (Molecular Cloning (1989), Methods in Enzymology 194 (1991)) 에 따라 게놈 DNA 에서 단리되거나, 프로모터 및 터미네이터 부위 함유 폴리뉴클레오티드가 게놈 DNA 에서 단리되는 경우, 게놈 DNA 는 수집원으로서 사용되는 유기체의 세포주로부터 추출되고, 폴리뉴클레오티드가 선택되고 단리된다. 게놈 DNA 의 추출은 예를 들어, Cryer 등의 방법 (Methods in Cell Biology, 12, 39-44 (1975)) 또는 P. Philippson 등의 방법 (Methods Enzymol., 194, 169-182 (1991)) 에 따라 수행될 수 있다.
- [0086] 외래 유전자 함유 표적 DNA 요소 또는 폴리뉴클레오티드는 또한 예를 들어, PCR 방법에 의해 수득될 수 있다 (PCR Technology. Henry A. Erlich, Astockton press (1989)). PCR 방법을 사용하는 폴리뉴클레오티드 증폭

에서, 20- 내지 30-mer 합성 단일 가닥 DNA 를 프라이머로서 사용하고, 계놈 DNA 를 주형으로서 사용한다. 유전자의 폴리뉴클레오티드 서열을 확인한 후, 증폭된 유전자를 사용한다. PCR 용 주형으로서, 계놈 DNA 라이브러리 예컨대 박테리아 인공 염색체 (BAC) 를 사용할 수 있다.

[0087] 한편, 서열이 알려져 있지 않은 외래 유전자 함유 폴리뉴클레오티드는 (a) 통상의 절차에 따라 유전자 라이브러리를 제조하고, (b) 제조된 유전자 라이브러리로부터 원하는 폴리뉴클레오티드를 선택하고, 상기 폴리뉴클레오티드를 증폭하여 수득될 수 있다. 유전자 라이브러리는 염색체 DNA 를 분해하기에 적절한 제한 효소를 사용하여, 수집원으로서 사용되는 유기체의 세포주로부터 통상적인 절차에 의해 수득한 염색체 DNA 를 부분적으로 소화시키고, 수득한 단편을 적절한 벡터에 라이게이션하고, 상기 벡터를 적절한 숙주 내로 도입하여 제조될 수 있다. 유전자 라이브러리는 또한 세포로부터 mRNA 를 추출하고, mRNA 로부터 cDNA 를 합성하고, cDNA 를 적절한 벡터에 라이게이션하고, 상기 벡터를 적절한 숙주 내로 도입하여 제조될 수 있다. 이러한 제조에서 사용되는 벡터로서, 유전자 라이브러리 제조용 벡터, 파지 벡터, 코스미드 등으로 일반적으로 알려져 있는 플라스미드가 또한 사용될 수 있다. 형질전환 또는 트랜스펙션 숙주로서, 상기 언급한 벡터의 유형에 대해 적합한 숙주가 사용될 수 있다. 외래 유전자 함유 폴리뉴클레오티드는 외래 유전자에 대해 특이적인 서열을 함유하는 표지된 프로브를 사용하여 콜로니 하이브리드화 방법, 플라크 하이브리드화 방법 등에 의해 상기 언급한 유전자 라이브러리에서 선택된다.

[0088] 또한, 외래 유전자 함유 폴리뉴클레오티드는 전 화학 합성 (total chemical synthesis) 에 의해 제조될 수도 있다. 예를 들어, 유전자는 두 쌍의 상보적 올리고뉴클레오티드가 제조되고 어닐링되는 방법, 여러 어닐링된 DNA 가닥이 DNA 리가아제에 의해 라이게이션되는 방법, 여러 부분 상보적 폴리뉴클레오티드가 제조되고 갭 (gap) 이 PCR 에 의해 채워지는 방법 등에 의해 합성될 수 있다.

[0089] 폴리뉴클레오티드 서열 결정은 종래의 기술, 예를 들어 디데옥시 방법 (Sanger et al., Proc. Natl. Acad. Sci., USA, 74, 5463-5467, (1977)) 등에 의해 수행될 수 있다. 또한, 상기 폴리뉴클레오티드 서열 결정은 시판되는 서열분석 키트 등을 사용하여 용이하게 수행될 수 있다.

### 3. 외래 유전자 발현 벡터, 요소 벡터

[0091] 본 발명의 외래 유전자 발현 벡터로서, 상기 언급한 DNA 요소 1 개 유형, 상기 언급한 DNA 요소 1 개 유형의 2 카피 이상, 또는 상기 언급한 DNA 요소의 2 개 이상의 상이한 유형을 조합으로 함유하고, 외래 유전자 발현 단위체를 추가로 함유하는 벡터가 제공된다. 외래 유전자가 상기 언급한 외래 유전자 발현 벡터를 사용하여 숙주 세포에서 발현되는 경우, DNA 요소는 유전자 발현 단위체의 바로 업스트림 또는 바로 다운스트림에 위치할 수 있거나, 유전자 발현 단위체에서 멀리 떨어진 장소에 위치할 수 있다. 또한, 다수의 상기 DNA 요소를 함유하는 하나의 외래 유전자 발현 벡터가 사용될 수 있다. 더욱이, DNA 요소는 유전자 발현 단위체에 대해 정방향 또는 역방향으로 삽입될 수 있다.

[0092] 또한, 본 발명에서 사용될 벡터로서, 상기 언급한 DNA 요소 1 개 유형, 상기 언급한 DNA 요소 1 개 유형의 2 카피 이상, 또는 상기 언급한 DNA 요소의 2 개 이상의 상이한 유형을 조합으로 함유하고, 유전자 발현 단위체 (이하, "요소 벡터"로서 또한 지칭함) 를 함유하지 않는 벡터가 또한 포함된다. 이러한 요소 벡터는 DNA 요소를 함유하는 상기 언급한 외래 유전자 발현 벡터 또는 DNA 요소를 함유하지 않으며 외래 유전자 발현 단위체만을 함유하는 외래 유전자 발현 벡터와 조합으로 사용될 수 있다. 요소 벡터를 외래 유전자 발현 벡터와 공존시킴으로써 외래 유전자 발현 벡터가 단독으로 사용되는 경우와 비교하여 외래 유전자의 발현이 증강되므로, 상기 언급한 벡터의 조합이 또한 본 발명의 외래 유전자 발현 벡터에 포함된다.

[0093] 외래 단백질 인코딩 유전자는 특히 제한되지는 않으나, 이의 예는 리포터 유전자 예컨대 분비 알칼리 포스파타아제 (SEAP), 녹색 형광 단백질 (GFP) 및 루시퍼라아제; 다양한 효소 유전자 예컨대  $\alpha$ -아밀라아제 유전자 및  $\alpha$ -갈락토시다아제 유전자; 약학적으로 유용하고 생리적으로 활성인 단백질인 다양한 인터페론 예컨대 인터페론  $\alpha$  및 인터페론  $\gamma$  의 유전자; 다양한 인터류킨 예컨대 IL-1 및 IL-2 의 유전자; 다양한 사이토카인 유전자 예컨대 에리트로포이에틴 (EPO) 유전자 및 과립구 콜로니-자극 인자 (G-CSF) 유전자; 성장 인자 유전자; 및 항체 유전자를 포함한다. 이를 유전자는 임의의 방법으로 수득될 수 있다.

[0094] 본 발명은 매우 소수성인 단백질 및 복합체 형성으로 인해 분비되고 제조되기 어려운 단백질에 대해 특히 유효하다. 따라서, 상기 언급한 외래 단백질은 다양체 단백질 예컨대 항체 또는 이의 기능적 단편인 이종다량체를 포함한다. "항체의 기능적 단편"은 항원-결합 활성을 갖는 항체의 부분적 단편을 지칭하며 Fab, F(ab')<sub>2</sub>, Fv, scFv, 디아바디, 선형 항체, 다중특이적 항체 (항체 단편으로부터 형성됨) 등을 포함한다. 항

체의 기능적 단편은 또한 환원 조건 하에  $F(ab')_2$  를 처리하여 수득한 항체의 가변부에서의 1 가 단편인  $Fab'$  를 포함한다. 그러나 기능적 단편은, 단편이 항원에 대한 결합 친화성을 갖는 것이면, 이들 분자에 제한되지 않는다. 또한, 이들 기능적 단편은 항체 단백질의 전체 길이 분자를 적절한 효소로 처리하여 수득한 단편 뿐 아니라 유전자 조작 항체 유전자를 사용하여 적절한 숙주 세포에서 생성된 단백질을 포함한다.

[0095]

유전자 발현 단위체는 전사의 판독 프레임 방향으로, 적어도 프로모터 부위, 외래 유전자 및 전사 터미네이터 부위 (폴리 (A) 부가 신호) 를 갖는다. 여기서 사용할 수 있는 프로모터는 항시성 (constitutive) 발현 프로모터 또는 유도성 (inducible) 발현 프로모터일 수 있다. 항시성 발현 프로모터의 예는 다양한 자연적 프로모터 예컨대 SV40 초기 프로모터, 아데노바이러스 E1A 프로모터, CMV (사이토메갈로바이러스) 프로모터, EF-1 $\alpha$  (인간 신장 인자-1 $\alpha$ ) 프로모터, HSP70 프로모터, MT 프로모터, RSV 프로모터, UBC 프로모터 및 액틴 프로모터; 및 인공 (융합) 프로모터 예컨대 SR $\alpha$  프로모터 및 CAG 프로모터를 포함한다. 또한, 폴리 (A) 부가 서열은 프로모터로부터의 전사에 대한 전사 종결을 유도하는 활성을 갖는 서열일 수 있으며, 프로모터와 동일하거나 상이한 유전자로부터의 서열일 수 있다.

[0096]

외래 단백질 제조를 증가시키기 위해서 강력한 프로모터를 사용할 필요가 있다. 그러나, 접힘을 갖기 어려운 단백질 또는 매우 활성인 프로모터를 사용하여 분비되기 어려운 단백질을 제조하고자 하는 경우, 단백질은 그 대신에 분비되지 못할 수 있다. 이는 단백질이 번역이 수행되는 리보솜 및 접힘과 분비가 수행되는 소포체의 수용력을 초과하는 양으로 제조되는 경우, 과다 제조된 단백질이 변성되고, 축적되고, 세포 내에 산재된 후 프로테오좀에 의해 분해되기 때문이다. 따라서, 생성된 단백질이 변성 또는 응집되지 않거나 생성된 단백질의 양이 분비 수용력을 초과하지 않는 정도로 발현 수준을 유지할 수 있는 프로모터가 적절히 선택되는 것이 바람직하다. 대안적으로는, 프로모터는 프로모터의 활성을 조정 (예를 들어, 감소시킴) 하여 사용된다. 다량체 단백질 중에서, 이종다량체 형성 분자, 및 특히 이종사량체인 항체와 같은 분자가 상기 언급한 효과에 민감하다. 항체는 서로 연관되는 2 개의 중쇄 분자 및 2 개의 경쇄 분자를 가지므로, 분자를 적절히 연관시키기 위해서는 이의 발현 수준이 중요한 인자가 된다.

[0097]

또한, 본 발명의 외래 유전자 발현 벡터 및 요소 벡터는 각각 형질전환체를 선별하기 위한 선별 마커를 함유할 수 있다. 예를 들어, 약물 예컨대 세룰레닌 (cerulenin), 오레오바시딘 (aureobasidin), 제오신 (Zeocin), 카나바닌 (canavanine), 시클로헥시미드 (cycloheximide), 히드로마이신 (hygromycin), 퓨로마이신 (puromycin), 블라스티시딘 (blasticidin), 테트라사이클린 (tetracycline), 카나마이신 (kanamycin), 암페실린 (ampicillin) 또는 네오마이신 (neomycin) 에 대한 내성을 부여하는 약물-내성 마커를 사용하여, 형질전환체가 선별될 수 있다. 또한, 에탄올과 같은 용매에 대한 내성, 글리세롤, 염 등의 삼투압에 대한 내성, 구리 이온 등과 같은 금속 이온에 대한 내성을 부여하는 유전자가 마커로서 사용되는 경우, 형질전환체가 또한 선별될 수 있다.

[0098]

본 발명의 외래 유전자 발현 벡터 및 요소 벡터는 각각 염색체 DNA 내로 혼입되지 않는 벡터일 수 있다. 일 반적으로, 외래 유전자 발현 벡터는 숙주 세포에 트랜스펙션된 후, 염색체 내로 무작위하게 혼입된다. 그러나 포유류 바이러스 예컨대 유인원 바이러스 40 (SV40), 파필로마바이러스 (papillomavirus) (BPV, HPV) 또는 EBV 에서 유래한 구성 성분을 사용하여, 벡터는 트랜스펙션된 숙주 세포 내에서 자가 복제 가능한 에피솜 벡터로서 사용될 수 있다. 예를 들어, SV40-유래 복제 기점 (oriP) 및 트랜스-작용 (trans-acting) 인자인 SV40 거대 T 항원 인코딩 서열을 함유하는 벡터, EBV-유래 oriP 및 EBNA-1 인코딩 서열을 함유하는 벡터 등이 사용될 수 있다. DNA 요소의 효과는 벡터 유형 또는 염색체 내로의 이의 혼입의 존재 또는 부재에 관계없이 외래 유전자 발현 증강 활성에 의해 표시될 수 있다.

[0099]

#### 4. 형질전환 세포

[0100]

본 발명의 형질전환 세포는 상기 "1" 에서 기재된 DNA 요소를 함유하는 상기 "3" 에서 기재된 외래 유전자 발현 벡터가 도입된 형질전환 세포이다. 외래 유전자 발현 벡터로서, DNA 요소를 함유하는 외래 유전자 발현 벡터만이 도입될 수 있거나 (A), DNA 요소를 함유하는 외래 유전자 발현 벡터 및 또한 상기 "3" 에서 기재된 요소 벡터가 조합으로 도입될 수 있다 (B). 대안적으로, DNA 요소를 함유하지 않는 외래 유전자 발현 벡터 및 요소 벡터가 조합으로 도입될 수 있다 (C).

[0101]

(B) 또는 (C) 의 상기 조합을 사용하는 숙주 세포 내 외래 유전자 발현은 예를 들어, Girod 등의 방법 (Biotechnology and Bioengineering, 91, 2-11 (2005)) 및 Otte 등의 방법 (Biotechnol. Prog., 2007, 23, 801-807 (2007)) 에 따라 수행될 수 있다.

- [0102] 형질전환되는 숙주 세포의 예는 진핵 세포를 포함하며, 이의 바람직한 예는 포유류 세포를 포함하고, 보다 바람직한 예는 인간, 마우스, 랙트, 햄스터, 원숭이 또는 소 유래의 세포를 포함한다. 이러한 포유류 세포의 예는 COS-1 세포, 293 세포 및 CHO 세포 (CHO-K1, DG44, CHO dhfr-, CHO-S) 를 포함하지만, 숙주 세포가 이에 제한되는 것은 아니다.
- [0103] 본 발명에서, 도입된 유전자가 숙주 세포에서 안정적으로 존재하고 충분히 발현되도록 하는 방법이면, 임의의 방법을 사용하여 발현 벡터를 숙주 세포 내로 도입할 수 있다. 일반적으로 사용되는 방법의 예는 인산칼슘법 (Ito et al., (1984) Agric. Biol. Chem., 48, 341), 전기천공법 (Becker, D. M. et al., 1990; Methods Enzymol., 194, 182-187), 스페로플라스트 (spheroplast) 법 (Creggh et al., Mol. Cell. Biol., 5, 3376 (1985)), 리튬 아세테이트법 (Itoh, H. (1983) J. Bacteriol. 153, 163-168) 및 리포펙션법을 포함한다.
- [0104] 5. 외래 단백질 제조 방법
- [0105] 본 발명에서 외래 단백질은, 알려져 있는 방법에 의해 상기 "3"에서 기재된 벡터를 사용하여 외래 단백질 인코딩 유전자가 도입된 상기 "4"에서 기재된 형질전환 세포를 배양하고, 생성된 배양 산물로부터 단백질을 수집한 후, 상기 단백질을 정제하여 제조될 수 있다. 본원에서 사용하는 바와 같은 용어 "배양 산물"은 배양 상청액 이외에 배양된 세포 또는 세포 균질액을 지칭한다. 더욱이, 상기 "4"에서 기재된 형질전환 세포를 사용하여 제조될 수 있는 외래 단백질로서, 단량체 단백질 뿐 아니라 다량체 단백질이 선택될 수 있다. 다수의 상이한 소단위체로 형성된 이종다량체 단백질이 제조되는 경우, 이를 소단위체를 인코딩하는 다수의 유전자를 상기 "4"에서 기재된 숙주 세포 내로 각각 도입하는 것이 필요하다.
- [0106] 형질전환 세포를 배양하는 방법은 종래의 숙주 세포 배양 방법에 따라 수행될 수 있다.
- [0107] 형질전환 세포가 포유류 세포인 경우, 세포는 예를 들어, 약 24 내지 1000 시간의 배양 시간 동안 37°C 및 5% 또는 8% CO<sub>2</sub> 의 조건 하에 배양된다. 배양은 고정, 진탕, 교반 또는 통기 조건 하 회분식 배양, 유가식 배양, 연속 배양 등을 통해 수행될 수 있다.
- [0108] 상기 언급한 배양 산물 (배양액)로부터의 외래 단백질 인코딩 유전자의 발현 산물 확인은 SDS-PAGE, 웨스턴 분석, ELISA 등에 의해 수행될 수 있다. 제조된 단백질을 단리하고 정제하기 위해서, 종래의 단백질 단리 및 정제 방법이 사용될 수 있다. 배양 완료 후, 표적 단백질이 세포에서 제조되는 경우, 상기 세포는 초음파 균질기, 프랜치 프레스 (French press), 맨톤-골린 (Manton-Gaulin) 균질기, 디노밀 (Dinomil) 등을 사용하여 균질화됨으로써, 표적 단백질이 수득된다. 또한, 표적 단백질이 세포 외부에서 제조되는 경우, 배양액은 그 자체로서 사용되거나, 세포가 원심분리 등에 의해 제거된다. 그 후, 표적 단백질은 유기 용매를 사용하여 추출 등에 의해 수집된 후, 수집된 표적 단백질은 다양한 크로마토그래피 기술 (소수성 크로마토그래피, 역상 크로마토그래피, 친화성 크로마토그래피, 이온 교환 크로마토그래피 등), 분자 체를 사용하는 겔 여과, 및 폴리아크릴아미드 겔을 사용하는 전기영동 등과 같은 기술을 단독으로 또는 필요에 따라 조합으로 사용하여 단리되고 정제될 수 있다.
- [0109] 상기 언급한 배양 방법 및 정제 방법은 단지 예시일 뿐이며, 방법이 이에 제한되는 것은 아니다. 정제된 유전자 산물의 아미노산 서열은 알려져 있는 아미노산 분석 기술, 예컨대 에드만 (Edman) 분해법을 사용하는 자동화 아미노산 서열 결정에 의해 확인될 수 있다.
- [0110] 6. 항체 단백질 제조 방법
- [0111] 상기 "5"에서 기재된 제조 방법을 사용하여 제조되는 이종다량체 단백질로서, 항체 단백질을 예로 들 수 있다. 항체 단백질은 중쇄 폴리펩티드 두 분자 및 경쇄 폴리펩티드 두 분자를 포함하는 사량체 단백질이다. 따라서, 항원-결합 친화성 유지 상태에서 이러한 항체 단백질을 수득하기 위해서, 중쇄 및 경쇄 유전자 모두를 상기 "4"에서 기재된 형질전환 세포 내로 도입하는 것이 필요하다. 이러한 경우, 중쇄 및 경쇄 유전자 발현 단위체는 동일한 발현 벡터 또는 상이한 발현 벡터에 존재할 수 있다.
- [0112] 본 발명에서 제조되는 항체로서, 토끼, 마우스 또는 랙트와 같은 실험 동물을 원하는 항원으로 면역화시켜 제조된 항체를 예로 들 수 있다. 또한, 출발 물질로서 상기 언급한 항체를 사용하여 수득한 키메라 항체 및 인간화 항체를 또한 본 발명에서 제조되는 항체로서 예로 들 수 있다. 또한, 유전자 조작 동물 또는 파지 디스플레이법을 사용하여 수득한 인간 항체가 또한 본 발명에서 제조되는 항체에 포함된다.
- [0113] 항체 제조에 사용되는 항체 유전자는 항체 유전자로부터 전사되고 번역되는 중쇄 폴리펩티드 및 경쇄 폴리펩티드의 조합이 주어진 항원 단백질에 대한 결합 활성을 갖는 것이면, 특정 폴리뉴클레오티드 서열을 갖는 항체 유

전자에 제한되지 않는다.

[0114] 또한, 항체 유전자가 항체의 전체 길이 분자를 인코딩할 필요가 없으며, 항체의 기능적 단편을 인코딩하는 유전자가 사용될 수 있다. 이러한 이의 기능적 단편 인코딩 유전자는 항체 단백질의 전체 길이 분자 인코딩 유전자를 유전자 조작하여 수득될 수 있다.

## 7. 그 외의 외래 단백질 제조 방법

[0116] 본 발명의 제조 방법을 사용하여 제조되는 외래 단백질의 예는, 상기 언급한 항체에 추가로, 인간 또는 비(non)-인간 유래의 다양한 단백질, 이의 기능적 단편 및 이의 변형된 산물을 포함한다. 이러한 단백질 등의 예는 웨პ티드 호르몬 예컨대 심방성 나트륨이뇨 웨პ티드 (ANP), 뇌 나트륨이뇨 웨პ티드 (BNP), C-형 나트륨이뇨 웨პ티드 (CNP), 바소프레신, 소마토스타틴, 성장 호르몬 (GH), 인슐린, 옥시토신, 그렐린, 랩틴, 아디포넥틴, 레닌, 칼시토닌, 오스테오프로테이린 및 인슐린-유사 성장 인자 (IGF); 사이토카인 예컨대 인터류킨, 캐모카인, 인터페론, 중앙 괴사 인자 (예컨대 TNF- $\alpha$ , TNF- $\beta$  및 TNF 슈퍼 패밀리), 신경 성장 인자 (예컨대 NGF), 세포 성장 인자 (예컨대 EGF, FGF, PDGF, HGF 및 TGF), 조혈 성장 인자 (예컨대 CSF, G-CSF 및 에리트로포이에틴) 및 아디포킨; 수용체 예컨대 TNF 수용체; 효소 예컨대 리소자임, 프로테아제, 프로테이나아제 및 웨პ티다아제; 이의 기능적 단편 (본래 단백질의 생물학적 활성의 일부 또는 전부를 갖는 단편), 및 임의의 이들 단백질을 포함하는 융합 단백질을 포함한다. 그러나, 상기 단백질이 이에 제한되는 것은 아니다.

## 실시예

[0118] 이하, 실시예를 참조로 하여 본 발명을 구체적으로 설명할 것이다. 그러나, 이들 실시예는 본 발명의 기술적 범주를 제한하지 않는다. 본 발명의 실시예에서 사용되는 플라스미드, 제한 효소, DNA 변형 효소 등은 시판 제품이며 통상적인 절차에 따라 사용될 수 있다. 또한, DNA 클로닝, 폴리뉴클레오티드 서열 결정, 숙주 세포 형질전환, 형질전환된 숙주 세포 배양, 생성된 배양 산물로부터의 단백질 수집, 단백질 정제 등에 사용된 절차는 또한 당업자에게 잘 알려져 있거나 문헌에서 발견될 수 있다.

## 실시예 1

### DNA 요소의 추출

#### (1-1) 항-아세틸화 히스톤 H3 항체를 사용하는 염색질 면역침전법

[0122] 항-아세틸화 히스톤 항체를 사용하는 ChIP 을, 하기 절차에 따라 EZ ChIP (Upstate) 을 사용하여 수행하였다. 더욱이, 다르게 언급하지 않는 한, 하기 절차에서 사용한 항체, 완충액 등으로서 Upstate 의 제품을 사용하였다.

[0123] 먼저, 293F 세포 (Invitrogen) 를 GIBCO (등록 상표) Freestyle™ 293 배지 (Invitrogen) 를 사용하여 37°C 및 8% CO<sub>2</sub> 의 조건 하에서 배양한 후, 원심분리 (1000 rpm, 5 분, 실온) 함으로써, 생장상에 있는 세포를 수집하였다. 2 x 10<sup>7</sup> 세포를 1% 포름알데히드 함유 배지에서 10 분 동안 고정한 후, 10x 글리신을 이에 추가한 후, 실온에서 5 분 동안 인큐베이션하였다. 원심분리 (3000 rpm, 5 분, 4°C) 후, 상청액을 제거하고, PBS 를 세포 펠렛에 추가하여 세포를 혼탁하였다. 그런 다음, 세포 혼탁액을 다시 원심분리하여 PBS 를 제거하고, 이후 SDS 용해 완충액을 세포 펠렛에 추가하여 세포를 혼탁하고 용해하였다. 세포 용해에 의해 수득한 각각의 샘플을, 샘플을 냉수로 냉각시키면서 초음파 균질기 (BRANSON) 를 사용하여 DNA 파쇄를 수행하고, 단백질 G-고정 아가로오스 및 프로테아제 저해제 칙테일 함유 희석 완충액을 이에 추가하였다. 생성된 혼합물을 4°C 에서 1 시간 동안 회전시킨 후, 원심분리하고, 이후 상청액을 수집하였다.

[0124] 이후, 10 μg 의 정상 토키 IgG 또는 α-아세틸 히스톤 H3 항체를 이에 추가한 후, 밤새 4°C 에서 회전시켰다. 생성된 용액에, 단백질 G-고정 아가로오스를 추가하고, 생성된 혼합물을 4°C 에서 1 시간 동안 회전시킨 후, 원심분리를 수행하고, 이후 펠렛을 수집하였다. 이에 따라 수득한 펠렛을 저염 면역 복합체 세척 완충액 (Low Salt Immune Complex Wash Buffer)으로 2 회, 고염 면역 복합체 세척 완충액 (High Salt Immune Complex Wash Buffer)으로 2 회, LiCl 면역 복합체 세척 완충액으로 2 회, 그리고 마지막으로 TE 완충액으로 4 회 세척하였다. 이후 용리 완충액 (20 μl 의 1 M 탄산수소나트륨, 10 μl 의 SDS 및 170 μl 의 멸균수 함유) 을 이에 추가하였다. 30 분 후, 혼합물을 원심분리하고, 상청액을 수집하였다.

[0125] 이후, 5 M 염화나트륨을 상청액에 추가하고, 생성된 혼합물을 밤새 65°C 에서 가열하였다. 이후 RNase A 를 이에 추가하고, 생성된 혼합물을 37°C 에서 30 분 동안 인큐베이션하였다. 그런 다음 0.5 M EDTA, 1 M

Tris-HCl 및 프로테이나아제 K 를 이에 추가하고, 생성된 혼합물을 45°C 에서 2 시간 동안 인큐베이션하였다.

[0126] 마지막으로, 시약 A, B 및 C 를 프로테이나아제 K 로의 처리에 의해 수득한 용액량의 5 배 초과의 양으로 이에 추가한 후, 스픈 필터를 사용하여 원심분리 (10000 rpm, 30 초, 실온) 함으로써, 염색질-면역침전 DNA 를 정제하였다.

[0127] (1-2) 마이크로어레이 분석

[0128] GenomePlex Complete Whole Genome Amplification (WGA) 키트 (Sigma) 를 사용하여, (1-1) 에서 수득한 각각의 ChIP 샘플을 증폭하였다. 절차는 키트에 수반되는 Sigma 의 프로토콜에 따라 수행하였다.

[0129] ChIP 를 확인하기 위해서, WGA 에 의해 증폭된 각각의 DNA 320 ng 을 주형으로서 사용하고 하기의 프라이머 및 SYBR (등록 상표) Premix Ex Taq™ (Perfect Real Time) (TAKARA) 을 또한 사용하여, 글리세르알데히드-3-포스 페이트 데히드로게나아제 (GAPDH) 내부 유전자를 PCR 방법 (95°C 에서 5 초, 그리고 60°C 에서 20 초 x 45 사이클) 에 의해 증폭하였다. 더욱이, GAPDH 는 DNA 요소가 ChIP 에 의해 풍부해졌는지 여부를 확인하기 위한 양성 대조군으로서 사용되는 하우스 키핑 유전자 (house keeping gene) 이며, PCR 방법을 EZ ChIP (Upstate) 에 부착된 프라이머를 사용하여 수행하였다.

5' -TACTAGCGGTTTTACGGGCG-3'

[0130] 5' -TCGAACAGGAGGAGCAGAGAGCGA-3'

[0131] 도 1 에서 나타낸 바와 같이, 항-아세틸화 히스톤 H3 항체로 면역침전 처리한 샘플에서 GAPDH 가 특이적으로 증폭된 것이 확인되었다. WGA 에 의해 증폭된 각각의 DNA 샘플을 마이크로어레이 분석 (NimbleGen) 하여 염색질 면역침전-온-칩 {Chromatin Immunoprecipitation-on-chip (칩-온-칩 (ChIP-on-chip))} 을 수행하였다. "칩-온-칩 (ChIP-on-chip)" 은 (1-1) 에서 풍부해진 각각의 DNA 를 마이크로어레이 분석하여 각각의 DNA 요소를 확인하기 위한 기술이다.

[0132] (1-3) DNA 요소의 추출

[0133] (1-2) 에서 수득한 칩-온-칩 (ChIP-on-chip) 분석 결과를 근거로 하여, AT 함량이 62% 이상인 5 개 서열을 추출하였다.

A2: 염색체 15 (80966429 ~ 80974878)

A7: 염색체 11 (88992123 ~ 89000542)

A18: 염색체 4 (111275976 ~ 111284450)

B5: 염색체 1 (143034684 ~ 143043084)

C14: 염색체 11 (46089056 ~ 46097482)

## 실시예 2

[0136] 지표로서 분비 알칼리 포스파타아제 (SEAP) 의 발현을 사용하는 DNA 요소의 효과

[0137] (2-1) SEAP 발현 벡터의 구축

[0138] 주형으로서 pSEAP2-대조군 (Clontech) 을 사용하여, SEAP 유전자를 하기의 프라이머 및 KOD-plus- (TOYOB) 를 사용하여 PCR 방법 (94°C 에서 30 초, 그리고 68°C 에서 2 분 x 40 사이클) 에 의해 증폭하였다.

5' -AAAGCTAGCATGCTGCTGCTGCTGCTGGGCC-3'

5' -AAAAGATCTTCATGTCTGCTCGAACGGCCGCCGC-3'

[0140] 이후, 증폭된 SEAP 단편을 겔 전기영동에 의해 분리하고 겔에서 절단한 후, QIAquick 겔 추출 키트 (Qiagen) 를 사용하여 정제하였다. 이에 따라 수득한 DNA 단편을 삽입물로서 사용하였다. 삽입물을 제한 효소 NheI 및 BgII 로 소화시키고, 벡터 pIREs hyg3 (Clontech) 을 제한 효소 NheI 및 BamHI 으로 소화시켰다. 생성된 DNA 단편을 각각 아가로오스 겔 전기영동 처리하여 표적 단편을 분리하고, 표적 단편을 겔에서 절단한 후, 정제하였다. 그런 다음, 라이케이션 반응 및 형질전환을 수행하였다. 라이케이션 반응을 LigaFast Rapid DNA Ligation System (Promega) 을 사용하여 수행하였다. 형질전환을 하기와 같이 수행하였다. 먼저, 동결된 컴피턴트 (competent) 세포 JM109 (TAKARA) 를 해동하고, 라이케이션 반응 후 수득한 10 μl 의 용

액을 해동 세포의 용액에 추가하고, 생성된 혼합물을 얼음 상에 30 분 동안 정치하였다. 그후, 열 충격 (42 °C, 45 초) 을 혼합물에 가하고, 혼합물을 얼음 상에서 5 분 동안 냉각하였다. 이러한 세포 혼탁액에, 1 ml 의 LB 배지를 추가하고, 생성된 혼합물을 37°C에서 1 시간 동안 진탕하였다. 이후, 혼합물을 0.1 mg/ml 암 피실린을 함유하는 LB 플레이트 상에 플레이팅하고, 상기 플레이트를 37°C에서 14 내지 16 시간 동안 인큐베이션하였다. 그후, 알칼리 용해에 의해, 표적 플라스미드를 LB 플레이트 상에서 배양된 콜로니로부터 수집하였다. 마지막으로, 알칼리 용해에 의해 수득한 플라스미드 내 SEAP 의 폴리뉴클레오티드 서열을 결정함으로써 pCMV/SEAP ires Hygro 를 구축하였다.

[0141] (2-2) DNA 요소의 클로닝

[0142] 이후, 실시예 1에서 추출한 각각의 DNA 요소를, 각각의 DNA 요소에 상응하는 폴리뉴클레오티드 서열을 함유하는 박테리아 인공 염색체 (BAC)로부터 BAC SUBCLONING 키트 (Gene Bridges)를 사용하여 (2-1)에서 수득한 SEAP 발현 벡터 내로 클로닝하였다.

[0143] 먼저, (2-1)에서 수득한 pCMV/SEAP ires Hygro 를 제한 효소 SpeI 으로 수 시간 동안 소화시킨 후, 에탄올 침전을 수행하고, 침전물을 멸균수에 용해하였다. SpeI 으로 소화시킨 벡터를 주형으로서 사용하여, 하기 프라이머 및 KOD-plus- (TOYOBIO) 를 사용하여 PCR 방법 (94°C에서 15 초, 55°C에서 30 초, 그리고 68°C에서 10 분 x 30 사이클) 을 수행하였다.

[0144] A2D:

5' -GGAAATTGAGAAGTATCATTCAACAGTACCACAAACATGAAATAATGTGGAT

CCTATTAATAGTAATCAATTACG-3'

A2R:

5' -CTCATTCTGTGGTTGTCATTCACTTCCCTTGATGCTATCCTTTCAAGCAAAATC

CTAGTCAAATAATCAATGTCAACG-3'

A7D:

5' - CTTATTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAACCTGGATC  
CTATTAATAGTAATCAATTACG - 3'

A7R :

5' - CTCTTCCCATTCTCATTTGAATCTACTTCAAAAGGTTTACCCATACTAAGACCTAG  
TCAATAATCAATGTCAACG - 3'

A18D :

5' - CGCCTGTAATCCCAGCACTTGGGAGGCTGAGGCAGGTGGATCACCTGAGGTCGA  
TCCTATTAATAGTAATCAATTACG - 3'

A18R :

5' - CATAACAGAACGCCAGTTGAAC TGAGACCTCACTCCATTCTACAAGTTATGCC  
TAGTCAATAATCAATGTCAACG - 3'

B5D :

5' - ACCGTTTATATTGTTAACGCATTCTAGACATATTGGCTACAAATCTAGATC  
CTATTAATAGTAATCAATTACG - 3'

B5R :

5' - GATCTTAGGGGGCTGATTATATAAAACAATAGAAATGTAGTCTTAGATGAAACC  
TAGTCAATAATCAATGTCAACG - 3'

C14D :

5' - CACAAAGTTCACTGTCAAGGCCAGGTGATGAGGCCACACATGCCGGACCTTGA  
TCCTATTAATAGTAATCAATTACG - 3'

C14R :

[0146] 5' - CAAAACCTCATCTCTACTGAAAATAGAAAATTAGCTGGCGTGGTGGCAGGTGCC

[0147] CTAGTCAATAATCAATGTCAACG - 3'

[0148] 반응 용액 일부를 사용하여 아가로오스 겔 전기영동에 의해 증폭을 확인한 후, 나머지 반응 용액을 에탄올 침전 처리하였다. 침전물을 멸균수에 용해하고, 생성된 용액을 형질전환용 DNA로서 사용하였다.

[0149] 이후, 형질전환용 대장균 (*Escherichia coli*)의 제조를 수행하였다.

[0150] 실시예 1에서 추출한 5 개 서열에 상응하는 BAC 클론은 하기와 같다:

추출 서열	상응 BAC 클론
A2	RP11-152F13
A7	RP11-643G5
A18	RP11-115A14
B5	RP11-640M9
C14	RP11-702F3

[0151] 해동한 10  $\mu\text{l}$  의 상기 언급한 BAC (Advanced GenoTechs Co.) 를 1  $\text{m}\text{l}$  의 배지 (15  $\mu\text{g}/\text{m}\text{l}$  의 최종 농도로 클로람페니콜 함유) 에 접종하고 밤새 37°C에서 인큐베이션하였다. 30  $\mu\text{l}$  의 배양액을 1.4  $\text{m}\text{l}$  의 배지 (15  $\mu\text{g}/\text{m}\text{l}$ 의 최종 농도로 클로람페니콜 함유)에 옮기고 37°C에서 2 시간 동안 인큐베이션하였다. 원심분리 및 멸균수로의 세척을 2 회 반복하고, 세포를 20  $\mu\text{l}$  의 멸균수에 혼탁하였다. 냉각된 큐벳 (0.1 cm)에, 1  $\mu\text{l}$ 의 pRED/ET (Gene Bridges) 및 대장균을 추가한 후, 전기천공 (1350 V, 10  $\mu\text{F}$ ) 하였다. 이후, 1  $\text{m}\text{l}$ 의 SOC 배지를 이에 추가하고, 생성된 혼합물을 30°C에서 70 분 동안 인큐베이션하였다. 100  $\mu\text{l}$ 의 배양액을

LB 플레이트 (각각 3  $\mu\text{g}/\text{ml}$  및 15  $\mu\text{g}/\text{ml}$  의 최종 농도로 테트라사이클린 및 클로람페니콜 함유) 상에 플레이팅하고, 밤새 30°C에서 인큐베이션하였다. 다음날, 이에 따라 수득한 각각의 콜로니를 1 ml의 배지 (각각 3  $\mu\text{g}/\text{ml}$  및 15  $\mu\text{g}/\text{ml}$ 의 최종 농도로 테트라사이클린 및 클로람페니콜 함유)에 접종하고, 밤새 30°C에서 인큐베이션하였다. 30  $\mu\text{l}$ 의 배양액을 1.4 ml의 배지 (각각 3  $\mu\text{g}/\text{ml}$  및 15  $\mu\text{g}/\text{ml}$ 의 최종 농도로 테트라사이클린 및 클로람페니콜 함유)에 옮기고, 30°C에서 2시간 동안 인큐베이션하였다. 이후, 50  $\mu\text{l}$ 의 10% L-아라비노오스를 이에 추가하고, 37°C에서 1시간 동안 인큐베이션을 추가 수행하였다. 그후, 멸균수로의 세척을 2회 반복하고, 30  $\mu\text{l}$ 의 멸균수에 혼탁한 대장균 및 1  $\mu\text{l}$ 의 형질전환용 DNA를 냉각한 큐벳 (0.1 cm)에 추가한 후, 전기천공 (1350 V, 10  $\mu\text{F}$ ) 하였다. 이후, 1 ml의 SOC 배지를 이에 추가하고, 생성된 혼합물을 37°C에서 90분 동안 인큐베이션하였다. 배양액의 총량을 LB 플레이트 (100  $\mu\text{g}/\text{ml}$  암피실린 함유) 상에 플레이팅하고, 상기 플레이트를 인큐베이션하였다. 그후, 표적 플라스미드를 알칼리 용해에 의해 수득하였다.

마지막으로, 수득한 플라스미드의 서열 및 이의 제한 효소 위치를 확인함으로써, 표적 플라스미드를 구축하였다. 벡터 구축물을 도 2에 나타낸다.

[0153] (2-3) 지표로서 SEAP 발현을 사용하는 평가

[0154] (2-2)에서 구축한 각각의 플라스미드를 숙주 세포 CHO-K1 (ATCC) 및 트랜스펙션 시약 Lipofectamine 2000 (Invitrogen)을 사용하여 평가하였다.

[0155] 800  $\mu\text{g}/\text{ml}$ 에서의 히그로마이신으로의 항생제 선별을 트랜스펙션 2일 후 시작하여 약 2주 동안 수행함으로써, 안정적으로 발현하는 다클론 세포주를 확립하였다. 이에 따라 확립된 세포주를 측정 전날에 배지 교체하고, 일정 세포수를 24-웰 플레이트 (IWAKI)에 시딩하였다. 세포 플레이팅 후 24시간에, 배양 상청액을 수집하고, SEAP의 활성을 측정하였다. 배양 상청액에서의 SEAP 활성을 SensoLyte™ pNPP 분비 알칼리 포스파타제 리포터 검정 (ANASPEC)을 사용하여 측정하였다.

[0156] 측정 결과를 도 3에 나타낸다. 요소를 갖지 않는 대조군의 SEAP 활성을 1로 정규화한 경우, DNA 요소 A2, A7, A18, B5 또는 C14를 갖는 안정적으로 발현하는 CHO 세포주의 배양 상청액에서의 SEAP 활성은 대조군보다 5배 이상 더 높은 수치 값을 나타내었다. 결과를 근거로 하여, 5개 유형의 DNA 요소 모두가 SEAP 발현을 급격히 증강시키는 것이 확인되었다. 더욱이, 상기 5개 유형의 DNA 요소의 폴리뉴클레오티드 서열을 각각 서열 목록에서 SEQ ID NOS: 1 내지 5로 표기한다.

### 실시예 3

[0158] 조합으로 사용되는 프로모터의 일반성

[0159] 실시예 2에서의 DNA 요소 평가에 사용한 벡터에 대한 프로모터는 CMV 프로모터였으므로, 다른 일반적인 프로모터와 조합된 DNA 요소의 사용을 실시예 3에서 연구하였다.

[0160] (3-1) EF-1 $\alpha$  및 SV40 프로모터를 사용하는 SEAP 발현 벡터의 구축

[0161] pSEAP2-대조군 (Clontech)을 주형으로서 사용하여, SEAP 유전자를 (2-1)에서 기재된 프라이머 및 KOD-plus-를 사용하여 PCR 방법 (94°C에서 30초, 그리고 68°C에서 2분 x 40 사이클)에 의해 증폭하였다. 증폭한 SEAP를 (2-1)에서와 동일한 방식으로 삽입물로서 제조하였다. 삽입물을 제한 효소 NheI 및 BglII로 소화시키고, 벡터 pIRES puro3(Clontech)을 제한 효소 NheI 및 BamHI으로 소화시키고, pCMV/SEAP ires Puro를 (2-1)에서와 동일한 방식으로 구축하였다.

[0162] 이후, pEF1/V5-His A (Invitrogen)를 주형으로서 사용하여, EF-1 $\alpha$  프로모터를 하기 프라이머 및 KOD-plus-를 사용하여 PCR 방법 (94°C에서 15초, 60°C에서 30초, 그리고 68°C에서 2분 x 30 사이클)에 의해 증폭하였다.

5' -AAAACCTAGTCAGAGAGGAATCTTGAGCTAATGGACC-3'

5' -AAAGATATCCCTAGCCAGCTTGGGTGGTACCAAGC-3'

[0164] 상기 구축한 pCMV/SEAP ires Puro를 벡터로서 사용하여, 제한 효소 SpeI 및 EcoRV로의 소화를 벡터 및 프로모터에 대해 수행하고, pEF/SEAP ires Puro를 (2-1)에서 기재된 방법에 따라 구축하였다.

[0165] 유사하게, pcDNA3.1+ (Invitrogen)을 주형으로서 사용하여, SV40 프로모터를 하기 프라이머 및 KOD-plus-를 사용하여 PCR 방법 (94°C에서 15초, 60°C에서 30초, 그리고 68°C에서 1분 x 30 사이클)에 의해 증폭하였다.

5' -AAAAACTAGTCTGTGGAATGTGTGTCAGTTAGGGTG-3'

[0166] 5' -AAAGATATCAGTTTGCAAAAGCCTAGGCCTC-3'

[0167] 상기 구축한 pCMV/SEAP ires Puro를 벡터로서 사용하여, 제한 효소 SpeI 및 EcoRV로의 소화를 벡터 및 프로모터에 대해 수행하고, pSV40/SEAP ires Puro를 (2-1)에서 기재된 방법에 따라 구축하였다.

[0168] (3-2) DNA 요소 A2 또는 A7의 클로닝

[0169] 이 후, DNA 요소 A2 또는 A7의 클로닝을, (3-1)에서 구축한 pEF/SEAP ires Puro 및 pSV40/SEAP ires Puro를 기본 구조로서 사용하여 수행하였다.

[0170] 먼저, pEF/SEAP ires Puro 및 pSV40/SEAP ires Puro를 제한 효소 SpeI으로 수 시간 동안 소화시킨 후, 에탄올 침전시키고, 침전물을 멸균수에 용해하였다. SpeI으로 소화시킨 각각의 벡터를 주형으로서 사용하여, 형질전환용 DNA를 하기 프라이머 및 KOD-plus-를 사용하여 PCR 방법 (94°C에서 15초, 55°C에서 30초, 그리고 68°C에서 10분 x 30 사이클)에 의해 제조하였다.

A2 (EF/D) :

5' -GGAAATTGAGAAGTATCATTACAAACAGTACCCACAAACATGAAATAATGTGCTA  
GTCAGAGAGGAATCTTGCAGC-3'

A2 (SV40/D) :

5' -GGAAATTGAGAAGTATCATTACAAACAGTACCCACAAACATGAAATAATGTGCTA  
GTCTGTGGAATGTGTGTCAGTTAG-3'

A2 (EF 및 SV40/R) :

5' -CTCATTCTGTGGTTGTCATTCACTTCCTTGATGCTATCCTTCAAGCAAATT  
TTAAAACTTATCCATCTTGCA-3'

A7 (EF/D) :

5' -CTTATTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAACTTGCTAG  
TCAGAGAGGAATCTTGCAGC-3'

A7 (SV40/D) :

5' -CTTATTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAACTTGCTAG  
TCTGTGGAATGTGTGTCAGTTAG-3'

A7 (EF 및 SV40/R) :

5' -CTCTTCCCATTCTCATTTGAATCTACTTCAAAAGGTTACCATAAGAACTAG  
TTTAAAACTTATCCATCTTGCA-3'

[0171] [0172] 이에 따라 제조된 형질전환용 DNA 및 pRed/ET로 트랜스펙션된 BAC를 사용하여, DNA 요소 A2 또는 A7을 (3-1)에서 기재된 벡터 내로 클로닝하였다. 벡터 구축물을 도 2에 나타낸다. 더욱이, 절차를 (2-2)에서 기재된 방법에 따라 수행하였다.

[0173] (3-3) 지표로서 SEAP 발현을 사용하는 평가

[0174] (3-2)에서 구축한 각각의 플라스미드를 숙주 세포 CHO-K1 (ATCC) 및 트랜스펙션 시약 Lipofectamine 2000 (Invitrogen)을 사용하여 평가하였다.

[0175] 8 μg/ml에서의 퓨로마이신으로의 항생제 선별을 트랜스펙션 2일 후 시작하여 약 2주 동안 수행함으로써, 안정적으로 발현하는 디클론 세포주를 확립하였다. 이에 따라 확립된 세포주를 측정 전날에 배지 교체하고, 일정 세포수를 24-웰 플레이트에 시딩하였다. 세포 플레이팅 후 24시간에, 배양 상청액을 수집하고, SEAP의 활성을 측정하였다. 배양 상청액에서의 SEAP 활성을 SensoLyte™ pNPP 분비 알칼리 포스파타아제 리포터 검정 (ANASPEC)을 사용하여 측정하였다.

- [0176] 측정 결과를 도 4 에 나타낸다. 요소를 갖지 않는 대조군의 SEAP 활성을 1 로 정규화한 경우, DNA 요소 A2 또는 A7 은 발현 증강에 대한 효과를 나타내어, SEAP 의 활성이 EF-1 $\alpha$  프로모터 사용의 경우 대조군보다 2 배 이상 더 높고, SV40 프로모터 사용의 경우 대조군보다 4 배 이상 더 높았다. 결과를 근거로 하여, 이들 DNA 요소가 일반적인 프로모터와 조합으로 사용된 경우 외래 유전자 발현 증강의 효과를 나타낸다는 것이 확인되었다.
- [0177] 실시예 4
- [0178] 지표로서 항체 발현을 사용하는 평가
- [0179] (4-1) 인간 경쇄 발현 벡터 pEF6KCL 의 구축
- [0180] 플라스미드 pEF6/V5-HisB (Invitrogen) 를 주형으로서 사용하여, 위치 2174 (BGHPA 의 바로 다운스트림) 와 위치 2958 (SmaI) 사이의 DNA 단편 (f1 복제 기점 및 SV40 프로모터 및 기점을 포함하는 DNA 단편, 이하 "단편 A" 로서 지칭함, 단편 A 의 폴리뉴클레오티드 서열은 서열 목록에서 SEQ ID NO: 6 으로 표기됨) 을 하기 프라이머 및 KOD-plus- 를 사용하여 PCR 방법에 의해 수득하였다.
- 5' -CCACGCGCCCTGTAGCGCGCATTAAAGC-3'  
5' -AAACCCGGGAGCTTTGCAAAAGCCTAGG-3'
- [0181]
- [0182] 수득한 단편 A, 및 인간  $\kappa$  사슬 분비 신호, 인간  $\kappa$  사슬 불변부 및 인간 폴리 (A) 부가 신호를 인코딩하는 DNA 서열 함유 DNA 단편 (이하, "단편 B" 로서 지칭함) 을 중첩 (overlapping) PCR 에 의해 라이게이션하였다. 이에 따라 수득한, 단편 A 및 단편 B 가 라이게이션된 DNA 단편을 제한 효소 KpnI 및 SmaI 으로 소화시키고, 생성된 단편을 제한 효소 KpnI 및 SmaI 으로 소화시킨 플라스미드 pEF6/V5-HisB (Invitrogen) 에 라이게이션함으로써, EF-1 $\alpha$  프로모터의 다운스트림에 신호 서열, 클로닝 부위, 인간  $\kappa$  사슬 불변부 및 인간 폴리 (A) 부가 신호 서열을 갖는 인간 경쇄 발현 벡터 pEF6KCL 을 구축하였다.
- [0183] 상기 언급한 방법에 의해 수득한 pEF6KCL 을 제한 효소 KpnI 및 SmaI 으로 절단하여 수득한 DNA 단편을, KpnI 및 SmaI 으로 소화시킨 pEF1/myc-HisB (Invitrogen) 에 라이게이션한 후, 형질전환, 알칼리 용해하고 이의 서열 확인을 수행함으로써, 플라스미드 pEF1KCL 을 구축하였다.
- [0184] (4-2) 인간 중쇄 발현 벡터 pEF1FCCU 의 구축
- [0185] 인간 IgG1 신호 서열 및 불변부 아미노산 서열을 인코딩하는 DNA 서열을 함유하는 DNA 단편 (이 DNA 단편의 폴리뉴클레오티드 서열을 서열 목록에서 SEQ ID NO: 7 로 표기함) 을 제한 효소 NheI 및 PmeI 으로 소화시키고, 생성된 단편을 NheI 및 PmeI 으로 소화시킨 플라스미드 pEF1KCL 에 라이게이션함으로써, EF-1 $\alpha$  프로모터의 다운스트림에 신호 서열, 클로닝 부위, 인간 중쇄 불변부 및 인간 폴리 (A) 부가 신호 서열을 갖는 인간 중쇄 발현 벡터 pEF1FCCU 를 구축하였다.
- [0186] (4-3) 단일 인간화 항체 유전자 X 발현 벡터 (인간화 항체 유전자 X/pEF\_LHN#) 의 구축
- [0187] (4-1) 또는 (4-2) 에서 구축한 L-사슬 또는 H-사슬 발현 벡터를 라이게이션하여, 단일 인간화 항체 발현 벡터 (pEF\_LHN (가변부 결여)) 를 구축하였다.
- [0188] PCR 방법에 의해, 제한 효소 SalI 부위를 프로모터의 업스트림에서 pEF1KCL 의 폴리 (A) 의 다운스트림까지의 유전자 발현 단위체의 양끝에 추가하였다. 아가로오스 겔 전기영동을 수행하여, 겔에서 원하는 DNA 단편을 절단하고, DNA 단편을 정제함으로써, 삽입물을 제조하였다. (4-2) 에서 구축한 pEF1FCCU 를 제한 효소 SalI 으로 소화시켜, 벡터를 유전자 발현 단위체의 업스트림에 위치한 SalI 부위에서 선형화하였다. 이후, 선형화된 벡터를 상기 삽입물에 라이게이션한 후, 형질전환, 알칼리 용해하고 서열 확인함으로써, 단일 인간화 항체 발현 벡터 (pEF\_LHN (가변부 결여)) 를 구축하였다.
- [0189] 이후, 하기의 올리고뉴클레오티드를 벡터 pEF\_LHN (가변부 결여) 의 AatII 부위에 도입하였다.
- 5' -CGCGGCCGCACTAGTGACGT-3'  
5' -CACTAGTGCGGCCGCGACGT-3'
- [0190]
- [0191] 각각의 올리고뉴클레오티드를 5 pmol 로 희석하고, T4 폴리뉴클레오티드 키나아제 (TAKARA) 를 사용하여, 반응 을 37°C 에서 1 시간 동안 진행시켰다. 그런 다음, 10x 완충액 (TAKARA) 을 이에 추가하고, 어닐링을 96°C 에서 1 분 동안 실온에서 수행하였다. 이들 올리고뉴클레오티드 및 제한 효소 AatII 로 소화시킨 벡터

pEF\_LHN 을 라이제이션한 후, 형질전환, 알칼리 용해하고, 서열 확인함으로써, pEF\_LHN# (가변부 결여) 을 구축하였다.

[0192] 인간화 항체 유전자 X 의 가변부를 상기 구축한 범용 (universal)벡터 (pEF\_LHN# (가변부 결여)) 내로 통합시켜, 인간화 항체 유전자 X 발현 단일 벡터 (인간화 항체 유전자 X/pEF\_LHN#) 의 구축을 완결하였다.

[0193] 먼저, 하기 프라이머 및 KOD-plus- 를 사용하여, 인간화 항체 유전자 X 의 L-사슬 가변부를 PCR 방법 (94°C 에서 15 초, 55°C 에서 30 초, 그리고 68°C 에서 1 분 x 30 사이클) 에 의해 증폭하였다.

[0194] L-사슬 가변부:

5' -AACATATGGCGACATCCAGATGAC-3'

[0195] 5' -AACGTACGCTTGATCTCCACCTTGG-3'

[0196] 증폭된 L-사슬 가변부 단편 및 범용 벡터 (pEF\_LHN# (가변부 결여)) 를 제한 효소 NdeI 및 BsiWI 으로 소화시킨 후, 아가로오스 겔 전기영동을 수행하여, 겔에서 원하는 단편을 절단하고, 정제하고, 라이제이션 반응시키고, 형질전환, 알칼리 용해하고, 서열 확인함으로써, L-사슬 가변부를 벡터 내로 통합하였다. 동일한 방식으로, 하기 프라이머 및 KOD-plus- 를 사용하여, 인간화 항체 유전자 X 의 H-사슬 가변부를 PCR 방법 (94°C 에서 15 초, 55°C 에서 30 초, 그리고 68°C 에서 1 분 x 30 사이클) 에 의해 증폭하였다.

[0197] H-사슬 가변부:

5' -AAAGCTGAGGCCAGGTGCAGCTGCAGG-3'

[0198] 5' -AAAGCTGAGCTCACGGTCACCAGGGTTC-3'

[0199] 증폭된 H-사슬 가변부 단편 및 삽입된 L-사슬 가변부를 갖는 벡터를 제한 효소 BspI 으로 소화시킨 후, 아가로오스 겔 전기영동을 수행하여, 겔에서 원하는 단편을 절단하고, 정제하고, 라이제이션 반응시키고, 형질전환, 알칼리 용해하고, 서열 확인함으로써, H-사슬 가변부를 벡터 내로 통합하고 단일 인간화 항체 유전자 X 발현 벡터 (인간화 항체 유전자 X/pEF\_LHN#) 를 구축하였다.

[0200] (4-4) 단일 인간화 항체 유전자 X 발현 벡터 (인간화 항체 유전자 X/pCMV\_LHN#) 의 구축

[0201] (4-3)에서 구축한 단일 인간화 항체 유전자 X 발현 벡터 (인간화 항체 유전자 X/pEF\_LHN#) 를 기본 벡터 구조로서 사용하여, 하기 절차에 따라 프로모터를 교체하여 또 다른 단일 인간화 항체 유전자 X 발현 벡터 (인간화 항체 유전자 X/pCMV\_LHN#) 를 구축하였다.

[0202] pIRES puro3 을 주형으로서 사용하여, CMV 프로모터 단편을 하기 프라이머 및 KOD-plus- 를 사용하여 PCR 방법 (94°C 에서 30 초, 그리고 68°C 에서 3 분 x 40 사이클) 에 의해 증폭하였다.

[0203] H-사슬의 업스트립:

5' -CTTTTGCAAAAGCTTCGGTTACATAACTTACGGTAAATGGCC-3'

[0204] 5' -TTCATGGTGGCGCTAGCCCCGAGATATCGATCCGAGCTCGGTA-3'

[0205] L-사슬의 업스트립:

5' -TGACGTCGACAAGCTTCGGTTACATAACTTACGGTAAATGGCC-3'

5' -CTGGATGTCGCCATATGCGCCGGAGATCCACAGCAGCAGGGAGATGAACACCTGG

[0206] GTCTGCAGCACCATGGTGGCGCTAGCCCCGAGATATCGATCCGAGCTCGGTA-3'

[0207] PCR 반응 용액에, 제한 효소 DpnI 을 추가하고, 반응을 37°C 에서 1 시간 동안 진행시킨 후, miniElute 반응 정화 (Cleanup) 키트 (Qiagen) 를 사용하여 정제함으로써, In-Fusion 에서 사용하기 위한 샘플을 제조하였다. 그 동안, 인간화 항체 유전자 X/pEF\_LHN# 을 제한 효소 HindIII, NheI, NdeI 및 FseI 으로 소화시킨 후, 아가로오스 겔 전기영동을 수행함으로써, 생성된 단편 중에서 2 개의 거대 단편을 분리하였다. 각각의 단편을 겔에서 절단하고, DNA 를 겔에서 추출하여, In-Fusion 에서 사용하기 위한 샘플을 제조하였다. In-Fusion 에서 사용하기 위한 모든 샘플을 모으고, In-Fusion™ Advantage PCR 클로닝 키트 (TAKARA) 를 사용하여 클로닝 을 수행한 후, 형질전환, 알칼리 용해하고, 서열 확인함으로써, 단일 인간화 항체 유전자 X 발현 벡터 (인간화 항체 유전자 X/pCMV\_LHN#) 를 구축하였다.

[0208] (4-5) DNA 요소 A7 의 클로닝

[0209] A7 을 SEAP 발현 증강 효과를 갖는 것으로 확인된 DNA 요소 5 개 유형에서 선택하고, 항체 발현 벡터 내로 클로닝하였다.

[0210] (2-2) 에서와 동일한 방식으로, 제한 효소 NotI 으로 소화시킨 각각의 인간화 항체 유전자 X 발현 단일 벡터 (인간화 항체 유전자 X/pEF\_LHN# 및 인간화 항체 유전자 X/pCMV\_LHN#) 를 주형으로서 사용하여, 형질전환용 DNA 를 하기 프라이머 및 KOD-plus- 를 사용하여 수행한 PCR 방법 (94°C 에서 15 초, 55°C 에서 30 초, 그리고 68°C 에서 11 분 x 30 사이클) 에 의해 제조하였다.

[0211] 인간화 항체 유전자 X/pEF\_LHN# D:

5' -CTCTTCCCATTCTCATTGAATCTACTTCAAAAGGTTACCATAACTAAGACTCGA

GGCACTAGTGACGTCAGGTGGCACT- 3'

[0213] 인간화 항체 유전자 X/pEF\_LHN# R:

5' -CTCTTCCCATTCTCATTGAATCTACTTCAAAAGGTTACCATAACTAAGAGCACT

AGTGACGTCAGGTGGCACTTTTCGG- 3'

[0215] 인간화 항체 유전자 X/pCMV\_LHN# D:

인간화 항체 유전자 X/pEF\_LHN# D 사용.

[0217] 인간화 항체 유전자 X/pCMV\_LHN# R:

인간화 항체 유전자 X/pEF\_LHN# R 사용.

[0219] 상기 제조한 형질전환용 DNA 및 pRed/ET 로 트랜스펙션된 BAC 를 사용하여, DNA 요소 A7 을 (4-3) 및 (4-4) 에서 기재된 단일 인간화 항체 유전자 X 발현 벡터 내로 클로닝하였다. 벡터 구축물을 도 5 에서 나타낸다. 더욱이, 절차를 (2-2) 에서 기재된 방법에 따라 수행하였다.

[0220] (4-6) 지표로서 항체 발현을 사용하는 평가

[0221] (4-5) 에서 구축한 각각의 플라스미드를 숙주 세포 CHO-K1 (ATCC) 및 트랜스펙션 시약 Lipofectamine 2000 (Invitrogen) 을 사용하여 평가하였다.

[0222] 800  $\mu$ g/ml 에서의 제네티신 (Roche) 으로의 항생제 선별을 트랜스펙션 2 일 후 시작하여 약 2 주 동안 수행함으로써, 안정적으로 발현하는 디클론 세포주를 확립하였다. 이에 따라 확립된 세포주를 측정 전날에 배지 교체하고, 일정 세포수를 24-웰 플레이트에 시딩하였다. 세포 플레이팅 후 24 시간에, 배양 상청액을 수집하고, 배양 상청액에서의 항체의 발현 수준을 ELISA 방법에 의해 측정하였다. 더욱이, ELISA 를 하기와 같이 수행하였다. 50 ng/웰에서의 항-카파 경쇄로 코팅된 96-웰 플레이트에, 100  $\mu$ l 의 세포-미함유 배양 상청액을 각각의 웰에 추가하고, 상기 플레이트를 37°C 에서 1 시간 동안 인큐베이션하였다. 이후, 샘플 (배양 상청액) 을 제거하고, 각각의 웰을 200  $\mu$ l 의 PBS-Tween (0.05%) 으로 세척하였다. 그런 다음, 100  $\mu$ l 의 HRP-표지 항-인간 IgG (Fc) 를 각각의 웰에 추가하고, 플레이트를 37°C 에서 추가 1 시간 동안 인큐베이션하였다. 그후, HRP-표지 항-인간 IgG (Fc) 를 제거하고, 각각의 웰을 PBS-Tween (0.05%) 으로 세척하였다. 그런 다음, POD 기질 ABTS 키트 (Nacalai) 를 사용하여 발색시키고, 405 nm 의 측정 파장에서 흡광도를 측정하였다. 항-카파 경쇄, 항-인간 IgG (Fc) 및 샘플의 희석을 위해서, PBS-Tween (0.05%) 을 사용하였다. 12 ng, 6 ng, 3 ng, 1.5 ng, 0.75 ng, 0.375 ng 및 0.1875 ng 으로 연속 희석한 인간 IgG 를 표준으로서 사용하여, 샘플 농도를 계산하였다.

[0223] 결과를 도 6 에서 나타낸다. EF-1 $\alpha$  프로모터 또는 CMV 프로모터를 항체 발현 벡터에서 사용한 경우, 요소를 갖지 않는 대조군과 비교하여 DNA 요소 A7 을 갖는 샘플이 더 높은 항체 생성 증강 효과를 갖는다는 것이 확인되었다.

[0224] 실시예 5

[0225] 외래 유전자 발현 증강 활성을 나타내는 서열의 길이

[0226] (5-1) 상이한 서열 길이를 갖는 DNA 요소의 클로닝

[0227] 실시예 2에서 사용한 서열의 길이를 기준으로, 상이한 서열 길이를 갖는 각각의 DNA 요소를 함유하는 벡터를 구축하였다.

[0228] 각각의 DNA 요소 A2, A7, A18, B5 및 C14의 전체 길이를 기준으로 설계한 상이한 서열 길이를 갖는 DNA 요소의 세부 사항을 각각 도 7, 9, 11, 13, 15, 18 및 19에 나타낸다. (2-1)에서 기재된 pCMV/SEAP ires Hygro를 수 시간 동안 제한 효소 SpeI으로 소화시킨 후, 에탄올 침전 처리하고, 침전물을 멸균수에 용해하였다. SpeI으로 소화시킨 벡터를 주형으로서 사용하여, 형질전환용 DNA를 하기 프라이머 및 KOD-plus-를 사용하여 PCR 방법(94°C에서 15초, 55°C에서 30초, 그리고 68°C에서 10분 x 30사이클)에 의해 제조하였다. 이에 따라 제조된 형질전환용 DNA 및 pRed/ET로 트랜스펙션된 상응하는 BAC를 사용하여, 상이한 서열 길이를 갖는 각각의 DNA 요소를 (2-1)에서 기재된 pCMV/SEAP ires Hygro내로 클로닝하였다. 벡터 구축물을 도 2에 나타낸다. 더욱이, 절차를 (2-2)에서 기재된 방법에 따라 수행하였다.

A2-1D:

5' -CATGCACAGATTAGCCATTTAGTACTTAACTCAAACCTCAATTCTGAAGTCT  
AGTTATTAATAGTAATCAATTACG-3'

A2-1R:

5' -CTCATTCTGTGGGTTGTCAATTCACTTCCTTGATGCTATCCTTCAAGCAAAATT  
CAATAATCAATGTCAACGCGTATAT-3'

A2-2D:

5' -ACACTGGTCAAAGGGACAGGTCAATTGTTATGCTGGCAATGCAGGCTGCTGAAAAC  
TAGTTATTAATAGTAATCAATTACG-3'

A2-2R:

5' -ACTGTAGCTTCTTATTTTTACCTGCAGTGCATTCCCTGTAAAAGTAGTGTGGAGT  
CAATAATCAATGTCAACGCGTATAT-3'

A2-3D:

5' -CTGGAAATTGAGAAGTATCATTACAACAGTACCAACAAACATGAAATAAATGTGC  
TAGTTATTAATAGTAATCAATTACG-3'

A2-3R:

5' -CCAAGCTTGTCCAACCGCGGCCTGCAGGCTGCATGCAGCCTGTGAAGGCTTGAT  
CAATAATCAATGTCAACGCGTATAT-3'

A2-4D:

5' -TCAATCATTATCAATTATCTTCAAAGTCCCTCACTTCAGGGAGATGATATAC  
TAGTTATTAATAGTAATCAATTACG-3'

A2-4R:

5' -ATATATAAAAGTCATGTATATATAAAATCATGCAATACACGGCTTTGTGACT  
CAATAATCAATGTCAACGCGTATAT-3'

A2-5D:

5' -CGCATAAAAGGAAAGCATCCTAAAATAAACACCCTCAATGGCTCCTCGGTGGC  
TAGTTATTAATAGTAATCAATTACG-3'

[0235]

A2-5R:

- [0236] A2-4R 사용.  
A2 - 6D :  
5' - GGGAGGCTACAGCTTGCCTCTCTAACCACTAAAGGCATGACCCTCCTCAAAGCT
- [0237] AGTTATTAAATAGTAATCAATTACG - 3'
- [0238] A2-6R:
- [0239] A2-4R 사용.  
A2 - 7D :  
5' - TCTGGCTTCCTGGGCCACGCTGGAAGAAGAATTGTCTTGCGCCACACATAAAAC  
TAGTTATTAAATAGTAATCAATTACG - 3'
- A2 - 7R :  
5' - AGCTGATTTTACGTTAAATGTAACATGTAAGAAATATATGTGTGTTTAGAT  
CAATAATCAATGTCAACGCGTATAT - 3'
- [0240] A2 - 8D :  
5' - GTGAAGAGGAGGAGATGTCAAAATTCAAAGTCTTAAATGATGTAGTTAAAGTAC  
TAGTTATTAAATAGTAATCAATTACG - 3'
- [0241] A2 - 8R :  
5' - ATGACACTTGATATTGTTGTTATATTGCTGGTTAGTATGTGCCTTCATTACCT  
CAATAATCAATGTCAACGCGTATAT - 3'
- [0242] A2-9D:
- [0243] A2-9R:
- [0244] A2-6D 사용.
- [0245] A2-10R:
- [0246] A2R 사용.
- [0247] A2-10D:
- [0248] A2-2D 사용.
- [0249] A2-10R:
- [0250] A2-7R 사용.
- [0251] A2-11D:
- [0252] A2-8D 사용.
- [0253] A2-11R:
- [0254] A2-2R 사용.
- [0255] A2-12D:
- [0256] A2-2D 사용.
- [0257] A2-12R:
- [0258] A2-4R 사용.
- [0259] A2-13D:
- [0260] A2-8D 사용.
- [0261] A2-13R:

[0262] A2-7R 사용.

[0263] A2-14D:

[0264] A2D 사용.

[0265] A2-14R:

[0266] A2-2R 사용.

[0267] A2-15D:

[0268] A2-2D 사용.

[0269] A2-15R:

[0270] A2R 사용.

[0271] A2-16D:

[0272] A2-8D 사용.

[0273] A2-16R:

[0274] A2-4R 사용.

[0275] A2-17D:

[0276] A2D 사용.

[0277] A2-17R:

[0278] A2-7R 사용.

[0279] A7-1D:

5' -AAAAACAAAATGGAGTAAACAAGATGAATTGTTTAATAGAGGCAGTGATTAC  
TAGTTATTAATAGTAATCAATTACG- 3'

A7-1R:

5' -ATACAATGTTCCATGTATTCTGTGCCTGAACCTATGCAGCTGATGTAGCTGAAGT  
CAATAATCAATGTCAACGCGTATAT- 3'

[0280]

A7-2D:

5' -GATCTTATTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAACCTTGC  
TAGTTATTAATAGTAATCAATTACG- 3'

A7-2R:

5' -TGTTGTTTCAGCCACTAAGTTGAGGTGATTGTTCTGGCAGTCCTAGGAACT

[0281]

CAATAATCAATGTCAACGCGTATAT- 3'

[0282]

A7-3D:

[0283]

A7-2D 사용.

A7-3R:

5' -AGCCTACACTACCCTTGAGCCTTGGTAACTATCCTCTGCTGTCTACCTCCT  
CAATAATCAATGTCAACGCGTATAT- 3'

[0284]

A7-4D:

5' - AGGAGCTCCTGAATGAAGGACATCACTCAGCTGTGTTAAGTATCTGGAAACAATAC  
TAGTTATTAATAGTAATCAATTACG - 3'

A7-4R:

5' - GACATAAAATGTAAGATATGATATGCTATGTAAGATATGATACTGCCTTAAAAT  
CAATAATCAATGTCAACGCGTATAT - 3'

[0285]

A7-5D:

5' - CACTGCTTGATACCTACTGTGGACTTGAAAATTATGAATGTGTGTGTGTGTC  
TAGTTATTAATAGTAATCAATTACG - 3'

A7-5R:

5' - CAATTACATTCCAGTGATCTGCTACTTAGAATGCATGACTGAACCTGGTGGT  
CAATAATCAATGTCAACGCGTATAT - 3'

[0286]

A7-6D:

5' - TTATTTGAAGAGAAACTCCTGGTCCCACTTAAAATCCTTCTGTTCCAAGC  
TAGTTATTAATAGTAATCAATTACG - 3'

A7-6R:

5' - AAGCAGTGTGTGTTACCTGCATGTATGTGAATTAACCTGTTCTGAGGCAT  
CAATAATCAATGTCAACGCGTATAT - 3'

[0287]

A7-7D:

5' - ATTGCATGTTCTCATTTATTTGTGGATGAAAAATCAAACAAATAGAACGTATC  
TAGTTATTAATAGTAATCAATTACG - 3'

A7-7R:

5' - TTGGGAGGCCGAGCTGGTAGATCACTTGAGGCCACGAATTGACACCAGCAGGT  
CAATAATCAATGTCAACGCGTATAT - 3'

[0288]

A7-8D:

[0289]

A7-1D 사용.

[0290]

A7-8R:

[0291]

A7R 사용.

[0292]

A7-9D:

[0293]

A7-7D 사용.

[0294]

A7-9R:

[0295]

A7-5R 사용.

[0296]

A7-10D:

[0297]

A7-4D 사용.

[0298]

A7-10R:

[0299]

A7-7R 사용.

[0300]

A7-11D:

- [0302] A7-6D 사용.
- [0303] A7-11R:
- [0304] A7-4R 사용.
- [0305] A7-12D:
- [0306] A7-2D 사용.
- [0307] A7-12R:
- [0308] A7-6R 사용.
- [0309] A7-13D:
- [0310] A7-7D 사용.
- [0311] A7-13R:
- [0312] A7R 사용.
- [0313] A7-14D:
- [0314] A7-4D 사용.
- [0315] A7-14R:
- [0316] A7-5R 사용.
- [0317] A7-15D:
- [0318] A7-6D 사용.
- [0319] A7-15R:
- [0320] A7-7R 사용.
- [0321] A7-16D:
- [0322] A7-2D 사용.
- [0323] A7-16R:
- [0324] A7-4R 사용.
- [0325] A7-17D:
- [0326] A7-4D 사용.
- [0327] A7-17R:
- [0328] A7R 사용.
- [0329] A7-18D:
- [0330] A7-6D 사용.
- [0331] A7-18R:
- [0332] A7-5R 사용.

A18-1:

5' -ATCCCCTGCTCTGCTAAAAAAGAATGGATGTTGACTCTCAGGCCCTAGTTCTTGA

TCCTATTAATAGTAATCAATTACG-3'

- [0333] A18-1R:

[0335] A18R 사용.

A18-2D:

5' -CTAAAGTGTGGATTACAGGCATAAGCCACCGTGCCCCGCTGGAGCATTGGAT  
CCTATTAATAGTAATCAATTACG-3'

A18-2R:

5' -ACTACTTACACATTCGAGTTAAATAAGGCCTCAATATAGAGTGAACACCTA  
GTCAATAATCAATGTCAACG-3'

[0336]

A18-3D:

5' -CAGGCATAAGCCACCGCACCCGCCACCCCTTACTAATTTAGTAACGTGATC  
CTATTAATAGTAATCAATTACG-3'

A18-3R:

5' -CTGATTGACTTGACCTCTGCTTCCAACTTGCCCAAAGAAAGTTAGTCACCT  
AGTCAATAATCAATGTCAACG-3'

[0337]

[0338]

A18-4D:

[0339]

A18-3D 사용.

[0340]

A18-4R:

5' -TTCAATGAAACAAGCTCTGTGAGGCTCATTTGTACCCATTTGTTCACTGACTGCC  
TAGTCAATAATCAATGTCAACG-3'

[0341]

B5-1D:

5' -ACATACCCAGAGACACTGAGAGAGACAGACAGACAGTAAACAGAGGAGCACGATC  
CTATTAATAGTAATCAATTACG-3'

[0342]

B5-1R:

[0343]

B5R 사용.

B5-2D:

5' -GCTCAATTGTATCTTATGAAAACAATTTTCAAAATAAAACAAGAGATATGATCC  
TATTAATAGTAATCAATTACG-3'

[0345]

B5-2R:

[0346]

B5R 사용.

B5-3D:

5' -CCTGTGCTGAATACCGTCTGCATATGTATAGGAAAGGGTTAACTCAGCAGGGATC  
CTATTAATAGTAATCAATTACG-3'

B5-3R:

5' -TATGTGAATGAAATAAAATAATCAAGCTTGTAGAATTGTGTTCAATGACCC  
TAGTCAATAATCAATGTCAACG-3'

[0348]

B5-4D:

[0349]

B5D 사용.

[0350]

B5-4R:

5' -GAAAGTCTACAATTTTCAGTTAAAATGGTATTATTTGTAACATGTACCCCA

[0352] GTCATAATCAATGTCAACG-3'

[0353] B5-5D:

[0354] B5-1D 사용.

B5-5R:

5' -CAAAGATGAAGGATGAGAGTGACTTCTGCCTTCATTATGTTATGTGTTCATATCC

[0355] TAGTCAATAATCAATGTCAACG-3'

B5-6D:

5' -CAGTGAATTATTCACTTTGCTTAGTTAAGTAAAAAATAAAATCTGACTGTGATCC

TATTAATAGTAATCAATTACG-3'

B5-6R:

5' -GAACAGACAGGTGAATGAGCACAGAGGTCAATTGTAACCGTTGTGGTTAGCCT

[0356] AGTCAATAATCAATGTCAACG-3'

C14-1D:

5' -CTTTTGCGTTCTGTGTTAAGTTATTTCCCCTAGGCCACAAACAGAGTCGA

TCCTATTAATAGTAATCAATTACG-3'

C14-1R:

5' -AACCTTGGAAAAATTCTGTTGTTAGAAGCATGTACCAATCTATCACTCCTAG

[0357] TCAATAATCAATGTCAACG-3'

C14-2D:

5' -CTATTCACTGCTGTAGGATGAAAAAGTTAATAACACCCTGAGAGGTTTCGATCC

TATTAATAGTAATCAATTACG-3'

C14-2R:

5' -CCTTAGATTAGTTATTGTATTTTATCAGCTACTATAAGGTTACACACCCTA

[0359] GTCAATAATCAATGTCAACG-3'

C14-3D:

5' -CAAGACCCTCAAAATTCAAAAATTCCTTATCTGCTGTAGCACCTCCTGCGAT

CCTATTAATAGTAATCAATTACG-3'

C14-3R:

5' -GGAGGGGATAGGAAGGGATGAGGCCTAACAGGTTGATGATCTAGGTTACCTA

[0360] GTCAATAATCAATGTCAACG-3'

C14-4D:

5' -CTCAAAAGGAGATAATTCCAGCCCCCTGCCTTAAAGAATCCCTATCAAGTGATC

[0361] CTATTAATAGTAATCAATTACG-3'

[0362] C14-4R:

- [0363] C14-1R 사용.  
C14-5D:  
5' -CGCTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCCGAGATCACGCCGTTGGATCC
- [0364] TATTAAATAGTAATCAATTACG-3'
- [0365] C14-5R:
- [0366] C14-1R 사용.
- [0367] C14-6D:
- [0368] C14-4D 사용.  
C14-6R:  
5' -TTAACCTTTCATCCTACAGACAGTGAATAGTAAAGCTTCTGTGAAGACATACC
- [0369] CTAGTCAAATAATCAATGTCAAACG-3'
- [0370] C14-7D:
- [0371] C14-2D 사용.
- [0372] C14-7R:
- [0373] C14-1R 사용.
- [0374] C14-8D:
- [0375] C14-3D 사용.  
C14-8R:  
5' -AAATTATTCCTGGTGGCAATATTAGAATATGGGAATGTTGCTTCTGAGCCT
- [0376] AGTCAAATAATCAATGTCAAACG-3'
- [0377] C14-9D:
- [0378] C14-4D 사용.
- [0379] C14-9R:
- [0380] C14-3R 사용.
- [0381] C14-10D:
- [0382] C14-2D 사용.
- [0383] C14-10R:
- [0384] C14R 사용.
- [0385] C14-11D:
- [0386] C14-3D 사용.
- [0387] C14-11R:
- [0388] C14-2R 사용.
- [0389] C14-12D:
- [0390] C14-4D 사용.
- [0391] C14-12R:
- [0392] C14-8R 사용.

- [0393] C14-13D:
- [0394] C14-3D 사용.
- [0395] C14-13R:
- [0396] C14-1R 사용.
- [0397] C14-14D:
- [0398] C14-4D 사용.
- [0399] C14-14R:
- [0400] C14-2R 사용.
- [0401] A2 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, A2-1 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 1 ~ 3000 의 폴리뉴클레오티드 서열에 상응하고; A2-2 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2801 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-3 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 5401 ~ 8450 의 폴리뉴클레오티드 서열에 상응하고; A2-4 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 2700 의 폴리뉴클레오티드 서열에 상응하고; A2-5 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 2200 의 폴리뉴클레오티드 서열에 상응하고; A2-6 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 3700 의 폴리뉴클레오티드 서열에 상응하고; A2-7 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 5000 의 폴리뉴클레오티드 서열에 상응하고; A2-8 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 4001 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-9 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 1 ~ 3700 의 폴리뉴클레오티드 서열에 상응하고; A2-10 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-11 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2801 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-12 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-13 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-14 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2801 ~ 8450 의 폴리뉴클레오티드 서열에 상응하고; A2-15 는 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 1 ~ 5800 의 폴리뉴클레오티드 서열에 상응하고; A2-16 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 701 ~ 7000 의 폴리뉴클레오티드 서열에 상응하고; A2-17 은 서열 목록에서 SEQ ID NO: 1 의 뉴클레오티드 2001 ~ 8450 의 폴리뉴클레오티드 서열에 상응한다.
- [0402] A7 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, A7-1 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 601 ~ 3600 의 폴리뉴클레오티드 서열에 상응하고; A7-2 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3601 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-3 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 5401 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-4 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3401 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-5 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 4500 의 폴리뉴클레오티드 서열에 상응하고; A7-6 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 4401 ~ 7400 의 폴리뉴클레오티드 서열에 상응하고; A7-7 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 2401 ~ 5400 의 폴리뉴클레오티드 서열에 상응하고; A7-8 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1 ~ 3600 의 폴리뉴클레오티드 서열에 상응하고; A7-9 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 5400 의 폴리뉴클레오티드 서열에 상응하고; A7-10 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 2401 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-11 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3401 ~ 7400 의 폴리뉴클레오티드 서열에 상응하고; A7-12 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 4401 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-13 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1 ~ 5400 의 폴리뉴클레오티드 서열에 상응하고; A7-14 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-15 는 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 2401 ~ 7400 의 폴리뉴클레오티드 서열에 상응하고; A7-16 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 3401 ~ 8420 의 폴리뉴클레오티드 서열에 상응하고; A7-17 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1 ~ 6400 의 폴리뉴클레오티드 서열에 상응하고; A7-18 은 서열 목록에서 SEQ ID NO: 2 의 뉴클레오티드 1501 ~ 7400 의 폴리뉴클레오티드 서열에 상응한다.
- [0403] A18 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, A18-1 은 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 1 ~ 5040 의 폴리뉴클레오티드 서열에 상응하고; A18-2 는 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 1001 ~ 6002 의 폴리뉴클레오티드 서열에 상응하고; A18-3 은 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 2001 ~

7000 의 폴리뉴클레오티드 서열에 상응하고; A18-4 는 서열 목록에서 SEQ ID NO: 3 의 뉴클레오티드 3000 ~ 7000 의 폴리뉴클레오티드 서열에 상응한다.

[0404] B5 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, B5-1 은 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 1 ~ 4001 의 폴리뉴클레오티드 서열에 상응하고; B5-2 는 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 1 ~ 3200 의 폴리뉴클레오티드 서열에 상응하고; B5-3 은 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 2491 ~ 5601 의 폴리뉴클레오티드 서열에 상응하고; B5-4 는 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 5373 ~ 8401 의 폴리뉴클레오티드 서열에 상응하고; B5-5 는 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 901 ~ 4001 의 폴리뉴클레오티드 서열에 상응하고; B5-6 은 서열 목록에서 SEQ ID NO: 4 의 뉴클레오티드 4001 ~ 7000 의 폴리뉴클레오티드 서열에 상응한다.

[0405] C14 의 각 단편의 폴리뉴클레오티드 서열에 대해서는, C14-1 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 4015 의 폴리뉴클레오티드 서열에 상응하고; C14-2 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1987 ~ 5014 의 폴리뉴클레오티드 서열에 상응하고; C14-3 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 4020 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-4 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-5 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 6011 의 폴리뉴클레오티드 서열에 상응하고; C14-6 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 4939 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-7 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 5014 의 폴리뉴클레오티드 서열에 상응하고; C14-8 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 2994 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-9 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 4020 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-10 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1 ~ 5014 의 폴리뉴클레오티드 서열에 상응하고; C14-11 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1987 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-12 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 2994 ~ 8141 의 폴리뉴클레오티드 서열에 상응하고; C14-13 은 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 960 ~ 7119 의 폴리뉴클레오티드 서열에 상응하고; C14-14 는 서열 목록에서 SEQ ID NO: 5 의 뉴클레오티드 1987 ~ 8141 의 폴리뉴클레오티드 서열에 상응한다.

[0406] 전체 길이 서열에서의 각 단편의 출발점 및 종료점을 또한 도 18 및 19 에 나타낸다.

[0407] (5-2) 상이한 서열 길이를 갖는 DNA 요소의 평가

[0408] (5-1) 에서 구축한 각각의 플라스미드를 숙주 세포 CHO-K1 (ATCC) 및 트랜스펙션 시약 Lipofectamine 2000 (Invitrogen) 을 사용하여 평가하였다.

[0409] (2-3) 에서와 동일한 방식으로, 히그로마이신으로의 항생제 선별을 트랜스펙션 후 수행함으로써, 안정적으로 발현하는 다클론 세포주를 확립하였다. 이에 따라 확립된 세포주를 측정 전날에 배지 교체하고, 일정 세포수를 24-웰 플레이트에 시딩하였다. 세포 플레이팅 후 24 시간에, 배양 상청액을 수집하고, SEAP 의 활성을 측정하였다.

[0410] 측정 결과를 도 8, 10, 12, 14 및 16 에 나타낸다. 전체 길이 DNA 요소 뿐 아니라 전체 길이보다 짧은 서열 길이를 갖는 클론이 발현 증강 효과를 갖는다는 것이 확인되었다. 결과를 근거로 하여, DNA 요소 A2, A7, A18, B5 및 C14 가 전체 길이보다 짧은 서열 길이를 갖는 경우에도 외래 유전자 발현 증강 활성을 갖는다는 것이 확인되었다. 그러나, 이는 서열 길이가 전체 길이인 경우 최고 효과를 나타낸다.

#### 실시예 6

[0412] CHO 세포주 외의 숙주 세포 사용의 효과

[0413] CHO 세포주를 실시예 2 내지 5 에서의 평가에서 세포주로서 사용하였다. 그러나, 실시예 6 에서는 HEK293 세포주를 CHO 세포주 외의 세포주로서 선택하였다. HEK293 세포주를 10% FCS 를 함유하는 DMEM 배지 (Invitrogen) 를 사용하여 37°C 에서 5% CO<sub>2</sub> 의 존재 하에 정지 배양하고, 일정 세포수를 트랜스펙션 전날에 6-웰 플레이트에 시딩하였다. (3-2) 에서 구축한 각각의 DNA 요소를 함유하는 SEAP 발현 벡터를 평가하기 위해서, 각각의 플라스미드 및 트랜스펙션 시약 Lipofectamine 2000 (Invitrogen) 을 사용하여 트랜스펙션을 수행하였다. 히그로마이신으로의 항생제 선별을 트랜스펙션 2 일 후에 시작하여 약 2 주 동안 수행함으로써, 안정적으로 발현하는 다클론 세포주를 확립하였다. 이에 따라 확립된 세포주를 측정 전날에 배지 교체하고, 일정 세포수를 24-웰 플레이트에 시딩하였다. 세포 플레이팅 후 24 시간에, 세포 상청액을 수집하고, SEAP

의 활성을 측정하였다. 배양 상청액에서의 SEAP 활성을 SensoLyte™ pNPP 분비 알칼리 포스파타아제 리포터 검정 (ANASPEC) 을 사용하여 측정하였다.

[0414] 측정 결과를 도 17 에 나타낸다. 실시예 3 에서와 동일한 방식으로, HEK293 세포주에서의 외래 유전자 (SEAP) 발현 증강에 있어서, 각각의 요소가 또한 매우 유효하다는 것이 확인되었다.

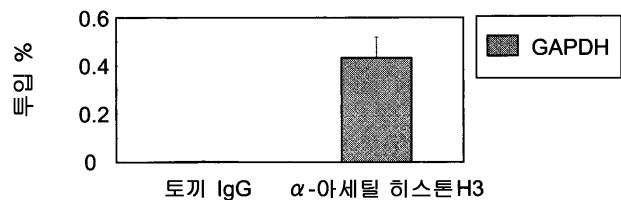
[0415] [산업상 이용가능성]

[0416] 본 발명에 따른 DNA 요소를 사용하여 외래 유전자 발현 벡터를 포유류 숙주 세포에 도입함으로써, 치료 단백질, 항체 등의 외래 유전자의 생산성이 향상될 수 있다.

## 도면

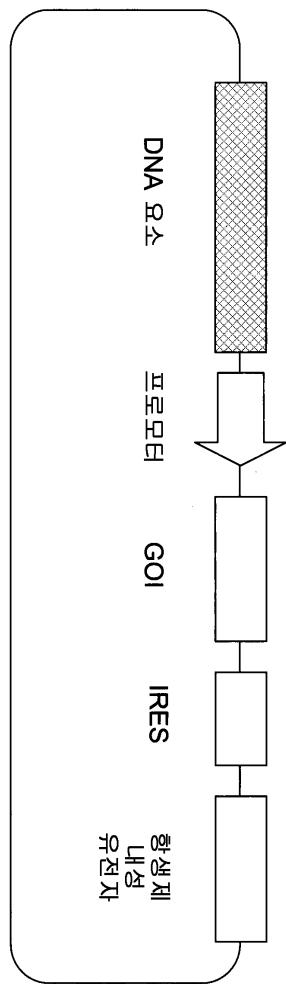
### 도면1

염색질 면역침진법

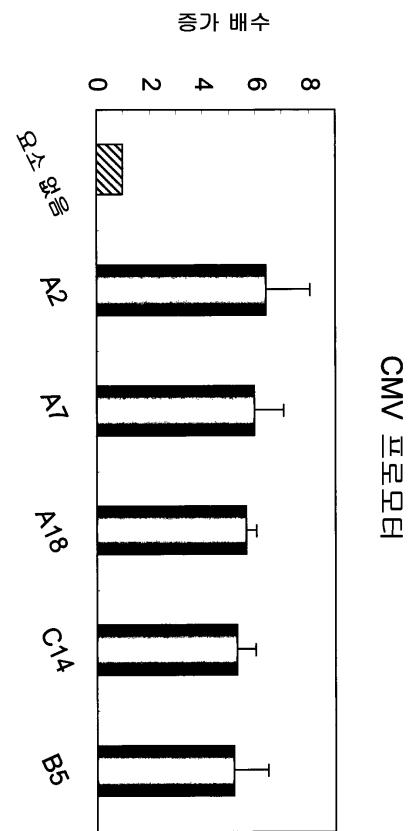


### 도면2

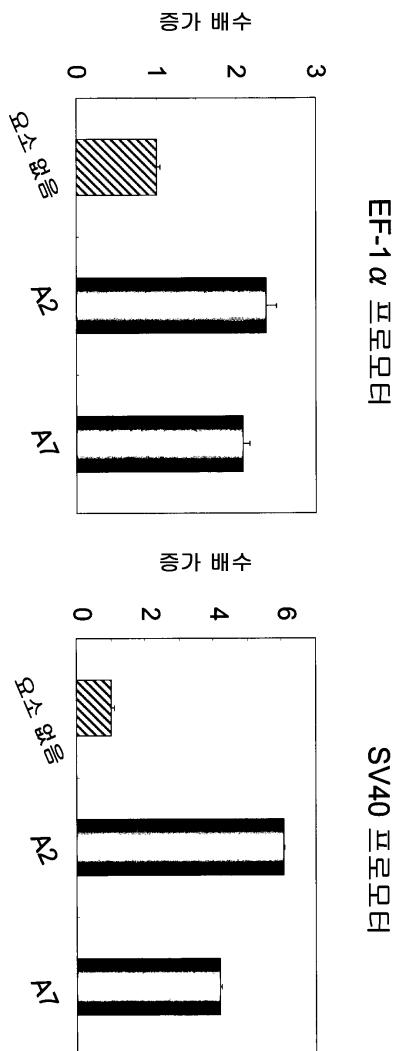
GOI : 관심 유전자 (SEAP)



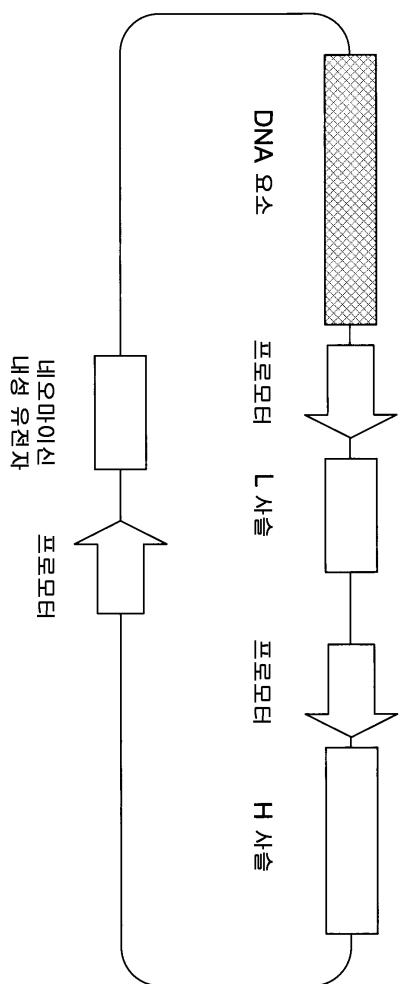
## 도면3



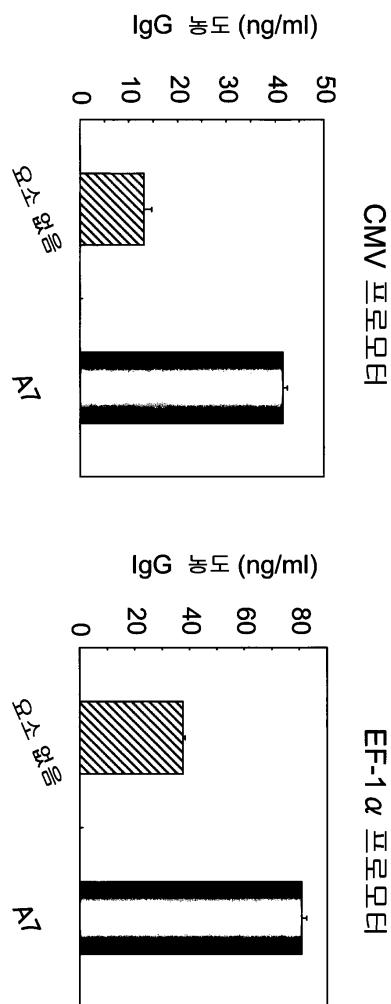
도면4



도면5



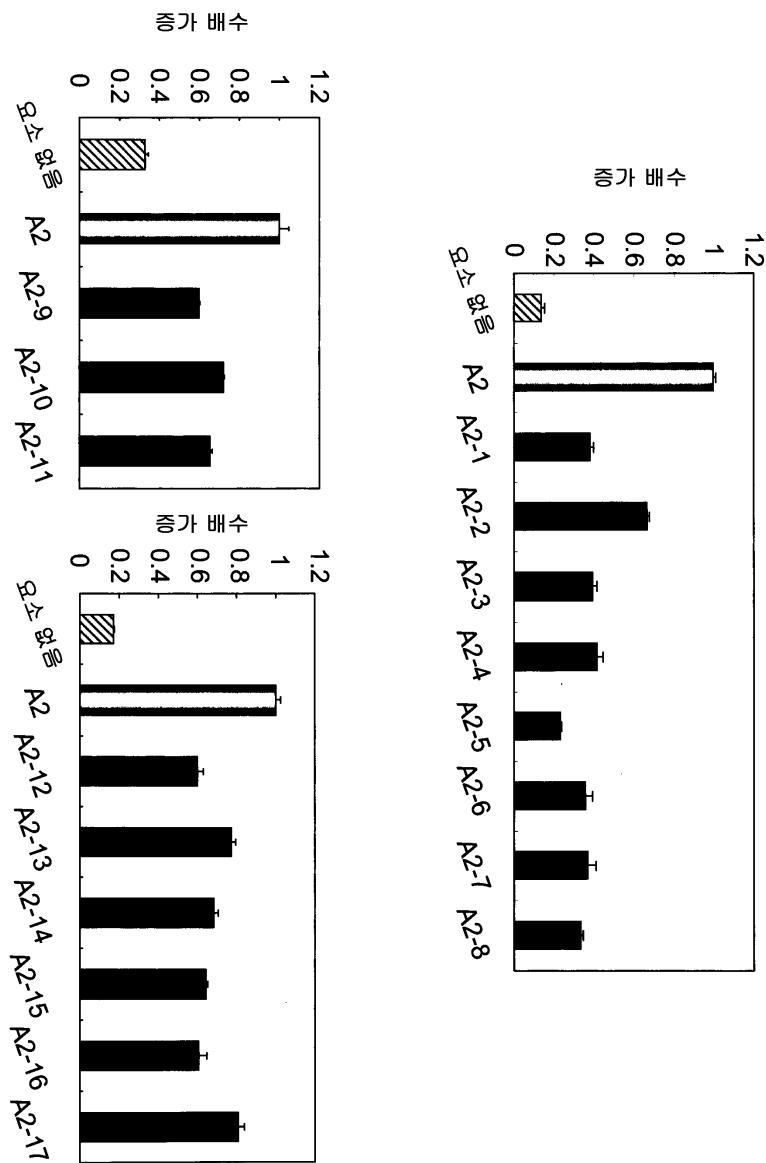
도면6



## 도면7

ID	출발 위치	종료 위치	길이 (bp)
A2	80966429	80974878	8450
A2-1	80966429	80969428	3000
A2-2	80969229	80972228	3000
A2-3	80971829	80974878	3050
A2-4	80967129	80969128	2000
A2-5	80967129	80968628	1500
A2-6	80967129	80970128	3000
A2-7	80968429	80971428	3000
A2-8	80970429	80973428	3000
A2-9	80966429	80970128	3700
A2-10	80968429	80972228	3800
A2-11	80969229	80973428	4200
A2-12	80967129	80972228	5100
A2-13	80968429	80973428	5000
A2-14	80969229	80974878	5650
A2-15	80966429	80972228	5800
A2-16	80967129	80973428	6300
A2-17	80968429	80974878	6450

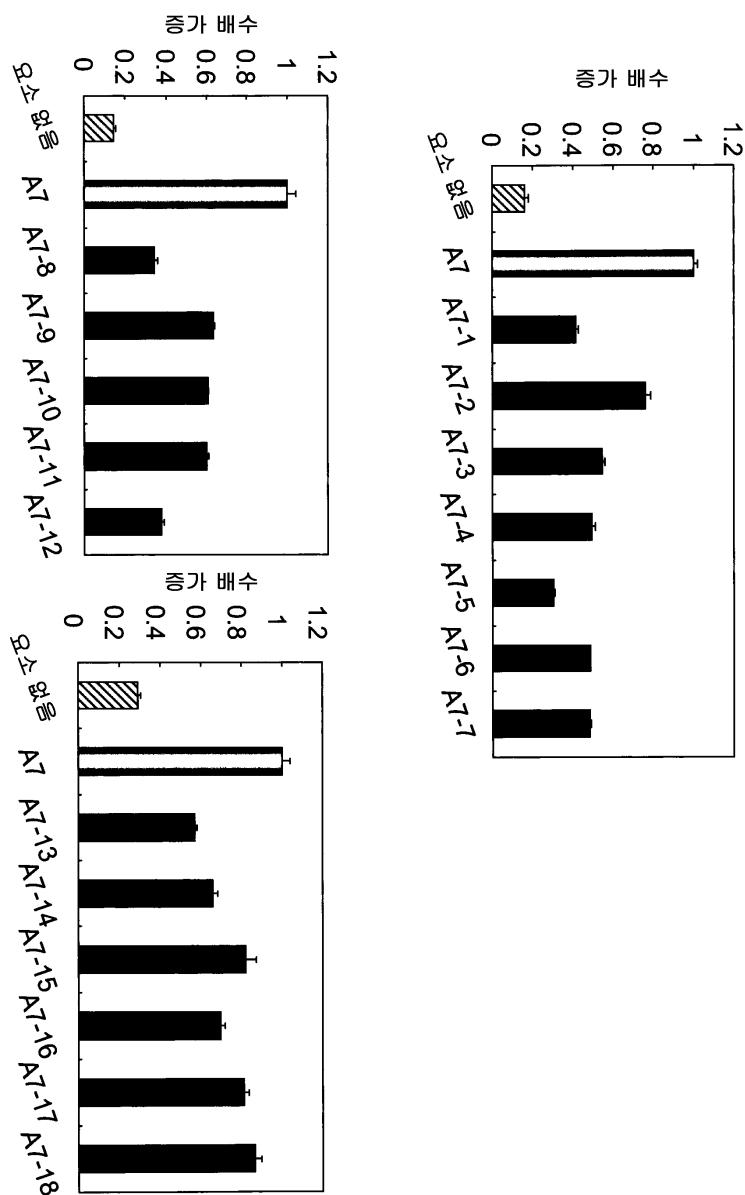
## 도면8



## 도면9

ID	출발 위치	종료 위치	길이 (bp)
A7	88992123	89000542	8420
A7-1	88992723	88995722	3000
A7-2	88995723	89000542	4820
A7-3	88997523	89000542	3020
A7-4	88995523	88998522	3000
A7-5	88993623	88996622	3000
A7-6	88996523	88999522	3000
A7-7	88994523	88997522	3000
A7-8	88992123	88995722	3600
A7-9	88993623	88997522	3900
A7-10	88994523	88998522	4000
A7-11	88995523	88999522	4000
A7-12	88996523	89000542	4020
A7-13	88992123	88997522	5400
A7-14	88993623	88998522	4900
A7-15	88994523	88999522	5000
A7-16	88995523	89000542	5020
A7-17	88992123	88998522	6400
A7-18	88993623	88999522	5900

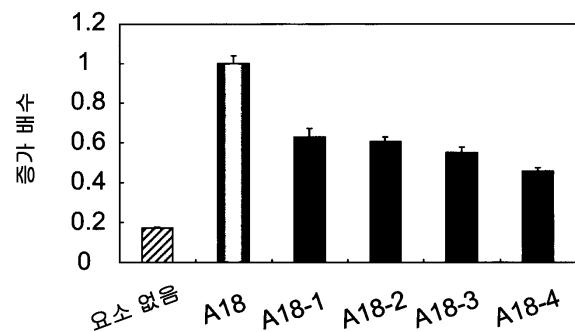
도면10



도면11

ID	출발 위치	종료 위치	길이 (bp)
A18	111275976	111284450	8475
A18-1	111275976	111281015	5040
A18-2	111276976	111281977	5002
A18-3	111277976	111282975	5000
A18-4	111278975	111282975	4001

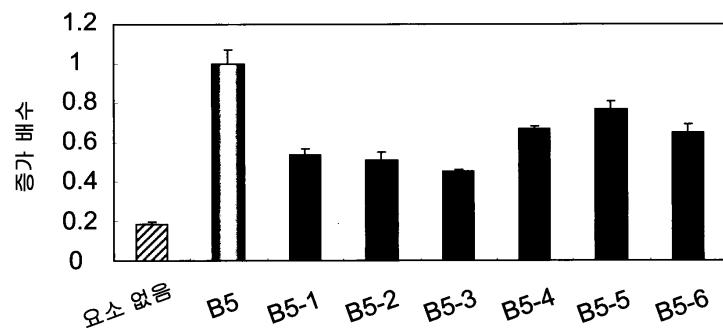
도면12



도면13

ID	출발 위치	종료 위치	길이 (bp)
B5	143034684	143043084	8401
B5-1	143034684	143038684	4001
B5-2	143034684	143037883	3200
B5-3	143037174	143040284	3111
B5-4	143040056	143043084	3029
B5-5	143035584	143038684	3101
B5-6	143038684	143041683	3000

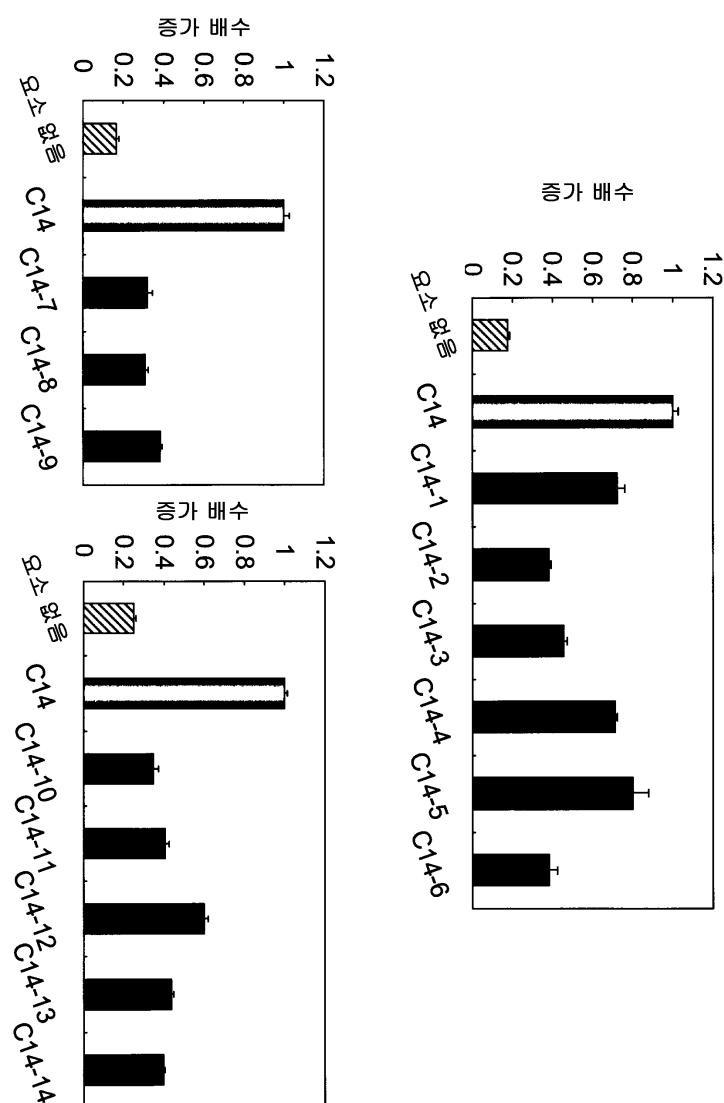
도면14



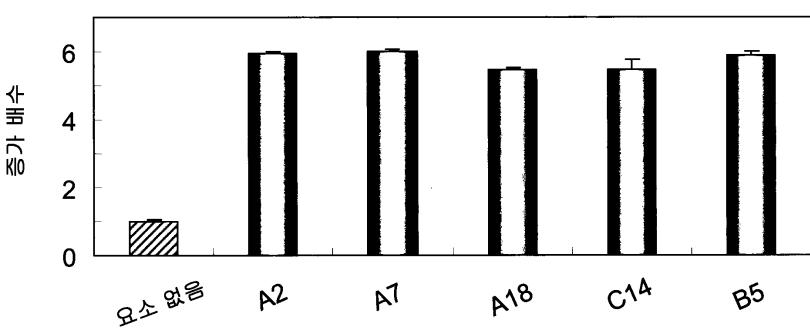
**도면15**

ID	출발 위치	종료 위치	길이 (bp)
C14	46089056	46097482	8427
C14-1	46090015	46093070	3056
C14-2	46091042	46094069	3028
C14-3	46093075	46096174	3100
C14-4	46090015	46097196	7182
C14-5	46090015	46095066	5052
C14-6	46093994	46097196	3203
C14-7	46090015	46094069	4055
C14-8	46092049	46096174	4126
C14-9	46093075	46097196	4122
C14-10	46089056	46094069	5014
C14-11	46091042	46096174	5133
C14-12	46092049	46097196	5148
C14-13	46090015	46096174	6160
C14-14	46091042	46097196	6155

도면16



도면17



도면18

DNA 요소 A2의 전체 길이 서열을 기준으로 하는 출발점 및 종료점			DNA 요소 A7의 전체 길이 서열을 기준으로 하는 출발점 및 종료점			DNA 요소 A18의 전체 길이 서열을 기준으로 하는 출발점 및 종료점		
A2	출발점	종료점	A7	출발점	종료점	A18	출발점	종료점
A2-1	1	8450	A7-1	1	8420	A18-1	1	5040
A2-2	2801	5800	A7-2	3601	8420	A18-2	1001	6002
A2-3	5401	8450	A7-3	5401	8420	A18-3	2001	7000
A2-4	701	2700	A7-4	3401	6400	A18-4	3000	7000
A2-5	701	2200	A7-5	1501	4500			
A2-6	701	3700	A7-6	4401	7400			
A2-7	2001	5000	A7-7	2401	5400			
A2-8	4001	7000	A7-8	1	3600			
A2-9	1	3700	A7-9	1501	5400			
A2-10	2001	5800	A7-10	2401	6400			
A2-11	2801	7000	A7-11	3401	7400			
A2-12	701	5800	A7-12	4401	8420			
A2-13	2001	7000	A7-13	1	5400			
A2-14	2801	8450	A7-14	1501	6400			
A2-15	1	5800	A7-15	2401	7400			
A2-16	701	7000	A7-16	3401	8420			
A2-17	2001	8450	A7-17	1	6400			
A7-18	1501	7400						

## 도면19

DNA 요소 B5의 전체 길이 서열을 기준으로 하는 출발점 및 종료점			DNA 요소 C14의 전체 길이 서열을 기준으로 하는 출발점 및 종료점		
	출발점	종료점		출발점	종료점
<b>B5</b>	1	8401	<b>C14</b>	1	8427
<b>B5-1</b>	1	4001	<b>C14-1</b>	960	4015
<b>B5-2</b>	1	3200	<b>C14-2</b>	1987	5014
<b>B5-3</b>	2491	5601	<b>C14-3</b>	4020	7119
<b>B5-4</b>	5373	8401	<b>C14-4</b>	960	8141
<b>B5-5</b>	901	4001	<b>C14-5</b>	960	6011
<b>B5-6</b>	4001	7000	<b>C14-6</b>	4939	8141
			<b>C14-7</b>	960	5014
			<b>C14-8</b>	2994	7119
			<b>C14-9</b>	4020	8141
			<b>C14-10</b>	1	5014
			<b>C14-11</b>	1987	7119
			<b>C14-12</b>	2994	8141
			<b>C14-13</b>	960	7119
			<b>C14-14</b>	1987	8141

## 서열 목록

<110>	DAIICHI SANKYO COMPANY, LIMITED					
<120>	DNA elements that have an effect on increasing transgene expression					
<130>	DSPCT-FP1124					
<150>	JP2010-154782					
<151>	2010-07-07					
<160>	7					
<170>	PatentIn version 3.4					
<210>	1					
<211>	8450					
<212>	DNA					
<213>	Homo sapiens					
<400>	1					
attttgcttg	aaaggatagc	atcaaggaag	tgaaatgaca	acccacagaa	tgagagataa	60
tttttgcaaa	tcatgtatct	gataagggac	ctgttagtcag	aatatgcaaa	gaacccttac	120
aattcaataaa	gacaacccaa	tttaaaaaca	ggcaaaggat	gtgaataggc	atttctccaa	180
agatacggaa	aaacggccaa	taagcacata	aaaagatgct	caaaatcatt	tgccatttgg	240

gaaatgcaat caaaaccaca atgaggtatc acttcacgcc cattagggtg gctatagatc	300
agaagaatcg ataacatgtg ttggcaagca catggaaaca ctgaagtct tacacactgc	360
tggtaggaat gtaaaatggt gcagccactg tggaaaacag tttccaatt tctcaaaatg	420
ttaaacacag ttatcataca cccaagcaat tctactcta ggtatatacc caagagaat	480
gaaaacatat gtcttcacca gaacttgctg ttcacagcag cattatgcat aatagaccaa	540
aagtggaaac aactcaactg cccatcaact ggtgaatgga taagtaaaat gtgtatgtaac	600
cagtcattgg actgtcattc attaataaaa agaacaaggt actgattcat gttctaacat	660
gagtgaatct taaaaacact atgctaaatt aaagaagcca gtcacaaaag gccgtgtatt	720
gcatgattt atatatacat gaactttat atatatataa ttatatatat tatataaat	780
tttatataa taaatttcta tatataaata tataaaatca tatatatgtat atatatttt	840
tcatatacat catatatatt tacaaaaatt atatatcata tatcatatga tatatgagat	900
atatatcatg atatatatga tatatgtat atatcatatg agatatatga tatcatgaga	960
tatatgtat catatgtat atatgtatata gatatcatat gatatatata taatatatat	1020
atgatagata tattatata gatagatatg atagatatca tattatata gatagatatg	1080
atagatatca tattatata gatagatata gatatcatat tatatatgtat agatatgata	1140
gatatcatat tatatatgtat agatatgata gatatcatat tatatatgtat agatatgata	1200
gatatcatat tatatatgtat agatatgata gatatcatat tatatatgtat agatatgata	1260
gatatcatat tatatatgtat agatatgata gatatcatat tatatatgtat agatatgata	1320
gatatcatat tatatatgtat agatatgata gatatcatat tatatatgtat agatatgata	1380
gatatcatat tatatatgtat agatatgata gatatcatat tatatatgtat agatatgata	1440
gatatcatat tatatatgtat atcatatata taccacatac atcatatata catcatatat	1500
acatcatata tattatatacat atatatgaac tttccagaat aggtatatca ataaagacag	1560
gaagtataca agtgggtgcc acagcctgag aggagcaggg aatgggtgagt gactgctaatt	1620
ggatatggca ctttttgg ggggtgatga aatgttctg gtcagacaat ggcaattaca	1680
aaactgtata cacacgaaaa accaaagaat cacacacttt aaaaggagg atttagctcg	1740
gcatggtgcc atgcgcctgt actcccagtt actcgggagg ctgaaggcagg actgcttgc	1800
gcccaggact tcaaggctgc agcgagctat gatcgctcca ctgcactcca acaaggatga	1860
cagtgcgaga cccgtttct aaataataat aataataata ataataaaata acccaaggta	1920
cccgatccac atgcaaaacc actggtaaac ataaattatc tccaagtaat ctggaaagaa	1980
aatgagcaca taagacgtct tctaaaaaca cacatataatt tctttacatg ttacattaa	2040
cgtaaaaatc agctatgcag aagttacatg aacatttat gttggaaagg taaatgacta	2100



gagtttgttataaaagtgtatctatcctctaataagatcgagaaacagaaccaagcgact	3960
tcactataatttgatcttagagaagttcttactcacaataggtaaatgaaaggcacatact aaccagcaataaaaacaacaatatacaagtgtcattcacacatgcaaaaaaacagacaaaat cccaactctgtgttctaacaatcgcaaaaacctcaactacaataattgaaatgacca aatgtttgga ctgaaaagcaatgccttggtagcctagccatgcctaactcaaataacaga accatctcgatgttaaaatctcacagatcaagctgtatgtctcggtcaagacttcg ccaaaaagca gtgagcacacacttaagaggaaaaaatctacctcagcctcctaaatgca atcatctcta caccgatgtcaggccccaaagcttcaacgtttctgctgga caacgcagta	4020 4080 4140 4200 4260 4320 4380
gaaagctgac aagcaggtgg cttcccaca ctgactgaac cacccatg cccatgtcca ttcattttct tgcccacccatgtgctata acagacccatggctcaggg cactcttcc ttcctgactgcctcaacttatagactttgtacttttagtgcaaaaattatctgcagaaa tccacactga aaaccaagcttgagaaaggc agcaataacc aacattttcaagaagaac aaggtaata tcaagccat cagattcaaa tagcaagcat ggtgaaaat gaaagattga aaggcttgatgcctttaatgtattttatccatttattacaatt aagctcaact tgctcaactggccttaatc agcttccag gtcctgctca gacttgccta ggacatggga	4440 4500 4560 4620 4680 4740 4800
atgaaaagaac ctatacattt atggaccaat ctacctaactaacttgtca agtgttcctg catcaagcag aagaacatc agtggaaactg atacaggaaat taacccttg ttaatccata aaacttaaaggagcggatc caatctctg gcttcctgg gccacgctgg aagaagaatt gtctgcgcc acacataaaa tacacgaaca ctaataatag ctgctaagctttaaaaaat tgcaaaaaag gaaaatcttaatttttgttggatgtggagcct cactctgtca cccaggccgg agtgcagtgg caccatcttgcctactgca acctctgcctcctgggttca agccattctc ctgcctcagccctccgatgatgtggatgttggatgtccaccatgcc	4860 4920 4980 5040 5100 5160 5220
cagctaattttcgatgtttttt agtagagacgggttcaccatgtggccaggctggctc aaactcctgacctcaggatgatccacccaccatcgccctccc aaagtgtggattacaggat gtgagccaccgtgcccccccaatgttttaa gaacgtttac gaatttgtat tggccacat tcaaaggcttcacaggctgc atgcagccctgcaggccgg ttggacaagcttggattaga gaaatctaca gagacaaact agtgcacttag tagccctctg atagctcatg atttgcaaga aacttagat gactatgtttaaagaccaca aacatcaatttaactgaatgttccgc cactggaatg aggaagctgatcagaaactcaggactctaa gaaaggcgtatgtcatctg	5280 5340 5400 5460 5520 5580 5640

aactgttcgg aattataaac tcctctaaac atgttcaaa gccagaacct gtaggagttg	5700
ttctgataca cggttaaaaa gagggatgac aaagtgtctg tcccccacac tggtcaaagg	5760
gacaggtcat tggtatgctg gcaatgcagg ctgctgaaaa gaatgtatct gtcaaaagta	5820
atcaaagtaa tgacccaga aggctccaga aacagactgg taaattcagg ttgcttcag	5880
acttccacaa tgctggcaca caagggaaa gacaaaacta acatttacag agcattatat	5940
ttgatattac attaatccc cattaaaag atactattc cggttact agtaaaaaag	6000
ttgatttc aaaggtaaa ttatthaaca ccaaggtaaa agggtaagtt ggagagacca	6060
gattcaaacc cagtctgaca taaaacatg tgtttcccc ccacatgta tcctgcta	6120
aacctcaaat ctaaaaactg acttgcccta caccttgagc cccatctac aaactctccc	6180
tgacgttatt aattcagctg tcactgtgca cctacaacgt gccagacacc atactcctca	6240
acactctgta ggcacagaag gaacagataa aaatccctac ctcatagat attattctag	6300
gggtaacaca ggttaataaa acattaaaat agtttcaca tagtagcaaa ttccatata	6360
caaataaaaa cagaagaagg aatagcaaatt gaggagatg ccctctaaa catgtgctg	6420
aggaaaggcc tccctgagaa agatatcatt tacccaaaa ataaaaaagc aagtaataga	6480
aaaaacaggt aaaagggttt ctagacactt aaacctgcca cattgagaac tcagggttct	6540
gatgcaaaac ctcgctgcat agaatgcatt aacttattt tatacatat aacaaacaaa	6600
ctctacttaa gaactgtgtt ctaaaggaag gagcatat caggaaggca attttggtc	6660
agagtagaca cactaaaaa ctaaacctat tgaaagacca agaacaactg aaagtcttg	6720
ctttgtcaga ttttgacca aaaggaaaat taaagaaaca caccgtgcc atccaatgat	6780
ttcaccaagg aattnaaga gagaaaatcc tacttcttc tcacccagta gccagtgaaa	6840
tgactgagca aattcacaag ttcaactggg ctgcttcat gtaacacagg gacaacacat	6900
gacagacaca gtggAACCT acaggttgcc tagtatttga aagactgtga agaggaggag	6960
atgtcaaaat tcaaagtctt aatgtatgtt gtttaagta tggtcagcaa tttcaccact	7020
cagtagtaaa gccagctaca gttgaaagca atcagaatt tgaggggtgt gaaataagca	7080
gaagcacaga agttaaggat ttgtattttt cccacattt ccactttttt ttatactgct	7140
gagaaaaaac aaatttaata gtttctgt gtataagaga gacacattca ctttatgtca	7200
cagtaagagt cactcaattt taatacaact atctcaatgt ataaattaac attctcccc	7260
ctgcccacac atagtaagtc tcttatgtt gttgctgat gagaagcaaa agttgccct	7320
acaattctct tcctgcattt taatataaac aatcatcagt ctttcttca tagagtgcag	7380
tgtggcact atcatcagaa tgtaccagca ctgggtgtgc aaagttaca aagattagca	7440
agagcaaaag ttttgagatt ttgaaattc atgctgctgc aaagaagttt gtaaaaactc	7500

actaccata gaggaccaca cagaaactca ggcatgaagt tatatggctg tgtgagtgg	7560
ttgggagaag gaacgaaaag cactccacc aacctatatg cctgagcaa ttaatgcaa	7620
acctcagaag ctacaaaaaa gtttatctac ctaaattaaa attggtgtcc acagcagtag	7680
ccagaaaaat gcctgcgaag cgcaaaagtgg taaatatgg agggtctgtatgg	7740

tctctgttaa acaaatatgta aatgaatggg tgtggctgtg ttccaataaa acttcattta 7800  
taaaaagagg cagcatggta catccagtca gcaagctata atgtaccaac ccccggtcta 7860  
acactaacca aataacctt aataagccaa agaaaactgtg tcctcttagg ccggaagcgg 7920  
tggctcacac ctataatccc agcattttgg gagggccgagg cggggagatc acctgaggtc 7980  
aggagtttga gaccatctg gccaacatgg tgaaaacccta tttctactaa aaatacaaaa 8040  
attagccagg cgtgctggcg ggcgcctgta atgcacaacta ctggggagggc tgaagcacga 8100  
gaatcgctg aacccaggag gcagaggttg cagcgcgcct agatcacgcc attgcactcc 8160

agcctggcaca acaagagaga aactccgtct caaaaaaaaaaa aaaggaaata aaagtataca 8220  
aagtgaaaac aaagaatatta aactgcctt atttgccagt gacattactg tctatgcaca 8280  
aaattccaaa aatctacaaa aaagcttcta gtactaaaaa tgagtttagc aaggttgtag 8340  
aatccaaggt cagcatataa cataaaatca ctttcctata tactagcaat caccaactgg 8400  
aaattgagaa gtatcattca caacagtacc acaaacatga aataaaatgt 8450

<210> 2

<211> 8420

212 <212> DNA

<213> Homo sapiens

<400> 2

tcttagtatg gtaaacctt tgaagtatgt tcaaatacgt atggaaagag agaaaaggga 60

gagaagcaac ataagaatac tcttttaagg aattttatat agagagaaac agaggaatca	120
gttgatagtt ggaatttatt ttaaagaaaa tgggttattt taaagaaaaa aggtattaca	180
acatgttgc actattgtgg gaataatcaa gttgagacag aaaatttattt tttaggaag	240
agtctaattg ctgaagtgaa agagaatgaa tgagaccctg tgcataagtg tgatcagata	300
ggagcatgta cagctcaagt aagaacagga agaaagagac aataaacatg tacagatagg	360
atgggcttgt cgatgtggtg gtgaaaagac atgcgagtttta ttactgatta cttctatttc	420
cccaatgtaaa taggaagcca ggttcataaa cccaaatgaa gaggagcgag gcagttttgg	480

aagttcagga aaagtataat gtgtaaaaat atgtaaagta gaattaccag ggagtgatgaa 540  
gatacatttc caattaagga tgaagaattt aaagtggcc cagccaaatccctgtttg 600

cttcagctac atcagctgca taggttcagg cacagaatac atggaacatt gtatthaat	660
agggcctgga ttttacaaaa gtaacacaat gaagaagaga gatgcaaggc tatttgagg	720
tgttgtgg agagattgta aaatattgc taagtaagaa gggactgca aattttatg	780
gtataaagga atgagaaaaa gtgtaaatac agtgggtca aagaatgtt ggagccaagg	840
cactagaggc aattagctga aaatgttaggt gattattggt gagtgacatg gttaaatga	900
aaagtataga aggtaatac tatccatcat gaaaagtctt agggtaacaac taagatctga	960
gtagctgaag tagaatgaaa gtagaatgga cctttccata tccagccagg ttcagtgaca	1020
gaaggtagg aaacaatta taaaccactt gagagaacat atcccctaag ttgttttgc	1080
tattttctt tcagcatata ttgttggaa tgccaaactat gttcaggta attaatatgg	1140
gcttctaaa taagggctcc agcactggat aatcctgcca tttattttga tacattccat	1200
cctgctgctc agatctattt gcatctacag gatgtttt gagaagatgg gcattcacat	1260
ccctatgtcc tagcaaattt ccaactcaga aaaccacatt aggcttctct atatatcttc	1320
caactattc aatggaaaat acaattctct gatttcttcc tatgatattt atcaaagaga	1380
atggtgctcg ccagttctag ggtggggaa ctcaatacaa atcaccaacc ttttagatgac	1440
accctgtctt caaagtgttt tcaaagtctg gcagaaaaaa agtacccagt ggctataaga	1500
ccacccagga gttcagtc catgattctaa tagcagatca ctggaaatgtt attggctgt	1560
gagttcattt tactttctc ttcttggtca catgttaccg cccttgacc ctgcacgttc	1620
tctttccag acttacaaag catgttctct tgaattcggtt ctcttttaa attcacacag	1680
tcttaatgt tcttcttca caagagtctt tcactttac aattcagttc aagtcatcca	1740
catgcttatt atgagcaagg gttctggact tagggaaaaa gggataaaaa agatgaatga	1800
aatgtgatcc ctgcagtcca agagcttgct gtggaaaaagg aagtttggt tacattgcct	1860
ccctaattccc ttggcttaggc cagaacagaa tattgtctaa aacctcctca cgtcagcagt	1920
cctctgggtt ggtgactgga agttagattt aaacaaaaat ataattgaca cataataatt	1980
gtgcatactt atagggtaca atctgatgtt tcgatgttgc tttaatggg tgcattgtgt	2040
aatgatcaa ttgaggtaat ttatccacca ccttgaagag agattttca atattctcat	2100
tgcgaagaag caggaatttt tagcagacaa ctgagatgct tcttggcac actaagtcat	2160
tctgacgatg gatttacata acttgggtt tttttgtgt gtgtgtttt gagacagagt	2220
cttactttgt cgactaggct gaagtgcagt ggcacaatct cggctcaactg caacccctcac	2280
ctccgggtt caaacgattc tcctgcctca gcctcctgag tagctggat tacaggtgca	2340
tgcaactagg cctggctaat ttttatattt ttaatacaga tggatttca ccatgttggc	2400
cctgctgggtg tcaaattcgt ggcctcaagt gatctaccag ctgcggcctc ccaaagtgca	2460

gggattacag gtgtgagaca ccaaggctgg tacattaca tttcttatct ggatttcc	2520
tttagtaagt gctaaggaat cctacttccc ccaatattt ttcctattc aatgttttag	2580
catgtatcat gtactactt tgccagacatt tgatttccc cttgtttac tgtaaagtagat	2640
atttttagt ccttgata agaagtattc taaaatctgc ctgcaaccta tcttctgac	2700
tctgcattt agggataat tctctgtgt ggaatgaaaa aaaaaacaga gcctgtggag	2760
tcagagatct cattcaaata tatagttac ccttagaata aatctgagtg acaggttagta	2820
tagtataata ataagtataa agctatggtt aaggaaaact caacaacctt atctgttaat	2880
tgggatgaca acagcctacg taaaaaaaaat gtgaaggtaa atgagataat gtaaggctga	2940
tacttagtaa gcaatttaaa aacacccaaa aaactattgc catgattact ctacttactc	3000
tatttctcta tgctccaggc aaatgaacta ctaatgaccc aggggtcctt cccattctc	3060
ttcttcacaa ggaatattc tctctgtgt tgctgttat taaaatctac tgccctttt	3120
agaagcctt ccagatcatc ccatggccaa gaacgatcgc tgctcctct tcttacata	3180
cagatgttt tccctgctt gacaattatt tttgtcaat tatttcctt ttgattgtgt	3240
tttaatgtc ccccccaccc cacaatttc cagactgtt gtcacgag agaggagacc	3300
atcatctcg tgctcacgt tggatgacca gtatctgag gagtggtgt tacataatta	3360
catcaggcac tcaataaaaaa ttgtatgaat aaacactgga ttttaaggca ggtatcatat	3420
cttacatagc atatcatatc ttacattttt tgcctcac ataaatacca cagagtgaa	3480
tatatgacag ataaggtcat ttctttgtat aagtacatag tccagtcga aacagatag	3540
ccaaaaaaaaa acaaaactgg agtaaacaag atgaattttt ttaatagagg cactgttata	3600
gtttcctagg actgccagaa caaatcacct caaacttagt ggctgaaaac aacaaaaatt	3660
tattgtctca cagttataga tggatgaaatg ataaaattaa ggtgtcagtggatgggttc	3720
cttctggggg ctgtggaga gaatctgtcc caagccttca cactgtaaag tacagtactg	3780
gaggatagg acttcaactt gctctatctc agatagagag gagccattttt ttgtgaattt	3840
agaagagggg tggatgaaat ccataataag cacataaaaaa ctggctgg tcataggaga	3900
agtaacatgt ttccagctt agtaaaaaac aaattgaagt ggcctataaa aaggtacaga	3960
gtacgacaga atgaaaaata aatgaacaag aatacagaga ggatgtggta aattatcatg	4020
ttccctaat atgttattgg acactaaatg gtattagaat tatttatcaa taataattct	4080
aaactgttgc aatgttgc aatgttgc atatattaag tggatgttata tgagaatgtc cagggcatc	4140
tcatttctgtt ccaatgggag aaacatttc gtttgagacc tccgtgaata atacagtctt	4200

ttagtttagga gagctgcatt ttgagtggc caggcagaat ggcgatctt caccacaca	4260
aacactaaga tagagagaga cagagacaga gacagagaca gcagagagag acagagaaag	4320
gaagtacagg tactcagata gagataaagcc atttcttgc attaagaaat aaagttagaat	4380
ccattggagg gaaataaaaac tgcctcagga acagagttt ttcacatata catgcaggtt	4440
aacacacact gcttgatact tactgtggac ttgaaaatt atgaatgtgt gtgtgtgtgt	4500
gtgtgtacat tcagccctcc atatccatgg atttgcatt cacagattca accaaccatg	4560
aattaaaaac atttggaaat aacaacatt aaaatataac aatacaacaa taaaaataat	4620
acaataaaaa aatatgtgt aacaactgtt tacatagcat gtatgtgt ttaagtagta	4680
taaatctaga gattacttaa tgtataccag aggtgcata ggctatatgc aaatactatg	4740
ccactttaaa ctgataagaa cagataactaa acttcatctt agccaaaagt cagagaaaca	4800
atataactat gccatttac ataaggact tgagctgagc atcctcagat ttcagtatct	4860
ttggagttcc tggaaacaat tccttggttt atatatatat atgtgtgtgt atatatatat	4920
atatatatac acacatatac atatatatat atatatgata gctactgagt gacaggtgat	4980
attataccat accacttgc actcagtagc tgtatatgca tatgtatata tatacatata	5040
catatatgtg tgtatgtgta tgtgtgtgtg tgtgtgtgtg tgtgtgtatg ctgtcttcc	5100
tcggatcac aggaaattgg agatatatat attctttca gtacaaaaaa aattgaacac	5160
agatgggtat ggtaccagaa cagaaggtaa agacacatga aaaaaatttg caacaacatg	5220
aatgaaactg gagatcatta ttggggaga aataatccag gcacagaaaa acaagcattt	5280
tattatTTTtta ggtggaaagac aaacatTTta tttaggtga aataatccag gcacagaaaa	5340
acaaacatttgc catgttctca ttatTTgtg ggatgtaaaa atcaaaacaa tagaacgtat	5400
ggaggttagac agcagaagga tagttacca aggctgaaaa gggtgtgtg ggcttgggg	5460
gtgaggtgg gatggttatt gggtacaaaa aatagttaga aagaataat aatatctgt	5520
atTTaatagc acaacagggtt gactatagtc aaaataacat aattgtacaa ttAAatatg	5580
aaatTTaaata tatatacaag actagaacac caagttgaat gactccagct tgcgaaaccc	5640
acattgtatca ccatgcttc cccaaaggaa gctgtacaat gtctggctcg tccagaaccc	5700
catcattttt cactagcaat ctattgtcca taatcatgtt taaattataa gcatTTTaaa	5760
ggtacaaata ttTTTaaa aacaataat tatttaatTC gcctttaaa agctttttaa	5820
aaacgttttt aaaaactttt taaagtccct gaggactatt ttctttaaag tgctcagttt	5880
cagagctcca tatattggc tatgatagcc ttacctgatt ctggccaaga atcttagtgc	5940

cagaaaatgc aaatacaaag taagcaactg aaaaataaaac aaataagttt gaggtatgct	6000
acctgttcaa atatgaccta gcgcaaacac ctatgccact tgcttatgaa atcatatatgg	6060
tttcgggtgt gcagtttga ctgaatgagg gagttacgc tggaccacaa gggggccct	6120
ctgtcaataa cgtactccat ttgtgttata agtcaaaaat gaaatgaaag agaaaagaaa	6180
catcgatgac cccaagtctc ttaattgaa tggaggtaaa agggaaacaa cgaatgagaa	6240
aagtactctg ccctttaag aatcttgcatt tcacattctt gatgaagttt ttttcctcc	6300
tctcactgtat tcccatttca ctctattaca tagcacccgtt ttccccagga gctcctgaat	6360
gaaggacatc actcagctgt gttaagtatc tggaacaata aatatactat tttcaatgtc	6420
taggctatgg gtattccctt ttactgaagg tatgacatat agtgccag gcctgactaa	6480
attaatagta ataataatata ataatggcaa atttttattt tattaagttt cttggcttga	6540
ctttagaaaa tagcaacattt catctgaaat gccccctctt acacttatgtt ctaaggacaa	6600
atcccacata caccacagat aacttcattt tacatgtttt attctgttac caaactaaat	6660
ttttatcata tagtctgttgc ttcaactgaa tcttcagttt ttctcaacat accatgtaaa	6720
gcattaagca cagttccaaac acagagcaaa tgagcaataa ctgttagttt ttataacattt	6780
attatgtgtt ttcaactgcat taaaccactg gtctgatacc tagcccaaca ttcttataaa	6840
ccacataatc cagttgaata atatatgata atataataaa atggcgataa gtgctaaata	6900
tccagataga aacacagatg gaatcagaca gctttccaa gaaatagaga aaatagtaga	6960
taggcgatct aggcttaagc actctaaagca gaagcttactt tatcacagga tatcttggca	7020
atctgtggca cgtgaaccctt ttcttctgg agtctggaaat tttttttttt ctctacttt	7080
ctccctatctt agagacttagt ttgttccctt tttttttttt ttctcaacat accatgtaaa	7140
accttctgaa aggactactt tttaattttataatataat atttaaaaata catacttttta	7200
tatataatat atattnaat atataatatt taaattaata tatattttaa tatataatat	7260
ttaaattat atatattttaa ataaataat ttatattttaa atatataata attaaaatat	7320
atttttatg aacagagatg aaaggattttt tttttttttt tttttttttt tttttttttt	7380
aatcctttctt tttttttttt tttttttttt ggagccctt ttaccagctt gccccctcag	7440
agataagctg ttccctactt tttttttttt tttttttttt tttttttttt tttttttttt	7500
ttttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	7560
tgtttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	7620
aacttctcaa actatgtttt gaccggcttag cataatgtttt cagtctttt ggagaatgcc	7680
ccttggaaactt gttttttttt acacaacttc ctcccttctt ttgactttcc ttgtctggaa	7740
gggaagaaca ggaagaggac agatcaaattt actcaagagg aaggacaaga aataaggaac	7800

caaattatca acaattggag aaagaaaagct gatgtcagta tcatttcata tatgattatg	7860
ttagtgcag gtggataagg caatccgtt gaatagcata ctttcctgc tactcctgaa	7920
gggttaaagag gtctttctt tacaaggccg tcctagctag taatcttaca ggtgaaaaaa	7980
gcttgtttc atgttatttc ttagtaactc aaaataccctc taaagtatata catatgtga	8040

aagtactaca gtcacagtgc tgagaaaagg agttaataag acaatgtata taaaaacact	8100
tggctcagcc cctggctctg tggttgataa atattaagtt agtattcatt attattataa	8160
tttccaaaga gtcattaaa agatatagaa gaagggaggc agcaataaca ctaagagaaa	8220
attccattat ctccaactat ttatcctcta gcccaaata attgccatta gaaagagcaa	8280
ctttaacaaa aatittaagt tgcaatagat gttcaacttt aaatccatcc cagaaaaatt	8340
tctaacaaa ggagcataga agatttgatc ttatcccata agtagtatag acttaattgt	8400
gagaacaaaa taaaaacttg	8420

<210> 3	
<211> 8475	
<212> DNA	
<213> Homo sapiens	
<400> 3	
gcataacttg taagaaatgg agtgaggctc cagttcaaac tggcttctgt atgacttcaa	60
agccaaagtc agcaacttag aaggcaaaaa ttataattta gttggcaaat acgagaaaag	120
gtcagaaaca catgaaatga agctcaatag gaacacttac agggtagcag ggttagtagcc	180
tagggaaaaa agtcagacac taaaattgtt taaataggtt agttcaaggg acaggtaag	240
accttagtgg gtaagaagcc aatcagcaga cgaactgcaa gcaagcactg tctctttc	300
ccttctgtct cctcttgtag taactgacca caattaaggc tgcctaggg aataatgaag	360

taatccctt attatcgca atggctgtat ccagtccag gcaccacaga caacttggtg	420
ttcagagaag atccttcaag atgaacaaag ggtcaaaata aaaaattcttta gaagagagaa	480
gactgatcac aatttatgt aaggcttggaa aggaactgat ctctacccctt cttacatct	540
caagaacttc ctcagattca ttggatgtt agtgtgtgt agtcttagtag aaaaatgaat	600
ttttgtttct taacttggat atgtgattag gatgttaata attaagtctg ggctaatatt	660
gaaggatct tatgtatggc ttcttaaagc attgatcaca aagactgcat gttcataaac	720
tgagctgcac ttgttaggat tctagatgtt tggaaatttct tggtttattt tggctcaga	780

tttctagaca aattttctca aattccattt tcacttttg acatatcatg agtgactcaa	840
atgtttgccctt tagtgcga aaacacccag cattaggaat aggcacataa acataataact	900

tcaagctca gatccaatc caattataaa gtgtttaag gctgtgctga tagttttct	960
gagtagaatt cctacaacta tgggttgct tataataaaa tgtcactct atatgtacg	1020
ccttattaa aactcgaaat gtgtaaatgc taataaagaa aatatgtcct cctgtaaacc	1080
aagcttaggac cgattacatg ttcaatgtac tgacagatac aatcacctat attaggagca	1140
atcagcactt ccttacaaac taacaacttg agatgtatg ttccattgg ctatgaagat	1200
tttcttatt tactcagaat agtctgttagg atctgccagc tgcccgtat tataccagct	1260
gcacccaatg atcacagtga acattatTTT acattctaaa taactggtgc aaggtgagcc	1320
atggtttct gagtttccta tcaccttgtt gttcaggc ctcaaagtta aatttgtaaa	1380
gctgcgttt caggcaaaac taacaaaatt agcatctaat caataaccat actatgtcca	1440
cccatatcct ataacacaga agtagggaa gagtgagaaa ggtgaaatgt gagaataga	1500
ggcccaaaaa gaaagttta tcacaggaat atctagatgt ctctggat tgtctgttaa	1560
agagctgtga cactcatata aatgcagaat tactctctt ctccctgtt ggtagaagg	1620
ccaagggtgc catggtataa ctaccaaaca tatataaag ctggcagga aaaatggta	1680
cttcagaat ttataatct gatatcaat aggtcaagaa atataataaa actagttct	1740
ttggttcct tagaaacctg gaaaactta aattagaac tttagaaatgt ttaatcaga	1800
cttttagtt aaaaaaggaa attttagttc ctccagcat tagaattccg tgattctcg	1860
actctgagcc tggattaaat ctggccagc tgagtggaaa cttaaatgtac tagctggtt	1920
cctttagtga tcttccactt tatggctgtc tccgcctaa aagttcatca tcgtgactta	1980
ctttcttgg ggcaaaatcg tgactaactt tctttgggc aaagttggaa agcagaggc	2040
aaagtcaatc agaaatggga caaactcaat tcctactgcc tggtaaggg gccatTTCA	2100
gtagccccctt ttcaagatta gtttcatca agatttgata agctgtttt actttactat	2160
agatcttatt atccatgtca gttaaatgtt tgcttccactt aaatctatct gaattcaaaa	2220
ggtaaaaaagc taatgttcag tcttatcaga ttatcttat ttatataatgt aatgtggatt	2280
tttttaagca tataacaata atagtaatgt taggaccata aatgtggatg gctcttaca	2340
agtcaactaac attacataaa tccctcaaca acacactctg aggccataac aaacttttag	2400
aaataacaca attggctacg gaactccagc catctagctt catggctcc cacttaatt	2460
tcaaaaacaac agaactgtgc acattcattt acatgattag ggcagagctt aactgtatct	2520
catgttagcac ctacatcattt cttcagacaa acttattgcc ttttacagac aagaaaactg	2580
gggctcaaaa aaggacttgc ttataactgg ctaataaaga ggaactctgg gttcaaagtgc	2640
agtccaaatc ttcttccac ccacagcttc tgctaaatgtt attacagaaa tgcatagac	2700
agttcttcca cgttattgtc taggtttcta aagagcgtg acctaataaca acatgtctta	2760

taattttata ctgatttaac tatttcacta aggattcaact ttaactttt aacttgtaaa	2820
tatgtctaat aaacaccact gaaatagcaa cctcttctt catggccttg tggtgtaaa	2880
gcaagctagt aatatatgtc tgtggatttg tgctaataaa gttctataca cctcattaat	2940
tccacaatac ctactggta tttcttatct gccagatcct acgctaggta ctggatacac	3000
agtactgaac aaaatggta caaatgagcc tcacagagct tgttcattt gaaaggcagag	3060
agatacacac taatcaacaa attaatagta acacactacg atgtgtttt aaggaaaatt	3120
agagcatcaa agagacggtg ttagcaggtg gaggggagct cttttagatg gagaatgaga	3180
atgcctccct aaagacatgg gaataaattt agatcacaaa aatgagaaaa tagccagcct	3240
tgagaagagc agaaggaaga acattcaaag gaaaagaaag tgcatactgg aaagcctgaa	3300
cactagagtt tggtgtatgt aaggagctga gcaatggta cttgtgtat aagatgtgtg	3360
gatgtgggt gggggcagg ggtgagtccc acgcagctct taagtgtgtc ctcagactcc	3420
tgtggttcc atcagccaca acctgaataa ctgtgtggta atccaaaaat gattacagat	3480
taaacatata aaaatatcat tacaccata gtacctaagc caaggacaca gtattctatc	3540
tttcaatga agatctgcat gaagtaaaaat tattatataat aattttagt attgatata	3600
atacatcagt ggatagatag atagatgtgt ctctggata gaaaaaaagtt taaaaggat	3660
attaaaagtt cttatcttgc aggggttgaag attgtggcaa ctttcatttc ttttaattt	3720
taagaaaaaa gtggatttat gggggattag catgtttgtg ggtatataat tattttat	3780
taaaaaataa acaacaaaat gaaaacgttt ttcttctatg aaagcctaataa aagaagaaaat	3840
ttcagctgtt ttaactttagg gagctaaaaa catcaatcc aagaatgttc tctggaaactg	3900
agctcaatac atttttatgg gagtaagaat tggatacatt tccatccc tggggctcca	3960
gtctgtcaat attttacttt tcagcgataa aaagacacat gtagataatc acagtgcac	4020
cagtaacttt ctttcttta tttaagttt ttttatttctt atcgttagttt tccctgttaa	4080
agattttttc tttttgttta catatataat ttttagagaat aacaatgcac acacaaaaaa	4140
ttccctttgt tctgctagac ctggactttt tctctaataat atatctccat tttttgttt	4200
ttttcagacg tattttggaa gcaaaggaga gaattgctat atagctgact tcctttctc	4260
atcaacagt ttttaacagt ttttaagcaa aagtcagctt tgtttatcta agatttttt	4320
tgctggcatt taacctaccc ctgcctcccc tttcccaagt ccacttcagc caacctctca	4380
ttcgacaggt accaccctct aacataactg aaataatgtc taccattact ggatcttgct	4440
agcaagaat ctcaaattttt cccacttggt tgtaaattat ttgtatct ctatgtttt	4500

aggcgcgtt	gtcctatcta	atccccccc	tggcaggaca	ccttacagaa	cctaccctt	4560
acactagtca	ttaagcacca	tcagggacgg	atggctgtgt	cactggctg	tttggtattc	4620
cctactgatc	ctaccatgtg	tgatgttatct	atgacttccc	taatccctgg	ctgccttagc	4680
tgggactggc	tgacatgctt	ctcaggttgc	cgctggctt	acagtcctt	actgccatg	4740
ccactttgga	gataggcagg	gctagtaactt	ttctatataa	gcccccaaac	ttgactttgt	4800
gtttcacagt	aggtgaaaaa	gttgggtctc	tttctttta	ctttcttgc	cacaagatga	4860
taaagctagg	ggaagcctgt	ggacatggtt	tatttctgca	actgcaatga	ttgattggtg	4920
cttcctgctg	cttacttcct	aaactttgtg	ctcagtgta	gatcccttagc	agttctatc	4980
ccctgctctg	ctaaaaaaga	atggatgttgc	actctcaggc	cctagttctt	tttaattaaa	5040
ttgtatTTT	gttatcatta	ttattattat	tatTTTgaga	tgggtctta	ctctgtcgcc	5100
caggctgaag	tgcagtggtg	caatcacagc	tcactgtttt	agcctcctga	gtagctggaa	5160
ctacaagcgt	catgccacca	tgcttctttt	taatTTTta	aatggtttt	ctgccttcaa	5220
ttctaagcac	ttctcaattt	taaccaagag	ataatactttt	ttatgaattt	ttaaagttt	5280
caacagatac	tcaaagtTTT	agcaaagtct	aaatgatatt	aagcttgtcc	ttattgccc	5340
atgtacttca	atgacttattt	gttaatttgca	accaagggtc	atTTTTiaaa	tgaatatata	5400
tttattattat	atataataata	ttaaggctt	caaataccta	aaagtttagc	aaaatctaaa	5460
taatattgt	catattcttt	tattactgta	tttagtccgtt	ttcatgttgc	tgataaagac	5520
atacccaaga	ctggcaattt	tacaaaagaa	agaggttcac	tggactcaca	gttccacgt	5580
gctggggagg	cctcacaatc	acggcagctt	acgggattgt	tgagaaatga	cacttctcaa	5640
gctggggcta	aactatctt	gtggtagttt	ttctgattca	agtattgaat	tggTTTTTT	5700
tgtttttttt	gagatggagt	ttcgttcttgc	ttgcccaggc	tggagtgc	ttggcacgatc	5760
ttagctcacc	gcaacctctg	cctccgggt	tcaagtgatt	ctcctgcttgc	agcctcccaa	5820
gttagctggga	ctacaggcat	gagccaccac	acccagctaa	ttttgtatTTT	tttagtagaga	5880
catggTTTCT	ccatgttgg	caggctggtc	tcaaactccc	aacctcagg	gatccacctg	5940
ccttggcctc	ctaaagtgt	gggattacag	gtataagcca	ccgtgcccgg	ctgggacatt	6000
ggtatataaa	agctgcctag	gtaaactctaa	ccttggccc	catacatctg	aaggataacct	6060
acaatgcacc	tgaaaaatgc	aactgaaaca	gttagttccct	gggaccacac	actcagaaag	6120
gggggtgtatc	aggagatcta	gggaccagga	gggtggaaga	cctaaggcag	cactacagat	6180
gatggagaaaa	aacccactgg	ggaggggcga	tcctaaccctt	gagaatca	gagatcatgc	6240



gaataaccca tcctaagaaa tacatttctt attatatttc ttatcttata tttcttagga	8160
caatgagcta ttcaaagggt gatgataacc agcaccatca gtcagcatta tctaagaata	8220
agaatctgtg tttctacata cagacctctt aaaaaggaac ctacacttaa caggattccc	8280
caggcaattt ggatgcacat taaagctiga gcaacactgc attagaaagt tagtttcca	8340
tcacaaaaac agtaacaaaa ggaatataaa gtaagttact ttaataaatat aagaagaggg	8400
gcaggccggg cgcatggct cacgcctgta atcccagcac tttggaggc tgaggcgggt	8460
ggatcacctg aggtc	8475
<210> 4	
<211> 8401	
<212> DNA	
<213> Homo sapiens	
<400> 4	
tttcatctaa gactacattt ctattgttt atataatcg cccccctaag atcaacatgt	60
ccacatttt tggcaaagac aaagcctact gatttcagga tcattattt ccttttcaa	120
aagcacaaac ccaaactgag aaataaatca agagaaatc tccttttgc tatgtaatt	180
tagaagtaga gtcttattt ctttcaaacc ccaaagagaa tcagacatac aatatgaatt	240
tatctacttt cgcttgctca gactgagagg aaagattaat atttcaggc tgtagtcaa	300
aactgttcat tcaaataattt ttaataaaaa tccaagaacc agctaaaaag tcgcttaagc	360
taagaaacct tcaccagcct catggaaat tgtgtacgt tttccactag aatagcctat	420
aatgcttac taaaatgtc taagttcata tcttggttaac taacattttta attcaatctg	480
cagaataata tatgcttctt tagtgctaag atatgaatat tagaggcatt ctttcttaaa	540
atttcttattt agttataactt tcacaaataa ctatataata ttaaaattct gcatgtggca	600
taaaaacatat ttatggag aaggtaatgt gtagggagtt tatttcgtt tgctattaga	660
acttgttattt attcttggtt aaaaaaactg cagattacaa catagaaaaa aacaaaagta	720
tgttgatata ctcttacagt agaagataaa gagtagttct aaattttagaa aggaaaaata	780
aatatacaca gtgaaaatat gtgtcagtga gatgttaatc aaagatcaac tattgctgag	840
accagcaata ttaaatccct gcacaattac tcatattata atgagaattt taaaaagaaaa	900
atatgaacac ataacataat gaaggcagaa gtcactctca tccttcatct ttgtattccc	960
aattcaggaa gctggtagat tatcttcatt ataattacta ttcaacaaac atttgtaaaa	1020
tgaatgaata aggaatgaat gatgagaaaa atgataaaca tctccctctg tctcctggga	1080
gttaactgca ctactttctt ttaatttaa ttaatcctca atgccttgt aaaatagcca	1140

aaggaaaaat gtatttacat tactctaaat attgatgcaa tctacaaaaa gtgttaaca	1200
acttcctcaa agtaaataaa atgttcacaa tccagctagg ataaaaggat ttaaatcatt	1260
tccttaggtag agggcttca attagagccc ctgctgcatt aaccatggga actcatctca	1320
ctctctcat gaiggagccc tgagtgtgc tgctaatctg tactctacca ttctaatgct	1380
ttaaggttc ctttcagcc cttcctcctc gtaatccaca aatactgaga ccaaggcatt	1440
ttttgggtca gtcctaattt caagcattct atcctgcctt cccaaatga actcacactt	1500
attagaccat atgttcctat attagttcag gaagggggaa aaaatgttaa tcacacttgt	1560
atataaagaga tcatagaaaa acagttact aacctgtgaa aataccattc attctctgtt	1620
tacctctggt ccacagctaa gcaatcagca ggatataat gtaccctatg ttcactatcc	1680
agtattcata agtatactac ttatgaattt gaaatctgac acaacattt catgacctaa	1740
ttttgaaaat taaaatagt gtaaggcccc taggettaat ttacagggg aaagattaaa	1800
gggacacaag caaacatata ttctctctt gtgctgtggg acactggtaa tttttgact	1860
taaaatattt gatacttaaa atgccaaact tctacatttc tgcaatcaca aggcaattat	1920
catattgaat accatttctt tctctccagt aagttaggtt aatatttagca catgaactga	1980
aaatattaag tgattataaa aacgtccaaa taaattcatt aaaatttagc ttggcaaaat	2040
gttagttca tggcttggtt agaagtcctt ttatatttat attcaaataatgaaatgaaat	2100
ttacaagcaa aggaaatggc atcaaataattt tcacaccctg cctcccaagg tgtattgatt	2160
catgctttt gctcagatct aggtttctcc actcaggaaa agaggagaat gtaccatac	2220
ttggaaaaac aagttccga tggcacagct ttgtcaaaac agcaaaattt tatccatcta	2280
tgtattgcca tctgacagta tgacaaatgg tcccatgtgc gatattcaca ctgcattgca	2340
gtcaaacctg taagtcaaag gatatgaaat aatgttaact atacattaag cacagaagaa	2400
aatgaaacaa acaaaaaggat tttaaaccaa ccaaaaatat gtcttatttt ggttctca	2460
tatgttctta cattctctca ggtttttgtt gtcattatgaa acacaattctt aacaagcttgc	2520
attattttat ttccattcac atattacagg caacaagctg aaaaagtaga acgggtgtta	2580
gagagacagg acaaagtaca gattaggct tgaagtgcc ctgaccagtc gacagcaacc	2640
acatggaata atgactcatg tgcattaaatg atcacactaa atgatatttgc ttttttacc	2700
tagtcctca actgacagct taaagaactt caggttgttc tgatttttgc gcctcctcta	2760
cagcttcaga gaggacttcc attttattttt ggatcaaattt ctccacaact agttgaaact	2820
ggaattttat ttatatgaa gttccttagat gatTTAAAGC TGTAGAAGA AGAATAATGAA	2880
atcataagaa aacttgctgc tacagatatac aaaaaggaaat gttaccatcc ctcatgctaa	2940
tcctttcat tttaaataaa caggatctaa aaaaataat gctggaaatg cctaaaccaca	3000

tcaagaatgc cttagatcg tgacccaggg aacttccag aatggatgaa atagaccaa	3060
agctgaattc acctaatttt agggccaaaa acccaaaaa caaaacaaga ccaaaaaat	3120
cttcagatac tggagaaca aatctcaatt gctcaattgt atcttatgaa aacaatttt	3180
caaaataaaa caagagatat ttaagattca ttaagtttt gcatttcaa atttaagaa	3240
aatattttc taatggaatt acatataattt atatgattct tctagtata tccatggtaa	3300
taaatactct ttcatgttg aaataaaacc catttgtct atattattag ggaaaatatc	3360
tacataaatt agttttaat ttaactaaag tctatcttt gaattcataa gcataaaatt	3420
ttaaccactt gcaaaattta taacacactt aaggtagtca gatgccttgt caagtagttt	3480
aacaaaagt atittcacct gtttgttta ataacagtgc atcgatttt tgaaaatcag	3540
gcatgccctc gggcctaac aaagtatacg aagctgaatg gatctatgcc aaatatgcc	3600
gattttactt tctgagtctg atttatact tctgtctt ttcttaccac atggcttcca	3660
gtatcactta cagactaacc cttcaaaagg agaaggctaa gttactaaca tttggaaggc	3720
ttatgaaagt gaagcatagt tatgagccag caatgtttt atttagggaa tgtgtgcaaa	3780
ccatacactt aagcaagctc tggggaatga gagttgggg gaatcaactc ttttattgc	3840
taattggat ttcccttaaa agatagagtt cttccagatt ttaactgtgt taatagttac	3900
tctagaaaaa ttggagattt gtgtgcataat ttttatgtt gtaaacagac acataccag	3960
agacactgag agagacagac agacagtaaa cagaggagca ctaaccacaa acggtttaca	4020
aatgacctct gtgctcatc acctgtctgt tccccacctt gcctttata gcaactatag	4080
caacagccat gagagtcat gtggaaagaa ataaaataaa attaaaaat cctggaagct	4140
tgtaaagaat gtgagcaaag gggaggaagt tgtgaaaaaa atgaataaag ggcaccgatc	4200
cagagtattt aagaaggcag agtggagagc ctagtaatga gtatctggta ccccgatc	4260
ctctcccaca gaatctgtac agctctccgt ttatgacagt ttaaacttaa tttaaattat	4320
caaacagaca ctccctcaa acatataaaat gatgaggcag ttcatcagg ctgtatgtat	4380
aaagttgttc cagccacctt ttctaatgg cttctctata tctttacat ggagacaatg	4440
agagatttgc ttaggacaat ttgactgtaa tttagaagta gggaaatggaa agtatttgc	4500
tcttcttgc ctaactcaca ttagttactc aagtaagcat ttcttcgtt attgcatttt	4560
cctgattaca agttttatgt ttctctaaa acacatataca aaagaaatgt cctaagcact	4620
atgcaggggg aagccatgac atttatccac cactgtcagc aaaaacatga acttagccct	4680
caacagaata ttctacttca ttcttagtgc acctctgcgt cacctgcact ggagtccca	4740

cttgccgtt gggtaagacc agatgcacc gctgaaataa aaagggtca gacaatacaa	4800
gaaaagccag tagaaattgc caaatgtatc agaatacaca caggcttct aaggatatgg	4860
ccaaagagga aggctctaga gcccaccccg aaacaggatt ttgacttca cagataaatt	4920
atttatccc caataacaca attcaattaa agaaaggaa atacaaggct aaacaataaa	4980
gaaatgaaga caaaaaccca accttcaaa tctaaagaaa ataatctttt ttaaagacac	5040
agatgaagat caggaaccca aaacagaaga aaggaaaggc aattaacgct ggcacatgtat	5100
aacaacgaaa agtatggagt ctggagaatc gctagactct aaaaattata aaggtttaga	5160
cttggacttt gtacactgaa gaaaagaaaa ctgcacatgc ttatactgac caatgtacac	5220
tattgctgct tttaacttt tgtgtatag taggtagat tttttttaa gtgaaagcaa	5280
gcttattaag aaagtaaaag aataaaaagg tggcttctcc ataggcagaa aactagcgtat	5340
gttttttat tagaaattgt tattcaataa tagtacatgt tacaataaa taccatTTA	5400
aactgaaaaa atttagact tcaaatcag ttaggttgtt caccctaaaa aaggcattt	5460
tttcccccta gtctccttgc tcattgttgc cacaacaaga aatggctaa tgctatgaat	5520
aataataaca aacactgcct tctgtcaggc cctgtgtga ataccgtctg catatgtata	5580
ggaaagggtt aactcagcag gtctgtttt cccagactct gtacatttc aagaaaggc	5640
tgccttagg actggcctt ggccagctcc tggagaatga gctctcagct tttagaaaaat	5700
tctatctgct aagaatagtt ttgcattgtct caggcttgg gccacaaaaat atcagttaa	5760
tcagatggtt tatgttaaca agttagtattt atggcaaca tagatctcta atctccattt	5820
ctctctcata tatctatattt tatctatcca tatatatgtt cctatatata tcaaataatga	5880
agatatgttt atagcaattt catataataa gagagatgtt atgttagttagg aagagagaca	5940
tagatattttt ttttcattttt agaatgtttt ctggatgtt ttaaaaggaa aaacttaaga	6000
tgtgttgc aa ttgcattgtt agtttcaggat atgtacatgtt tatgtgtgtg tgtgagagac	6060
acacacaaac acattcaaa catgtttat gttttagctc aatattcaaa cacagaaata	6120
taacatctat tcttaatatg ttttatgtaa gtacagcagc agcatttata aataactgtat	6180
ttctatggtg attggaaattt agtaggcaga gaattttgtt aatggttctt aataattttt	6240
gtaatagtaa atgattactt ttgttttagt atgttttattt aatctatataa tgaataaagt	6300
ggatattctt attcatatag aaatgtgatt tactctcatg tacttatcta catgtaaaaa	6360
ccataagtttcaattttag ttctgtgccaa aggcactttt actgaataaa aataatcagc	6420
taatTTTATA tttcctgtat tcaaattttt atgcccgtgtt aatgttccgg ggtttttttt	6480

ttaattct gtaaatcaga atattcagat gttgaaaaag tcttcgcctt cagattnaa	6540
agataccctt gaaatgtac atatccaaa atgcaaccca gaggctggca atgtcaacat	6600
ttttctgttt taaaaaacct cttatgaaaa ctattgccat actaaattt ttacttgctg	6660
atgacitaca gcigggaaagg attctgtaca tataagacat caaatattga ggatactgga	6720
acttttaat taatggcaa gaaagtcaac aaaggaagtt catatgaaat caaactagta	6780
atatgattac aaaaaaaaaa gtttaaaatt tttcttgcc ccagtcttat catttctgag	6840
ccaaatacaa ttctatcgaa atcacctgaa actgaaatca ccattctagg ctggtttcc	6900
cataaagatg gactgctcca aaaagaggaa tcaagaaaga atttggctca cagtgaatta	6960
ttcacttgt cttagttaag taaaaataaa atctgactgt taactacaga aatcattca	7020
aattctgtgg tgataataaa gtaatgacca ctttcagct ggagggacta acttctttt	7080
ttttttgt gcataatatag ctgtggtaca ttttaatgtg aaatgtatgac tgcatcagct	7140
tatatccatg gagcagattt tagcatttag cttgggtctc ccagtcaata tctacgagtc	7200
tcttcttaag gagatcgatg acacagatac atacagacta acaaatgtga taccaataat	7260
caagaattca ctcagttaa atttgccccca ctgatttcca cacaagaaac ctagaatttta	7320
ctagattctt gtgcctgtga ggctccactc atttccctga atcacaaaag ctacagagta	7380
tttagataga aatataccta ctcttaacat gaaccatttt aaatatatgt attactgtgt	7440
ccacaggagt acactttaaa gcagggactt cactttcaa tctctccaaat cacgtgttac	7500
ctaaagtggc atgtggttcc ctaaagctta ataactgaca ttgccttaaa aaaggggttt	7560
gcttccgac taatgtggaa aaagtctgaa aaatgatttt aaatcttca ctaaatttct	7620
catttggtca cgtggaggaa aatgatttca ccaaataatgactt actctcattttaat	7680
gtaatttatac aaagaatga aatattttaga taaattccag atttccccca ccatgagctt	7740
ctccgaaagt atactccatc acagactgct cactaagaag ctctactgca gtcaaagtga	7800
ccgaatttaa ggggacataa tgactacttc tgctacacag aaacatttac catctctaa	7860
acttccctat gagatggaa acggacttct aatcaggtac cagagggc tctgccaact	7920
tcagggcttt gatgaataag aatgggttag agcgctcatc ataaatgaat tcagtataac	7980
tgagtgagaa agtgagagaa ccagagaaat aaatcctcat gtagaaaaatt taggggtatg	8040
aaatgccaaa tgccagttaa ccaaagctt ctttgcata aagcaacttc tataaaaatt	8100
gctgaaaata aattcttcat ggctcaatgt gaatcagtaa ttccatctc tattacactg	8160
ttgtttaccc aaaaactatt ttaatgact aagactcaga gtttgcaga gtgtttcca	8220
caaacaact gtttgagat actccagatc tgtaatcaag taagtctgaa aaacccaaa	8280
tacctcactc acctcttgaa tatgcataaa gcacactaat atataacgtt ctaaaaagcc	8340

aatcattaaa accgttttat attgttaag catttcctag acatatttgg ctacaaatct	8400
a	8401
<210> 5	
<211> 8427	
<212> DNA	
<213> Homo sapiens	
<400> 5	
gcacactgcca ccacgcccag ctaatttct atttcagta gagatgaggt tttgccatgt	60
tggccaggct ggtctcgAAC tcttgaccTC aggtgatCCA cccgcCTCAG cctaccaAAAG	120
agctgggatt acaggcgtGA gCcaccgcGC ctggccatat taacaAAATT taaatcacAA	180
ctatgtgggg ggggaggcta gtattattAC agcagatTTGG tttgctataAT aaacaAGTAC	240
tttaaaaaAT atttcttggg ccaggcgtgg tggctacgc ctgtaatCCC agcactttGG	300
gaggccgagg tggcagatC acttgaggCC aagagttaAG agaccagCCT ggccaacatG	360
gtgaaACCCC atctctactA aaaatataAC aattagccAG gcatggagGT gcatgcctGT	420
aattccagCT gtcgagagg ctgaggcAtG agaactgctG gatcctggA ggcagaggtT	480
gcagttagct gatattgcgc cactgcactC catatccAGC ctggcaacaA tggcaagact	540
ccgtctcaaA taaataaATA aataaataAA taaaactAAAG ggcagagtTT tcttaataAA	600
acatggtagC cctcagcaAC aatattgtAA gaactcctCG caagagAAAA agctggAAATA	660
agatactggC taagcaAGTA agaaaggcAC tgccctgCTT ctgcataCAT tcaaactAAAG	720
acatatacat TgcagcttAC acttacattT tccaaatATCC ccaggcatCC ctTccCTTC	780
tcaaacAGCC AAAAGGAACC agccatgCAA ataaaaataAC aagttaAGA gcctAAAAGA	840
agtcaGTGTC ctAAAAGAGA AAATTATGT aaagaattAA gatTTTGAA aactacACTT	900
tcttctggg gcgtttact ggcctccaAT acatcaatCC tgaacactG tgaactacAG	960
tgatagattG gtacatgCTT ctaaacacaA cagaattttt ccaaggTTAC atacactGTA	1020
acaaaaggGGG catTTGcAG catTTTATTt tccttaatCA actagTTGG atattctaAC	1080
agtGcaaACA ttgtAAACAA taaatttCA ttacTTTG aacttCTGA agtcaacCAA	1140
aggcttGTTG tatggatGCA atgagtACTA gacaggcAGA gctgaatACT agtcaaAAATA	1200
ttcagttact ggtgtgatAG tcctttGGG ggcataCATC acttagggAG aaactgagGT	1260
gcaaggacAT ttacacaca GcaaaaACAT tctcagGAAT ttgtcacatC attaccataA	1320
gccccAAATC tcaaggTCTT agaacAGCCT gagttCTGA tcaaattATA ttgtAAAAG	1380
agaggaaaaAA AATGtGAAGC gtGCTATTT taaaataAC agtaactACT actactGCTG	1440

ctgctgctaa ttctaaacgt ttactgagcc cttattatgt gccaaggcacc gtgcttagta	1500
cggcataga ttttacaat taatccctgt aacaaccctc tgatattagt taataaaatt	1560
aaagtagaat cctcaccaaa aaaattaaa ctttccaaat aaaaatataa ataaattatt	1620
aaagacattt caccttttc tctgcctcg actacattt caagtattaa atttacacta	1680
aaaccacatt tatttcagg aattccagtt aaagcgtaca gatattcaag atgttgacaa	1740
tttattacaga agaatcacag aactctgaaa ttaaatactg gcacagaaaa cttccatcc	1800
aaccttacgg aacaactatc cccatttaa aaaaaaagga acagcatata tatcaggctt	1860
gataataaga ggcttctcat gcccacacta gcaatgaatg atgccataat tataaagaga	1920
cctgtatcgc cacatgcata aaaataattt acatctgcta agtcaagttt tcaatatattt	1980
attttgtgtg taaaccttat agtagctgat aaaaataca ataaactaat ctaaggtaaa	2040
ctaaaacact aggttggttc tgaagactca cttagaatt tgagcagcat aataatcata	2100
atatttagtaa tcaaactact tagcagaaag ttcttagagg gctggaaagc tgtgtataat	2160
aaaatggagc agacaagaag gaagggttt ccgtactgtt taaatcaact acaggtccca	2220
gcatgcagt ctctaatctg aagttaaagca aaaactgcaa tgcatactgg gacttgttagt	2280
aagtaaacca cgttatcaca gcaagttca agaaagtctg aactatctag cacaatttga	2340
ctatatctta ttatcagagt ctaatcaaattttaatcaaa tttgtatgtt ctctgatgtg	2400
gcacacagtt tctctagcac ataccggaaa aagtatcaat atttagacca acatttcac	2460
attagaaaaa tcttacgttag gagaagcaca gaaaaaaaaatg ctgaaaaagc aaaaaaactt	2520
gatgaataaa aaatataattttttaatag ttttttaaag tttgaatgga tccatttcaa	2580
cattctctaa tcctccccca caaaaagttt aattgttttgcggggcgcg gtggctcacg	2640
cctgtaatcc caacacttta ggaggctgag gcgggtgaat tacgagatca agagatcgag	2700
accatcctgg ccaacatggt gaaaccatct ctactaaaaa tacaaaaattt agttggcggt	2760
ggtgccgcac gcctgttagtc ccagctactc aggaggctga gacaggagaa ttgttigaac	2820
ctggaggtg gaggctgcag ttagctataa tcgcaccact gcactccagc ctggtgacag	2880
tgtgagattt attctcaaaa aaaaaaaaaaaa aaaaagtttta attgttttaa cagtttgctt	2940
tttaacaattt attcaagatg tattttataa ataatttttc ttgaagaaaa ttctcagaag	3000
caaacattcc ccatattcta atattgccca ccagggaaata attttttag taatacgac	3060
acacccatc acaaaaacaa acaaaaaaca ctgaagttct gcttttgtca agtcttact	3120
caatatttat gccctccatt ctcacccctt aattccctac acacacacac acacacgcac	3180
acatccccac acacacacgc ttctacaaag aacacttaga aaaacagttt tccaaactaca	3240
agcccacttc tctcatccac tgacctttc tgaaaacaca aaagatttt taagctatca	3300

gtaacacgtc caaacacaag ctgataagtt tgagctagaa ttacatata tacagttgt acacaccctc ctatccctg caagtctgtg gaaggaggct gggaaagaac taagtcaat	3360 3420
ctgcatcagg aggctaaca caggtggtgg gtttttca ggcaacagca cttcacaaa catgtttgg aatatagtc aagaattcc taacaaggaa agataagctg gcacacaat ttaacgcaat ccagctaaaa atcatctgc acacatgtca ctacattca ccataaaagt gacggctac tataaaggat ttgaagctt gtcaatacaa catactgtcc ataaggccag agatagcgt tgccatggtt actataccca ctttatcag gaaattactg tcattaccc aaagtttgg gtacttattt aaaattaaa aaaaacacac acaatttagg gttcgactg ttaatigagt gaaataatca actactgtt gattgttaag tatgtcgctt tggagatgca	3480 3540 3600 3660 3720 3780 3840
catggtaac aatacttgga tctgcagcag aaaaaaaatc aattccccc tgctgctc tctcctcaag tactgacagt ttgtattctc aatgcagcca aaacaataaa acaaaaccca tcttttggc ttctgtgtt aagttatcc tcccccttaggc ccacaacag agtcaaaata aagcctagat catcaacctg ttaggcctca tccccctcc atccccctca tactggc ctttcttgac tacttagaaa aggcagaaaa cattctgtc actgattcca aagtatagaa aagaatagtt gcctcaact gagatattt caccaaaatc ttttttattt actttttt taaggcaggg agaggggaga gacttgagg gtactgaaag ggagaagtgg aggagtattc	3900 3960 4020 4080 4140 4200 4260
aaatgccac acaagtctag tgtaagaaag ttgcctttaga agagtccaa ggtggc acctcacata taattctaa aagcttgga agagttcacc ataattttaa gactgaattt agggacaagt aatagaaaag ttattctaa agtctacttc aacattttt caaaagataa ctattcaaaa attaacaca catataagaa ttatacgaaa gcctacaaaa tagtatggcc acatatacac acaaacatac aaagtagaaa acataagctt tttaaagaaat aattatctac aataaattca atgcaatgtt aacatattat ctcttttta aaaaatcgca aagcagcaaa aacatacacc tgagaaaatt aatgtgatca aaacgttaaa gaattcttag gcctataaaa	4320 4380 4440 4500 4560 4620 4680
aaagcccatg tacaaaagct cctgagaagt caacataaat cattaatatt tccagcaca aaataatatg aaaattcaaa catgtttcaa gaaatcgtt ctagatatag atataaaaga attccattaa agtcagaga cctaaaactt taattccctc cttctctgt ttgaatagta attaaataca aaagcctca gcaataaaat actaaggata caaaattaa aagcacatca atataagctt aacttcgttca tgtcttcaca gaaagcttta ctattcactg tctgttaggat gaaaaagtttta ataacaccct gagaggtttcc attttatct aaacgttaa gtgtttct	4740 4800 4860 4920 4980 5040

caccgttcac agaagcaagt ttctatattt actttctaaa gggggcaatt tcaaaagaat	5100
agtcaacttct aaaatthaag atactatacc ttttgatagg ctcataaaca cagggttcct	5160
aattatctat atttacttt aaaatgttc tattccaaat ttgtgagcag agtttataag	5220
aaagctgaaa ctcaaggcct taaactttg ggttatTTT acacaaaaat attcagtgc	5280
actcctctag atttgagtag tcatttcctt gtgcattcctt ctaaaataga aaaacaaaaa	5340
tgatatacc atatatacct aatactaaca catacagata tacatcttt tcactgtcaa	5400
acaagcttga aagctttagg cagtaagaat tttcagaaa gttagcagag tcagtcaaaa	5460
cattcaaaac ttgaaccatg acatctgtta ctctgtcaat aagagtctat agaagaatca	5520
gggaacttac atactcacta aaatcaacta ctatcacatc acatcaatgg agaaatgaag	5580
aaaaactgta atagggaca tacaattcac aggatctca aaaggaaaaa tgatctttt	5640
ttttttta aattatgaga aactgactag gcagcatttt ttcaaaagca gcttcaaaac	5700
tataacaag acatTTTgg taaccacagc agtatttaaa aaacaaaaat ttaggccgg	5760
cgtggcgtt cacgcctata atcccagcac ttgggaggc caaggcaggt ggatcacctg	5820
agttaggat tcaagaccag cctgaccaac atggtgatac cccgtctcta ctcaaaatac	5880
aaaacttagc cggcgtagt ggccggacacc tctataatca cagctactca ggaggctgag	5940
aggcaggaga atcgcttcaa cctgggaggc agaggttgca gtgagccgag atcacccgt	6000
tgcactccag ccggaaac agagcgagac tccgtctcaa aaaataaaaa aataaaaaaa	6060
ctatagtgtc cagggtgac tttaaatgtt ttacttctc aactgatatg gaaaaagtta	6120
gcatttaag acagaagctt ctgtccatgt attaattagt tacctatctc aacaacttaa	6180
tatctgcattt ctttcttacc atttatgaag aactttata tgtattatct cattggct	6240
tactgagaaa acagtatttt gcctacaaaa tagacaaaaat tcaaagcaga tttatcaaac	6300
tttcttagcat ccccaattt ttAAAacttc gacacaaaaac ttacaagca accacagtgg	6360
catgatattt tcagtgtataa tcaattcacc taacactaac agagttcaa aggaccatgt	6420
gtctataatg ctatgaaact gttaaagtag ctatattcat cttatgcag ttactgttac	6480
atcaacaatg acctaccact gatacaactt gacttacagt tcaagaatct cagtcttgc	6540
aggctaactt aagtacatca accatatgtt ttataaaagc cgagtgccta aaaattgatc	6600
tatattagaa tcatagtctg taaatccgag gggaaaaac tacaagaatg ctAAAatTTT	6660
ttcaacacac tataccccctt tccaaaatct caactactt atatcctatt tgttataata	6720
ttatagggat gataacaagg cttaaagccc taaatcatac caactactt tgTTTataac	6780

aattacaat aatttttaa aatacatgtc caacatccc ctcataaca caagactaat	6840
tcccttcca aataaaataa ttctaaacag tgctctgtac caagggccag aatcctata	6900
ctatccgaa tcgcacatct actttgtaca gtcaaagact tcactttcaa gtagcaaaca	6960
ttatTTatga atgaaatTT taaatggact tactcaaaat ctttctggaa cttaaggtg	7020
ttaatcctgt tgcttagctg aagctaagca gagctgtaat aagtagcaag accctaaaa	7080
ttcaaaaatt tccttatct tgctgttagca ctcctgctg gatagcattt agagatctc	7140
atgtaagcag aagaagagta ttccagaggc agctccttcc agaagactga ataggaaaaa	7200
ggatggaccc ttcaagcta aaagaaatag gccccatcca tcacttatac cttctaaaaa	7260
tacaatttag cccaggtagg tgtcttttc atctattact actccagttc cacaaagact	7320
tgcctcagtc caaaatacaa catgctaaa taaagcctgc aaaattgtct aaaaactaag	7380
ttaaaaagca ttcaatagca cccaaagcaaa acactttatt atggcagcc aagcaatgtc	7440
agtcaaactg taaatactat tatgttacca aaagcaaaag tctgatgtt aaaaaaaaaa	7500
aaaaaaaaagcc cctggaatat tcgtaacatg tttagccagat gtttgttt tgagaacttt	7560
gtgcactatt actatgctct tcacttaagg atagttgtac atctacaac gtttaagta	7620
cagaaatttt ttataaaca ttagcataac tgtacacaaa attcctctt tgccatgaaa	7680
agataggc tgggatttga aatgttattt ttccagacatt ttaatgacc ccctaaaata	7740
aactagttt aagcccacaa caccgattcc ataaacaagt aaagacagaa gaagagaata	7800
agaaggaact taccaaaatt aaaatgaata atagtatttc cagtaaaaat gttagaacag	7860
tttccaacaa tgctgttaac caaataaatt gtgaaactta aaaaaggaag gagggggcca	7920
gtcttcaag accaaagca aagctgacct atttatttct attgctttaga gtgaacacca	7980
gatgtaaaca aatatcataa acactgaaaa gtacgcttac atggtttagc ctcaattca	8040
gtacccttac caggccctca ataaagctac agatgttgg gagaactcgc tcaaaaagga	8100
gataattcca gccctcgcc ttaagaatc cctatcaagt gaacctgtga aaagacttcc	8160
ttccctcagat gcacaactgc ttaaaaaaaaaaaacttccatcataatctga	8220
ttctaatatt caactatcca ttatTTatata ataaatgttc ttccctctt aactttccca	8280
gctcgagcat ctacattcct gacaccgact attagcaaaa atgcacaact cttccccag	8340
ctatggggca aatcttggaa atctgaaaca cagccacaaa gttcactgtc aaggccaggt	8400
gatgaggccc acacatgccc ggacctt	8427

&lt;210&gt; 6

&lt;211&gt; 1704

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Fragment A

&lt;400&gt; 6

ggtaccaccc aagtggcta ggtaagctg ctaggccac caiggtgcg cagacccagg	60
tgttcatctc cctgctgctg tggatctcg ggcataatgg cgatatcgat atgattaaac	120
gtacggtgtc cgccccctcc gtgtcatct tccccccctc cgacgagcag ctgaagtccg	180
gcaccgcctc cgtgggtgc ctgctgaata acttctaccc cagagaggcc aaggtgcagt	240
ggaagggtgga caacgcctg cagtcggga actcccagga gagcgtgacc gagcaggaca	300
gcaaggacag cacctacagc ctgagcagca ccctgaccct gagcaaagcc gactacgaga	360
agcacaagggt gtacgcctgc gaggtgaccc accagggcct gagtcffff gtcaccaaga	420
gcttcaacag gggggagtgt tagggggcccg tttaaacggg tggcatccct gtgaccctc	480
cccagtgcct ctccctggccc tggaagttgc cactccatgt cccaccagcc ttgtcctaatt	540
aaaattaagt tgcatttgcattt tgtctgacta ggtgccttc tataatatta tgggtggag	600
gggggtggta tggagcaagg ggcaagttgg gaagacaacc tggtaggcct gcggggctta	660
ttgggaacca agtggagtg cagtggcaca atctggctc actgcaatct ccgcctctg	720
ggttcaagcg attctcctgc ctcagcctcc cgagttgtg ggattccagg catgcattac	780
caggctcacc taattttgt tttttggta gagacggggt ttcaccatat tggccaggct	840
ggcttccaac tcctaatctc aggtgatcta cccaccttg cctcccaaatt tgctggatt	900
acaggcgtga accactgctc cacgcgcct gtacggcgc attaagcgcg gcgggtgtgg	960
tggtaacgcg cagcgtgacc gctacacttg ccagcgcct agcgcgcct ccttcgcct	1020
tcttccttc ctttcgcgc acgttcgcgc gctttcccg tcaagctcta aatcgggggc	1080
tccctttagg gtccgattt agtgcttac ggcacctcga cccaaaaaaa cttgattagg	1140
gtgatggttc acgttgtgg ccatgcgcct gatagacggt tttcgcctt ttgacgttgg	1200
agtccacgtt cttaatagt ggactttgt tccaaactgg aacaacactc aaccctatct	1260
cggcttatttc tttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg	1320
agctgattta aaaaaattt aacgcgaatt aattctgtgg aatgtgtgtc agttagggtg	1380
tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcattc tcaatttagtc	1440
agcaaccagg tggaaagt cccaggctc cccagcaggc agaagtatgc aaagcatgca	1500
tctcaattttt tcagcaacca tagtcccgc cctaactccg cccatcccgc ccctaactcc	1560
gcccaatttcc gccattctc cgcccatgg ctgactaatt ttttttattt atgcagggc	1620
cgaggccccc tctgcctctg agctattcca gaagtagtga ggaggcttt ttggaggcct	1680

aggctttgc	aaaaagctcc	cggg	1704			
<210>	7					
<211>	1120					
<212>	DNA					
<213>	Artificial Sequence					
<220><223>	human IgG1 signal + human IgG1 constant					
<400>	7					
tgcttaggcc	accatgaaac	acctgtggtt	cttcctcctg	ctggtggcag	ctcccagatg	60
ggtgctgagc	caggtgaat	tgtgcaggcg	gttagctcag	cctccaccaa	gggccaagc	120
gtcttcccc	tggcacccctc	ctccaagagc	acctctggcg	gcacagccgc	cctgggtgc	180
ctggtaagg	actacttccc	cgaaccgtg	accgtgagct	ggaactcagg	cgccctgacc	240
agcggcgtgc	acaccttccc	cgctgtcctg	cagtctcag	gactctactc	cctcagcagc	300
gtggtgaccg	tgcctccag	cagttggc	acccagacct	acatctgcaa	cgtgaatcac	360
aagcccagca	acaccaaggt	ggacaagaga	gttgagcca	aatcttgta	caaaaactcac	420
acatgcccac	cctgcccagc	acctgaactc	ctggggggac	cctcagttt	ccttttcccc	480
ccaaaaccca	aggacacccct	catgatctcc	cggaccctg	aggtcacatg	cgtggtggtg	540
gacgtgagcc	acgaagaccc	tgaggtcaag	ttcaacttgt	acgtggacgg	cgtggaggtg	600
cataatgcc	agacaaagcc	ccgggaggag	cagtacaaca	gcacgtaccg	ggtggtcagc	660
gtcctcaccg	tcctgaccca	ggactggcig	aatggcaagg	agtacaagtg	caaggcttcc	720
aacaaagccc	tcccagcccc	catcgagaaa	accatctca	aagccaaagg	ccagccccgg	780
gaaccacagg	tgtacaccct	gccccatcc	cgggaggaga	tgaccaagaa	ccaggtcagc	840
ctgacacctcc	tggtaaaagg	ttcttatccc	agcgacatcg	ccgtggagtg	ggagagcaat	900
ggccagcccg	agaacaacta	caagaccacc	cctccgtgc	tggactccga	cggctccctc	960
ttcctctaca	gcaagctcac	cgtggacaag	agcagggtgc	agcaggccaa	cgtttctca	1020
tgctccgtga	tgcatgaggc	tctgcacaac	cactacaccc	agaagagcct	ctccctgtct	1080
cccgccaaat	gagatatcg	gcccgtttaa	acgggtggca			1120