



등록특허 10-2135745



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년07월21일
(11) 등록번호 10-2135745
(24) 등록일자 2020년07월14일

(51) 국제특허분류(Int. Cl.)
C12N 15/85 (2006.01) *C07K 16/00* (2006.01)
(21) 출원번호 10-2014-7018672
(22) 출원일자(국제) 2012년12월05일
 심사청구일자 2017년12월04일
(85) 번역문제출일자 2014년07월04일
(65) 공개번호 10-2014-0113666
(43) 공개일자 2014년09월24일
(86) 국제출원번호 PCT/IB2012/056977
(87) 국제공개번호 WO 2013/084157
 국제공개일자 2013년06월13일
(30) 우선권주장
 61/567,675 2011년12월07일 미국(US)
(56) 선행기술조사문현
 EP1308510 A

(73) 특허권자
 아이크노스 사이언스 에스. 아.
 스위스, 2300 라 쇼드퐁, 슈명 드 라 콤베타 5
(72) 발명자
 로쉐, 테니얼
 스위스, 2300 라 쇼드퐁, 슈명 드 라 콤베타 5
 에비셔-거미, 크리스텔
 스위스, 2300 라 쇼드퐁, 슈명 드 라 콤베타 5
 (뒷면에 계속)
(74) 대리인
 이원희

전체 청구항 수 : 총 30 항

심사관 : 이효진

(54) 발명의 명칭 발현 카세트

(57) 요 약

본 발명은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열의 발현에 유용한 발현 카세트에 관한 것이다.

(72) 발명자

모레띠, 뼈에르

스위스, 에프-2300 라 쇼드퐁, 슈명 드 라 꼼베타

5

버츠친거, 마틴

스위스, 2300 라 쇼드퐁, 슈명 드 라 꼼베타 5

명세서

청구범위

청구항 1

프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 포유동물(eukaryotic) 글리세르알데하이드 3-포스페이트 디하이드로제네즈(Glyceraldehyde 3-phosphate dehydrogenase(GAPDH)) 프로모터의 비번역된 게놈 DNA 서열 다운스트림을 포함하는 발현 증강 요소를 포함하는 발현 카세트(expression cassette)로서,

여기서, 상기 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호 8인 발현 카세트.

청구항 2

제 1항에 있어서, 상기 발현 카세트는 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림(non-translated genomic DNA sequence upstream)을 추가적으로 포함하고,

여기서, 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호 7인 발현 카세트.

청구항 3

프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열 및 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림을 포함하는 발현 카세트로서,

여기서, 상기 발현 카세트는 포유동물 GAPDH 프로모터 또는 이의 단편을 포함하지 않는다는 것에 기반하여, 상기 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호 7인 발현 카세트.

청구항 4

제 3항에 있어서, 상기 발현 카세트는 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림을 추가적으로 포함하고,

여기서, 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호 8인 발현 카세트.

청구항 5

제 1항 내지 제 4항 중 어느 한 항에 있어서, 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열에 실시가능하게(operably) 연결되지 않는 발현 카세트.

청구항 6

제 1항 내지 제 4항 중 어느 한 항에 있어서, 상기 발현 카세트는 폴리아데닐레이션 자리(polyadenylation site)를 포함하는 발현 카세트.

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

제 1항 내지 제 4항 중 어느 한 항에 있어서, 상기 프로모터 및 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열은 작동가능하게 연결된(operatively linked) 발현 카세트.

청구항 22

제 1항 또는 제 4항에 있어서, 상기 포유동물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 동일한 방향으로 지향되는(orientated) 발현 카세트.

청구항 23

제 1항 또는 제 4항에 있어서, 상기 포유동물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 반대 방향으로 지향되는 발현 카세트.

청구항 24

제 2항 또는 제 3항에 있어서, 상기 포유동물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 동일한 방향으로 지향되는 발현 카세트.

청구항 25

제 2항 또는 제 3항에 있어서, 상기 포유동물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 반대 방향으로 지향되는 발현 카세트.

청구항 26

제 1항 내지 제 4항 중 어느 한 항에 있어서, 상기 프로모터는 SV40 프로모터, MPSV 프로모터, 마우스 CMV, 인간 tk, 인간 CMV, 래트 CMV, 인간 EF1알파, 차이니즈 햄스터 EF1알파, 인간 GAPDH, MYC, HYK 및 CX 프로모터를 포함하는 하이브리드(hybrid) 프로모터로 구성된 군으로부터 선택되는 발현 카세트.

청구항 27

제 1항 내지 제 4항 중 어느 한 항에 있어서, 상기 폴리펩티드는 항체, 항체 단편(antibody fragments) 또는 항체 유도체(antibody derivates)로 구성된 군으로부터 선택되는 발현 카세트.

청구항 28

제 6항에 있어서, 상기 폴리아데닐레이션 자리는 BGH 폴리(A) 및 SV40 폴리(A)로 구성된 군으로부터 선택되는 발현 카세트.

청구항 29

제 1항 내지 제 4항 중 어느 한 항에 있어서, 추가 프로모터(additional promoter), 인핸서(enhancer), 전사조절 요소(transcriptional control elements), 및 선별가능한 마커(selectable marker)로 구성된 군으로부터 선택되는 유전적 요소(genetic element)를 추가적으로 포함하는 발현 카세트.

청구항 30

제 29항에 있어서, 상기 유전적 요소는 선별가능한 마커를 암호화하는 폴리뉴클레오티드 서열에 포함된 CpG 자리의 함량(content)이 45 또는 그 미만인 선별감능한 마커인 발현 카세트.

청구항 31

제 1항 내지 제 4항 중 어느 한 항의 발현 카세트를 포함하는 발현 벡터.

청구항 32

하기의 순서를 포함하는, 발현 벡터로서:

- a) 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림
 - b) 프로모터
 - c) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열
 - d) 폴리아데닐레이션 자리
 - e) 인핸서
 - f) 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림, 또는
- a) 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림
 - b) 인핸서
 - c) 프로모터
 - d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열
 - e) 폴리아데닐레이션 자리
 - f) 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림, 또는
- a) 인핸서
 - b) 포유동물 GAPDH의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림
 - c) 프로모터
 - d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열
 - e) 폴리아데닐레이션 자리
 - f) 포유동물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림,

만약 a) 또는 b)가 포유동물 GAPDH의 비번역된 게놈 DNA 서열 업스트림이면 f)는 포유동물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림이고 만약 a) 또는 b)가 포유동물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림이면 f)는 포유동물 GAPDH의 비번역된 게놈 DNA 서열 업스트림인 조건으로, 상기 인핸서의 포함은 선택적이고, 상기 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호 8이고 상기 포유동물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호 7인 발현 벡터.

청구항 33

제 31항에 있어서, 상기 발현 벡터는 추가 프로모터, 인핸서, 전사 조절 요소, 복제의 기원(origin of replication) 및 선별가능한 마커로 구성된 군으로부터 선택되는 유전적 요소를 추가적으로 포함하는 발현 벡터.

청구항 34

제 31항에 있어서, 상기 발현 벡터는 복제의 기원 및 선별가능한 마커를 암호화하는 발현 벡터의 폴리뉴클레오티드 서열에 포함된 CpG 자리의 함량이 200 또는 미만인 복제의 기원 및 선별가능한 마커를 추가적으로 포함하는 발현 벡터.

청구항 35

삭제

청구항 36

제 31항의 발현 벡터를 포함하는 숙주 세포.

청구항 37

삭제

청구항 38

제 1항 내지 제 4항 중 어느 한 항의 발현 카세트 또는 제 31항 내지 제 34항 중 어느 한 항의 발현 벡터로 숙주 세포를 형질전환 및 폴리펩티드를 회수하는 것을 포함하는, 폴리펩티드 발현을 위한 시험관 내(*in vitro*) 방법.

청구항 39

제 38항에 있어서, 상기 발현 카세트는 안정적으로 형질전환되는 방법.

청구항 40

제 38항에 있어서, 상기 발현 카세트는 일시적으로 형질전환되는 방법.

청구항 41

제 1항 내지 제 4항 중 어느 한 항의 발현 카세트를 포함하는 포유동물 숙주 세포(mammalian host cell)로부터 이종폴리펩티드(heterologous polypeptide)의 발현을 위한 조성물.

청구항 42

유전자 치료를 위한 제 31항의 발현 벡터.

청구항 43

- 1) 제 31항의 발현 벡터를 가진 숙주세포를 형질전환 시키는 단계; 및
- 2) 폴리펩티드를 회수하는 단계; 를 포함하는 폴리펩티드의 발현을 위한 시험관 내 방법(*in vitro* method).

청구항 44

제 43항의 방법에 있어서, 상기 발현 벡터는 안정적으로 형질전환되는 방법.

청구항 45

제 43항의 방법에 있어서, 상기 발현 벡터는 일시적으로 형질전환되는 방법.

청구항 46

제 31항의 발현 벡터를 포함하는 포유동물 숙주 세포(mammalian host cell)로부터 이종폴리펩티드(heterologous polypeptide)의 발현을 위한 조성물.

발명의 설명

기술 분야

[0001] 본 발명은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열의 발현에 유용한 발현 카세트(expression cassette)에 관한 것이다. 또한, 본 발명은 숙주 세포로부터 폴리펩티드의 생성을 위한 발현 카세트 및 발현 카세트의 용도를 포함하는 벡터 및 숙주 세포에 관한 것이다.

[0002]

배경 기술

[0003] 재조합 폴리펩티드의 생성을 위한 발현 시스템은 당업계에서 잘 알려져 있고, 예를 들면, Marino MH (1989) Biopharm, 2: 18-33; Goeddel DV *et al.*, (1990) Methods Enzymol 185: 3-7; Wurm F & Bernard A (1999) Curr Opin Biotechnol 10: 156-159에 의해 설명되어 있다. 일반적으로, 약학적 적용에서 용도를 위한 폴리펩티드는 바람직하게는 CHO 세포, NSO 세포, SP2/0 세포, COS 세포, HEK 세포, BHK 세포 또는 등과 같은 포유동물 세포에서 생성된다. 상기 목적을 위해 사용된 발현 벡터의 필수적 요소는 복제의 원핵생물 기원(origin) 및 원핵생물 선별가능한 마커(selectable marker), 선택적으로 진핵생물 선별가능한 마커, 및 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 선택적으로 폴리아데닐레이션 신호를 포함하는 전사 종결자를 포함하는 각각의 관심의 구조적 유전자의 발현을 위한 하나 또는 그 이상의 발현 카세트를 포함하는 원핵생물 플라스미드 증식 유닛(propagation unit), 예를 들면, *E.coli*로부터 선택된다. 포유동물 세포에서 일시적 발현을 위하여, SV40 Ori 또는 OriP와 같은 복제의 포유동물 기원이 포함될 수 있다. 프로모터로서 구조적(constitutive) 또는 유발하는(inducible) 프로모터가 선택될 수 있다. 최적화된 전사를 위해서, 코작(Kozak) 서열이 5' 비번역 영역에 포함될 수 있다. mRNA 진행, 특히 mRNA 스플라이싱(splicing) 및 전사 종료(transcription termination)를 위해서, 폴리아데닐레이션 신호뿐만 아니라 구조적 유전자(엑손/인트론 조합)의 조합에 따른 mRNA 스플라이싱 신호가 포함될 수 있다. 유전자의 발현은 일시적 또는 안정 세포주를 사용하여 수행될 수 있다. 생성 세포주에서 안정 수준 및 폴리펩티드의 고발현은 재조합 폴리펩티드의 생성의 전체 진행에 중대하다. 단백질과 같은 생물 분자 및 특정 항체 또는 항체단편을 위한 요구는 지난 수년에 걸쳐 현저히 증가하고 있다. 고비용과 낮은 수율은 생물 분자의 활용성에 있어서 제한요인이 되어왔고 산업적 규모에서 원하는 생물 분자의 수율을 증가시키는 강건한 진행을 발달시키는데 주요 도전이 되어왔다. 따라서, 재조합 폴리펩티드 생성에서 고발현을 얻기 위하여 발현 벡터의 효율을 개선시키기 위한 필요성이 아직 있다.

[0004]

발명의 내용

해결하려는 과제

과제의 해결 수단

발명의 효과

도면의 간단한 설명

[0005] 도 1은 마우스 거대세포바이러스(cytomegalovirus) 프로모터(mCMV), 첫번째 인트론을 포함하는 Ig 기증 수용체 단편(IgDA), IgG1 항체 경쇄(IgG1 LC), 뇌심근염바이러스(Encephalomyocarditis)로부터 유래된 내부 리보좀 입구부위(Internal Ribosomal Entry Sites)(IRES), IgG1 항체 중쇄(IgG1 HC), 녹색 형광 단백질(GFP) 및 시미안

바이러스 40 폴리아데닐레이션 신호(poly A)를 구성하는 리포터 발현 구조체(REP)를 나타낸다.

도 2는 형질전환 후 5일째에 CHO-S 세포에서 IgG1 항체의 일시적 발현을 나타낸다(IgG 적정량(titers)의 평균이 두 개 별개의 형질전환에 대해 플롯됨). 세포를 GAPDH_A 및 GAPDH_B 벡터(GAPDH_A 및 GAPDH_B), GAPDH 업스트림 및 다운스트림 요소 (A 및 B)가 없는 동일한 벡터 및 대조군(pGLEX41)으로서 pGLEX41 벡터를 사용하여 형질전환하였다. 상층액에 축적된 IgG1 항체의 농도를 옥테트(Octet) 기구(Fortebio, Menlo, CA, USA)를 사용하여 결정하였다.

도 3은 HEK293 EBNA 세포에서 IgG1 항체의 발현을 나타낸다. 세포를 GAPDH_A 및 GAPDH_B 벡터(GAPDH_A 및 GAPDH_B) 및 대조군(pGLEX41)으로서 pGLEX41 벡터를 사용하여 형질전환하였다. 상층액을 옥테트(Octet) 기구를 사용하여 형질전환 후 10일째에 수확하고 분석하였다. 데이터는 벡터 당 튜브스핀에서 N = 3 독립 형질전환을 나타낸다.

도 4는 세포 풀(pools)을 사용하여 일괄(batch) 생성에 대한 발현 수준 연구를 나타낸다. 세포를 형질전환하고 안정 세포의 풀을 GAPDH_A 및 GAPDH_B 벡터(GAPDH_A(1), GAPDH_A(2), GAPDH_B(1) 및 GAPDH_B(2)), GAPDH 업스트림 및 다운스트림 요소(A(1) 및 A(2))가 없는 동일한 벡터 및 대조군(pGLEX41)으로서 pGLEX41 벡터를 사용하여 제조하였다. 배양 7일 후에 상층액을 상층액에 축적된 항체에 대해 옥테트 기구를 사용하여 분석하였다. IgG 적정량의 평균이 각각 풀에 대해 ($\mu\text{g}/\text{ml}$)주어졌다. 데이터는 풀 당 N = 2 일괄(batches)을 나타낸다.

도 5는 안정 형질전환 및 제한 희석에 의해 생성된 집단(populations) 상에서 발현 수준 연구를 나타낸 도이다. 세포를 GAPDH_A 및 GAPDH_B 벡터(GAPDH_A 및 GAPDH_B), GAPDH 업스트림 및 다운스트림 요소가 없는 동일한 벡터(A 및 B) 및 대조군(pGLEX41)으로서 pGLEX41 벡터를 사용하여 형질전환하였다. 안정 형질전환으로부터 클론 및 미니풀에 의해 발현된 GFP 형광의 평균 값을 형질전환 후에 14일째에 측정하였다. 세포를 96-웰 플레이트에서 선별 압력(selection pressure)하에서 배양하였다. 데이터는 벡터당 N = 48 클론 또는 미니풀을 나타낸다.

도 6은 상층액에서 IgG1 항체의 발현 상에 대한 배지 첨가제인 인슐린 및 PMA (포르볼 12-미리스테이트 13-아세테이트, 포르볼 에스테르)(phorbol 12-myristate 13-acetate, a phorbol ester)의 효과를 나타낸다. GAPDH_A 벡터(GAPDH_A)와 대조군(pGLEX41)로서 pGLEX41 벡터로 형질전환 후에 세포를 PowerCHO2 배지, 4mM Gln, +/- 인슐린 및 PowerCHO2, 4mM Gln, PMA +/- 인슐린에서 희석하였다. pGLEX41(채워진 바(filled bars)) 또는 GAPDH_A(공개 바(open bars))에 대하여 표준 배지와 비교하여 발현에서 차이가 없음을 관찰하였다.

도 7은 인간 GAPDH 유전자 자리(locus)의 개요를 나타낸다. GAPDH 유전자는 유전자 NCAPD2 및 IFF01 옆에 있다(flanked).

도 8은 인간 GAPDH 유전자, GAPDH 업- 및 다운스트림 요소 및 GAPDH 업스트림 단편화 연구 분석을 위해 제조된 단편의 세부를 나타낸다. NruI 제한 자리는 클로닝 단계를 용이하게 하기 위하여 도입되었고 계놈 5' GAPDH 업스트림 서열(이는 별표(asterisk)를 사용하여 강조됨)의 일부가 아니다. 단편의 크기는 하기와 같다: 단편 1(서열번호: 9): 511 bps, 단편 2 (서열번호: 10): 2653 bps, 단편 3 (서열번호: 11): 1966 bps, 단편 4 (서열번호: 12): 1198 bps, 단편 8 (서열번호: 13): 259 bps, 단편 9 (서열번호: 14): 1947 bps, 단편 11 (서열번호: 15): 1436 bps, 및 단편 17 (서열번호: 16): 1177 bps.

도 9는 GAPDH 업스트림 및 다운스트림 요소의 단편화의 발현 결과를 나타낸다. 발현 결과를 형질전환 후에 10일째에 CHO 세포에서 일시적 형질전환으로 얻었다. 정량화를 옥테트 기구를 사용하여 수행하였다. 벡터 pGLEX41는 음성 대조군으로서 이용되었다. 또한, pGLEX41-ampiA는 GAPDH 인접 요소가 없는 벡터의 기저 발현을 나타내는 음성 대조군이다. pGLEX41-업/다운은 전장 인접(업스트림 및 다운스트림) 영역을 포함하고 양성 대조군으로서 이용되었다. pGLEX41-업은 단지 업스트림 인접 영역을 포함하고 pGLEX41-다운은 단지 다운스트림 인접 영역을 포함한다. 모든 다른 구조체는 도 8에 설명된 단편을 포함한다. 단편 2 및 3은 IgG1 LC 및 IgG1 HC로서 동일한 방향으로 클론되거나 또는 IgG1 LC 및 IgG1 HC(AS)와 관련하여 반대 방향으로 클론되었다.

도 10은 형질전환 후 8일째에 CHO-S 세포에서 IgG1 항체의 일시적 발현을 나타낸다(IgG의 적정량의 평균은 세 개의 독립 형질전환에 대하여 플롯됨; 에러바: SD +/-). 세포를 마우스 CMV(A_GAPDH_UP) 또는 차이니즈 햄스터 GAPDH 프로모터 (A_GAPDH_UP_PR)와 조합된 차이니즈 햄스터 GAPDH 업스트림 요소를 갖는 벡터를 사용하여 형질전환하였다. 단지 마우스 CMV (A) 또는 차이니즈 햄스터 GAPDH 프로모터 (A_PR)를 갖는 플라스미드는 대조군으로서 형질전환되었다. 상층액에 축적된 IgG1 항체의 농도는 옥테트 QK 기구를 사용하여 결정되었다(Fortebio, Menlo, CA, USA).

발명을 실시하기 위한 구체적인 내용

[0006]

발명요약

[0007]

본 발명은 일반적으로 재조합 폴리펩티드 생성에서 증가된 발현을 얻기 위하여 사용될 수 있는 발현 카세트 및 발현 벡터와 같은 발현 시스템에 관한 것이다. 하나의 측면에서, 본 개시는 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 진핵생물 글리세르알데하이드 3-포스페이트 디하드로제네이즈(Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)) 프로모터의 비번역된 게놈 DNA 서열 다운스트림을 포함하는 발현 카세트를 제공하고, 여기서 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치로부터 +7000 주위의 뉴클레오티드 위치까지 영역거리(region spanning)내에서 시작하고, 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 100 주위부터 15000 주위의 뉴클레오티드까지이다.

[0008]

추가적 측면에서, 본 개시는 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열 및 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림을 제공하고, 여기서 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터 5' 말단 주위로부터 -3500 주위의 뉴클레오티드 위치까지 영역 거리내에서 시작하고, 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 발현 카세트는 진핵생물 GAPDH 프로모터 또는 이의 단편을 포함하지 않는다는 조건으로 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 100부터 15000 주위의 뉴클레오티드까지이다.

[0009]

추가적 측면에서, 본 개시는 발현 카세트를 포함하는 발현 벡터 및 발현 카세트 또는 발현 카세트를 포함하는 발현 벡터를 포함하는 숙주세포를 제공한다

[0010]

더욱 추가적 측면에서, 본 개시는 발현 카세트 또는 발현 벡터를 갖는 숙주세포 형질전환 및 폴리펩티드 회수를 포함하는, 폴리펩티드 발현을 위한 시험관 내 방법 및 포유동물 숙주 세포로부터 이종 폴리펩티드의 발현을 위한 발현 카세트 또는 발현 벡터의 용도를 제공한다.

[0011]

본 발명의 상세한 설명

[0012]

본 개시는 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열 및 진핵생물 글리세르알데하이드 3-포스페이트 디하드로제네이즈(Glyceraldehyde 3-phosphate dehydrogenase(GAPDH)) 프로모터의 비번역된 게놈 DNA 서열 다운스트림을 포함하는 발현 카세트 및 발현 벡터에 관한 것이고, 여기서 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치로부터 +7000 주위의 뉴클레오티드 위치까지 영역거리(region spanning)내에서 시작하고, 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 100 주위부터 15000 주위의 뉴클레오티드까지이다.

[0013]

본 개시는 추가적으로 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열 및 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림을 포함하는 발현 카세트에 관한 것이고, 여기서 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고,

[0014]

진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터의 5' 말단 주위로부터 -3500 주위의 뉴클레오티드 위치까지 영역 거리내에서 시작하고, 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 발현 카세트는 진핵생물 GAPDH 프로모터 또는 이의 단편을 포함하지 않는다는 조건으로 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 100부터 15000 주위 뉴클레오티드까지이다.

[0015]

본 명세서에서 사용된 용어 "발현 카세트"는 시스-액팅 전사 조절 요소(cis-acting transcriptional control

elements)의 임의 조합을 포함하는 발현되록 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열 및 프로모터 및 선택적으로 인핸서(enhancer) 서열과 같은 자체의 발현을 조절하는 서열을 포함한다. 유전자 발현을 조절하는 서열은, 예를 들면, 전사 생성물의 자체의 전사 및 번역은 조절유닛(regulatory unit)으로서 일반적으로 언급된다. 조절유닛의 대부분은 유전자의 코딩 서열의 업스트림에 위치해 있고, 실시가능하도록 거기에 연결되어 있다. 발현카세트는 폴리아데닐레이션 자리를 포함하는 다운스트림 3' 비번역 영역을 포함할 수 있다. 본 발명의 조절 유닛은 전사유닛과 같은 발현되는 유전자에 실시가능하도록 연결되거나 또는 예를 들면, 이종 유전자의 5'-비번역된 영역과 같은 간접 DNA에 의해 거기로부터 분리된다. 바람직하게는 발현 카세트는 벡터로 발현 카세트의 삽입 및/또는 벡터로부터 자체의 절제를 가능하도록 하기 위하여 하나 또는 그 이상의 적절한 제한효소에 의해 옆에 있다. 따라서, 본 발명에 따른 발현 카세트는 발현 벡터, 특히, 포유동물 발현 벡터의 구조체에 사용될 수 있다. 본 발명의 발현 카세트는 하나 또는 그 이상, 예를 들면, 진핵생물 GAPDH 프로모터 또는 이의 단편의 2개, 3개 또는 그 이상의 비번역된 게놈 DNA 서열 다운스트림, 및/또는 하나 또는 그 이상, 예를 들면, 진핵생물 GAPDH 프로모터 또는 이의 단편의 2개, 3개 또는 그 이상의 비번역된 게놈 DNA 서열 업스트림을 포함할 수 있다. 만약 본 발명의 발현 카세트가 진핵생물 GAPDH 프로모터 또는 이의 단편의 하나 이상의 DNA 서열 다운스트림 및/또는 업스트림을 포함한다면, 상기 DNA 서열은 직접적으로 연결되는, 즉, 예를 들면, 5' 및 3' 말단에 부착되고 서열 또는 이의 단편의 일련의 클로닝을 수월하게 하는 제한효소 자리를 포함하는 링커서열을 포함할 수 있다. 진핵생물 GAPDH 프로모터 또는 이의 단편은 DNA 서열 다운스트림 및/또는 업스트림은 직접적으로 연결되지 않을 수 있다, 즉, 간접 DNA 서열로 클론될 수 있다.

[0016] 본 명세서에서 사용된 용어 "폴리펩티드를 암호화하는 폴리뉴클레오티드 서열"은 유전자, 바람직하게는 폴리펩티드를 발현하는 이종 유전자를 위한 DNA 코딩을 포함한다.

[0017] 용어 "이종 코딩 서열", "이종 유전자 서열", "이종 유전자", "재조합 유전자", 또는 "유전자"는 교체사용될 수 있다. 상기 용어는 재조합, 특히 숙주세포, 바람직하게는 포유동물 세포에서 발현되고 수화되어야만 하는 재조합 이종 단백질 생성물에 대해 코드하는 DNA 서열을 의미한다. 상기 유전자의 생성물은 폴리펩티드일 수 있다. 상기 이종 유전자 서열은 숙주 세포에 자연적으로 존재하지 않고 동일한 또는 상이한 종으로부터 유래되고 유전적으로 변형될 수 있다.

[0018] 상기 용어 "단백질" 및 "폴리펩티드"는 붙어있는 잔기의 알파-아미노와 카르복시기 사이의 펩티드 결합에 의해 다른 것과 연결된 일련의 아미노산 잔기를 포함하도록 교체될 수 있다.

[0019] 본 명세서에서 사용된 상기 용어 "비번역 게놈 DNA 서열"은 유기체의 유전자 정보를 구성하는 DNA를 포함한다. 대부분의 모든 유기체의 게놈은 DNA이고, 단지 RNA 게놈을 갖는 일부 바이러스는 예외이다. 대부분의 유기체에서 게놈 DNA 문자는 염색체라고 불리는 DNA-단백질 복합체로 구성체로 조직화되어 있다. 크기, 염색체의 수, 게놈 DNA의 성질은 다른 유기체 사이에서 다양하다. 바이러스 DNA 게놈은 단일- 또는 이중-가닥, 직선 또는 원형일 수 있다. 박테리아는 단일, 원형 염색체를 가진다. 진핵생물에서, 대부분의 게놈 DNA는 다른 크기의 다수의 직선 염색체로서 핵(핵의 DNA) 내에 위치해 있다. 진핵생물 세포는 미토콘드리아 및, 식물 및 저등 진핵생물, 클로로플라스트(chlorop마지막s)에서 추가적으로 게놈 DNA를 포함한다. 상기 DNA는 원형분자이고 상기 세포기관 내에 다수의 복사본으로 존재한다. 비번역된 게놈 DNA 서열은 일반적으로 프로모터에 실시가능하도록 연결되어 있지 않으므로 번역되지 않는다. 상기는 예를 들면, 발현되지 않는 단백질을 암호화하는 유전자, 따라서, 번역되지 않는 유전자를 포함할 수 있다.

[0020] 본 명세서에서 사용된 상기 용어 "진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림"은 진핵생물 GAPDH 프로모터의 비번역된 진핵생물 게놈 DNA 3'에 해당한다. 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치, 바람직하게는 +1 뉴클레오티드 위치에서 시작하고, 여기서, 뉴클레오티드 위치는 GAPDH mRNA의 전사시작에 비례, 즉, GAPDH에 대한 진핵생물 유전자 코딩의 전사 시작의 기원(origin)에 비례한다. 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 항상 진핵생물 GAPDH 프로모터로서 동일한 기원이고, 예를 들면, 만약 GAPDH 프로모터가 인간 기원이면, 인간 GAPDH 프로모터

의 비번역된 게놈 DNA 서열 다운스트림은 더욱이 인간 기원이고 자연적으로 발생하는 인간 GAPDH 프로모터의 인간 게놈 DNA 서열 다운스트림에 해당함으로써, 항상 동일한 기원이다.

- [0021] 본 명세서에서 사용된 용어 "진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터의 비번역된 진핵생물 게놈 DNA 5'에 해당한다. 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 일반적으로 진핵생물 GAPDH 프로모터의 5' 주위의 뉴클레오티드 위치, 바람직하게는 진핵생물 GAPDH 프로모터 5' 말단 후 즉시의 뉴클레오티드 위치에서 시작한다. 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터, 예를 들면, 만약 GAPDH 프로모터가 인간 기원이면, 인간 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 더욱이 인간 기원이고 자연적으로 발생하는 인간 GAPDH 프로모터의 인간 게놈 DNA 서열 업스트림에 해당함으로써, 항상 동일한 기원이다.
- [0022] 본 명세서에서 나타낸 진핵생물 GAPDH 프로모터의 위치, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 또는 업스트림 및 다른 DNA 서열은 GAPDH mRNA의 전사 시작에 비례, 예를 들면, 특정적으로 지적하지 않는 한, 진핵생물 GAPDH의 전사 시작의 기원에 비례한다.
- [0023] 상기 용어 "진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 확장한다" 또는 "진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 확장한다"는 시작으로부터 특정 유전적 요소까지 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림의 길이 확장(extension), 예를 들면, 인트론(intron)으로 확장을 정의하는데 사용된다. 상기 확장은 유전적 요소를 암호화하는 전장의 DNA 서열, 예를 들면, 인트론 또는 이의 일부를 포함한다.
- [0024] 진핵생물 GAPDH 프로모터 및 GAPDH 프로모터의 진핵생물 게놈 DNA 업스트림 and/or 다운스트림은 NCBI 공동 데이터은행(public databank)에 있는 인간, 래트 및 마우스에 대하여 발견될 수 있고(인간, 마우스, 래트 및 차이니즈 햄스터 GAPDH 유전자 가입은 각각 유전자 IDs 2597(mRNA: NM_002046.3), 14433(mRNA: NM_008084.2), 24383(mRNA: NM_017008.3) 및 100736557(mRNA: NM_001244854.2); 생물기술 정보를 위한 국가센터(NCBI): <http://www.ncbi.nlm.nih.gov/>), 인간 GAPDH 유전자에 대하여 도 7 및 8에서 예로써 나타낸다.
- [0025] 진핵생물 GAPDH 프로모터는 GAPDH mRNA의 전사 시작에 비례하여 bps -500 주위로부터 +50 주위까지 스트레치 되도록 항상 고려된다. 인간 GAPDH 프로모터는 염색체 12 상에 위치해 있다. 인간 GAPDH 프로모터는 단편화(fragmentation) 연구에 기반한 GAPDH mRNA의 전서 시작에 비례하여 bps -488 까지 +20 스트레치(stretch) 되도록 Graven *et al.* (Graven *et al.*, (1999) *Biochimica et Biophysica Acta*, 147: 203-218)에 의해 고려된다. NCBI 공동 데이터은행에 따라 인간 GAPDH 프로모터는 NCBI 공동 데이터은행에 의해 정의된 GAPDH mRNA의 전사 시작에 비례하여 bps -462부터 +46까지 스트레치한다. 특정적으로 달리 지적되지 않는 한, 본 명세서에서 언급된 인간 GAPDH 프로모터는 4071부터 서열번호: 17의 4578까지 스트레칭하는 서열에 해당하는 GAPDH mRNA의 전사 시작에 비례하여 -462부터 +46 위치까지 스트레치한다.
- [0026] 본 명세서에서 언급된 인간, 마우스 및 래트 기원의 GAPDH 유전자, IFF01 유전자 및 NCAPD2 유전자의 DNA에 대해 사용된 숫자는 NCBI 공동 데이터 은행(<http://www.ncbi.nlm.nih.gov/>)에 있는 상기 유전자에 대해 사용된 숫자에 해당한다.
- [0027] 본 명세서에서 사용된 용어 "프로모터"는 DNA에 결합하는 RNA 중합효소를 감독하거나 RNA 합성을 개시하여 전사의 개시를 매개하는 유전자의 업스트림에 일반적으로 위치한 조절 DNA 서열을 정의한다.
- [0028] 본 명세서에서 사용된 용어 "인핸서(enhancer)"는 유전자의 정체성(identity), 유전자에 관련한 서열 위치 또는 서열의 방향(orientation)에 독립적으로 유전자의 전사를 강력하게 행하는 뉴클레오티드 서열을 정의한다.

[0029]

[0030]

상기 용어 "기능적으로 연결된(functionally linked)" 및 "실시가능하게 연결된(operably linked)"는 교체적으로 사용되고 두 개 이상의 DNA 시그먼트(segments), 특히, 발현되는 유전자 서열과 자체의 발현을 조절하는 상기 서열 사이에서의 기능적 관계를 언급한다. 예를 들면, 만약 프로모터 및/또는 인핸서 서열이 적절한 숙주 세포 또는 다른 발현 시스템에서 코딩 서열의 전사를 촉진 또는 조절하면 시스-액팅(cis-acting) 전사 조절요소의 임의 조합을 포함하는, 프로모터 및/또는 인핸서 서열은 코딩 서열에 실시가능하게 연결된다. 전사된 유전자 서열에 실시가능하게 연결된 프로모터 조절 서열은 전사된 서열에 물리적으로 근접하다.

[0031]

"방향(Orientation)"은 주어진 DNA 서열에서 뉴클레오티드 순서를 언급한다. 예를 들면, 얻어진 서열로부터의 DNA에서 참조 포인트와 비교하였을 때 다른 DNA 서열과 관련하여 반대 방향으로 DNA 서열의 방향은 다른 서열이 역방향된 것과 관련하는 서열의 5'에서 3'으로의 순서인 것이다. 상기 참조 포인트는 원천 DNA에 있는 다른 특정된 DNA 서열의 전사 방향 및/또는 서열을 포함하는 반복가능한 벡터의 복제의 기원을 포함한다.

[0032]

본 명세서에서 사용된 용어 "발현 벡터"는 적절한 숙주 세포로 형질전환이 숙주 세포 내에 재조합 유전자 생성물의 고수준의 발현을 제공하는 분리되고 정제된 DNA 분자를 포함한다. 재조합 또는 유전자 생성물에 대한 DNA 서열 코딩에 추가하여, 발현 벡터는 mRNA로 DNA 코딩 서열의 효율적 전사 및 숙주 세포주의 단백질의 mRNA의 효율적 번역에 대해 요구되는 조절 DNA 서열을 포함한다.

[0033]

본 명세서에서 사용된 용어 "숙주 세포" 또는 "숙주 세포주"는 배양에서 자라게 하고 원하는 재조합 생성 단백질을 발현하는 것이 가능한 임의 세포, 특히 포유동물 세포를 포함한다.

[0034]

본 명세서에서 사용된 용어 "단편(fragment)"은 각각의 뉴클레오티드 서열, 예를 들면, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림의 일부 또는 프로모터와 같은 특정 유전적 요소를 암호화하는 뉴클레오티드 서열의 일부를 포함한다. 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림의 단편은 생물 활성을 유지할 수 있고 따라서 변경, 예를 들면, 프로모터에 실시가능하게 연결된 코딩 서열의 발현 패턴을 증가할 수 있다. 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림의 단편은 적어도 약 100부터 약 3000 bp까지, 바람직하게는 약 200부터 약 2800 bp까지, 더욱 바람직하게는 약 300부터 약 2000 bp까지 뉴클레오티드, 특히 약 500부터 약 1500까지 bp 뉴클레오티드의 범위일 수 있다. 본 발명의 발현 카세트에서 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림의 단편을 클론하기 위하여, 클로닝을 수월하게 하는 제한효소 자리를 포함하는 링커 서열은 단편의 5' 및 3' 말단에 항상 부착된다.

[0035]

본 명세서에서 사용된 용어 "뉴클레오티드 서열 정체성" 또는 "동일한 뉴클레오티드 서열"은 가능하면, 최대 퍼센트 서열 정체성(identity)를 얻기 위하여 서열 및 도입 갭(introducing gap)을 정렬시킨 후에, 예를 들면, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림의 뉴클레오티드 서열과 동일한 후보서열에서 뉴클레오티드의 퍼센트를 포함한다. 따라서, 서열의 정체성은 두개의 뉴클레오티드 서열의 뉴클레오티드 위치에 있는 유사성을 비교하기 위하여 사용되는 표준 방법에 의해 결정될 수 있다. 항상 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림에 대한 후보 서열의 뉴클레오티드 서열의 정체성은 예를 들면, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 및 100%를 포함하여, 적어도 80%, 바람직하게는 적어도 85%, 더욱 바람직하게는 적어도 90%, 및 가장 바람직하게는 적어도 95%, 특히 96%, 더욱 특히 97%, 더욱 더 특히 98%, 가장 특히 99%에 있다.

[0036]

본 명세서에서 사용된 용어 "CpG 자리"는 시토신(cytosine) 뉴클레오티드가 자체의 길이를 따라 베이스(bases)의 직선서열에서 구아닌(guanine) 뉴클레오티드 옆에서 일어나는 DNA 영역을 포함한다. "CpG"는 "-C-포스페이트-G-"에 대해 함축되고, 즉, 시토신 및 구아닌이 단지 하나의 포스페이트에 의해 분리된 것; 포스페이트는 DNA

에서 임의 두 개의 뉴클레오사이드와 함께 연결된다. "CpG" 표기법은 시토신 및 구아닌의 CG 베이스-짝짓기로부터 상기 직선 서열을 구별하도록 사용된다.

[0037] 본 명세서에서 사용된 용어 "선택적 코돈 사용(alternative codon usage)"은 CpG 서열 모티프를 피하기 위하여 동일한 아미노산을 위한 선택적 코돈 코딩의 사용을 포함한다. 상기는 새로운 CpG 자리를 유도하는 두 개의 코돈 결합을 회피하는 것뿐만 아니라 내부의 CpG 자리를 가지지 않는 바람직한 코돈 사용을 포함한다(예를 들면, 알라닌(Alanine)을 위한 GCG 코딩 및 CpG 자리를 포함하는 것은 GCT, GCC 또는 GCA에 의해 교체될 수 있다)

[0038] GAPDH mRNA의 전사시작에 비례한, 예를 들면, 진핵생물의 전사 시작의 기원에 비례하는 DNA 서열의 길이와 관련 및 뉴클레오티드 위치와 관련한 본 명세서에서 사용된 용어 "주위(around)"는 최대 ±50 % 편차 값을 포함하고, 기술된 값, 예를 들면, "3000 뉴클레오티드 주위"의 항상 최대 ±10 %는 2700 내지 3300 뉴클레오티드, 바람직하게는 2900 내지 3100 뉴클레오티드, 더욱 바람직하게는 2995 내지 3005 뉴클레오티드를 포함하고, "100 뉴클레오티드 주위"는 50 내지 150 뉴클레오티드, 바람직하게는 90 내지 110 뉴클레오티드, 더욱 바람직하게는 95 내지 105 뉴클레오티드의 값을 포함하고, "15000 뉴클레오티드 주위"는 13500 내지 16500 뉴클레오티드, 바람직하게는 14500 내지 15500 뉴클레오티드, 더욱 바람직하게는 14990 내지 15010 뉴클레오티드, 가장 바람직하게는 14995 내지 15005 뉴클레오티드 값을 포함하고, "200 뉴클레오티드 주위"는 50 to 250 뉴클레오티드s, 바람직하게는 190 to 210 뉴클레오티드s, 더욱 바람직하게는 195 to 205 뉴클레오티드 값을 포함하고, "8000 뉴클레오티드 주위"는 7200 내지 8800, 바람직하게는 7500 내지 8500 뉴클레오티드, 더욱 바람직하게는 7990 내지 8010 뉴클레오티드, 가장 바람직하게는 7995 내지 8005 뉴클레오티드의 값을 포함하고, "500 뉴클레오티드 주위"는 50 내지 550 뉴클레오티드, 바람직하게는 475 내지 525, 더욱 바람직하게는 490 내지 510, 가장 바람직하게는 495 내지 505 뉴클레오티드의 값을 포함하고, "5000 뉴클레오티드 주위"는 4500 내지 5500 뉴클레오티드, 바람직하게는 4750 내지 5250, 더욱 바람직하게는 4990 내지 5010, 가장 바람직하게는 4995 내지 5005 뉴클레오티드의 값을 포함하고, "1000 뉴클레오티드 주위"는 900 내지 1100 뉴클레오티드, 바람직하게는 950 내지 1050, 더욱 바람직하게는 990 내지 1010, 가장 바람직하게는 995 내지 1005 뉴클레오티드의 값을 포함하고, "4500 뉴클레오티드 주위"는 4050 내지 4950 뉴클레오티드, 바람직하게는 4250 내지 4750, 더욱 바람직하게는 4490 내지 4510, 가장 바람직하게는 4495 내지 4505 뉴클레오티드의 값을 포함하고, "1500 뉴클레오티드 주위"는 1350 내지 1650 뉴클레오티드, 바람직하게는 1450 내지 1550, 더욱 바람직하게는 1490 내지 1510, 가장 바람직하게는 1495 내지 1505 뉴클레오티드의 값을 포함하고, "4000 뉴클레오티드 주위"는 3600 내지 4400 뉴클레오티드, 바람직하게는 3800 내지 4200, 더욱 바람직하게는 3990 내지 4010, 더욱 바람직하게는 3995 내지 4005 뉴클레오티드의 값을 포함하고, "2000 뉴클레오티드 주위"는 1800 내지 2200 뉴클레오티드, 바람직하게는 1900 내지 2100, 더욱 바람직하게는 1990 내지 2010, 가장 바람직하게는 1995 내지 2005 뉴클레오티드의 값을 포함하고, "2700 뉴클레오티드 주위"는 2430 내지 2970 뉴클레오티드, 바람직하게는 2600 내지 2800, 더욱 바람직하게는 2690 내지 2710, 가장 바람직하게는 2695 내지 2705 뉴클레오티드의 값을 포함하고, "3300 뉴클레오티드 주위"는 2970 내지 3630 뉴클레오티드s 바람직하게는 3100 내지 3500, 더욱 바람직하게는 3290 내지 3310, 가장 바람직하게는 3295 내지 3305 뉴클레오티드의 값을 포함하고, "3200 뉴클레오티드 주위"는 2880 내지 3520 뉴클레오티드s, 바람직하게는 3000 내지 3400, 더욱 바람직하게는 3190 내지 3210, 가장 바람직하게는 3195 내지 3205 뉴클레오티드의 값을 포함하고, +7000 주위 또는 위치 +7000 주위는 위치 +6300 내지 +7700, 바람직하게는 위치 +6700 내지 +7300, 더욱 바람직하게는 위치 +6990 내지 +7010, 가장 바람직하게는 위치 +6995 내지 +7005를 포함하고, +1 주위 또는 위치 +1 주위는 위치 -10 내지 +10, 바람직하게는 위치 -5 내지 +5, 더욱 바람직하게는 위치 -1 내지 +2를 포함하고, -3500 주위 또는 위치 -3500 주위는 위치 -3150 내지 -3850, 바람직하게는 위치 -3300 내지 -3700, 더욱 바람직하게는 위치 -3490 내지 -5010, 가장 바람직하게는 위치 -3495 내지 -3505을 포함한다.

[0039] SEQ ID 번호의 서열에서 위치와 관련하여 본 명세서에서 언급되거나 사용된 인간, 마우스 및 래트 기원의 GAPDH 유전자, IFF01 유전자 및 NCAPD2 유전자의 DNA에 사용된 번호와 관련하여 본 명세서에서 사용된 용어 "주위(round)"는 최대 ± 500 bps, 바람직하게는 ± 100 bps, 더욱 바람직하게는 ± 10 bps, 가장 바람직하게는 ± 5 bps의 편차 값을 포함한다.

[0040] 일실시형태에서, 본 개시는 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 진핵생물 GAPDH 프로모터의 비변역된 계놈 DNA 서열 다운스트림을 포함하는 발현 카세트를 제공하고, 여기서, 폴리뉴클레오티드

서열에 의해 암호화되는 폴리뉴클레오티드 서열은 GAPDH가 아니고, 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치부터 +7000 주위의 뉴클레오티드 위치까지의 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 100 주위부터 15000 주위 뉴클레오티드까지이다.

[0041] 일실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 적어도 100 뉴클레오티드 주위이고 자체 최대에서 IFF01 유전자의 두 번째 인트론 또는 이의 일부까지 연장한다. 일실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 적어도 100 뉴클레오티드 주위이고 자체 최대에서 IFF01 유전자의 마지막 인트론까지 연장한다.

[0042] 인간 IFF01 유전자는 염색체 12 (NCBI 유전자 ID: 25900)의 bps 6665249 내지 6648694 주위의 인간 DNA에 위치해 있다. 일실시형태에서, 인간의 IFF01 유전자의 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 인간의 IFF01 유전자(위치 +7021)에 대한 염색체 12 코딩의 bps 6650677 주위까지 자체의 최대에서 스트레치한다. 일실시형태에서, 인간의 IFF01 유전자의 두 번째 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 인간(위치 + 13574)의 IFF01 유전자에 대한 염색체 12 코딩의 bps 6657230 주위까지 자체의 최대에서 스트레치한다. 각각 인간의 IFF01 유전자의 마지막 인트론 및 IFF01 유전자의 두 번째 마지막까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 염색체 12 (NCBI 유전자 ID: 25900)의 bps 6657230 내지 6639125를 나타내는 서열번호: 17에 포함된다. 상기 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 17에 의해 나타낸 뉴클레오티드 서열의 bps 11553 주위까지 스트레치하고 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 17에 의해 나타낸 뉴클레오티드 서열의 bps 18106 주위까지 스트레치한다.

[0043] 마우스 IFF01 유전자 (NCBI 유전자 ID: 320678)는 염색체 6의 bps 125095259 내지 125111800 주위의 마우스 DNA에 위치해 있다. 일실시형태에서, 마우스의 IFF01 유전자의 마지막 인트론까지 자체 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 마우스의 IFF01 유전자 (위치 + 6391)에 대한 염색체 6 코딩의 bps 125109211 주위의 자체 최대에서 스트레치한다. 일실시형태에서, 마우스의 IFF01 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 마우스 IFF01 유전자(위치 +12081)에 대한 염색체 6 코딩의 bps 125103521 주위까지 자체의 최대에서 한다. 마우스의 IFF01 유전자의 마지막 인트론 및 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 염색체 6(NCBI 유전자 ID: 320678)의 bps 125103521 내지 125119832를 나타내는 서열번호: 18에서 각각 포함된다. 마우스의 IFF01 유전자의 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 18에 의해 나타내는 뉴클레오티드 서열의 bps 10622 주위까지 스트레치하고 마우스의 IFF01 유전자의 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 18에 의해 나타내는 뉴클레오티드 서열의 bps 16312 주위까지 스트레치한다.

[0044] 래트 IFF01 유전자(NCBI 유전자 ID: 362437)는 염색체 4의 bps 161264966 내지 161282150 주위의 래트 DNA에 위치해 있다. 일실시형태에서, 래트의 IFF01 유전자의 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 래트의 IFF01 유전자(위치 + 5154)에 대한 염색체 4 코딩의 bps 161280937 주위까지 자체의 최대에서 스트레치한다. 일실시형태에서, 래트의 IFF01 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 래트의 IFF01 유전자(위치 +6640)에 대한 염색체 4 코딩의 bps 161279451 주위까지 자체의 최대에서 스트레치한다.

[0045] 래트의 IFF01 유전자의 마지막 인트론 및 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH

프로모터의 비번역된 게놈 DNA 서열 다운스트림은 염색체 4(NCBI 유전자 ID: 362437)의 bps 161279451 내지 161290508를 나타내는 서열번호: 19에서 각각 포함된다. IFF01 유전자의 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 19에 의해 나타내는 뉴클레오티드 서열의 bps 9572 주위까지 스트레치하고 IFF01 유전자의 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 19에 의해 나타내는 뉴클레오티드 서열의 bps 11058까지 스트레치한다.

[0046] 차이니즈 햄스터 IFF01 유전자(NCBI 유전자 ID: 100753382)는 bps 3577293 내지 3593683 주위의 차이니즈 햄스터 DNA에 위치해 있다. 일실시형태에서, 차이니즈 햄스터에서 IFF01 유전자의 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 차이니즈 햄스터 (위치 +6883)에서 IFF01 유전자에 대한 bps 3579014 코딩 주위까지 자체의 최대에서 스트레치한다. 일실시형태에서, 차이니즈 햄스터의 IFF01 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 차이니즈 햄스터의 IFF01 유전자(위치 +12930)에 대한 bps 3585061 코딩 주위에 자체의 최대에서 스트레치한다. 염색체의 위치는 염색체의 데이터뱅크에서 아직 주석이 달리지 않았고 현재 서열정보는 많은 알려지지 않은 베이스를 포함한다. 그러므로 한계의 정확한 표기법은 더 많은 정확한 서열 정보의 활용가능성으로 변할 수 있다.

[0047] 차이니즈 햄스터의 IFF01 유전자의 마지막 및 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 bps 3567932 내지 3585061를 나타내는 서열번호: 29에 각각 포함된다. IFF01 유전자의 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 29에 의해 나타낸 바와 같이 뉴클레오티드 서열의 bps 11083 주위까지 스트레치하고 IFF01 유전자의 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 29에 의해 나타낸 바와 같이 뉴클레오티드 서열의 bps 17130 주위까지 스트레치한다.

[0048] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 진핵생물 GAPDH 폴리아데닐레이션 자리에서 시작하고, 예를 들면, 진핵생물 GAPDH 폴리아데닐레이션 자리를 암호화하는 첫 번째 뉴클레오티드에서 시작한다. 바람직하게는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 진핵생물 GAPDH 폴리아데닐레이션 자리의 다운스트림에서 시작하고, 예를 들면, 진핵생물 GAPDH 폴리아데닐레이션 자리를 암호화하는 마지막 뉴클레오티드 후에 즉시 시작한다. 더욱 더 바람직하게는, 진핵생물 GAPDH 프로모터의 진핵생물 GAPDH 프로모터는 진핵생물 GAPDH 폴리아데닐레이션 자리의 다운스트림에서 시작하고 진핵생물 GAPDH 프로모터는 비번역된 게놈 DNA 서열 다운스트림의 길이는 적어도 100 뉴클레오티드 주위이고 IFF01 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장한다.

[0049] 일실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 +3881 주위의 뉴클레오티드 위치로부터 +5000 주위의 뉴클레오티드 위치까지 내에서 시작하고, 바람직하게는 +3931 주위의 뉴클레오티드 위치로부터 +5000 주위의 뉴클레오티드 위치까지 영역거리 내에서, 더욱 바람직하게는 +4070 주위의 뉴클레오티드 위치로부터 +5000 주위의 뉴클레오티드 위치까지 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례한다.

[0050] 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은, 예를 들면, 본 발명에서 사용된 진핵생물 GAPDH 폴리아데닐레이션 자리의 다운스트림은 항상 위치 +3931 주위의 뉴클레오티드 위치에서 시작하고, 바람직하게는 +4070 주위의 뉴클레오티드 위치에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례한다.

[0051] 인간 GAPDH 폴리아데닐레이션 자리의 인간 비번역된 게놈 DNA 서열 다운스트림은 뉴클레오티드 위치 +3931(서열 번호: 17에서 나타낸 바와 같이 bp 8463에 해당하는 GAPDH mRNA의 전사 시작에 비례하여) 주위에서 시작한다. 바람직하게는, 만약 GAPDH 폴리아데닐레이션 자리의 비번역된 게놈 DNA 서열 다운스트림은 인간으로부터 이면,

GAPDH 폴리아데닐레이션 자리의 비번역된 게놈 DNA 서열 다운스트림은 +3931 주위에서 시작하고(서열번호: 17에 나타낸 바와 같이 bp 8463에 해당하는 GAPDH mRNA의 전사 시작에 비례하여) 자체의 길이는 서열번호: 17에서 나타낸 바와 같이 bps 8463 주위로부터 11819 주위까지의 서열에 해당하는 3357 bps 주위이고 더욱 바람직하게는 이것은 +4070 주위에서 시작하고(서열번호: 17에서 나타낸 바와 같이 bps 8602에 해당하는 GAPDH mRNA의 전사 시작에 비례하여) 자체의 길이는 서열번호: 17에 나타낸 바와 같이 bps 8602 주위부터 11819 주위까지의 서열에 해당하는 3218 bps 주위이다.

[0052] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 8 및 21 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열을 포함한다.

[0053] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 8 및 21 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 상보적인 뉴클레오티드 서열을 포함한다.

[0054] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 8 및 21 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 적어도 80% 일치하는 뉴클레오티드 서열을 포함한다.

[0055] 일부 실시형태에서, 서열번호: 8 및 21 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열은 5개 또는 미만, 4개 또는 미만, 더욱 바람직하게는 3개 또는 미만, 가장 바람직하게는 2개 또는 미만, 특히, 1개 핵산 변형, 상기 핵산 변형은 바람직하게는 핵산 치환이다.

[0056] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 바람직하게는 200 주위부터 8000 주위까지 뉴클레오티드이고, 더욱 바람직하게는 500 주위부터 5000 주위까지 뉴클레오티드이고, 더욱 더 바람직하게는 1000 주위부터 4500 주위까지 뉴클레오티드이고, 가장 바람직하게는 1500 주위부터 4000 주위까지 뉴클레오티드이고, 특히 2000 주위부터 3500 주위까지 뉴클레오티드, 더욱 특히 2700 주위부터 3300 주위까지, 더욱 더 특히 3200 주위, 가장 특히 3218 뉴클레오티드이다. 본 명세서에서 정의된 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 비번역된 게놈 DNA 서열에 첨가된 임의 링커 서열을 포함한다.

[0057] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열로서 동일한 방향으로 지향되었다.

[0058] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열에 관하여 반대 방향으로 지향되었다.

[0059] 일부 실시형태에서, 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림을 포함하는 발현 카세트는 추가적으로 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림을 포함하고, 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터의 5' 말단 주위부터 -3500 주위의 뉴클레오티드 위치까지 영역내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 100 주위부터 15000 주위까지 뉴클레오티드이다.

[0060] 다른 실시형태에서, 발현 카세트는 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 진핵생물

GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림을 포함하고, 상기 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 상기 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터의 5' 말단 주위부터 -3500 주위의 뉴클레오티드 위치까지 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 발현 카세트가 진핵생물 GAPDH 프로모터 또는 이의 단편을 포함하지 않는 것에 기반하여 상기 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 100부터 15000 주위의 뉴클레오티드까지이다.

[0061] 일부 실시형태에서, 상기 발현 카세트는 추가적으로 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림을 포함하고, 상기 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치부터 +7000 주위의 뉴클레오티드 위치까지의 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고 상기 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림의 길이는 100 주위부터 15000 주위까지의 뉴클레오티드이다. 상기 실시형태에서 사용된 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림은, 예를 들면, 상기에 설명되어 있다.

[0062] 일부 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 바람직하게는 200 주위부터 8000주위까지 뉴클레오티드이고, 더욱 바람직하게는 500 주위부터 5000 주위까지 뉴클레오티드이고, 더욱 더 바람직하게는

[0063] 1000 주위부터 4500 주위의 뉴클레오티드, 가장 바람직하게는 1500 주위부터 4000 주위의 뉴클레오티드, 특히 길이에서 2000 주위부터 3500 주위의 뉴클레오티드, 더욱 특히 2700 주위부터 3300 주위까지, 더욱 더 특히 3200 주위, 가장 특히 3158 뉴클레오티드이다. 본 명세서에서 정의된 바와 같이 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 비번역된 계놈 DNA 서열에 첨가된 임의 링커 서열을 포함하지 않는다.

[0064] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 적어도 100 뉴클레오티드 주위이고 NCAPD2 유전자의 시작 코돈까지 자체의 최대에서 확장한다. 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 적어도 100 뉴클레오티드 주위이고 NCAPD2 유전자의 세 번째 마지막 인트론까지 자체의 최대에서 확장한다. 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 적어도 100 뉴클레오티드 주위이고 NCAPD2 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장한다. 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 적어도 100 뉴클레오티드 주위이고 NCAPD2 유전자의 마지막 인트론까지 자체의 최대에서 확장한다.

[0065] 인간 NCAPD2 유전자 (NCBI 유전자 ID: 9918)는 염색체 12의 bps 6603298 내지 6641132 주위의 인간 DNA에 위치해 있다. 일실시형태에서, 인간 NCAPD2 유전자의 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 인간의 NCAPD2 유전자에 대한 염색체 12 코딩의 6640243 bps 주위까지 자체의 최대에서 스트레치한다(서열번호: 17에서 bp 1119에 해당하는 GAPDH 유전자의 전사 시작에 비례하여 -3414위치).

[0066] 일실시형태에서, 인간의 NCAPD2 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림이 길이는 인간의 NCAPD2 유전자에 대한 염색체 12 코딩의 6639984 bps 주위까지 자체의 최대에서 스트레치한다(서열번호: 17에서 bp 860에 해당하는 GAPDH 유전자의 전사 시작에 비례하여 -3673 위치).

[0067] 일실시형태에서, 인간의 NCAPD2 유전자의 세 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 인간의 NCAPD2 유전자에 대한 염색체 12 코딩의 6639125 bps 주위까지 자체의 최대에서 스트레치한다(서열번호: 17에서 bp 1에 해당하는 GAPDH 유전자의 전사 시작에 비

례하여 -4532 위치).

[0068] 인간의 NCAPD2 유전자의 마지막 인트론, 두 번째 마지막 인트론 및 세 번째 마지막 인트론까지, 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호: 17에 각각 포함되고, 이는 염색체 12의 bps 6657230 내지 6639125(NCBI 유전자 ID: 9918)를 나타낸다.

[0069] 마우스 NCAPD2 유전자(Gene ID: 68298)는 염색체 6의 125118025 내지 125141604 위치 주위의 마우스 DNA에 위치해 있다. 일실시형태에서, 마우스의 NCAPD2 유전자의 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는(전사 시작의 500 bps 업스트림의 길이를 가지는 것으로 추정됨) 마우스의 NCAPD2 유전자에 대한 염색체 6 코딩의 bps 125118607 주위까지 자체의 최대에서 스트레치한다.

[0070] 일실시형태에서, 마우스의 NCAPD2 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 마우스의 NCAPD2 유전자에 대한 염색체 6 코딩의 125118880 bps 주위까지 자체의 최대에서 스트레치한다.

[0071] 일실시형태에서, 마우스의 NCAPD2 유전자의 세 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 마우스의 NCAPD2 유전자에 대한 염색체 6 코딩의 125119832 bps 주위까지 자체의 최대에서 확장한다.

[0072] 마우스의 NCAPD2 유전자의 마지막 인트론, 두 번째 마지막 인트론 및 세 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호: 18에 각각 포함되고, 이는 염색체 6의 bps 125103521 내지 125119832(NCBI 유전자 ID: 68298)를 나타낸다. 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호: 18(마우스 GAPDH mRNA의 전사 시작에 비례하여 -3006)에 의해 나타낸 바와 같이 뉴클레오티드 서열의 bp1226 주위까지 스트레치한다. 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 18에 의해 나타낸 바와 같이 뉴클레오티드 서열의 bps 953 주위까지 스트레치한다(마우스 GAPDH mRNA의 전사 시작에 비례하여 -3279). 세 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 18에 나타낸 바와 같이 뉴클레오티드 서열의 bp 1 주위까지(마우스 GAPDH mRNA의 전사 시작에 비례하여 -4231)스트레치한다.

[0073] 래트 NCAPD2 유전자(Gene ID: 362438)는 염색체 4의 161288671 내지 161310417 위치 주위의 진핵생물 DNA에 위치해 있다. 일실시형태에서, 래트의 NCAPD2 유전자의 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 래트의 NCAPD2 유전자에 대한 염색체 4 코딩의 161289191 bps 주위까지 자체의 최대에서 스트레치한다. 일실시형태에서, 래트의 NCAPD2 유전자의 두 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 래트의 NCAPD2 유전자에 대한 염색체 4 코딩의 161289446 bps 주위까지 자체의 최대에서 스트레치한다. 일실시형태에서, 래트의 NCAPD2 유전자의 세 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 래트의 NCAPD2 유전자에 대한 염색체 4 코딩의 161290508 bps 주위까지 자체의 최대에서 스트레치한다.

[0074] 래트의 NCAPD2 유전자의 마지막 인트론, 두 번째 마지막 인트론 및 세 번째 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호: 19에 각각 포함되고, 염색체 4의 bps 161279451 내지 161290508(NCBI 유전자 ID: 362438)를 나타낸다. 마지막 인트론을 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호: 19에 나타낸 바와 같이 뉴클레오티드 서

열의 bps 1318 주위까지 스트레치한다(래트 GAPDH mRNA의 전사 시작에 비례하여 -3101). 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 19에 의해 나타낸 바와 같은 뉴클레오티드 서열의 bps 1063 주위까지 스트레치한다(래트 GAPDH mRNA의 전사시작에 비례하여 -3356 위치). 세 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 19에 의해 나타낸 바와 같은 뉴클레오티드 서열의 bp 1 주위까지 스트레치한다(래트 GAPDH mRNA의 전사 시작에 비례하여 -4418 위치).

[0075] 차이니즈 햄스터 NCAPD2 유전자(Gene ID: 100753087)는 3544184 내지 3569879 위치 주위의 진핵생물 DNA에 위치해 있다. 염색체의 위치는 NCBI 데이타베이스 상에서 활용되지 않는다. 일실시형태에서, 차이니즈 햄스터의 NCAPD2 유전자의 마지막 인트론까지 자체의 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 차이니즈 햄스터에서 3569380 bps 주위에서 자체 최대에서 스트레치한다. 일실시형태에서, 차이니즈 햄스터의 NCAPD2 유전자의 두 번째 마지막 인트론까지 자체 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 차이니즈 햄스터의 3569131 bps 주위까지 자체 최대에서 스트레치한다. 일실시형태에서, 차이니즈 햄스터의 NCAPD2 유전자의 세 번째 마지막 인트론까지 자체 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 차이니즈 햄스터의 3567932 bps 주위까지 자체 최대에서 스트레치한다.

[0076] 차이니즈 햄스터의 NCAPD2 유전자의 마지막 인트론, 두 번째 마지막 인트론 및 세 번째 마지막 인트론까지 자체 최대에서 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 각각 서열번호: 29에 포함되고, 이는 bps 3567932 내지 3585061를 나타낸다. 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 서열번호: 29에 의해 나타낸 바와 같이 뉴클레오티드 서열의 bps 1449 주위까지 스트레치한다(차이니즈 햄스터 GAPDH mRNA의 전사시작에 비례하여 -2752). 두 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 29(차이니즈 햄스터 GAPDH mRNA의 전사시작에 비례하여 -3001 위치)에 의해 나타낸 바와 같은 뉴클레오티드 서열의 bps 1200 주위까지 스트레치한다. 세 번째 마지막 인트론까지 확장하는 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 서열번호: 29에 의해 나타낸 바와 같은 뉴클레오티드 서열의 bp 1까지 스트레치한다(차이니즈 햄스터 GAPDH mRNA의 전사시작에 비례하여 -4200 위치).

[0077] 일부 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 -500 주위 위치의 뉴클레오티드부터 -3500주위 위치의 뉴클레오티드까지 영역거리 내, 바람직하게는 -576 주위 위치의 뉴클레오티드부터 -3500 주위 위치의 뉴클레오티드까지 영역거리 내에서 항상 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사시작에 비례한다.

[0078] 일부 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 -500 주위 뉴클레오티드 위치에서, 바람직하게는 -576 주위 뉴클레오티드 위치에서 항상 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사시작에 비례한다.

[0079] 인간에서 인간 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 뉴클레오티드 위치 -463 주위에서 시작한다(서열번호: 17에 나타낸 바와 같은 bp 4533에 해당하는 GAPDH mRNA의 전사시작에 비례하여). 바람직하게는, 만약 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림이 인간으로부터이면, GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 -500 주위에서 시작한다(서열번호: 17에 나타낸 바와 같이 bp 4533에 해당하는; GAPDH mRNA의 전사시작에 비례하여). 더욱 바람직하게는, 만약 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림이 인간으로부터이면, GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 -576 주위에서 시작하고(GAPDH mRNA의 전사시작에 비례하여; 서열번호: 17에 나타낸 바와 같이 bp 4533에 해당하는) 자체의 길이는 서열번호: 17에 나타낸 바와 같이 bps 800 주위부터 3957주위까지 서열에 해당하는 3158 bps 주위이다.

[0080]

추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 서열번호: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 및 28 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열, 바람직하게는 서열번호: 7, 9, 10, 11, 12, 13, 14, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열, 또는 서열번호: 20, 22, 23, 24, 25, 26, 27, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열을 포함한다. 더욱 바람직한 것은 서열번호: 10, 12, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열, 더욱 바람직하게는 서열번호: 10, 12, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열이고, 상기 서열번호: 10 및/또는 16을 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 관련하여 반대 방향으로 지향되고, 서열번호: 12 및/또는 15를 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열로서 동일한 방향으로 지향된다. 마찬가지로 더욱 바람직한 것은 서열번호: 23, 25, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열, 더욱 바람직하게는 서열번호: 23, 25, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열이고, 상기 서열번호: 23 및/또는 16을 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 관련하여 반대 방향으로 지향되고 서열번호: 25 및/또는 28을 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열로서 동일한 방향으로 지향된다.

[0081]

추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 서열번호: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 및 28 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 상호보완라는 뉴클레오티드 서열, 바람직하게는 서열번호: 7, 9, 10, 11, 12, 13, 14, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 상호보완적인 뉴클레오티드 서열, 또는 서열번호: 20, 22, 23, 24, 25, 26, 27, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 상호보완적인 뉴클레오티드 서열을 포함한다. 더욱 바람직한 것은 서열번호: 10, 12, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 상호보완되는 뉴클레오티드 서열이다. 마찬가지로 더욱 바람직한 것은 서열번호: 23, 25, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 상호보완적인 뉴클레오티드 서열이다.

[0082]

추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 서열번호: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 및 28 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 적어도 80% 일치하는 뉴클레오티드 서열을 포함하고, 바람직하게는 서열번호: 7, 9, 10, 11, 12, 13, 14, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 적어도 80% 일치하는 뉴클레오티드 서열, 또는 서열번호: 20, 22, 23, 24, 25, 26, 27, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 적어도 80% 일치하는 뉴클레오티드 서열을 포함한다. 더욱 바람직한 것은 서열번호: 10, 12, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 80% 일치하는 뉴클레오티드 서열, 더욱 바람직하게는 서열번호: 10, 12, 15 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 80% 일치하는 뉴클레오티드 서열이고, 상기 서열번호: 10 및/또는 16을 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 관련하여 반대 방향으로 지향되고, 서열번호: 12 및/또는 15를 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열로서 동일한 방향으로 지향된다. 마찬가지로 더욱 바람직한 것은 서열번호: 23, 25, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 적어도 80% 일치하는 뉴클레오티드 서열이고, 더욱 바람직하게는 서열번호: 23, 25, 28 및 16 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열에 적어도 80% 일치하는 뉴클레오티드 서열이고, 상기 서열번호: 23 및/또는 16을 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열과 관련하여 반대 방향으로 지향되고, 서열번호: 25 및/또는 28을 포함하는 뉴클레오티드 서열은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열로서 동일한 방향으로 지향된다.

[0083]

일부 실시형태에서, 서열번호: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 및 28 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열은 5개 또는 미만, 바람직하게는 4개 또는 미만, 더욱 바람직하게는 3개 또는 미만, 가장 바람직하게는 2개 또는 미만, 특히 핵산 변형이고, 상기 핵산 변형은 바람직하게는 핵산 치환이다.

- [0084] 일부 실시형태에서, 서열번호: 7, 9, 11, 14, 20, 22, 24 및 27 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열은 5개 또는 미만, 바람직하게는 4개 또는 미만, 더욱 바람직하게는 3개 또는 미만, 가장 바람직하게는 2개 또는 미만, 특히 핵산 변형을 포함하고, 상기 핵산 변형은 바람직하게는 핵산 치환이다.
- [0085] 일부 실시형태에서, 서열번호: 7, 9, 11, 14, 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열은 서열번호: 7, 9, 11, 14의 뉴클레오티드 서열의 시작에 비례하여 16 위치에서 하나의 핵산 치환을 포함한다. 바람직하게는 뉴클레오티드 서열의 시작에 비례하는 16 위치에서 G는 T로 대체될 수 있다.
- [0086] 일부 실시형태에서, 서열번호: 20, 22, 24 및 27 또는 이의 단편으로 구성된 군으로부터 선택되는 뉴클레오티드 서열은 서열번호: 20, 22, 24 및 27의 뉴클레오티드 서열의 시작에 비례하여 13 위치에서 하나의 핵산 치환을 포함한다. 바람직하게는 뉴클레오티드 서열의 전사에 비례하여 13 위치의 G는 T로 대체된다.
- [0087] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열로서 동일한 방향으로 지향된다.
- [0088] 추가 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열에 비례하여 반대 방향으로 지향된다.
- [0089] 바람직한 실시형태에서, 발현 카세트는 상기에서 설명된 바와 같이 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림을 포함한다. 바람직하게는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 기원은 동일하고, 예를 들면, 동일한 종(species)이다. 더욱 바람직하게는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림, 진핵생물 GAPDH 프로모터 및 숙주 세포의 비번역된 계놈 DNA 서열 업스트림의 기원은 동일하고, 예를 들면, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림 및 숙주 세포의 기원은 동일한 포유동물이고, 예를 들면, 인간으로부터이다.
- [0090] 일부 실시형태에서, 만약 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림이 하나의 종으로부터 비번역된 계놈 DNA 서열이면, 발현 카세트의 프로모터는 동일한 동일한 종으로부터의 GAPDH 프로모터가 아니다.
- [0091] 일부 실시형태에서, 만약 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림이 인간 기원의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림이면, 발현 카세트의 프로모터는 인간 GAPDH 프로모터가 아니다.
- [0092] 일부 실시형태에서, 발현 카세트의 프로모터는 GAPDH 프로모터가 아니다.
- [0093] 일실시형태에서, 만약 발현 카세트가 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 및 진핵생물 글리세르알데하이드 3-포스페이트 디하드로제네즈(Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)) 프로모터를 포함한다면, 상기 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치부터 +7000 주위의 뉴클레오티드 위치까지 영역거리 내에서 시작하고, 뉴클레오티드 위치는 GAPDH mRNA의 전사시작에 비례하고, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림의 길이는 100 주위부터 15000 주위의 뉴클레오티드까지이

고, 상기 발현 카세트는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림을 추가적으로 포함하고, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터의 5' 말단 주위부터 -3500 주위의 뉴클레오티드 위치까지의 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림의 길이는 100 주위부터 15000 주위의 뉴클레오티드까지이면, 발현 카세트의 프로모터는 진핵생물 GAPDH 프로모터, 바람직하게는 포유동물 GAPDH 프로모터, 더욱 바람직하게는 설치류(설치류) 또는 인간 GAPDH 프로모터일 수 있다. 상기 실시형태에서, 진핵생물 GAPDH 프로모터의 5' 말단 주위부터 -3500 주위의 뉴클레오티드 위치까지 영역거리 내에서 시작하는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 업스트림은 바람직하게는 진핵생물 GAPDH 프로모터의 업스트림에 직접적으로 위치해 있고, 더욱 바람직하게는 상기 실시형태에서 발현 카세트는 진핵생물 GAPDH 프로모터를 포함하는 자연적으로 발생하는 계놈 DNA 서열 및 -3500 주위의 뉴클레오티드 위치까지 확장을 포함하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례한다.

[0094] 일부 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림은 포유동물 기원이고, 예를 들면, 진핵생물 GAPDH 프로모터는 포유동물 GAPDH 프로모터이고 포유동물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림은 본 명세서에서 설명된 바와 같이 사용될 수 있다.

[0095] 일부 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림은 설치류 또는 인간 기원이고, 예를 들면, 진핵생물 GAPDH 프로모터는 설치류 또는 인간 GAPDH 프로모터이고 설치류 또는 인간 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림은 본 명세서에서 설명된 바와 같이 사용된다.

[0096] 바람직하게는 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림은 인간, 래트 또는 마우스 기원, 더욱 바람직하게는 인간 또는 마우스 기원으로부터, 가장 바람직하게는 인간 기원으로부터 선택된다.

[0097] 일부 실시형태에서, 진핵생물 GAPDH 프로모터의 비번역된 계놈 DNA 서열 다운스트림 및/또는 업스트림은 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열에 실시가능하게 연결되지 않는다.

[0098] 일부 실시형태에서, 발현 카세트는 폴리아데닐레이션 자리를 포함한다. 바람직하게는 폴리아데닐레이션 자리는 SV40 poly(A) 및 BGH(Bovine Growth Hormone) poly(A)로 구성된 군으로부터 선택된다.

[0099] 일부 실시형태에서, 발현 카세트의 프로모터 및 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열은 실시가능하도록 연결된다.

[0100] 일부 실시형태에서, 발현 카세트의 프로모터는 SV40 프로모터, 인간 tk 프로모터, MPSV 프로모터, 마우스 CMV, 인간 CMV, 래트 CMV, 인간 EF1알파, 차이니즈 햄스터 EF1알파, 인간 GAPDH, MYC, HYK 및 CX 프로모터를 포함하는 하이브리드 프로모터로 구성된 군으로부터 선택된다.

[0101] 일부 실시형태에서, 발현 카세트에 의해 암호화되는 폴리펩티드는 비당화(non-glycosylated) 및 당화된 (glycosylated) 폴리펩티드일 수 있다. 당화된 폴리펩티드는 적어도 하나의 올리고사카라이드 (oligosaccharide) 사슬을 갖는 폴리펩티드를 의미한다.

[0102] 비당화된 단백질의 예로는, 예를 들면, 비당화된 호르몬; 비당화된 효소; 신경 성장 요인(NGF) 페밀리, 상피 성

장 요인(EGF) 및 섬유모세포 성장요인(fibroblast growth factor(FGF))페밀리의 비당화된 성장 요인 및 호르몬 및 성장요인에 대한 비당화된 수용체이다.

[0103] 당화된 단백질의 예로는 호르몬 및 호르몬 방출요인, 응고요인(clotting factors), 항응고 요인, 호르몬 또는 성장요인에 대한 수용체, 신경영양성 요인 사이토카인(neurotrophic factors cytokines) 및 이의 수용체, T-세포 수용체, 표면 메브린 단백질, 운반 단백질, 귀소수용체(homing receptors), 어드레신(addressins), 조절 단백질, 항체, 면역접합체(immunoadhesins)와 같은 키메릭 단백질, 및 당화된 단백질의 임의 단편이다. 바람직하게는 항체, 항체 단편 또는 항체 유도체(예를 들면, Fc 융합 단백질 및 양특이성 항체(bispecific antibodies)와 같은 특정 항체 포맷)로 구성된 군으로부터 선택된다. 본 명세서에서 사용된 항체 단편은 (i) 도메인, (ii) Fab' 및 Fab'-SH를 포함하여, VL, VH, CL 또는 CK 및 CH1 도메인을 구성하는 Fab 단편, (iii) VH 및 CH1 도메인을 구성하는 Fd 단편, (iv) 단일 가변 도메인(variable domain)으로 구성된 dAb 단편(ard ES *et al.*, (1989) *Nature*, 341(6242): 544-6) (v) F(ab')2 단편, 2 개 연결된 Fab 단편을 포함하는 이가단편(bivalent fragment) (vi) 단일 사슬 Fv 분자 (scFv), 상기 VH 도메인 및 VL 도메인은 항원결합 자리를 형성하기 위해 연합된 두 개의 도메인을 가능하게 하는 결합 링커에 의해 연결된다(Bird RE *et al.*, (1988) *Science*, 242(4877): 423-6; Huston JS *et al.*, (1988) *Proc Natl Acad Sci U S A*, 85(16): 5879-83), (vii) "이중체(diabodies)" 또는 "삼중체(triabodies)", 다가(multivalent) 또는 유전자 융합에 의한 다중특정 단편 구조체 (Holliger P *et al.*, (1993) *Proc Natl Acad Sci U S A*, 90(14): 6444-8; Holliger P *et al.*, (2000) *Methods Enzymol*, 326: 461-79), (viii) scFv, 이중체 또는 Fc 영역에 융합된 도메인 항체 및 (ix) 동일한 또는 상이한 항체에 융합된 scFv을 포함하나, 이에 한정하지 않는다.

[0104] 일부 실시형태에서 발현 카세트는 프로모터, 인핸서, 전사 대조군 요소, 및 선별가능한 마커, 바람직하게는 동물세포에서 발현되는 선별가능한 마커로 구성된 군으로부터 선택되는 유전적 요소를 추가적으로 포함한다. 전사 대조군 요소, 예를 들면, 코작(Kozak) 서열 또는 전사 종결자 요소(transcriptional terminator element)이다.

[0105] 일실시형태에서, 유전적 요소는 선별가능한 마커를 암호화하는 폴리뉴클레오티드 서열에 포함된 CpG 자리의 내용이 45 또는 미만, 바람직하게는 40 또는 미만, 더욱 바람직하게는 20 또는 미만, 특히 10 또는 미만, 더욱 특히 5 또는 미만, 가장 특히 0(CpG 자리가 완전히 제거됨)인 선별가능한 마커이다.

[0106] 추가적 측면에서, 본 개시는 발현 벡터, 바람직하게는 상기에서 설명된 바와 같이 발현 카세트를 포함하는 포유동물 발현 벡터를 제공한다. 일부 실시형태에서, 발현 벡터는 적어도 두 개의 별도의 전사 유닛을 포함한다. 두 개의 별도의 전사 유닛을 갖는 발현 벡터는 또한 이중-유전자 벡터로서 언급된다. 이의 실시예로는 첫 번째 전사 유닛은 항체 또는 이의 단편의 중쇄를 암호화하고 두 번째 전사 유닛은 항체의 경쇄를 암호화하는 벡터이다. 다른 실시예로는 두 개의 전사 유닛이 효소와 같은 단백질의 두 개의 상이한 서부유닛을 암호화하는 이중-유전자 벡터이다. 그러나, 본 발명의 발현 벡터는 두 개 이상의 별도의 전사 유닛, 예를 들면, 3개, 4개 또는 그 이상의 별도의 상이한 폴리펩티드 사슬을 암호화하는 다른 뉴클레오티드 서열을 포함하는 각각의 전사유닛을 포함하는 것이 또한 가능하다. 따라서, 실시예는 각각은 4개의 상이한 서브유닛으로 구성된 효소의 하나의 서브유닛을 암호화하는 상이한 뉴클레오티드 서열을 포함하는 4 개의 별도의 전사 유닛을 갖는 벡터이다.

[0107] 일부 실시형태에서, 발현 벡터는 추가 프로모터, 인핸서, 전사 대조군 요소, 복제 기원 및 선별가능한 마커로 구성된 군으로부터 선택되는 유전적 요소를 추가적으로 포함한다.

[0108] 일부 실시형태에서, 발현 벡터는 복제 기원 및 선별가능한 마커를 암호화하는 발현 벡터의 폴리뉴클레오티드 서열에 포함된 CpG 자리의 내용이 200 또는 미만, 바람직하게는 150 또는 미만, 특히 100 또는 미만, 더욱 특히 50 또는 미만, 가장 특히 30 또는 미만인 복제 기원 및 선별가능한 마커를 추가적으로 포함한다.

- [0109] 티미딘 키나제(thymidine kinase(tk)), 디하이드로풀레이트 리덕타제(dihydrofolate reductase (DHFR)), 퓨로마이신, 네오마이신 또는 글루타민 합성효소(glutamine synthetase(GS))와 같은 일반적으로 사용되는 임의 선별 가능한 마커는 본 발명의 발현 카세트 또는 발현 벡터에 사용될 수 있다. 바람직하게는, 본 발명의 발현 벡터는 본 발명의 이종 단백질의 배출(secretion)에 대한 발현 카세트의 삽입체를 위한 유용한 제한효소 자리의 제한된 수를 또한 구성한다. 단지 일시적/유전차 부체(episomal) 발현에 특히 사용되는 경우에, 본 발명의 발현 벡터는 진핵생물 숙주 세포에서 자율복제/유전자 부체 유지를 위한 앱스타인-바바이러스(Epstein Barr Virus(EBV)) 또는 SV40 바이러스의 or iP 기원과 같은 복제의 기원이 추가적으로 포함될 수 있으나 선별 가능한 마커가 결핍될 수 있다. 벡터의 복제를 용이하게 하는 관련 요인이 부족한 세포에서 일시적 발현이 또한 가능하다.
- [0110] 발현 카세트를 정박하는 발현 벡터는 형광마커, ncRNA를 코딩하는 발현 카세트, 항세포사멸 단백질을 코딩하는 발현 카세트, 또는 배출 경로의 용량을 증가시키는 단백질을 코딩하는 발현 카세트를 추가적으로 포함할 수 있다.
- [0111]
- [0112] 추가적 측면에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:
- [0113] a) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림
- [0114] b) 프로모터
- [0115] c) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열
- [0116] d) 폴리아데닐레이션 자리
- [0117] e) 인핸서
- [0118] f) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림, 또는
- [0119] a) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림
- [0120] b) 인핸서
- [0121] c) 프로모터
- [0122] d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열
- [0123] e) 폴리아데닐레이션 자리
- [0124] f) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림, 또는
- [0125] a) 인핸서
- [0126] b) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림 및/또는 다운스트림
- [0127] c) 프로모터
- [0128] d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열
- [0129] e) 폴리아데닐레이션 자리
- [0130] f) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림,
- [0131] 만약 a) 또는 b)가 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림이면 f)는 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림이고 만약 a) 또는 b)가 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림이면 f)는 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림이라는 조건으로, 상기 인핸서의 포함은 선택적이고, 상기 폴리뉴클레오티드 서열에 의해 암호화되는 폴리펩티드는 GAPDH가 아니고, 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은 +1 주위의 뉴클레오티드 위치부터 +7000 주위의 뉴클레오티드 위치까지의 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사시작에 비례하고, 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림의 길이는 100 주위부터 15000 주위 뉴클레오티드까지이고, 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림은 진핵생물 GAPDH 프로모터의 5' 말단 주위부터 -3500

주위의 뉴클레오티드 위치까지 영역거리 내에서 시작하고, 상기 뉴클레오티드 위치는 GAPDH mRNA의 전사 시작에 비례하고, 상기 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림의 길이는 100 주위부터 15000 주위의 뉴클레오티드까지이다.

[0132] 일부 실시형태에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:

a) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림

b) 프로모터

c) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열

d) 폴리아데닐레이션 자리

e) 인핸서

f) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림, 상기 인핸서의 포함은 선택적이다.

[0139] 추가적 측면에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:

a) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림

b) 인핸서

c) 프로모터

d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열

e) 폴리아데닐레이션 자리

f) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림, 상기 인핸서의 포함은 선택적이다.

[0146] 추가적 측면에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:

a) 인핸서

b) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림

c) 프로모터

d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열

e) 폴리아데닐레이션 자리

f) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림, 상기 인핸서의 포함은 선택적이다.

[0153] 추가적 측면에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:

a) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림

b) 프로모터

c) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열

d) 폴리아데닐레이션 자리

e) 인핸서

f) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 업스트림, 상기 인핸서의 포함은 선택적이다.

[0160] 추가적 측면에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:

a) 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림

b) 인핸서

c) 프로모터

d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열

e) 폴리아데닐레이션 자리

f) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림, 상기 인핸서의 포함은 선택적이다.

[0167] 추가적 측면에서, 본 개시는 하기의 순서를 포함하는 발현 벡터를 제공한다:

a) 인핸서

b) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 다운스트림

c) 프로모터

d) 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열

e) 폴리아데닐레이션 자리

f) 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림, 상기 인핸서의 포함은 선택적이다.

[0174] 진핵생물 GAPDH의 비번역된 게놈 DNA 서열 업스트림, 인핸서, 프로모터, 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열, 폴리아데닐레이션 자리 및 발현 벡터의 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림은, 예를 들면, 상기에서 설명된 바와 같다.

[0175] 추가적 측면에서, 본 개시는 상기에서 설명된 바와 같이 발현 카세트 또는 발현 벡터를 포함하는 숙주 세포를 제공한다. 숙주 세포는 인간 또는 비인간 세포일 수 있다. 바람직한 숙주 세포는 포유동물 세포이다. 포유동물 숙주 세포의 바람직한 실시예는, 제한 없이, 인간 배아 신장(human embryonic kidney) 세포(Graham FL et al., J. Gen. Virol. 36: 59-74), MRC5 인간 섬유모세포(fibroblasts), 983M 인간 흑색종(melanoma) 세포, MDCK 카닌 신장(canine kidney)세포, 스프라그-다우리 래트로부터 분리된 RF 배양된 래트 폐 섬유모세포, B16BL6 마우스 흑색종(murine melanoma) 세포, P815 마우스 비만 세포, MT1 A2 마우스 유방 샘암종(murine mammary adenocarcinoma) 세포, PER:C6 세포(Leiden, Netherlands) 및 차이니즈 햄스터 난소(CHO) 세포 또는 세포주(Puck et al., 1958, J. Exp. Med. 108: 945-955)를 포함한다.

[0176] 특히 바람직한 실시형태에서 숙주 세포는 차이니즈 햄스터 난소(CHO) 세포 또는 세포주이다. 적합한 CHO 세포 주는, 예를 들면, CHO-S(Invitrogen, Carlsbad, CA, USA), CHO K1(ATCC CCL-61), CHO pro3-, CHO DG44, CHO P12 또는 dhfr- CHO 세포주 DUK-BII(Chasin et al., PNAS 77, 1980, 4216-4220), DUXBI 1(Simonsen et al., PNAS 80, 1983, 2495-2499), 또는 CHO-K1SV(Lonza, 베이스1, Switzerland)를 포함한다.

[0177] 추가적 측면에서, 본 개시는 상기에서 설명된 바와 같이 발현 카세트 또는 발현 벡터를 갖는 숙주 세포를 형질 전환시키는 것 및 폴리펩티드를 회수하는 것을 포함하는, 폴리펩티드의 발현을 위한 시험관 내 방법을 제공한다. 상기 폴리펩티드는 바람직하게는 이종(heterologous), 더욱 바람직하게는 인간 폴리펩티드이다.

[0178] 발현 카세트 또는 발현 벡터를 본 발명에 따른 숙주세포로 형질전환을 위하여, 당업계에서 잘 알려진 기술, 예를 들면, 전기천공법(electroporation), 칼슘포스페이트 공용침전법(calcium phosphate co-precipitation), DEAE-덱스트란 형질전환, 리포펙틴과 같은 임의 형질전환 기술은 만약 주어진 숙주 세포 형태가 적절하다면 사

용될 수 있다. 본 발명의 발현 카세트 또는 발현 벡터를 갖는 형질전환된 숙주 세포는 일시적 또는 안정적으로 형질전환된 세포주로서 해석되는 것을 알려져 있다. 따라서, 본 발명에 따라 본 발명의 발현 카세트 또는 발현 벡터는 유전자 부체(episomally)로 유지될 수 있고, 예를 들면, 일시적으로 형질전환될 수 있거나 숙주 세포의 게놈에서 안정적으로 통합(integration)될 수 있고, 예를 들면, 안정적으로 형질전환될 수 있다.

[0179] 일시적 형질전환은 벡터 운반 선별가능한 마커를 위한 임의 선별 압력(selection pressure)의 비장치(non-appliance)에 의해 특징될 수 있다. 형질전환 후 통상적으로 마지막 2일부터 10일 까지 지속하는 일시적 발현 실험에서, 형질전환된 발현 카세트 또는 발현 벡터는 유전자 부체 요소로서 유지되고 게놈으로 통합되지 않는다. 즉 형질전환된 DNA는 항상 숙주 세포 게놈으로 통합되지 않는다. 숙주 세포는 형질전환된 DNA를 손실하는 경향이 있고 일시적으로 형질전환된 세포 풀의 배양에 따른 집단에서 형질전환된 세포를 과성장시키는 경향이 있다. 따라서 발현은 형질전환 즈후에 가장 강하고 시간이 지나면서 감소한다. 바람직하게는, 본 발명에 따른 일시적 형질전환체는 형질전환 후 2 내지 10일까지 선별 압력의 부재에서 세포 배양에서 유지되는 세포로서 이해된다.

[0180] 본 발명의 바람직한 실시형태에서 숙주 세포 예를 들면, CHO 숙주 세포는 본 발명의 발현 카세트 또는 발현 벡터를 같이 안정적으로 형질전환된다. 안정적 형질전환은 벡터 DNA와 같은 새로이 도입된 외부 DNA가 게놈 DNA로 항상 무작위, 비상동재조합 사건에 의해 혼합되는 것이다. 벡터 DNA의 복제수(copy number) 및 동시에 유전자 생성의 수는 벡터 서열이 숙주 세포의 DNA로 통합 후에 증폭되어 왔다. 따라서, 상기 안정적 통합(integration)은, 유전자 증폭을 위한 선별 압력에서 추가적으로 증가하는 유출에 따라, CHO 세포에서 이중분열색체로 발생한다. 추가적으로, 안정적 형질전환은 예를 들면, 박테리아 복제수 조절 영역이 게놈 통합에 따른 과잉(superfluous)을 표현하는 것과 같은 재조합 유전자 생성의 발현과 직접적으로 관련되지 않은 벡터 서열 부분의 손실을 초래할 수 있다. 따라서, 형질전환된 숙주 세포는 게놈으로 발현 카세트 또는 발현 벡터의 적어도 일부 또는 상이한 일부를 통합한다.

[0181] 추가적 측면에서, 본 개시는 포유동물 숙주 세포로부터 이종 폴리펩티드의 발현을 위한, 특히 포유동물 숙주 세포로부터 이종 폴리펩티드의 시험관 내 발현을 위한 상기에서 설명된 발현 카세트 또는 발현 벡터의 용도를 제공한다.

[0182] 단백질의 발현 및 회수는 당업자에 알려진 방법에 따라 수행될 수 있다.

[0183] 폴리펩티드의 발현을 위하여, 상기에서 설명된 발현 카세트 또는 발현 벡터의 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림 및 상기에서 설명된 숙주 세포사 사용되고 항상 동일한 기원이다. 놀랍게도 발현의 증가는 만약 발현 카세트 또는 발현 벡터 및 숙주 세포의 진핵생물 GAPDH 프로모터의 비번역된 게놈 DNA 서열 다운스트림 및/또는 업스트림이 상이한 기원이라면, 예를 들면, 진핵생물 GAPDH 프로모터의 인간 DNA 서열 다운스트림 및/또는 업스트림이 CHO 세포에서 사용되면 얻어진다.

[0184] 추가적 측면에서, 본 개시는 장애의 치료를 위한 의약 제조를 위한 상기에서 설명된 발현 카세트 또는 발현 벡터의 용도를 제공한다.

[0185] 추가적 측면에서, 본 개시는 장애의 치료를 위한 의약용으로서 상기에서 설명된 발현 카세트 또는 발현 벡터를 제공한다.

[0186] 추가적 측면에서, 본 개시는 유전자 치료에서 용도를 위한 상기에서 언급된 발현 카세트 또는 발현 벡터를 제공한다.

- [0187] 실시예
- [0188] 실시예 1: 발현 벡터의 클로닝:
- [0189] I. 재료 및 방법
- [0190] I.1. 플라스미드 구조체
- [0191] I.1.1. LB 배양 플레이트
- [0192] 500 ml의 물을 16 g의 LB 아가(Invitrogen, Carlsbad, CA, USA)와 혼합하고 끓였다(1 litre의 LB는 10 g 트립톤, 5 g 이스트 추출물 및 10 g NaCl을 포함한다). 냉각 후, 각각의 항생물질을 플레이트 웰 용액(100 µg/ml 엠피실린 플레이트 및 50 µg/ml 카나마이신 플레이트)으로 첨가하였다.
- [0193] I.1.2. 중합효소연쇄반응 (PCR)
- [0194] 모든 PCR은 50 µl 최종 부페에서 1 µl의 dNTPs(각각의 dNTP에 대해 10 mM; Invitrogen, Carlsbad, CA, USA), 2 유닛의 Phusion®DNA 중합효소 (Finnzymes Oy, Espoo, Finland), 25 nmol의 프라이머 A(Mycrosynth, Balgach, Switzerland), 25 nmol의 프라이머 B (Mycrosynth, Balgach, Switzerland), 1.5 µl의 디메틸суль포사이드(DMSO, Finnzymes, Espoo, Finland) 및 1-3 µl의 템플레이트(1-2 µg)을 사용하여 수행하였다. 모든 프라이머는 표1에 나타냈다.
- [0195] 상기 PCR은 98°C에서 3분 동안 초기 변성(denaturation)에 의해 시작되었고, 98°C에서 30초 변성, 프라이머-특정 온도(CG 함량에 따라)에서 30초 풀림(annealing) 및 72°C에서 2 분 연장(elongation)의 35 사이클을 하였다. 10 분 동안 72°C에서 최종 연장을 냉각 전에 수행하였고 4°C에 보관하였다.
- [0196] 표 1: PCRs에 사용된 프라이머의 요약 GAPDH: 글리세르알데하이드 3-포스페이트 디하이드로제네즈 (glyceraldehyde 3-phosphate dehydrogenase) 서열, 5': 업스트림 서열, 3': 다운스트림 서열 프라이머 GlnPr1172에 있는 "T" (밀줄침)은 프라이머 다이머의 형성을 피하기 위하여 도입되었다.
- 표 1**
- | 프라이머 | 프라이머 서열 | 서열 종목 | Seq ID |
|------------|--|----------|---------|
| GlnPr 1171 | ATTATTCCGCGATGGCTCCTGGCATCTCTGGGACCGAGGC | 5' GAPDH | 서열번호: 1 |
| GlnPr 1172 | ATCGTCGCGAAGCTTGAGATTGTCCAAGCAGGTAGCCAG | | 서열번호: 2 |
| GlnPr 1173 | AGCAAGTACTTCTGAGCCTTCAGTAATGGCTGCCTG | 3' GAPDH | 서열번호: 3 |
| GlnPr 1174 | TGGCAGTACTAAGCTGGCACCCTACTTCAGAGAACAAAG | | 서열번호: 4 |
- [0198] I.1.3. 제한 분해(Restriction digest)
- [0199] 모든 제한 분해를 위하여, 약 1 µg의 플라스미드 DNA(나노드롭으로 정량화, ND-1000 분광강도계(Thermo Scientific, Wilmington, DE, USA))를 각각의 10-20 유닛효소, 4 µl의 해당 10 X NE버퍼(NEB, Ipswich, MA, USA)와 혼합하였고 부피는 물로 40 µl로 채웠다. 추가 지시 없이, 분해를 37°C 1 시간 배양하였다.
- [0200] 각각의 백본(backbone)의 예비 분해 후에, 1 유닛의 송아지 장 알칼린 포스파타제(CIP; NEB, Ipswich, MA, USA)를 첨가하고 37°C에서 30 분 배양하였다.
- [0201] 만약 분해가 NEB버퍼 3(NEB, Ipswich, MA, USA)에서 완성되면, 상기 효소는 상기 버퍼에서 강한 활성을 가지고 외부 말단의 뉴클레오티드의 일부를 또한 분해할 수 있기 때문에 상기 버퍼를 CIP 첨가전에 NEB 버퍼 4로 변경

하였다

[0202] I.1.4. PCR 정제 및 아가로우즈 젤 전기이동(agarose gel electrophoresis)

[0203] I.1.4.1. PCR 클린업(clean up)

[0204] 분해를 하기 위해서 모든 PCR 단편은 맥커리 나겔 추출 II 키트(Macherey Nagel Extract II kit)(Macherey Nagel, Oensingen, Switzerland)를 사용하여 40 μl 의 용출버퍼를 사용하여 제조사의 매뉴얼을 따라 제한 분해 전에 클린되었다. 상기 프로토콜은 DNA 시료의 변경버퍼(changing buffer)를 위해 또한 사용되었다.

[0205] I.1.4.2. DNA 추출

[0206] 젤 전기이동을 위하여, 1% 젤을 울트라퓨어TM 아가로우즈(UltraPureTM Agarose)(Invitrogen, Carlsbad, CA, USA) 및 50X 트리스아세트산(Tris Acetic Acid)EDTA 버퍼(TAE, pH 8.3; Bio RAD, Munich, Germany)을 사용하여 제조하였다. DNA 염색을 위하여 1 μl 의 젤 레드 다이(Gel Red Dye (Biotum, Hayward, CA, USA))를 100 $\text{m}\mu\text{l}$ 의 아가로우즈 젤에 첨가하였다. 사이즈 마커로서 2 μg 의 1 kb DNA 래더(NEB, Ipswich, MA, USA)를 사용하였다. 전기이동은 125 볼트에서 약 1시간 동안 실행하였다. 관심 밴드는 키트 추출 II(kit Extract II)(Macherey-Nagel, Oensingen, Switzerland)에 이어서 40 μl 의 용출버퍼를 사용하여 제조사의 매뉴얼을 사용하여 아가로우즈 젤로부터 잘랐다.

[0207] I.1.5. 결찰(Ligation)

[0208] 각각의 결찰을 위하여, 4 μl 의 삽입체를 10 μl 부피에서 1 μl 의 벡터, 400 유닛의 라이게이즈(T4 DNA 라이게이즈, NEB, Ipswich, MA, USA), 1 μl 의 10X 라이게이즈 버퍼(T4 DNA ligase buffer; NEB, Ipswich, MA, USA)에 혼합하였다. 혼합은 RT에서 1-2 시간 배양하였다.

[0209] I.1.6. 컴피턴트 박테리아로 결찰 생성물의 변형

[0210] pGLEX41-[REP]의 클로닝 및 복제의 표준 기원을 포함하는 pCR-블런트 벡터로 만들어진 구조체를 위하여, TOP 10(One Shot[®]TOP 10 컴피턴트 *E. coli*; Invitrogen, Carlsbad, CA, USA)을 사용하였다.

[0211] 복제의 R6K 기원을 포함하는 플라스미드의 복제 개시를 위하여, pir 서열에 의한 코드된 π 단백질의 발현이 요구된다. π 단백질은 One Shot[®]PIR1 컴피턴트 *E. coli*(Invitrogen, Carlsbad, CA, USA)에 발현된다. 상기 박테리아는 R6K 서열을 포함하는 모든 벡터에 사용된다.

[0212] 결찰 생성물을 갖는 컴피턴트 생성물을 갖는 컴피턴트 박테리아를 변형하기 위하여, 25-50 μl 의 박테리아는 5분 동안 열음 상에서 녹였다. 이후 3-5 μl 의 결찰 생성물을 컴피턴트 박테리아에 첨가하고 42°C에서 1분 동안 열충격(thermic shock) 전에 열음 상에서 20-30 분 동안 배양하였다. 이 후, 500 μl 의 S.O.C 배지 (Invitrogen, Carlsbad, CA, USA)를 투브 당 첨가하였고 교반하에서 37°C에서 1 시간 동안 배양하였다. 최종적으로, 박테리아를 엠피실린이 있는 LB 플레이트(Sigma-Aldrich, St. Louis, MO, USA) 상에 넣고 37°C에서 밤사이 배양하였다. pCR-Blunt 벡터에서 클로닝을 위하여, 카나마이신이 있는 플레이트(Sigma-Aldrich, St. Louis, MO, USA)를 사용하였다.

[0213] I.1.7. 소량 (미니(mini)) 및 중급 규모 (미디(midi))의 플라스미드 제조

[0214] I.1.7.1. 미니제조(Minipreparation)

[0215] 미니 제조를 위하여, 형질전환 박테리아의 콜로니를 37°C, 200 rpm에서 LB 및 엠피실린 또는 카나마이신의 2.5 $\text{m}\mu\text{l}$ 에서 6-16 시간 동안 자라게 하였다. DNA를 *E. coli* (QuickPure, Macherey Nagel, Oensingen, Switzerland)

를 위한 플라스미드 정제 키트에 이어서 제공된 매뉴얼로 추출하였다.

[0216] 미니 제조로부터 플라스미드 DNA를 260 nm에서 흡광도를 측정하여 나노드롭 ND-1000 분광강도계 (Thermo Scientific, Wilmington, DE, USA)로 일단 정량화하고 1.8과 2사이이어야만 하는 OD260 nm/OD280 nmA의 비율을 평가하였다. 조절 분해는 서열 확인을 위해 Fasteris SA(Geneva, Switzerland)로 시료를 보내기 전에 수행되었다.

[0217] BAC 추출을 위해, 쿼터 키트(QuickPure kit)(Macherey Nagel, Oensingen, Switzerland)를 프로토콜의 하기의 변형으로 사용하였다: LB 및 클로람페니콜(12.5 μ g/ml)(Sigma-Aldrich, St. Louis, MO, USA)은 pBACe3.6 벡터를 포함하는 박테리아로 분주하였다. 37°C에서 밤사이 동안 흔들기 플랫폼 상에서 배양한 후에 배양은 500 μ l의 A1 베페에서 재현탁하기 전에 배양은 13300 rpm에서 5 분 동안 원심분리하였다. 500 μ l의 A2 용해 베페를 첨가하였고 용액을 상온에서 5 분 동안 배양하였다. 이 후, 600 μ l의 A3 베페로 중성화하고 13300 rpm에서 10 분 원심분리하였다. 상층액을 컬럼상에 로드하고 상기 단계로부터 쿼터 미니프렙 키트의 표준 프로토콜에 따라서 사용하였다.

I.1.7.2. 미디제조(MidiPreparation)

[0219] 미디제조를 위하여, 형질전환된 박테리아를 LB 및 엠피실린(또는 kanamycin)의 200 내지 400 ml에서 밤사이 동안 37°C 밤사이 동안 자라게 하였다. 이 후, 배양은 20 분 동안 725 g에서 원심분리하였고 플라스미드를 상업적 키트(NucleoBond Xtra Midi; Macherey Nagel, Oensingen, Switzerland)에 이어서 제조사의 매뉴얼에 제공된 낫은 플라스미드 프로토콜을 사용하여 정제하였다.

[0220] 미디제조로부터 플라스미드-DNA를 나노드롭 ND-1000 분광강도계로 3회 정량화하고 제한 분해로 확인하고 최종적으로 시퀀싱을 보냈다(Fasteris SA, 유전자va, Switzerland).

II. 결과 및 토론

II.1. GAPDH 발현카세트(5' 및 3' GAPDH)의 DNA 영역 업스트립 및 다운스트립의 클로닝

[0223] BAC 클론 RPCIB753F11841Q을 이메진Imagene (Berlin, Germany)에 주문하였다. 상기 클론은 클로람페니콜(chloramphenicol) 저항 유전자를 포함하는 pBACe3.6 벡터 백본에서 인간 GAPDH 서열을 포함한다. 미디제조 (midipreparation)에 의해 DNA 추출 후에, 벡터 농도를 나노드롭에 의해 27 ng/ μ l으로 결정하였다.

[0224] 프로모터의 GAPDH 발현 카세트 업스트립 및 폴리-아네닐레이션 자리의 다운스트립을 둘러싼 직후의 DNA 서열은 템플레이트로서 27 ng의 정제된 클론 RPCIB753F11841Q을 사용하여 증폭되었다. 프로모터의 3 kb 단편 업스트립은 서열번호: 5을 갖는 엠플리콘으로 유도하는 프라이머 GlnPr1171(서열번호: 1) 및 GlnPr1172(서열번호: 2)으로 증폭되었다. 프라이머 GlnPr1172(서열번호: 2)는 템플레이트 서열에 비례하여 베이스(base) 변경(G에서 T로)을 읊기기 때문에, 상기 PCR 반응으로부터 유래된 모든 서열은 상기 베이스 변경을 역시 읊길 것이다. 상기 변경은 GAPDH 유전자(서열번호: 17의 bp 812, 서열번호: 5의 시작에 비례하여 위치 23)의 전사시작에 비례하여 위치 -3721에 위치해 있다. 폴리아데닐레이션 자리의 3kb 단편 다운스트립은 서열번호: 6(표 1)를 갖는 엠플리콘으로 유도하는 프라이머 GlnPr1173(서열번호: 3) 및 GlnPr1174(서열번호: 4)로 증폭되었다. PCRs에 사용된 풀럼 온도는 72°C이었다.

[0225] 5' 및 3'GAPDH 단편(서열번호: 5 및 6)을 pCR-Blunt, 상업적으로 활용가능한 PCR-생성물 클로닝 벡터(pCR-Blunt, PCR Zero Blunt 클로닝 키트, Invitrogen)에서 클론되었다. 결찰 생성물을 TOP10 컴파턴트 박테리아로 형질전환되었고 카나마이신 LB-아가 플레이트 상에서 분주하였다. 콜로니는 증폭되었고 플라스미드는 미니프렙에 의해 분리되었다. 조절 분해는 pCR-Blunt-5' GAPDH 및 pCR-Blunt-3' GAPDH 구조체를 수율하는 양성 클론을

식별하기 위하여 수행되었다.

[0226] II.2. 리포터 단백질 GFP 및 재조합 IgG1 단클론 항체(LC-IRES-HC-IRES-GFP)를 위한 DNA 단편 코딩의 제조

본 발명에서 사용된 리포터 구조체(REP)는 폴리시스트로닉 유전자로 구성된다: IgG1 단클론 항체 경쇄 (LC)-IRES- IgG1 단클론 항체 중쇄(HC)-IRES-녹색형광 단백질(GFP). 뇌심(장)근(육)염 바이러스 (Encephalomyocarditis virus)로부터 유래된 내부 리보솜 엔트리 사이트(Internal Ribosomal Entry Sites(IRES)(Gurtu *et al.*, Biochem Biophys Res Commun.; 229(1): 295-298, 1996))는 3 웨이드 IgG1 단클론 항체 경쇄 (LC), IgG1 단클론 항체 중쇄 (HC) 및 GFP (도 1)의 번역을 허가하게 한다. 따라서, 형질전환된 세포는 IgG1 단클론 항체를 배출할 것이고 의존적 방법으로 세포 내 GFP를 축적한다. 그러나, 폴리크리토닉 mRNA는 진핵생물 세포에서 일반적이지 않고 자체의 번역은 상대적 낮은 적정량의 IgG1 및 GFP 발현을 유도하여, 매우 효율적이지 않다. REP 구조체를 포함하는 벡터는 제한효소 NheI 및 BstBI(BstBI는 65°C에서 사용됨)을 사용하여 절단된다. 발현 구조체를 포함하는 REP 단편은 자르고, 추가적으로 클로닝 단계에 사용된다.

[0228] II.3. 발현 벡터의 클로닝

벡터 pGLEX41, pcDNA3.1 (+) (Invitrogen, Carlsbad, CA)로부터 유래된 발현 벡터는 안정 세포주 생성에 사용된다. GAPDH 서열이 있고 없이 벡터 A 및 B의 두 번째 생성을 생성시키는 변형된 초기 백본으로서 사용되었다. 모든 벡터를 위하여 동일한 프로모터-인트론 조합(mCMV 및 첫 번째 인트론 (IgDA)을 위한 기증자-수용체 단편 코딩)이 사용되었다(Gorman *et al.*, (1990) Proc Natl Acad Sci USA, 87: 5459-5463).

[0230] 중간 벡터 pGLEX41-HM-MCS-ampiA의 클로닝:

새로운 벡터 생성의 개발은 pGLEX41로부터 시작되었다. 상기 벡터를 엠피실린 저항 카세트를 방출하기 위하여 제한효소 NruI 및 BspHI를 사용하여 자른다. 백본 단편은 CIP되고 젤 전기이동에 의해 정제된다. 엠피실린 저항 유전자(블라(bla) 프로모터를 포함하여)의 코돈 최적화 버전(*E. coli*에서 발현을 위하여)을 위한 DNA 단편 코딩은 진아트(GeneArt)로부터 주문되었다. 삽입체는 제한효소 NruI 및 BspHI(백본으로서 사용된 동일한 효소)를 사용하여 진아트(GeneArt) 클로닝 벡터 #1013237 밖으로 잘렸고, 정제되었고 백본으로 클론되었다.

미니프렙은 제한 분해에 의해 분석되었다. 클론 pGLEX41-HM-MCS-ampiA#2는 기대된 제한 프로파일을 가졌고 정확한 단편의 통합은 시퀀싱으로 확인되었다.

[0233] 중간 벡터 pGLEX41-MCS-R6K-ampiA의 클로닝

벡터 pGLEX41-HM-MCS-ampiA#2의 복제의 pUC 기원을 교화하기 위하여 벡터는 PvuI 및 BspHI를 사용하여 절단되었다. 백본 단편은 CIP되고 정제되었다. 새로운 삽입체 단편은 발현 카세트의 일부로서 복제의 R6K 기원 및 변형된 SV40 poly(A) 서열을 포함한다. SV40 poly(A) 주위의 불필요한 박테리아 또는 바이러스 백본 서열은 삭제되어 왔다(하기의 표 2를 참조). 삽입체 단편은 진아트(GeneArt)로부터 주문되었고; 효소 PvuI 및 BspHI(백본을 위하여 사용된 바와 같이 동일함)을 사용하여 진아트 클로닝 벡터 밖으로 잘렸고, 정제되었고 백본 단편으로 클론되었다. 미니프렙을 준비하였고 서열 분석에 의해 확인되었다. 클론 pGLEX41-MCS-R6K-ampiA#1은 정확한 서열을 가졌다.

[0235] 표 2: 상이한 벡터에서 CpG 함량

표 2

	발현 벡터에서 CpG 함량		
	pGLEX41	코돈 최적화된 벡터: "A"	CpG 감소된 벡터 "B"
엠피실린(Ampicillin) 저항	49	43	19

퓨로마이신(Puromycin) 저항	93	36	1
제네티신(Geneticin) 저항	74	51	0
복제의 기원 (Origin of replication)	45	9	9
합계:	261	139	29

[0237] 중간 벡터 pGLEX41-MCS-R6K-эмпиB의 클로닝

[0238] 벡터 pGLEX41-MCS-R6K-ampiA#1은 제한효소 BspHI를 사용하여 열렸고 엠피실린 저항을 방출하기 위하여 CIP되었다. 새로운 삽입체 단편은 *E. coli*에서 발현을 위하여 엠피실린 저항 코돈 최적화를 포함하나, 선택적 코돈 사용에 의해 삭제될 수 있는 모든 CpG 서열은 대체될 수 있다(상기 표 2를 참조). 상기 단편은 진아트에서 주문되었다. 삽입체 단편을 방출하기 위하여, 진아트 클로닝 벡터 #1016138는 BspHI를 사용하여 절단되었다. 젤 전기이동에 의해 삽입체 및 백본 단편 모두의 정제 후에, 이들은 결찰되었고 PIR1 박테리아로 형질전환되었다. 미니프렙은 시퀀싱을 위해 직접 보내졌다. pGLEX41-R6K-MCS-эмпиB#1은 정확한 서열을 가졌고 추가 클로닝 단계에 사용되었다.

[0239] pGLEX41-derived 발현 벡터에서 리포터 구조체의 클로닝

[0240] 발현 벡터 pGLEX41-MCS-R6K-ampiA 및 pGLEX41-MCS-R6K-эмпиB에서 리포터 구조체 REP를 클론하기 위하여, 벡터를 제한효소 NheI 및 ClaI을 사용하여 잘랐다. 발현 벡터 pGLEX41-HM-MCS는 제한효소 NheI 및 BstBI를 사용하여(65°C에서) 오픈하였다. 모든 벡터 백본은 분해 후에 CIP로 처리되었고 백본은 젤 전기이동에 의해 정제되었다. 백본은 리포터 구조체 REP를 위하여 코딩하는 NheI/BstBI 단편으로 결찰되었다(BstBI은 ClaI으로 호환됨). 결찰 생성물은 PIR1 또는 TOP 10 캄피던트 박테리아로 형질전환되었고 엠피실린 LB-아가 플레이트 상에 플레이트되었다. 콜로니는 증폭되었고 플라스미드는 미니프렙에 의해 분리되었다. 양성 클론은 미니프렙의 제한 분해에 의해 식별될 수 있고 후속의 서열은 Fasteris SA에 의해 확인되었다.

[0241] pGLEX41 유래된 발현 벡터에서 인접 GAPDH 서열의 추가

[0242] 상기 단락의 모든 제한 분해는 80 μ l 최종 부피에서 수행되고 37°C에서 밤사이 배양되었다.

[0243] 5' GAPDH 서열 (서열번호: 7)은 제한효소 NruI을 사용하여 pCR-blunt-5' GAPDH로부터 절단되고 NruI을 사용하여 직선화된 발현 벡터 pGLEX41-R6K-ampiA-[REP] 및 pGLEX41-R6K-эмпиB-[REP]에서 결찰되었고 재순환을 피하기 위하여 CIP 처리되었다. PIR1 콜로니의 증폭 후에(결찰 생성물의 변형에 의해 얻어짐), 미니프렙을 제한 분해에 의해 분석하였다. 클론 pGLEX41-R6K-ampiA-5'GAPDH-[REP] #2 및 pGLEX41-R6K-эмпиB-5' GAPDH-[REP] #1은 제한 분석에서 예상된 사이즈의 밴드를 나타냈고 후속적으로 시퀀싱에 의해 확인되었고 추가 클로닝 단계에 사용되었다. 상기 새로운 벡터는 ScaI을 오픈되었고 CIP로 처리하였다. 3' GAPDH 단편(서열번호: 8)은 동일한 효소를 사용하여 pCR-Blunt-3'GAPDH로부터 절단되고 pGLEX41-R6K-ampiA-GAPDH-[REP] 및 pGLEX41-R6K-эмпиB-GAPDH-[REP] 발현 벡터를 생성하기 위하여 2개의 백본으로 결찰되었다.

[0244] 클론 pGLEX41-R6K-ampiA-GAPDH-[REP] #2 및 pGLEX41-R6K-эмпиB-GAPDH-[REP] #8은 조절 분해는 제한 분석에서 예상되는 사이즈의 밴드를 나타냈다. 정확한 방향으로 3' GAPDH 단편의 삽입은 후속적으로 시퀀싱에 의해 확인되었다(Fasteris).

[0245] II.4. 저항 벡터의 클로닝

[0246] 저항 벡터의 클로닝을 위한 시작점은 벡터 pGLEX-MCS-R6K-ampiA#1이다. 저항 유전자의 발현에 있어서 약한 프

로모터는 충분하고, mCMV 프로모터는 SV40 프로모터에 의해 대체될 수 있다. 저항 유전자를 위해 코딩하는 유전자는 진아트 SA(Regensburg, Germany)로부터 주문되었고 선별 코돈 사용에 의해(퓨로마이신: 퓨로B 및 네오마이신: neoB) 차이니즈 햄스터(퓨로마이신: 퓨로A 및 네오마이신: neoA)에서 최적화되거나 CpG 함량에서 감소된다.

[0247] **pGLEX-R6K-ampiA-puroA/puroB의 클로닝:**

발현 카세트에서 퓨로마이신 저항을 클론하기 위하여, 벡터 pGLEX41-MCS-R6K-ampiA#1는 제한효소 NruI 및 XbaI을 사용하여 오픈되었고 이어서 CIP 처리되었다. 삽입체 단편은 진아트로부터 주문되었고 진아트 클로닝 벡터 #1013239에서 삽입체로서 제공되었다. 이는 퓨로마이신 저항(CH0 세포의 코돈 사용을 위하여)을 위한 SV40 프로모터 및 코돈 최적화 유전자를 포함한다. 삽입체는 효소 NruI 및 XbaI(백본을 위해 사용된 바와 같이 동일함)을 사용하여 진아트 클로닝 벡터 밖으로 잘랐고, 정제되었고 백본 단편으로 클론되었다. 미니프렙을 준비하고 제한 분해로 분석하였다. 클론 pGLEX-MCS-R6K-ampiA-퓨로A#1은 정확한 프로파일을 나타냈고 시퀀싱에 의해 확인할 수 있었다.

상기 벡터는 퓨로마이신 저항 유전자를 위한 코딩 영역의 교환에 의해 벡터 pGLEX-MCS-R6K-ampiA-퓨로B의 클로닝을 위해 사용되었고 반면에 SV40 프로모터를 남겼다. 새로운 삽입체 단편은 선택적 코돈 사용에 의해 삭제될 수 있는 모든 CpG 서열이 대체될 수 있는 퓨로마이신 유전자의 코돈-최적화된 버전을 포함한다. 상기 단편은 진아트에 의해 주문되고 삽입체 단편을 방출하기 위하여 클로닝 벡터 # 1016139에 운반되었고, 진아트 벡터는 제한효소 XbaI 및 NotI을 사용하여 절단되었다. 삽입체 단편은 젤 전기이동에 의해 정제되었고 XbaI 및 NotI을 사용하여 제한 분해에 의해 퓨로마이신 오픈 리딩 프레임으로 방출 후, CIP 처리에 의해 pGLEX-MCS-R6K-ampiA-퓨로A의 백본으로 클론되었다. 얻어진 벡터 pGLEX-MCS-R6K-ampiA-퓨로B#1은 서열 분석에 의해 직접 확인되었다.

[0250] **벡터 pGLEX-R6K-ampiA-NeoA 및 pGLEX-R6K-ampiA-NeoB의 클로닝**

발현 카세트에서 네오마이신 저항을 클론하기 위하여, 벡터 pGLEX-R6K-puroA#1를 제한효소 XbaI 및 NotI을 사용하여 오픈되었고, 이어서 CIP처리 되었다. 삽입체 단편은 진아트로부터 주문되었고 진아트 클로닝 벡터 #1013242(neoA) 및 #1026894(neoB)에서 삽입체로서 제공되었다. 이들은 각각 CHO 세포에서 코돈 사용에 대한 네오마이신 저항을 위한 코돈 최적화 유전자 및 CpG가 감소된 네오마이신 저항을 포함한다. 삽입체는 효소 XbaI 및 NotI(백본을 위해 사용된 바와 같이 동일함)을 사용하여 진아트 클로닝 벡터 밖으로 자르고, 정제되고, 백본 단편으로 클론되었다. 미니프렙을 제조하고 클론을 시퀀싱으로 확인하였다.

[0252] **벡터 pGLEX-R6K-ampiB-NeoB 및 pGLEX41-R6K-ampiB-puroB의 클로닝:**

벡터 pGLEX41-R6K-puroB#1은 제한효소 BspHI을 사용하여 오픈되었고 후속적으로 CIP되었다. 삽입체 단편은 *E. coli*에서 발현을 위한 코돈 최적화된 엠피실린 저항 유전자를 포함하고, 반면에 선택적 코돈에 의해 삭제될 수 있는 모든 CpG 서열은 대체되어 왔다. 상기 단편은 진아트에서 주문되었고 클로닝 벡터 #1016138에서 도달되었다. 삽입체 단편을 방출하기 위하여 진아트 클로닝 벡터는 BspHI을 사용하여 절단되었다. 젤 전기이동에 의해 삽입체 및 백본 단편 모두의 정제 후에, 이들은 결찰되었고 PIR1 박테리아로 형질전환되었다. 미니프렙은 시퀀싱을 위해 직접 보내졌고 확인할 수 있었다(pGLEX41-ampiB-R6K-puroB#1).

벡터 pGLEX-R6K-neoB-엠피B로 유도하는 클로닝은 백본 단편을 제조하기 위하여 제한효소 BspHI을 사용하여 pGLEX-R6K-neoB-ampiA을 오픈하여 행하여졌다. 동일한 제한효소 조합을 사용하여 pGLEX-R6K-ampiB-hygroB의 분해는 ampib를 위하여 삽입체 단편 코딩을 수율하였다. 상기 ampib 삽입체는 pGLEX-R6K-neoB-ampiA 백본으로 클론되었다.

[0255] **II.5 저항 벡터로 인간 GAPDH 유전자의 서열 업스트림 및 다운스트림의 첨가**

[0256] 5' GAPDH 삽입체(3164 bps)를 얻기 위하여 벡터 pCR-blunt-5' GAPDH를 NruI으로 절단하였다. 저항 유전자를 코딩하는 벡터는 NruI로 절단되었고, 백본 단편을 제조하기 위하여 후속적으로 CIP 처리되었다(Calf intestinal phosphatase, NEB, Ipswich, MA). 4개의 상이한 백본 단편(pGLEX-R6K-neoA-ampiA, pGLEX-R6K-neoB-ampiB, pGLEX-R6K-puroA-ampiA 및 pGLEX-puroB-ampiB)은 3164 bps 5' GAPDH 삽입체로 결찰되었고 PIR1 컴피컨트 박테리아로 형질전환되었다. ApaI을 사용하여 미니프렙의 제한분해는 클론 pGLEX-R6K-neoB-ampiB-5'GAPDH #5, pGLEX-R6K-neoA-ampiA-5'GAPDH #6, pGLEX-R6K-puroA-ampiA-5'GAPDH #16 및 pGLEX-puroB-ampiB-5'GAPDH #5의 식별을 가능하게 하였다.

[0257] 상기 중간 벡터를 제한효소 ScaI로 자르고 결찰을 위한 백본을 제조하기 위하여 CIP 처리하였다. 두 번째 삽입체 단편을 운반하는 벡터, pCR-Blunt-3'GAPDH를 삽입체 단편 (3224 bps) GAPDH 다운스트림 인접 영역을 방출하기 위하여 ScaI을 사용하여 잘랐다. 4 개의 상이한 백본 분자는 정제된 3224 bps 삽입체 단편으로 결찰되고 PIR1 컴피턴트 세포로 형질전환되었다. 미니프렙은 제한 분해에 의해 분석되었다. 예상되는 사이즈의 제한 단편을 나타내는 클론은 pGLEX-R6K-neoB-ampiB-GAPDH #8, pGLEX-R6K-neoA-ampiA-GAPDH #1, pGLEX-R6K-puroA-ampiA-GAPDH #1 및 pGLEX-puroB-ampiB-GAPDH #4이었다. 후속적으로 클론을 시퀀싱 분석에 의해 확인하였다 (Fasteris, geneva, Switzerland).

II.1.5. 형질전환을 위한 플라스미드 클론의 미디제조

[0259] 충분한 적정량의 플라스미드를 가지기 위하여, 미니프렙은 맥케레이나겔 키트(NucléoBond Xtra Midi; Macherey Nagel, Oensingen, Switzerland)를 사용하여 제조되었다. 제한 분해에 의해 확인하고 시퀀싱한 후에, 플라스미드를 직선화하고 CHO-S 세포에서 형질전환을 위해 사용되었다. 표 3은 미디제조에서 얻어진 플라스미드 DNA 일괄의 농도를 요약하고 서열을 각각의 플라스미드의 식별 및 서열 정보를 확인하는 Fasteris SA로부터 파일한다. 모든 미니프렙은 형질전환에 사용되기 전에 시퀀싱에 의해 확인되었다.

[0260] 표 3: 클론된 플라스미드의 요약. DNA 미디제조의 농도 및 직선화된 미디제조(해당 효소와 같이). GSC 숫자는 각각의 플라스미드에 대해 코드하고 관련있는 서열파일을 식별하는 것을 허가한다.

표 3

플라스미드	미디제조의 농도($\mu\text{g}/\text{ml}$)	직선화를 위한 효소	직선화된 플라스미드의 농도($\mu\text{g}/\text{ml}$)	글렌마크(Glenmark) 플라스미드 코드
pGLEX41-R6K-ampiA-[REP]-GAPDH	1538	EcoRV	1019	GSC 2774
pGLEX41-R6K-ampiB-[REP]-GAPDH	1243	EcoRV	1233	GSC 2775
pGLEX-R6K-ampiA-neoA- GAPDH	890	AseI	766	GSC 2776
pGLEX-R6K-ampiB-neoB- GAPDH	594	AseI	979	GSC 2777
pGLEX-R6K-ampiA-puroA- GAPDH	917	AseI	859	GSC 2778
pGLEX-ampiB-퓨로B- GAPDH	869	AseI	1049	GSC 2779
pGLEX41-[REP]	2119	BspHI	868	GSC 2239
pGLEX41-R6K-ampiA-[REP]	865	BspHI	779	GSC 2240
pGLEX41-R6K-ampiB-[REP]	1751	BspHI	806	GSC 2249
pGLEX-R6K-ampiA-neoA	890	BspHI	764	GSC 2214
pGLEX-R6K-ampiB-neoB	767	BspHI	654	GSC 2244
pGLEX-R6K-ampiA-puroA	708	BspHI	659	GSC 2220

pGLEX-R6K-amp iB-puroB	574	BspHI	746	GSC 2213
------------------------	-----	-------	-----	----------

[0262] 실시예 2: 발현 벡터를 갖는 세포의 형질전환 :

[0263] 1. 재료 및 방법

[0264] CHO-S 세포 및 HEK293 세포

[0265] 포유동물 세포는 재조합 단백질의 정확한 폴딩, 접합 및 전사후 변형을 할 수 있기 때문에 단백질을 발현시키는 바람직한 숙주이다. CHO 세포주는 잘 특징되어있고 대부분의 인간 발병의 바이러스를 위한 숙주로서 제공하지 않기 때문에, 안정된 치료 단백질생성을 위한 상대적으로 안전한 숙주를 만드는데 사용된다. 차이니즈 햄스터 Ovary 세포 (CHO-S, Invitrogen, Carlsbad, CA, USA)는 4 mM L-glutamine (Applichem, Germany)으로 보충되는 PowerCHO-2 CD 배지 (Lonza, Verviers, Belgium)에서 혼탁으로 배양되고 37°C, 5% CO₂ 및 80% 습도에서 진탕 배양기(2.5 cm의 원형 스트로크를 갖는 200 rpm)에서 배양된다. HEK293 세포는 형질전환되기 쉽고 낮은 그램량 까지 재조합 단백질의 빠른 생성을 하게 하기 때문에 사용된다. 사용되는 세포는 HEK293-EBNA 세포 (ATCC, Manassas, VA)이고 Ex-cell 293 배지에서 혼탁으로 일상적으로 배양된다(Sigma-Aldrich, St. Louis, MI).

[0266] CHO-S 및 HEK293 EBNA 세포의 계대배양은 신선한 배지에 0.5×10^6 살아있는 세포/ml의 분주 밀도를 사용하여 매 일 3-4일 일상적으로 수행하였다. 세포를 가스교환할 수 있도록 침투할 수 있는 필터를 포함하는 50 ml 생물반응장치(bioreactor) 투브에서 10 ml의 배지를 사용하여 배양하였다(Tubespin Bioreactor 50; TPP, Trasadingen, Switzerland). 세포 생존율 및 농도는 트립판 블루 세포 배제 방법을 사용하여 카운티스 자동화된 세포 카운터(Countess automated cell counter)로 결정되었다. 세포 농도는 CHO-S 세포에 대하여 PCV 투브 (TPP, Trasadingen, Switzerland)를 사용하여 농축 세포 부피(PCV) 방법의 결정으로 확인되었다.

[0267] 농축 세포 부피 (PCV)

[0268] PCV 방법은 5000 rpm에서 1 분 동안 미니-PCV 투브 (PCV 팩된 세포 부피 투브; TPP, Trasadingen, Switzerland)에서 배양 액체의 특정 부피의 원심분리 상에서 기초되었다. 원심분리 동안에, 세포는 투브의 베이스에서 눈금 측정기로 펠렛되었다. 농축 세포 부피의 퍼센트는 원심분리된 세포 배양액에 관계하여 펠렛의 부피를 측정하여 결정된다. 예를 들면, 1% PCV는 세포 펠렛의 10 μl는 1 ml의 배양액에 존재하는 것을 나타낸다.

[0269] 세포의 수를 세는 일상적인 세포를 위하여, 각각의 시료의 200 μl를 PCV 투브에 파이펫하고 해당 펠렛의 부피 (μl에서)는 자(ruler)로 읽었다("easy read" 측정 기구; TPP, Trasadingen, Switzerland). 상기 부피는 1 ml에 대해 값을 갖도록 5를 곱한 후 살아있는 세포의 농도 측정을 연도록 세포 특정 연관 요인을 사용하여 곱하였다(세포의 백만/ml에서)

[0270] "자동화" 세포 측정

[0271] 세포 농도 및 생존율은 트립판 블루의 동일한 양으로 시료를 혼합하여 카운티스[®] 자동화된 세포 카운터(Countess[®] Automated cell Counter)로 결정하였다. 용액은 기구에 의해 읽혀지기 전에 카운티스[®] 챔버 슬라이드로 파이펫되었다. 상기 기구는, 교정 후에, 뉴바우어 챔버의 자동화 읽기를 허가하였고 죽고 살아있는 세포의 세포 생존율 및 농도를 결정한다.

[0272]

[0273] 유동세포측정 분석

[0274] 유동세포측정은 각각의 세포의 다수 파라미터의 분석에 대한 기술이다. 상기 기술은 서로 표현형적으로 다른

세포, 예를 들면, 살아있는 세포로부터 죽은(세포의 크기 및 입상에 따른)세포의 정량적 및 질적 분석을 허가한다. 이는 또한 GFP와 같은 관심 단백질을 발현하는 세포의 정량을 허가한다. 세포는 300 μl 의 시료를 멸균 파이펫팅에 의해 배양으로부터 수득되고 488 nm에서 공기-냉각된 아르곤 레이저 방출이 장착된 형광-연관된 세포 구분(FACS) 칼리브라 유동세포측정기(Becton, Dickinson 및 Company, Franklin Lakes, NJ, USA)로 분석하였다. 분석체는 셀퀘스트 소프트웨어로 만들어졌다. GFP 방출은 530/30-nm 밴드 패스 필터를 사용하여, FL-1로 검출되었다.

[0275] 첫 번째 문에서, 죽은 세포뿐만 아니라 세포 부스러기(debris)는 직선 규모 상의 SSC/FSC 점선에서 분석으로부터 배제되었다. 이후, 살아있는 세포의 GFP 형광은 대수(logarithmic) 규모상에서 히스토그램에서 디스플레이 되었다. 형광 분포의 정중 값(median value)은 분석된 세포 집단의 GFP 발현 수준을 평가하는데 사용되었다.

[0276] IgG 정량적 방법: OCTET QK

[0277] 옥테트 QK 시스템(ForteBio, Menlo Park, CA, USA)은 항체, 단백질, 결합, DNA 및 다른 항체의 라벨-프리 정량을 수행하고 생분자 결합 상호작용의 운동특성(kinetic characterization)을 제공한다. 시료의 결합 비율(binding rate) 및 측정된 IgG1 농도 ($\mu\text{g}/\text{mL}$) 사이에 연관성은 표준곡선을 갖는 IgG 적정량의 정량화를 허가한다.

[0278] 세포 시료는 300 g에서 5 분간 원심분리한다. 이후 웰 당 항체 농도를 얻기 위하여 단백질 A 바이오센서 (Protein A DIP 및 READTM Biosensor, Forte Bio, USA)를 사용하여 옥테트(Octet)로 분석되기 전에 상층액을 96 웰 플레이트에서 옥테트버퍼로 희석(IgG1 항체에 대하여 1/5)하였다.

[0279] JetPEI을 사용하여 일시적 형질전환

[0280] CHO-S 및 HEK293 EBNA 세포의 일시적 및 안정적 형질전환은 폴리에틸렌이민(PEI; JetPEI, Polyplus-형질전환, Illkirch, France)을 사용하여 수행되었다. PEI는 DNA와 같은 음이온 전하된 분자를 갖도록 복합체를 형성할 수 있는 양이온 중합체(cationic polymer)이다. 양이온 전하된 DNA-PEI 복합체는 음이온으로 전하된 세포표면에 결합하고 엔도시토시스(endocytosis)에 의해 내부화(internalized)된다. 이는 융해(lysis)에 의해 핵으로 방출되는 곳으로부터 라이이소좀(lysosome) 구획에 도달한다. DNA-PEI 복합체를 갖는 높은 형질전환 효율은 라이소좀 분해로부터 DNA를 보호하기 위하여 PEI의 능력때문이다. 상기 세포는 제조사에 의해 제공된 매뉴얼에 따라 형질전환된다.

[0281] 모든 플라스미드는 안정 형질전환 (100 μl Tris-EDTA, pH 7.5에 재현탁된 100 μg 의 DNA)전에 직선화되었다. 일시적 형질전환을 위하여, 원형 플라스미드는 미디제조 DNA로부터 직접 사용된다. 상기 연구에서 일시적 형질전환은 50 mL 생물반응기 튜브에서 보존하고 항생제를 첨가하지 않았다.

[0282]

[0283] IgG1 및 GFP를 발현하는 안정적 CHO-S 클론은 하나의 발현 벡터 및 두 개의 저항 벡터를 공동-형질전환에 위하여 얻어질 수 있다(각각 퓨로마이신 또는 네오마이신 저항에 대하여 코딩).

[0284] 안정적 풀 및 미니 풀(minipools)의 선별

[0285] 형질전환 효율은 세포 내 GFP 발현을 분석함으로써 유동세포측정 (BD FACS Calibur cytometer, #1293)에 의해 형질전환 후 24시간에 결정되었다. 만약 GFP 양성 세포의 퍼센트가 20 %보다 높으면, 형질전환된 세포는 선택 배지로 희석되고 96 웰 플레이트로(분리된 안정적 미니 풀을 생성하기 위해 제한 희석을 위하여) 또는 T-플라스크(안정적 풀을 생성하기 위하여)에 분포되었다. 사용된 선택 배지는 제네티신 및 퓨로마이신의 상이한 농도로 보충되는, PowerCHO-2, 4 mM 글루타민이었다.

[0286] 형질전환 후 7일째에, 선별 엄중성(selection stringency)은 세포로 선별배지를 첨가하여 새롭게 하였다. 96 웰 플레이트에 있는 콜로니가 융합되자마자, 플레이트는 형광 리더를 사용하여 측정되었다.

[0287] T-플라스크에 있는 풀은 항생제가 없는 PowerCHO-2, 4 mM L-글루타민을 사용하여 투브스핀(tubespin) 규모로 확장되었다. 이의 생존율 및 농도는 카운티스 자동화된 세포 카운터로 평가되었다(Invitrogen, Carlsbad, CA, USA). 세포농도가 상기를 허락하자마자, 씨드 연쇄(seed train)는 50 ml 생물반응기 투브(5% CO₂, 37°C 및 80% 습도에서 진탕기(200 rpm)에서 배양)에서 10 ml 배지의 0.5×10^6 세포/ml의 농도에서 세포 분주에 의해 각각의 풀에 대하여 시작되었다. 각각의 씨드 연쇄(seed train)는 성장 배지에서 0.5×10^6 세포/ml에서 세포를 분주하여 주 당 2회 계대하였다(세포농도는 PCV 분석에 의해 결정되었다). 씨드 연쇄는 모든 생성물의 실행(productions runs)의 접종원(inoculum)에 대하여 사용되었다(일괄).

[0288] 다음 4-5 주 동안에 생성물의 실행은 중복(duplicates)으로 주 당 한번 분주 되었다. 풀의 안정성은 클론의 집단에 대하여 상기에서 언급된 바와 같이 FACS 및 IgG 발현에 의해 평가되었다.

생성물 실행(Production runs) (비연속 발효(batch fermentation))

[0289] 세포 풀의 일괄 실행은 접종을 위한 씨드 연쇄(seed train)를 사용하여 0.5×10^6 세포/ml의 농도에서 분주되었고 세포는 공급배지(Feed media)에서 7일 동안 배양되었다. 4 및 8일째에, 200 μ l의 세포를 300 g에서 5 분 동안 원심분리하였고 상층액을 옥텟(Octet)를 사용하여 측정된 IgG에 대하여 분석하였다. 추가적으로, 각각 일괄에 대한 GFP 발현을 FACS로 분석하였다.

2. 결과

2.1 CHO 세포에서 일시적 발현:

[0290] 상기 연구와 비교하여 벡터는 자체의 백본에서 주요하게 다르다. 전체 발현 카세트(프로모터, 첫 번째 인트론, 발현 구조체, 폴리(A))는 모든 벡터에 대하여 정확히 동일하다. 상기 벡터는 실시에 1에서 설명된 바와 같이 벡터 pGLEX41로부터 유래되었다. 하나의 벡터에서, 엠피실린 저항 유전자는 *E. coli*에서 발현을 위해 코돈 최적화되었고 박테리아 백본은 최소로 감소되었다: pGLEX41-R6K-ampiA-[REP](짧은 A에서). 두 번째 벡터에서, 엠피실린 저항 유전자는 *E. coli*에서 발현을 위해 코돈 최적화되었으나, 모든 CpG 서열은 선택적 코돈 사용에 의해(가능할 때), 회피되었다: 상기 벡터는 pGLEX41-R6K-AmpiB-[REP] (짧은 B에서)로 불린다. 세 번째 변형은 벡터 pGLEX41-R6K-ampiA-[REP]-GAPDH(짧은 GAPDH_A에서) 및 pGLEX41-R6K-AmpiB-[REP]-GAPDH(짧은 GAPDH_B에서)를 주는 벡터 A 및 B의 발현 카세트의 업스트림 및 다운스트림이 클론되는 GAPDH 인접 서열의 용도를 포함한다.

[0291] 상이한 플라스미드 백본의 맥락에서 발현되는 리포터 단백질의 발현 수준을 비교하기 위하여 CHO-S 세포(Invitrogen)의 일시적 형질전환이 행하여졌다. 상기 형질전환(중복)을 10 ml의 최종 배지 부피를 사용하여 50 ml 생물 반응기 투브(TPP, Trasadingen, Switzerland)에서 수행하였고 옥텟에 의해 형질전환 후 5일째에 분석하였다(도 2).

[0292] 정확한 백본이 있는 모든 벡터(A 및 B)는 대조군 벡터 pGLEX41보다 약간 높은 발현 수준을 나타낸다. 벡터 A 및 B 사이에는 단지 최소 차이만 있다. 이는 예상되는데, 왜내하면 백본에서 단지 차이가 일시적 발현 상에서 영향을 갖지 않아야만 하는 엠피실린 저항이기 때문이다.

[0293] 가장 현저한 관찰은 발현 상에서 GAPDH 서열의 긍정적 효과이다. 2-배 높은 발현 수준은 GAPDH 서열이 없는 하

나에 비교하여 GAPDH 인접 서열에 정착하는 플라스미드와 같이 얻어진다. 이는 A 및 B 구조체 모두에 대해 사실이다. pGLEX41 벡터와 비교하여, 3-배 높은 발현이 관찰되었다. 상기는 만약 플라스미드의 사이즈가 고려된다면 더욱 더 놀랍다. 벡터 A (7048 bps)는 벡터 GAPDH-A (13436 bps)와 비교하여 거의 절반 사이즈이다. 따라서, 일시적 형질전환 과정 동안에 운반된 DNA의 양은 모든 플라스미드에 대하여 동일하고, GAPDH-A의 몰라 양의 단지 절반만이 핵으로 운반된다고 추정된다.

[0297] 2.2 HEK293 세포에서 일시적 발현

[0298] 상이한 플라스미드 백본의 맥락에서 발현된 리포터 단백질의 발현 수준과 비교하기 위하여 HEK293 EBNA 세포의 일시적 형질전환을 수행하였다. 형질전환 (중복)을 10 ml의 최종 배지 부피를 사용하여 50 ml 생물반응기 튜브 (TPP, Trasadingen, Switzerland)에서 수행하였고 옥텟에 의해 형질전환 후 10일째에 분석하였다 (도 3).

[0299] 도 3에 나타낸 결과는 HEK293 EBNA 세포에서 GAPDH 인접 영역을 사용하여 얻어질 수 있는 발현에서 현저한 증가를 나타냈다. GAPDH-B 벡터은 발현에서 3배 증가를 나타내는 반면, GAPDH-A 벡터는 5 배의 발현에서 더욱 높은 증가를 나타낸다. 상기 벡터는 or iP 요소를 포함하지 않고 따라서 더욱 높은 적정량에 대한 잠재성을 가질 수 있다.

[0300]

[0301] 2.3 안정적 CHO 세포주에서 발현

[0302] 안정적 형질전환된 세포의 구축

[0303] 안정적 집단은 저항 유전자에 대하여 코딩하는 발현 벡터 및 벡터 코딩에 의하여, 이어서 항생제에 의해 매개되는 선별 압력에 의해 생성된다. 선별압력은 형질전환 후 14일째에 제거되었다. 상기 단계는 상이한 구조체 및 발현의 안정성의 리포터 단백질 (IgG1 항체 및 GFP)의 발현 수준과 비교하기 위하여 생성물 실행에서 정기적 간격으로 배양되는 안정적 미니 풀 및 안정적 풀의 생성을 허가하였다

[0304] 세포 풀로 수행된 생성물 실행 상에서 리포터 단백질 발현 연구

[0305] 풀은 안정적 형질전환에 의해 생성되었다. 선별 절차 동안 (형질전환 후 첫 번째 14일째)에 풀은 FACS에 의해 분석되었다. 배양의 생존율과 같이 GFP 양성세포 분획의 증가는 시간이 지남에 따라 관찰된다. 항생제에 의해 매개되는 선별압력은 14일 후에 풀로부터 제거되었다. 상기 접근법을 사용하여 "B" 플라스미드로 형질전환된 세포 풀은 얻을 수 없었다. 생성된 풀의 발현 수준은 세포가 50 ml 생물반응기 튜브에서 배양될 수 있자마자 분석되었다. 일괄은 중복으로 행하여졌다. 세포는 GFP 발현을 위해 FACS에 의해 분석되었고 상층액에서 IgG의 축적은 발현 8일째 후에 옥텟에 의해 분석되었다.

[0306] 비례적 관계를 IgG 적정량 및 풀의 GFP 발현 사이에서 관찰할 수 있다. 따라서, 단지 IgG 데이터를 도 4에 나타냈다. GAPDH 서열을 포함하는 벡터를 갖는 형질전환된 모든 풀은 벡터 pGLEX41 또는 GAPDH 서열 없이 동일한 벡터 (A 및 A-GAPDH 사이의 2.8 요소, B 풀이 생존하지 않았기 때문에 B 및 B-GAPDH 사이에서 결론을 나타낼 수 없음)와 비교하여 높은 발현을 나타낸다. A-GAPDH 및 B-GAPDH로 수행된 형질전환은 pGLEX41 형질전환 (일괄-2에 대하여)보다 높은 발현의 IgG (각각 2.7 및 3.5 폴드 이상)를 유도하였다. 따라서, 풀에서, GAPDH 인접 서열은 단백질 생성에 대하여 호의적으로 보인다.

[0307] 최종적으로, B-GAPDH 벡터로 수행된 형질전환은 A-GAPDH로 수행된 형질전환보다 높은 발현의 IgG를 유도하였다 (1.25 요소). 따라서, 저항 유전자에서 CpG 감소는 역시 단백질의 안정적 생성을 위해 호의적으로 보인다.

[0308] of 1.25).

[0309] 클론의 집단 상에서 발현 수준 연구

[0310] 세포를 형질전환하고 클론 또는 올리고클론 집단을 얻기 위하여 선택 배지에서 96 웰 플레이트에 분포하였다. 7일 후에 선별압력을 세포로 선택 배지 첨가에 의해 새롭게하였다. GFP의 발현을 ELISA-플레이트 리더를 사용하여 형질전환 후에 14일째에 평가하였다. 상기 결과를 도 5에 나타냈다.

[0311] 세포 내 풀에서 얻어진 결과를 확인함으로써, GAPDH 인접 서열을 포함하는 벡터로 형질전환된 세포는 GAPDH 업- 및 다운스트림 서열(1.7부터 2 폴드까지 요소) 없는 동일한 백본 또는 대조군으로서 사용된 다른 벡터(pGLEX41: 2.5 폴드)보다 현저히 더 많은 GFP를 발현한다(도 5). 추가적으로, 단지 코돈 최적화된 해당 벡터 (A)보다 CpG 감소된 저항 서열을 포함하는 벡터(B)를 갖는 집단은 높은 발현을 유도하였다(A 및 B사이에서 1.5 폴드; B 및 B-GAPDH 사이에서 1.2 폴드).

[0312] 발현연구로부터 몇 가지 결론을 그릴 수 있다. 첫째, GAPDH 업- 및 다운스트림 서열은 벤치마크(pGLEX41)로서 사용된 표준 벡터보다 높은 발현을 허가하였다. 또한 세포가 GAPDH 서열이 없는 동일한 백본으로 형질전환될 때 낮은 발현 수준이 얻어짐으로써 발현 상에서의 유익한 효과는 삽입된 GAPDH 인접 서열과 관련됨을 확인하였다. 추가적으로, 플라스미드의 발현 및 선별에서 CpG 수의 감소는 발현에 대해 역시 약간 호의적으로 보인다.

실시예 3: 새롭게 디자인된 벡터로 형질전환된 CHO-S GMP 세포의 일시적 발현 수준

[0314] GAPDH 프로모터의 5' 영역이 포볼 에스테르 응답 요소(phorbol ester response element) 뿐만 아니라 잠재적 인슐린을 정착한다는 것은 문현에서 설명되어 있다(Alexander-Bridges *et al.*, (1992) *Advan Enzyme Regul*, 32: 149-159). 포볼 에스테르 응답 요소(-1040 -1010 bps)는 GAPDH 프로모터 (-488 - +20)로서 항상 언급되는 것의 업스트림에 위치해 있다. 안정적 H35 간암 세포주에서 수행된 삭제연구에서, 발명자는 베이스짝 (basepairs) -1200 내지 -488(전사시작 점에 비례하여)의 삭제의 현저한 효과를 나타낼 수 없었다. 따라서, 포볼 에스테르 응답 요소는 GAPDH 프로모터로부터 유도된 발현과 기능적으로 연결될 수 없다. 그럼에도 불구하고 GAPDH 인접 요소를 포함하는 플라스미드를 사용하여 관찰되는 일시적 및 안정적 발현의 증가에서 인슐린 및 PMA(phorbol-12-myristate-13-acetate, the most common phorbol ester)의 기여를 평가하기 위하여 일시적 형질전환 실험을 수행하였다.

[0315] 인슐린이 없는 성장 배지를 얻기 위하여, PowerCHO2를 분말 배지로부터 제조하고 인슐린을 첨가하지 않았다. PMA를 시그마(St. Louis, MO)로부터 구입하였고, PowerCHO2(+/- 인슐린)에서 1.6 μ M의 최종농도(H35 간암 세포주 상에서 알렉산더-브릿지(Alexander-Bridges)에 의하여 사용된 농도에 해당함)로 투여하였다.

[0316] 형질전환을 상기에서 설명한 바와 같이 50 ml 생물반응기(bioreactor) 튜브(Tubespins, TPP, Trasadingen, Switzerland)에서 수행하였다. 옵티MEM(OptiMEM) (Life technologies, Carlsbad, CA)에 의해 제공된 인슐린의 존재를 피하기 위하여, 형질전환 배지를 4 mM Gln 및 25 mM 헤페스로 보충된 RPMI1640 (PAA, Pasching, Austria)로 변경하였다. 형질전환 후, 세포를 12 웰 플레이트에 분포하고 4개의 상이한 배지의 1 ml를 첨가하였다(PowerCHO2, 4mM Gln, +/-인슐린; PowerCHO2, 4mM Gln, 1.6 μ M PMA, +/-인슐린). 다시, 2개의 IRES를 사용하여 IgG1 및 GFP를 발현하는 리포터 구조체를 사용하였다(실시예 2에서 설명됨). 상기 벡터는 형질전환 효율을 검증하는것을 허가하였다. 형질전환된 세포의 퍼센트 및 생존율을 모든 4 개의 상이한 배지 제조에서 유사함을 발견하였다.

[0317] 도 6에 나타낸 바와 같이, 인슐린 고갈 및/또는 PMA 첨가의 현저한 효과를 상기 실험 동안에 관찰할 수 없었다. 유사한 적정량을 발현에 사용된 모든 배지에서 얻었다. 상기는 GAPDH 유전자와 업스트림 인접 서열에 존재하는 잠재적 포볼 에스테르 및 인슐린 응답요소가 일시적 외래유전자(transgene) 발현에 영향을 주지 않는다는 것을 제시한다.

[0318] 실시예 4: 리포터 유전자 발현 상의 효과를 연구하기 위하여 프로모터의 업스트림 및 polyA 자리의 다운스트림 GAPDH 발현 카세트 인접하는 DNA의 단편화 분석

[0319] 인간 GAPDH 자리(locus)는 인간 게놈의 염색체 12 상에 위치해 있다. GAPDH는 효소가 글루코오스의 대사(metabolism)에서 핵심역할인 것처럼, 포유동물 기원의 모든 세포에서 계속적으로 활성되도록 설명된다. 프로모터의 업스트림, GAPDH 유전자는 NCAPD2, 30000 bp 보다 이상으로 스트레치하는 유전자에 의해 옆에 있다. 폴리아데닐레이션 자리의 다운스트림, GAPDH 유전자는 IFF01에 의해 옆에 있다(상세설명을 위해 도 7을 참조).

[0320] GAPDH 및 프로모터 뿐만아니라, 또한 인접 영역은 상이한 종(species) 사이에서 잘 보존된다(표 4를 참조).

[0321] 표 4: 인간, 래트 및 마우스 GAPDH 인접 영역 사이에서 높은 상동(homologies)의 스트레치. 분석을 클론 매니저 9(manager 9)(ScieED, Cary, NC, USA)을 사용하여 행하였다. 숫자화는 각각 업스트림 또는 다운스트림 인접 요소(각각 서열 ID NO: 7 및 서열 ID NO: 8)의 첫 번째 베이스에 비례한다. 정렬에 사용된 서열은 서열 ID No 18의 마우스 베이스 532-3731 (업스트림) 및 8164-11364 (다운스트림) 및 서열 ID No 19의 래트 베이스 719-3918 (업스트림) 및 8495-11058 (다운스트림)에 대하여 사용된다.

표 4

업스트림 영역		다운스트림					
상동(homology)의 서열 [래트]		상동(homology)의 서열 [마우스]		상동(homology)의 서열 [래트]		상동(homology)의 서열 [마우스]	
>80 %	>90 %	>80 %	>90 %	>80 %	>90 %	>80 %	>90 %
161-249	279-331	15-69	278-329	1608-1764	1706-1764	1614-1671	1904-2061
256-338	554-623	159-249	546-626	1894-2067	1912-2061	1888-2072	2927-3071
515-659		273-342				2918-3082	
2296-2349		515-647					
2381-2513		1143-1223					
2736-2818		1957-2009					
		2029-2080					
		2375-2485					
		2730-2821					

[0323] 설치류 및 인간 사이에서 DNA 상동(homology)의 비교는 38%의 최소 DNA 보존을 나타낸다. 프로모터 영역 또는 유전자에 대해 코딩하는 영역에 대한 DNA 외부의 보존된 스트레치의 존재는 DNA 서열을 유지하거나 단지 특정/소수 변경을 허가하는 세포 상에서의 선별 압력일 수 있음을 나타낸다. 본 발명의 특정 경우에, GAPDH 인접 영역은 GAPDH 유전자의 높은 발현 수준을 유지하기 때문에 세포에 대해 중요할 수 있다. 발현 감소를 유도하는 DNA 서열에서의 변경은 이에 대하여 선별될 수 있다.

[0324] 발현에서 증가가 관찰된 업스트림 및 다운스트림 GAPDH 요소의 기여를 평가하기 위하여, 구조체는 단지 업스트림 GAPDH 인접 영역(서열번호: 7), 업스트림 GAPDH 인접 영역 또는 다운스트림 GAPDH 인접 영역(서열번호: 8)의 단편을 포함하도록 만들어져졌다. 리포터 IgG1 탑입 항체는 RES 구조체 (경쇄-IRES-중쇄)에 의해 발현되고, 따라서 다수의 플라스미드의 공동-형질전환을 피한다. GAPDH 업스트림 단편의 단편화에 대한 상세설명을 도 8에 나타냈다. 하기의 업스트림 GAPDH 인접 영역의 단편이 사용되었다: 단편 1 (서열번호: 9), 단편 2 (서열번호: 10), 단편 3 (서열번호: 11), 단편 4 (서열번호: 12), 단편 8 (서열번호: 13), 단편 9 (서열번호: 14), 단편 11 (서열번호: 15), 단편 17 (서열번호: 16).

[0325] 사용된 업스트림 GAPDH 인접 영역(서열번호: 7)은 3개가 자체의 5'에 게놈 DNA에 연결되고 3 개가 자체의 3' 말단에 게놈 DNA에 연결된 NruI 제한 자리의 2 곱하기 3(총 6)의 뉴클레오티드를 포함한다. 사용된 다운스트림

GAPDH 인접 영역(서열번호: 8)은 3개가 자체의 5'에 계놈 DNA에 연결되고 3 개가 자체의 3'말단에 계놈 DNA에 연결된 ScaI 제한 자리의 2 곱하기 3(총 6)의 뉴클레오티드를 포함한다. 각각의 제한 자리의 뉴클레오티드가 없는 업스트림 GAPDH 인접 영역 및 다운스트림 GAPDH 인접 영역은 서열번호: 20(제한효소 자리 없는 업스트림 GAPDH 인접 영역) 및 서열번호: 21(제한효소 자리 없는 다운스트림 GAPDH 인접 영역)에 나타낸다. 사용된 업스트림 GAPDH 인접 영역의 단편은 계놈 DNA(단편 1은 자체의 5' 말단에 NruI 제한 자리(restriction site)의 3 뉴클레오티드를 포함하고; 단편 2는 자체의 3' 말단에 NruI 제한 자리의 3 뉴클레오티드를 포함하고; 단편 3은 자체의 5' 말단에 NruI 제한 자리의 3 뉴클레오티드를 포함하고; 단편 4는 자체의 3' 말단에 NruI 제한 자리의 3 뉴클레오티드를 포함하고; 단편 8은 자체의 3'말단에 NruI 제한 자리의 3 뉴클레오티드를 포함하고; 단편 9는 자체의 5'말단에 NruI 제한 자리의 3 뉴클레오티드 및 자체의 3' 말단에 NruI 제한 자리의 3 뉴클레오티드를 포함하고; 단편 11은 자체의 3'말단에 NruI 제한 자리의 3 뉴클레오티드를 포함한다)에 연결된 자체의 5' 및/또는 자체의 3' 말단에 있는 각각의 제한 자리의 3 뉴클레오티드를 각각 포함한다. 단편 17은 제한 자리의 뉴클레오티드를 포함하지 않는다. 각각의 제한 자리의 뉴클레오티드가 없는 업스트림 GAPDH 인접 영역의 단편은 서열번호: 22 (제한 자리 없는 단편 1), 서열번호: 23 (제한 자리 없는 단편 2) 서열번호: 24 (제한 자리 없는 단편 3), 서열번호: 25 (제한 자리 없는 단편 4), 서열번호: 26 (제한 자리 없는 단편 8), 서열번호: 27 (제한 자리 없는 단편 9), 서열번호: 28 (제한 자리 없는 단편 11)에 나타냈다.

[0326] 발현 상의 업스트림 및 다운스트림 GAPDH 요소의 효과는 상층액에서 배출된 IgG1의 양을 정량화하기 위하여 옥테트(Fortebio, Menlo, CA, USA)를 사용하여 형질전환 후 10일 째에 평가되었다(도 9 참조). pGLE41, 본래 벡터는 pGLE41-ampiA 벡본에 사용된 개선된 새로운 벡터 디자인과 비교하여 낮은 발현 결과(80%)를 준다. 본래 pGLE41 벡본과 비교하여 새로운 디자인은 복제의 상이한 기원(복제의 pUC 기원 대신에 R6K), *E. coli*에서 엠피실린 저항에 필요한 ampiA 유전자의 코돈 최적화 및 박테리아 기원의 불필요한 링커(또는 스페이서)서열의 제거를 포함한다. 벡터 모두는 대략 동일한 사이즈를 가진다.

[0327] 놀랍게도, 업스트림(서열번호: 7) 및 다운스트림 요소(서열번호: 8)을 포함하는 pGLE41-ampiA (발현 결과를 나타내는 도 9에서 pGLE41-업/다운으로 명명)는, 업스트림 및 다운스트림 서열없는 동일한 벡터와 비교하여 높은 발현(1.5 지수)을 준다. 만약 누군가 사이즈에서 차이(업/다운 단편은 대략 6000 bps에 의해 플라스미드의 사이즈를 증가시킴) 및 따라서 형질전환 동안에 운반된 플라스미드 복제에서 차이를 고려하면, 효과는 플라스미드 기본 당(per plasmid basis)에서 더욱 더 중요할 수 있다

[0328] 단지 업스트림 단편(업)을 포함하는 벡터는 본래 발현 구조체 pGLE41-ampiA에 유사한 발현 수준을 나타낸다. 단지 다운스트림 단편(다운)을 포함하는 벡터는 본래 발현 구조체 pGLE41-ampiA과 비교하여 발현에서 현저한 증가(1.2 지수)를 나타내고 있다. 발현에서 추가 증가는 만약 업- 및 다운스트림 단편 모두가 존재한다면, 관찰될 수 있다. 상기는 업스트림 단편의 단편화에 의해 확인된다. 단편 9 및 프로모터 근부(proximal) 단편 8은 pGLE41-ampiA와 비교하여 발현에서 임의 차이를 나타내지 않는다. 단편 1, 11 및 17은 발현에서 증가를 나타낸다. 가장 높은 증가는 단편 4에서 관찰되었다. 프로모터 근부(proximal) 단편 8이 임의 효과를 나타내지 않음이 강조되어야만 한다. 따라서 발현에서 증가는 기존에 게재된 서열에 의해 설명될 수 없다(Alexander-Bridges *et al.*, (1992) *Advan. Enzymol. Regul.*, 32: 149-159), Graven *et al.*, (1999) *Biochimica et Biophysica Acta* 147: 203-218).

[0329] 흥미롭게도, 단편 2 및 3은 발현에서 현저한 감소를 유도한다. 상기는 예상되지 않았고, 특히 상기 단편이 반대 방향(도 9에서 안티센스(AS))으로 클론되는 사실의 관점에서 상기 효과를 초래하지 않는다. 단편 1, 8, 9, 11 및 17에 대하여 센스 또는 안티센스 방향(orientation)(데이터를 나타내지 않음)으로 통합된 단편에 대하여 발현에서 차이가 관찰되지 않았다. 비록 단편 2의 일부이지만, 단편 11은 상기 효과를 나타내지 않는다. 따라서 발현에 해를 끼치는 것처럼 보이는 서열 요소는 단편 11을 얻기 위하여 단편 2에서 삭제된 BstBI-BstBI 단편 상에 적어도 부분적으로 있어야만 한다.

[0330] 추가적으로, 부정적 요소가 BstBI-BstBI 단편 상에 위치한다는 가설은 단편 3 (BstBI-BstBI 단편을 포함하는)

및 단편 1 사이에 관찰된 발현에서 증가에 의해 지지된다.

[0331] 부정적 효과(BstBI-BstBI)를 가지는 단편을 국한시키기에 쉬운 것으로 보이는 반면에, 상기 연구로부터 단편 2 및 3에서 관찰되는 부정적 효과가 어떻게 완전한 업스트림 단편에서 서열 요소 존재에 의해 보상되는지는 덜 분명하다. 상기 부정적 효과는 단편 1 및 단편 4(그러나 단편 1에 대한 발현에서 증가는 단편 4보다 약함)에 의해 관찰된 적은 긍정적 효과에 의해 균형될 수 있다. 그럼에도 불구하고 상기 관찰된 단편 4에 대한 긍정적 효과(1.25 지수)가 부정적 효과(0.4 지수)와 비교하여 덜 중요한 것처럼 보인다. 추가적으로 BstBI-BstBI 단편이 없는 전체 업스트림 영역인 단편 9은 전체 GAPDH 업스트림 인접 영역(nevertheless, 단편 9 포함한다 the EcoRV-BstBI 단편 which is part of 단편 2 및 3 및 might have a 부정적 효과 on 발현)과 비교하여 증가된 발현을 나타내지 않는다(그럼에도 불구하고, 단편 9은 단편 2 및 3의 일부인 EcoRV-BstBI 단편을 포함하고 발현 상에서 정적 효과를 가질 수 있다)

[0332] 관찰된 효과 뒤의 기전에 대하여 단지 추측될 수 있다. 단편 2 및 3으로 관찰된 발현 상의 부정적 효과의 방향의 존성은 단지 하나의 방향의 발현을 유발할 수 있는 둘러싸는 프로모터가 없기 때문에, 비식별된 오픈리딩프레임의 발현(예를 들면, ncRNA의 발현)을 배제한다. 발현이 기저 수준 밑으로 감소된 사실은 긍정적 효과(예를 들면, 인핸서 활성)의 부재 뿐만아니라 부정적 효과에 의존하는 방향의 존재도 나타낸다.

[0333] 요약하면, CHO 세포에서 일시적 발현의 놀라운 증가는 만약 인접 영역, 업스트림 및 다운스트림 영역 모두가 발현 플라스미드에서 존재하면, 관찰된다. 비록 단편 4는 발현상에서 현저한 긍정적 효과를 가지지만, 단일 단편은 관찰되는 발현의 전체 증가에 책임되는 것으로 확인될 수 없다. 발현 벡터 GLEX41-ampiA (up/down) 발현 증가는 up- 및 다운스트림 인접 영역 모두의 효과를 요약하는 것으로 보인다.

[0334] 실시예 5: 차이니즈 햄스터 GAPDH 유전자 및 차이니즈 햄스터 프로모터의 비변역된 게놈 DNA 서열 업스트림의 클로닝

1.1 발현 벡터로 차이니즈 햄스터 GAPDH 유전자의 비변역된 게놈 DNA 서열 업스트림의 클로닝

차이니즈 햄스터 GAPDH 유전자의 비변역된 게놈 DNA 서열 업스트림은 PCR에 의해 CHO-S (Life Technologies) 세포의 게놈 DNA로부터 증폭되었다. 게놈 DNA는 실시예 1에서 설명된 바와 같이 추출되었다. 구조체는 실시예 1에서 설명된 바와 같이 리포터 유전자 구조체 [REP]의 발현을 위하여 마우스 CMV 프로모터 또는 차이니즈 햄스터 GAPDH 프로모터를 사용하여 제조된다.

[0337] 마우스 CMV 프로모터와 혼합하여 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림의 클로닝을 위하여, 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림 및 서열번호 30으로 엠플리콘을 유도하여 3 kbs 단편 (bps 672 to 3671 of 서열번호 29)의 증폭에 사용되었다. 엠플리콘은 프라이머에 의해 도입된 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림 및 5' 및 3' 제한효소 자리를 포함한다.

[0338] 차이니즈 햄스터 GAPDH 프로모터와 혼합하여 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림의 클로닝을 위하여, 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림 및 서열번호 31로 엠플리콘을 유도하는 GAPDH 프로모터 (bps 672 to 4179 of 서열번호 29)를 포함하는 게놈 DNA 서열을 포함하는 3508 bps 단편을 증폭하기 위하여 프라이머 GlnPr1902 및 GlnPr1905가 사용되었다. 두 번째 PCR에서, GlnPr1901 및 GlnPr1902는 서열번호 32로 유도하는, 단지 프로모터 영역(bps 3672 to 4179 of 서열번호 29)을 포함하는 508 bps 단편의 증폭에 사용되었다. 벡터 "A"에서 사용된 인트론(실시예 1에서 설명됨)은 프라이머 GlnPr1903 및 GlnPr1904를 사용하여 증폭되었다.

[0339] 첫번째 융합 PCR은 템플레이트로서 서열번호: 32를 갖는 엠플리콘 및 인트론 서열을 갖는 엠플리콘을 사용하여

프라이머 GlnPr1904 및 GlnPr1901로 수행되었다. 상기 엠플리콘은 차이니즈 햄스터 GAPDH 프로모터, 인트론 및 프라이머에 의해 도입된 5' 및 3' 제한효소 자리를 포함한다. 모든 프라이머는 표5에 나타냈다.

[0340] 두 번째 융합 PCR은 템플레이트로서 서열번호 31을 갖는 엠플리콘 및 인트론 서열을 갖는 엠플리콘을 사용하여 프라이머 GlnPr1905 및 GlnPr1904으로 수행되었다. 엠플리콘은 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림, 차이니즈 햄스터 GAPDH 프로모터, 인트론 및 프라이머에 의해 도입된 5' 및 3' 제한효소 자리를 포함한다.

[0341] 1% 아가로우즈 젤에서 정제 후에, "뉴클레오스핀 젤 및 PCR 클린업(NucleoSpin Gel and PCR Clean-up)" (Macherey Nagel, Oensingen, Switzerland)를 사용하여 관심의 밴드를 자르고 정제하였다. 정제된 단편은 제로블런트 PCR 클로닝 키트(Zero Blunt PCR cloning kit) (Invitrogen, Carlsbad, CA, USA)를 사용하여 플라스미드 pCR_Blunt로 클론되었다. 결찰 생성물(Ligation products)은 컴피턴트(competent) *E.Coli* TOP10(One Shot®TOP 10 Competent *E.Coli*; Invitrogen, Carlsbad, CA, USA)으로 형질전환되었고 미니프렙의 제한 분석으로 분석되었다. 상기는 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림을 포함하는 플라스미드 pCR_blunt [CHO-업스트림GAPDH], 벡터"A"로부터 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림 및 GAPDH 프로모터 및 인트론을 포함하는 pCR_Blunt [CHO-upstreamGAPDH_GAPDH프로모터] 및 벡터"B"로부터 GAPDH 프로모터 및 인트론을 포함하는 pCR_Blunt [CHO-GAPDH프로모터]로 유도한다.

[0342] 배출된 유전자의 발현에 대한 자체 효과상에서 엠플리콘의 평가를 위하여, 벡터"A"(실시예 1에서 설명됨)가 사용되었다. 이전에 설명한 바와 같이, 상기 벡터에 사용된 발현 카세트는 배출된 IgG1 및 GFP를 위한 폴리시스트로닉(polycistronic) 유전자 코딩을 포함한다(실시예 1 참조). 따라서 형질전환된 세포는 IgG1 단클론 항체를 배출할 것이고 의존적 방식으로 세포 내 GFP를 측정할 것이다.

[0343] 차이니즈 햄스터 GAPDH 유전자의 게놈 DNA 서열 업스트림을 포함하는 3 kb 삽입체 단편을 배출하기 위하여, 플라스미드 pCR_Blunt [CHO-upstreamGAPDH]를 제한효소 NaeI를 사용하여 절단하였다. 상기 삽입체는 "A"의 백본 상에서 클론되고 제한효소 NruI를 사용하여 절단되고 CIP처리되었다(CIP; NEB, Ipswich, MA, USA). 백본 및 삽입체는 T4 DNA 라이케이즈를 사용하여 결찰되었고(T4 DNA ligase, NEB, Ipswich, MA, USA) 후속적으로 컴피턴트 *E.coli* PIR1으로 형질전환되었다. 클론은 미니프렙 제조를 위하여 선택되고 후속적으로 제한분석을 하였다. 얻어진 플라스미드는 "A_GAPDH_UP"으로 명명하였고, 시퀀싱 분석으로 확인하였고 뉴클레오본드 엑스트라 미디키트(NucleoBond Xtra Midi kit)(Macherey Nagel, Oensingen, Switzerland)를 사용하여 미니프렙 규모로 생성되었다.

[0344] 차이니즈 햄스터 GAPDH 프로모터를 사용하여 발현 구조체의 클로닝을 위하여, 삽입체 단편은 제한효소 NheI 및 NruI를 사용하여 분해에 의해 플라스미드 pCR_Blunt [CHO-업스트림GAPDH_GAPDH 프로모터] 및 pCR_Blunt [CHO-GAPDH프로모터]로부터 방출되었다. 얻어진 단편은 벡터 "A"의 백본에서 클론되었고 동일한 효소를 사용하여 오픈되었고 CIP처리되었다. T4 DNA 라이케이즈로 결찰 및 컴피턴트 *E.coli* PIR1로 전환 후에, 클론은 미니 제한 분석을 위해 선택되었다. 얻어진 플라스미드는 "A_GAPDH_UP_Prom" (차이니즈 햄스터 GAPDH 및 프로모터의 비번역된 게놈 DNA 서열 업스트림을 갖는 플라스미드)으로 명명되었고 "A_PR"(단지 프로모터를 갖는 플라스미드)는 시퀀싱 분석에 의해 확인되었고 키트 뉴클레오본드 엑스트라 미디키트(NucleoBond Xtra Midi)(Macherey Nagel, Oensingen, Switzerland)를 사용하여 미니프렙 규모로 생성되었다.

[0345] 2. 리포터 유전자 구조체의 발현 상에서 차이니즈 햄스터 GAPDH 유전자의 비번역된 게놈 DNA 서열 업스트림의 평가

[0346] CHO-S 세포는 10 ml의 배지 부피(실시예 2에서 설명됨)를 사용하여튜브스핀생물반응기(tubespins bioreactors)에서 형질전환되었다. 형질전환된 세포는 37°C, 5 % CO₂ 및 80 % 습도에서 200 rpm 진탕으로 진

탕 배양기에서 배양되었다. 세포의 상층액은 단백질 A 바이오센서 (ForteBio, Menlo Park, CA, USA)가 있는 옥테트 QK 시스템을 사용하여 IgG1 발현에 대하여 분석되었다. 결과는 도 10에 나타냈다.

[0347] 마우스 CMV 프로모터 (A)와 비교하여 GAPDH 프로모터 ("A_PR")를 포함하는 플라스미드의 발현 수준은 50 %에 의해 감소되었고, 차이니즈 햄스터 GAPDH 프로모터는 바이러스 프로모터만큼 강하지 않다. 차이니즈 햄스터 GAPDH 프로모터 ("A_GAPDH_UP_Prom")와 혼합하여 차이니즈 햄스터 GAPDH 유전자의 비변역된 게놈 DNA 서열 업스트림을 포함하는 플라스미드는 단지 GAPDH프로모터 ("A_PR")을 갖는 구조체와 비교하여 발현에서 2 폴드 증가를 나타낸다. 차이니즈 햄스터 GAPDH 유전자의 비변역된 게놈 DNA 서열 업스트림 및 마우스 CMV 프로모터 ("A_GAPDH_UP")을 포함하는 플라스미드는 단지 마우스 CMV 프로모터("A")를 포함하는 플라스미드에서 40% 이상의 가장 높은 발현 및 증가를 나타낸다("A"). 상기는 차이니즈 햄스터 GAPDH 유전자의 비변역된 게놈 DNA 서열 업스트림은 리포터 단백질의 발현 상에서 인핸서 효과를 가진다.

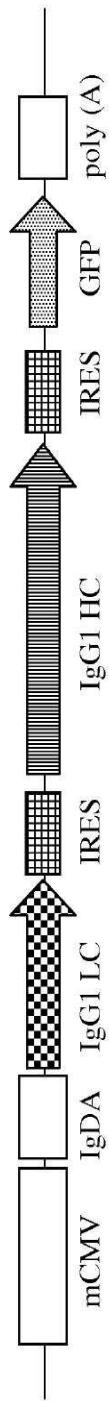
[0348] 표 5: 실시예 5에서 클로닝에 사용된 프라이머

표 5

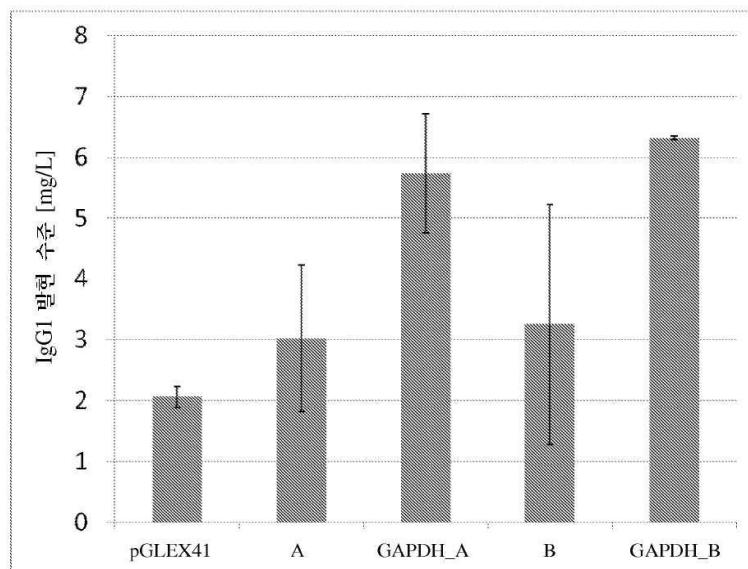
프라이머	서열식별번호 (SEQ ID No)	서열	방향 (Orientation)	제한자리 (Restrn. site)
GlnPr 1896	서열식별번호 33	TACGGCCGGCTTCACTGTACAGTGGCACAT	정방향	NaeI
GlnPr 1897	서열식별번호 34	TCAGGCCGGCGTGGTTCTCGGTAGTGAC	역방향	NaeI
GlnPr 1901	서열식별번호 35	TACTCGCGAAGAAGATCCTCAACTTTCCACAGCC	정방향	NruI
GlnPr 1902	서열식별번호 36	GTTCACTAAACGAGCTCTGCTATTATAGGAAC TG GGGTG	역방향	/
GlnPr 1903	서열식별번호 37	CACCCCAGTTCCCTATAAATAGCAGAGCTCGTTAG TGAAC	정방향	/
GlnPr 1904	서열식별번호 38	CGCTAGCACCGGTCGATCGA	역방향	NheI
GlnPr 1905	서열식별번호 39	TACTCGCGATTCACTGTACAGTGGCACATAC	정방향	NruI

도면

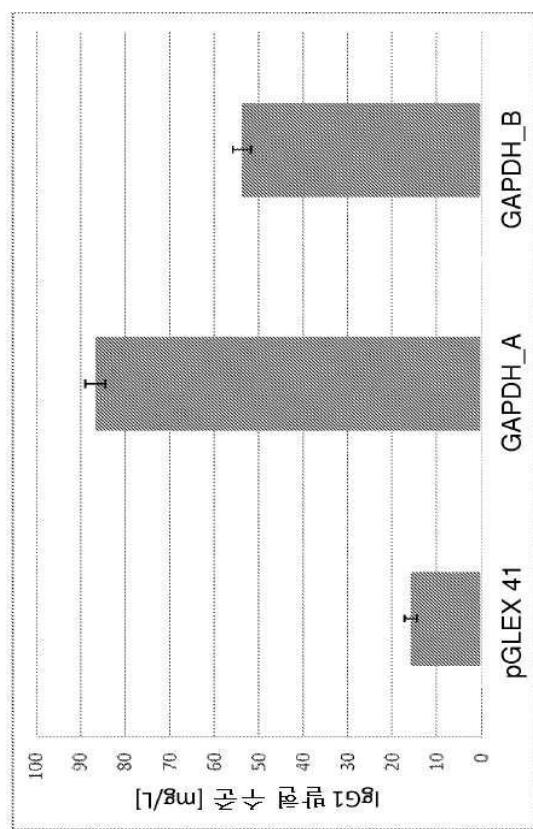
도면1



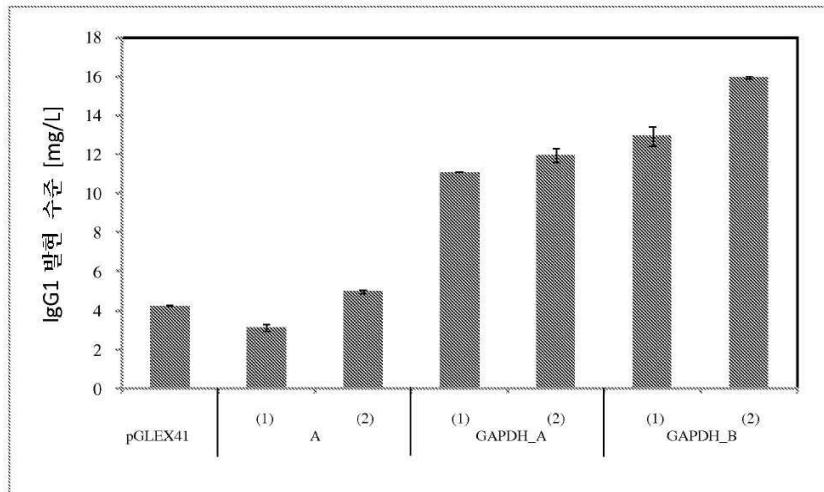
도면2



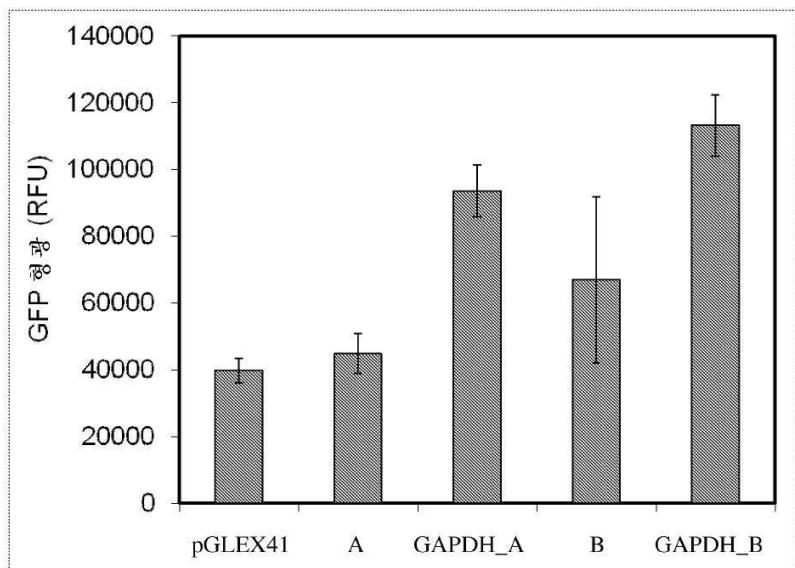
도면3



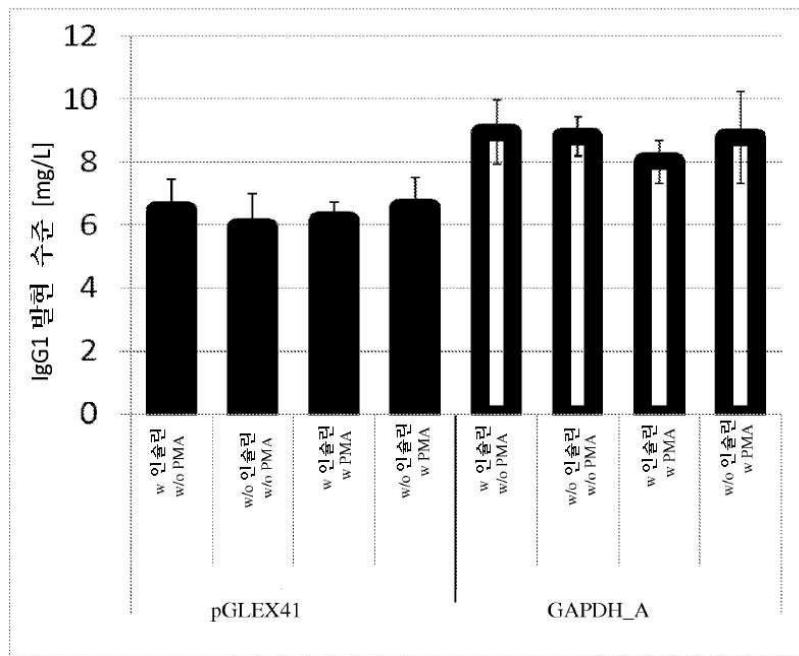
도면4



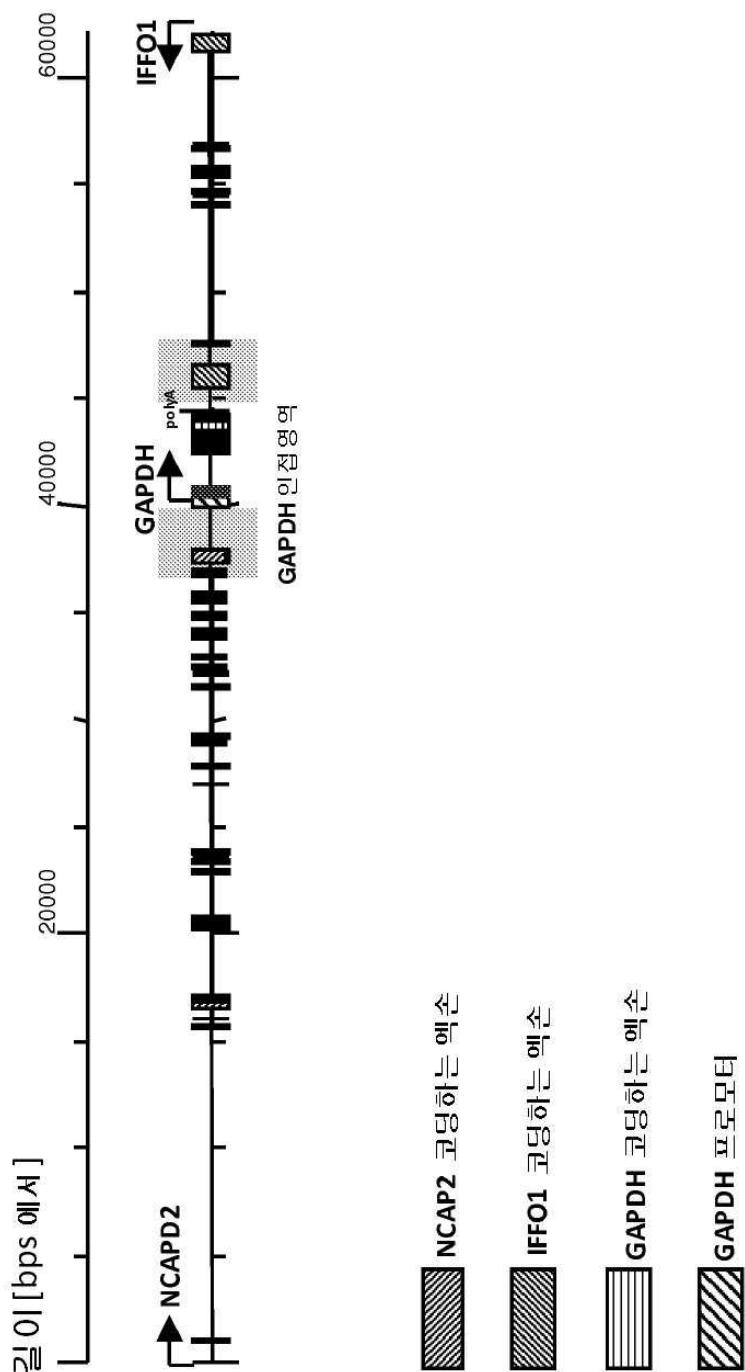
도면5



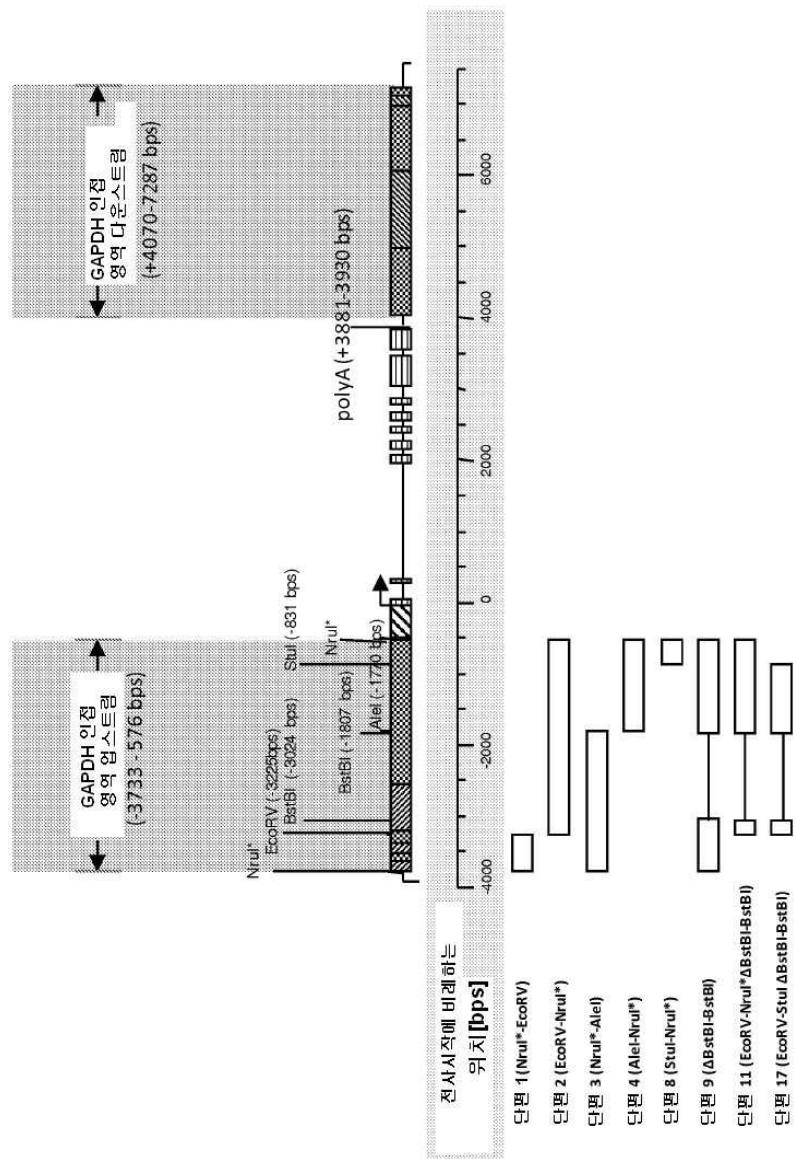
도면6



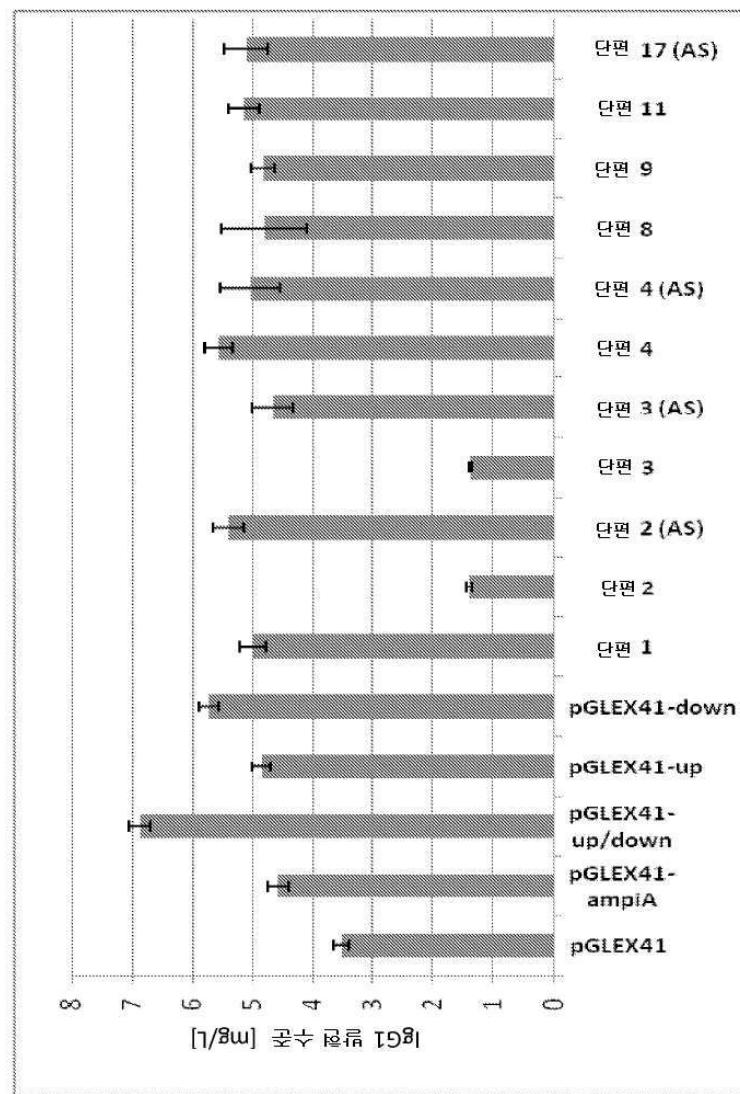
도면7



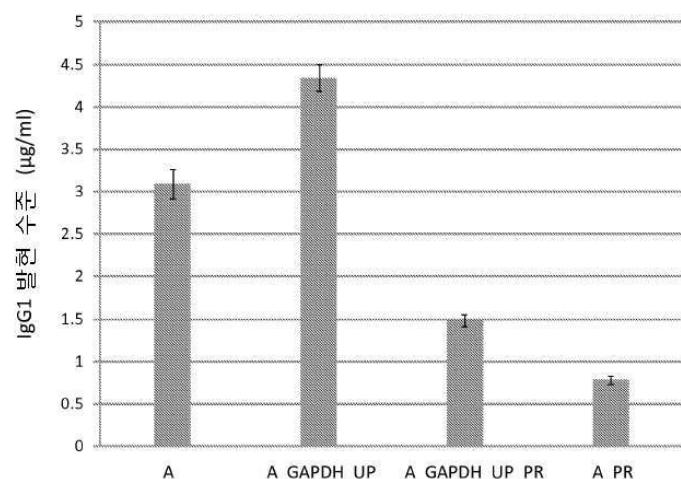
도면8



도면9



도면10



서 열 목록

<110> Glenmark Pharmaceuticals SA

<120>	Expression Cassette	
<130>	2014fpi-06-003	
<150>	PCT/US US 61/567,675	
<151>	2011-12-07	
<160>	39	
<170>	Kopatent In 2.0	
<210>	1	
<211>	39	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer 1	
<400>	1	
	attatttcgat atggctcctg gcatctctgg gaccgaggc	39
<210>	2	
<211>	39	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer 2	
<400>	2	
	atcgatcgatc agctttagat tgtccaaagca ggttagccag	39
<210>	3	
<211>	36	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer 3	
<400>	3	
	agcaagtttt tcttggcctt cagttatggc tggctt	36
<210>	4	
<211>	39	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer 4	
<400>	4	
	tggcgtttttt aagctggcac cactacttca gagaacaag	39
<210>	5	

<211> 3179

<212> DNA

<213> Artificial Sequence

<220><223> 5' GAPDH fragment

<400> 5

atcgtcgcga agctttagat tgtccaagca ggttagccaga gagcgccatc agccaagaaa 60

ccatccactg gtacgtaagg cagcctgtgc gggcagacc agactggcc ctccctcct 120

gcagtgattt gttcttctt ctttttaaa tcacgtttc ctgcctttc tagttctag 180

gtaccagcct ctggcttcta cagcctcaga caatgacttt gtacaccag agccccgccc 240

tactaccgt cgccatccaa acacccagca gcgagcttcc aaaaagaaac ccaaagtgt 300

cttctcaagt gatgagtcga gtgaggaagg tatgatgctc ccgcctgttc cggccgaga 360

aggcacacag cttagggtcga gaggcgttgt ttccatagga cctgctgcgg gggcctgagt 420

gtagatgctc tgccccactg ccgcagaagg gcctctcctg tacagcttgg attttatttc 480

ttctgtcggt tggggattt tctcaactgt tctctgatat ctattttc accatcttg 540

tgactcagct ttttcttatt ctttaattt tttcataga tcttcagca gagatgacag 600

aagacgagac accaagaaa acaactccca ttctcagagc atcggtcgc aggacagat 660

ccttaggaagt ctgttctgt cttccctgtc caggatattcc tgggtatgtt cctggattc 720

gaattctgtt tccctgtta aatatttgc tgccttttt tttaaaaaaa aaaaaaggcc 780

gggcactgtg gtcacgcct gtaatccag cacttgcga taccaggcg ggtggataac 840

ctgaggttagg gagttcgaga ccagcgtac caacatggag aaacccatc tctactaaaa 900

ataaaaaatt agccggcgt attggcgtgc gcctgtatc ccagctactc aagaggctga 960

ggcaggagaa tcgcctgaac ccagaggcgg aggttgtgt gggccaaat cacaccattt 1020

cactccagct tggcaacaa tagcgtaccc ctatctcaaa tttttttttt aatgcctaca 1080

cgctctttaa aatgcgttgc ttctcttaa attgcctaa ctgcactgcg ttggctgtct 1140

tcaactttgg aatatatgtt tgcctatctc ctgttttttctt aatgcataaa tttttttata 1200

tactttttaga catttttcc taagcttgc tttgtttcat ctgcacttgcg agcccgat 1260

catgcgtcgag agagagggtt atcgtgcgt agagagatga gtgagcccg agtccctagg 1320

cctgtcccggt gatggcagat gagcttctgt cccgtcact gccacccccc ccctctcaac 1380

ctctggaccc tgcacagtga ccagacagcc tctctggggaa gaattatgca gtgcctaggc 1440

tccagatcg tgcgttgcgttcc cggggggcaa ttttgcgtgc cagaggacat ctgcacacac 1500

ctggggctgttgcgttgcgttcc atgcctata gggagaat gtcaccagca tttgtggaa 1560

gaggccaggg atgtggctca acatcctgca gtgcacagga tggccctca acaaagaatc	1620
acacggccca caatgtcaat agcgtcacag ttgagaaaac ctgctctaga ccaagggttg	1680
ctttctgccc tgtgcctcac cccacccca ctcgtgttcc ctaatccat ctccaaaggt	1740
tggcagcaga ccggccagg ctcgtggaag ttcatatcat gatccctcc agctctgcag	1800
gagacaagac ctgtctcca gcattcctca ttgtccgg gtctcagag ggcgtgagct	1860
atgctgcagg cgggctgccc cctgaaggct gcgcacccct ctccagctcc tcaagtcttc	1920
tctgctgagt cacccctgaa ccggaggctg tgagctggct gtcgtgacca cactggtgc	1980
tctgctgtca tgacaacagc acactacgtc agtagtgctc cctggcact gagctccctc	2040
tttgcgggaa gaagacagta atgaaaaatg acaagcatga ggcagagggg aagatcacgc	2100
ttgggtggtg caggagcatg gaggtgtct taatgtctc aatgagaaag ggttaacgg	2160
cctggttgca ggaatagctg agtcagaggt ggggcttcct ccactcccc accccaccc	2220
tttaccatt agggacccctc ttgccttgc ctgtactc tgctctgggt ggtcatttg	2280
aaaagccgc accaaccatg ccagtggcag ccagacgagg acacagctg gctctggtc	2340
ccagcaggaa aggcaatccc agaaaggcag gtcaggac tggagtcctg tgggtgttt	2400
ttaagcaaag attatcacca ggcaggctaa acttagcaac cggcttttag ctagaaggc	2460
agggggctgg tgtcaggta tgctggcca gcaaagaggc cccggatccc cctccatgc	2520
acctgctgat gggcaaggc caccccaccc caccccttc cttacaagtg ttcagcaccc	2580
tccatccca cactcacaaa cctggccctc tgccctcta ccagaagaat ggatccctg	2640
tgggaggggg caggggaccc tttccaccc tggtcccaag acctcttttc ccacttttc	2700
cctttcttg actcaccctg ccctcaatat ccccccggcgc agccagtgaa agggagtccc	2760
tggctctgg ctgcctgca cgtccctgg cggggaggga ctcccccct cacgtccgc	2820
tcttcgcccc aggctggatg gaatgaaagg cacactgtct ctctccctag gcagcacagc	2880
ccacagggtt ccaggagtgc ctttgtggg ggcctctggg ccccccacccag ccatccctg	2940
ctccgcctgg gggcccgagcc cggagagagc cgctgtgca cacagggccg ggattgtctg	3000
ccctaattat caggtccagg ctacagggt gcaggacatc gtgaccccttcc gtgcagaaac	3060
ctccccctcc ccctcaagcc gcctcccgag cctcccttcc ctccagggcc ccagtgc	3120
gtgcccagtg cccagcccg gcctcggtcc cagagatgcc aggagccatc gcgaataat	3179
<210> 6	
<211> 3238	
<212> DNA	
<213> Artificial Sequence	

<220><223> 3'-GAPDH fragment

<400> 6

tggcagtaact aagctggcac cactacttca gagaacaagg cctttcctc tcctcgctcc	60
agtccctaggc tatctgtgt tggccaaaca tggaaagaac tattctgtgg gcagccccag	120

ggaggctgac aggtggagga agtcaggcgt cgcaactggc tctgacgctg actggtagt	180
ggagctcagc ctggagctga gctgcagcgg gcaattccag ctggcctcc gcagctgtga	240
ggtcttgagc acgtgtctca ttgccttctg tgccctcggt tcttacatcga ggacatcg	300
gccagccccct aaggcttca agcaggattc atcttagtta accaagtacc taaaaccatg	360
cccaaggcgg taaggactat ataatgtttt aaaaatcggtt aaaaatgcctt cctcgcatag	420
ttttgaggaa gatgaactga gatgtgtcag ggtgacttat ttccatcatc gtccttaggg	480
gaacttgggtt agggcaagg cgtgttagtgg ggaccttaggtt ccagaccctt ggctctgcca	540

ctgaacggct cagttgcctt gggcagttac tcccgccctt cacttgcac gtgtgtttac	600
ctagtggaga caaaagtaca tacctcggtt gaggcgcac gcctgttaacc ccagca	660
gggaggccaa ggtgggtgtt tcaacctgagg tcaggagttt gagaccagcc tggccaacat	720
ggtaaaactc cgtctctact aaaattacaa aaatcagccca ggcttcatgg cacatgc	780
tagtcccage tacaggcatg ctgaaggcagg agaattcggtt gcaccccgaa ggcagaggct	840
gcagttagtgg gagaccacac cactgcactc cagccttaggc aacagatgtt gagactccat	900
ctcaaaaaaaa aaaaaaagtac ctacctcaga gttcaaacta gtgaatatta ggaagtgttt	960

gagacagtga caccaaaagtgcacaataat actcgccagt ttcatttattttttaaagaatc	1020
catttgaatgtcagtcacac acagcctctt ataccgaggc attgtgaacc gcatctcccc	1080
agcttctcca ggctttcca agaatcaggg acactgttagc ctgttggctt cagtgatgtt	1140
cagacacggaa ggaaggcacat cttagtgcatactttaaaca gagaccctgttgcgcacatac	1200
acccgcgcac acaatgcatttgg agcttccactt tctctgtcat tctgcgtgttgc ccaggagagc	1260
aagagctccc acctcccttc aaaacactgtt gcccattccgg ggcactaagg cctctttaaa	1320
gcacggcacc tccacgaggg agggccacag ccacatacac tccacctggc aggtggacag	1380

cgtgagcacg tggaccatag cagggacaag gtgccccggc cagccccaaac gcccctgtcc	1440
gctgacaggg acagaagccc tctccagctg cgtgtgttgc agaggccatg cgtagccctcc	1500
agctgcatttccatcactc cagtcctgg gccagtttgc accagttgtgg aagacagtgtt	1560
gctggctccg gacaacaggg atggaggaaa ggtcccacat tcacatttgc gatacgttgg	1620
caaggtgagg gcccgcatac gctctggcag cattttaaag atggggaaatgttgcacacacc	1680
cacgcgttggaa ggcaggagag ccccaactgtt ggtggaaatgttgcacacacc ggttagggcca	1740

agcctagctc cagacacccc agagccctgg agaagccaag actgagggag aaagcctgag	1800
ggaggagcgc cccagtc(ccc) agggaccggc ctggcaga gctgcagctg atgttcccct	1860
ctgtgcagcc ccaccctctg cctcgctgag ctccctgctg cgaggcctc gggtgcaagg	1920
gggaggcagg tctctatctc atggagctgt cagatgagac atcgatcg gagtcctcag	1980
cctcgcttgg cggcgccgc gggtcgctaa gcgggaccgc agtgaagca ggagacttc	2040
tagaaaaaaa caccagttgt caaccttggg gcaggcagga atcctgaaga cggacggcac	2100
tcctcctcct gctgcctcac cctctggcag cccgtgagaa gtaccgaaag cgagggcggg	2160
gccgcgggat ggcgagggag cggcagggac tgaactctc ccaaaccac cctgacaggg	2220
aaatgggccc cgccctgtgtc ttggaaactc agaggctgag gt caggcatg atggctcact	2280
cctgttaattc cagcactttgg gggggcagag gcgggtggat cacgacgtca ggagttaaag	2340
gcaaggcctgg ccaacatggt gaaacccat ctctactaaa aatgaaaaaa ttagccaggt	2400
gtggtggcgg ggccttggg ggctgaggca agataatcac ttgaacctgg gaggcggagg	2460
ttgttagttag ccaaaaaaaaaaaaaaaa aatagctgaa gtacagtag gagagaagct	2520
gctgagcctc cagcacccctg actctaggc cttggctta tgtctatctg cagtatTTT	2580
gtgatTTTaaatTTTact ttcttgttgc ggtgtactt acacagggtc aaatgcacaa	2640
atcatggccc tgactttaga taaaaatctg ccccccacaaac cttctgttc cttggcagg	2700
ttaaaactgc ctctaacccag gggaaaccacc agagctggtg gccttgggag gtttccagcc	2760
tcccgcatg aatggacata gtcataccaa ctgccaagg agagagctgt gggcttggc	2820
cagccccacc aggttaactcc caaaggcagccccacagca agatgtgacc cagtcattgc	2880
ctgagggtct ctggggctgt gttccaaacct ttctccccgc tgggtcccccc tggaaggccc	2940
catgcccagg ggaggcgcct accggctcc agactgggtg atgagccggc ggcaggctc	3000
catctgcacg tccaggccgc gttcatgtc gcacatctcc atgtactctgt gcagggtccg	3060
gttcatgtcg ttcttggccg tggccaaactc cagctgtggt gggcaggca gggccatcg	3120
ggtttaacagg tggtttcac agcgcccttg ttggcccccgc caggaggcca acacgccaag	3180
agcagtggct gggccggggg cccaggcagc cattactgaa ggctcagaag tacttgc	3238
<210> 7	
<211> 3164	
<212> DNA	
<213> Artificial Sequence	
<220><223> Upstream GAPDH flanking region	
<400> 7	

cgaagcttga gattgtccaa gcaggttagcc agagagcgcc atcagccaag aaaccatcca	60
ctggtagtca aggccgcctg tgcggcgag accagactgg gccctccct cctgcagtga	120
tttggtttctt ctctttttt aaatcacgtt ttccctgcctt ttcttaggttc taggttaccag	180
cctctggctt ctacagccctc agacaatgac tttgtcacac cagagccccg ccgtactacc	240
cgtcgccatc caaacaccca gcagcgagct tccaaaaga aacccaaagt tgtcttca	300
agtgtatgagt ccagtgagga aggtatgatg ctcccgctg ttcccgccg agaaggcaca	360
cagcttagggt gcagagggct ggtttccata ggacctgctg cggggccctg agttagatg	420
ctctgccccca ctgcccaga agggcctctc ctgtacagct tggattttat ttcttctgtg	480
cggtgtggaa ttgtctact ttgtctctga tatctatttt ttccacatct ttgtgactca	540
gccttttctt attcctttaa ttctttgcat agatcttca gcagagatga cagaagacga	600
gacacccaag aaaacaactc ccattcttag agcatcggt cgccaggcaca gatccttagga	660
agtctgttcc tgtccctccct gtgcagggtt tcctgttaggg tgacctggaa ttcaattct	720
gtttcccttg taaaatattt gtctgtctt ttttttaaa aaaaaaaaag gccggcact	780
gtggctcacg cctgtaatcc cagcaacttg cgataccaag gcgggtggat aacctgaggt	840
agggagttcg agaccagct gaccaacatg gagaaccccc atctctacta aaaataaaaa	900
attagccggg cgtattggcg tgccctgtt atccctacta ctcaagaggc tgaggcagga	960
gaatcgccgt aacccagagg cggaggttgt agtgagccga aatcacacca ttgcactcca	1020
gcttggcaa caatagcgaa cctccatctc aaattaaaaa aaaaatgcct acacgtctt	1080
taaaatgcaa ggcttctct taaattagcc taactgaact gcgttgagct gcttcaactt	1140
tggaatataat gtttgccat ctccttgc ttctaatgaat aaatgtttt atatactttt	1200
agacatttt tcctaagctt gtctttgtt catcttac attagccag tttcatgcag	1260
cagagagagg gttatcagtg cagagagaga tgagtgagcc cagactcta gggcctgtcc	1320
cggatggca gatgagcttc ctgccccgtc actgcccaccc ttccctctc aacctctggaa	1380
ccctgcacag tgaccagaca gcctctctgg ggagaattat gcagtgccta ggctccagat	1440
cagtgccttctt gaaccggggg caatttgtc tgccagagga catctgacaa cacctgggc	1500
ctgtttgtt gtcatacgct atagggaaat aatgtacca gcattgtgg gaagaggcca	1560
gggatgtggc tcaacatcct gcagtgcaca ggatggcccc tcaacaaaga atcacacggc	1620
ccacaatgtc aatagcgtca cagttgagaa aacctgcctt agaccaaggg ttgccttctg	1680
ccgtgtgcctt caccctaccc ccactcgtgt tccctaatcc catctccaaa ggttggcagc	1740

agaccggccc aggctcggtgg aagttcagat catgatcccc tccagctctg caggagacaa	1800
gacctgtctc ccagcattcc tcattgttcc cgggtctgca gagggcgtga gctatgtgc	1860
aggcgggctg cccctgaag cctgcgcacc cctctccagc tcctcaagtc ttctctgctg	1920
agtcaccttc gaaccggagg ctgtgagctg gctgtctgca ccacactggt gcctctgctg	1980
tcatgacaac agcacactac gtcagtagtg ctccctggc actgagctcc ctcttgccg	2040
ggagaagaca gtaatgaaaa atgacaagca tgaggcagag gggaaagatca cgcttgggtg	2100
gtgcaggagc atggaggtgc tcttaatgct ctcaatgaga aagggttaac ggtcctgggtt	2160
gcaggaatag ctgagtcaga ggtggggctt cctccactcc cccacccac cccttcacc	2220
attagggacc ttttgcctt gctttgtcta ctctgtctg ggtggtcatt gtgaaaagcc	2280
cgcaccaacc atgccagtgg cagccagacg aggacacagc ctggctctgg gtcccagcag	2340
gaaaggcaat cccagaaagg cagggtcagg gactggagtc ctgtgggtgc ttttaagca	2400
aagattatca ccaggcaggc taaacttagc aaccggctt tagctagaag ggcaggggc	2460
tgggtcagg ttatgctggg ccagcaaaga ggcccggtat cccctccca tgcacctgct	2520
gatggccaa gcccacccca ccccacccca ttcccttacaa gtgttcagca ccctccatc	2580
ccacactcac aaacctggcc ctctgccctc ctaccagaag aatggatccc ctgtgggagg	2640
ggcaggggga cctgttccca ccgtgtgcc aagaccttt ttccctttt ttcccttcc	2700
ttgactcacc ctgcctcaa tatccccgg cgccagccagt gaaaggaggat ccctggctcc	2760
tggctgcct gcacgtccca gggcgggag ggacttccgc ctcacgtcc cgcttgc	2820
cccaggctgg atggaatgaa aggcacactg tctctctcc taggcagcac agccacagg	2880
tttccaggag tgcctttgtg ggaggcctct gggcccccac cagccatcct gtcctccgc	2940
tggggccca gcccgagag agccgctggt gcacacaggg ccgggattgt ctgcctaat	3000
tatcagggcc aggtacacagg gctgcaggac atcgatgaccc tccgtgcaga aacctccccc	3060
tccccctcaa gccgcctccc gagcctcctt cctctccagg ccccccagtgc ccagtgc	3120
gtgcggcc caggcctcggtt cccagagat gccaggagcc atcg	3164
<210> 8	
<211> 3224	
<212> DNA	
<213> Artificial Sequence	
<220><223> Downstream GAPDH flanking region	
<400> 8	
actaagctgg caccactact tcagagaaca aggcctttc ctctcctcgc tccagtcata	60

ggctatctgc tggccaa acatggaaga agctattctg tggcagccc cagggaggct	120
gacaggtgga ggaagt cagg gctcgactg ggctctgacg ctgactgggtt agtggagctc	180
agcctggagc ttagctgcag cggcaattc cagctggcc tccgcagctg tgaggcttg	240
agcacgtgct ctattgctt ctgtgccctc gtgtcttatac tgaggacatc gtggccagcc	300
cctaaggct tcaaggcagga ttcatctagg taaaccaagt acctaaaacc atgccaagg	360
cggtaaggac tatataatgt taaaaaatcg gtaaaaatgc ccacctcgca tagtttgag	420
gaagatgaac ttagatgtgt cagggtgact tattccatc atcgtccta gggaaacttg	480
ggtagggca aggcgtgtag ctgggaccta ggtccagacc cctggctctg ccactgaacg	540
gctcagttgc ttggcagt tactccggg ctcactttg cacgtgtgct tacctagttg	600
agacaaaagt acatacctcg ttagagcgcg cacgcctgt accccagcac tttggggaggc	660
caaggtgggt gtatcacctg aggtcaggag tttgagacca gcctggccaa catggtaaaa	720
ctccgtctct actaaaatta caaaaatcg ccaggctca tggcacatgc ctatagccc	780
agctacaggc atgtgaagc aggagaatcg cttgcacccc ggaggcagag gctgcagtga	840
gctgagacca caccactgca ctccagccta ggcaacagag tatgagactc catctaaaa	900
aaaaaaaaaaag tacctacctc agagttcaaa ctagtgaata ttaggaagtg cttgagacag	960
tgacacaaaaa gtgcacaata aatactcgcc agtttcattta ttattaaaga atccatttgaa	1020
atgtcagtc aacacagcct cctataccga ggcattgtga accgcacatc cccagcttct	1080
ccagggcttt ccaagaatca gggacactgt agcctgttg tctcagtgta tgacagacac	1140
ggaggaagca catcttagc tgatacttaa acagagaccc tgagcgcaca tacacccg	1200
cacacatgca tggagcttca cttctctgt cattctgcag tgaccaggag agcaagagct	1260
cccacctccc ttcaaaacac tggccatc cgggcaacta aggcccttt aaagcacggc	1320
acctccacga gggagggcca cagccacata cactccaccc ggcaggtgga cagcgtgagc	1380
acgtggacca tagcaggac aaggtcccc ggcagcccc aacgcctct gcccgtgaca	1440
gggacagaag cccttcagg ctgcgtgtgc tgcagaggcc atgcgttagcc tccagctgca	1500
ttctattcca ctccagtgcc tggccagtt agcaccagtg tggaaagacag tgagctggct	1560
ccggacaaca gggatggagg aaaggcccattt cctgatacgt ggacaagggtg	1620
agggccgca atcgtctgg cagcatttt aagatggga agtagcagac acccacgcgt	1680
gaagggcagga gagcccaac tgggtggaa atggccctt aatggtaggg ccaagcctag	1740
ctccagacac cccagagccc tggagaagcc aagactgagg gagaagcct gaggaggag	1800

cgccccagtc cccagggacc ggcctggtgc agagctgcag ctgatgttcc cctctgtca	1860
gccccaccct ctgcctcgct gagctccctg ctgcgagggc ctgggtgca agggggaggc	1920
aggctctat ctcatggagc tgtcagatga gacatcgca tcggagtct cagcctcgct	1980
tggcggcggc ggcgggtcgc taagcggac cgcaatgaaa gcaggagact ttctagaaaa	2040
aaacaccagt tgtcaacatt gggcaggca ggaatcctga agacggacgg cactcctct	2100
cctgctgcct caccctctgg cagccctgtga gaagtaccgg aagcgagggc gggccgcgg	2160
gatggcgagg gagcggcagg gactgaactc tctccaaacc caccctgaca gggaaatggg	2220
ccccgcctgt gtctggaa ctcagaggct gaggtcaggc atgatggctc acgcctgtaa	2280
ttccagcact ttgggaggca gaggcgggtg gatcacgacg tcaggagttc aaggcaagcc	2340
tggccaacat ggtgaaaccc catctctact aaaaatgcaaa aaattagcca ggtgtgg	2400
cgggcgcctt ggaggctgag gcaagataat cacttgaacc tgggaggcgg aggttgttagt	2460
gagccaaaaa aaaaaaaaaa agaaatagct gaagtcacag taggagagaa gctgtgagc	2520
ctccagcacc ctgactctag ggccttgct ttatgtctat ctgcagtatt tttgtgattt	2580
ttaaaaattc acttcttgt tgccgtgtaa cttacacagg gtcaaatgca caaatcatgg	2640
ccctgacttt agataaaaat ctgccccac aaccctctg ttccttgcca gttttaaac	2700
tgccctctaacc caggggaacc accagagctg gtggccttgg gaggtttcag ccctccgtc	2760
atgaatggac atagctcatc caactgccaa gggagagagc tgtggcttg gcccagcccc	2820
accaggtaac tccaaaggg cagccccaca gcaagatgtg acccagtcat tgcctgaggg	2880
tctctgggc tgtgttccaa ctttctccc cgctgtgtcc ccctgaaagg cccatgccc	2940
aggggaggcg cttaccggtc tccagactgg gtgatgagcc ggcggcaggt ctccatctgc	3000
acgtccaggc cgcgcttcat gtcacatc tccatgtact cgtgcaggtg ccgttcatg	3060
tcgttcttgg ccgtggccaa ctccagctgt ggtggggcag gcagggccat cgggttaac	3120
aggtggcggt cacagcgct ctgttgcacc cggcaggagg ccaacacgcc aagagcagtg	3180
gctggccgg gggcccgagc agccattact gaaggctcag aagt	3224
<210> 9	
<211> 511	
<212> DNA	
<213> Artificial Sequence	
<220><223> Fragment 1	
<400> 9	
cgaagcttga gattgtccaa gcaggttagcc agagagcgcc atcagccaag aaaccatcca	60

atgttaccag catttgtggg aagaggccag ggtatgtggct caacatcctg cagtgcacag	1080
gatggccctt caacaagaa tcacacggcc cacaatgtca atagcgtcac agttgagaaa	1140
acctgtctta gaccaagggt tgctttctgc cgtgtgcctc accccacccc cactcgttt	1200
ccctaattccc atcccaaaag gttggcagca gaccggccca ggctcggtt agttcagatc	1260
atgtatccctt ccagctctgc aggagacaag acctgtctcc cagcatctt cattgttccc	1320
gggtctgcag agggcgtgag ctatgctgca ggcgggctgc cccctgaagc ctgcgcaccc	1380
ctctccagct cctcaagtct tctctgctga gtcacccctcg aaccggaggc tgtgagctgg	1440
ctgtcgtgac cacactgggt cctctgctgt catgacaaca gcacactacg tcagtagtgc	1500
tccctggca ctgagctccc tctttgcggg gagaagacag taatgaaaaa tgacaagcat	1560
gaggcagagg ggaagatcac gcttgggtgg tgcaggagca tggaggtgt cttaatgctc	1620
tcaatgagaa agggtaacg gtcctgggtt caggaatagc tgagtcaagag gtggggcttc	1680
ctccactccc ccacccacc ccttcacca ttagggacct tcttgccttgc tcttgcctac	1740
tctgctctgg gtggtcattt tgaaaagccc gcaccaacca tgccagtggc agccagacga	1800
ggacacagcc tggctctggg tcccgccagg aaaggcaatc ccagaaaggc agggtcaggg	1860
actggagtcc tgggggtgtt tttaagcaa agattatcac caggcaggct aaacttagca	1920
accggctttt agctagaagg gcagggggct ggtgtcaggt tatgctggc cagcaaagag	1980
gccccggatc cccctccat gcacctgctg atggccaag gccacccac cccacccct	2040
tccttacaag tggcagcac cctccatcc cacactcaca aacctggccc tctggcccttcc	2100
taccagaaga atggatcccc tgggggggg ggcaggggac ctgttccac cgtgtgcacca	2160
agacactttt tccactttt tcccttttct tgactcaccc tgccctcaat atccccggc	2220
gcagccagtaaaggagtc cctggctctt ggctcgccctg cacgtcccag ggcggggagg	2280
gacttccccc ctcacgtccc gctttcgcc ccaggctgaa tggaatgaaa ggcacactgt	2340
ctctctccctt aggccgcaca gcccacagggt ttccaggagt gcctttgtgg gaggcctctg	2400
ggcccccacc agccatcctg tccctccctt gggcccccag cccggagaga gccgctgggt	2460
cacacaggcgc cgggattgtc tgccttaattt atcagggtcca ggctacaggc ctgcaggaca	2520
tcgtgacctt ccgtgcagaa acctccccctt cccctcaag ccgcctcccg agcctcccttcc	2580
ctctccaggc ccccaagtggcc cagtggccag tgcccagccc aggccctcggtt cccagagatg	2640
ccaggagccatcg	2653
<210> 11	
<211> 1966	
<212> DNA	

<213> Artificial Sequence

<220><223> Fragment 3

<400> 11

cgaagcttga gattgtccaa gcaggttagcc agagagcgcc atcagccaag aaaccatcca	60
ctggtagtta aggcagcctg tgcgggcgag accagactgg gccctccct cctgcagtga	120
tttggtttctt ctctttttt aatcacgtt ttccctgcctt ttcttaggttc taggtaccag	180
cctctggctt ctacagcctc agacaatgac tttgtcacac cagagccccg ccgtactacc	240
cgtcgccatc caaacaccca gcagcgagct tccaaaaga aacccaaagt tgtttctca	300
agtgtatgagt ccagtggagga aggtatgtatg ctccgcctg ttccggccg agaaggcaca	360
cagcttagggt gcagagggtt gggttccata ggacctgctg cggggccctg agttagatg	420
ctctgcctca ctggcgaga agggccttc ctgtacagct tggattttat ttcttctgtg	480
cgggtgtggga ttgtctcaact ttttctgtat tatctatttt ttccatctt ttgtgactca	540
gtttttctt attccttaa ttcttgcat agatcttca gcagagatga cagaagacga	600
gacacccaag aaaacaactc ccattctcag agcatcggtt cgcaggcaca gatcttagga	660
agtctgttcc tttctccctt gtgcagggtt tcctgttaggg tgacctggaa ttcaattct	720
gtttcccttg taaaatattt gtctgtctt tttttttttt aaaaaaaaaaag gcccggcact	780
gtggctcacg cctgtatcc cagcaatggcgatcgatccaaag gcccgggtt aacatgggt	840
aggagatcccg agaccaggctt gaccaacatg gagaaccccc atctctacta aaaataaaaa	900
attagccggg cgtattggcg tgcgcctgtatccatccatca gacccaggaggatcccgat	960
gaatcgcttgc aacccagagg cggagggtt agtgagccga aatcacacca ttgcactcca	1020
gcttggcaaa caatagcgaa cctccatctc aaattaaaaaaa aaaaatgcctt acacgtt	1080
taaaaatgcaaa ggcttctct taaaatttgc taaactgaact gcgttggact gcttcaactt	1140
ttggaaatataat gtttgcctt ctccctgtttt tctaaatgtat aatgtttttt atataactttt	1200
agacatttt tccatagctt gtctttgtttt catcttccat attagcccgat tttcatgcag	1260
cagagagagg gttatcgatg cagagagaga tgagtggatcc cagatccctt gggcctgtcc	1320
cgggatggca gatgagcttc ctggcccgatc actgcccaccc ttccctctc aacctctggaa	1380
ccctgcacatc tgaccagaca gcctctctgg ggagaattat gcgtgcctt ggctccagat	1440
cagtgccttctt gaaccggggg caatttgtc tgccagagga catctgacaa cacctggggc	1500
ctgtttgtt gtcatacgctt ataggaaag aatgttacca gcattttgtgg gaagaggcca	1560
ggggatgtggc tcaacatcctt gcagtgcaca ggatggcccc tcaacaaaga atcacacggc	1620
ccacaatgtc aatagcgatca cagttggaa aacctgtctt agaccaagggtt tgctttctg	1680

ccgtgtgcct cacccaccc ccactcggt tccctaattcc catctccaaa ggttggcagc	1740
agaccggccc aggctcggtgg aagttcagat catgatcccc tccagctctg caggagacaa	1800
gacctgtctc ccagcattcc tcattgttcc cgggtctgca gagggcgtga gctatgctgc	1860
aggcggctg cccctgaag cctgcgcacc cctccagc tcctcaagtc ttctctgtcg	1920
agtcaccttc gaaccggagg ctgtgagctg gctgtcgta ccacac	1966
<210> 12	
<211> 1198	
<212> DNA	
<213> Artificial Sequence	
<220><223> Fragment 4	
<400> 12	
tggtgcctct gctgtcatga caacagcaca ctacgtcagt agtgctccct gggcactgag	60
ctccctctt gccccggagaa gacagtaatg aaaaatgaca agcatgaggc agaggggaag	120
atcacgcttg ggtggtgcag gagcatggag gtgctttaa tgctctcaat gagaaagggt	180
taacggctt ggttgagga atagctgagt cagaggtgg gttcccca ctccccacc	240
ccacccctt caccattagg gaccttcttg cttgtctt gctactctgc tctgggttgt	300
cattgtgaaa agccgcacc aaccatgccca gtggcagcca gacgaggaca cagcctggct	360
ctgggtccca gcaggaaagg caatcccaga aaggcaggtt cagggactgg agtccctgtgg	420
gtgcctttta agcaaagatt atcaccaggc aggctaaact tagcaaccgg cttagcta	480
gaagggcagg gggctgggtt caggttatgc tggccagca aagaggcccg gatccccct	540
cccatgcacc tgcgtatggg ccaaggccac cccacccac ccccttcctt acaagtgttc	600
agcacccctcc cttccacac tcacaaacct gcccctctgc ctcctacca gaagaatgga	660
tccctgtgg gagggggcag gggacctgtt cccaccgtgt gccaaagacc tctttccca	720
cttttcctt ctcttgact caccctgccc tcaatatccc cggcgcagc cagtgaaagg	780
gagtccctgg ctctggctc gcctgcacgt cccagggcgg ggagggactt cggccctcac	840
gtcccgctct tcgccccagg ctggatggaa tgaaaggcac actgtctctc tccctaggca	900
gcacagccca caggtttcca ggagtgcctt tgtgggaggc ctctggccccc ccaccagcca	960
tcctgtctc cgcctggggc cccagccgg agagagccgc tggcgcacac agggccggga	1020
ttgtctgccc taattatcg gtccaggcta caggctgca ggacatcgac accttcgtg	1080
cagaaacctc cccctcccc tcaagccgc tcccggaccc cttccctctc cagggcccca	1140
gtgcccagtg cccagtgcacc agcccaggcc tcggtcccag agatgccagg agccatcg	1198

<210> 13
<211> 259
<212> DNA
<213> Artificial Sequence
<220><223> Fragment 8
<400> 13
cctctggccc cccaccagcc atccgtcct ccgcctgggg ccccagcccg gagagagccg 60
ctgggtgcaca cagggccggg attgtctgcc ctaattatca ggtccaggct acagggctgc 120
aggacatcgt gaccttcgt gcagaaaacct cccctcccc ctaagccgc ctcccgagcc 180
tccttcctct ccaggccccc agtgcggcagt gcccagtgcc cagcccaggc ctcggtccca 240
gagatgccag gagccatcg 259

<210> 14
<211> 1947
<212> DNA
<213> Artificial Sequence
<220><223> Fragment 9
<400> 14
cgaagcttga gattgtccaa gcaggtagcc agagagcgcc atcagccaag aaaccatcca 60
ctggtagcgt aaggcagctg tgcgggcag accagactgg gccctcccc cctgcagtga 120
tttggttttt cttctttttt aaatcacgtt ttctgcctt ttcttaggttc taggttaccag 180
cctctggctt ctacagcctc agacaatgac tttgtcacac cagagcccg ccgtactacc 240
cgtcgccatc caaacaccca gcagcgagct tccaaaaaga aacccaaagt tgtttctca 300
agtgtatgagt ccagtgagga aggtatgatg ctccgcctg ttccggccg agaaggcaca 360

cagcttagggt gcagagggt gcgttccata ggacctgctg cggggccctg agtgttagatg 420
ctctgccccca ctgccgcaga agggcctctc ctgtacagct tggattttat ttctctgtg 480
cggtgtggaa ttgtctcaact tggtctctga tatctatttt ttcaccatct ttgtgactca 540
gctttttctt attcctttaa ttcttgcat agatcttca gcagagatga cagaagacga 600
gacacccaag aaaacaactc ccattctcag agcatcggt cgcaggcaca gatccttagga 660
agtctgttcc tgtcctccct gtgcaggta tcctgttaggg tgacctggaa ttcaaccgg 720
aggctgtgag ctggctgtcg tgaccacact ggtgcctctg ctgtcatgac aacagcacac 780

taatgtcagta gtgctccctg ggcactgagc tcccttttgc cggggagaag acagtaatga 840
aaaatgacaa gcatgaggca gagggaaaga tcacgcttgg tggtgcagg agcatggagg 900

tgctcttaat gctctcaatg agaaagggtt aacggtcctg gttgcagga tagctgagtc	960
agaggtgggg ctccctccac tcccccaccc cacccttac accattaggg accttcttgc	1020
cttgctcttgc ctactctgct ctgggtggtc atttgaaaa gcccgcacca accatgccag	1080
tggcagccag acgaggacac agcctggctc tgggtccag caggaaggc aatcccagaa	1140
aggcagggtc agggactgga gtcctgtggg tgcttttaa gcaaagatta tcaccaggca	1200

ggctaaactt agcaaccggc ttttagctag aagggcaggg ggctgggtgc aggttatgct	1260
gggcaccaa agaggcccg gatccccctc ccatgcacct gctgatggc caaggccacc	1320
ccacccacc cccttcctta caagtgtca gcaccctccc atcccacact cacaacactg	1380
gcccctgccc ctccattcag aagaatggat cccctgtggg agggggcagg ggaccctgtc	1440
ccaccgtgtg cccaagacct ctttccac ttttccctc ttcttgactc accctgcct	1500
caatatcccc cgcgccagcc agtcaaaggg agtccctggc tctggctcg cctgcacgtc	1560
ccagggcggg gagggacttc cgccctcagc tcccgtctt cgccccaggc tggatgaaat	1620

gaaaggcaca ctgtctctt ccctaggcag cacageccac aggtttccag gagtgcctt	1680
tggggaggcc tctggggccc caccagccat cctgtctcc gcctggggcc ccagcccgga	1740
gagagccgct ggtgcacaca gggccggat tgtctgcct aattatcagg tccaggctac	1800
agggctgcag gacatcgtga cttccgtgc agaaacctcc ccctccctt caagccgcct	1860
cccgagcctc ctccctctcc aggccccag tgcccagtgc ccagtgccta gcccaggcct	1920
cggtcccaga gatgccagga gccatcg	1947

<210> 15

<211> 1436

<212> DNA

<213> Artificial Sequence

<220><223> Fragment 11

<400> 15

atctatttt tcaccatctt tgtgacttag cttttctta ttccttaat tctttgcata	60
gatctttcag cagagatgac agaagacgag acacccaaga aaacaactcc catttcaga	120
gcatcggtc gcaggcacag atcctaggaa gtctgttcc gtctccctg tgcagggtat	180
cctgttaggt gacctgaaat tgcgacccgaa ggctgtgagc tggctgtcgt gaccacactg	240
gtgcctctgc tgtcatgaca acagcacact acgtcagtag tgcctccgg gcactgagct	300
ccctcttgc gggagaaga cagtaatgaa aaatgacaag catgaggcag agggaaagat	360

cacgcttggg tggtgccagga gcatggaggt gctcttaatg ctctcaatga gaaagggtta	420
--	-----

acggtcctgg ttgcaggaat agctgagtca gaggtggggc ttcctccact ccccccaccc 480
accccttca ccattaggga ccttcttgcc ttgctctgc tactctgtc tgggtggtca 540
ttgtgaaaag cccgcaccaa ccatgccagt ggcagccaga cgaggacaca gcctggctct 600
gggtcccagc aggaaaggca atcccagaaa ggcagggtca gggactggag tcctgtgggt 660
gcttttaag caaagattat caccaggcag gctaaactta gcaaccggct tttagctaga 720
agggcagggg gctgggtgtca gggtatgctg ggcgcagcaaa gagggccggg atccccctcc 780

catgcacccgt ctgatgggcc aaggccaccc caccggaccc ctttccttac aagtgttcag	840
caccctccca tccccacactc acaaaccctgg ccctctgccc tcctaccaga agaatggatc	900
ccctgtggga gggggcaggg gacctgttcc caccgtgtgc ccaagaccc ttttccact	960
ttttccctct tcttgactca ccctggccctc aatatcccc ggcgcagcca gtgaaaggga	1020
gtccccggct cctggctcgcc ctgcacgtcc cagggcgggg agggacttcc gcctcacgt	1080
cccgctcttc gccccaggtt ggatggaatg aaaggcacac tgtctcttc cctaggcagc	1140
acagccccaca gggttccagg agtgccttg tggggggccct ctggggccccc accagccatc	1200

ctgtcctccg cctggggccc cagcccgag agagccgctg gtgcacacag ggccgggatt 1260
gtctgcccta attatcaggt ccaggctaca gggctgcagg acatcgtgac cttccgtgca 1320
gaaacctccc cctccccctc aagccgcctc ccgagcctcc ttccctctcca ggccccca 1380
ggcccaagtgcc cagtgcggcag cccaggcctc ggtcccagag atgccaggag ccatcg 1436

<210> 16

<211> 1177

<212> DNA

<213> Artificial Sequence

<220><223> Fragment 17

<400> 16

atctatttt tcaccatctt tgtgactcag cttttctta ttccctttaat tctttgcata 60

gatctttcag cagagatgac agaagacgag acacccaaga aaacaactcc cattctcaga 120
gcatcggtc gcaggcacag atccttagaa gtcgttctt gtcctccctg tgcagggtat 180
cctcttaggtt gacctggaa tcgaaccgga ggctgtgagc tggctgtcgt gaccacactg 240
gtgcctctgc tgtcatgaca acagcacact acgtcagtag tgctccctgg gcactgagct 300
ccctcttgc ggggagaaga cagtaatgaa aaatgacaag catgaggcag aggggaagat 360
cacgcttggg tggtcagga gcatggaggt gctcttaatg ctctcaatga gaaagggtta 420
acggttctgg ttgcaggaaat agctgagtca gaggtggggc ttcttccact ccccccccccc 480

accccttca ccattaggga cttcttgcc ttgctttgc tactctgctc tgggtggta	540
ttgtgaaaag cccgcaccaa ccatgccagt ggcagccaga cgaggacaca gcctggctc	600
gggtcccagc aggaaaggca atcccagaaa ggcagggta gggactggag tcctgtgggt	660
gcttttaag caaagattat caccaggcg gctaaactta gcaaccggct tttagctaga	720
agggcagggg gctgggtca ggttatgctg ggccagcaaa gaggccggg atccccctcc	780
catgcacctg ctgatggcc aaggccaccc cacccaccc cttccttac aagtgtttag	840
caccctccca tcccacactc acaaacctgg ccctctgccc tcttaccaga agaatggatc	900
ccctgtggga gggggcaggg gacctgttcc caccgtgtc ccaagacctc ttttccact	960
ttttccctct tcctgactca ccctgcctc aatatcccc ggccagcca gtgaaaggaa	1020
gtccctggct cctggctcgc ctgcacgtcc cagggcgggg agggacttcc gcccctacgt	1080
cccgctttc gccccaggct ggtatgaaatg aaaggcacac tgtctctc cctaggcagc	1140
acagccccaca gtttccagg agtgccttg tggagg	1177
<210> 17	
<211> 18106	
<212> DNA	
<213> Homo sapiens	
<400> 17	
gtgagcagca caggacactt caatgcctgt tgggtctgg gctggctaag acatctgccc	60
gccctggca gcatacggt cttgcagtca cttcccgcc ctccttatcc ccagctgggt	120
tgcaacaaa ttgccaggt gacctaagac cagatcttg tctccagttc tttttttatt	180
actccaaaaa cacaacaaa gcagcatctc atccaattct tgtttggttt ttttaatag	240
tttttatttt tcagagcagt tttaggtca aagcaaaatt gaggcagaaag tacagggagt	300
tccttctac ccctgcccc tacacatcac agcctcccc acctcaaca tcctgcacca	360
gggtggcaca ttgttacag ctgaacctac acttacacat catctctaa agtcatggtt	420
taccttggag ttactgcac gtaatgacat gtacccacca ttgcagttac atacagaaga	480
gtttcactgc ttacaaatc ccctgcactc cacattttt tccctctc cccacaaccc	540
ctgattttt tactgttgcc atcactttgt ctttccaga atgtatcatt ggaatgatcc	600
ggatggagc ctctcacct tggcttctta gtaatgtgct ttaaggcct ccatgtctc	660
catggcttg ttctttta atcagaagta actgtttca ggcctgtct gaatctctt	720
ttctccctcc aggtataat agatgaattt gaggcagaagc ttggccctg tcataccaga	780
ggtttggatg gaatcaagga gtttgagatt ggccaaaggcag gtggccagag agcccatca	840

gccaagaaac catccactgg tacgttaaggc agccctgtgcg ggcgagacca gactggcc 900

tccccctcctg cagtgatttgc ttcttcttc ttttttaat cacgtttcc tgcctttct 960
aggttctagg taccagccctc tggcttctac agcctcagac aatgactttg tcacaccaga 1020
gccccggcgt actaccgcgtc ggcattccaaa caccgcgcg cgagcttcca aaaagaaacc 1080
caaagttgtc ttctcaagtgc atgagtccag tgaggaaggt atgatgttcc cgcctgttcc 1140
cgccgcgagaa ggcacacagc tagggtgcag agggctgggtt tccataggac ctgctgcggg 1200
ggcctgagtg tagatgtctt gccccactgc cgcagaaggg cctctctgtt acagcttgaa 1260
tttttatttct tctgtgcgggt gtgggattgtt ctcacttgc ttctgtatc tatttttca 1320

ccatctttgt gactcagctt ttcttatttc cttaattct ttgcataatc cttcagcag 1380
agatgacaga agacgagaca cccaaagaaaa caactccat tctcagagca tcggctcgca 1440
ggcacagatc ctaggaagtc tggtctgtc ctccctgtgc agggatcct gttagggtgac 1500
ctggaatttcg aattctgtttt ccctgtaaa atattgtctt gtctctttt tttaaaaaaa 1560
aaaaaggccg ggcactgtgg ctcacgcgtg taatcccagc actttgcgtt accaaggcgg 1620
gtggataacc tgaggttaggg agttcgagac cagcctgacc aacatggaga aacccatct 1680
ctactaaaaaa taaaaaattta gcccggcgtt ttggcgtgcg cctgtatcc cagctactca 1740

agaggctgag gcaggagaat cgcctgaacc cagaggcggg ggttgtatgt agccgaaatc 1800
acaccattgc actccagctt gggcaacaat agcgaacctc catctcaat tttaaaaaaa 1860
atgcctacac gctctttaaa atgcaaggctt ttctcttaaa ttagcctaac tgaactgcgt 1920
ttagctgcctt caacttggaa atatatgtttt gccaatctcc ttgtttctat atgaataat 1980
gtttttatataatcttttagac atttttctt aagcttgcgtt ttgtttctat tttcacat 2040
gcccagttc atgcagcaga gagagggta tcagtcaga gagagatgag tgagcccaga 2100
gtccttagggc ctgtcccggtt atggcagatg agcttcgtgc cccgtactg ccacctttcc 2160

cctctcaacc tctggaccct gcacagtgcac cagacagccct ctctggggag aattatgcag 2220
tgccttaggtt ccagatcagt gcttctgaac cggggggcaat ttgttgcctt agaggacatc 2280
tgacaaacacc tggggcgtt tttgttgcata tagcctatag gggaaagatg ctaccagcat 2340
ttgtggaaag aggccaggaa tggcgtcaat ctcctgcag tgcacaggat ggccttcaa 2400
caaagaatca cacggccac aatgtcaata ggttcacatg tgagaaaaacc tgctctagac 2460
caagggttgc ttctgcgtt gtgcctcacc ccaccccccac tcgtgttccc taatccatc 2520
tccaaagggtt ggcagcagac cggcccgaggc tcgtggaaatg tcaatgtatc atccctcc 2580

gctctgcagg agacaagacc tgtctccag cattcctcat tttccggg tctgcagagg	2640
gcgtgagcta tgcgcaggc gggctcccc ctgaagcctg cgccccctc tccagctct	2700
caagtcttct ctgcgtgagtc accttcgaac cggaggctgt gagctggctg tcgtgaccac	2760
actggggct ctgcgtcat gacaacagca cactacgtca gtgtgtctcc ctggcactg	2820
agctccctct ttgcggggag aagacagtaa tgaaaaatga caagcatgag gcagagggga	2880
agatcacgct tgggtggtgc aggagcatgg aggtgtctt aatgtctca atgagaaagg	2940
gttaacggtc ctggttgcag gaatagctga gtcagaggtg gggcttcctc cactccccca	3000
ccccacccct ttaccatttta gggaccttct tgccttgctc ttgtactct gctctgggtg	3060
gtcatgtga aaagccgc acaaccatgc cagtgccagc cagacgagga cacagccgg	3120
ctctgggtcc cagcaggaaa ggcaatccca gaaaggcagg gttaggact ggagtctgt	3180
gggtgtttt taagcaaaga ttatcaccag gcaggtaaa cttagcaacc ggcttttagc	3240
tagaaggca gggggctgtgt gtcaggttat gctggccag caaagaggcc cggatcccc	3300
ctcccatgca cctgctgatg ggccaaggcc accccacccc acccccttcc ttacaagtgt	3360
tcagcacccct cccatccac actcacaaac ctggccctct gcccctctac cagaagaatg	3420
gatccccgtt gggagggggc aggggacctg ttcccaccgt gtgcacaaga ccttttcc	3480
cacttttcc ctcttcttga ctcaccctgc cctcaatatac cccggcgca gccagtgaaa	3540
gggagtcctt ggctcctggc tgcctgcac gtcccaggc ggggaggagttccgcctc	3600
acgtccccgtt ctgcggggca ggctggatgg aatgaaaggc acactgtctc tctccctagg	3660
cagcacagcc cacaggttcc caggagtgcc tttgtggag gcctctggc cccaccagc	3720
catcctgtcc tccgcctggg gcccagccc ggagagagcc gctggcgcac acagggccgg	3780
gattgtctgc cctaattatac aggtccaggc tacagggctg caggacatcg tgaccttccg	3840
tgcagaaacc tccccctccc cctcaagccg cctcccgagc ctccttcctc tccaggcccc	3900
cagtggccag tgcccagtgc ccagcccagg cctcggtccc agagatgcca ggagccagga	3960
gatggggagg gggaaatggg ggctggaaag gaaccacggg ccccccggc agggccatgg	4020
gcccctctta gccccttgc tgagcgtcc ggtgtacta ccgcagagcc tcgaggagaa	4080
gttcccaac ttccccgcct ctcagcctt gaaagaaaga aaggggaggg ggcaggccgc	4140
gtgcagccgc gagcggtgct gggctccggc tccaaattccc catctcagtc gttcccaag	4200
tcctcctgtt tcatccaagc gtgttaagggt ccccgctt gactccctag tgtcctgctg	4260
cccacagtcc agtctggga accagcacccg atcacccccc atcggccaa tctcagtc	4320
ttccccctta cgtcgccggc cacacgctcg gtgcgtcccc agttgaacca ggcggctcg	4380
gaaaaaaaaa agcggggaga aagtagggcc cggctactag cggtttacg ggcgcacgt	4440

gctcaggcct caagacccctg ggctgggact ggctgagcct ggccggaggc ggggtccag	4500
tcaccgcctg ccgcgcgc cccggttct ataaatttag cccgcagcct cccgcgttcgc	4560
tctctgtcc tcctgttcga cagtcagccg catttcttt tgcgtcgcca ggtgaagacg	4620
ggcggagaga aaccgggag gctagggacg gccigaaggc ggcagggcgc ggcgcagcc	4680
ggatgtgttc gcgcgcgtgc ggggtggcc cgggcggcct ccgcattgca gggcggcgc	4740
gaggacgtga tgcggcgcgg gctgggcattt gaggccttgtt gggggagggg aggggaggcg	4800
tgtgtgtcgcc cggggccac taggcgcgtca ctgttcttc cctccgcgc gccgagccac	4860
atcgctcaga caccatgggg aaggtgaagg tcggagtcaa cgggtgagtt cgcgggtggc	4920
tggggggccc tgggtgcga ccgcggccaa accgcgtcta cgacccttgc gggctccgg	4980
tcttgcagt cgtatgggg cagggtagct gttcccgca aggagagctc aaggtcagcg	5040
ctcgacccctg gcgagcccc gcacccagcc tggcgccccc tgtgcagctc cgcccttgcg	5100
gcgcacatcg cccggggcct cttccctta gtcccccaga acaggaggc cctactcccg	5160
cccgagatcc cgacccggac ccctagggtgg gggacgcattt cttcccttc ggcgtctgcg	5220
gggtcacgt tcgcagagga gcccctcccc cacggcctcc ggacccgcag gccccggat	5280
gctagtgcgc agcgggtgca tccctgtccg gatgctgcgc ctgcggtaga gggccgcga	5340
tgttgcacc gggaggaaa tgaatggca gccgttagga aagcctgccc gtgactaacc	5400
ctgcgtctt gcctcgatgg gtggagtgcg gtgtggcggg gaagtcaggt ggagcgaggc	5460
tagctggccc gatttctctt ccgggtgtatg cttttccat attattctctt ggttaatcaa	5520
agaagtgggt ttatggaggt cctttgtgtt cccctccccg cagagggtgtg gtggctgtgg	5580
catggtgcca agccggaga agctgagtca tggtagttt gaaaaggaca tttccaccgc	5640
aaaatggccc ctctgggtt gggcccttcc tgcagcgcgc gtcacccatca cggccccc	5700
cttccctgc cagcctagcg ttgacccgac cccaaaggcc aggctgtaaa tgtcaccggg	5760
aggatgggt gtctggcgc ctcgggaac ctgccttct ccccatccg tttccggaa	5820
accagatctc ccacccgacc ctggcttgag gttaatata gtcgtgacc tttctgtac	5880
tggggccctg ggctgggct ctctccatc ctttctcccc acacacatgc acttacctgt	5940
gctcccaactc ctgatttctg gaaaagagct aggaaggaca ggcaacttgg caaatcaaag	6000
ccctgggact agggggtaa aatacagctt cccctttcc caccggcccc agtctctgtc	6060
ccttttgtttag gaggactta gagaagggtt gggcttgcct tgcgtgttta attttctgacc	6120
tttactctgtt cccttgagt ttgatgtgc tgagtgtaca agcggtttct ccctaaagg	6180
tgcagctgag ctggcagca gcaagcattc ctggggcgc atagtgggtt ggtgaatacc	6240
atgtacaaag ctgtgccc gactgtgggtt ggcagtgcgc cacatggccc cttctctgg	6300

aaggcctcg tatgactggg ggtgttggc agccctggag cttcagttt cagccatgcc	6360
ttaagccagg ccagcctggc agggaaagctc aaggagata aaattcaacc tcttgggccc	6420
tcctgggggt aaggagatgc tgcattcgcc ctcttaatgg ggaggtggcc tagggctgct	6480
cacatattct ggaggagcct cccctctca tgccttctt cctttgtct ctttagatttg	6540
gtcgatattgg gcgcctggc accaggcgtc cttaactc tggtaaagtg gatatttttg	6600
ccatcaatga ccccttcatt gacctaact acatggtgag tgctacatgg tgagcccaa	6660
agctgggtgtg ggaggagcca cctggctgat gggcagcccc ttccataccct cacgtattcc	6720
cccaggttta catgttccaa tatgattcca cccatggcaa attccatggc accgtcaagg	6780
ctgagaacgg gaagcttgtc atcaatggaa atcccatcac catcttccag gagtgagtgg	6840
aagacagaat ggaagaaatg tgctttgggg aggcaactag gatggtgtgg ctcccttgg	6900
tataatggtaa ctttgtgtcc ctcaaatatgg tcctgtcccc atctcccccc cacccccata	6960
ggcgagatcc ctccaaaatc aagtggggcg atgctggcgc tgagtacgtc gtggagtcca	7020
ctggcgtctt caccaccatg gagaaggcgt gggtgagtgc aggagggccc gcgggagggg	7080
aagctgactc agccctgcaa aggcaggacc cgggttcata actgtctgct tctctgtgt	7140
aggctcattt gcagggggga gccaaaaggg tcatcatctc tgcccccctct gctgatgccc	7200
ccatgttcgt catgggtgtg aaccatgaga agtatgacaa cagcctcaag atcatcaggt	7260
gaggaaggca gggccctgg agaagcggcc agccctggcac cctatggaca cgctccccgt	7320
acttgcgcccg cgctccctt ttctttgcag caatgcctcc tgcaccacca actgcttagc	7380
acccctggcc aaggcatcc atgacaactt tggatcggt gaaggactca tggatgaga	7440
gctgggaat gggactgagg ctcccacctt tctcatccaa gactggctcc tccctgcccgg	7500
ggctgcgtgc aaccctgggg ttgggggttc tggggactgg ctccatccata atttcccttc	7560
aagggtggga gggaggtaga ggggtgtatgt ggggagtgacg ctgcaggccc tcactcctt	7620
tgcagaccac agtccatgcc atcaactgcca cccagaagac tgtggatggc ccctccggga	7680
aactgtggcg tgaatggccgc ggggctctcc agaacatcat ccctgcctct actggcgctg	7740
ccaaggctgt gggcaagggtc atccctgagc tgaacggaa gctcactggc atggccttcc	7800
gtgtccccac tgccaacgtg tcagtgggtt acctgacccgt ccgtctagaa aaacctgcca	7860
aatatgtga catcaagaag gtggtaaagc aggcgtcgga gggcccccctc aaggcatcc	7920
tgggctacac tgagcaccag gtggctctt ctgacttcaa cagcgacacc cactccctca	7980
ccttgacgc tggggctggc attgcctca acgaccactt tgtcaagctc atttccctgg	8040

atgtggctgg gcccagagac tggctttaa aaagtgcagg gtcggcgcc ctctgggtgc	8100
tggctcagaa aaaggcccct gacaactttt ttcatcttct aggtatgaca acgaatttgg	8160
ctacagcaac aggggtggtgg acctcatggc ccacatggcc tccaaggagt aagacccctg	8220
gaccaccagg cccagcaaga gcacaagagg aagagagaga ccctcactgc tggggagtcc	8280
ctgccacact cagtccccca ccacactgaa tctccctcc tcacagttgc catgttagacc	8340
ccttgaagag gggaggggcc tagggagccg cacctgtca tgtaccatca ataaagtacc	8400
ctgtgctcaa ccagttactt gtcctgtctt attctaggtt ctggggcaga ggggaggaa	8460
gctggcttg tgtcaaggtg agacattttt gctggggagg gacctggat gttccctca	8520
gactgagggt agggctcca aacagccttg ctgttgcga gaaccatttg ctcccgtc	8580
agacgtctt ggtgtacag gaagctggca ccactacttc agagaacaag gcctttct	8640
ctcctcgctc cagtcctagg ctatctgtt ttggccaaac atggagaag ctattcttg	8700
ggcagccca gggaggctga caggtggagg aagttagggc tcgcactgg ctctgacgt	8760
gactggtag tggagcttag cctggagctg agctgcagcg ggcaattcca gcttggcctc	8820
cgcagctgtg aggttttagt cacgtgtctt attgttttctt gtccctcgatgttatctg	8880
aggacatcgt gcccagcccc taaggtttt aagcaggatt catctaggta aaccaagtac	8940
ctaaaaccat gcccaaggcg gtaaggacta tataatgttt aaaaatcggt aaaaatgccc	9000
acctcgata gtttgagga agatgtactg agatgtgtca gggtgactta tttccatcat	9060
cgtccttagg ggaacttggg tagggcaag gcgtgttagt gggaccttagg tccagacccc	9120
tggctctgcc actgaacggc tcagttgtt tggcagttt ctccggcc tcactttgtca	9180
cgtgtgttta cctagtggag aaaaaatgtt atacctcggt agagcgcgca cgcctgttaac	9240
cccagcaattt tgggaggcca aggtgggtgtt atcacctgag gttaggagtt tgagaccaggc	9300
ctggccaaca tggtaactt ccgtcttac taaaattaca aaaaatcagcc aggcttcatg	9360
gcacatgcct atagtcccg ctacaggcat gtcggcgac gagaatcgct tgccatccccgg	9420
aggcagaggc tcgactgagc tgagaccaca ccactgcact ccagcttagg caacagagta	9480
ttagacttcca tctcaaaaaaaa aaaaaatgtt cctacctcgat agtcaactt agttaatatt	9540
aggaagtgtt tgagacagtgc acaccaaaatgtt gcacaataaa tactcgccag tttcattttt	9600
attaaagaat ccatttgaat gtcagcttca cacagcttcc tataccggagg cattgttaac	9660
cgcacatccc cagtttcttcc aggctttcc aagaatcagg gacactgttag cctgttggtc	9720
tcagtgtatg acagacacgg aggaaggcaca tcttttagctg atacttaaac agagaccctg	9780
agcgcacata cacccgcgca cacatgcatg gagcttccacc ttctctgtca ttctgcgttg	9840
accaggagag caagagctcc cacctccctt caaaaactg tgcccatccc gggcactaag	9900

gcctcttaa agcacggcac ctccacgagg gagggccaca gccacataca ctccacctgg	9960
caggtggaca gcgtgagcac gtggaccata gcagggacaa ggtgccccgg ccagcccaa	10020
cgccctctgc cgctgacagg gacagaagcc ctctccagct gcgtgtgctg cagaggccat	10080
gcgttagcctc cagtcattt ctattccact ccagtgcctg ggccagttag caccagtgt	10140
gaagacagtg agctggctcc ggacaacagg gatggggaa aggtcccaca ttcacattcc	10200
tgatacgtgg acaaggtgag gggccgcaat cgctctggca gcattttaaa gatggggaa	10260
tagcagacac ccacgcgtga aggcaggaga gccccaaactg tggggaaat gcccccaaa	10320
tggtagggcc aagcttagct ccagacaccc cagagccctg gagaagccaa gactgaggaa	10380
gaaaggctga gggaggagcg ccccgatccc cagggccgg cctggcag agctgcagct	10440
gatgttcccc tctgtgcagc cccaccctct gcctcgctga gctccctgct gcggggcct	10500
cgggtgcaag ggggaggcag gtctctatct catggagctg tcagatgaga catcgatc	10560
ggagtccctca gcctcgcttg gcggcgccgg cgggtcgcta agcgggaccg cagtggaa	10620
aggagacttt ctagaaaaaa acaccagttt tcaaccttgg ggcaggcagg aatcttggaa	10680
acggacggca ctccctctcc tgctgcctca ccctctggca gcccgtgaga agtaccggaa	10740
gcgagggccgg gcccggggta tggcgaggga gcccggggta ctgaactctc tccaaaccca	10800
ccctgacagg gaaatggcc cccctgtgt ctggggact cagaggctga ggtcaggcat	10860
gatggctcac gcctgtattt ccagcacttt gggaggcaga ggcgggtggta tcacgacgtc	10920
aggagttcaa ggcaaggctg gccaacatgg tgaaacccca tctctactaa aaatgcaaaa	10980
attagccagg tgtggtggcg ggcgccttgg aggctgaggc aagataatca cttgaacctg	11040
ggaggcggag gttgttgtga gccaaaaaaa aaaaaaaaaaag aaatagctga agtcacagta	11100
ggagagaagc tgctgagccct ccagcacccct gactctaggc cttggcttt atgtctatct	11160
gcagtatttt tgtgatttt aaaaattcac ttcttgttgc cgggttaact tacacagggt	11220
caaattgcaca aatcatggcc ctgacttttag ataaaaatct gccccacaa ccctctgtt	11280
ccttgccagt tttaactg cctctaaccg ggggaaccac cagagctgtt ggccttggta	11340
gtttcagcc ctcccgatcat gaatggacat agtcatcca actgccaagg gagagagctg	11400
tgggtctggg ccagcccccac caggttaactc cccaaaggca gccccacagc aagatgtgac	11460
ccagtcattt cctgagggtc tctggggctg ttttccaaacc tttctccccc ctgtgtcccc	11520
ctggaaaggcc ccatgcccac gggaggccgc taccggctc cagactgggt gatggccgg	11580
cggcagggtct ccatctgcac gtccaggccg cgcttcatgc tgccatctc catgtactcg	11640
tgcaggtgcc gtttcatgtc gtttctggcc gtggccact ccagctgtgg tggggcaggc	11700
aggccatcg gggtaacag gtggcgatca cagcgctct gttggccccc ccaggaggcc	11760

aacacgcca a gacgatggc tggccggg gcccaggcag ccattactga aggctcagat	11820
tttaaaataa accagaccaa ggcagaaggc agcgaataca ccattgttc ctctgcttt	11880
cctggaaatc ttgcctatg gaggaatgca aaggtaactgc tgcaggctgt gttctttaa	11940
gactgataag agtggtgagg aaagttagtgc tggcggat ggagcagatc cacaggctcc	12000
cagggccaca caggttcata tggggcccg ctcaccctc atccctgta ctgcgttgt	12060
cctggctgt cccctactg cactgtgac tccgttaagg gttagcgaca gttttcctc	12120
atcctggcag acccagaccc cgaaggaaat taactcttc tggaccgt ggcataacaag	12180
cctggcta ac tggtaaagg ggcataatcc tcttaccatt ctcccttat cactttattt	12240
tctcacctaa aagccattt cagacactaa aaggaagact cccatattgg aggtgccatt	12300
gattctgaga cctgtcctga tttcagaaat gtcacacgtg gaaaacacgg ggctcagaat	12360
caaagaaatc cggctccct cccggggccc attctgcagc ccacactctg aggcaggcgg	12420
agaggcaggg cccaggcagc acaaactcag tgccaggac agtgcactgc ctctcttgc	12480
tttgggtaca aagggcaggg gcccgaagg gaagcaaaat gaaacttctt ggagatgacg	12540
ttcgtgactg cctgagtgaa gagaacgggg aggataccag ctgcataat gtggattgga	12600
cagagagggt gggacttag agaaagaagg aaacaggtt cagggccaga cagcagagat	12660
ggtcaggaca gtgacatttgggtgagatg gcccgtggct gggtgagaag ccctagtgg	12720
gacagcttcc ctggcaggct tggatccaca gctggggccc cacagtagtgc ggaggagcag	12780
gacacggta ctgtcaacc cagggggaaa atacagtgc accagccctt gttcagatca	12840
gctgcctaca ccccaatttc tccacattgg aagaaccgga gtgcaggag ctgggttctg	12900
gaaccatcac tcagggagac tggagaagaag gaaatgaacc aaatatattt tactgaaata	12960
caaggctagt actgccacca gtagtctctg catatgtcta tggccact taattcaca	13020
atttacagca ggaacaaggt caggaggtga acgtgcacgt caggagatgt ctttgacaag	13080
ccacattgaa aactgcaggc tcgctgtgtg aacagacttg ggggtgggac gggaggagac	13140
agagagaggt gaggagcggt ggggaagaac tgcctccgc ccaccactt gggccttggc	13200
tttacattgc tccatcttag tgacccttag agatcaaagg tcacaaagat aaaggcacct	13260
tggtaactcc tgggtcagc aagtcttcc aacggttccc ctaatgcaga agtccccaaa	13320
ccccaggcca cggaccagta cctgtctgtg gcctgttagg aaccaggcca aatagcagga	13380
ggtgaggcgtt gggcaagtga gtgaagcttc atctgtatcc acagtcgtc cccatctactg	13440
gcattacagc ctgagctcca ctcctgtca gatccggc ggcatttagat tctcacagga	13500

tcctgaaccc taccttacgt gaactgtgca tgcgaggat ctaggctgt tgctcctat	13560
gagaatctaa tgcctgatga tctgtcacta tctccatca cccccagatg ggaccgtcta	13620
gttgcggaaa acaagttcat ggctcccact gattctacat tatggtgaga tgtataatta	13680
tttcattata tattataatg taataataat agaaataaag tgacacaggcc gggtgcggtg	13740
gctcatgcct gtaatcccag cactttggga ggccgaggtg gtggatcacc tgaggctggg	13800
agttcgagac cagcttgacc aacatggaga aaccccatct ctactaaaaa tacaaaatta	13860
gccaggcgtg gtggcacatg cccgtaatcc cagctacttg ggaggctgag gcaggagaat	13920
tgcttgagcc cgggaggcgg aggctgcagt gagccaagat tgccattg cactccagcc	13980
tggcaacaa gagcaaaact ctgtctctaa aaataataat aataattttt aaaaatagta	14040
ataaagtaca caataaatgt aatgtttgaa tcatctcaa atcattccc ccgggtcagt	14100
ggaaaaactg tctccacaa accagtccct ggtaccaaaa aggttggga ctgctgccta	14160
aagtcatgc actgggagag aggaggagtc agatgattt aaaaactttt tttttttttt	14220
tttcctgaga tggaaatcttgc tctgtcacc caggctgtgg tgcaatggca caatctcagc	14280
tcactgcagc ctccacactcc cagattccag caattctcct gcctcagcct cccaaatgc	14340
tgggattaca ggtgcccacc acaatgcctg gctactttt gtattttag tagagacggg	14400
gtttcacatg ttggtgccca ggctggctc acactcctga cctcaagtga tctgcccgcg	14460
ttggcctccc aaagtgcgtt gattacagggc gagagccacc acgcccagcc aaaaatagct	14520
taaaaggcat gattccattt aagatattttt tttcgttctt tcagatcatg ttttacactg	14580
ggcacttgcgatgtt gtcaggaa gacttagggag agatggcagc ccgagatctc ttctcacgaa	14640
gatggccgca caaaggcccc tagttgggtt aaggccagttt gaaaacacaa accttaaatc	14700
caccatccaa aagagggaga caagaatcag gtcaggcagg aggaggaaac agtgagccgc	14760
ggtagacggg agggtgggaa gtactggttt cccttgcctt gtcttagct agcagcacaa	14820
atccaggaaa cctcatctac ctggccaaga aaaccagggtt tgatgcagga gcaggacagt	14880
ttgatgtgtt ttgggagaat gaggatataat gaaaataatt ttcgcacttc ttctcaaac	14940
acctccaaatg tgacagcaat aagaataaaag tcaaggccgg ggcgggtggc tcacgcctgt	15000
aatcccagca ctggaaagg ccgaggcggg tggatcacga ggtcaggaga tcgagaccat	15060
cctggctaac atggtaaaac cccgtctcta ctaaaaaata caaaaatttta gccggcgta	15120
gtggcaggcg cctgttagtcc cagctacttggg ggaggctgag gcaggagaat ggcgtgaacc	15180
tgggaggtgg agcttgcagt gagctgagat cgccgcactg cactccagcc tggcgcacac	15240
agcgagacta cgtctaaaaa aaataataat ataaataat aatgaaaag aataaaatgtca	15300
ctctctcattt atccactgtt ttcactttcc acggcttcag gtacccaagg tgaaccttga	15360

tctgaaaata ttaaatgaaa tattccagaa ataaacaatt cctaagttt aaattgcacg 15420
ctgttctgag tggcctgata aaatctcaact ctctctgcgg ggcgcggtgg ctcacgccta 15480
taatcccagc actttgggag gccgagggtgg gcggatcatg aggtcaggag atcgagatca 15540
tcctggctaa cacggtgaaa ccccatctct actaaaaaat acaaaggaaa tttagccggc 15600

gtgatggcgg gcgcctgttag tcccagctac tcaggaggct gaggcaggag aatggcgtga 15660
acccaggagg tggagcttgc agtgagccga gttcaagcca ctgcactcca gcctgggtga 15720
cacagcgaga ctccgtctcg gaaaaaaaaaaa aaaaaacctc actctatcca gaccaggac 15780
tgaaccctcc ctctgtccca aggatccaca ctgtctatac tgcctgccta ttagttactt 15840
aggagccatc tcagtgtatgattgactgt cgtggatgg cagtgcgggt gttcaaggaa 15900
cccttatttt tacttcataa tggcccaaa gcacaagagt agtgatgtg gcaatttggaa 15960
tatgtgaaag agaagccaaa atatgtccc tttaagtgaa aaggtaaaag tttcaactt 16020

aataagaaaa agaatccgat gctgagggtt gtaagctcta cagtaagaat gaatttctaa 16080
tccgtgaat tggaaagaag gaaaaagaaa ttgtgccag tttgtgtc acacctcgaa 16140
ctgcaaaagg taccgccaca gcgcacgcta agtgcttagt taagatagaa aaggcattga 16200
atgtgtgggt gggagacatg aacagaaaca cgttccagtt gacagcaact ggacccaagc 16260
tatcaggagg tcagtgtat cttgggttc aggcattcac tgggttctt ggaacacatg 16320
ccctgaagat aaggggggac tacgtatgaa tctccatgt gggctgaag agaggctgag 16380
acgcagagaa ggctccatgc ctctccaaa gtcagagcgt cgggaaagga ccacctcacc 16440

aggccaggag agctcacggg ggtcccggtcc atccctctc ccacagcaac agcaaccc 16500
attctgtgtt tccacaaatc cttttccctt tgggttattttttagacgg agtcttgctc 16560
tgtcgcccgag gctggagtgc agtggtgcga tctcgctca ctgcaagctc cgcctccgg 16620
gttacgcacca ttctctgccc tcagccccc gagtagctgg gactacaggt gcctgccacc 16680
acggccagct aattttttttttttagtattttttagacggatggatggatggatggatggatgg 16740
caggatggtc tcgatctctt gacctgtatgaa tctcccgcc tggccctccc aaagtgtgg 16800
gattacagggc gtgagccacc ggcggccggcc tccacaaatc cttttgcag ctccctaaagc 16860

ctagagttt agctaaagca ccagctccac ttgcctctag gtacctctag aaggtaacttc 16920
aaggcttggg aaaaagcatg gagatttctc tggatctca aaggctgagt ggaagaaaca 16980
ggaaagtccc agtccatctc tgagtgcctt ggcactgata gtgttaataa taataaaaca 17040
gtccccacta tgctacactc gctatgagcc agacactgccc ctaagctcta tatatgcatt 17100
atttcatctg atcttctaaa caaccctata ggtgaggat tattaaattc ctcatttac 17160
aaataaaaag gcccggcaca gtggctcatg cctgtaatcc caacacttttggggccgag 17220

ttaggaggac tgcttgaatc cagaagttt agggcagcct ggaaaacata gtgacaccct	17280
gttctctgta aaaaaatca aaaaattagc cggcggttgt agcacacgcc tgtggtcca	17340
gctacttggg aggctgaggt aggaggattt ctggatcca gaagatcacg gctgcagtgg	17400
gcctggagtg acagagcaag accctgtctc aaaaaaagaa agaaagaaag aaagaagaa	17460
aaagaaacag gttagggct ggacacatgg ctcatgtctc taattctagt acttggag	17520
gccgaggcag gaggatcaact taagcccaag agttaagac ctgcctggaa aacacaggaa	17580
gacaccatct ctacaaaaaa atacaaataa aaattagctg ggtatggtgg cacacacctg	17640
tggcccccagc tactcagaag gctgggtgg gaggatcaact ccagcctagg caacagagt	17700
aaacagagtg agacttgtc taaaaaaga aaggacagaa agacaaggaa gagagggagg	17760
gaggataga gggggagggg aagagagaga gagagagaaa ggaaggaagg aaggaaggaa	17820
agaaaggaaa gaaaggaaa gaaaggagg ggctcagagg gaacaagtaa cttattacaa	17880
gttcagggtg gtgttaagta acagagctga gatttgaacc gatatctgtc tgttccaga	17940
gcctgaggc ttcagagccc accaggagga ctatagctga gagatcaact ctcctccaa	18000
gaccttggtc tctagggtcc tgcagcaggg atgaggccac tgccctgca gcccactca	18060
aaaccctctg ggacaccacg cccctggct ccgctctgc cttac	18106
<210> 18	
<211> 16312	
<212> DNA	
<213> Mus sp.	
<400> 18	
gtgagttggtc ctgttagtt ccggtgcaat aaacttcatc tgcctcccta agtgcctgt	60
gactcctgtc tgaactttct gccttgctag cctcaactaca gctatggaac cagagtgact	120
aagaccagat cttaaattcct ggtccctttt ttaaaggcca ttcatttctt tttttgttaa	180
catgcattgg tattttgcct gtgttatatgt ttgcattgaag gtgcctgtc cccttggaaat	240
tggagttaga ctgttgggg ccctgagaac tgaacacagg tctttgtaa gagcagccag	300
tgctcctaact cactatgcca cttcccagc cccagtgaag gccttgatgat agggcttgc	360
<210> 420	
<211> 480	
<212> 540	
<213> 600	
<400> 660	
tatgttagacc cagctgcctc agactaacag agattcccgc gcctctgtct cctggtagta	420
gaggatgtgcc gcacatgtttt gcactggact gcttctttag agtagttca tgttaaagc	480
aaaagtcaat gaaatggact cccatatgtt ctgtccacaca cccgcattgc ctccacccca	540
tgacattctt gtcctggat acgtggctt ggtatggaa acgtgattga cagatcagca	600
tccaaagtcc actcaactgga cacagtggca catacttacc acttgggtca tactggcat	660

tttcaggact ttgaggtcac cttacccccc acccctgtgg tggcatcagt tagttggaat	720
cacagttca gattggtttc tttagtttc tctctttca tggctatttc ttttaagca	780
gggagataat gttccagcc ctgtcttaa cttctgtcc tccaggctat aatagatgaa	840
tttgagcaga agttcgtgc ctgtcacacc agaggcatgg atgaaataga ggagttgaa	900
actggccagg gaggcagcca gcgagccctg tctgccaaga aaccctccgc tggtatgaga	960
gggatgcaag cccacatggg gtcctatga gtggcccca ccctttttt gttccttgc	1020
gttttgcataatggccct tgccttttt ttctagtc aaggctacag	1080
cctctgactt ctgtagactc agacaatgac ttgtcacac ctaagccgac acgtacaaa	1140
cctggtcgtc cacaactca gcagcgaag aagtccaga gaaagccaa agttgtttc	1200
ttgagtgacg agtccagtga ggacggatg atacaccctc ccaacagatc agctgcgt	1260
gtggggctt ccagtggggc ttgccttagct gtggattcat tcttcctggc catgtttcc	1320
tcactatgca gctcggctga cttcacagt gtgcgtgtg cgtgtgttt tgtgtgtgt	1380
tttgtgtgtg tgtgtgcaca cgtgtgcgc tgccacgcg cgcacatgttt tactttgaga	1440
tccatttgc cccatctta tgacttagct ttgttccta taatttttgc aatagaactt	1500
tcagcagaga tgacagaaga agagacaccc aagagaacca ccccatccg cagagcatct	1560
ggcgaagac acaggctcta ggagctgct ttgcctactc ccgtccctgc cgtcagggt	1620
atcttactgg atgatcttgg attcagtgcc cccccccct tggaaatgt ctttttgtct	1680
gtggtttaa catggaaatg gccttgcattt tgaggcttt ctcttagagc cctgcgtgc	1740
ctttccctt ctccgagct gtccaaattt ctgtttctta atgaaataat gttttatag	1800
acttttaagc atccgtcct ggctggctg tcccatcctt agccctgagc tgtgtgttt	1860
tcttcatggc ccctgtggtg cctcatcctt gcacaactga aactctgtgt tgcaatatt	1920
tgcgtttctt aggttacgtt gaacagttctt cagcgtctg gggacttttgcacccagaa	1980
aacattcggtt aatacctaga gacttatttc tgtttgcctt gtggggcagg gggctgcagt	2040
gaaagccta agatgaaggg cacattcatg cacattgtgc cccataaacac agcatgtcag	2100
gagtagaccaa cttgagaaat gctgctctag acctaagggtt tggttgcctt gtgcgc	2160
ccccatcccc cgaagactgg caaatgactc agcctggcta atggcagaga ttatgaaccc	2220
cactcagttc tataagggtac aaggccatc gccccagac cagtccttca ctacactaag	2280
gggtctgcag gaggtccagc tgtaagccat gctgtgtgcc caccatcccc aagccctaa	2340
tctgctgagt cacttggagc aggaggcaca actgctgttt ctgctgccat gacaacagca	2400

cagtgcgtc gaagtgcgtcc ctacctcttgc tgaggggac agtaatgata gatgacaggc	2460
gtgaggcaaa agggaaagggtg cacctctgaa acactaggtt agggctggc tatggaaggg	2520
ctcttgggtt aggaaaagggtt ttaacgcgtgt gcatagagcc tcggtaggaa ggcagaaaag	2580
ggactctgaa tctgcgtc ctctctccc caccactgctg ctggggccac tcccaccctt	2640
cccaccctgt tcatctggcc tgctccccgc tcgtggaaag cctgggcacg ggccacacag	2700
ctgtcagtct aacacagcct tggctctggg gccccaaagga ggtaggcaaa tctggaaggg	2760
caaggagcca agactagatt tgggtgcag cccagcgtgc tccctgcctt ttaagcaag	2820
gttatcacca gcccagctaa acttagcaat aggcttgcata gctagatgag cagggggctg	2880
gtttcatgtt gcactggcct agcaaagagg cgccagagtc cccctgcct gcacctgcta	2940
cagtgcgtccc aagcccttcc acccttccca gcctccctac cttcgagggt gaagaatccc	3000
ggtctcgaaa agacaaaata aaaccaaaga atgccttttc tcccttccac ttgtggcaag	3060
aggctagggg ctccctgtc ctggctcagg ggtataatat ttccctcttgc ttgttctccc	3120
ctcaactgatc tcaccctgtg tccacgaggg cagccaaaaa aggagagtcc ctggctccag	3180
ggccagtctg cactccctg gatggggagg gactccgccc ctacgtccc aactctccac	3240
cctggggctt cagtgggtga aaggggcagt gtctcttagc ctgggggggg cagctctcag	3300
gttccgagga gggatactat aggaggcctc tgagccctt ccaattcaac ccttaagagg	3360
gatgcgtccc ttaccccggtt gtcccagtt aggttcatca ggttaactca ggagagtgtt	3420
tgttaagtctc aattatcagg tttccaggtt gcagggcatc ctgacccatg gcgttagcaat	3480
ctccctttca agcgtcttcc tggccacccc ctggccctca cagcgcggta cacaagccag	3540
gcctctggcc aaagacagaa gccaggaggg gggggggggc agataggaa atggggctcc	3600
tgagggtcttcc cgccgggaggaa aaagtgccttcc caccctggca ttttcttcca ctctttttt	3660
tttcgggggggg ggggggttgcgtt gtgtcaactac cgaagaacaa cgaggagaag atcctcaact	3720
tttccgcagc ctttcaata atggggagag gttcgatgtat gcagtggcag ggagacccac	3780
actttctccat ttccctgtt ctcccatttt actcggttgcac actccaaaggatc ctcgttgcac	3840
cctggatgttc ctgggtgcac actccaaaggatc ctcgttgcac ttaagttcat agtctgtatt	3900
ccctgagtcc tttccctggatcc accatcaccc ggtcacccatcc tgagcggggc aatctcgttgc	3960
ccctccccc tatcgttgcac gagccacac gcttgggtgcg tgcacatttc aaaaatgagg	4020
cgggtccaaa gagagggagg aggaaaatggc agagaggccc agtactcgc ggcttacgg	4080
gtgcacgttgcac ctcaggcctc tgcgttgcac agttaggact ggataagcag ggcggggaggc	4140

ggggcgcgcg tcatcagctc ccccccacca tccgggttcc tataaatacg gactgcagcc	4200
ctccctggtg ctctctgctc ctccctgttc cagagacggc cgcatcttct tgtgcagtgc	4260
caggtgaaaa tcgcggagtg ggccgcagga ggccggggac agtcgaaac tggaaagggg	4320
agtgggcact gtacgggtct agggatgctg gtgcgaagtg tgcagccgg acccaggctc	4380
cgcattgcag ggggggatg atggaggacg tgcgtggcg cacggcggga atggaggcgg	4440
ggtgggggag gggactgcct ggtgtcctc gggccacgct aatctcattt tcttcctcg	4500
cagcctcgctc ccgtagacaa aatggtgaag gtcgggtgta acgggtgagt tccagggcgg	4560
ggccctgctc cgttgccctac gcaggtcttg ctgaccggg ggctctgcag tactgtgggg	4620
aggtggatga ggtggccgaa gcgcacaagg agacctaag gtacgcgtc ggacctggcg	4680
atggctcgca ctgcggccc ccagctgctg cacctctggt aactccgcct ttgcgggat	4740
gagcggcccg gagtcctaag tattaggaac aacccacgc gccggtcag acccatcccg	4800
taatccccag tccggggctt cttctttac tttcgcgcctc tgaggagtca cgtgccagga	4860
gggaagcccc tccccatct ccccttcct cgggtcggt ggcattggcgt tgtgaggtgc	4920
atacccttgc gcatcatctc ccaggcttgg gcttcctta gggtaactgg ccgcgcct	4980
gttgcaaacg ggaaggaaat gaatgaaccc cggttatgaa atttgctta ggccttcctt	5040
cttcctagct tgtgactaac ctcattcctc tcggctgggt ggagtgtcct ttatcctgta	5100
ggccaggtga tgcaggcctt ccgtgcgtc gagagagttc tacctcacaa tctgtctcac	5160
cttattagcc taaaagccc ttgagccta ttgtcctcg gcataatgct tattctagat	5220
tattctgta aaatcaaagc ggacttacag aggtccgctt gaccccaaa cccagaggt	5280
agttatggcg tagtgcagag ccgtggatg gggagctgag tcatgggtt tctgaaaaga	5340
aattttcac cacaatgg ctccgttagt agcagccct tccatccct gcacttccca	5400
tcacagcctc gcactgaccc aggcctata ggccaggatg taaaggcat taaggaggatt	5460
gggtgtccct ggcctcaga atcctgcct tctcccggtt ccatccctca gaaaccagat	5520
ctctccactc cgcctgtatc tgaggtaaa tttagccgtg tgaccttct ggatctgggg	5580
tctgagcggg ctctccaccc tgctccccct acacacatct gttgctccgg ctctcatttt	5640
tgcggagaa gaacaggtgt ttgcgaac agccctggaa tttaggttgg aaaccccca	5700
catgtttct cagtcttcc ctttagttcg agggacttgg aggacacagg tggccccc	5760
ctgtgctgct cacgctgacc tttagccttgc cccttgagc ttgctgatga atgagttcac	5820
aggctgtcccc tgtccagggg gtgttagcctg aagtccagcc atgctggaac aaacttccca	5880
ggccatgagt gatgggggtg atgtgccaag ctgtaccca gactggggca aactgccact	5940
tcttaagaga cttagaatga ctggaggag gtttgcgtgg caagcaatca cctctggac	6000

aggaaagaaa cctccactt ataaccgtgc tataaaagcc ctgccaggcc tcggctgctc	6060
aaagaatata aaatttagatc tctttggact ttcttagggtg ggaacagtct atattgggt	6120
tgtacatcca agcattcaac tagctttatt aaaggaaatt ctgaacaaa catgaacttc	6180
ctgatgcata catttcaat gtactgtgc tgcataaagg ctgttacccit tttgtggta	6240
cgtgcatacg tcatggctgc aggttctcca cacctatggt gcaacagtat tccactctga	6300
agaacatgag atagcctggg gctcaactaca gacccatgag gagttctgtat ctcagctccc	6360
ctgtttcttg tcttcagat ttggccgtat tggcgccctg gtcaccaggctg cttccatgg	6420
cagtgccaaa gtggagattt ttgccatcaa cgacccttc attgacacta actacatgg	6480
ctacatgttc cagttatgact ccactcacgg caaatcaac ggcacagtca aggccgagaa	6540
tggaaagctt gtcataacg ggaagccat caccatctt caggagcgag accccactaa	6600
catcaaatgg ggtgaggccg gttgtgagta tttgtggag tctactgggt ttttaccac	6660
catggagaag gccggggtaa gtggccggaa agctgaaggt gacgggcacc cttgatatgg	6720
tgcaacctga aaaccaagaa ctgagtctga aatcaacttc ttcccttaa acagggccac	6780
ttgaagggtg gagccaaaag ggtcatcatc tccgcccctt ctggcgatgc ccccatgttt	6840
gtgatgggtg tgaaccacga gaaatatgac aactcaactca agattgtcag caatgcatcc	6900
tgcaccacca actgtttagc cccctggcc aaggtcatcc atgacaactt tggcattgtg	6960
gaagggctca tggatgttag gcagtgggaa gacagctcat gcatttctta tcttaccctg	7020
ccatgagtgg accctctt ttaggtgtcc cttttggta gaggggtgcc gtgcaggacc	7080
tcactcattt ccccggtt ttcttagacca cagtcacatgc catcaactgcc acccagaaga	7140
ctgtggatgg cccctctgga aagctgtggc gtgatggccg tggggctgcc cagaacatca	7200
tccctgcac cactgggtct gccaaggctg tggcaaggt catccagag ctgaacggga	7260
agctcaactgg catggcccttc cgtttccata ccccaatgt gtccgtcgt gatctgacgt	7320
gccgcctgga gaaacctgta tttatgggaa gagctggct tttctgtt gtgacagtga	7380
cttggacaa ggtatgtcat tttgggttt gttttatca ggcaagat gatgacatca	7440
agaaggttgtt gaagcaggca tctgaggcc cactgaaggg catctggc tacactgagg	7500
accaggttgtt ctctcgac ttcaacagca actcccactc ttccacccatc gatgccgggg	7560
ctggcattgc tctcaatgac aactttgtca agctcatttc ctggatgtg gggatggaa	7620
acctgacttt taagagcaac tgggggttt gtgcctctg gtggctagct cagaaaagaa	7680
acccaaacta acagttgtcc caatttggc taggtatgac aatgaatacg gctacagcaa	7740
caggggttgtt gacccatgg cctacatgg ctccaaggag taagaaaccc tggaccaccc	7800
accccaacaa ggacactgag caagagaggc cctatccaa ctggcccccc aacactgagc	7860

atctccctca caattccat cccagacccc cataataaca ggagggcct agggagccct	7920
ccctactctc ttgaataccca tcaataaaagt tcgctgcacc cacctctcct tggttttaat	7980
gcgggtgtgg gggagggagg gtgtgtatgg ggaggtgaag caggctaat caactttgt	8040
gcatgtatct ttattggctc tgacatccag ctgggtgccg gaagtccagg gctacattct	8100
atccaagtcc cagccaggc cgcctgtcct gaacagctga tatgtctggc aatgcggagg	8160
gtggggcttg acaatggac agggaaatgca gtctgagact cacttggc acagcactga	8220
cttccagagc tcactgctct tgagaaagat cggtgtcctt ttatctgcaa acaaaggta	8280
attgtaccca gccctcctgc agacagacgt aaaatttagat aaaattggaa gggcgccag	8340
tacagtaaaa cctagtggtt tccatgaaga ggggctgttag cactggtcga tgctttggtt	8400
cgtcgctcac acagtgactc agtgcttggg aagtattcca caccgtcatt gcttttaag	8460
gttccattcc aatagttcaa catagcgtct cctgggaggt cgtgactgac tccctaggtc	8520
agggacatct taaccttttgg ttttgtcag tgacagccat ggcggcatac cctcagctgc	8580
tcactggaga aacaggcccc gagcaccagt aacctgcagg ccacacccat gtgcggcag	8640
cagcttacc tcctcagccg cgacctggc ccagcacagg acttccaaa ggcgagcctg	8700
cacggccatg ttactgcaa gaggcaagtg aacagctggg agcatgtgga cagagcaccc	8760
cacaccctcc aaagccttac tagggagcaa agccctatcc caaggtacat gatgaatctg	8820
cttcacacga gtgcctaggc aagtttagcac cagtacaaaa tccaagatac ccaaggaata	8880
caggcaggc cctgtatcac attccttaca aagtggaaaa gggaggctca caagtgtcct	8940
ggcagcattc cgaagatgga gaagtagcag atacccacat gtgaggaagg caggaaggag	9000
tgtcccaatg tcatggcgt tgtacaggga ggggtgggct gagcctagct ccagacaccc	9060
cagaacctgg ggagggacag cccagggcca ggcttcatgc agtggcacag aagacccaca	9120
cccctgcaga cccacacccc tgccccagtg agctctggg tgccagacac tggttttat	9180
ctcatggagc tatcatggatga gacgtctcga tcggagtctt cagtctcact cggcggtggc	9240
ggcggtcgtaa agcgccggac cgcagtgaaa gcaggagact ttctagaaaa aaacaccagc	9300
tgtcaccttg gaccaagcag gaaggctgca gacttctaag ctgttttttcatggccttc	9360
tgaggtacag ggggtgccag gcaaggcgt gggagggaaag aagacacttc cctttcaac	9420
ttgatggga gcagtgcctc tcttatcttggggctctg agaggctcag agatcagaaaa	9480
agtgcgaagt cacagcacgg agatgcccc tgagccacgc tgccagacac tggtttgtc	9540
cagcagtaat ctggatggg ggtggggct gctggacccc aaaactcctg ggtgcaagca	9600

atccctttgc ctccaaagtg gctgggtgg agccacagat ctcttgaag catgcacact	9660
atactcagct tggatgttatt tggagattt gggtgggttt tttgttgtct ttgagacagg	9720
gcctgtacgc attccgtctt ggcctgaaac tctatacaga ggtgaggctg gcctggact	9780
tggagttaccc gcctgttgtct tctatgcatt tgaatgacag atatgcgtta gcacacciga	9840
cttactgttag tttcttggat tttgttgtt gttgttattt ggtttgttggacaaattt	9900
ttgctgttgtt cgtccaggac agcctaaaac tcatgacaat cttctgcct ctgccttcca	9960
agactgaaac tgcagacatg taccactctg cctgggttgc agtaaatttt aaaaaaaattt	10020
ttaatgtgt gtgtgttgtt gtatgtgtat ctgtctgtgt gtctgtgtgt gtatgtgtgt	10080
gtctgttgtt gtatctgtct gtgtgtgttt atctgtgtc ctgtatctgt ctgtgttgtat	10140
gtctgttgtt atctgtgtgt ctgtgtctgt ctgtgtgtgtt atctgtgtgt gtgtgttgtt	10200
atctgtgtgt gtgtgtgtgt gtctgactgt ctgcctgtct gtctctgtgt gtgtgtgtgt	10260
gtgtacacag gcccccacac aggctaggca agtcctctac cactaagcta tactccgaaa	10320
ccccagattt actttacac aaggtcaaacc acataaaaccc gtgatttggtaa agatcctg	10380
tccccgtatc ttggccctt gcctgggtta ggttctaaga cactcagagt tggcacactg	10440
gtcagttca tctgtatggcc aaggcaggga gaggaagctt cattctggc cacacctaca	10500
ggcgcactccc aaggacagct ggacataaag atgtgacctg gtccctgccc acagaccatc	10560
ggggcctatt ccagccttcc tccctggag agagacccag cccgctcgg gcaggccctt	10620
accggcccccc ggactgtgtg atgagccggc ggcaggcttc catctgcacg tccaggcccc	10680
gcttcatgtct gcacatctcc atgtactcat gcagggtgtcg gttcatgtcg ttcttgctg	10740
tggccagctc cagctgcagt aaggtgggtt cagggaggac cacggtagc aaacccaggc	10800
aaggccctgg gggcaagca gtcaggaga caagtaagct acaactgcct gtttccccca	10860
ggatatggag gagggtggg gcaactggcg atctgacttc aggcctcac tgtggctagg	10920
caagggtctca ggacagtgtct cccgcctcg ccctgtccctg ccaatccact cctactgcca	10980
cccaagaatc aatcagctga tgtctgagtt ctctgtttaa acagcaagag cacacaacac	11040
agtaaggatg gtgcgggtcg gtgcgggtcg gggaggcacgc atccctccag cgcttactcg	11100
gctcaggcag cagggtcccg aaggcaaaacc taaatataaa accggcttc aaaaagggtt	11160
aaggaggcca caccaggcgg ctcagtgat aaggctctg cgcatctggc aatgtgaatt	11220
tgttctctaa cttccacaca tgcaccgtgg catgtggaga gagagagaca cacacacact	11280
aataacaata ataataaaat ttagtatttt gggggctggt gagatggctc agtggtaag	11340
agcacccgac tgccttcca aaggtccaga gttcaaattcc cagcaaccac atggtggtctc	11400
acaaccatcc gtaacgagat ctgacgccct cttctgggtgt gtctgaagac agctacagtg	11460

tacttacata taataaataa ataaatcta aaaaaaaaaa aaaaaattta gtatttgtt	11520
tgttgttt tcgaaacagg gtctctccat gtagccctgg ctgtcctgga actcactcta	11580
tagaccaggc tggcctgaa ctcagaaatg cgcctgcctc tgcctcccga gtgtggat	11640
taaagggtgt caccaccact cccggcaagg cagcagttct taacctgtga gtctcaatcc	11700
ccttgggggt tgactatcg acactctgca tgtctgtatt tatgttataa ttcgtaacag	11760
tagcaaatt agttatgaag cagcaatgaa atagttatg cctggggcca ccacagcgtg	11820
aagcactgtg ttaagggtca cagcgtctgg aaggctgagc accactgctt taaggacaga	11880
cagggccact catctgtcca tctttgccac tgagcctccc ctcaactgaa gctaattcac	11940
tgtgaaatat tccaaaggct cgctgcaatg gagggaacct ggctgcttct gtttatacg	12000
tttcattcca catcatcttc tcccttataat cctgctcgcc cacataagag agagacctc	12060
ctcctgagtg attacagggt aaatccccaa caaaagaaaat gggggggggg agtctctgta	12120
tcaagaaaag tctcccttcc agagctgact ctgcagccag gttgaaggc agcaaaaaaa	12180
gtaggctgag gcggaaaagt atctaggcca gagacaaact gggcccttc ttttttg	12240
atacaaaggc caggggccag atagggagaa tcagactgctg gtggggagat ttgcttgct	12300
gcctgagggc aaagacaagg aggccgagtg gccttgaagc atgtggatca gctgtggaa	12360
gtcacccgaa acaaattgtga actggaccgg gtgggagttgg gccccaggag cagacaaagg	12420
agacagccag gacactggca tttgggtca tctgctgaac agctgggtgg taagccacag	12480
aggagatggc ttgaaaggca tgcttggtcc cacagcagac tgaccattt gtgagaaaaa	12540
catgacactg aggatcagag ctcagctgcc accccccaca gacaaatgca cacagcatc	12600
caggtcagag atcagctgcc acccatgtac atcaacacgc acacacacac acactattac	12660
aggtcagaga gcagctgtca cccatacata ccaacacaca cacacactat tacaggtcag	12720
agagcagctg tcacccatac ataccaacac acacacacac acacacacac acacactatt	12780
acaggtcaga gaggcagctgt cacccattca taccacacaca cacacacaca cacacacact	12840
attacaggtc gagaggcagct gtcacccata aatacaagca taccctcac tccctccc	12900
accccacctc cattatagtc tctcaaagca taaggcagta ctaccagcca tgtctctaag	12960
tgtcatatgt caccctgatt tacacactga aggggaagat gcacattcta gaagatgcct	13020
tcaacaaatc accccaaat ttcagactca gtgtggaga ggcagggtta gaggtagata	13080
ggagcgtatga agggggggggg ccatggaaga aaactccacc caccactctg catcaccatt	13140
caataatggg ggctggtcta tctctatgcc aataaccctg agatcaaggg tcatggagat	13200
aaaggtctta cagatgctcc agggagaccc aggagctact ctgctaccaa cgctaccaag	13260
tgagcatcct ttggcttct cctcaactca ttctcccagg agagaacaga agtcatacaa	13320

ccttaggaaca gcttcaaaca ttttccaaat tttttggga gggcaggggg agagtagagg	13380
tttgaagcag ggcttctctg tgcagccctg gccatcctgg aagttagttct gttccccagg	13440
ctgtcctcga actcacagag atcctcctgc ctgtttccc actaccagga ctaaagggtgg	13500
caccaccaca ctccattta taggtactgg tatcaaacct agggcttcac gcatgaaggt	13560
aggcaggcat tctaccatga gagcacagtc ccagccagg gtgtcttct ctgtcctacc	13620
acatcatgtt tacacttagt tatttgtggc ctaaggatg aaccacagaa gggcagcagc	13680
tcctgatcac tgtgctaaga ctccctacctg gctgacggta gggaaaaaaag gaccaagcct	13740
taaatctacc gtctgtaaga ggtggacaca acccagggag gggacagga aggaggaagg	13800
agcagtaaga ggagcggaca ggaggagatg ggagatcatt tgcgtccagg caaagcttaa	13860
ctgggtggcag aacactggct ggggtggaaat tctgggttaa aaaccaaag cactgggttt	13920
cctcagagta gcacggccat ctgacatgtg ctggccactg cagaagagtt acttccaaag	13980
ttcccttca aatgcaaacc agagtgcaccc ccatggcac gggaatgcac cacaggc	14040
gcagggagcc caatggcatt agtctcctgc ctcatccctg catcagtgtc taagacatag	14100
agccctgttc actgtccaaat ggccgacccccc ttccctgcagt ccctgcctgc agtccagcag	14160
gacaatctaa accccagccc caccaccatt catacccgat gccccaaaggc tcaaaaaaca	14220
aaagagggag acacccctccag gagctccaag gctgagccaa ggagacagga gtcctctgg	14280
gtgctacggc acttagggtg ttcccagtaa taaatgcgc aactaccgc tgtcaacctg	14340
ccagatcctg tccaaagagc cacgcgcgc tcactctatc tgcgtccatctt aataaccctg	14400
cctgcctacc agccgtcgg gagagattat gaacatctt attggacaaa taaagagaca	14460
ggcccacccct tagcaggta gaaaccatct ctaaccctaa caggctgagg caggaggatt	14520
gcctcaagtt tgactggcta cacagaaatc tggctgggt gggggcttag gatggctcag	14580
tagttaagag catgccctgt tcttgagag gacccaaatg tattttccag cgcccacatc	14640
aggttcacaa acacccgtgta gcgcgcgtc ctggaaacct gatgccacc tctgaccc	14700
aaggagagag ggggggggggg ggaggggaagg aaggagggag ggagggaggg agggagggag	14760
ggagggaggg agagagagag aaagagagag agagagaaag agagaggaga gatggggcgc	14820
tggatctgcg aggaaggcag gcaggcaaga tgaatgtgt tgagggctca caacaaaaga	14880
ataaaaaagga atgaaggagc tagctgagga acacccatc agtggaggc ttactgcaca	14940
gcatgaagac ctgagtttga tcttagaacc catctgaaaa ataagctagg gctgaagaga	15000
tggctcagca gttaaagagca ctgtgtttgc tcttgcattt cccagcaacc acatggtagc	15060

tcacaacctg ggacctgatg ccctttctg gcatgcagg gtacatacag atagaacact 15120
 catatattaa ataaataaat aaataaataa atcttaaaaa aataaaataa actagccata 15180
 gtagaaacaca cttgcaatcc caaagctaag aatgtacaga cagacagaca gacagacaga 15240
 cagacaaaag atgcctaggc ctccctggcc agctcgccct actcagcaag tttcagacca 15300
 gtgaaagacc ctgtctcaat taaaaaaaaa aaacaaaaaa cccttaacca cactcaaaga 15360
 acactaccca gtttgcctt ctggcctcca cacttgtatg tgtacacatg tgcgctgca 15420
 cgcacacaca cacacacaca cacacacaca cacagaggaa aggaggaaga 15480

ccaaaggaag gataaaaagc tggtcagtca aagctggctt aaagctcctg ccgtaatcct 15540
 atcgcacagg aggctggc agaagcacca ggagctttag gccaacctag actacaagac 15600
 aagatcttga aaagaaacag agaggcaggg gcaggaagga cactcacaag tgcgagagag 15660
 cttgcagttt tgatgcacac ttttaatccc aggagaggca cagtaacttc tccagagtga 15720
 gttccaggac agccaggcgt acagagaagc cctgtctgtc agggggaaaa gaaaaaaaaa 15780
 agaaaggggg ggacagaaaa gaaaaagata gtctggtgg aaatgattca cgggtaaag 15840
 gcatgagcca ccaagtccaa agacctaaac ttgattgctg tcaccccat ggtaaaagag 15900

aacttcctcc cacacgttgc tctctgaccc ctacacatgt gcatggactg cgcatgctct 15960
 ctgacactcc atccatccat ccatccatcc atccatccat ccatccatgt tttgtgtgt 16020
 tgtgtatgtg tttgtgtgtg tttgtgtaaa ataaataaaa aggaaggagg gagggacagc 16080
 aaaagatttgc ttccatgtgg tgacagagcc gggatttggaa cccacatcta tgtgcctcca 16140
 gatccctgaag tctcaaggac ctgctacagc taccacaact gagaggatca ctactgacaa 16200
 cagatggaag ctggggccct ggggttcagc gaggaggagg ccagcaacac acctacagcg 16260
 tggctctgaa gcccggacc tacagcgtgg ctccctgaggc cggcaccctt ac 16312

<210> 19
 <211> 11058
 <212> DNA
 <213> Rattus sp.
 <400> 19

gtgagtgac ccggatcact ggcctgttg gttccgggt tagaaacgg aatccgctct 60
 ccctaagtgc tcagagatcc ctgtctggac ttgctagcct cactacagcc gtgcaaccga 120
 agtacttaa ggcggaccc ttttttttt tttttctga gctggggacc gaacccagg 180
 cttgcgtt gctaggcaag tgctctacca ctgagctaaa tcccaaccc cagggcagac 240
 ctttttaaa ggcattcgt ttccctttt atttatttt tagtgcac tggtatgg 300

cctatgtta tgtctgtgaa ggagccagt gccctggagt tggagccaca gacagtttg	360
ggtgctgaga actaaacaca ggtcttaacc actatgccgc cccccccccccc cccagcctca	420
gtaaaggcca tttctactt gcgatagggt ctgtctatgt agacccagct ggcctaaac	480
taacagagat tccctgct ctgtctccag tggctggta gtaaagggtgt accccatgt	540
tccggactag actgtttct tagagtggc tcgtttaa agtaacactg agttccatc	600
tgcctgtccc acaccctcat ggcctctacc ccatcacctt atcctcggtt cgtgatccta	660
caatctacat tgactggat gaaggaatga gattgacaga tcagcatcca aagtccactc	720
actgaatcca gtggcacatt ctggccacta gtgtcataact gggcatttc aggactgtga	780
ggtcaccta ctccccaccc tgggtggta tcagtcattt ggaatcacag tttcaggtt	840
tgtttctta acttgctgta tcttcatgg ctgtttta aataggaga taatgctcc	900
gaccccgctt ttacctttt ctccaggctg taatagatga atttggcag aagttcgtg	960
cctgtcacac cagaggcatg gatggaaatg aggagcttga aactggccaa ggaggcagcc	1020
agcggccct gtctgccaag aaaccctcg ccgggtgtgag aggcaagccc acgctggcc	1080
ccacaggcag tccctcttt tctttttt ctttgcgtt ttgttttgg ttttggaaatc	1140
atgtccctct gcatattcca gtctcaaggc aacagcctct gacttcgtta gactcagaca	1200
atgactttgt cacacctaag cctcgacgta ccaaaccgg tcgtccacaa actcagcaac	1260
gcaagtccca gaggaaagcc aaagttgtct tctcaagtga cgagtccagt gaagatggta	1320
agatacaccc tcccaacagg acaactgtt gtgtgggtc tccagccggg cttgcctggc	1380
tgtggattca ttctccctgg cagtgcttc cccactgtgc agctgtggg ctgactttca	1440
ctctgtgtgt gtgtttgtgg tgggtgtcca gatccgttt gccccatctt tatgactcag	1500
cgttggttcc tataattctt tgaatagaac ttccagcgaa gatgaccgaa gaagagacac	1560
ccaagaaaac caccccatc cgcagagcat ctgcgcaaaag acacagggtcc taggagctgc	1620
ctttgcctat tcccatccct gctgtgcagg gtatcttact gaatgatcga gaattcaacg	1680
cccccccccc cttggaaaa tactgttgc ctccctgttt ttaacatgta aatggcttt	1740
tcgagcaagg ctttctctt agagccctg actgtactt tccaagatcc aatcaccatg	1800
ttttctatgt aataaatgtt ttatataact tttaaggatt ctggccggct ggtctgttcc	1860
attatttaggc agcagatagc agtccagaga gagaacccca ggcacccatg gccatctctg	1920
gaccctcatg gaccctgtgg tgcttcatct ttgcacaact gctgagattt gctgttagaga	1980
gttctgtggc ttcttaggttc acataagcac ttctcagcct ggggtactt tgccacccag	2040

aacacgttg gcaatatcta gagaccagggt gttggagaga tggctcagcg attaggagca	2100
cttactgccc ttccagaggt cctgagttca attcccaagca accacatgggt ggctcacagc	2160
catctgtaat gagatctgggt gccccttcc ggtgtgtcta aagacagcta tataataat	2220
aaataaaatct ttaaaaaaaa aatctagaga cctattttt ctgcccacct tgtggggcag	2280
gggctgtgtt aggcacccag tcagtcagag cctaaggagg aagactcctt catccatagc	2340
actcccaaca cagcatgtcg ggactgtgca actagagggtg ctgctctaga cctgcgggtt	2400
gtttctgtg tgccacaccc catccagaa gactggcaaa tgagttagcc tggctgggtt	2460
cagagatcat gccattagct ctgttggta caaggtccac ctcccaagag tcagtcctcc	2520
actatgctaa ggggtctctc gctcagctgc aggcccaggc tggctatgt gctgccacc	2580
ttcccaagc ccttaatctg ctgagtcact tggagcagga ggcaggactg ttgttccgc	2640
tgtcatgaca acagcacagt gcagctgagc ctctccctac ctttggat gaaagagta	2700
atgacaggcg tgaggcagaa gaagatgcac ctcagaaaca cttagttagg ggtggactgt	2760
ggaagggctc ttgagacctc tggtaggaa aagggttaac tcggtgata gagcctcagt	2820
aggaagacag aaaagggaca gctgagtcg agatgcctgt ctcccaagca ctgcactggg	2880
gccactctca cccttccac cctgttttg gcctgtccc tgctcatgaa aagcctaggc	2940
atggaccaca cagctgtcag tctaaacaca gctttggctc tggggccca aggaggtagg	3000
gcaatcccag aaggcaagg agccaggact agatttgggg tgcagccag cgtgtccct	3060
gcctttaag caaaggat caccaggcca gctaaactta gcaataggct cttagctag	3120
aagtgcaggg ggctggctc aagttgcact gggctagcaa agaagccca gagtccccct	3180
ggccgcacc tgcgtatggc ccgcacccctc caatcaaccc ctccacccct ttctagccct	3240
cctaccctgc aaagcaaaca agtcccttcc tccctccact gtgcaaagag gctagggtt	3300
gcctgttctg gctcggata ctcccttct gctttttcc cttctgtatc tcaccctgt	3360
tccacgaggg cagccaaagg aagagagtcc ttggctctg ggccagtcag caggtcgct	3420
ggatggggag ggacttccgc cctcacgtcc cagctctca ccctgggct gcggtgggt	3480
aaagggacac tgcgtccat cctggcggtt acaactcagg ttccggggag ggaaggcctc	3540
tgagccctt ccaacccaaac cctcaacagg gatgttacc ccggggccccc ggtatagtt	3600
cagcaggtaa actcaggaa gtgttgcact aatcagggtt ccagactgca gggatcccc	3660
gacctatggc gtaggaatct cttttcaag cctctagtc cttccctgtt ttcacatcg	3720
ggtacacaag ccaggactct ggcctaagag atagaagcca ggccgtgggg agagaggaaa	3780
atggggctgc gaaggccccc cacggaggaa aaagtgcctt ccccgccctt ccccgccct	3840
tttctctgc tcctttttc gggcggggg gggcgtgtt gtcactacc gaagaacaac	3900

gagaagatcc tcaacttctc ctaagcctt tcactaatag ggagaagttc gatggggcag	3960
ccttggcag acccacactt ctgctccatt tccctggtc ctgcagctc cagattctcc	4020
cattttattc gggaaagcagc tttctggttt ctgggtcctg gatgtcctg gtgcacactc	4080
caaggactcc tcgtccttaa tccatagtcgtatcc tggaaacctc	4140
catccggtca cttectcgcc gggacaatct cagtccttc cccctctca ggtcgagcc	4200
cacacgcttg gtgcgtgcac atttcaaaaa cgaggcgggt ccaaaaagag ggaggggggg	4260
aatgagagag gcccagctac tcgcggctt acgggtgcac gtagctcagg cctctgcgc	4320
ctttagctgg gactggatga gccgagcggg aggccggcg cgcgtcatca gctccccca	4380
ccatccagtt cctataaata cggactgcag ccctccctgg tgcgtctgc tcctccctgt	4440
tcttagagaca gccgcatctt cttgtgcagt gccaggtgaa aaagaaaaaga aaaaaaagga	4500
ctggccgca ggaggccgga gacgaatgga aatttaggaat ggggggaagg acgctgtacg	4560
ggttaggg cgctggtgcg aggtccggaa gccgagccca ggctccgcattgcagaggat	4620
ggtagaggac gtgatggggc atgcggcggg aatggaggcg ggtggggga gggactggc	4680
cacgctaattc tgactttctt ctcccgacgc ctcgtctcat agacaagatg gtgaaggatcg	4740
gtgtgaacgg gtgagttcca gggcgccggc ccgcgtccgtt gcgcattcccg gtctggctaa	4800
attcgggct gtgtgttagc gtgcggggag gtggataggg tggccgaagt acccaaggag	4860
acctaaggat cagcgcctgg acctggggaa ggctcgact taacgggacc ccagctgttg	4920
ctcccttgt aactccgcct ttgtggga gcggccggaa gtctaaagta gcaggagcac	4980
ccccgacact ggccgcacc cgctcaggat ccgaccgtatccccaggc tggggctct	5040
ttctttactt tcgcgcctcaggaggatcactg tgccagaagg ggagccctc cccatcg	5100
cccttcctca gggtcgggtgg cacgggggtg tgaggtcat accttgcgc atcttctaa	5160
ggcttcactt atggtaactg gcccgcaca tggtcaaaac ggaaaggaaa tgaatgaacc	5220
accgttaaga aatttccctt cggcttctt catttttagc ttgtactaa ctttcttatt	5280
cctctcagtt gggtgagatg tcctttattc tgtaggcac gatgtggatg gcagggttc	5340
cgtgtctca tttagagctt actatacaat ccgtctctcg tcagccctta taggcctcca	5400
gccttatttc ctcggccat aacttatatt ctagatttt ctctggaaat cgaagtgcatt	5460
ttacagaggt ctacttggcc tcccagcccc ggaggtgcgg tagcaatggc gtatgtccgaa	5520
gccccgggggtt ggggtggggaa gctgatgtcat gatggccctg aaaagaaatt ttctgccaca	5580
aaatggctcc ggtggtagca gcccctccc tccaggccctt ccacttccat cccagctcag	5640
cactgaccca aaccctatag gccaggatgt aaaggtcact aagaggattt ggtgtctcg	5700
agcctcagga tcctaccctt ttcccccaacc catcctccag aaaccagatc tccccattcc	5760

ggccctgatct ggggtaaat ttagctgtct gacccttctg tatctgggt ctgagctggg	5820
ctctctgccc tggcatccc tccacacate tgggtgcct gctccgattt tttttttt	5880
tttttttgg ccggaaaga cagggtttt gcaaatgagt cctgggatta gggctggaaa	5940
atcaactggtt ttcttgatct ttccccaga aaggacttg gaggaagcag gtgggctggc	6000
cctgtcctgc tcactctgac cttagcctt gcccttgag ctgctgatga atgagttctc	6060
tcctgtccgg gagttagcc tgaagtcag ccatgctgga accggcttcc caggcctgtg	6120
tggtggggt gatgaatacc atgtgccaag ctgttaccca gaccgggca aaccgccact	6180
tcttaagaga cttaaatga ctgggttgc cgggcaagg ctgtggcct aatcacctct	6240
tggacaggaa aggaacctcc actttatagc tgctataaaa gtcctgccag ccctgggtgg	6300
ctcaaggaat ataaaattag atctcttgg actttctagg gagggaacat tctatatttgc	6360
ggttgcacat ccaagcatc aaccggctt attaaaggaa attctggaca aaacatgaac	6420
ttcctgctgc ttacgttct gatgctgtac tgaatcagcg taagaacttg caccttttgc	6480
tgtatgcgtgt gtacgggct gctgttagat ctccacgccc atggtcagc gatgcattac	6540
tttctgaaga acatgcgttgc cccaggcgt gactacaaac ccaggagggg ttcactgatc	6600
ccaaactaact cgcctatttc ttgcctcaga ttggccgta tcggacgcct gttaccagg	6660
gctgccttct ctgtgacaa agtggacatt gttgcctca acgaccctt cattgaccc	6720
aactacatgg tctacatgtt ccagtatgac tctacccacg gcaagtcaa cggcacagtc	6780
aaggctgaga atgggaagct ggtcatcaac gggaaaccca tcaccatctt ccaggagttac	6840
gtatggaaac atgcacaggg tacttcgagg agatggtaag gggcggcatt aactctgtcc	6900
tgttccccgc ctgttagcga gatccgcta acatcaaatg gggtgatgct ggtgctgagt	6960
atgtcgtgga gtctactggc gtcttaccca ccatggagaa ggctgggta agtggccagg	7020
aagctggagg taagggaaac cttgtatgt ggtcaacct ccaaactgaa gagctgagtc	7080
tggaaatcaac ttctcctgca caggctcacc tgaagggtgg ggccaaaagg gtcatcatct	7140
ccgcccccttc cgctgatgcc cccatgtttg tggatgggtgt gaaccacgag aaatatgaca	7200
actccctcaa gattgtcagc aatgcacccctt gcaccaccaa ctgcttagcc cccctggcca	7260
aggctcatcca tgacaacttt ggcacgtgg aaggctcat ggtatgtagg caatggagac	7320
agctcatgca tgagtggacc tttcttggaa gatgtccctt tgggttagg ggctggccctg	7380
caagacccca cccattgcct ctatgtttc tagaccacag tccatgcct cactgcccact	7440
cagaagactg tggatggccc ctctggaaag ctgtggcgtg atggccgtgg ggcagcccg	7500

aacatcatcc ctgcatccac tggtgctgcc aaggctgtgg gcaaggcat cccagagctg	7560
aacgggaagc tcactggcat ggccttccgt gttcctaccc ccaatgtatc cgttgtggat	7620
ctgacatgcc gcctggagaa acctgtacgt agggggagag ctgggttgt tttgtctatg	7680
gtatggtgtt gacttgggc aagagcagtc atcttgggt tttcccttac aaggcaagt	7740
atgatgacat caagaaggtg gtgaagcagg cggccgaggg cccactaaag ggcattctgg	7800
gctacactga ggaccagggtt gtctccctgtg acttcaacag caactcccat tcttccacct	7860
ttgatgctgg ggctggcatt gctctcaatg acaactttgt gaagctcatt tcctggatg	7920
tggggaccgc aagcctggct cttgagagta actgaaggtt tggtgccctc tggggccag	7980
cttagaaaag aagccaaac taaccgttgt cccaaatctgt tctaggtatg acaatgaata	8040
tggctacagc aacagggtgg tggacctcat ggcctacatg gcctccaagg agtaagaaac	8100
cctggaccac ccagccccc aaggatactg agagcaagag agaggccctc agttgtgag	8160
gagtccccat cccaactcag ccccaacac tgagcatctc cctcacaatt ccatcccaga	8220
ccccataaca acaggagggg cctggggagc cctcccttct ctgaataacc atcaataaaag	8280
ttcgcgtcac cctttcttg tgtgtgaatg cagtgttgtg ggggttatgg gcaggtgaag	8340
gaggctccat tcgggtttgt gcacatatct ttattggctc tgccatccag ctgggtgcta	8400
gaggtccagg gtcacattct aaccgagtgc cagccaggc cgcctgtcca gacacctgat	8460
gtgtctggca gtgtggaggg tggggcttga cactggaca gaaggtgcag tctttgaaca	8520
gcactgactt tcagagcttgc ctgctttga gaaagatctg tgcctttat ctgcaaaca	8580
aggttaattt tacccagctc tcaaagtctc cagatgttaa attagttaa gttggaaaggg	8640
tgcccaatgtc agtaggacta attgtgtcca tgaagggga aggtccctgc tgagctgttc	8700
acctgtttc aggtgtggc acttgaaggt gctttgggtc tggggcagtg actagccagc	8760
gctcactgtt gttggaaat gaaatgaagg tgcattacac ggtgactcga ctcaacgcct	8820
tcatgtttt caaaggccc attcaaattgt cagctctgca tagcctctcc tgggagatcc	8880
tagctcccta ggaccaggga cattgttaacc ttttggttt gtcagtgaca gccatggagg	8940
cattccctca gctggtaact gaggagacac cctgagccacc cctaacccttgc tgcaggccac	9000
acccaagtgt ggcagcagct tcatctcgtt ggcgcacccggcccggacttccca	9060
aaggcacacg catgttcaact gcagtggca ggtggggcagc tagctgtgag catgtgcagt	9120
gtccacgccc tcaaggcctt cctaggagc aaggcttat ccaaaggatc atgcggaaatc	9180
tgtttcacat gagtgccctag gccagtttgc accagtacaa aaccaagga aaacaggcag	9240
gtcccaataac cacattcattt acaaattggaa aaggggaggctt ccaagttgtgc tggcagcatt	9300
ccaaagatgg agaagtagca gataccaca tgtgaggaag gcaggagtgc cccaaatgtgt	9360

tgggagttt acagggggtg gggccgagcc tagctccaga cacccagaa ccctggagag	9420
gacaagccag gcccaggctt catgcgggtgg cacagaagat ggtcctctgc agaccacac	9480
ccctgccccca gagagctctg gggcaccagg ccctggttcc tatctcatgg agctgtcaga	9540
tgagacatct cgtacggagt cctcagtc gcttggcgc ggccggggt cgctaagcgg	9600
gaccgcagt aaagcaggag actttctaga aagaataacc agttgtcacc ttggggcaaa	9660
gcagggaaagc tcctgagact tccaaacttg ctttctctgg ctttatcttt tgaggtacag	9720
ggggatgaca ggcaggacg ggggagggaa gaaaaacttg tttcaactt gatggggaga	9780
acagtttcgc tcttggac tcagagatca gaaaagtctg aagtacacgc atggagatgt	9840
tccctgagcc acactaaggc cctagctct tatccagcaa taattcggtg tgcgggggc	9900
tgtggggact gtaggaggtg agactgggg agtatgggg ttttaggggc tgttcactc	9960
tgttgtcaact gctgacccca aattcctggg tgcaagcaat cctttgcct tcagttgatt	10020
ttctttttt ttttaattta ttttattttt atgagtacac tgtaactgtt ttcagacata	10080
ccagaagagg gcatcagatc ccattacaga tgggtgtgag ccaccatgtg gttgctggaa	10140
attgaactca ggacctctgg aagagcagtc agtgcctta accgctgagc catctctcca	10200
gccccctttg cttcaaaagt ggctgggtg gccacaaacc cagtttagcgt gtcaccata	10260
caatgcaatt tgggggtcg ggttgggttgg tttttttgt tcttttgag atgagtca	10320
ttgcattctt gcctggcctt gaactcctca cagaggcag gctggcttag tacttgtggc	10380
aattcacctg cccgttgctt ctgtgcattt gaatgatagg tatgcgtca taaggctgac	10440
ttaactgcca ttcttgat tttgtttgt tttgtgtct gttgggttg tgtggacaa	10500
ggtttactg ctgtccagaa cagcctaaaa ctcatgacaa tcctctgcc tccgcctcc	10560
aagactgaga ctgcagacat gtaccgctag gcctggtttta gagattcta aaaaaatctt	10620
agtgtgtgtg tttgtgtctg tttgtgtgtg tttgtgtgtc tttgtgtgtg tttgtgtgtg	10680
tctgtgtgtg tctgtgtgtg tctgtgtgtg tgcttagacag gctagacaag gactctacca	10740
ctgagctata ttccaaaacc ccagattac tcttaaacaa ggtcaaaacag ctaaactgtg	10800
atttggattt ggataatctt gtccctctat ttctttgcctt ctgtccaggt ttggctctc	10860
agaccctcag agttggcac tagtcggttt catctgacgg ccaaggcagg gagaggaagc	10920
ttcattctag gccacaccta cagggactcc caaggacagc tggacataaa gatgtgaccc	10980
ggtcactccc acaggccact gggccctgtc cagccttcc ccctgggaga cagccacccc	11040
actcaggcgc gcccttac	11058
<210> 20	
<211> 3158	

<212> DNA
<213> Homo sapiens
<400> 20
agcttgagat tgtccaagca ggtagccaga gagcgccatc agccaagaaa ccatccactg 60
gtacgtaagg cagcctgtgc gggcgagacc agactggcc ctcccctct gcagtgattt 120

gtttcttctt ctttttaaa tcacgtttc ctgcctttc tagttctag gtaccagcct 180
ctggcttcta cagcctcaga caatgacttt gtcacaccag agccccgccc tactaccgt 240
cggcatccaa acaccagca gcgagttcc aaaaagaaac ccaaagttgt cttctcaagt 300
gatgagtcca gtgaggaagg tatgatgctc ccgcctgttc cggccgaga aggcacacag 360
ctagggtgca gagggcttgtt ttccatagga cctgctgcgg gggcctgagt gtagatgctc 420
tgccccactg ccgcagaagg gcctctctg tacagcttgg attttatttc ttctgtgcgg 480
tgtgggatgt tctcacttgt tctctgatat ctatttttc accatcttg tgactcagct 540

ttttcttatt ccttaattt tttgcataga tcttcagca gagatgacag aagacgagac 600
acccaagaaa acaactccca ttctcagagc atcggtcgcc aggcacagat ccttaggaagt 660
ctgttccctgt cctccctgt cagggatcc tggtaggtga cctggaaattc gaattcttt 720
tcccttgtaa aatatttgc tggcttttt tttaaaaaaa aaaaaggcc gggcactgt 780
gtcacaacgt gtaatcccg cacttgcga taccaggcg ggtggataac ctgaggtagg 840
gagttcgaga ccagctgac caacatggag aaacccatc tctactaaaa ataaaaaatt 900
agccggcgt attggcgtgc gcctgtaatc ccagctactc aagaggctga ggcaggagaa 960

tcgcctgaac ccagaggcg 1020
agtttgtagt gagccgaaat cacaccattc cactccagct 1080
tgggcaacaa tagcgaacct ccatctcaaa ttaaaaaaa aatgcctaca cgcttttaa 1140
aatgcaaggc ttctcttaa attagctaa ctgaaactgcg ttgagctgct tcaactttgg 1200
aatatatgtt tgccaatctc cttgtttct aatgaataaa tggggataa tacttttaga 1260
catttttcc taagcttgtc ttgtttcat cttcacatt agcccgatgg catgcagcag 1320
agagagggtt atcagtgcag agagagatga gtgagcccg agtccctaggg cctgtcccg 1380
gatggcagat gagcttcctg ccccgtaact gccaccttc ccctctcaac ctctggaccc

tgcacagtgaa ccagacagcc tctctggga gaattatgca gtgccttaggc tccagatcag 1440
tgcttctgaa ccggggccaa ttttgtctgc cagaggacat ctgacaacac ctggggccctg 1500
ttttgttgtc atagccataa gggagaagaat gctaccagca ttgtggaa gaggccaggg 1560
atgtggctca acatctgca gtgcacagga tggccctca acaaagaatc acacggccca 1620
caatgtcaat agcgtcacag ttgagaaaac ctgctctaga ccaagggtt cttctgcgg 1680

tgtgcctcac cccacccca ctcgtgtcc ctaatccat ctccaaaggt tggcagcaga	1740
ccggcccccagg ctcgtgaaag ttcaagatcat gatccctcc agctctgcag gagacaagac	1800
ctgtctccca gcattcctca ttgttcccg gtctgcagag ggcgtgagct atgctgcagg	1860
cgggctgccc cctgaaggct ggcacccct ctccagctcc tcaagtc tctgtgagt	1920
caccccgaa cccggaggctg tgagctggct gtcgtgacca cactggtgcc tctgtgtca	1980
tgacaacagc acactacgac agtagtgctc cctgggact gagctccctc tttgcgggaa	2040
gaagacagta atgaaaaatg acaagcatga ggcagagggg aagatcacgc ttgggtggtg	2100
caggagcatg gaggtgctct taatgctctc aatgagaaag ggttaacggt cctggttgca	2160
ggaatagctg agtcagaggt ggggcttctt ccactcccc accccacccc tttcaccatt	2220
agggaccccttc ttgccttgct ctgtctactc tgctctgggt ggtcatgtg aaaagccgc	2280
accaaccatg ccagtggcag ccagacgagg acacagccctg gctctgggtc ccagcaggaa	2340
aggcaatccc agaaaggcag ggtcagggac tggagtccctg tgggtgcttt ttaagcaaag	2400
attatcacca ggcaggctaa acttagcaac cggcttttag ctagaaggc agggggctgg	2460
tgtcaggtta tgctggcca gcaaagaggg ccggatccc cttccatgc acctgctgat	2520
gggccaaggc caccccaccc caccccttc cttacaagtg ttcagcaccc tcccatccca	2580
cactcacaaa cctggccctc tgccctcta ccagaagaat ggatccctg tgggaggggg	2640
caggggacct gtccccaccc tggcccaag accttttc ccacttttc ccttttttg	2700
actcaccctg ccctcaatat ccccccggcgc agccagtgaa agggagtccc tggctctgg	2760
ctcgccctgca cgtccccagg cggggaggga cttccgcct cacgtcccgc tcttcgcccc	2820
aggctggatg gaatgaaagg cacactgtct ctctccctag gcagcacagc ccacaggttt	2880
ccaggagtgc ctttgtggga ggcctctggg ccccccacccag ccatccctgca ctccgcctgg	2940
ggcccccagcc cggagagagc cgctggtgca cacagggccg ggattgtctg ccctaattat	3000
cagggtccagg ctacagggtc gcaggacatc gtgaccccttcc gtgcagaaac ctcccccctcc	3060
ccctcaagcc gcctcccgag cctccttctt ctccaggccc ccagtgcaca gtgcggcagtg	3120
cccgcccgag gcctcggtcc cagagatgcc aggagccca	3158
<210> 21	
<211> 3218	
<212> DNA	
<213> Homo sapiens	
<400> 21	
aagctggcac cactacttca gagaacaagg cttttccctc tcctcgctcc agtccttaggc	60

tatctgtgt tggccaaaca tggagaaggc tattctgtgg gcagccccag ggaggctgac	120
agggtggagga agtcagggtc cgcaactggc tctgacgctg actggtagt ggagctcagc	180
ctggagctga gctgcagcgg gcaattccag ctggccctcc gcagctgtga ggtcttgagc	240
acgtgctcta ttgtttctg tgccctcgta tcttatctga ggacatcgta gccagcccc	300
aaggcttca agcaggattc atcttaggtaa accaagtacc taaaaccatg cccaaggcgg	360
taaggactat ataatgttta aaaatcgta aaaatgccc cctgcatacg ttttggaa	420
gatgaactga gatgtgtcag ggtgacttat ttccatcatc gtccttaggg gaacttgggt	480
agggcaagg cgtgtagctg ggacctaggt ccagaccct ggctctgcca ctgaacggct	540
cagtgcctt gggcagttac tcccgccct cacttgcac gtgtgcctac ctatggaga	600
caaaagtaca tacctcgta gagcgcgcac gcctgttaacc ccagcactt gggaggccaa	660
ggtgggtgtt tcaacctgagg tcaggagttt gagaccagcc tggccaaacat ggtgaaactc	720
cgtctctact aaaattacaa aaatcagcca ggcttcattt cacatgccta tagtcccac	780
tacaggcatg ctgaaggcagg agaatcgctt gcaccccgga ggcagaggct gcagtgagct	840
gagaccacac cactgcactc cagccttaggc aacagagttt gagactccat ctcaaaaaaa	900
aaaaaaagtac ctacctcaga gttcaaaacta gtgaatattt ggaagtgcctt gagacagtga	960
caccaaaatgtt cacaataat actcgccatg ttcatttattt ttaagaatc catttgaatg	1020
tcaagtcac acagcctctt ataccgaggc attgtgaacc gcatctcccc agttctcca	1080
ggctttcca agaatcaggg acactgttagc ctgttggctc cagtgtatga cagacacgg	1140
ggaagcacat cttagctga tacttaaaca gagaccctga ggcacatac acccgccac	1200
acatgcatttgg agttcacct tctctgtcat tctgcagtgcc ccaggagagc aagagctccc	1260
accccttc aaaacactgt gcccattccg ggcactaagg ccttttaaa gcacggcacc	1320
tccacgaggg agggccacag ccacatacac tccacctggc aggtggacag cgtgagcagc	1380
tggaccatag cagggacaag gtgccccggc cagccccaaac gcctctgccc gctgacaggg	1440
acagaagccc tctccagctc cgtgtgcgtc agaggccatg cgtgcctcc agtcgcattt	1500
tattccactc cagtgcctgg gccagtttagc accagtgtgg aagacagtga gctggctccg	1560
gacaacaggg atggaggaaa ggtccacat tcacattctt gatacggttga caaggtgagg	1620
ggccgcaatc gctctggcag cattttaaag atggggaaatg agcagacacc cacgcgtgaa	1680
ggcaggagag ccccaactgt ggtggaaatg gccccagaat ggttagggca agccttagctc	1740
cagacacccc agagccctgg agaagccaaag actgaggggag aaaggcttgcgg gggaggcgc	1800
cccaagtcccc agggaccggc ctgggtgcaga gctgcagctg atgttccctt ctgtgcagcc	1860

ccaccctctg cctcgctgag ctccctgctg cgagggccctc gggtgcaagg gggaggcagg	1920
tctctatctc atggagctgt cagatgagac atcgcgatcg gagtcctcag cctcgcttgg	1980
cggcggcggc gggtcgctaa gcgggaccgc agtgaaagca ggagacttgc tagaaaaaaa	2040
caccagtgt caaccttggg gcaggcagga atcctgaaga cggacggcac tcctccctc	2100
gctgcctcac cctctggcag cccgtgagaa gtacccgaaag cgagggcggg gccgcgggat	2160
ggcgaggggag cggcaggac tgaactctc ccaaaccac cctgacaggg aaatggccc	2220
cgcctgtgtc ttgggaactc agaggctgag gtcaggcatg atggctcacg cctgttaattc	2280
cagcactttg ggaggcagag gcgggtggat cacgacgtca ggagttcaag gcaagcctgg	2340
ccaacatggt gaaacccat ctctactaaa aatgcaaaaa ttagccaggt gtggggcgg	2400
gcgccttggg ggcgtgaggca agataatcac ttgaacctgg gaggcggagg ttgttagttag	2460
ccaaaaaaa aaaaaaaaga aatagctgaa gtcacagtag gagagaagct gctgagcctc	2520
cagcacccctg actctaggc cttggctta tgtctatctg cagtattttt gtgatttttta	2580
aaaatttact ttcttggc ggtgttaactt acacagggtc aaatgcacaa atcatggccc	2640
tgactttaga taaaaatctg cccccacaaac cttctgttc cttgccagtt tttaaactgc	2700
ctctaaccag gggaaaccacc agagctggtg gccttgggag gttcagccc tcccgatcg	2760
aatggacata gctcatccaa ctgccaaggg agagagctgt gggcttggc cagccccacc	2820
aggttaactcc caaagggcag ccccacagca agatgtgacc cagtcattgc ctgagggct	2880
ctggggctgt gtccaaacct ttctccccgc tgtgtcccc tggaaaggccc catgcccagg	2940
ggaggcgcct accggcttcc agactgggtg atgagccggc ggaggcttc catctgcacg	3000
tccaggccgc gtttcatgtc gcacatctcc atgtactcgt gcagggtccg gttcatgtcg	3060
ttcttggccg tggcaactc cagctgtggt gggcaggca gggccatcg ggttaacagg	3120
tggcgttcac agccctctg ttggcccccgc caggaggcca acacgccaag agcagtggt	3180
ggccgggggg cccaggcagc cattactgaa ggctcaga	3218
<210> 22	
<211> 508	
<212> DNA	
<213> Homo sapiens	
<400> 22	
agctttagat tgtccaaagca ggttagccaga gagcgccatc agccaagaaaa ccatccactg	60
gtacgtttagg cagcctgtgc gggcgagacc agactggcc ctccctcct gcagtgtttt	120
gtttttttttt cttttttaaa tacgttttc ctgcctttc taggttctag gtaccaggct	180

ctggcttcta cagcctcaga caatgactt gtcacaccag agccccccg tactaccgt	240
cggcatccaa acacccagca gcgagcttcc aaaaagaaac ccaaagttgt cttctcaagt	300
gatgagtcca gtgaggaagg tatgatgctc ccgcctgtc cggccgaga aggcacacag	360
ctagggtgca gagggtcggt ttccatagga cctgctgcgg gggctgagt gtatgtctc	420
tgcccactg ccgcagaagg gcctctcctg tacagttgg atttatttc ttctgtgcgg	480
tgtgggattg tctcacttgt tctctgat	508
<210> 23	
<211> 2650	
<212> DNA	
<213> Homo sapiens	
<400> 23	
atctatttt tcaccatctt tgtgacttag cttttctta ttccttaat tcttgata	60
gatcttcag cagagatgac agaagacgag acacccaaga aaacaactcc catttcaga	120
gcatcggtc gcaggcacag atcctaggaa gtctgtcct gtccctcctg tgcatggat	180
cctgtagggt gacctgaaat tcgaattctg tttcccttgt aaaatatttgc tctgtcttt	240
tttttaaaa aaaaaaaagg ccggcactg tggctcacgc ctgtatccc agcaacttgc	300
gataccaagg cgggtggata acctgaggta gggagttcga gaccagcctg accaacatgg	360
agaaaccca tctctactaa aaataaaaaa ttagccggc gtattggcgt ggcctgtaa	420
tcccaagctac tcaagaggct gaggcaggag aatcgctga acccagaggc ggaggttgt	480
gtgagccgaa atcacaccat tgcactccag ctgggcaac aatagcgaac ctccatctca	540
aattaaaaaa aaaatgccta cagcgtctt aaaatgcaag gctttcttt aaattagcct	600
aactgaactg ctttgagctg cttcaacttt ggaatataatg tttgccaatc tcctgtttt	660
ctaataata aatgtttta tatactttt gacatttt cctaaatgttgc tctttgtttc	720
atcttcaca ttagccagt ttcatgcagc agagagaggg ttatcgtgc agagagagat	780
gagttagccc agagtcctag ggcctgtccc gggatggcag atgagcttcc tgccctgtca	840
ctgcccaccc tcccctctca acctctggac cctgcacagt gaccagacag cctctgtgg	900
gagaattatg cagtgcctag gctccagatc agtgcttgc aaccggggc aattttgtct	960
gccagaggac atctgacaac acctggggcc tgtttggc tcatgccta tagggaaaga	1020
atgctaccag catttgtgg aagaggccag ggatgtggct caacatcctg cagtgcacag	1080
gatggccct caacaaagaa tcacacggcc cacaatgtca atagcgtcac agttgagaaa	1140
acctgctcta gaccaagggt tgcttctgc cgtgtgcctc accccacccc cactcgttt	1200

ccctaattccc atctccaaag gttggcagca gaccggccca ggctcggttga agttcagatc	1260
atgatccctt ccagctctgc aggagacaag acctgtctcc cagcatccctt cattgttccc	1320
gggtctgcag agggcgttagt ctatgctgca ggcgggctgc cccctgaagc ctgcgcaccc	1380
ctctccagct cctcaagtct tctctgctga gtcacccctcg aaccggaggc tgtgagctgg	1440
ctgtcgtgac cacactggtg cctctgctgt catgacaaca gcacactacg tcagtagtgc	1500
tccctggca ctgagctccc tcttgcggg gagaagacag taatgaaaaa tgacaagcat	1560
gaggcagagg ggaagatcac gcttgggtgg tgcaggagca tggaggtgct cttaatgctc	1620
tcaatgagaa aggttaacg gtcctggtt caggaatagc tgagtcaagag gtggggcttc	1680
ctccactccc ccacccacc ctttacca ttagggacct tcttgcttgc ctctgttac	1740
tctgctctgg gtggtcattt tgaaaagccc gcaccaacca tgccagtgcc agccagacga	1800
ggacacagcc tggctctggg tcccagcagg aaaggcaatc ccagaaaggc agggtcaggg	1860
actggagtcc tgggggtgt ttttaagcaa agattatcac caggcaggct aaacttagca	1920
accggctttt agctagaagg gcagggggct ggtgtcaggt tatgctggc cagcaaagag	1980
gccccggatc cccctccat gcacctgttgc atggccaag gccacccac cccacccct	2040
tccttacaag tggcagcac cttccatcc cacactcaca aacctggccc tctgcccctcc	2100
taccagaaga atggatcccc tgggggggg ggcaggggac ctgttccac cgtgtgcacca	2160
agacactttt tcccactttt tcccttttct tgactcaccc tgccctcaat atccccggc	2220
gcagccagtg aaaggaggc cctggctct ggctcgctg cacgtcccag ggcggggagg	2280
gacttccgcc ctcacgtccc gctttcgcc ccaggctgga tggaaatgaaa ggcacactgt	2340
ctctctccctt aggacgacaca gcccacagggt ttccaggagt gccttgcgttgg gaggcctctg	2400
ggccccccacc agccatcctg tcctccgcgtt gggccccag cccggagaga gccgctggtg	2460
cacacagggc cgggattgtc tgccttaattt atcaggtcca ggctacaggc ctgcaggaca	2520
tcgtgacctt ccgtgcagaa acctccccctt cccctcaag cccctcccg agcccttcc	2580
ctctccaggc ccccaagtgc cagtgccctt gcccagccc aggctcggtt cccagagatg	2640
ccaggagcca	2650
<210> 24	
<211> 1963	
<212> DNA	
<213> Homo sapiens	
<400> 24	
agctttagat tgtccaagca ggttagccaga gagcgccatc agccaagaaa ccatccactg	60

gtacgttaagg cagcctgtgc gggcgagacc agactggcc ctcccccct gcagtgattt	120
gtttcttctt ctttttaaa tcacgtttc ctgcctttc taggttctag gtaccagcc	180
ctggcttcta cagcctcaga caatgacttt gtcacaccag agccccccg tactaccgt	240
cggcatccaa acacccagca gcgagcttc aaaaagaaac ccaaaggttt cttctcaagt	300
gatgagtcca gtgaggaagg tatgtatgtc ccgcctgtc ccggccgaga aggcacacag	360
ctagggtgca gagggtggg ttccatagga cctgctgcgg gggcctgagt gtatgtgtc	420
tgccccactg ccgcagaagg gcctctcctg tacagcttg attttatttc ttctgtgcgg	480
tgtgggattt tctcacttgt tctctgatat ctatttttc accatcttg tgactcagct	540
ttttcttatt ccittaattt ttgcataaga tctttagca gagatgacag aagacgagac	600
acccaagaaa acaactccca ttctcagagc atcggtcgcc aggcacagat cctaggaagt	660
ctgttcttgtt cctccctgtg cagggtatcc tgggttgc cctggaaattt gaattctgtt	720
tcccttgtaa aatatttgtc tgcgtttttt ttttaaaaaa aaaaaggcc gggcactgtg	780
gctcacgcct gtaatcccaag cactttgcga taccaggcg ggtggataac ctgaggtagg	840
gagttcgaga ccagcctgac caacatggag aaacccatc tctactaaaa ataaaaattt	900
agccggcggtt attggcgtgc gcctgtatc ccagctactc aagaggctga ggcaggagaa	960
tcgcctgaac ccagaggcggtt aggtttagt gagccgaaat cacaccattt cactccagct	1020
tggcaacaa tagcgaacctt ccatctcaaa ttaaaaaaa aatgcctaca cgctcttaa	1080
aatgcaaggc ttctcttaa attagcctaa ctgaaactgctt ttgagctgtc tcaactttgg	1140
aatataatgtt tgccaatctc cttgtttctt aatgaataaa tttttttata tacttttaga	1200
catttttcc taagcttgc tttgtttcat ctgttacatt agcccgatttt catgcagcag	1260
agagagggtt atcagtgcag agagagatga gtgagcccg agtccctaggc cctgtcccg	1320
gatggcagat gagttccctg ccccgtaact gcccaccccttcc cctctcaac ctctggacc	1380
tgcacagtga ccagacagcc tctctggga gaattatgca gtgcctaggc tccagatcag	1440
tgcttctgaa ccggggcataa tttgtctgc cagaggacat ctgacaacac ctggggcctg	1500
ttttgttgc atagcctata ggggaagaat gctaccagca ttgtggaa gaggccagg	1560
atgtggctca acatcctgca gtgcacagga tggccctca acaaagaatc acacggccca	1620
caatgtcaat agcgtcacag ttgagaaaac ctgctctaga ccaagggttg ctgttgcgg	1680
tgtgcctcac cccacccca ctgcgttcc ctaatcccat ctccaaaggt tggcagcaga	1740
ccggcccgagg ctgcgttccat gatccctcc agtctgcag gagacaagac	1800
ctgtctccca gcattcctca ttgttcccggtt gtcgttccat ggcgttgcagat atgttgcagg	1860
cgggctcccc cctgaaggctt ggcacccctt ctccagctcc tcaagtcttc tctgttgcagg	1920

cacccatcgaa ccggaggctg tgagctggct gtcgtgacca cac 1963

<210> 25

<211> 1195

<212> DNA

<213> Homo sapiens

<400> 25

tggtgcctct gctgtcatga caacagcaca ctacgtcagt agtgctccct gggcactgag 60

ctccctcttt gcggggagaa gacagtaatg aaaaatgaca agcatgaggc agaggggaag 120

atcacgcctt ggtggtgcag gagcatggag gtgccttaa tgcctcaat gagaaagggt 180

taacggtcct ggttgcagga atagctgagt cagaggtggg gcttcctcca ctccccacc 240

ccaccccttt caccattagg gaccccttg cttgtctt gtactctgc tctgggtgg 300

cattgtgaaa agccgcacc aaccatgcca gtggcagcca gacgaggaca cagcctggct 360

ctgggtccca gcagggaaagg caatcccaga aaggcagggc cagggactgg agtccctgtgg 420

gtgcctttta agcaagatt atcaccaggc aggctaaact tagcaaccgg cttttagcta 480

gaagggcagg gggctgggt caggttatgc tggccagca aagaggcccg ggatccccct 540

cccatgcacc tgcgtatggg ccaaggccac cccacccac ccccttcctt acaagtgttc 600

agcacccctcc catcccacac tcacaaacct gcccctctgc cctctacca gaagaatgga 660

tccctgtgg gagggggcag gggacctgtt cccaccgtgt gccaagacc tctttccca 720

cttttcctt ctcttgact caccctgccc tcaatatccc cggcgcagc cagtgaaagg 780

gagtcctgg ctccctggcgc cctgcacgt cccagggcgg ggagggactt ccgcctcac 840

gtcccgctc tcgccccagg ctggatggaa tgaaaggcac actgtctctc tccctaggca 900

gcacagccca caggttcca ggagtgcctt tgtggaggc ctctggccccc ccaccagccca 960

tcctgtctc cgcctgggc cccagcccg agagagccgc tggtgacac agggccggga 1020

ttgtctgccc taattatcg gtccaggcta cagggtgca ggacatctg accttcgt 1080

cagaaacctc cccctcccc tcaagccgccc tcccggctt cttccctctc caggccccca 1140

gtgcccagtg cccagtgc cccaggcc tcggtcccag agatgccagg agccca 1195

<210> 26

<211> 256

<212> DNA

<213> Homo sapiens

<400> 26

cctctggcc cccaccagcc atcctgtct ccgcctggg ccccagccg gagagagccg 60
 ctggcaca cagggccggg attgtctgc ctaattatca ggtccaggct acagggctgc 120
 aggacatcgt gacccatcg 180
 tccttcctc ccagcccccc agtgcctcact gcccactgccc cagccaggc ctcggccca 240
 gagatgccag gagcca 256

<210> 27

<211> 1941

<212> DNA

<213> Homo sapiens

<400> 27

agcttgatgttgtccaa ggttagccaga gagcgccatc agccaagaaa ccatccactg 60
 gtacgttgc cagccgtgc gggcgagacc agactggcc ctccctcct gcagtgattt 120
 gtttcttctt ctttttaaa tacgttttc ctgcctttc tagttctag gtaccagct 180
 ctggctcta cagcctcaga caatgacttt gtcacaccag agcccccgg tactaccgt 240
 cggcatccaa acacccagca gcgagcttc aaaaagaaaac ccaaagtgtt ctttcaagt 300
 gatgagtcca gtgaggaagg tatgtatgctc ccgcctgttc cggccggaga aggcacacag 360
 cttaggtgca gaggcgttgtt tccatagga cctgctggg gggctgagt gtatgtctc 420

tgccccactg ccgcagaagg gcctctcctg tacagcttg attttatttc ttctgtgcgg 480
 tgtggatttgc ttcacttgtt tctctgatat ctatttttc accatcttgc tgactcagct 540
 ttttcttatt ccttaatttcc ttgcataaga tcttcagca gagatgacag aagacgagac 600
 acccaagaaa acaactccca ttctcagagc atcggctcgc aggcacagat cctaggaagt 660
 ctgttctgtt ctcctgtt cagggtatcc tggatgttgc cctggatttgc gaaccggagg 720
 ctgtgagctg gctgtcgtga ccacacttgtt gcctctgttgc tcatgacaac agcacactac 780
 gtcagtagtg ctccctggc actgagctcc ctcttgcgg ggagaagaca gtaatgaaaa 840

atgacaagca tgaggcagag gggaaatca cgcttgggtt gtgcaggagc atggagggtgc 900
 tcttaatgtt ctaatgaga aagggttaac ggtcctgggtt gcaggaatag ctgagtcaga 960
 ggtggggctt ctcctactcc cccacccac cccttcacc attaggacc ttcttgcctt 1020
 gctttgttca ctctgtctg ggtgggttgc tggaaaagcc cgccaccaacc atgcccgtgg 1080
 cagccagacg aggacacagc ctggctctgg gtcccagcag gaaaggcaat cccagaaagg 1140
 cagggtcagg gactggagtc ctgtgggtgc ttttaagca aagattatca ccaggcaggc 1200
 taaacttagc aaccggcttt tagctagaag ggcaggggc tgggtcagg ttatgtctgg 1260

ccagcaaaga ggccgggat ccccctcca tgcacctgct gatggccaa ggccacccca	1320
ccccacccctt tccttacaa gtgttcagca ccctccatc ccacactcac aaacctggcc	1380
ctctgccctc ctaccagaag aatggatccc ctgtggagg gggcagggga cctgttcca	1440
ccgtgtgcc aagacctttt tccttactt ttccctcttc tigactcacc ctgcctcaa	1500
tatccccgg cgccagccagt gaaaggagt ccctggctcc tggctcgct gcacgtccca	1560
ggccggggag ggacttcgc cctcacgtcc cgctttcgc cccaggctgg atggaatgaa	1620
aggcacactg tctctctccc taggcagcac agcccacagg ttccaggag tgccttg	1680
ggaggccctt gggcccccac cagccatctt gtcctccgcc tggggcccca gcccggagag	1740
agccgcgtgt gcacacaggg ccgggatgt ctgcctaat ttcagggtcc aggctacagg	1800
gctgcaggac atcgtgacct tccgtgcaga aacctcccc tccccctcaa gccgcctccc	1860
gagcctcctt cctctccagg cccccagtgc ccagtgcaca gtgcggcagcc caggcctcgg	1920
tcccaagat gcccaggagcc a	1941
<210> 28	
<211> 1433	
<212> DNA	
<213> Homo sapiens	
<400> 28	
atctatttt tcaccatctt tgtgacttag cttttctta ttccttaat tctttgcata	60
gatcttcag cagagatgac agaagacgag acacccaaga aaacaactcc catttcaga	120
gcatcggtc gcaggcacag atccttagaa gtctgttctt gtctccctg tgcagggtat	180
cctgttagggt gacctgaaat tcgaaccgga ggctgtgagc tggctgtcgt gaccacactg	240
gtgcctctgc tgcgtgaca acagcacact acgttagtag tgctccctgg gcactgagct	300
ccctcttgc ggggagaaga cagtaatgaa aaatgacaag catgaggcag agggaaagat	360
cacgcttggg tggtgccagga gcatggaggt gctttaatg ctctcaatga gaaagggtta	420
acggcttggg ttgcaggaat agctgagtca gaggtggggc ttccctccact cccccaccc	480
accccttca ccattaggga ctttcttgc ttgtcttgc tactctgctc tgggtggta	540
ttgtgaaaag cccgcaccaa ccatgccagt ggcagccaga cgaggacaca gcctggctct	600
gggtcccagc agggaaaggca atcccagaaa ggcagggta gggactggag tcctgtgggt	660
gttttttaag caaagattat caccaggcag gctaaactta gcaaccggct ttttagctaga	720
agggcagggg gctgggtgtca ggttatgtc ggccagcaaa gaggccggg atccccctcc	780
catgcacctg ctgtatggcc aaggccaccc caccaccccttac aagtgttcag	840

caccctccca tcccacactc acaaacctgg ccctctgccc tctaccaga agaatggatc	900
ccctgtggga gggggcaggg gacctgttcc caccgtgtc ccaagacctc ttttccact	960
tttccctct tcttactca ccctgcctc aatatcccc ggcgcagcca gtgaaaggga	1020
gtccctggct cctggctgc ctgcacgtcc cagggcgggg agggacttcc gccctcacgt	1080
cccgcttc gccccaggct ggatggaatg aaaggcacac tgtctctc cctaggcagc	1140
acagccccaca gtttccagg agtgccttg tgggaggcct ctgggcccc accagccatc	1200
ctgtccctcg cctggggccc cagcccgag agagccgctg gtgcacacag ggccgggatt	1260
gtctgccta attatcaggt ccaggctaca gggctgcagg acatcgtgac cttccgtgca	1320
gaaacctccc cctccccctc aagccgcctc ccgagccctc ttctctcca ggccccag	1380
gcccagtgcc cagtgcctc cccaggccctc ggtccagag atgccaggag cca	1433
<210> 29	
<211> 17130	
<212> DNA	
<213> Cricetus griseus	
<400> 29	
gtgagtggtg cagggtgact ggccttattt ggttctgtt cagtgaacac agtctgttct	60
tcgagtgctg tggactctt aacttactgc cgtgctggcc tcaactacagc catttaacta	120
gagtgacttg agatcacttt ctatgtccac ttccgggtcc taaaaaaaaa atttattttg	180
tatacattag tattttgcct gtgtgtatgt ccatgtgaag gtgccagatc tggaacttga	240
gttacaggca gttgtgagct gccacatggg tgctggaaac tgaaccagg tcttctgcag	300
gagcagcaag tgccatctct ctagccccca gtaaaagccg atttctagtt gagacatgg	360
ctttctgtgt agctctagct ggcttagaat ttgctaccta gaccaggctg gccccaaatt	420
cagattcccc tgcctctgcc tctcatgtgc tggtagaaaa ggcgtgagtc acatgcttgg	480
cactagacta gcttgtgtt tagagcaaaa ctaagtaagt agagttcccg tttgtccctg	540
tcctctgcctt catcatcttc atacccttgg gtactgtat cttacaatct acagtgacat	600
ggatgaatga gtgtgagtga caatgaatga cagagcagta cccagagtcc acagttcaca	660
ctgatactgg gttcactgtt cagtggcaca tacttgtcac cacagtgtca tactggacat	720
tttcatgacc ctaaggccac atcacccttg tggtgtcact cagttgtat ttccaccstat	780
atagtgattt caggttggct tttacttag aatataacat ttaaattatg tttcttttt	840
ggttttgagt ttgttttc gagacagggt ttctctgtat tgctttggag gttgtgtgg	900
aacttgcctt ggagaccagg ctggcttcga actcacagag atcctccac ctctgcctcc	960

cgagtgtgg gattaaaggc gtgcgccacc aatgcccg catttctg tttctgtt	1020
tttaaagagg gagataatgt ttccagtgtc gttttaact ttttcgtt aggctaccat	1080
agatgaattt gaggcagaagc ttccggcatg tcataccaga ggcattggatg gaatagagga	1140
gcttggaaact ggcagaagaag gcagccagca agccctgtct gccaagaaac cctctgtgg	1200
tatgagagac aageccacac ggactccca caaatcatcc ctgtcatttt ttgggttttg	1260
tttttaaaat cacatctccc tgccatttct agtctcaagg cgacagccctc tgacttctat	1320
agattcagac aatgactttg tcacaccta gcctcgacgc accaaccgtc atcgtccaa	1380
cactcagcaa cggaaagtcca agaagaaagc caaagttgtc tttcaagtg atgagtcag	1440
tgaggatggt atgactcgcc ctccggcag ctgcttgctt gccttgggt ctccctgggg	1500
cttgcctgg ttttaggtca ttcttcctgg taatgtgtc cccactgtgc acgttggc	1560
acctttgcag ctctatgttt taagatccac ttacaccat ctttgtact cagctttgt	1620
tcctataatt ctgttgcac aactttcagc ggaaatgacg gaagaagaga ctcccaagag	1680
aaccacccca atccgcagat catctgccc aagacacagg tcttaggagt tgcctcacct	1740
attcccatcc ctgtgtgca gggtatctt attgatttt aattcatttc ctccctgtt	1800
aaaatgttgc ctccctgttt ttaatatgtt aatggtctt gcatcaagg ctttcct	1860
agagcccatc tgactgctgc ttctacttca caacaaacac ttaaccaatc accatgttt	1920
ctaagaataa atgttttat atactttat tctttgtgg tctgttcat ctcccacctt	1980
agcccgatt caggcagcaa tagagggtt tcaaaccaga gagaagatac accagcccg	2040
gcacccctgag gctgtctctg ggtcttcatg ttcccttatt gattcatctt gcacaactga	2100
tgagactcca ctgttagagaa ttctcacttc tcagectggg ggtactttct acccagagac	2160
ctatttctgt catagccctt tggcacat aaagttctt tccgtcatag aacacccac	2220
aacagtgtat caggagtgtt caagttgaca aacaatgtc tagacctacg gttttttcc	2280
tctgtgtgcc tcatccaga agagatcatg actcccgaga gtcagccctt actatgggt	2340
ctgcaggggc gtcagcccc ttagccggca gccatgccca ccctcccaa gtccttaatc	2400
tgctgagtca ttggaaacag gagacactga ttctgtgtc atgacaacag cacattgcca	2460
tagaaatgtt ccctacctt tacgtgtgt gtggggaaac agtaatgaca aaccataggc	2520
aggaggcaaa agggaaagacg gcacccatcaga aacatgtttt aggttagggc agaactatgg	2580
aggggctctt gagactctt gatggaaag ggttaatgtc gtcctgaaa cctctgttgg	2640
aaggcagaaa agggacaggg ctgagtcctt gcactggac cattccatc ctctgcatcc	2700

tgccccggc tcatggaaag cctggcatg ggccacacag ctgtcagtct tggctctgg	2760
gccccaaagga ggtaggcaaa tcccagaatg gcaaggagcc aggactggat ttgggggtgca	2820
gcccagctg ctccctgcct ttaagcaaa ggttatcacc aggccagcta aacttagcaa	2880
ttaggctctt cagctaaaag agcagggggc tggctcaag ttgcactgac ctagcaaaga	2940
ggcccagga tccccctgcc cagcacctgt ggctgagctc ccaagccctt cccgagagct	3000
caggatccac ctttccacc ctccctactc ttcaaggag gaacccctt tctccttccc	3060
acttggtaa ggggctggg gccaggctgt tctggcttg ggtataatac cccctacccc	3120
ttctactttc ccctcctctc agacctcacc ctgcctccac gagggcagcc aagaaaggag	3180
agtccctggc tgcaaggcca gtaggcacgt cccaggacgg ggagggactt ccgcctcac	3240
gtccagctct ccgcctggg gctgcagtgg gtgaaagggg cagtgtctcc tagcctggc	3300
ggtgcaaccc tcaggttccg aggaggaacg ctctggagg ctctttgcc tcctccaacc	3360
caacccacaa ccaggacatt gtcctcaccc cggggccca acctagacct taactgagga	3420
acacagaggc cagttgtaa gtctcaatta tgcaaggcat cccgacctgt ggcgttaggaa	3480
gcgccttcc aggccgcttc cctagccctc tcctggccct cacagccag gcctctggcc	3540
caagaaatgg aagtgggggt ggggatgga actgcgaatg cgaaggccc ccgcaggagg	3600
caaagtgacc cctccgggc ctttctgtcc cgagacttg ttttgcctg tgtcactacc	3660
gaagaaccac gagaagatcc tcaacttttcc cacagcttt gcataaaggg gagagggctg	3720
gcgggtgcagc tgtggcacac acgcacttct gctcaacccg ccccccccg ccccggttcc	3780
tgttccttcc caggttctcc ccattttac gggggccaa ctttaggtc cctgggtctt	3840
ggaagtccctt agtacacact cttcgctt aagtccatag tctgtattcc ctcggctcta	3900
tcctgtcccc catcacccggg tcacccccc agcgaagcaa tctcagttcc cctccccctc	3960
tcagccccga gcccacacgt ttggcgttg cacattcaa aaacgaggcg ggtccaaaga	4020
gaggggggtgg ggaggtgccg agtggccag ctactcgccg cttacgggt gcacgtagct	4080
caggcctcag cgcctttagt ctgtgactgg atggatgagc gggggggggag gggggggcag	4140
cgtcctcgcc gctccccacc accccagttc ctataaatac ggactgcagc cctccccgt	4200
gctctctgtt cctccctgtt ctagagacag ccgcacatcc cctgtcagtg ccaggtgaaa	4260
accgcagagt gggccgcagg tggccgggaa cggtcggaaa cggggaaaggg gggcgctcag	4320
cccgggactg cgggcgtgg ggcgagctcc actgcccggag cccggctcc gcattgcaga	4380
ggctggaggg ggaaggatg gggggcgcgg cggggggcac gctgacccctt ctttcttctc	4440

tctccctccg cagccctcgct ccggagacgc aatggtaag gtcggcgtga acgggtgagt	4500
tcgtggctgg gctagggtgg ggctccgggt cccgctccgt cgcgtatgca ggtctacccc	4560
accccgggc tctgcgggag cgtggggtgg ccgggtgggtg gccgcagcac ccaaggagac	4620
ctcaaggta cgcagccacc tcctcccttg cgggatgag cagcgccggag tccnnnnnnn	4680
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	4740
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	4800
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	4860
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnaggtcggt	4920
ggcacccggc cgtgcggtgc ccgcctttagc gcatccatca tctcccaagg gcttcctta	4980
gggtggctgg ccgcgcctat gttgcaaacg ggaaggaaat gaatgaacca cggttagaa	5040
acctcccttc ggcttcctc cttcttagcc cgtgactaac ctccccactc cctcccccggg	5100
tggagtcgcc tctgtactgt aagccaggtg atgcaaggct tccgtgctct cgagagagct	5160
ctacctcgcc agctgtctca tattattagc ctcaaagcag ccctaagcc tcatttacct	5220
tgagcatatg atatatttg tagattctct gagaatcgaa gcggacttgg agaggtctgc	5280
ttgtccttct cccagccaa aggtggtagc tatggcgtag cgccggaggg gggagtgcccc	5340
ggggagctga gtcatggtgg ttctgaaaag aaaattcca ccacaaaatg gtcgcgtgc	5400
tagcateccc ttccccccat aacctctgct tcccatcaca ccctgaccca aaccctgttag	5460
gccagactgt aaaggtaact aagaggattt agtgtctgag cctcgaaacc ctgccttct	5520
cccatccca tcctctggaa accagatctc ccccgatcca ccctaactg aggttatatt	5580
tagccggctg accttctgtg atttgggtc tggccctta cacacatctg ttgcctcgc	5640
tcctgatttt tagtagcaa attcaagtgc ttgcaaatt agagccagg gattaggggt	5700
tggaaagctc agtggttttc tcagtccttc ccttagggg gagggacttg gaggaagcag	5760
gtggccgac ccctgtccta ctcatctga ccttaacct tgccttga gcttgatgat	5820
gctgagtgca cgatcttc ctgtccagg ggttagcct gaagccagcc caggctagaa	5880
caaacttccc aggggtggg ggttagtgaat gccttgtcc cacacagggg cacactgcca	5940
cctttggag acttggaaatg actgggtggg gggttggaca aggcttgag cccaaatcacc	6000
tcttggacag gaaagtaacc cccactttat ggcctgctg taaaagccca gtcaaaccc	6060
atttgtccaa ggaagataga cctttgggg ctccctaagg ataggggtgt tctatatttgc	6120
ggccctgctt ctaagcatc agccagctt attaaaggaa attcataaca aaacttgaat	6180
ttcctgttcc taaaatacta atagtgtgt ggtatctccat taaaatgtt gtcttgacca	6240
gttaggctatg gttctgtgg gctcttaca gctatggac aactggattc tgtttctga	6300

aggcatgtg tcagtcagt actgactata gacatatgag ttctctgacc ccctaactca	6360
cctttttt tctgcctca gattggccg tattggacgc ctggttacca gggctgcctt	6420
cactctggc aaagtggaaat ttgttgcatt caatgacccc ttcatgtacc tcaactacat	6480
gtctacatgttccagatgt actctaccca tggcaagttc aaaggcacag tcaaggcgtga	6540
gaatggaaat ctgtcatca acggaaaggc catcaccatc ttccaggagc gagatccgc	6600
caacatcaaa tgggtgtatg ctggccgcga gtatgttg gaatctactg gcgttccac	6660
caccatggg aaggctgggg cccacttcaa gggcggggcc aagagggtca tcatctccgc	6720
ccctctgct gatccccca tgggtgtatg ggggtgtgaac caagacaagt atgacaactc	6780
cctcaagatt gtcaaggtagt gatggcagag ggctgtggca aagtggcaaa gcagggcaaa	6840
gttacaggt gggcgagct cctaaccgtt ctctctt cagcaatgcg tcctgcacca	6900
ccaactgtt aaaaaaaaaaaaaa gccaagggtca tccatgacaa ctttggcatt gtggaaaggac	6960
tcatggtag tagttcatct gtttcatctt gccagcagtggcgctgtgg tggggccct	7020
gcaagaccc actccctgcc tctgtgtt tcagaccacg gtccatgcca tcactgccac	7080
ccagaagact gtggatggcc cctccggaa gctgtggcgt gatggccgtg gggctgccc	7140
gaacatcacc cctgcatcca ctggcgctgc caaggctgtg ggcaaaagtca gcccagagct	7200
gaacggaaat ctgactggca tggccttccg tggcttacc cccaaatgtt ccgtgtggaa	7260
tctgacatgt cgcctggaga aacctgtatg tctgggtgg gctgagggtt gtctctatgt	7320
gtgagggttgg ggctttagt gtcacccatgtt ttttgcct taataggcca agtatgagga	7380
catcaagaag gtggtaagc aggcatctga gggccactg aaggcatcc tggctacac	7440
cgaggaccag gttgtctctt gcgacttcaa cagtactcc cactttcca ctttgc	7500
tgggctggc attgtctca atgacaactt tggtaatgtc atttctgtt atgtggatc	7560
agaaacatgc tttaaaatg aacttgggtt ttgtcaccc tgggtgtcta actcaaaaag	7620
taaaccccttc ctaactgccc tcccaatgtt ttcttaggtt gacaatgaaat tggctacag	7680
caacagatgt gtggacccatca tggcctacat ggcctcaag gatgtggaaag cccaccctgg	7740
accatccacc ccagcaagga ctggcgacaa agggaggccc tggctgtga gcaatccctg	7800
tccaaataacc cccacaccga tcatctccct cacagtgtcc atccagacc cccagaataa	7860
ggagggcctt agggagccct actctcttgc ataccatcaa taaatgtcac tgcacccatc	7920
ttcccttggcc ttcaatgtt aggttgggg agggaggctg tgacctagca aggttggaa	7980
tcctctgtt cactttcaa acagggcact agccacatgc cagccaggt ttccctgtt	8040
gaacagatgtt aaatttgcctt aaggtgtttt ggtgtggaa gggttgggg ttgacaatgg	8100
gaccagatgtt atggtcttag aggttggc tggactgcat caagttccag ggctgtgtgt	8160

gtgttatctg caaacaagg tcaattgtgt ctggaggcct tagtaaaat tggaaggatg	8220
ccccacatag taaaaatgta tcagccaggg gaagtgacta cactgtatct aacctgaaac	8280
agctgagctg taagccagca gctgtacta tggcagggt tggtccgctg gttctgggt	8340
ggtcacttgt atccagttt taggaagtg ttgtcattgc ttgttaggaa gacaacacat	8400
ctcaggctgg gcagtggtag tgcgtgcctc taatgccagc attcccagca cggaaatggc	8460
agaggcagga ggacagcctg gaataacaac ccagggcaga gccccatct cggaaaagac	8520
aaaaaccaga aagtgcattaa aacattgaca caaagggtct caaatattcc ttcatgttt	8580
tttagagattc cactgtcagc ttggcatggc ctctagttag acatccgtac ttggtcccct	8640
gcgttccaag gtcaaggagaa tgatagccac agaacgttcc ctcaagctgtat ggctggagaa	8700
ccggggtccc tgagccccca ccctcacacc catgtgcagg agggagcttc acctttccct	8760
ccgagcagtg tctgccttca ggacaccgct ctcagccag acactgagtc ttctttgtgt	8820
gcatttcccc gggaaaagc tacacattcc ctgtacagca ggcaggaggg cagctgttag	8880
ctcatggaca cgggacagaa ggccaggtac cctgcctccc tgcggtaacc cgcctccctt	8940
cgagccctta ctaggaagca aagccttctc caggtgtacg tgcgtcgaa agtctgcaga	9000
aagtccccggaa ctcccaagatg agtctgctgc acagtgcctg ggccagttttag caccagaaca	9060
aaagtccaag atagccaagg agaacaggca ggtcccacaa tcacattcct tacaaacgga	9120
gaagggagga tcacaagtgt tctggcagcg tttcaagat agagaagtag caaacagtca	9180
gagtgaaagaa ggcaggagtg tcccaatgtc atgggatat ccaggaaggg tggggccaag	9240
cctagctcca gacacccagg gagccccggc gaggagacag ccccaggggc aggcttcatg	9300
cagtggcaca aaagaccccc accccggct cggtgagctc cagggagccca ggcactggc	9360
cctatctcat ggagctatca gatgagacat ggcaatcgga gtcctcagtt tgcgttggcg	9420
ctggcggcgg gtcgctaagc gggaccgcag tgaaaggcagg agactttcta gaaaaaaaca	9480
ccagatgtca cttggggca agcaggaagc ccaagcaggc ctgcgtgc gctctcctt	9540
ccatggcctt atttctggg gtgcagctgg atgcaggca gggaaatggg tggaaagact	9600
tgtccccatt cacctctatg aggaacagtt cccctttgg aggctctgag aggctcagag	9660
atcaaaaatg tctcaagtca cagcatggag atgcttcctg agcccacagt gtgcgtgacac	9720
cagggccagc ttcttgcctt gctgtaaagg ggggtgggct gtctcaactat gttcccaagg	9780
ctgacccccc acacttggta caaagcaatc ctctgtcc agagtagctg ggtatgacaag	9840
caaaggccac taaacatgtt gaaatgtc cacactgtgc ctagcttaat gtaattgggg	9900

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13680
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13740
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13800
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13860
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13920
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	13980
nnnnnnnnnn nnnnnnnnnnac acacacacac acacacacac acacacttac ttccgtgagga	14040
agactgcagg gaacaaggct agatagaacc cattattcca ggagacaaaa gtgaagcaaa	14100
aatggctct taaagcacaa agcagcacta ccaacagtgg tctcttccag cctatatatc	
accctagttt acacattgaa gggaaagatg tacattctag aacatttca acaagttagc	14160
ccccaaattc cagactcagc atgggagatc agagcgttag aggttagatgg aggagataga	14220
ggggccaggg ccatggagga aaactccac ctgtcatcag ctaccatc caataatgg	14280
cagctctacc tcagtaacac tgagagatta agggccatac agacaaaagg aactttgcc	14340
acaaagctac caagtgaact ttcttacact ttctctcaa gtcatttttc caggaaggac	14400
agaagtacaca caacttagaa acaacttcc atccatgtgc agtatacata ttttatgt	14460
tgtggataca gatgtgcattgtgtggta catgggaag cctgatgtcg atgcaggaa	14520
tcttcctagg tggttcttat actctgagtc tggtaacgg cattgggtgt cctggactt	14580
gatctgtaga ccagactggc ctcaaactca cagagatcta cctgtcttg cctccgaat	14640
gctgggatta aaagcgtgtg cctcagccgg acgttggtagt cgcattgcctt taatcccagc	14700
actcgggagg cagaggcagg agatctctg tgagttctag accagcctgg tctacannn	14760
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	14820
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn ctatccag	14880
ggcagcctcc aaagccacag agaaaccctg tctcaaaaaa caaaacaaaa ataaataaat	14940
aaataaataaa aatctgcgc ctccactgcc cggctcagaa tattttaaa gacaaggatt	15000
atgttagccca ggctggcctc aaacttggta tggtaactag atgtggtagtgg cctctgtatcc	15060
tcctgcctct acctcccaag tggtaggact ataggctggc actaccacac ccggtttcat	15120
gtgggctagg atcacactca ggcaggcatt catgcacatt gggcaggcat tctaccaact	15180
gagttacaca gtcccagccc aggtgtctc accagatcac ggtagcact aggacacgt	15240
ggctctgaggt ttgcaccaca tggcagcccc agtcaactgag ctaaggtgaa actcatactt	15300
gactaaaggc acggaatctg aaagggtgg gtacaaccca ggaggaaat cggcggaaagc	15360

aataagaggc gtcaacaggc aatggtcaact tttcccaga caagtctcag ctggaaggcag	15420
cacacgggct gggtatctga aactcagaag tttcctgggt taaaaacaaa agccctggtt	15480
tttcttagga tagtatgttc atttgatgtg tgctggaca ctgagaacac aggaaaataa	15540
cttccaaagt ttgtctcaa aatgcagaca gggtaaccc ttacaggct tgcagggagg	15600
ctgataacaag aggtctcctg cctcattcac ccatcagagt ccaagacaca gtcctcacac	15660
agccccaccc aaccaactgt cctccgacgg tgacccctc cctacaaacc tgcctgcagc	15720
ccagaaggac agtcagagcc ccagctccat agccatttg aaaggactcc ccaggctcaa	15780
aaacaaaagg gagactcctc caggagctca caggctgagc caaggaggca ggagtgtcc	15840
ggcactttagg gtgtccgggt gataatgcac aatatgtatg atgtgtcaac ctgcctcagatc	15900
ctgtcctaag cggcacgtgt ctatctgatc atcctaataa cccagcctgc ggggagagac	15960
atgcacatct cttctgaca aataaagagg caggeccaag ctttagcatgg cagtacacat	16020
ctgtaaccca gcactcaaca ggcagagcag gaggattgac tgcctcctat tccaggccag	16080
tttgctaca cagaaagacc ctgtactgat gctaagagag atggctcaga gttaaagaca	16140
cttccctgctc ctggagagga cccaggttt cttccagca cccatgtcag gtcacatgtc	16200
gcctgtgaac accagctcca ggggatctga tgcccttcc tggcctccag gggcacgcgc	16260
gtgtgcgcctc acacacacac acacacacac acacacacac acacacacac	16320
acagagtgca aagctgaagt gcagctcaca cccataatcc tagcaaattgg agacaaaagg	16380
accaggagtt cgaggccaaac cttagactaca agacaagtac ttgtaaataa atacatatat	16440
acgaaaataa acaaaccaaa aataatgaaa caggcactca gatcaaggga gccagtattg	16500
gtggcacacg ctttaatcc cagcactcag gaggcagagg caggtggatc tctgagttt	16560
aggccagcct agtctacaga atgagatcca ggactacaca gagaaccct gtctcaccct	16620
cctctccctt aaaaaaaatt aagaaagaaa agaaagagcc agggctgggt aagcaatgga	16680
tcagcaggtt aaggcacaag ccaccaagcc caacaaccctt aatttgcattt ctggacacta	16740
catggtagaa ggagagaact cactgcaagc tgcctctga cttccacaca cacccatgg	16800
caggcacatg ctctccccgc cacatacata catacataca tacatacata catacataca	16860
tacatacatg taaacagaaa agaaagttagg aagacaaaag gttcatttca agtggtgaca	16920
gagccaagat ttgaacccac atctatctgc gtctagatcc tgaggtctca agggccatt	16980
agagctacca aaactggggt ggggtgcact ctcctgtca acaggtggaa gtcctggc	17040
ttggccttg gctggagga ggccagccag cagcacaccc ccagcctgggt ttgtgagccc	17100
tgggaacaga gccctgcca aacccttac	17130

<211> 3020

<212> DNA

<213> Artificial Sequence

<220><223> amplicon

<400> 30

tacggccggc ttcactgtac agtggcacat acttgtacc acagtgtcat actggacatt	60
ttcatgaccc taaggccaca tcacccctgt ggtgtcactc agttgttaatt tcaccctata	120
tagtgatttc agttggctt tttacttaga atataacatt taaattatgt ttcttttg	180
gttttgagtt ttgttttcg agacagggtt tctctgtatt gctttggagg ttgtgctgga	240
acttgctctg gagaccaggc tggcttcgaa ctcacagaga tcctccacc tctgcctccc	300
gagtgctggg attaaaggcg tgcgccacca atgcccagc atcttctgt ttctctgttt	360
ttaaagaggg agataatgtt tccagtgtc ttttaactt tttcctgtta ggctaccata	420
gatgaatttgc agcagaagct tcggcatgt cataccagag gcatggatgg aatagaggag	480
cttggaaactg gccagaagg cagccagcaa gccctgtctg ccaagaaacc ctctgctggt	540
atgagagaca agcccacacg gactccccac aaatcatccc tgcattttat tggttttgt	600
ttttaaaatc acatctccct gccatttcta gtctcaaggc gacagcctct gacttctata	660
gattcagaca atgactttgt cacacctaag cctcgacgc ccaaccgtca tcgtccaaac	720
actcagcaac gaaagtccaa gaagaaagcc aaagttgtct tttcaagtga tgagtccagt	780
gaggatggta tgactcgccc tcccgccagc tgctgtctg cttgggtc tcctggg	840
ttgcctgggt gttagagtcat tcttcctggta aatgtgttcc ccactgtgca cgttggctca	900
cctttgcagc tctatgtttt aagatccact ttacaccatc tttgtgactc agttttgtt	960
cctataattc ttgaacaga actttcagcg gaaatgacgg aagaagagac tcccaagaga	1020
accacccca tcccgagatc atctgcccga agacacaggt cttaggagtt gcctcaccta	1080
ttcccatccc tgctgtgcag ggtatcttaa ttgattttta attcattcc tcccttgaa	1140
aaatgttgcc tcctgtttt taatatgtaa atggcttttgc cattcaaggc ctttctctta	1200
gagcccatct gactgctgtct ttcactttcc aacaaacact taaccaatca ccatgtttc	1260
taagaataaa tggggatata tacttttatt ctggctgtt ctgtttcatc ttccaccta	1320
gcccagattc aggcagcaat agagggttat caaaccagag agaagataca ccagccagg	1380
catccctgagg ctgtctctgg gtcttcatttgc attcatcttgc cacaactgt	1440
gagactccac ttagagaat tctcacttct cagcctgggg gtactttcta cccagagacc	1500
tatttctgtc atagcctttt gggacacata aagttccctt cctgcata gacacccaca	1560

acagtgtatc aggagtgtac aagttgacaa acaatgctct agacctacgg ttttcttc	1620
ctgtgtgcct catccagaa gagatcatga ctccaggag tcagcctta ctatgggtc	1680
tgcagggcg tccagccct cagcggcaag ccatgcccac cctcccaag tcctaatct	1740
gctgagtac ttggAACAGG agacactgtat tctgtgtca tgacaacagc acatgtccat	1800
agaaatgttc cctaccttt acgtgtgtt tggggaaca gtaatgacaa accataggca	1860
ggaggcaaaa gggaaGACGG cacctcagaa acatgtgtt ggtagggca gaactatgga	1920

ggggctctg agactcttg atggaaagg gttaatgtc ctctgaaac ctctgttgg	1980
aggcagaaaa gggacaggc tgagtccccg cactggacc attccatcc tctgcaccc	2040
ccccccggct catggaaagc ctggcatgg gccacacagc tgcgtctt ggctctggg	2100
ccccaggag gttaggcaat cccagaatgg caaggagcca ggactggatt tgggtgcag	2160
cccaagctgc tccctgcctt ttaagcaagg gttatcacca ggccagctaa acttagcaat	2220
taggctttc agctaaaaga gcaggggct ggtctcaagt tgcactgacc tagcaaagag	2280
cccccaggat cccctgccc agcacctgtg gctgagctcc caagcccttc ccgagagctc	2340

aggatccacc cttccaccc tccctactct tcagaggagg aaccccttt ctcctccca	2400
cttggggag ggggctgggg ccaggctgtt ctggcttggg gtataatacc ccctaccct	2460
tctactttcc cctcctctca gacctcaccc tgcctccacg agggcagcca agaaaggaga	2520
gtccctggct gcagggccag taggcacgtc ccaggacggg gagggacttc cgcctcacg	2580
tccagctctc cggccctgggg ctgcagtggg tgaaagggc agtgtctctt agcctggcg	2640
gtgcaaccct caggttccga ggaggaacgc tctggagggc ttctttgcct cctccaaccc	2700
aaccacaac caggacattt tctcacccccc gggcccaaa cctagacctt aactgagggaa	2760

cacagaggcc agttttaag tctcaattat gcagggcatc ccgacctgtg gcgtagggag	2820
cgccttcca gggcgcttcc ctagcctctt cctggccctc acagcccagg cctctggccc	2880
aagaaatgga agtgggggtg gggatggaa ctgcgaatgc gaagggcccc cgcaggaggg	2940
aaagtgaccc ctccggggcc tttctgttc cgagacttgt tttgcctgt gtcactaccg	3000
aagaaccacg gccggcctga	3020

<210> 31

<211> 3537

<212> DNA

<213> Artificial Sequence

<220><223> amplicon

<400> 31

tactcgcgat tcactgtaca gtggcacata cttgtcacca cagtgtcata ctggacattt	60
tcatgaccct aaggccacat caccctgtg gtgtcactca gttgttaattt caccctataat	120
agtgatttca gtttgcctt ttacttagaa tataacattt aaattatgtt tctttttgg	180
ttttgagttt tggtttcgaa gacagggtt ctctgtattt cttggaggt tggctggaa	240
cttgctctgg agaccaggct ggcttcgaac tcacagagat cctcccacct ctgcctcccg	300
agtgtggaa ttaaaggcgt gcccacaa tgccccagca tcttctgtt tctctgtttt	360
taaagaggaa gataatgttt ccagtgcgtt tttaacttt tttcctgttag gctaccatag	420
atgaatttga gcagaagctt cgggcatgtc ataccagagg catggatgga atagaggagc	480
ttgaaactgg ccaagaaggc agccagcaag ccctgtctgc caagaaaccc tctgtggta	540
tgagagacaa gcccacacgg actccccaca aatcatccct gtcatttattt gttttttttt	600
tttaaaatca catctccctg ccatttctag tctcaaggcg acgcctctg acttctatag	660
attcagacaa tgactttgtc acacctaagc ctgcacgcac caaccgtcat cgtccaaaca	720
ctcagcaacg aaagtccaag aagaaagcca aagttgtctt ttcaagtgtat gagtccagtg	780
aggatggat gactgcctt cccggcagct gcttgcgttgc ctgggtct cttgggct	840
tgcctggttt tagagtcattt cttctggta atgtgttccc cactgtgcac gttggctcac	900
ctttgcagct ctatgttttta agatccactt tacaccatct ttgtgactca gctttgttc	960
ctataattct ttgaacagaa ctttcagcgaa aatgacgga agaagagact cccaaagagaa	1020
ccaccccat ccgcagatca tctgcccga gacacaggctt ttaggatgttgc cttcacctat	1080
tcccatccct gctgtgcagg gttatcttaat tgattttaa ttcatcttccct ccctgtaaa	1140
aatgttgctt cttgtttttt aatatgtaaa tggtcttgc attcaaggcc tttctctatag	1200
agcccatctg actgctgctt tcactttcca acaaacactt aaccaatcac catgtttct	1260
aagaataaaat gttttatattt acttttattt tttgctggtc tgttcatct tccaccttag	1320
cccgaggattca ggcagcaata gagggttattt aaaccagaga gaagatacac cagcccaggc	1380
atcctgaggc tgcgtctggg tttcatgtt tcctcatttga ttcatcttgc acaactgtatg	1440
agactccact gtagagaattt ctcacttctc agcctggggg tactttctac ccagagacct	1500
attttctgtca tagccttttggg gacacataa agcttccttc ctgcataagaa caccccaaa	1560
cagtgtatca ggagtgtaca agttgacaaa caatgtctt gacctacggt tttcttctc	1620
tgtgtgcctc atccagaag agatcatgac tcccgaggat cagccttac tatgggtct	1680
gcagggcgtt ccagcccttc agccggcaagc catggccacc ctccccaaatg ccttaatctg	1740

ctgagtcact tggAACAGGA gacactgatt ctgctgtcat gacaacagca cattgccata	1800
gaaatgctcc ctaccttta cgtgtgtggt gggggAACAG taatgacaaa ccataggcag	1860
gaggcaaaag ggaAGACGGC acctcagaaa catgtgttag gtttagggcag aactatggag	1920
gggccttiga gactcttga tggAAAGGG ttaatgctgc tcctgaaacc tctgtigaa	1980
ggcagaaaag ggacaggcgt gagtccccgc actgggacca tttccatct ctgcacctg	2040
cccccggtc atggAAAGCC tggcatggg ccacacagct gtcagtcgt gctctgggc	2100
cccaaggagg tagggcaatc ccagaatggc aaggagccag gactggattt ggggtgcagc	2160
ccagcctgct cccctgcctt taagcaaagg ttatcaccag gccagctaaa ctttagcaatt	2220
aggctcttca gctaaaagag cagggggctg gtcicaagtt gcactgaccc agcaaaggagg	2280
ccccaggatc cccctgccc gcacctgtgg ctgagctccc aagcccttcc cgagagctca	2340
ggatccaccc ttccacccct ccctactttt cagaggagga acccccttcc tccttccac	2400
ttgttggagg gggctgggc caggctgttc tggctgggg tataataccct cctaccctt	2460
ctactttccc ctccctctc tagtcacccct gcctccacga gggcagccaa gaaaggagag	2520
tccctggctg cagggccagt aggcacgtcc caggacgggg agggacttcc gcccctacgt	2580
ccagctctcc gccctgggc tgcagtgggt gaaAGGGCA gtgtctccta gcctggcgg	2640
tgcaaccctc aggttccgag gaggaacgct ctggggaggct tctttgcctc ctccaaaccca	2700
acccacaacc aggacattgt cctcaccccg gggccccaac cttagacccatc actgaggaac	2760
acagaggcca gtttgaatgt ctcaattatg caggcatcc cgacctgtgg cgttagggagc	2820
gccctccag gcccttccc tagcctcctc ctggccctca cagcccaggc ctctggccca	2880
agaaatggaa gtgggggtgg gggatgaaac tgcgaatgca aaggcccccc gcaggaggca	2940
aagtgacccc tccgggcct ttctgctcc gagacttgg tttgcctgtc tcactaccga	3000
agaaccacga gaagatcctc aactttcca cagccttgc ataaaggggaa gagggctggc	3060
ggcgcagctg tggcacacac gcacttcgtc tcaacccgc cccccccgc cccgttccctg	3120
tcccttccca gtttccccc atttatcg ggcggcaact ttttaggtccc tgggtctgg	3180
aagtcccttag tacacactt tcgtccttaa gtccatagtc tttttccctt cggtcctatc	3240
ctgtccccca tcaccgggtc acctccccag cgaagcaatc tcaaggcccc tcccccttc	3300
agccccgagc ccacacgttt ggtgcgtgca cattcaaaa acgaggcggg tccaaagaga	3360
gggggtgggg aggtgccgag tggcccaagct actcgccggtt ttacgggtgc acgtagctca	3420
ggcctcagcg cccttgcgt gtgactggat ggatgagcgg ggcgggaggc gggcgcagcg	3480
tcctcgccgc tccccaccac cccagttctt ataaatagca gagctcggtt agtgaac	3537

<211>	537	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	amplicon	
<400>	32	
tactcgcgaa gaagatcctc aactttcca cagccttgc ataaagggga gagggtcggc		60
gggcgcagctg tggcacacac gcacttctgc tcaacccgcc cccccccgcc cccgttctgc		120
ttccttccca ggttctcccc atttatcg ggcggcaact tttaggtccc tgggtctgg		180
aagtcccttag tacacactct tcgtccttaa gtcctatgtc tgtattccct cggctctatc		240
ctgtcccca tcaccgggtc acctccccag cgaagcaatc tcaagttcccc tccccctctc		300
agccccgagc ccacacgtt ggtgcgtgca catttcaaaa acgaggcgaa tccaaagaga		360
gggggtgggg aggtgcgcag tggcccgact actcgcggct ttaggtccc acgttagctca		420
ggcctcagcg cccttgcgt gtgactggat ggatgagcgg ggcgggagggc gggcgagcg		480
tcctcggcgc tccccaccac cccagttcct ataaatagca gagctcggtt agtgaac		537
<210>	33	
<211>	30	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer GlnPr1896	
<400>	33	
tacggccggc ttcaactgtac agtggcacat		30
<210>	34	
<211>	30	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer GlnPr1897	
<400>	34	
tcaggccggc cgtggttctt cggttagtgac		30
<210>	35	
<211>	35	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Primer GlnPr1901	

<400>	35	
tactcgcgaa gaagatcctc aactttcca cagcc		35
<210>	36	
<211>	40	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> Primer GInPr1902		
<400>	36	
gttcaactaaa cgagctctgc tatttatagg aactgggtg		40
<210>	37	
<211>	40	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> Primer GInPr1903		
<400>	37	
caccccagtt cctataaata gcagagctcg tttagtgaac		40
<210>	38	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> Primer GInPr1904		
<400>	38	
cgcttagcacc ggtcgatcga		20
<210>	39	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> Primer GInPr1905		
<400>	39	
tactcgcgat tcactgtaca gtggcacata c		31