



(19) 대한민국특허청(KR)

(12) 공개특허공보(A)

(11) 공개번호

10-2019-0092471

(43) 공개일자

2019년08월07일

(51) 국제특허분류(Int. Cl.)

C07K 14/005 (2006.01)

A61K 39/00 (2006.01)

A61K 39/235 (2006.01)

A61P 31/14 (2006.01)

A61P 31/18 (2006.01)

C12N 15/86 (2006.01)

(52) CPC특허분류

C07K 14/005 (2013.01)

A61K 39/235 (2013.01)

(21) 출원번호

10-2019-7018883

(22) 출원일자(국제)

2017년12월07일

심사청구일자

없음

(85) 번역문제출일자

2019년06월28일

(86) 국제출원번호

PCT/IB2017/057738

(87) 국제공개번호

WO 2018/104911

국제공개일자

2018년06월14일

(30) 우선권주장

1620968.6 2016년12월09일 영국(GB)

(71) 출원인

글락소스미스클라인 바이오로지칼즈 에스.에이.

벨기에왕국 릭센사르트 (비-1330) 루 드 린스티튜트 89

(72) 발명자

아멘돌라, 버지니아

이탈리아 00128 로마 비아 디 카스텔 로마노 100

카쁘네, 스테파니아

이탈리아 00128 로마 비아 디 카스텔 로마노 100

(뒷면에 계속)

(74) 대리인

특허법인 남앤남

전체 청구항 수 : 총 71 항

(54) 발명의 명칭 아데노바이러스 폴리뉴클레오티드 및 폴리펩티드

(57) 요약

본 발명은 신규한 침팬지 아데노바이러스 ChAd157로부터 유래된 단리된 폴리뉴클레오티드 및 폴리펩티드 서열, 및 상기 폴리뉴클레오티드 및 폴리펩티드 서열을 포함하는 재조합 폴리뉴클레오티드, 벡터, 아데노바이러스, 세포 및 조성물에 관한 것이다.

(52) CPC특허분류

A61P 31/14 (2018.01)

A61P 31/18 (2018.01)

C12N 15/86 (2013.01)

A61K 2039/572 (2013.01)

C12N 2710/10343 (2013.01)

C12N 2740/16234 (2013.01)

C12N 2760/20134 (2013.01)

(72) 발명자

폴로까, 스테파노

이탈리아 00198 로마 비아 살라리아 73

폴고리, 안토넬라

이탈리아 00128 로마 비아 디 카스텔 로마노 100

메로네, 로셀라

이탈리아 00128 로마 비아 디 카스텔 로마노 100

명세서

청구범위

청구항 1

하기로 구성된 군으로부터 선택되는 폴리펩티드를 인코딩하는 단리된 폴리뉴클레오티드:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드, 및
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 2

하기로 구성된 군으로부터 선택되는 폴리뉴클레오티드를 포함하는 재조합 폴리뉴클레오티드:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드, 및
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드.

청구항 3

하기로 구성된 군으로부터 선택되는 폴리뉴클레오티드를 포함하는 재조합 벡터:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드, 및
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드.

청구항 4

하기로 구성된 군으로부터 선택되는 적어도 하나의 폴리뉴클레오티드 또는 폴리펩티드를 포함하는 재조합 아데노바이러스:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드,
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드,
- (c) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드, 및
- (d) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 5

하기 중 적어도 하나 및 약학적으로 허용되는 부형제를 포함하는 조성물:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드,
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드,
- (c) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드,
- (d) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체,

- (e) 제3항에 따른 벡터, 및
- (h) 제4항에 따른 재조합 아데노바이러스.

청구항 6

하기 중 적어도 하나를 포함하는 세포:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드,
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드,
- (c) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드,
- (d) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체,
- (e) 제3항에 따른 벡터, 및
- (f) 제4항에 따른 재조합 아데노바이러스.

청구항 7

하기로 구성된 군으로부터 선택되는 단리된 아데노바이러스 폴리펩티드:

- (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드, 및
- (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 8

제1항 내지 제6항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하고, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 9

제1항 내지 제6항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 10

제9항에 있어서, 폴리뉴클레오티드가 SEQ ID NO: 2에 따른 서열을 갖는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 하기를 인코딩하는 폴리뉴클레오티드를 추가로 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포:

- (a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드; 또는
 - (b) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 60% 동일한 아미노산 서열을 갖는, 기능성 유도체,
- 또는
- (a) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드; 또는
 - (b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 60% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 12

제11항에 있어서, 하기를 인코딩하는 폴리뉴클레오티드를 추가로 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포:

(a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드; 또는

(b) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 98% 동일한 아미노산 서열을 갖는, 기능성 유도체,
또는

(a) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드; 또는

(b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 98% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 13

제1항 내지 제10항 중 어느 한 항에 있어서, 하기를 인코딩하는 폴리뉴클레오티드를 추가로 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포:

(a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드; 또는

(b) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 98% 동일한 아미노산 서열을 갖는, 기능성 유도체,
및

(a) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드; 또는

(b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 98% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 14

제13항에 있어서, 하기를 인코딩하는 폴리뉴클레오티드를 추가로 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포:

(a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드; 또는

(b) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 98% 동일한 아미노산 서열을 갖는, 기능성 유도체,
및

(a) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드; 또는

(b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 98% 동일한 아미노산 서열을 갖는, 기능성 유도체.

청구항 15

제11항 내지 제14항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 SEQ ID NO: 4에 따른 서열을 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 16

제11항 내지 제15항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 SEQ ID NO: 6에 따른 서열을 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 하기 중 적어도 하나를 포함하는 폴리뉴클레오

티드, 벡터, 아데노바이러스, 조성물 또는 세포:

- (a) 아데노바이러스 5'-말단, 바람직하게는 아데노바이러스 5' 역 말단 반복;
- (b) 아데노바이러스 E1A 영역, 또는 E1A_280R 및 E1A_243R 영역으로부터 선택되는 이의 단편;
- (c) 아데노바이러스 E1B 또는 IX 영역, 또는 E1B_19K, E1B_55K 또는 IX 영역으로 구성된 군으로부터 선택되는 이의 단편;
- (d) 아데노바이러스 E2b 영역; 또는 E2B_pTP, E2B_중합효소 및 E2B_IVa2 영역으로 구성된 군으로부터 선택되는 이의 단편;
- (e) 아데노바이러스 L1 영역, 또는 이의 단편으로서, 상기 단편이 L1_13.6k 단백질, L1_52k 및 L1_IIIa 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L1 영역, 또는 이의 단편;
- (f) 아데노바이러스 L2 영역, 또는 이의 단편으로서, 상기 단편이 제3항에 따른 L2_헨톤 단백질, L2_pVII, L2_V, 및 L2_pX 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L2 영역, 또는 이의 단편;
- (g) 아데노바이러스 L3 영역, 또는 이의 단편으로서, 상기 단편이 L3_pVI 단백질, 제2항에 따른 L3_핵소 단백질 및 L3_프로테아제로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L3 영역, 또는 이의 단편;
- (h) 아데노바이러스 E2A 영역;
- (i) 아데노바이러스 L4 영역, 또는 이의 단편으로서, 상기 단편이 L4_100k 단백질, L4_33k 단백질 및 단백질 L4_VIII로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L4 영역, 또는 이의 단편;
- (j) 아데노바이러스 E3 영역, 또는 E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, 및 E3 ORF9로 구성된 군으로부터 선택되는 이의 단편;
- (k) 아데노바이러스 L5 영역, 또는 이의 단편으로서, 상기 단편이 제1항에 따른 L5_섬유 섬유 단백질을 인코딩하는, 아데노바이러스 L5 영역, 또는 이의 단편;
- (l) 아데노바이러스 E4 영역, 또는 E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, 및 E4 ORF1으로 구성된 군으로부터 선택되는 이의 단편;
- (m) 아데노바이러스 3'-말단, 바람직하게는 아데노바이러스 3' 역 말단 반복; 및/또는
- (n) 아데노바이러스 VAI 또는 VAII RNA 영역, 바람직하게는 ChAd157가 아닌 아데노바이러스, 더욱 바람직하게는 Ad5로부터의 아데노바이러스 VAI 또는 VAII RNA 영역.

청구항 18

제17항에 있어서, 폴리뉴클레오티드가 하기 중 적어도 하나를 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포:

- (a) 아데노바이러스 5'-말단, 바람직하게는 아데노바이러스 5' 역 말단 반복;
- (e) 아데노바이러스 L1 영역, 또는 이의 단편으로서, 상기 단편이 L1_13.6k 단백질, L1_52k 및 L1_IIIa 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L1 영역, 또는 이의 단편;
- (f) 아데노바이러스 L2 영역, 또는 이의 단편으로서, 상기 단편이 제3항에 따른 L2_헨톤 단백질, L2_pVII, L2_V, 및 L2_pX 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L2 영역, 또는 이의 단편;
- (g) 아데노바이러스 L3 영역, 또는 이의 단편으로서, 상기 단편이 L3_pVI 단백질, 제2항에 따른 L3_핵소 핵소 단백질 및 L3_프로테아제로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L3 영역, 또는 이의 단편;

(i) 아데노바이러스 L4 영역, 또는 이의 단편으로서, 상기 단편이 L4_100k 단백질, L4_33k 단백질 및 단백질 L4_VIII로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L4 영역, 또는 이의 단편;

(k) 아데노바이러스 L5 영역, 또는 이의 단편으로서, 상기 단편이 제1항에 따른 L5_섬유 섬유 단백질을 인코딩하는, 아데노바이러스 L5 영역, 또는 이의 단편;

(m) 아데노바이러스 3'-말단, 바람직하게는 아데노바이러스 3' 역 말단 반복.

청구항 19

제1항 내지 제16항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 아데노바이러스 VAI 또는 VAII RNA 영역을 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 20

제19항에 있어서, VAI 또는 VAII RNA 영역이 ChAd157이 아닌 아데노바이러스로부터 유래되는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 21

제20항에 있어서, VAI 또는 VAII RNA 영역이 Ad5로부터 유래되는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 22

제1항 내지 제16항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 이의 전장에 걸쳐 SEQ ID NO: 15 또는 22로 본질적으로 구성되는 참조 서열과 적어도 95% 동일한 폴리뉴클레오티드를 포함하거나 이로 구성되는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 23

제22항에 있어서, 폴리뉴클레오티드가 이의 전장에 걸쳐 참조 서열과 적어도 99% 동일한 폴리뉴클레오티드를 포함하거나 이로 구성되는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 24

제23항에 있어서, 폴리뉴클레오티드가 이의 전장에 걸쳐 참조 서열과 적어도 99.5% 동일한 폴리뉴클레오티드를 포함하거나 이로 구성되는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 25

제22항 내지 제24항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 이의 전장에 걸쳐 참조 서열과 동일한 폴리뉴클레오티드를 포함하거나 이로 구성되는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 26

제25항에 있어서, 참조 서열이 SEQ ID NO: 15인 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 27

제25항에 있어서, 참조 서열이 SEQ ID NO: 22인 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 28

제1항 내지 제16항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 E1A, E1B, E2A, E2B, E3 및 E4로 구성된 군으로부터 선택되는 유전체 영역의 적어도 하나의 유전자를 비-기능성이 되도록 하는 돌연변이 또는 결실을 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 29

제28항에 있어서, 폴리뉴클레오티드가 E1A, E1B, E2A, E2B, E3 및/또는 E4로 구성된 군으로부터 선택되는 유전체 영역의 적어도 하나의 유전자가 결합된 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 30

제28항 또는 제29항에 있어서, 유전체 영역이 E1A 및/또는 E1B인 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 31

제1항 내지 제16항 중 어느 한 항에 있어서, 폴리뉴클레오티드가 E1 유전체 영역의 결실을 포함하는 폴리뉴클레오티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 32

제4항 및 제8항 내지 제31항 중 어느 한 항에 있어서, 재조합 아데노바이러스가 복제-적격인 아데노바이러스.

청구항 33

제4항 및 제8항 내지 제32항 중 어느 한 항에 있어서, 재조합 아데노바이러스가 복제-부적격인 아데노바이러스.

청구항 34

제4항 및 제8항 내지 제33항 중 어느 한 항에 있어서, 재조합 아데노바이러스가 단백질을 인코딩하는 핵산 서열을 포함하고, 상기 핵산 서열이 숙주 세포에서 상기 단백질의 발현을 유도하는 하나 이상의 서열에 작동 가능하게 연결되는 아데노바이러스.

청구항 35

제34항에 있어서, 단백질이 항원성 단백질 또는 이의 단편인 아데노바이러스.

청구항 36

제35항에 있어서, 단백질이 이중성 단백질 또는 이의 단편인 아데노바이러스.

청구항 37

제36항에 있어서, 단백질이 바이러스로부터 유래되는 아데노바이러스.

청구항 38

제34항 내지 제37항 중 어느 한 항에 있어서, 숙주 세포에서 상기 생성물의 발현을 유도하는 하나 이상의 서열이 전사 개시, 전사 종료, 프로모터 및 인핸서 서열로 구성된 군 중 하나 이상으로부터 선택되는 서열을 포함하는 아데노바이러스.

청구항 39

제38항에 있어서, 숙주 세포에서 상기 생성물의 발현을 유도하는 하나 이상의 서열이 프로모터 서열을 포함하는 아데노바이러스.

청구항 40

제39항에 있어서, 프로모터 서열이 내부 프로모터, 천연 프로모터, RSV LTR 프로모터, CMV 프로모터, SV40 프로모터, 디하이드로폴레이트 환원효소 프로모터, β -액틴 프로모터, PGK 프로모터, EF1a 프로모터 및 CASI 프로모터로 구성된 군으로부터 선택되는 아데노바이러스.

청구항 41

제39항에 있어서, 프로모터 서열이, 예를 들어, SEQ ID NO: 42에 제공된 증강된 hCMV 프로모터인 아데노바이러스.

청구항 42

제4항 및 제8항 내지 제41항 중 어느 한 항에 있어서, 아데노바이러스가 인간 대상체에서 10% 미만의 혈청유행률(seroprevalence)을 갖고, 바람직하게는 인간 대상체에서 혈청유행률을 갖지 않는 아데노바이러스.

청구항 43

제4항 및 제8항 내지 제42항 중 어느 한 항에 있어서, 아데노바이러스가 포유동물 세포를 감염시킬 수 있는 아데노바이러스.

청구항 44

제5항 및 제8항 내지 제43항 중 어느 한 항에 있어서, 무기 애주번트(예를 들어, 무기 금속 염, 예를 들어, 알루미늄 포스페이트 또는 알루미늄 하이드록사이드), 유기 애주번트(예를 들어, 사포닌, 예를 들어, QS21, 또는 스쿠알렌), 오일-기반 애주번트(예를 들어, 프로인트 완전 애주번트 및 프로인트 불완전 애주번트), 사이토카인(예를 들어, IL-1 β , IL-2, IL-7, IL-12, IL-18, GM-CSF, 및 INF- γ) 미립자 애주번트(예를 들어, 면역-자극성 복합체(ISCOMS), 리포솜, 또는 생물분해성 미세구), 바이로솜, 박테리아 애주번트(예를 들어, 모노포스포릴 지질 A, 예를 들어, 3-데-O-아실화 모노포스포릴 지질 A(3D-MPL), 또는 뮤라밀 펩티드), 합성 애주번트(예를 들어, 비-이온성 블록 공중합체, 뮤라밀 펩티드 유사체, 또는 합성 지질 A), 합성 폴리뉴클레오티드 애주번트(예를 들어, 폴리아르기닌 또는 폴리리신) 및 메틸화되지 않은 CpG 디뉴클레오티드("CpG")를 함유하는 면역자극성 올리고뉴클레오티드로 구성된 목록으로부터 선택되는 애주번트를 포함하는 조성물.

청구항 45

제44항에 있어서, 애주번트가 3D-MPL 및/또는 QS21인 조성물.

청구항 46

제6항 및 제8항 내지 제31항 중 어느 한 항에 있어서, 세포가 E1A, E1B, E2A, E2B, E3 E4, L1, L2, L3, L4 및 L5로 구성된 군으로부터 선택되는 적어도 하나의 아데노바이러스 유전자를 발현하는 숙주 세포인 세포.

청구항 47

제45항에 있어서, 숙주 세포가 현탁액에서 성장되는 세포.

청구항 48

제1항 내지 제47항 중 어느 한 항에 있어서, 약제로서 사용하기 위한 폴리뉴클레오티드, 폴리펩티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 49

제48항에 있어서, 백신으로서 사용하기 위한 폴리뉴클레오티드, 폴리펩티드, 벡터, 아데노바이러스, 조성물 또는 세포.

청구항 50

질병의 치료 또는 예방을 위한 제1항 내지 제47항 중 어느 한 항에 따른 폴리뉴클레오티드, 폴리펩티드, 벡터, 아데노바이러스, 조성물 또는 세포의 용도.

청구항 51

제1항 내지 제47항 중 어느 한 항에 따른 폴리뉴클레오티드, 폴리펩티드, 벡터, 아데노바이러스, 조성물 또는 세포를 대상체에게 투여하는 것을 포함하는 대상에서 면역 반응을 유도하는 방법.

청구항 52

SEQ ID NO: 2에 따른 서열을 포함하거나 이로 구성되는 단리된 폴리뉴클레오티드.

청구항 53

제4항 및 제8항 내지 제43항 중 어느 한 항에 있어서, 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터에 이미 노출된 대상체에게 투여하기 위한 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 54

제53항에 있어서, 대상체가 ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터에 이미 노출된 대상체인 재조합 아데노바이러스 벡터.

청구항 55

제54항에 있어서, 대상체가 ChAd155 섬유, 핵손 및 펜톤을 포함하는 재조합 아데노바이러스 벡터에 이미 노출된 대상체인 재조합 아데노바이러스 벡터.

청구항 56

제54항 또는 제55항에 있어서, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터가 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 상이한 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 57

제54항 또는 제55항에 있어서, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터가 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 동일한 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 58

제54항 또는 제55항에 있어서, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터가 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 동일한 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 59

제53항 내지 제58항 중 어느 한 항에 있어서, 트랜스진이 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 60

제4항 및 제8항 내지 제43항 중 어느 한 항에 있어서, 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터에 후속하여 노출될 수 있는 대상체에게 투여하기 위한 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 61

제60항에 있어서, 대상체가 ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터에 후속하여 노출될 수 있는 대상체인 재조합 아데노바이러스 벡터.

청구항 62

제61항에 있어서, 대상체가 ChAd155 섬유, 핵손 및 펜톤을 포함하는 재조합 아데노바이러스 벡터에 후속하여 노출될 수 있는 대상체인 재조합 아데노바이러스 벡터.

청구항 63

제61항 또는 제62항에 있어서, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터가 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 상이한 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 64

제61항 또는 제62항에 있어서, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터가 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 동일한 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 65

제61항 또는 제62항에 있어서, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터가 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 동일한 트랜스진을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 66

제60항 내지 제65항 중 어느 한 항에 있어서, 트랜스진이 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩하는 재조합 아데노바이러스 벡터.

청구항 67

대상체에서 면역 반응을 유도하는 방법으로서, 상기 방법이,

(a) 대상체에게 제1 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터를 투여하는 단계; 및

(b) 대상체에게 ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 벡터를 투여하는 단계로서, 상기 벡터가 제2 트랜스진을 인코딩하는, 단계를 포함하고,

여기서 단계 (a) 및 (b)가 어느 한 순서로 수행될 수 있고 상기 제1 및 상기 제2 트랜스진이 동일하거나 상이할 수 있는 방법.

청구항 68

대상체를 예방 또는 치료하기 위한 방법으로서, 상기 방법이,

(a) 대상체에게 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩하는 제1 트랜스진을 인코딩하는 제4항 및 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터를 투여하는 단계; 및

(b) 대상체에게 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터를 투여하는 단계로서, 상기 벡터가 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 상이한 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩하는 제2 트랜스진을 인코딩하는, 단계를 포함하고,

여기서 단계 (a) 및 (b)가 어느 한 순서로 수행될 수 있는 방법.

청구항 69

제53항 내지 제68항 중 어느 한 항에 있어서, 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터가 제4항 또는 제8항 내지 제43항 중 어느 한 항에 따른 재조합 아데노바이러스 벡터와 낮은 교차-반응성을 갖는 재조합 아데노바이러스 벡터 또는 방법.

청구항 70

제69항에 있어서, 제1 벡터에 의한 면역화가 제2 벡터에 의한 면역화로부터 발생하는 수준의 평균 50% 미만의 중화 역가를 유도하는 재조합 아데노바이러스 벡터 또는 방법.

청구항 71

제53항 내지 제70항 중 어느 한 항에 있어서, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 벡터가

ChAd157 섬유, ChAd157 핵소 또는 ChAd157 섬유 또는 이와 적어도 98% 동일성을 갖는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터 또는 방법.

발명의 설명

기술 분야

[0001] 발명의 분야

[0002] 본 발명은 신규한 침팬지 아데노바이러스 ChAd157로부터 유래된 단리된 폴리뉴클레오티드 및 폴리펩티드 서열, 및 상기 폴리뉴클레오티드 및 폴리펩티드 서열을 포함하는 재조합 폴리뉴클레오티드, 벡터, 아데노바이러스, 세포 및 조성물에 관한 것이다.

배경 기술

[0003] 발명의 배경

[0004] 아데노바이러스는 다양한 표적 조직에서 매우 효과적인 유전자 전달 및 큰 트랜스진(transgene) 용량을 달성하는 이의 능력으로 인해 유전자 전달 용도로 널리 사용되어 왔다. 통상적으로, 아데노바이러스의 E1 유전자는 결실되고, 선택 프로모터, 관심 유전자의 cDNA 서열 및 poly A 신호로 구성된 트랜스진 카세트에 대체되어, 복제 결합 재조합 바이러스를 발생시킨다.

[0005] 재조합 아데노바이러스는 유전자 요법에서 및 백신으로서 유용하다. 침팬지 아데노바이러스에 기반한 바이러스 벡터는 유전학적 백신의 개발을 위한 인간 유래된 아데노바이러스 벡터의 사용에 대한 대안을 나타낸다. 침팬지로부터 단리된 아데노바이러스는 인간 기원의 세포에서의 이들의 효율적인 증식에 의해 입증된 바와 같이 인간으로부터 단리된 아데노바이러스와 밀접한 관련이 있다. 그러나, 인간 및 침팬지 아데노바이러스는 밀접하게 관련이 있어, 두 바이러스 중 사이의 혈청학적 교차 반응성이 가능하다.

[0006] 분자를 표적에 효과적으로 전달하고, 집단 내 선택된 아데노바이러스 혈청형에 대한 기존 면역의 효과를 최소화하는 벡터에 대한 요구가 존재한다. 인간에서 관찰되는 기존 면역의 한 양태는 체액성 면역이며, 이는 아데노바이러스 단백질에 특이적인 항체의 생성 및 지속을 발생시킬 수 있다. 아데노바이러스에 의해 유도되는 체액성 반응은 주로 섬유, 펜톤 및 핵소의 3개의 주요 구조 캡시드 단백질에 대해 유도된다.

발명의 내용

[0007] 발명의 개요

[0008] 단리된 폴리뉴클레오티드가 제공되며, 상기 폴리뉴클레오티드는 하기로 구성된 군으로부터 선택되는 폴리펩티드를 인코딩한다:

[0009] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드; 및

[0010] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체.

[0011] 또한, 하기로 구성된 군으로부터 선택되는 폴리뉴클레오티드를 포함하는 재조합 폴리뉴클레오티드가 제공된다:

[0012] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드; 및

[0013] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드.

[0014] 또한, 하기로 구성된 군으로부터 선택되는 폴리뉴클레오티드를 포함하는 재조합 벡터가 제공된다:

[0015] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드; 및

[0016] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드.

- [0017] 또한, 하기로 구성된 군으로부터 선택되는 적어도 하나의 폴리뉴클레오티드 또는 폴리펩티드를 포함하는 재조합 아데노바이러스가 제공된다:
- [0018] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드;
- [0019] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 폴리뉴클레오티드;
- [0020] (c) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드; 및
- [0021] (d) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체.
- [0022] 또한, 하기 중 적어도 하나 및 약학적으로 허용되는 부형제를 포함하는 조성물이 제공된다:
- [0023] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 단리된 폴리뉴클레오티드;
- [0024] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 단리된 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 단리된 폴리뉴클레오티드;
- [0025] (c) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 단리된 폴리펩티드;
- [0026] (d) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 단리된 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 단리된 기능성 유도체;
- [0027] (e) 상기 (a) 또는 (b)에 기재된 폴리뉴클레오티드를 포함하는 벡터; 및
- [0028] (f) 상기 (a) 또는 (b)에 기재된 폴리뉴클레오티드를 포함하는 재조합 아데노바이러스.
- [0029] 또한, 하기 중 적어도 하나를 포함하는 세포가 제공된다:
- [0030] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 단리된 폴리뉴클레오티드;
- [0031] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 단리된 폴리뉴클레오티드로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 단리된 폴리뉴클레오티드;
- [0032] (c) 상기 (a) 또는 (b)에 기재된 폴리뉴클레오티드를 포함하는 벡터, 및
- [0033] (d) 상기 (a) 또는 (b)에 기재된 폴리뉴클레오티드를 포함하는 재조합 아데노바이러스.
- [0034] 또한, 하기로 구성된 군으로부터 선택되는 단리된 아데노바이러스 폴리펩티드가 제공된다:
- [0035] (a) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드; 및
- [0036] (b) SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는, 기능성 유도체.

도면의 간단한 설명

[0037] 도면의 설명

도 1a-1d - 표시된 유인원 아데노바이러스로부터의 섬유 단백질 서열의 정렬.

ChAd157 (SEQ ID NO:1)

ChAd3 (SEQ ID NO:27)

PanAd3 (SEQ ID NO:28)

ChAd17 (SEQ ID NO:29)

ChAd19 (SEQ ID NO:30)

ChAd24 (SEQ ID NO:31)

ChAd155 (SEQ ID NO:7)

ChAd11 (SEQ ID NO:32)

ChAd20 (SEQ ID NO:33)

ChAd31 (SEQ ID NO:34)

PanAd1 (SEQ ID NO:35)

PanAd2 (SEQ ID NO:36)

도 2 - 하위그룹 C BAC 셔틀(Shuttle) 개략도

도 3 - 하위그룹 C 플라스미드 셔틀 개략도

도 4 - pChAd157 ΔE1/TetO hCMV GAG 벡터 개략도

도 5 - pARS 종C Ad5orf6-2 셔틀 개략도

도 6 - ChAd157 RG를 운반하는 플라스미드 개략도

도 7 - ChAd157/GAG, ChAd19/GAG 및 ChAd155/GAG에 의한 트랜스진 발현

도 8 - ChAd155/RG 및 ChAd157/RG로 감염된 HeLa 세포 용해물의 웨스턴 블롯 분석

도 9 - BALB/c 마우스에서 ChAd157/GAG, ChAd155/GAG 및 ChAd19 GAG의 면역학적 효능

도 10 - BALB/c 마우스에서 ChAd157/RG 및 ChAd155/RG의 면역학적 효능

도 11 - 상이한 ChAd 벡터로 마우스를 사전면역시킨 후 중화 역가

도 12 - 다양한 사전면역 요법 후 ChAd157/GAG로 마우스의 백신화 후 IFN-γ ELISpot

서열의 설명

SEQ ID NO: 1 - ChAd157 섬유의 폴리펩티드 서열

SEQ ID NO: 2 - ChAd157 섬유를 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 3 - ChAd157 펜톤의 폴리펩티드 서열

SEQ ID NO: 4 - ChAd157 펜톤을 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 5 - ChAd157 핵손의 폴리펩티드 서열

SEQ ID NO: 6 - ChAd157 핵손을 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 7 - ChAd155 섬유의 폴리펩티드 서열

SEQ ID NO: 8 - ChAd155 섬유를 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 9 - ChAd155 펜톤의 폴리펩티드 서열

SEQ ID NO: 10 - ChAd155 펜톤을 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 11 - ChAd155 핵손의 폴리펩티드 서열

SEQ ID NO: 12 - ChAd155 핵손을 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 13 - 야생형 ChAd155를 인코딩하는 폴리뉴클레오티드 서열

SEQ ID NO: 14 - 하위그룹 C BAC 셔틀((#1365)의 폴리뉴클레오티드 서열

SEQ ID NO: 15 - pChAd157 ΔE1 TetO hCMV RpsL-Kana#1551의 폴리뉴클레오티드 서열

SEQ ID NO: 16 - HIV Gag 폴리뉴클레오티드 서열

SEQ ID NO: 17 - pChAd157 ΔE1/TetO hCMV GAG#1557의 폴리뉴클레오티드 서열

SEQ ID NO: 18 - Ad5orf6 프라이머 1 폴리뉴클레오티드 서열
 SEQ ID NO: 19 - Ad5orf6 프라이머 2 폴리뉴클레오티드 서열
 SEQ ID NO: 20 - 섬유-E4 polyA 프라이머 1 폴리뉴클레오티드 서열
 SEQ ID NO: 21 - 섬유-E4 polyA 프라이머 2 폴리뉴클레오티드 서열
 SEQ ID NO: 22 - ChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana#1594의 폴리뉴클레오티드 서열
 SEQ ID NO: 23 - 광견병 당단백질 폴리뉴클레오티드 서열
 SEQ ID NO: 24 - pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RG#1559의 폴리뉴클레오티드 서열
 SEQ ID NO: 25 - CMVfor 프라이머 폴리뉴클레오티드 서열
 SEQ ID NO: 26 - CMVrev 프라이머 폴리뉴클레오티드 서열
 SEQ ID NO: 27 - ChAd3의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 28 - PanAd3의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 29 - ChAd17의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 30 - ChAd19의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 31 - ChAd24의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 32 - ChAd11의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 33 - ChAd20의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 34 - ChAd31의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 35 - PanAd1의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 36 - PanAd2의 섬유 단백질에 대한 아미노산 서열
 SEQ ID NO: 37 - hCMV(tetO)의 폴리뉴클레오티드 서열
 SEQ ID NO: 38 - 하위그룹 C 플라스미드 서열#1376의 폴리뉴클레오티드 서열
 SEQ ID NO: 39 - BGH polyA의 폴리뉴클레오티드 서열
 SEQ ID NO: 40 - pARS 종C Ad5orf6-2의 폴리뉴클레오티드 서열
 SEQ ID NO: 41 - CMVFAM-TAMRA 프로브의 폴리뉴클레오티드 서열
 SEQ ID NO: 42 - 증강된 hCMV 프로모터를 인코딩하는 폴리뉴클레오티드 서열

발명을 실시하기 위한 구체적인 내용

[0038]

발명의 상세한 설명

[0039]

본 발명의 벡터, 조성물 및 방법은 높은 생산성, 개선된 면역원성, 증가된 트랜스진 발현 또는 독특한 혈청 교차 반응 프로파일을 비제한적으로 포함하는, 종래 기술에 비해 개선된 하나 이상의 특징을 가질 수 있다.

[0040]

본 발명의 벡터, 조성물 및 방법은 생산성, 면역원성, 트랜스진 발현 및/또는 혈청 교차 반응성과 같은 특성의 조합을 입증할 수 있는데, 이는 이들이 공지된 접근법에 대한 가치있는 대안을 제공한다는 것을 의미한다.

[0041]

아데노바이러스

[0042]

아데노바이러스는 3개의 주요 단백질인 헥손(II), 펜톤 베이스(III) 및 노브형 섬유(knobbed fiber)(IV)와 함께 다수의 다른 작은 단백질들인 VI, VIII, IX, IIIa 및 IVa2를 포함하는 정이십면체 캡시드를 갖는 특징적인 형태를 갖는다. 바이러스 유전체는 선형의 이중 가닥 DNA이다. 바이러스 DNA는 고도의 베이스 단백질 VII 및 작은 캡티드 pX(이전에는, 뮤로 언급됨)와 밀접하게 회합되어 있다. 또 다른 단백질인 V가 상기 DNA-단백질 복합체와 함께 패키징되고, 단백질 VI를 통해 캡시드에 대한 구조적 연결을 제공한다. 상기 바이러스는 또한 성숙한 감염

성 바이러스를 생성하기 위해 일부 구조 단백질의 처리에 필요한 바이러스-인코딩된 프로테아제를 함유한다.

- [0043] 아데노바이러스 유전체는 잘 특성화되어 있다. 유사하게 정위된 특정 오픈 리딩 프레임과 관련하여 아데노바이러스 유전체의 전체적인 구성, 예를 들어, 각각의 바이러스의 E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 및 L5 유전자의 위치에서의 전반적인 보존이 존재한다. 아데노바이러스 유전체의 각각의 말단은 바이러스 복제에 필요한 역 말단 반복(ITR)으로 공지된 서열을 포함한다. 상기 바이러스는 또한 감염성 비리온을 생성하는데 필요한 일부 구조 단백질을 처리하는데 필요한 바이러스-인코딩된 프로테아제를 포함한다. 아데노바이러스 유전체의 구조는 바이러스 유전자가 숙주 세포 형질도입 후에 발현되는 순서에 기초하여 기재된다. 더욱 특히, 바이러스 유전자는 전사가 DNA 복제의 개시 전 또는 후에 발생하는지에 따라 초기(E) 또는 후기(L) 유전자로 언급된다. 형질도입의 초기 단계에서, 아데노바이러스의 E1A, E1B, E2A, E2B, E3 및 E4 유전자가 발현되어 숙주 세포가 바이러스를 복제하도록 준비시킨다. 감염의 후기 단계 동안, 바이러스 입자의 구조 성분을 인코딩하는 후기 유전자 L1-L5의 발현이 활성화된다.
- [0044] 아데노바이러스는 종-특이적이며, 다양한 혈청형, 즉, 항체에 의해 교차-중화되지 않는 바이러스의 유형은 다양한 포유동물 종으로부터 단리되어 왔다. 예를 들어, 서열 상동성 및 적혈구 세포를 응집시키는 능력을 기초로 하여 6개의 하위그룹(A-F; B는 B1 및 B2로 세분됨)으로 나뉘는 50개 초과 혈청형이 인간으로부터 단리되었다(Tatsis and Ertl *Molecular Therapy* (2004) 10:616-629). 다수의 아데노바이러스가 비인간 유인원, 예를 들어, 침팬지, 보노보, 히말라야 원숭이(rhesus macaque) 및 고릴라로부터 단리되었고, 이들은 헥손 또는 섬유서열에 기반한 계통발생학적 관계를 기초로 하여 동일한 인간 그룹으로 분류된다(Colloca *et al.* (2012) *Science Translational Medicine* 4:1-9; Roy *et al.* (2004) *Virology* 324: 361-372; Roy *et al.* (2010) *Journal of Gene Medicine* 13:17-25).
- [0045] W02005071093호는 ChAd19를 포함하는 침팬지 아데노바이러스를 개시한다. ChAd155 유래된 벡터를 정의하기 위한 목적으로 본원에 참조로서 포함된 W02016198621호(PCT/EP2016/063329)는 침팬지 아데노바이러스 ChAd155를 개시한다.
- [0046] **섬유 단백질을 포함하는 아데노바이러스 캡시드 단백질 및 이들 단백질을 인코딩하는 폴리뉴클레오티드**
- [0047] 상기 개시된 바와 같이, 아데노바이러스 캡시드는 3개의 주요 단백질인 헥손, 펜톤 및 섬유를 포함한다. 헥손은 캡시드의 구조 성분의 대부분을 차지하며, 이는 240개의 삼량체 헥손 캡소미어 및 12개의 펜톤 베이스로 구성된다. 헥손은 3개의 보존된 이중 배럴을 가지며, 상단은 3개의 타워(tower)를 갖고, 각각의 타워는 캡시드 대부분을 형성하는 각각의 서브유닛으로부터의 루프를 함유한다. 헥손의 베이스는 아데노바이러스 혈청형 사이에 고도로 보존되어 있는 반면, 표면 루프는 가변적이다(Tatsis and Ertl *Molecular Therapy* (2004) 10:616-629).
- [0048] 펜톤은 섬유가 부착하는 오랑체 베이스를 형성하는 또 다른 아데노바이러스 캡시드 단백질이다. 삼량체 섬유 단백질은 캡시드의 12개의 꼭지점 각각에서 펜톤 베이스로부터 돌출되어 있으며, 노브형 막대-유사 구조이다. 대부분의 다른 정이십면체 바이러스의 것과 비교하여 아데노바이러스 캡시드의 표면에서의 현저한 차이는 길고 얇은 섬유 단백질의 존재이다. 섬유 단백질의 주요 역할은 세포 수용체와의 상호작용을 통해 세포 표면에 바이러스 캡시드를 테더링(tethering)시키는 것이다.
- [0049] 많은 아데노바이러스 혈청형의 섬유 단백질은 N-말단 꼬리, 반복 서열로 이루어진 중심 축, 및 C-말단 구형 노브 도메인(또는 "헤드")의 공통 구조를 공유한다. 중심 축 도메인은 다양한 수의 베타-반복으로 구성된다. 베타-반복은 연결되어 매우 견고하고 안정적인 3개의 얇은 나선형 가닥의 연장된 구조를 형성한다. 상기 축은 N-말단 꼬리와 구형 노브 구조를 연결하며, 이는 표적 세포 수용체와의 상호작용을 담당한다. 아데노바이러스 노브 도메인의 구형 특성은 수용체를 측면 및 정점으로 결합시키기 위한 큰 표면을 제공한다. 이러한 구조의 효과는 수용체-결합 부위를 바이러스 캡시드로부터 멀리 떨어져서 돌출시킴으로써 바이러스를 비교적 편평한 캡시드 표면에 의해 제공되는 입체적 구속으로부터 자유롭게 하는 것이다.
- [0050] 많은 아데노바이러스 혈청형의 섬유는 동일한 전반적인 구조를 갖지만, 이들은 이들의 기능뿐만 아니라 구조에 영향을 미치는 다양한 아미노산 서열을 갖는다. 예를 들어, 섬유 노브의 표면 상의 다수의 노출된 영역은 용이하게 적응 가능한 수용체 결합 부위를 제공한다. 섬유 노브의 구형 형태는 수용체가 노브의 측면 또는 섬유 노브의 상단에 결합되도록 한다. 이들 결합 부위는 통상적으로 인간 아데노바이러스 사이에서 불량하게 보존되는 베타-가닥을 연결하는 표면-노출 루프 상에 놓여 있다. 이들 루프 상의 노출된 측쇄는 3차 및 4차 구조를 유지하면서 노브에 다양한 표면 특징을 제공한다. 예를 들어, 노브 표면에서의 정전식 전위 및 전하 분포는 Ad 8, Ad 19 및 Ad 37에 대한 약 9의 pI로부터 하위그룹 B 아데노바이러스에 대한 약 5까지의 섬유 노브 서열에서의

광범위한 등전점으로 인해 다양할 수 있다. 구조적으로 복잡한 바이러스 리간드로서, 섬유 단백질은 바이러스 캡시드로부터 다수의 배향 및 거리(축)의 다양한 결합 표면(노브)의 제시를 가능하게 한다.

[0051] 일부 혈청형 사이의 가장 명백한 변화 중 하나는 섬유 길이이다. 섬유 축의 길이가 노브 및 바이러스와 이의 표적 수용체의 상호작용에 강하게 영향을 미치는 것이 연구에서 제시되었다. 또한, 혈청형 사이의 섬유 단백질은 또한 이들의 회는 능력에 있어 다양할 수 있다. 축의 베타-반복은 매우 안정적이고 규칙적인 구조를 형성하나, 전자현미경(EM) 연구는 섬유 내에 별개의 힌지를 나타내었다. 여러 아데노바이러스 혈청형 섬유로부터의 단백질 서열의 분석은 EM에 의해 관찰되는 바와 같이 축 내의 힌지 중 하나와 강하게 연관된 N-말단 꼬리로부터의 세 번째 베타-반복에서의 축의 반복 서열에서의 파괴를 정확하게 지적한다. 섬유 내의 힌지는 노브가 바이러스 캡시드와 관련된 다양한 배향을 채택하는 것을 가능하게 하며, 이는 노브 상의 수용체 결합 부위의 정확한 제시를 요구하는 수용체 연동에 대한 입체적 방해로 회피할 수 있다. 따라서, 예를 들어, 하위그룹 D Ad의 단단한 섬유는 이들이 스스로 휘 수 없으므로 바이러스 부착을 위해 가요성 수용체 또는 사전배치된 수용체를 필요로 한다. (Nicklin et al *Molecular Therapy* 2005 12:384-393).

[0052] 다양한 Ad 혈청형에 대한 특정 세포 수용체의 확인 및 이들이 조직 향성에 기여하는 방법에 대한 지식은 섬유 슈도타이핑(pseudotyping) 기술의 사용을 통해 달성되었다. 일부 하위그룹의 Ad가 일차 수용체로 CAR을 이용하나, 많은 Ad가 대안적 일차 수용체를 이용하여 시험관 내 및 생체 내에서 광범위하게 다양한 향성을 발생시키는 것이 명백해지고 있다. 이들 혈청형의 섬유는 이들의 일차 및 삼차 구조에서 섬유 축 강성, 섬유 축 길이, 및 CAR 결합 부위 및/또는 추정 HSPG 결합 모티프의 결핍과 같은 명백한 차이와 함께 섬유 노브 내에서의 순 전하의 차이를 나타낸다. 따라서, 대안적 섬유 축 및 노브로 Ad 5 입자를 슈도타이핑하는 것은 중요한 세포 결합 도메인을 제거할 기회를 제공하며, 또한 Ad 5로 달성된 것에 비해 규정된 세포 유형으로의 더욱 효과적인(및 잠재적으로 더욱 세포-선택적인) 트랜스진 전달을 가능하게 할 수 있다. 섬유-슈도타이핑된 Ad 입자의 중화는 또한 섬유가 인간 또는 실험 모델에서 더 낮은 혈청유병률(seroprevalence)을 갖는 Ad로부터 유래되는 경우에 감소될 수 있고, 이는 백신의 성공적인 투여를 촉진하는 상황이다(Nicklin et al *Molecular Therapy* (2005) 12:384-393). 또한, 핵소 또는 펜톤 단독은 아니지만 전장 섬유뿐만 아니라 단리된 섬유 노브 영역은 수지상 세포 성숙을 유도할 수 있고, 효능 있는 CD8+ T 세포 반응의 유도와 관련이 있다(Molinier-Frenkel et al. *J. Biol. Chem.* (2003) 278:37175-37182). 종합하면, 아데노바이러스 섬유는 아데노바이러스 백신의 적어도 수용체-결합 및 면역원성에서 중요한 역할을 한다.

[0053] 그룹 C 유인원 아데노바이러스의 섬유 단백질 사이의 차이를 예시하는 것이 도 1에 제공된 정렬이다. 이들 아데노바이러스의 섬유 서열이 ChAd157과 같이 짧은 섬유를 갖는 것, 또는 ChAd155와 같이 긴 섬유를 갖는 것으로 광범위하게 그룹화될 수 있다는 것이 두드러진 특징이다. 이러한 길이 차이는 긴 섬유에 비해 짧은 섬유에서 대략 위치 321에서의 36개의 아미노산 결실로 인한 것이다. 또한, 짧은 섬유 하위그룹과 긴 섬유 하위그룹 사이에서는 상이하나 각각의 하위그룹 내에서는 일치하는 다수의 아미노산 치환이 존재한다. 이들 차이의 정확한 기능은 아직 밝혀지지 않았으나, 섬유의 기능 및 면역원성을 고려할 때, 이들은 유의한 것으로 보인다. 바이러스 향성의 결정인자 중 하나가 섬유 축의 길이인 것이 밝혀졌다. 더 짧은 축을 갖는 Ad5 백신이 CAR 수용체에 대한 결합의 더 낮은 효율 및 더 낮은 감염성을 갖는 것이 입증되었다(Ambriovic-Ristov A. et al.: *Virology*. (2003) 312(2):425-33): 이러한 손상은 세포 수용체에 대한 덜 효율적인 부착을 발생시키는 더 짧은 섬유의 증가된 강성의 결과인 것으로 추측되었다(Wu, E et al.: *J Virol*. (2003) 77(13): 7225-7235).

[0054] 본 발명의 일 양태에서, 침팬지 아데노바이러스 ChAd157의 단리된 섬유 폴리펩티드 및 침팬지 아데노바이러스 ChAd157의 섬유 폴리펩티드를 인코딩하는 단리된 폴리뉴클레오티드가 제공된다.

[0055] 섬유 단백질은 낮은 혈청유병률에 기여할 것으로 예상되며, 따라서, 예를 들어, 감소된 혈청유병률을 갖는 재조합 아데노바이러스를 제조하기 위해, 섬유 단백질은 기존의 중화 항체에 대한 아데노바이러스의 친화도를 억제하기 위해 ChAd157로부터의 핵소 및 펜톤 폴리펩티드와 함께(핵소 및 펜톤 중 하나 또는 둘 모두와 함께) 또는 이와 독립적으로 사용될 수 있다. 상기 재조합 아데노바이러스는 ChAd157로부터의 적어도 하나의 섬유 단백질과 함께 상이한 혈청형으로부터의 캡시드 단백질을 갖는 키메라 아데노바이러스일 수 있다.

[0056] ChAd157 섬유 폴리펩티드 서열은 SEQ ID NO: 1에 제공된다.

[0057] ChAd157 펜톤 폴리펩티드 서열은 SEQ ID NO: 3에 제공된다.

[0058] ChAd157 핵소 폴리펩티드 서열은 SEQ ID NO: 5에 제공된다.

[0059] ChAd157 섬유의 폴리펩티드 서열 또는 이의 기능성 유도체를 포함하는 폴리펩티드, 재조합 아데노바이러스, 조

성물 또는 세포

- [0060] 적합하게는, 본 발명의 단리된 폴리펩티드, 재조합 아데노바이러스, 조성물 또는 세포는 SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 포함한다.
- [0061] 본 발명의 폴리펩티드, 재조합 아데노바이러스, 조성물 또는 세포는 SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체인 폴리펩티드를 포함할 수 있고, 상기 기능성 유도체는 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8% 동일한 아미노산 서열을 갖는다.
- [0062] 대안적으로, 기능성 유도체는 SEQ ID NO: 1과 비교하여 하나 초과와 첨가, 결실 또는 치환을 갖지 않고, 예를 들어, SEQ ID NO: 1과 비교하여 하나의 치환을 갖는다.
- [0063] 적합하게는, 본 발명에 따른 폴리펩티드, 재조합 아데노바이러스, 조성물 또는 세포는 하기를 추가로 포함한다:
- [0064] (a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드; 또는
- [0065] (b) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 60% 동일한 아미노산 서열을 갖는, 기능성 유도체,
- [0066] 및/또는
- [0067] (a) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드; 또는
- [0068] (b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 60% 동일한 아미노산 서열을 갖는, 기능성 유도체.
- [0069] 적합하게는, SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체는 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 70%, 예를 들어, 적어도 80%, 특히 적어도 90%, 예를 들어, 적어도 95% 또는 적어도 98% 동일한 아미노산 서열을 갖는다.
- [0070] 적합하게는, SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체는 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 70%, 예를 들어, 적어도 80%, 특히 적어도 90%, 예를 들어, 적어도 95% 또는 적어도 98% 동일한 아미노산 서열을 갖는다.
- [0071] 특히, 본 발명에 따른 폴리펩티드, 재조합 아데노바이러스, 조성물 또는 세포는 하기를 추가로 포함한다:
- [0072] (a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드;
- [0073] 및/또는
- [0074] (b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드.
- [0075] ChAd157 섬유 또는 이의 기능성 유도체를 인코딩하는 폴리뉴클레오티드를 포함하는 단리된 폴리뉴클레오티드, 벡터, 재조합 아데노바이러스, 조성물 또는 세포
- [0076] 적합하게는, 본 발명의 단리된 폴리뉴클레오티드, 벡터, 재조합 아데노바이러스, 조성물 또는 세포는 SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드를 인코딩하는 폴리뉴클레오티드를 포함한다. 적합하게는, 폴리뉴클레오티드는 SEQ ID NO: 2에 따른 서열을 갖는다.
- [0077] 본 발명의 단리된 폴리뉴클레오티드, 벡터, 재조합 아데노바이러스, 조성물 또는 세포가 SEQ ID NO: 1에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체를 인코딩하는 폴리뉴클레오티드를 포함하는 경우, 상기 기능성 유도체는 이의 전장에 걸쳐 SEQ ID NO: 1의 아미노산 서열과 적어도 99.8%가 동일한 아미노산 서열을 갖고, 적합하게는 상기 폴리뉴클레오티드는 하나의 코돈이 첨가, 결실 또는 변경된 SEQ ID NO: 2에 따른 서열을 가져 상이한 아미노산을 인코딩한다.
- [0078] 적합하게는, 본 발명의 폴리뉴클레오티드, 벡터, 재조합 아데노바이러스, 조성물 또는 세포는 하기를 인코딩하는 폴리뉴클레오티드를 추가로 포함한다:
- [0079] (a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드; 또는
- [0080] (b) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 60% 동일한 아미노산 서열을 갖는, 기능성 유도체,

- [0081] 및/또는
- [0082] (a) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드; 또는
- [0083] (b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체로서, 상기 기능성 유도체가 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 60% 동일한 아미노산 서열을 갖는, 기능성 유도체.
- [0084] 적합하게는, SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체는 이의 전장에 걸쳐 SEQ ID NO: 3의 아미노산 서열과 적어도 70%, 예를 들어, 적어도 80%, 특히 적어도 90%, 예를 들어, 적어도 95% 또는 적어도 98% 동일한 아미노산 서열을 갖는다.
- [0085] 적합하게는, SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드의 기능성 유도체는 이의 전장에 걸쳐 SEQ ID NO: 5의 아미노산 서열과 적어도 70%, 예를 들어, 적어도 80%, 특히 적어도 90%, 예를 들어, 적어도 95% 또는 적어도 98% 동일한 아미노산 서열을 갖는다.
- [0086] 특히, 본 발명의 폴리뉴클레오티드, 벡터, 재조합 아데노바이러스, 조성물 또는 세포는 하기를 인코딩하는 폴리뉴클레오티드를 추가로 포함한다:
- [0087] (a) SEQ ID NO: 3에 따른 아미노산 서열을 갖는 폴리펩티드;
- [0088] 및/또는
- [0089] (b) SEQ ID NO: 5에 따른 아미노산 서열을 갖는 폴리펩티드.
- [0090] 본 발명의 폴리뉴클레오티드, 벡터, 재조합 아데노바이러스, 조성물 또는 세포는 하기를 추가로 포함할 수 있다:
- [0091] (a) SEQ ID NO: 4에 따른 폴리뉴클레오티드;
- [0092] 및/또는
- [0093] (a) SEQ ID NO: 6에 따른 폴리뉴클레오티드.
- [0094] **ChAd157 백본**
- [0095] 본 발명은 야생형 비변형된 ChAd157 및 ChAd157의 변형된 백본 작제물을 포함하는 침팬지 아데노바이러스 ChAd157의 단리된 폴리뉴클레오티드 서열을 제공한다. 이들 변형된 백본 작제물은 본원에 예시된 것들, 예를 들어, pChAd157 Δ E1 TetO hCMV RpsL-Kana#1551(SEQ ID NO: 15) 및 ChAd157 Δ E1E4_Ad5E4orf6/TetO hCMV RpsL-Kana#1594(SEQ ID NO: 22)를 포함한다. ChAd157 백본은, 예를 들어, 트랜스진의 전달을 위한 재조합 복제-적격 또는 복제-부적격 아데노바이러스의 작제에서 사용될 수 있다.

[0096] pChAd157 ΔE1/TetO hCMV GAG(SEQ ID NO: 17) 서열의 주석은 하기에 제공된다.

```

Annotations ChAd157DE1_TetOhCMV_GAG
IX      3187..3651
IVa2    Complement (3710..5045,5325..5337)
Pol      Complement(4816..8397, 13762..13770)
VA RNAI  10230..10391
pTP      Complement(8196..10199,13762..13770)
48K      10652..11914
pIIIa    11938..13714
III      13807..15588
pVII     15603..16199
V        16275..17390
pX       17415..17660
pVI      17750..18508
Hexon    18623..21499
Protease 21529..22158
DBP      Complement(22274..23926)
92K      23976..26447
22K      26164..26739
33K      Join(26164..26473,26679..27061)
E2e promoter Complement(27027..27274)
pVIII    27136..27819
E3 12K    27820..28137
E3 CR1-alpha0 28635..28835
E3 gp18K  28838..29329
E3A 11K   30776..31072
E3 RID alpha 31084..31356
E3 RID beta 31359..31757
E3 15K    31750..32136
U exon    Complement(32167..32331)
fibre     32342..33973
E4 ORF6/7 Complement(34181..34456,35168..35341)
E4 ORF6    Complement(34457..35341)
E4 ORF4    Complement(35241..35606)
E4 ORF3    Complement(35622..35969)
E4 ORF2    Complement(35966..36358)
E4 ORF1    Complement(36411..36797)

```

[0097]

[0098] 일 구체예에서, SEQ ID NO: 15, 22의 서열의 단편 및 이들의 상보성 가닥, cDNA 및 이에 상보적인 RNA가 제공된다. 적합하게는, 단편은 적어도 15개의 뉴클레오타이드 길이, 더욱 적합하게는 30개의 뉴클레오타이드 길이, 더욱 적합하게는 60개의 뉴클레오타이드 길이, 더욱 적합하게는 120개의 뉴클레오타이드 길이, 더욱 적합하게는 240개, 더욱 적합하게는 480개의 뉴클레오타이드 길이이며, 기능성 단편, 즉, 생물학적 관심 단편을 포함한다. 예를 들어, 기능성 단편은 요망되는 아데노바이러스 생성물을 발현시킬 수 있거나, 재조합 바이러스 벡터의 생성에 유용할 수 있다. 상기 단편은 상기 나열된 유전자 서열을 포함한다. 특정 구체예에서, SEQ ID NO: 15, 22의 단리된 서열 및 이들의 상보성 가닥, cDNA 및 이에 상보적인 RNA가 제공된다.

[0099] 본원에 기재된 아데노바이러스 핵산 및 이들의 상기 언급된 단편에 의해 인코딩되는 ChAd157 아데노바이러스의 유전자 생성물, 예를 들어, 단백질, 효소 및 이들의 단편이 제공된다. 상기 단백질은 상기 확인된 오픈 리딩 프레임에 의해 인코딩되는 단백질 및 서열 목록에 제공된 폴리뉴클레오타이드에 의해 인코딩되는 단백질을 포함한다.

[0100] **추가 ChAd157 폴리뉴클레오타이드 및 폴리펩티드**

[0101] 일부 구체예에서, 본 발명의 폴리뉴클레오타이드는 본 발명의 섬유 폴리펩티드; 핵산 폴리펩티드 및 섬유 폴리펩티드; 펜톤 폴리펩티드 및 섬유 폴리펩티드; 또는 핵산 폴리펩티드, 펜톤 폴리펩티드 및 섬유 폴리펩티드를 인코딩하는 폴리뉴클레오타이드를 포함하고; 추가적인 아데노바이러스 폴리뉴클레오타이드, 적합하게는 ChAd157 폴리뉴클레오타이드를 추가로 포함할 수 있다. 따라서, 적합하게는, 본 발명에 따른 폴리뉴클레오타이드는 하기 중 하나 이상을 포함한다:

[0102] (a) 아데노바이러스 5'-역 말단 반복(ITR);

[0103] (b) 아데노바이러스 E1A 영역, 또는 E1A_280R 및 E1A_243R 영역으로부터 선택되는 이의 단편;

[0104] (c) 아데노바이러스 E1B 또는 IX 영역, 또는 E1B_19K, E1B_55K 및 IX 영역으로 구성된 군으로부터 선택되는 이의 단편;

[0105] (d) 아데노바이러스 E2B 영역; 또는 E2B_pTP, E2B_중합효소 및 E2B_IVa2 영역으로 구성된 군으로부터 선택되는 이의 단편;

- [0106] (e) 아데노바이러스 L1 영역, 또는 이의 단편으로서, 상기 단편이 L1_13.6K, L1_52K 및 L1_pIIIa 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L1 영역, 또는 이의 단편;
- [0107] (f) 아데노바이러스 L2 영역 또는 본 발명의 펜톤 단백질을 인코딩하는 폴리뉴클레오티드를 포함하는 L2 영역, 또는 이들의 단편으로서, 상기 단편이 L2_펜톤 단백질, L2_pVII 단백질, L2_V 단백질 및 L2_pX 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L2 영역 또는 본 발명의 펜톤 단백질을 인코딩하는 폴리뉴클레오티드를 포함하는 L2 영역, 또는 이들의 단편;
- [0108] (g) 아데노바이러스 L3 영역 또는 본 발명의 핵소 단백질을 인코딩하는 폴리뉴클레오티드를 포함하는 L3 영역, 또는 이들의 단편으로서, 상기 단편이 L3_pVI 단백질, L3_핵소 단백질 및 L3_프로테아제 단백질로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L3 영역 또는 본 발명의 핵소 단백질을 인코딩하는 폴리뉴클레오티드를 포함하는 L3 영역, 또는 이들의 단편;
- [0109] (h) 아데노바이러스 E2A 영역;
- [0110] (i) 아데노바이러스 L4 영역, 또는 이의 단편으로서, 상기 단편이 L4_100k 단백질, L4_33K 단백질, L4_22K 단백질 및 단백질 L4_VIII로 구성된 군으로부터 선택되는 아데노바이러스 단백질을 인코딩하는, 아데노바이러스 L4 영역, 또는 이의 단편;
- [0111] (j) E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, 및 E3 ORF9로 구성된 군으로부터 선택되는 아데노바이러스 E3 영역 또는 이의 단편;
- [0112] (k) 아데노바이러스 L5 영역 또는 본 발명의 L5_섬유 섬유 폴리펩티드를 인코딩하는 폴리뉴클레오티드를 포함하는 L5 영역;
- [0113] (l) 아데노바이러스(예를 들어, Ad5) E4 영역, 또는 E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 및 E4 ORF1으로 구성된 군으로부터 선택되는 이의 단편; 특히 상기 E4 영역의 ORF6;
- [0114] (m) 아데노바이러스 3'-ITR; 및/또는
- [0115] (n) 아데노바이러스 VAI 또는 VAII RNA 영역, 바람직하게는 ChAd157이 아닌 아데노바이러스, 더욱 바람직하게는 Ad5로부터의 아데노바이러스 VAI 또는 VAII RNA 영역.
- [0116] **정의**
- [0117] 적합하게는, 본 발명의 폴리뉴클레오티드 또는 폴리펩티드는 단리된다. "단리된" 폴리뉴클레오티드는 이의 본래의 환경으로부터 분리된 것이다. 예를 들어, 자연-발생 폴리뉴클레오티드는 자연계에서 공존하는 물질 중 일부 또는 전부로부터 분리되는 경우에 단리된다. 폴리뉴클레오티드는, 예를 들어, 이의 자연 환경의 일부가 아닌 벡터 내로 클로닝되거나, cDNA 내에 포함되는 경우에 단리된 것으로 간주된다.
- [0118] 적합하게는, 본 발명의 폴리뉴클레오티드는 재조합 폴리뉴클레오티드이다. 재조합은 폴리뉴클레오티드가 자연에서 발견되는 폴리뉴클레오티드와 다른 폴리뉴클레오티드를 발생시키는 클로닝, 제한 또는 라이게이션 단계, 또는 다른 절차 중 적어도 하나의 생성물인 것을 의미한다. 재조합 아데노바이러스는 재조합 폴리뉴클레오티드를 포함하는 아데노바이러스이다. 재조합 벡터는 재조합 폴리뉴클레오티드를 포함하는 벡터이다. '재조합 바이러스'는 본래의 재조합 바이러스의 프로제니(progeny)를 포함한다. '재조합 벡터'는 본래의 재조합 벡터의 복제물을 포함한다. '재조합 폴리뉴클레오티드'는 본래의 재조합 폴리뉴클레오티드의 복제물을 포함한다.
- [0119] 적합하게는, 본 발명의 폴리펩티드 서열은 천연 서열과 관련하여 적어도 하나의 변형을 함유한다. 적합하게는, 본 발명의 폴리뉴클레오티드 서열은 천연 서열과 관련하여 적어도 하나의 변형을 함유한다. 예를 들어, 유전 공학 기술에 의해 상이한 종(종종, 상이한 속, 아과 또는 과)으로부터 유래된 플라스미드 또는 벡터 내로 도입된 폴리뉴클레오티드는 이종성 폴리뉴클레오티드이다. 이의 천연 코딩 서열로부터 분리되고, 자연적으로 연결된 채로 발견되지 않는 코딩 서열에 작동 가능하게 연결된 프로모터는 이종성 프로모터이다. 바이러스 또는 바이러스 벡터의 유전체 내로 클로닝된 특정 재조합 부위(바이러스의 유전체는 이를 자연적으로 함유하지 않음)는 이종성 재조합 부위이다. 이종성 핵산 서열은 또한 아데노바이러스 유전체에서 자연적으로 발견되나, 아데노바이러스 벡터 내의 비-천연 위치에 위치한 서열을 포함한다.
- [0120] 통상적으로, "이종성"은 비교되는 존재물의 나머지 것과 유전자형적으로 다른 존재물로부터 유래된 것을 의미한다. 이종성 핵산 서열은 아데노바이러스 벡터의 자연 발생 핵산 서열로부터 단리되거나, 이로부터 유래되거나, 이를 기초로 하지 않는 임의의 핵산 서열을 나타낸다. 이종성 단백질 서열은 아데노바이러스 벡터의 자연 발생

단백질 서열로부터 단리되거나, 이로부터 유래되거나, 이에 기반하지 않은 임의의 단백질 서열을 나타낸다. "자연 발생"은 자연에서 발견되고, 합성적으로 제조되거나 변형되지 않은 서열을 의미한다. 서열은 공급원으로부터 분리되거나 변형(예를 들어, 결실, 치환(돌연변이), 삽입, 또는 다른 변형에 의해)되어, 적합하게는 공급원 유전자의 정상 기능을 붕괴시키지 않는 경우에 공급원으로부터 "유래"된 것이다.

- [0121] 폴리펩티드의 "기능성 유도체"는 적합하게는 폴리펩티드의 변형된 형태, 예를 들어, 폴리펩티드의 하나 이상의 아미노산이 결실되고/되거나, 삽입되고/되거나, 변형되고/되거나, 치환될 수 있는 폴리펩티드의 변형된 형태를 나타낸다. 변형되지 않은 아데노바이러스 캡시드 단백질의 유도체는, 예를 들어, 하기와 같은 경우 기능성인 것으로 간주된다:
- [0122] (a) 캡시드 내에 유도체 캡시드 단백질을 포함하는 아데노바이러스가 변형되지 않은 캡시드 단백질을 포함하는 아데노바이러스에 비해 실질적으로 동일하거나 더 낮은 혈청유병률을 보유하고/하거나
- [0123] (b) 캡시드 내에 유도체 캡시드 단백질을 포함하는 아데노바이러스가 변형되지 않은 캡시드 단백질을 포함하는 아데노바이러스에 비해 실질적으로 동일하거나 더 높은 숙주 세포 감염성을 보유하고/하거나
- [0124] (c) 캡시드 내에 유도체 캡시드 단백질을 포함하는 아데노바이러스가 변형되지 않은 캡시드 단백질을 포함하는 아데노바이러스에 비해 실질적으로 동일하거나 더 높은 면역원성을 보유하고/하거나
- [0125] (d) 캡시드 내에 유도체 캡시드 단백질을 포함하는 아데노바이러스가 변형되지 않은 캡시드 단백질을 포함하는 아데노바이러스에 비해 실질적으로 동일하거나 더 높은 수준의 트랜스진 생산성을 보유하는 경우.
- [0126] 상기 특성 (a)-(d)는 적합하게는 하기 실시예 섹션에 기재된 방법을 이용하여 측정될 수 있다.
- [0127] 적합하게는, 폴리펩티드, 벡터 또는 재조합 아데노바이러스는 인간 집단에서 낮은 혈청유병률을 갖는다. "낮은 혈청유병률"은 인간 아데노바이러스 5(Ad5)에 비해 감소된 기존 중화 항체 수준을 갖는 것을 의미할 수 있다. 유사하거나 대안적으로, "낮은 혈청유병률"은 약 20% 미만의 혈청유병률, 약 15% 미만의 혈청유병률, 약 10% 미만의 혈청유병률, 약 5% 미만의 혈청유병률, 약 4% 미만의 혈청유병률, 약 3% 미만의 혈청유병률, 약 2% 미만의 혈청유병률, 약 1% 미만의 혈청유병률 또는 검출 가능하지 않은 혈청유병률을 의미할 수 있다. 혈청유병률은 문헌[Aste-Amezaga et al., *Hum. Gene Ther.* (2004) 15(3):293-304]에 기재된 방법을 이용하여 임상적으로 관련된 중화 역가(200 초과와 50% 중화 역가로 정의됨)를 갖는 개체의 백분율로 측정될 수 있다.
- [0128] 용어 폴리펩티드, 펩티드 및 단백질은 본원에서 상호교환적으로 사용된다.
- [0129] 용어 "유인원"은 통상적으로 비인간 영장류, 예를 들어, 구세계 원숭이, 신세계 원숭이, 원숭이 및 긴팔원숭이를 포함하는 것을 의미한다. 특히, 유인원은 비인간 원숭이, 예를 들어, 침팬지(판 트로글로디테(*Pan troglodyte*)), 보노보(판 파니스쿠스(*Pan paniscus*)) 및 고릴라(고릴라 속)를 나타낼 수 있다. 비-원숭이 유인원은 히말라야 원숭이(마카카 몰라타(*Macaca mulatta*))를 포함할 수 있다.
- [0130] **서열 비교**
- [0131] 2개의 밀접하게 관련된 폴리뉴클레오티드 또는 폴리펩티드 서열을 비교하기 위해, 제1 서열과 제2 서열 사이의 "동일성 %"는 표준 설정을 이용한 정렬 프로그램, 예를 들어, BLAST®(blast.ncbi.nlm.nih.gov에서 이용 가능함, 2015년 3월 9일에 최종 액세스)을 이용하여 계산될 수 있다. 동일성 %는 동일한 잔기의 수를 참조 서열 내의 잔기의 수로 나누고, 100을 곱한 것이다. 상기 및 청구항에서 언급된 동일성 % 숫자는 이러한 방법에 의해 계산된 백분율이다. 동일성 %의 대안적 정의는 동일한 잔기의 수를 정렬된 잔기의 수로 나누고, 100을 곱한 것이다. 대안적 방법은 정렬 내의 갭, 예를 들어, 다른 서열에 비한 한 서열에서의 결실이 스코어링 파라미터에서 갭 스코어 또는 갭 비용으로 설명되는 갭 방법(gapped method)을 이용하는 것을 포함한다. 더 많은 정보를 위해, 2015년 3월 9일에 마지막으로 액세스된 ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_BLASTGuide.pdf에서 이용 가능한 BLAST® 팩트 시트(fact sheet)를 참조한다.
- [0132] 폴리뉴클레오티드 또는 이에 의해 인코딩된 폴리펩티드의 기능성을 보존하는 서열은 더욱 밀접하게 동일할 수 있다. 폴리펩티드 또는 폴리뉴클레오티드 서열은 이들이 이들의 전장에 걸쳐 100% 서열 동일성을 공유하는 경우에 다른 폴리펩티드 또는 폴리뉴클레오티드 서열과 동일하거나 일치하는 것으로 언급된다.
- [0133] 서열 사이의 "차이"는 제1 서열에 비해 제2 서열의 한 위치에서의 단일 아미노산 잔기의 삽입, 결실 또는 치환을 나타낸다. 2개의 폴리펩티드 서열은 1개, 2개 또는 그 초과와 상기 아미노산 차이를 함유할 수 있다. 제1 서열과 달리 동일한(100% 서열 동일성) 제2 서열에서의 삽입, 결실 또는 치환은 서열 동일성 퍼센트를

감소시킨다. 예를 들어, 동일한 서열이 9개의 아미노산 잔기 길이인 경우, 제2 서열에서의 하나의 치환은 88.9%의 서열 동일성을 발생시킨다. 동일한 서열이 17개의 아미노산 잔기 길이인 경우, 제2 서열에서의 2개의 치환은 88.2%의 서열 동일성을 발생시킨다. 동일한 서열이 7개의 아미노산 잔기 길이인 경우, 제2 서열에서의 3개의 치환은 57.1%의 서열 동일성을 발생시킨다. 제1 및 제2 폴리펩티드 서열이 9개의 아미노산 잔기 길이이고, 6개의 동일한 잔기를 공유하는 경우, 제1 및 제2 폴리펩티드 서열은 66% 초과와 동일성을 공유한다(제1 및 제2 폴리펩티드 서열은 66.7%의 동일성을 공유한다). 제1 및 제2 폴리펩티드 서열이 17개의 아미노산 잔기 길이이고, 16개의 동일한 잔기를 공유하는 경우, 제1 및 제2 폴리펩티드 서열은 94% 초과와 동일성을 공유한다(제1 및 제2 폴리펩티드 서열은 94.1%의 동일성을 공유한다). 제1 및 제2 폴리펩티드 서열이 7개의 아미노산 잔기 길이이고, 3개의 동일한 잔기를 공유하는 경우, 제1 및 제2 폴리펩티드 서열은 42% 초과와 동일성을 공유한다(제1 및 제2 폴리펩티드 서열은 42.9%의 동일성을 공유한다).

[0134] 대안적으로, 제1의 참조 폴리펩티드 서열을 제2의 비교 폴리펩티드 서열과 비교하기 위해, 제2 서열을 생성시키기 위해 제1 서열에서 이루어진 첨가, 치환 및/또는 결실의 수가 확인될 수 있다. 첨가는 제1 폴리펩티드의 서열로의 하나의 아미노산 잔기의 첨가(제1 폴리펩티드의 어느 한 말단에서의 첨가를 포함)이다. 치환은 제1 폴리펩티드의 서열 내의 하나의 아미노산 잔기의 하나의 상이한 아미노산 잔기에 의한 치환이다. 결실은 제1 폴리펩티드의 서열로부터의 하나의 아미노산 잔기의 결실(제1 폴리펩티드의 어느 한 말단에서의 결실을 포함)이다.

[0135] 제1의 참조 폴리뉴클레오티드 서열을 제2의 비교 폴리뉴클레오티드 서열과 비교하기 위해, 제2 서열을 생성시키기 위해 제1 서열에서 이루어진 첨가, 치환 및/또는 결실의 수가 확인될 수 있다. 첨가는 제1 폴리뉴클레오티드의 서열로의 하나의 뉴클레오티드 잔기의 첨가(제1 폴리뉴클레오티드의 어느 한 말단에서의 첨가를 포함)이다. 치환은 제1 폴리뉴클레오티드의 서열 내의 하나의 뉴클레오티드 잔기의 하나의 상이한 뉴클레오티드 잔기에 의한 치환이다. 결실은 제1 폴리뉴클레오티드의 서열로부터의 하나의 뉴클레오티드 잔기의 결실(제1 폴리뉴클레오티드의 어느 한 말단에서의 결실을 포함)이다.

[0136] 적합하게는, 본 발명의 서열 내에서의 치환은 보존성 치환일 수 있다. 보존성 치환은 한 아미노산의 치환되는 아미노산과 유사한 화학적 특성을 갖는 또 다른 아미노산에 의한 치환을 포함한다(예를 들어, 문헌[Stryer et al, *Biochemistry*, 5th Edition 2002, pages 44-49] 참조). 바람직하게는, 보존성 치환은 (i) 염기성 아미노산의 또 다른 상이한 염기성 아미노산에 의한 치환; (ii) 산성 아미노산의 또 다른 상이한 산성 아미노산에 의한 치환; (iii) 방향족 아미노산의 또 다른 상이한 방향족 아미노산에 의한 치환; (iv) 비-극성 지방족 아미노산의 또 다른 상이한 비-극성 지방족 아미노산에 의한 치환; 및 (v) 극성의 하전되지 않은 아미노산의 또 다른 상이한 극성의 하전되지 않은 아미노산에 의한 치환으로 구성된 군으로부터 선택된 치환이다. 염기성 아미노산은 바람직하게는 아르기닌, 히스티딘, 및 리신으로 구성된 군으로부터 선택된다. 산성 아미노산은 바람직하게는 아스파테이트 또는 글루타메이트이다. 방향족 아미노산은 바람직하게는 페닐알라닌, 티로신 및 트립토판으로 구성된 군으로부터 선택된다. 비-극성의 지방족 아미노산은 바람직하게는 글리신, 알라닌, 발린, 류신, 메티오닌 및 이소류신으로 구성된 군으로부터 선택된다. 극성의 하전되지 않은 아미노산은 바람직하게는 세린, 트레오닌, 시스테인, 프롤린, 아스파라긴 및 글루타민으로 구성된 군으로부터 선택된다. 보존성 아미노산 치환과 대조적으로, 비-보존성 아미노산 치환은 한 아미노산의 상이 개설했던 보존성 치환 (i) 내지 (v)에 속하지 않는 임의의 아미노산에 의한 교환이다.

[0137] **백터 및 재조합 아데노바이러스**

[0138] 본 발명의 ChAd157 서열은 치료제로서, 그리고 다양한 백터 시스템, 재조합 아데노바이러스 및 숙주 세포의 작제에서 유용하다. 적합하게는, 용어 "백터"는 야생형 서열에 비해 실질적으로 변경(예를 들어, 결실되고/되거나 비활성화된 유전자 또는 기능성 영역)되고/되었거나 이중성 서열, 즉, 상이한 공급원으로부터 획득된 핵산("삽입물"로도 언급됨)이 혼입되고, 세포(예를 들어, 숙주 세포)로 도입되는 경우 삽입된 폴리뉴클레오티드 서열을 복제하고/하거나 발현시킬 수 있는 핵산을 나타낸다. 예를 들어, 삽입물은 본원에 기재된 ChAd157 서열의 전부 또는 일부일 수 있다. 추가로 또는 대안적으로, ChAd157 백터는 바이러스 유전자, 예를 들어, E1 또는 본원에 기재된 다른 바이러스 유전자 또는 기능성 영역의 하나 이상의 결실 또는 비활성화를 포함하는 ChAd157 아데노바이러스일 수 있다. 이중성 서열을 포함하거나 포함하지 않을 수 있는 상기 ChAd157는 종종 "백본"으로 언급되며, 그대로 사용되거나, 백터에 대한 추가 변형을 위한 시작점으로 사용될 수 있다.

[0139] 백터는 네이키드 DNA, 플라스미드, 바이러스, 코스미드, 파지 백터, 예를 들어, 람다 백터, 인공 염색체, 예를 들어, BAC(박테리아 인공 염색체), 또는 에피솜을 포함하는 임의의 적합한 핵산 분자일 수 있다. 대안적으로, 백터는 세포 비합류 시험관 내 전사 또는 발현을 위한 전사 및/또는 발현 유닛, 예를 들어, T7-상용성 시스템일

수 있다. 벡터는 단독으로 또는 다른 아데노바이러스 서열 또는 단편과 함께, 또는 비-아데노바이러스 서열로부터의 요소와 함께 사용될 수 있다. ChAd157 서열은 또한 안티센스 전달 벡터, 유전자 요법 벡터, 또는 백신 벡터에서 유용할 수 있다. 따라서, 유전자 전달 벡터, 및 ChAd157 서열을 함유하는 숙주 세포가 추가로 제공된다.

[0140] 용어 "복제-적격" 아데노바이러스는 세포 내에 포함된 임의의 재조합 헬퍼 단백질의 부재 하에 숙주 세포에서 복제할 수 있는 아데노바이러스를 나타낸다. 적합하게는, "복제-적격" 아데노바이러스는 다음과 같은 무손상 또는 기능성 필수 초기 유전자를 포함한다: E1A, E1B, E2A, E2B, E3 및 E4. 특정 동물로부터 단리된 야생형 아데노바이러스는 그 동물에서 복제 적격일 것이다.

[0141] 용어 "복제-부적격" 또는 "복제-결함" 아데노바이러스는 적어도 기능적 결실(또는 "기능 손실" 돌연변이), 즉, 완전히 제거되지 않고 유전자의 기능을 손상시키는 결실 또는 돌연변이, 예를 들어, 인공 정지 코돈의 도입, 활성 부위 또는 상호작용 도메인의 결실 또는 돌연변이, 유전자의 조절 서열의 돌연변이 또는 결실 등, 또는 바이러스 복제에 필수적인 유전자 생성물을 인코딩하는 유전자, 예를 들어, E1A, E1B, E2A, E2B, E3 및 E4(예를 들어, E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, E3 ORF9, E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 및/또는 E4 ORF1)로부터 선택되는 아데노바이러스 유전자 중 하나 이상의 완전한 제거를 포함하도록 공학 처리됨으로 인해 복제 불가능한 아데노바이러스를 나타낸다. 특히, 적합하게는 E1 및 선택적으로 E3 및/또는 E4가 결실된다. 결실되는 경우, 상기 언급된 결실된 유전자 영역은 적합하게는 또 다른 서열과 관련하여 동일성 %를 결정하는 경우에 정렬에서 고려되지 않을 것이다.

[0142] 본 발명은 치료 또는 백신 목적으로 단백질, 적합하게는 이종성 단백질을 세포로 전달하는 벡터, 예를 들어, 재조합 아데노바이러스를 제공한다. 벡터는 네이키드 DNA, 파지, 트랜스포존(transposon), 코스미드, 에피솜, 플라스미드, 또는 바이러스를 포함하는 임의의 유전적 요소를 포함할 수 있다. 상기 벡터는 본원에 기재된 ChAd157의 DNA 및 미니유전자를 함유한다. "미니유전자"(또는 "발현 카세트")는 선택된 이종성 유전자(트랜스진) 및 숙주 세포에서 유전자 생성물의 번역, 전사 및/또는 발현을 유도하는데 필요한 다른 조절 요소의 조합물을 의미한다.

[0143] 통상적으로, ChAd157-유래된 아데노바이러스 벡터는 미니유전자가 선택된 아데노바이러스 유전자에 고유한 영역에서 다른 아데노바이러스 서열을 함유하는 핵산 분자 내에 위치하도록 설계된다. 미니유전자는 요망되는 경우 기존 유전자 영역에 삽입되어 그 유전자의 기능을 붕괴시킬 수 있다. 대안적으로, 미니유전자는 부분적 또는 완전히 결실된 아데노바이러스 유전자의 부위에 삽입될 수 있다. 예를 들어, 미니유전자는 E1A, E1B, E2A, E2B, E3 및 E4로 구성된 군으로부터 선택되는 유전체 영역의 적어도 하나의 유전자가 비-기능성이 되게 하는 돌연변이, 삽입 또는 결실의 부위에 위치할 수 있다. 용어 "비-기능성이 되게 한다"는 충분한 양의 유전자 영역이 제거되거나 달리 붕괴되어, 유전자 영역이 더 이상 유전자 발현의 기능성 생성물을 생산할 수 없는 것을 의미한다. 요망되는 경우, 전체 유전자 영역이 제거될 수 있다(그리고 적합하게는 미니유전자로 대체될 수 있다).

[0144] 예를 들어, 재조합 바이러스의 생성에 유용한 생산 벡터의 경우, 벡터는 미니유전자 및 아데노바이러스 유전체의 5' 말단 또는 아데노바이러스 유전체의 3' 말단 중 어느 하나, 또는 아데노바이러스 유전체의 5' 및 3' 말단 둘 모두를 함유할 수 있다. 아데노바이러스 유전체의 5' 말단은 패키징 및 복제에 필요한 5' 시스-요소, 즉, 5' ITR 서열(복제 기점으로 기능함) 및 천연 5' 패키징 인핸서 도메인(전형 Ad 유전체 및 E1 프로모터에 대한 인핸서 요소를 패키징하는데 필요한 서열을 함유함)을 함유한다. 아데노바이러스 유전체의 3' 말단은 패키징 및 캡시드화에 필요한 3' 시스-요소(ITR을 포함함)를 포함한다. 적합하게는, 재조합 아데노바이러스는 5' 및 3' 아데노바이러스 시스-요소 둘 모두를 함유하며, 미니유전자(적합하게는 트랜스진을 함유함)는 5' 및 3' 아데노바이러스 서열 사이에 위치한다. ChAd157-기반 아데노바이러스 벡터는 또한 추가적인 아데노바이러스 서열을 함유할 수 있다.

[0145] 적합하게는, ChAd157-기반 벡터는 본 발명의 아데노바이러스 ChAd157 유전체로부터 유래된 하나 이상의 아데노바이러스 요소를 함유한다. 일 구체예에서, 벡터는 ChAd157로부터의 아데노바이러스 ITR 및 동일 아데노바이러스 혈청형으로부터의 추가 아데노바이러스 서열을 함유한다. 또 다른 구체예에서, 벡터는 ITR을 제공한 것과 상이한 아데노바이러스 혈청형으로부터 유래되는 아데노바이러스 서열을 함유한다.

[0146] 본원에 정의된 바와 같이, 슈도타이핑된 아데노바이러스는 아데노바이러스의 캡시드 단백질이 ITR을 제공한 아데노바이러스와 상이한 아데노바이러스로부터 유래되는 아데노바이러스를 나타낸다.

[0147] 추가로, 키메라 또는 하이브리드 아데노바이러스는 당업자에게 공지된 기술(예를 들어, US 7,291,498호)을 이용

하여 본원에 기재된 아데노바이러스를 이용하여 작제될 수 있다.

[0148] 본 발명의 벡터에 존재하는 ITR 및 임의의 다른 아데노바이러스 서열은 많은 공급원으로부터 획득될 수 있다. 다양한 아데노바이러스 균주가 미국 미생물보존센터(American Type Culture Collection, Manassas, Virginia)로부터 이용 가능하거나, 다양한 상업 및 기관 공급처로부터 요청에 의해 이용 가능하다. 추가로, 많은 상기 균주의 서열은, 예를 들어, PubMed 및 GenBank를 포함하는 다양한 데이터베이스로부터 이용 가능하다. 다른 침팬지 또는 인간 아데노바이러스로부터 제조된 동종 아데노바이러스 벡터는 공개된 문헌(예를 들어, US 5,240,846호)에 기재되어 있다. 유형 Ad5(GenBank 등록 번호 M73370)를 포함하는 다수의 아데노바이러스 유형의 DNA 서열이 GenBank로부터 이용 가능하다. 아데노바이러스 서열은 현재 확인된 인간 유형 중 임의의 유형을 추가로 포함하여 임의의 공지된 아데노바이러스 혈청형, 예를 들어, 혈청형 2, 3, 4, 7, 12 및 40으로부터 획득될 수 있다. 유사하게, 비인간 동물(예를 들어, 유인원)을 감염시키는 것으로 공지된 아데노바이러스가 또한 본 발명의 벡터 작제물에서 사용될 수 있다(예를 들어, US 6,083,716호). 본원에 기재된 벡터의 작제에서 사용되는 바이러스 서열, 헬퍼 바이러스(필요한 경우), 및 재조합 바이러스 입자, 및 다른 벡터 성분 및 서열은 하기 기재되는 바와 같이 획득될 수 있다.

[0149] **서열, 벡터 및 아데노바이러스 생성**

[0150] 본 발명의 서열은 재조합 생성, 화학 합성, 또는 다른 합성 수단을 포함하는 임의의 적합한 수단에 의해 생성될 수 있다. 적합한 생성 기술은 당업자에게 널리 공지되어 있다. 대안적으로, 펩티드는 또한 널리 공지된 고체상 펩티드 합성 방법에 의해 합성될 수 있다.

[0151] 아데노바이러스 플라스미드(또는 다른 벡터)가 아데노바이러스 벡터를 생성시키기 위해 사용될 수 있다. 일 구체예에서, 아데노바이러스 벡터는 복제-부적격한 아데노바이러스 입자이다. 일 구체예에서, 아데노바이러스 입자는 E1A 및/또는 E1B 유전자, 특히 E1A 및 E1B의 결실에 의해 복제-부적격이 된다. 대안적으로, 아데노바이러스는 선택적으로 E1A 및/또는 E1B 유전자를 보존하면서 또 다른 수단에 의해 복제-부적격이 된다. 유사하게는, 일부 구체예에서, 벡터에 대한 면역 반응의 감소는 E2B 및/또는 DNA 중합효소 유전자에서의 결실에 의해 달성될 수 있다. 아데노바이러스 벡터는 또한 아데노바이러스 유전체에 대한 다른 돌연변이, 예를 들어, 온도-민감성 돌연변이 또는 다른 유전자에서의 결실을 함유할 수 있다. 다른 구체예에서, 아데노바이러스 벡터에서 무손상 E1A 및/또는 E1B 영역을 보유하는 것이 바람직하다. 상기 무손상 E1 영역은 아데노바이러스 유전체 내의 이의 천연 위치에 위치될 수 있거나, 천연 아데노바이러스 유전체 내의 결실 부위(예를 들어, E3 영역 내)에 배치될 수 있다.

[0152] 포유동물(예를 들어, 인간) 세포로의 유전자의 전달을 위한 아데노바이러스 벡터의 작제에서, 광범위한 변형된 아데노바이러스 핵산 서열이 벡터에서 사용될 수 있다. 예를 들어, 아데노바이러스 지연 초기 유전자 E3의 일부 또는 전부가 재조합 바이러스의 일부를 형성하는 아데노바이러스 서열로부터 제거될 수 있다. E3의 기능은 재조합 바이러스 입자의 기능 및 생성과 관련이 없는 것으로 생각된다. 아데노바이러스 벡터는 또한 E4 유전자의 적어도 ORF6 영역, 및 더욱 바람직하게는 이러한 영역의 기능에서의 중복성으로 인해 전체 E4 영역의 결실을 갖도록 작제될 수 있다. 본 발명의 또 다른 벡터는 지연 초기 유전자 E2A에서 결실을 함유한다. 결실은 또한 아데노바이러스 유전체의 후기 유전자 L1 내지 L5 중 임의의 것에서 이루어질 수 있다. 유사하게는, 중간체 유전자 IX 및 IVa2에서의 결실이 일부 목적에 유용할 수 있다. 다른 결실이 다른 구조 또는 비-구조 아데노바이러스 유전자에서 이루어질 수 있다. 상기 논의된 결실은 개별적으로 이용될 수 있으며, 즉, 본원에 기재된 바와 같은 사용을 위한 아데노바이러스 서열은 단지 하나의 영역에서 결실을 함유할 수 있다. 대안적으로, 전체 유전자 또는 이들의 생물학적 활성을 파괴하기에 효과적인 상기 유전자의 일부의 결실은 임의의 조합으로 사용될 수 있다. 예를 들어, 한 예시적인 벡터에서, 아데노바이러스 서열은 E1 유전자 및 E4 유전자의 결실, 또는 E1, E2A 및 E3 유전자의 결실, 또는 E1 및 E3 유전자의 결실, 또는 E3의 결실과 함께 또는 이의 결실 없이 E1, E2A 및 E4 유전자의 결실 등을 가질 수 있다. E 유전자 중 어느 하나 이상은 적합하게는 아데노바이러스의 상이한 균주로부터 유래된 E 유전자(또는 하나 이상의 E 유전자 오픈 리딩 프레임)로 대체될 수 있다. 특히 적합하게는, ChAd157 E1 및 E3 유전자는 결실되고, ChAd157E4 유전자는 E4Ad5orf6으로 대체된다. 상기 논의된 바와 같이, 상기 결실 및/또는 치환은 요망되는 결과를 달성하기 위해 다른 돌연변이, 예를 들어, 온도-민감성 돌연변이와 함께 사용될 수 있다.

[0153] 하나 이상의 필수 아데노바이러스 서열(예를 들어, E1A, E1B, E2A, E2B, E4 ORF6, L1, L2, L3, L4 및 L5)가 결여된 아데노바이러스 벡터는 아데노바이러스 입자의 바이러스 감염성 및 증식에 필요한 손실된 아데노바이러스 유전자 생성물의 존재 하에 배양될 수 있다. 이들 헬퍼 기능은 하나 이상의 헬퍼 작제물(예를 들어, 플라스미드

또는 바이러스) 또는 패키징 숙주 세포의 존재 하에 아데노바이러스 벡터를 배양함으로써 제공될 수 있다.

[0154] 복제-부적격 벡터의 상보성

[0155] 상기 기재된 유전자 중 임의의 유전자에서 결실된 재조합 아데노바이러스를 생성시키기 위해, 바이러스의 복제 및 감염성에 필수적인 경우 결실된 유전자 영역의 기능은 헬퍼 바이러스 또는 세포주, 즉, 상보성 또는 패키징 세포주에 의해 재조합 바이러스에 제공되어야 한다.

[0156] 헬퍼 바이러스

[0157] 미니유전자를 운반하는데 사용되는 바이러스 벡터의 아데노바이러스 유전자 내용물에 따라, 미니유전자를 함유하는 감염성 재조합 바이러스 입자를 생성시키는데 필요한 충분한 아데노바이러스 유전자 서열을 제공하기 위해 헬퍼 아데노바이러스 또는 비-복제 바이러스 단편이 사용될 수 있다. 유용한 헬퍼 바이러스는 아데노바이러스 벡터 작제물에 존재하지 않고/않거나 벡터가 트랜스팩션되는 패키징 세포주에 의해 발현되지 않는 선택된 아데노바이러스 유전자 서열을 함유한다. 일 구체예에서, 헬퍼 바이러스는 복제-결합성이며, 적합하게는 본원에 기재된 하나 이상의 서열에 더하여 아데노바이러스 유전자를 함유한다. 상기 헬퍼 바이러스는 적합하게는 E1 발현 (선택적으로, 추가로 E3 발현) 세포주와 함께 사용된다.

[0158] 헬퍼 바이러스는 선택적으로 리포터 유전자를 함유할 수 있다. 다수의 상기 리포터 유전자가 당 분야에 공지되어 있을뿐만 아니라 본원에 기재되어 있다. 아데노바이러스 벡터 상의 트랜스진과 상이한 헬퍼 바이러스 상의 리포터 유전자의 존재는 아데노바이러스 벡터 및 헬퍼 바이러스 둘 모두가 독립적으로 모니터링되는 것을 가능하게 한다. 이러한 리포터는 정제시 생성된 재조합 바이러스와 헬퍼 바이러스 사이의 분리를 가능하게 하기 위해 사용된다.

[0159] 상보성 세포주

[0160] 많은 상황에서, 바이러스의 복제 및 감염성에 필수적인 하나 이상의 손실된 유전자, 예를 들어, 인간 E1을 발현하는 세포주가 침팬지 아데노바이러스 벡터를 교차보충(transcomplement)하는데 사용될 수 있다. 이는 본 발명의 침팬지 아데노바이러스 서열과 현재 이용 가능한 패키징 세포에서 발견되는 인간 아데노바이러스 서열 사이의 다양성으로 인해 현재 인간 E1-함유 세포의 사용이 복제 및 생산 과정 동안 복제-적격 아데노바이러스의 발생을 방지함에 따라 특히 유리하다.

[0161] 대안적으로, 요망된다면, 선택된 부모 세포주에서의 발현을 위한 프로모터의 전사 조절 하에 ChAd 157 또는 다른 아데노바이러스(예를 들어, hAd5 E1과 같은 인간 아데노바이러스, 또는 다른 ChAd E1)로부터 E1 유전자를 최소로 발현하는 패키징 세포 또는 세포주를 생성시키기 위해 본원에서 제공된 서열을 이용할 수 있다. 이러한 목적을 위해 유도성 또는 항시성 프로모터가 사용될 수 있다. 상기 프로모터의 예는 본 문서의 다른 곳에 상세히 기재되어 있다. 임의의 요망되는 ChAd157 유전자를 발현시키는 신규한 세포주의 생성을 위한 부모 세포가 선택된다. 비제한적으로, 상기 부모 세포주는 다른 것 중에서도 HeLa[ATCC 등록 번호 CCL 2], A549[ATCC 등록 번호 CCL 185], HEK 293, KB[CCL 17], Detroit[예를 들어, Detroit 510, CCL 72] 및 WI-38[CCL 75] 세포일 수 있다. 이들 세포주는 미국 미생물보존센터(American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209)로부터 모두 이용 가능하다.

[0162] 상기 E1-발현 세포주는 재조합 아데노바이러스 E1 결실 벡터의 생성에서 유용하다. 추가로, 또는 대안적으로, 하나 이상의 아데노바이러스 유전자 생성물, 예를 들어, E1A, E1B, E2A, E3 및/또는 E4를 발현하는 세포주가 재조합 바이러스 벡터의 생성에서 사용되는 것과 본질적으로 동일한 절차를 이용하여 작제될 수 있다. 상기 세포주는 그러한 생성물을 인코딩하는 필수 유전자가 결실된 아데노바이러스 벡터를 교차보충하거나, 헬퍼-의존성 바이러스(예를 들어, 아데노-관련 바이러스)의 패키징에 필요한 헬퍼 기능을 제공하기 위해 사용될 수 있다. 숙주 세포의 제조는 선택된 DNA 서열의 어셈블리와 같은 기술을 포함한다.

[0163] 또 다른 대안에서, 필수 아데노바이러스 유전자 생성물은 아데노바이러스 벡터 및/또는 헬퍼 바이러스에 의해 교차하여(in trans) 제공된다. 상기 예에서, 적합한 숙주 세포는 임의의 생물학적 유기체, 예를 들어, 원핵생물(예를 들어, 박테리아) 세포, 및 진핵생물 세포, 예를 들어, 곤충 세포, 효모 세포 및 포유동물 세포로부터 선택될 수 있다.

[0164] 숙주 세포는 인간, 원숭이, 마우스, 래트, 토끼, 및 햄스터를 포함하는 포유동물로부터 유래된 세포, 예를 들어, A549, WEHI, 3T3, 10'I'I/2, HEK 293 세포 또는 Per.C6(둘 모두는 기능성 아데노바이러스 E1을 발현함)[Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L 세포, HT1080, HepG2 및 일차

섬유모세포, 간세포 및 근육모세포를 포함하나 이에 제한되지는 않는 임의의 포유동물 종으로부터 선택될 수 있다.

[0165] 특히 적합한 상보성 세포주는 Procell92 세포주이다. Procell92 세포주는 인간 포스포글리세레이트 키나제-1(PGK) 프로모터의 조절 하에 Tet 억제인자로 트랜스펙션된 아데노바이러스 E1 유전자, 및 G418-내성 유전자를 발현하는 HEK 293 세포에 기반한다(Vitelli et al. *PLoS One* (2013) 8(e55435):1-9). Procell92.S는 현탁 조건에서의 성장에 적합하며, 독성 단백질을 발현하는 아데노바이러스 벡터를 생성하는데 유용하다(www.okairos.com/e/inner.php?m=00084, 2015년 4월 13일에 마지막으로 액세스됨).

[0166] 바이러스 입자의 어셈블리 및 세포주의 트랜스펙션

[0167] 일반적으로, 트랜스펙션에 의해 미니유전자를 포함하는 벡터를 전달하는 경우, 벡터는 약 1×10^4 개의 세포 내지 약 1×10^{13} 개의 세포, 및 바람직하게는 약 10^5 개의 세포에 약 5 μ g 내지 약 100 μ g의 DNA, 바람직하게는 약 10 내지 약 50 μ g의 DNA의 양으로 전달된다. 그러나, 벡터 DNA 대 숙주 세포의 상대량은 선택된 벡터, 전달 방법 및 선택된 숙주 세포와 같은 요인을 고려하여 조정될 수 있다.

[0168] 벡터의 숙주 세포로의 도입은 트랜스펙션 및 감염을 포함하는 당 분야에 공지된 임의의 수단에 의해 달성될 수 있다. 아데노바이러스 유전자 중 하나 이상은 숙주 세포의 유전체로 안정적으로 통합되거나, 에피솜으로서 안정적으로 발현되거나, 일시적으로 발현될 수 있다. 유전자 생성물은 모두 에피솜 상태로 또는 안정적으로 통합된 상태로 일시적으로 발현될 수 있거나, 유전자 생성물 중 일부는 안정적으로 발현될 수 있는 반면 다른 일부는 일시적으로 발현된다.

[0169] 숙주 세포로의 벡터의 도입은 또한 당업자에게 공지된 기술을 이용하여 달성될 수 있다. 적합하게는, 표준 트랜스펙션 기술, 예를 들어, CaPC 트랜스펙션 또는 전기천공이 사용된다.

[0170] 다양한 중간체 플라스미드로의 아데노바이러스의 선택된 DNA 서열(및 트랜스진 및 다른 벡터 요소)의 어셈블리, 및 재조합 바이러스 입자를 생성시키기 위한 플라스미드 및 벡터의 사용은 모두 통상적인 기술을 이용하여 달성된다. 상기 기술은 cDNA의 통상적인 클로닝 기술, 아데노바이러스 유전체의 중첩 올리고뉴클레오타이드 서열의 이용, 중합효소 연쇄 반응, 및 요망되는 뉴클레오타이드 서열을 제공하는 임의의 적합한 방법을 포함한다. 표준 트랜스펙션 및 공동-트랜스펙션 기술, 예를 들어, CaPC 침전 기술이 사용된다. 사용되는 다른 통상적인 방법은 바이러스 유전체의 상동 재조합, 아가 오버레이(agar overlay)에서의 바이러스 플라크 형성(plaquing), 신호 발생을 측정하는 방법 등을 포함한다.

[0171] 예를 들어, 요망되는 미니유전자-함유 바이러스 벡터의 작제 및 어셈블리 후, 벡터는 헬퍼 바이러스의 존재 하에 패키징 세포주로 시험관 내에서 트랜스펙션된다. 상동 재조합이 헬퍼 서열과 벡터 서열 사이에서 발생하며, 이는 벡터 내의 아데노바이러스-트랜스진 서열이 복제되고 비리온 캡시드로 패키징되는 것을 가능하게 하여, 재조합 바이러스 벡터 입자를 생성시킨다. 생성된 재조합 아데노바이러스는 선택된 트랜스진을 선택된 세포로 전달하는데 유용하다. 패키징 세포주에서 성장된 재조합 바이러스를 이용한 생체 내 실험에서, 본 발명의 E1-결실된 재조합 아데노바이러스 벡터는 트랜스진을 비-유인원 포유동물, 바람직하게는 인간 세포로 전달하는데 있어서의 유용성을 입증한다.

[0172] 트랜스진

[0173] 트랜스진은 그 트랜스진에 측접한 벡터 서열과 이종성인 핵산 서열이며, 관심 단백질을 인코딩한다. 핵산 코딩 서열은 숙주 세포에서 트랜스진 전사, 번역, 및/또는 발현을 가능하게 하는 방식으로 조절 성분에 작동 가능하게 연결된다.

[0174] 트랜스진 서열의 조성은 생성되는 벡터가 사용될 용도에 좌우될 것이다. 예를 들어, 트랜스진은 치료용 트랜스진 또는 면역원성 트랜스진일 수 있다. 대안적으로, 트랜스진 서열은 발현시 검출 가능한 신호를 생성하는 리포터 서열을 포함할 수 있다. 상기 리포터 서열은, 비제한적인 예로, β -락타마제, β -갈락토시다제(LacZ), 알칼리성 포스파타제, 티미딘 키나제, 녹색 형광 단백질(GFP), 클로람페니콜 아세틸트랜스페라제(CAT), 루시페라제, 막 결합된 단백질, 예를 들어, CD2, CD4, CD8, 인플루엔자 헤마글루티닌 단백질, 및 특이적인 고 친화성 항체가 존재하거나 통상적인 수단에 의해 생성될 수 있는 당 분야에 널리 공지된 것들, 및 다른 것들 중에서 헤마글루티닌 또는 Myc로부터의 항원 태그 도메인에 적절하게 융합된 막 결합된 단백질을 포함하는 융합 단백질을 인코딩하는 DNA 서열을 포함한다. 이들 코딩 서열은 이들의 발현을 유도하는 조절 요소와 회합되는 경우 효소, 방사선, 비색, 형광 또는 다른 분광 검정, 형광 활성화 세포 분류 검정 및 면역학적 검정, 예를 들어, 효소 결합 면역

역흡착 측정(ELISA), 방사면역측정법(RIA) 및 면역조직화학을 포함하는 통상적인 수단에 의해 검출 가능한 신호를 제공한다.

- [0175] 일 구체예에서, 트랜스진은 생물학 및 의학에 유용한 생성물을 인코딩하는 비-마커 서열, 예를 들어, 치료용 트랜스진 또는 면역원성 트랜스진, 예를 들어, 단백질, RNA, 효소, 또는 촉매성 RNA이다. 바람직한 RNA 분자는 tRNA, dsRNA, 리보솜 RNA, 촉매성 RNA, 및 안티센스 RNA를 포함한다. 유용한 RNA 서열의 한 예는 치료되는 동물에서 표적화된 핵산 서열의 발현을 소멸시키는 서열이다.
- [0176] 트랜스진은 암 치료제 또는 백신으로서, 예를 들어, 유전적 결함의 치료를 위해, 면역 반응의 유도를 위해, 및/또는 예방 백신 목적을 위해 사용될 수 있다. 본원에서 사용되는 바와 같은 면역 반응의 유도는 단백질에 대한 T 세포 및/또는 체액성 면역 반응을 유도하는 단백질의 능력을 나타낸다.
- [0177] 예방이라는 용어는 미리 약제를 제공하는 것을 의미하며, 이는 병원체에 노출되기 전에(노출-전 예방) 또는 질병 증상이 발생하기 전에(노출-후 예방) 이루어질 수 있다. 치료 및 요법이라는 용어는 본원에서 상호교환적으로 사용되고 질병 중 약제의 투여를 의미한다.
- [0178] 질병이라는 용어는 대상체에서 구조 또는 기능의 장애, 특히 특정 증상을 유발하거나 특정 위치에 영향을 미치고 단순히 신체적 상해의 직접적인 결과가 아닌 것을 의미한다.
- [0179] 조절 요소
- [0180] 트랜스진에 더하여, 벡터는 또한 플라스미드 벡터로 트랜스팩션되거나, 본 발명에 의해 생성된 바이러스로 감염된 세포에서 트랜스진의 전사, 번역 및/또는 발현을 가능하게 하는 방식으로 트랜스진에 작동 가능하게 연결되는 통상적인 조절 요소를 포함한다. 본원에서 사용되는 "작동 가능하게 연결된" 서열은 관심 유전자와 인접한 발현 조절 서열 및 관심 유전자를 조절하기 위해 교차하여(in trans) 작용하거나 멀리서 작용하는 발현 조절 서열 둘 모두를 포함한다.
- [0181] 발현 조절 서열은 적절한 전사 개시, 종료, 프로모터 및 인핸서 서열; 효과적인 RNA 처리 신호, 예를 들어, 토끼 베타-글로빈 polyA를 포함하는 스플라이싱 및 폴리아데닐화(poly A) 신호; 세포질 mRNA를 안정화시키는 서열; 번역 효율을 향상시키는 서열(예를 들어, 코작(Kozak) 컨센서스 서열); 단백질 안정성을 향상시키는 서열; 및 요망되는 경우, 인코딩된 생성물의 분비를 향상시키는 서열을 포함한다. 다른 서열 중에서도, 키메라 인트론이 사용될 수 있다.
- [0182] 일부 구체예에서, 우드척 간염 바이러스 전사후 조절 요소(WPRE)(Zuffrey et al. (1999) *J Virol*; 73(4):2886-9)가 트랜스진에 작동 가능하게 연결될 수 있다. 예시적인 WPRE는 SEQ ID NO: 26에 제공된다.
- [0183] "프로모터"는 RNA 중합효소의 결합을 가능하게 하고, 유전자의 전사를 유도하는 뉴클레오티드 서열이다. 통상적으로, 프로모터는 유전자의 전사 시작 부위에 근접한 유전자의 5' 비-코딩 영역에 위치한다. 전사 개시에서 기능하는 프로모터 내의 서열 요소는 종종 컨센서스 뉴클레오티드 서열을 특징으로 한다. 프로모터의 예는 박테리아, 효모, 식물, 바이러스, 및 포유동물(인간을 포함함)로부터의 프로모터를 포함하나, 이에 제한되지는 않는다. 내부, 천연, 항시성, 유도성 및/또는 조직-특이적인 프로모터를 포함하는 매우 많은 수의 발현 조절 서열이 당 분야에 공지되어 있으며, 사용될 수 있다.
- [0184] 항시성 프로모터의 예는 TBG 프로모터, 레트로바이러스 라우스 육종 바이러스 LTR 프로모터(선택적으로, 인핸서를 가짐), 사이토메갈로바이러스(CMV) 프로모터(선택적으로, CMV 인핸서를 가짐, 예를 들어, 문헌[Boshart et al, *Cell*, 41:521-530 (1985)] 참조), CASI 프로모터, SV40 프로모터, 디하이드로폴레이트 환원효소 프로모터, β -액틴 프로모터, 포스포글리세롤 키나제(PGK) 프로모터, 및 EF1a 프로모터(Invitrogen)를 포함하나, 이에 제한되지는 않는다.
- [0185] 일부 구체예에서, 프로모터는 CASI 프로모터(예를 들어, W02012/115980호 참조)이다. CASI 프로모터는 CMV 인핸서의 일부, 닭 베타-액틴 프로모터의 일부, 및 UBC 인핸서의 일부를 함유하는 합성 프로모터이다. 일부 구체예에서, CASI 프로모터는 SEQ ID NO: 12와 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 또는 그 초과와 서열 동일성을 갖는 핵산 서열을 포함할 수 있다. 일부 구체예에서, 프로모터는 SEQ ID NO: 12의 핵산 서열을 포함하거나, 이로 구성된다.
- [0186] 유도성 프로모터는 유전자 발현의 조절을 가능하게 하며, 외인적으로 공급된 화합물, 환경 요인, 예를 들어, 온도, 또는 특정 생리학적 상태의 존재, 예를 들어, 급성기, 세포의 특정 분화 상태에 의해, 또는 복제하는 세포에서만 조절될 수 있다. 유도성 프로모터 및 유도성 시스템은 Invitrogen, Clontech 및 Ariad를 포함하나, 이에

제한되지는 않는 다양한 상업적 공급처로부터 이용 가능하다. 많은 다른 시스템이 기재되어 있으며, 당업자에 의해 용이하게 선택될 수 있다. 예를 들어, 유도성 프로모터는 아연-유도성 양 메탈로티오닌(MT) 프로모터 및 텍사메타손(Dex)-유도성 마우스 유암 바이러스(MMTV) 프로모터를 포함한다. 다른 유도성 시스템은 T7 중합효소 프로모터 시스템(WO 98/10088호); 액디슨 곤충 프로모터(No et al, Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), 테트라사이클린-억제성 시스템(Gossen et al, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)), 테트라사이클린-유도성 시스템(Gossen et al, Science, 378:1766-1769 (1995), 또한 Harvey et al, Curr. Opin. Chem. Biol, 2:512-518 (1998) 참조)을 포함한다. 다른 시스템은 FK506 이량체, 카스트라디올(castradiol), 디페놀 무리슬레론(diphenol murislerone)을 이용하는 VP16 또는 p65, RU486-유도성 시스템(Wang et al, Nat. Biotech., 15:239-243 (1997) 및 Wang et al, Gene Ther., 4:432-441 (1997)) 및 라파마이신-유도성 시스템(Magari et al, J. Clin. Invest., 100:2865-2872 (1997))을 포함한다. 일부 유도성 프로모터의 효과는 시간이 지남에 따라 증가한다. 이러한 경우에, 다수의 억제인자, 예를 들어, IRES에 의해 TetR에 연결된 TetR을 일렬로 삽입함으로써 상기 시스템의 효과를 향상시킬 수 있다.

[0187] 일부 구체예에서, 프로모터는 SEQ ID NO: 42에 제공된 것과 같은 증강된 hCMV 프로모터이다.

[0188] 또 다른 구체예에서, 트랜스진에 대한 천연 프로모터가 사용될 것이다. 천연 프로모터는 트랜스진의 발현이 천연 발현을 모방해야 하는 것이 요망되는 경우에 바람직할 수 있다. 천연 프로모터는 트랜스진의 발현이 시간적 또는 발생적으로, 또는 조직-특이적인 방식으로, 또는 특정 전사 자극에 반응하여 조절되어야 하는 경우에 사용될 수 있다. 추가의 구체예에서, 다른 천연 발현 조절 요소, 예를 들어, 인핸서 요소, 폴리아데닐화 부위 또는 코작 컨센서스 서열이 또한 천연 발현을 모방하기 위해 이용될 수 있다.

[0189] 트랜스진은 조직-특이적인 프로모터에 작동 가능하게 연결될 수 있다. 예를 들어, 골격근에서의 발현이 요망되는 경우, 근육에서 활성인 프로모터가 사용되어야 한다. 이들은 골격 β -액틴, 미오신 경쇄 2A, 디스트로핀, 근육 크레아틴 키나제를 인코딩하는 유전자로부터의 프로모터뿐만 아니라 자연 발생 프로모터보다 높은 활성을 갖는 합성 근육 프로모터를 포함한다(Li et al, Nat. Biotech., 17:241-245 (1999) 참조). 조직-특이적인 프로모터의 예는 다른 것들 중에서 간(알부민, Miyatake et al, J. Virol, 71:5124-32 (1997); B형 간염 바이러스 코어 프로모터, Sandig et al, Gene Ther., 3:1002-9 (1996); 알파-태아단백질(AFP), Arbuthnot et al., Hum. Gene Ther., 7: 1503-14 (1996)), 뼈 오스테오칼신(Stein et al, Mol. Biol. Rep., 24:185-96 (1997)); 뼈 시알로단백질(Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), 림프구(CD2, Hansal et al, J. Immunol, 161:1063-8 (1998); 면역글로불린 중쇄; T 세포 수용체 사슬), 신경세포, 예를 들어, 뉴런-특이적 에놀라제(NSE) 프로모터(Andersen et al, Cell. Mol. Neurobiol, 13:503-15 (1993)), 신경미세섬유 경쇄 유전자(Piccioli et al, Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), 및 뉴런-특이적 vgf 유전자(Piccioli et al, Neuron, 15:373-84 (1995))에 대해 공지되어 있다.

[0190] 선택적으로, 치료적으로 유용한 생성물 또는 면역원성 생성물을 인코딩하는 트랜스진을 운반하는 벡터는 다른 것들 중에서 제네티신, 하이그로마이신 또는 퓨로마이신 내성을 인코딩하는 서열을 포함할 수 있는 선택 가능한 마커 또는 리포터 유전자를 또한 포함할 수 있다. 앰피실린 내성과 같은 상기 선택 가능한 리포터 또는 마커 유전자(바람직하게는, 바이러스 입자로 패키징되는 바이러스 유전체 외부에 위치됨)는 박테리아 세포에서 플라스미드의 존재를 알리기 위해 사용될 수 있다. 벡터의 다른 성분은 복제 기점을 포함할 수 있다.

[0191] 이들 벡터는 당업자에게 공지된 기술과 함께 본원에 제공된 기술 및 서열을 이용하여 생성된다. 상기 기술은 본문에 기재된 것과 같은 cDNA의 통상적인 클로닝 기술, 아데노바이러스 유전체의 중첩 올리고뉴클레오티드 서열의 이용, 중합효소 연쇄 반응, 및 요망되는 뉴클레오티드 서열을 제공하는 임의의 적합한 방법을 포함한다.

[0192] 치료 및 예방

[0193] 재조합 ChAd157-기반 벡터는 시험관 내, 생체 외, 및 생체 내에서 인간 또는 비-유인원 포유동물로의 유전자 전달에 유용하다.

[0194] 본원에 기재된 재조합 아데노바이러스 벡터는 시험관 내에서 이중성 트랜스진에 의해 인코딩된 생성물의 생산을 위한 발현 벡터로 사용될 수 있다. 예를 들어, 트랜스진을 함유하는 재조합 복제-부적격 아데노바이러스는 상기 기재된 바와 같이 상보성 세포주로 트랜스펙션될 수 있다.

[0195] ChAd157-유래된 재조합 아데노바이러스 벡터는 유기체가 하나 이상의 아데노바이러스 혈청형에 대한 중화 항체를 갖는 경우에도 생체 내 또는 생체 외에서 선택된 숙주 세포로 선택된 트랜스진을 전달할 수 있는 효과적인 유전자 전달 비히클을 제공한다. 일 구체예에서, 벡터 및 세포는 생체 외에서 혼합되고; 감염된 세포는 통상적

인 방법을 이용하여 배양되고; 형질도입된 세포는 환자로 재주입된다. 이들 기술은 치료 목적 및 보호 면역 유도를 포함하는 면역화를 위한 유전자 전달에 특히 매우 적합하다.

[0196] 면역원성 트랜스진

[0197] 재조합 ChAd157 벡터는 또한 면역원성 조성물로 투여되는 바와 같을 수 있다. 본원에 기재된 바와 같은 면역원성 조성물은 포유동물, 적합하게는 인간으로의 전달 후 벡터에 의해 전달된 트랜스진 생성물에 대해 면역 반응, 예를 들어, 체액성(예를 들어, 항체) 및/또는 세포-매개(예를 들어, 세포독성 T 세포) 반응을 유도할 수 있는 하나 이상의 재조합 ChAd157 벡터를 포함하는 조성물이다. 재조합 아데노바이러스는 요망되는 면역원을 인코딩하는 유전자를 포함(적합하게는, 이의 유전자 결실 중 임의의 결실 내에 포함)할 수 있으며, 따라서 백신에서 사용될 수 있다. 재조합 아데노바이러스는 면역 반응의 유도에 중요하고, 병원체의 확산을 제한할 수 있는 항원(들)이 확인되었고, cDNA가 이용 가능한 임의의 병원체에 대한 예방 또는 치료 백신으로 사용될 수 있다.

[0198] 면역원이라는 용어는 면역 반응을 유도할 수 있는 폴리펩티드를 의미한다. 적합하게는, 면역원은 적어도 하나의 B 또는 T 세포 에피토프를 포함하는 항원이다. 유도된 면역 반응은 중화 항체를 생성하는 항원 특이적 B 세포 반응일 수 있다. 유도된 면역 반응은 항원 특이적 T 세포 반응일 수 있고, 이는 전신 및/또는 국소 반응일 수 있다. 항원 특이적 T 세포 반응은 CD4+ T 세포 반응, 예를 들어, IFN감마, TNF알파 및/또는 IL2와 같은 복수의 사이토카인을 발현하는 CD4+ T 세포를 수반하는 반응일 수 있다. 대안적으로 또는 추가로, 항원 특이적 T 세포 반응은 CD8+ T 세포 반응, 예를 들어, IFN감마, TNF알파 및/또는 IL2와 같은 복수의 사이토카인을 발현하는 CD8+ T 세포를 수반하는 반응을 포함한다.

[0199] 따라서 면역화한다는 용어는 면역 반응을 유도하기 위해 면역원(또는 상황에 적절하게 면역원을 인코딩하는 폴리뉴클레오티드)의 투여를 의미한다.

[0200] 상기 백신 또는 다른 면역원성 조성물은 적합한 전달 비히클에서 제형화될 수 있다. 일반적으로, 면역원성 조성물에 대한 용량은 하기 '전달 방법 및 투여량'에 정의된 범위 내이다. 선택된 유전자의 면역성의 수준은, 존재 시, 부스터에 대한 필요성을 결정하기 위해 모니터링될 수 있다. 혈청에서의 항체 역가의 평가 후, 선택적 부스터 면역화가 요망될 수 있다.

[0201] 선택적으로, 본 발명의 백신 또는 면역원성 조성물은, 예를 들어, 애주번트, 안정화제, pH 조정제, 보존제 등을 포함하는 다른 성분을 함유하도록 제형화될 수 있다. 적합한 애주번트의 예는 하기 '애주번트'에 제공된다. 상기 애주번트는 항원만 인코딩하는 DNA 백신으로 프라이밍시 생성된 면역 반응에 비해 항원-특이적 면역 반응을 향상시키기 위해 항원을 인코딩하는 프라이밍 DNA 백신과 함께 투여될 수 있다. 대안적으로, 상기 애주번트는 본 발명의 ChAd157 벡터를 포함하는 투여 요법에서 투여되는 폴리펩티드 항원과 함께 투여될 수 있다(하기 '투여 요법'에 기재됨).

[0202] 재조합 아데노바이러스는 면역원성 양, 즉, 요망되는 표적 세포를 트랜스펙션시키기 위해 투여 경로에서 효과적이고, 면역 반응을 유도하기 위해 선택된 유전자의 충분한 수준의 발현을 제공하는 재조합 아데노바이러스의 양으로 투여된다. 보호 면역이 제공되는 경우, 재조합 아데노바이러스는 감염 및/또는 재발성 질병을 예방하는데 유용한 백신 조성물인 것으로 간주된다.

[0203] 본원에 기재된 재조합 벡터는 세포용해성 T 세포 및 벡터에 의해 발현된 삽입된 이중 항원성 단백질에 대해 유도되는 항체를 유도하는데 매우 효과적일 것으로 예상된다.

[0204] 다른 병원체에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 본 발명의 벡터에 의해 발현되는 면역원은, 예를 들어, 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충을 포함하거나, 암세포 또는 종양 세포로부터 유래된다. 예를 들어, 면역원은 다양한 바이러스 과로부터 선택될 수 있다. 면역 반응이 요망되는 바이러스 과의 예는 리사바이러스, 예를 들어, 광견병 바이러스, 호흡기 바이러스, 예를 들어, 호흡기 세포 융합 바이러스(RSV) 및 다른 파라믹소바이러스, 예를 들어, 인간 메타뉴모바이러스, hMPV 및 파라인플루엔자 바이러스(PIV)를 포함한다.

[0205] 인간 또는 비인간 동물을 면역화하기 위한 면역원으로서 유용한 적합한 광견병 항원은 광견병 바이러스 당단백질(G), RNA 중합효소(L), 기질 단백질(M), 핵단백질(N) 및 인단백질(P)로부터 선택될 수 있다. 용어 "G 단백질" 또는 "당단백질" 또는 "G 단백질 폴리펩티드" 또는 "당단백질 폴리펩티드"는 광견병 당단백질 폴리펩티드의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다. 용어 "L 단백질" 또는 "RNA 중합효소 단백질" 또는 "L 단백질 폴리펩티드" 또는 "RNA 중합효소 단백질 폴리펩티드"는 광견병 RNA 중합효소 단백질 폴리펩티드의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다. 용어 "M 단백질" 또

는 "기질 단백질" 또는 "M 단백질 폴리펩티드" 또는 "기질 단백질 폴리펩티드"는 광견병 기질 단백질 폴리펩티드의 아미노산 서열의 일부 또는 전부를 갖는 폴리펩티드 또는 단백질을 나타낸다. 용어 "N 단백질" 또는 "핵단백질" 또는 "N 단백질 폴리펩티드" 또는 "핵단백질 폴리펩티드"는 광견병 핵단백질 폴리펩티드의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다. 용어 "P 단백질" 또는 "인단백질" 또는 "P 단백질 폴리펩티드" 또는 "인단백질 폴리펩티드"는 광견병 인단백질 폴리펩티드의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다.

[0206] 인간 또는 비인간 동물을 면역화시키기 위한 면역원으로 유용한 RSV의 적합한 항원은 융합 단백질(F), 부착 단백질(G), 기질 단백질(M2) 및 핵단백질(N)로부터 선택될 수 있다. 용어 "F 단백질" 또는 "융합 단백질" 또는 "F 단백질 폴리펩티드" 또는 "융합 단백질 폴리펩티드"는 RSV 융합 단백질 폴리펩티드의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다. 유사하게, 용어 "G 단백질" 또는 "G 단백질 폴리펩티드"는 RSV 부착 단백질 폴리펩티드의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다. 용어 "M 단백질" 또는 "기질 단백질" 또는 "M 단백질 폴리펩티드"는 RSV 기질 단백질의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타내며, M2-1(본원에서 M2.1로 기재될 수 있음) 및 M2-2 유전자 생성물 중 어느 하나 또는 둘 모두를 포함할 수 있다. 마찬가지로, 용어 "N 단백질" 또는 "뉴클레오캡시드 단백질" 또는 "N 단백질 폴리펩티드"는 RSV 핵단백질의 아미노산 서열의 전부 또는 일부를 갖는 폴리펩티드 또는 단백질을 나타낸다.

[0207] G 당단백질의 항원성에서의 차이를 주로 기초로 하여 A 및 B 그룹의 2개 그룹의 인간 RSV 균주가 기재되었다. RSV의 다수의 균주가 현재까지 단리되었고, 이 중 임의의 것이 본원에 개시된 면역원성 조합물의 항원의 상황에서 적합하다. GenBank 및/또는 EMBL 등록 번호에 의해 표시된 예시적인 균주는 US 공개 출원 번호 2010/0203071호(WO2008114149호)에서 발견될 수 있으며, 이는 본 발명에서 사용하기에 적합한 RSV F 및 G 단백질의 핵산 및 폴리펩티드 서열을 개시할 목적으로 참조로서 본원에 포함된다. 일 구체예에서, RSV F 단백질은 RSV F 단백질의 엑토도메인(F Δ TM)일 수 있다.

[0208] 예시적인 M 및 N 단백질 핵산 및 단백질 서열은, 예를 들어, US 공개 출원 번호 2014/0141042호(WO2012/089833호)에서 발견될 수 있으며, 이는 본 발명에서 사용하기에 적합한 RSV M 및 N 단백질의 핵산 및 폴리펩티드 서열을 개시할 목적으로 본원에 포함된다.

[0209] 적합하게는, 본 발명에서 사용하기 위해, 핵산은 RSV F 항원 및 RSV, M 및 N 항원을 인코딩한다. 더욱 특히, 핵산은 RSV F Δ TM 항원 및 RSV M2-1 및 N 항원을 인코딩하며, 여기서 자가-절단 부위가 RSV F Δ TM 항원과 RSV M2-1 사이에 포함되고, 가요성 링커가 RSV M2-1과 N 항원 사이에 포함된다. 일 구체예에서, 적합한 핵산은 SEQ ID NO:37에 의해 표시되는 폴리펩티드를 인코딩한다.

[0210] 일 구체예에서, 면역원은 레트로바이러스, 예를 들어, 렌티바이러스, 예를 들어, 인간 면역결핍 바이러스(HIV)로부터 유래될 수 있다. 상기 구체예에서, 면역원은 HIV-1 또는 HIV-2로부터 유래될 수 있다.

[0211] HIV 유전체는 다수의 상이한 단백질을 인코딩하며, 이들 각각은 본 발명의 벡터에 의해 발현되는 경우 전체적으로 또는 단편으로서 면역원성일 수 있다. 외피 단백질은, 예를 들어, gp120, gp41 및 Env 전구체 gp160을 포함한다. HIV의 비-외피 단백질은, 예를 들어, 내부 구조 단백질, 예를 들어, gag 및 pol 유전자의 생성물 및 다른 비-구조 단백질, 예를 들어, Rev, Nef, Vif 및 Tat을 포함한다. 일 구체예에서, 본 발명의 벡터는 HIV Gag를 포함하는 하나 이상의 폴리펩티드를 인코딩한다.

[0212] Gag 유전자는 전구체 다기능단백질로 번역되며, 이는 프로테아제에 의해 절단되어 기질 단백질(p17), 캡시드(p24), 뉴클레오캡시드(p9), p6 및 2개의 공간 펩티드인 p2 및 p1을 포함하는 생성물을 생성시키며, 이들 모두는 Gag 단편의 예이다.

[0213] Gag 유전자는 p55로도 언급되는 55-킬로달톤(kD) Gag 전구체 단백질을 발생시키며, 이는 스플라이싱되지 않은 바이러스 mRNA로부터 발현된다. 번역 동안, p55의 N 말단은 미리스토일화되어, 세포막의 세포질 측면과 이의 회합을 촉발시킨다. 막-회합된 Gag 다기능단백질은 감염된 세포의 표면으로부터 바이러스 입자의 발아를 촉발시키는 다른 바이러스 및 세포 단백질과 함께 바이러스 유전체 RNA의 2개 카피를 모집한다. 발아 후, p55는 바이러스 성숙 과정 동안 바이러스 인코딩된 프로테아제(pol 유전자의 생성물)에 의해 MA(기질 [p17]), CA(캡시드 [p24]), NC(뉴클레오캡시드 [p9]), 및 p6으로 명명된 4개의 더 작은 단백질로 절단되며, 이들 모두는 Gag 단편의 예이다. 일 구체예에서, 본 발명의 벡터는 SEQ ID NO: 16의 Gag 폴리펩티드를 포함한다.

[0214] 애주버트

- [0215] 본원에서 사용되는 "애주번트"는 면역원에 대한 면역 반응을 향상시키는 조성물을 나타낸다. 상기 애주번트의 예는 무기 애주번트(예를 들어, 무기 금속 염, 예를 들어, 알루미늄 포스페이트 또는 알루미늄 하이드록사이드), 유기 애주번트(예를 들어, 사포닌, 예를 들어, QS21, 또는 스쿠알렌), 오일-기반 애주번트(예를 들어, 프로인트 완전 애주번트 및 프로인트 불완전 애주번트), 사이토카인(예를 들어, IL-1 β , IL-2, IL-7, IL-12, IL-18, GM-CSF, 및 INF- γ) 미립자 애주번트(예를 들어, 면역-자극성 복합체(ISCOMS), 리포솜, 또는 생물분해성 미세구), 바이로솜, 박테리아 애주번트(예를 들어, 모노포스포릴 지질 A, 예를 들어, 3-데-O-아실화 모노포스포릴 지질 A(3D-MPL), 또는 뮤라일 펩티드), 합성 애주번트(예를 들어, 비-이온성 블록 공중합체, 뮤라일 펩티드 유사체, 또는 합성 지질 A), 합성 폴리뉴클레오티드 애주번트(예를 들어, 폴리아르기닌 또는 폴리리신) 및 메틸화되지 않은 CpG 디뉴클레오티드("CpG")를 함유하는 면역자극성 올리고뉴클레오티드를 포함하나, 이에 제한되지는 않는다.
- [0216] 한 적합한 애주번트는 모노포스포릴 지질 A(MPL), 특히 3-데-O-아실화 모노포스포릴 지질 A(3D-MPL)이다. 화학적으로, 이는 종종 3-데-O-아실화 모노포스포릴 지질 A와 4, 5 또는 6개의 아실화된 사슬의 혼합물로 제공된다. 이는 GB 2122204B호에 교시된 방법에 의해 정제되고 제조될 수 있으며, 상기 참조는 또한 디포스포릴 지질 A, 및 이의 3-O-데아실화된 변이체의 제조를 개시한다. 다른 정제된 지질다당류 및 합성 지질다당류가 기재되었다(U.S. 특허 번호 6,005,099호 및 EP 0 729 473 B1호; Hilgers et al., 1986, Int.Arch.Allergy.Immunol., 79(4):392-6; Hilgers et al., 1987, Immunology, 60(1):141-6; 및 EP 0 549 074 B1).
- [0217] 사포닌이 또한 적합한 애주번트이다(Lacaille-Dubois, M and Wagner H, A review of the biological and pharmacological activities of saponins. Phytomedicine vol 2 pp 363-386 (1996) 참조). 예를 들어, 사포닌 Quil A(남아메리카 나무 퀴라자 사포나리아 몰리나(Quillaja Saponaria Molina)의 껍질로부터 유래됨), 및 이의 분획이 U.S. 특허 번호 5,057,540호 및 문헌[Kensil, Crit. Rev. Ther. Drug Carrier Syst., 1996, 12:1-55]; 및 EP 0 362 279 B1호에 기재되어 있다. Quil A의 정제된 분획, 예를 들어, QS21 및 QS17이 또한 면역자극제로서 공지되어 있다; 이들의 제조 방법은 U.S. 특허 번호 5,057,540호 및 EP 0 362 279 B1호에 기재되어 있다. QS7(Quil-A의 비-용혈 분획)이 또한 이들 참고문헌에 기재되어 있다. QS21의 사용은 문헌[Kensil et al. (1991, J. Immunology, 146: 431-437)]에 추가로 기재되어 있다. QS21 및 폴리소르베이트 또는 사이클로덱스트린의 조합물이 또한 공지되어 있다(WO 99/10008호). QuilA의 분획, 예를 들어, QS21 및 QS7을 포함하는 미립자 애주번트 시스템이 WO 96/33739호 및 WO 96/11711호에 기재되어 있다.
- [0218] 또 다른 애주번트는 메틸화되지 않은 CpG 디뉴클레오티드("CpG")를 함유하는 면역자극성 올리고뉴클레오티드이다(Krieg, Nature 374:546 (1995)). CpG는 DNA에 존재하는 시토신-구아노신 디뉴클레오티드 모티프에 대한 약어이다. CpG는 전신 및 점막 경로 둘 모두로 투여되는 경우에 애주번트로 공지되어 있다(WO 96/02555호, EP 468520호, Davis et al, J.Immunol, 1998, 160:870-876; McCluskie and Davis, J.Immunol., 1998, 161:4463-6). CpG는 백신 내로 제형화될 때 유리 항원과 함께 유리 용액으로 투여되거나(WO 96/02555) 항원에 공유적으로 컨쥬게이션된 채로 투여되거나(WO 98/16247), 알루미늄 하이드록사이드와 같은 담체와 함께 제형화될 수 있다(Brazolot-Millan et al., Proc. Natl. Acad. Sci., USA, 1998, 95:15553-8).
- [0219] 상기 기재된 것과 같은 애주번트는 담체, 예를 들어, 리포솜, 수중유 에멀전, 및/또는 금속 염(알루미늄 하이드록사이드와 같은 알루미늄 염을 포함함)과 함께 제형화될 수 있다. 예를 들어, 3D-MPL은 알루미늄 하이드록사이드(EP 0 689 454호) 또는 수중유 에멀전(WO 95/17210호)과 함께 제형화될 수 있고; QS21은 콜레스테롤 함유 리포솜(WO 96/33739호), 수중유 에멀전(WO 95/17210호) 또는 백반(WO 98/15287호)과 함께 제형화될 수 있고; CpG는 백반(Brazolot-Millan, 상기) 또는 다른 양이온성 담체와 함께 제형화될 수 있다.
- [0220] 애주번트의 조합물, 특히 모노포스포릴 지질 A 및 사포닌 유도체의 조합물(예를 들어, WO 94/00153호; WO 95/17210호; WO 96/33739호; WO 98/56414호; WO 99/12565호; WO 99/11241호 참조), 더욱 특히 WO 94/00153호에 개시된 바와 같은 QS21 및 3D-MPL의 조합물, 또는 WO 96/33739호에 개시된 바와 같은 QS21이 콜레스테롤-함유 리포솜(DQ)에서 켄칭되는 조성물이 본 발명에서 사용될 수 있다. 대안적으로, CpG 및 사포닌, 예를 들어, QS21의 조합물이 본 발명에서 사용하기에 적합한 애주번트이다. 수중유 에멀전 중 QS21, 3D-MPL & 토코페롤을 포함하는 효능 있는 애주번트 제형이 WO 95/17210호에 기재되어 있으며, 본 발명에서 사용하기 위한 또 다른 제형이다. 사포닌 애주번트는 리포솜 내에서 제형화될 수 있고, 면역자극성 올리고뉴클레오티드와 조합될 수 있다. 따라서, 적합한 애주번트 시스템은, 예를 들어, 모노포스포릴 지질 A, 바람직하게는 3D-MPL과 함께 알루미늄 염의 조합물을 포함한다(예를 들어, WO 00/23105호에 기재됨). 추가의 예시적인 애주번트는 QS21 및/또는 MPL 및/또는 CpG를 포함한다. QS21은 WO 96/33739호에 개시된 바와 같이 콜레스테롤-함유 리포솜 내에서 켄칭될 수 있다.

- [0221] 다른 적합한 애주번트는 알킬 글루코사미니드 포스페이트(AGP), 예를 들어, W09850399호 또는 U.S. 특허 번호 6,303,347호에 개시된 것(AGP의 제조를 위한 공정이 또한 개시됨), 또는 U.S. 특허 번호 6,764,840호에 개시된 바와 같은 AGP의 약학적으로 허용되는 염을 포함한다. 일부 AGP는 TLR4 효능제이고, 일부는 TLR4 길항제이다. 둘 모두가 애주번트로서 유용한 것으로 생각된다.
- [0222] 백신화에서 사용되는 발현 시스템에 의해 포함되는 항원에 대한 불변 사슬의 융합이 아데노바이러스와 함께 투여되는 경우 상기 항원에 대한 면역 반응을 증가시키는 것으로 밝혀졌다(US 2011/0293704호로 공개되고, 불변 사슬 서열을 개시하기 위한 참조로서 포함되는 WO 2007/062656호). 따라서, 본 발명의 일 구체예에서, 면역원성 트랜스진은 재조합 ChAd157 바이러스 벡터에서 불변 사슬과 공동-발현될 수 있다.
- [0223] 또 다른 구체예에서, 본 발명은 면역조절 효과 반응을 유도하거나, 대상체에 ChAd157 캡시드를 전달함으로써 또 다른 활성제에 대한 세포독성 T 세포 반응을 향상시키거나 애주번트화시키기 위한 ChAd157의 캡시드(선택적으로, 무손상 또는 재조합 바이러스 입자 또는 빈 캡시드가 사용됨)의 용도를 제공한다. ChAd157 캡시드는 단독으로 또는 이에 대한 면역 반응을 향상시키기 위한 활성제와 함께 조합 요법으로 전달될 수 있다. 유리하게는, 요망되는 효과가 숙주를 아데노바이러스로 감염시키지 않고 달성될 수 있다.
- [0224] 투여 요법
- [0225] 일반적으로, ChAd157 재조합 아데노바이러스 벡터는 치료용 또는 면역원성 분자(예를 들어, 단백질)의 전달에 사용될 것이다. 본 발명의 재조합 아데노바이러스 벡터가 재조합 아데노바이러스 벡터의 반복 전달을 포함하는 요법에서 사용하기에 특히 매우 적합하다는 것이 둘 모두의 적용에서 용이하게 이해될 것이다. 상기 요법은 통상적으로 바이러스 캡시드가 교대로 반복되는 일련의 바이러스 벡터의 전달을 포함한다. 바이러스 캡시드는 각각의 후속 투여에 대해, 또는 특정 혈청형 캡시드의 미리 선택된 투여 횟수(예를 들어, 1회, 2회, 3회, 4회 또는 그 초과) 후에 변경될 수 있다. 따라서, 요법은 제1 캡시드를 갖는 재조합 아데노바이러스의 전달, 제2 캡시드를 갖는 재조합 아데노바이러스의 전달, 및 제3 캡시드를 갖는 재조합 아데노바이러스의 전달을 포함할 수 있다. 본 발명의 아데노바이러스 캡시드를 단독으로, 서로 조합하여, 또는 다른 아데노바이러스(바람직하게는, 면역학적으로 비-교차반응성인 아데노바이러스)와 조합하여 사용하는 다양한 다른 요법이 당업자에게 명백할 것이다. 선택적으로, 상기 요법은 다른 비인간 영장류 아데노바이러스, 인간 아데노바이러스, 또는 본원에 기재된 것과 같은 인공 서열의 캡시드를 갖는 재조합 아데노바이러스의 투여를 포함할 수 있다.
- [0226] 본 발명의 아데노바이러스 벡터는 트랜스진의 다수의 아데노바이러스-매개 전달이 요망되는 치료 요법, 예를 들어, 동일한 트랜스진의 재전달을 포함하는 요법 또는 다른 트랜스진의 전달을 포함하는 조합 요법에 특히 매우 적합하다. 상기 요법은 ChAd157 아데노바이러스 벡터의 투여 후, 동일 혈청형 아데노바이러스로부터의 벡터를 이용한 재투여를 포함할 수 있다. 특히 바람직한 요법은 ChAd157 아데노바이러스 벡터의 투여를 포함하며, 여기서 제1 투여에서 전달되는 벡터의 아데노바이러스 캡시드 서열의 공급원은 1회 이상의 후속 투여에서 사용되는 바이러스 벡터의 아데노바이러스 캡시드 서열의 공급원과 상이하다. 예를 들어, 치료 요법은 ChAd157 벡터의 투여 및 동일하거나 상이한 혈청형의 하나 이상의 아데노바이러스 벡터를 이용한 반복 투여를 포함한다.
- [0227] 또 다른 예에서, 치료 요법은 아데노바이러스 벡터의 투여 후, 첫 번째 전달된 아데노바이러스 벡터 내의 캡시드의 공급원과 상이한 캡시드를 갖는 ChAd157 벡터를 이용한 반복 투여, 및 선택적으로 이전 투여 단계에서 벡터의 아데노바이러스 캡시드의 공급원과 동일하거나 바람직하게는 상이한 또 다른 벡터를 이용한 추가 투여를 포함한다. 이들 요법은 ChAd157 서열을 이용하여 작제된 아데노바이러스 벡터의 전달로 제한되지 않는다. 오히려, 이들 요법은 ChAd157 벡터와 함께, 다른 비인간 영장류 아데노바이러스 서열, 또는 인간 아데노바이러스 서열을 포함하는 다른 아데노바이러스 서열을 비제한적으로 포함하는, 다른 아데노바이러스 서열을 용이하게 이용할 수 있다.
- [0228] 추가 예에서, 치료 요법은 (i) 하나 이상의 ChAd157 아데노바이러스 벡터 및 (ii) 추가 성분, 예를 들어, 비-아데노바이러스 벡터, 비-바이러스 벡터, 및/또는 애주번트와 선택적으로 동시에 투여되는 다양한 다른 치료적으로 유용한 화합물 또는 분자, 예를 들어, 항원성 단백질의 동시(예를 들어, 공동-투여) 또는 순차적(예를 들어, 프라임-부스트) 전달을 포함할 수 있다. 공동-투여의 예는 동측성 공동-투여 및 대측성 공동-투여(하기 '전달 방법 및 투여량'에 추가로 기재됨)를 포함한다.
- [0229] 하나 이상의 ChAd157 아데노바이러스 벡터와 동시 전달 또는 특히 순차적 전달(예를 들어, 프라임-부스트)에서 사용하기에 적합한 비-아데노바이러스 벡터는 하나 이상의 포스바이러스 벡터를 포함한다. 적합하게는, 포스바이러스 벡터는 아파 코르도포스바이러스아과, 더욱 적합하게는 오르토폭스(orthopox), 파라포스(parapox), 야타

폭스(yatapox), 아비폭스(avipox)(적합하게는, 카나리폭스(canarypox)(ALVAC) 또는 조류폭스(FPV)) 및 몰루시폭스(molluscipox)로 구성된 군으로부터 선택되는 상기 아과 내의 속에 속한다. 더욱 더 적합하게는, 폭스바이러스 백신은 오르토폭스에 속하며, 백신이나 바이러스, NYVAC(백신니아의 코펜하겐(Copenhagen) 군주로부터 유래됨), 변형된 백신니아 양카라(MVA), 우두바이러스 및 원숭이폭스 바이러스로 구성된 군으로부터 선택된다. 가장 적합하게는, 폭스바이러스 백신은 MVA이다.

[0230] "동시" 투여는 적합하게는 동일 진행 면역 반응을 나타낸다. 바람직하게는, 둘 모두의 성분은 동시에(예를 들어, DNA 및 단백질 둘 모두의 동시 투여) 투여되나, 한 성분은 수분 내에(예를 들어, 동일한 의학적 약속 또는 의사 방문시), 수시간 내에 투여될 수 있다. 상기 투여는 또한 공동-투여로 언급된다. 일부 구체예에서, 공동-투여는 아데노바이러스 백신, 애쥬번트 및 단백질 성분의 투여를 나타낼 수 있다. 다른 구체예에서, 공동-투여는 아데노바이러스 백신 및 또 다른 바이러스 백신, 예를 들어, 제2의 아데노바이러스 백신 또는 폭스바이러스, 예를 들어, MVA의 투여를 나타낸다. 다른 구체예에서, 공동-투여는 아데노바이러스 백신 및 선택적으로 애쥬번트화되는 단백질 성분의 투여를 나타낸다.

[0231] 프라이밍-부스트 요법이 이용될 수 있다. 프라이밍-부스트는 (i) 면역계의 최초 프라이밍 후, (ii) 일차 면역 반응이 확립된 후 수주 또는 수개월 후의 면역계의 이차 또는 부스팅이라는 두 별개의 면역 반응을 나타낸다.

[0232] 상기 요법은 면역계를 프라이밍하기 위한 재조합 ChAd157 백신의 투여와 전통적인 항원, 예를 들어, 단백질(선택적으로, 애쥬번트와 공동-투여됨), 또는 상기 항원을 인코딩하는 서열을 갖는 재조합 바이러스를 이용한 제2의 부스터 투여를 포함할 수 있다(예를 들어, WO 00/11140호). 대안적으로, 면역화 요법은 항원을 인코딩하는 백신(바이러스 또는 DNA-기반)에 대한 면역 반응을 부스팅하기 위해 재조합 ChAd157 백신의 투여를 포함할 수 있다. 또 다른 대안에서, 면역화 요법은 단백질의 투여 후, 항원을 인코딩하는 재조합 ChAd157 백신을 이용한 부스터를 포함한다. 한 예에서, 프라이밍-부스트 요법은 항원이 유래되는 바이러스, 박테리아 또는 다른 유기체에 대한 보호 면역 반응을 제공할 수 있다. 또 다른 구체예에서, 프라이밍-부스트 요법은 치료가 제공되는 질환의 존재의 검출을 위한 통상적인 검정을 이용하여 측정될 수 있는 치료 효과를 제공한다.

[0233] 바람직하게는, 부스팅 조성물은 대상체에 프라이밍 조성물을 투여한 후 약 2 내지 약 27주 후에 투여된다. 부스팅 조성물의 투여는 프라이밍 백신에 의해 투여되는 것과 동일한 항원 또는 상이한 항원을 전달할 수 있거나 이를 함유하는 유효량의 부스팅 조성물을 이용하여 달성된다. 부스팅 조성물은 동일한 바이러스 공급원 또는 또 다른 공급원으로부터 유래되는 재조합 바이러스 백신으로 구성될 수 있다. 대안적으로, 부스팅 조성물은 프라이밍 백신에서 인코딩된 것과 동일한 항원을 함유하나, 단백질 형태인 조성물일 수 있으며, 이러한 조성물은 숙주에서 면역 반응을 유도한다. 부스팅 조성물의 주요 필요조건은 조성물의 항원이 프라이밍 조성물에 의해 인코딩되는 것과 동일한 항원이거나 교차-반응성 항원이라는 것이다.

[0234] ChAd157에 대한 중화 항체와 ChAd155와 같은 특정 다른 아데노바이러스 백신 간의 낮은 교차-반응성은 다중 백신 투여가 요구되는 상황에서 유리하다. 다중 투여는 상이한 트랜스진(예를 들어, 상이한 의학적 적응증과 관련된 면역원을 인코딩함)의 분리된 전달 또는 동일하거나 유사한 트랜스진의 전달(예를 들어, 특정 의학적 적응증에 대한 면역 반응을 증가시키기 위한 프라이밍-부스트 요법에서)을 목적으로 할 수 있다.

[0235] 결과적으로, 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 백신에 이미 노출된 대상체에게 투여하기 위한 트랜스진을 인코딩하는 본 발명의 재조합 아데노바이러스 백신이 제공된다(예를 들어, 본원에 기재된 ChAd157 섬유, 핵소 또는 펜톤을 포함하지 않음, 예를 들어, ChAd155 섬유, 핵소 및/또는 펜톤을 포함하는 재조합 아데노바이러스 백신, 특히 ChAd155 섬유, 핵소 및 펜톤을 포함하는 재조합 아데노바이러스 백신). 특히, 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 백신이 이미 투여된 대상체에게 투여하기 위한 트랜스진을 인코딩하는 본 발명의 재조합 아데노바이러스 백신이 제공된다(예를 들어, 본원에 기재된 ChAd157 섬유, 핵소 또는 펜톤을 포함하지 않음, 예를 들어, ChAd155 섬유, 핵소 및/또는 펜톤을 포함하는 재조합 아데노바이러스 백신, 특히 ChAd155 섬유, 핵소 및 펜톤을 포함하는 재조합 아데노바이러스 백신). 적합하게는, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 백신은 ChAd157과 낮은 교차-반응성을 갖는 백신이다. 일 구체예에서, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 백신은 본 발명의 트랜스진의 재조합 아데노바이러스 백신과 다른 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩한다. 다른 구체예에서, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 백신은 본 발명의 트랜스진(예를 들어, 동일한 트랜스진)의 재조합 아데노바이러스 백신과 동일한 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩한다.

[0236] 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 백신에 후속하여

노출될 수 있는(즉, 의도되거나 예상되는) 대상체에게 투여하기 위한 트랜스진을 인코딩하는 본 발명의 재조합 아데노바이러스 벡터가 제공된다(예를 들어, 본원에 기재된 ChAd157 섬유, 핵손 또는 펜톤을 포함하지 않음, 예를 들어, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터, 특히 ChAd155 섬유, 핵손 및 펜톤을 포함하는 재조합 아데노바이러스 벡터). 특히, 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터가 후속하여 투여될 수 있는 대상체에게 투여하기 위한 트랜스진을 인코딩하는 본 발명의 재조합 아데노바이러스 벡터가 제공된다(예를 들어, 본원에 기재된 ChAd157 섬유, 핵손 또는 펜톤을 포함하지 않음, 예를 들어, ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터, 특히 ChAd155 섬유, 핵손 및 펜톤을 포함하는 재조합 아데노바이러스 벡터). 적합하게는, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 벡터는 ChAd157과 낮은 교차-반응성을 갖는 벡터이다. 일 구체예에서, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 벡터는 본 발명의 트랜스진의 재조합 아데노바이러스 벡터와 다른 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩한다. 다른 구체예에서, ChAd157 섬유를 포함하지 않는 재조합 아데노바이러스 벡터는 본 발명의 트랜스진(예를 들어, 동일한 트랜스진)의 재조합 아데노바이러스 벡터와 동일한 의학적 적응증 또는 적응증들을 겨냥하는 트랜스진을 인코딩한다.

- [0237] 따라서, 본 발명은 대상체에서 면역 반응을 유도하는 방법을 제공하고, 상기 방법은,
- [0238] (a) 대상체에게 제1 트랜스진을 인코딩하는 본 발명의 재조합 아데노바이러스 벡터를 투여하는 단계; 및
- [0239] (b) 대상체에게 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터를 투여하는 단계로서, 상기 벡터가 제2 트랜스진을 인코딩하는, 단계를 포함하고,
- [0240] 여기서 단계 (a) 및 (b)는 어느 한 순서로 수행될 수 있고 제1 및 제2 트랜스진은 동일하거나 상이할 수 있다.
- [0241] 제1 및 제2 트랜스진은 통상적으로 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩할 것이다. 제1 및 제2 트랜스진은 동일하거나 상이한 면역원을 인코딩할 수 있다. 상이한 면역원을 인코딩할 때, 이들은 동일하거나 상이한 병원체 또는 암 세포 또는 종양 세포에 대해 유도될 수 있다.
- [0242] 결과적으로, 대상체의 예방 또는 치료를 위한 방법이 또한 제공되고, 상기 방법은,
- [0243] (a) 대상체에게 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩하는 제1 트랜스진을 인코딩하는 본 발명의 재조합 아데노바이러스 벡터를 투여하는 단계; 및
- [0244] (b) 대상체에게 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터를 투여하는 단계로서, 상기 벡터가 인간 및 비인간 척추동물을 감염시키는 박테리아, 진균, 기생 미생물 또는 다세포 기생충과 같은 상이한 병원체에 대해, 또는 암세포 또는 종양 세포에 대해 인간 또는 비인간 동물을 면역화시키는데 유용한 면역원을 인코딩하는 제2 트랜스진을 인코딩하는, 단계를 포함하고,
- [0245] 여기서 단계 (a) 및 (b)는 어느 한 순서로 수행될 수 있다.
- [0246] 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터는 적합하게는 ChAd157 섬유, ChAd157 핵손 또는 ChAd157 섬유를 포함하지 않고, 예를 들어, ChAd157 섬유, ChAd157 핵손 또는 ChAd157 섬유 또는 이에 적어도 98% 동일성을 갖는 이의 기능성 유도체를 포함하지 않는다.
- [0247] 본원에 기재된 ChAd157 섬유 또는 이의 기능성 유도체를 포함하지 않는 재조합 아데노바이러스 벡터는 ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터, 특히 ChAd155 섬유, 핵손 및 펜톤을 포함하는 재조합 아데노바이러스 벡터일 수 있다.
- [0248] 언급된 대로, 본 발명의 재조합 아데노바이러스 벡터는 ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터와 함께 치료적 또는 면역원성 분자의 전달에 사용될 수 있다. ChAd155 섬유, 핵손 및/또는 펜톤을 포함하는 재조합 아데노바이러스 벡터는 SEQ ID NO: 7, 9 및 11에 따른 섬유, 펜톤 및/또는 핵손, 특히 SEQ ID NO: 7, 9 및 11에 따른 섬유, 펜톤 및 핵손을 포함할 것이다.
- [0249] 낮은 교차-반응성이라는 용어는 제1 벡터에 의한 면역화가 제2 벡터에 대한 주목할 만한 중화 항체 반응을 유도하지 않고, 즉, 제2 벡터의 면역학적 효능에 유의한 영향을 미치지 않는다는 것을 의미한다. 중화 항체 반응은 본원의 실시예 7과 유사한 방법으로 결정될 수 있다. 바람직하게는, 제1 벡터에 의한 면역화는 제2 벡터에 의한

면역화로부터 발생하는 수준의 평균 50% 미만, 예를 들어, 75% 미만, 적합하게는 90% 미만의 중화 역가를 유도한다.

[0250] 용어 "대상체"는 임의의 동물, 적합하게는 포유동물, 특히 인간을 의미한다.

[0251] 전달 방법 및 투여량

[0252] 벡터는 약학적으로 또는 생리학적으로 허용되는 담체, 예를 들어, 등장성 염수; 당업자에게 명백할 등장성 염 용액 또는 다른 제형에 현탁되거나 용해됨으로써 투여를 위해 제조될 수 있다. 적절한 담체는 당업자에게 명백할 것이며, 투여 경로에 대부분 좌우될 것이다. 본원에 기재된 조성물은 생물분해성 생체적합성 중합체를 이용하는 서방형 제형으로 포유동물에 투여될 수 있거나, 마이셀, 젤 및 리포솜을 이용하여 현장 전달될 수 있다.

[0253] 일부 구체예에서, 본 발명의 재조합 아데노바이러스는 근내 주사, 질내 투여, 정맥내 주사, 복강내 주사, 피하 주사, 경피 투여, 피내 투여, 코 투여, 직장 투여 또는 경구 투여에 의해 대상체에 투여된다. 설하 투여도 관심의 대상이 될 수 있다.

[0254] 치료 요법이 각각 상이한 조성물에 제형화된 하나 이상의 ChAd157 아데노바이러스 벡터 및 추가 성분의 공동-투여를 포함하는 경우, 이들은 바람직하게는 동일 부위 또는 그 부근에서 공동-국소적으로 투여된다. 예를 들어, 성분은 동일 측면 또는 말단("공동-측면" 투여) 또는 반대 측면 또는 말단("대측성" 투여)으로 투여(예를 들어, 근내, 경피, 피내, 피하로부터 선택된 투여 경로를 포함)될 수 있다.

[0255] 바이러스 벡터의 투여량은 주로 치료되는 질환, 환자의 연령, 체중 및 건강과 같은 요인에 좌우될 것이며, 따라서 환자마다 다를 수 있다. 예를 들어, 바이러스 벡터의 치료적으로 효과적인 성인 인간 또는 수의학적 투여량은 일반적으로 1×10^5 내지 1×10^{15} 개의 바이러스 입자, 예를 들어, 1×10^8 내지 1×10^{12} 개(예를 들어, 1×10^8 개, 2.5×10^8 개, 5×10^8 개, 1×10^9 개, 1.5×10^9 개, 2.5×10^9 개, 5×10^9 개, 1×10^{10} 개, 1.5×10^{10} 개, 2.5×10^{10} 개, 5×10^{10} 개, 1×10^{11} 개, 1.5×10^{11} 개, 2.5×10^{11} 개, 5×10^{11} 개, 1×10^{12} 개의 입자)를 함유한다. 대안적으로, 바이러스 벡터는 통상적으로 1×10^5 내지 1×10^{10} 플라크 형성 단위(PFU), 예를 들어, 1×10^5 PFU, 2.5×10^5 PFU, 5×10^5 PFU, 1×10^6 PFU, 2.5×10^6 PFU, 5×10^6 PFU, 1×10^7 PFU, 2.5×10^7 PFU, 5×10^7 PFU, 1×10^8 PFU, 2.5×10^8 PFU, 5×10^8 PFU, 1×10^9 PFU, 2.5×10^9 PFU, 5×10^9 PFU, 또는 1×10^{10} PFU의 용량으로 투여될 수 있다. 투여량은 동물의 크기 및 투여 경로에 따라 달라질 것이다. 예를 들어, 근내 주사를 위한 적합한 인간 또는 수의학적 투여량(약 80 kg 동물의 경우)은 단일 부위에 대해 1 mL 당 약 1×10^9 내지 약 5×10^{12} 개의 입자 범위 내이다. 선택적으로, 다수의 투여 부위가 이용될 수 있다. 또 다른 예에서, 적합한 인간 또는 수의학적 투여량은 경구 제형에 대해 약 1×10^{11} 내지 약 1×10^{15} 개의 입자 범위 내일 수 있다.

[0256] 바이러스 벡터는, 예를 들어, HCMV 프로모터를 포함하는 발현 카세트와 함께 벡터 유전체를 함유하는 플라스미드 DNA의 연속 희석액을 표준 곡선으로서 이용하여 CMV 프로모터 영역에 대해 설계된 프라이머 및 프로브를 이용한 정량적 PCR 분석(Q-PCR)에 의해 정량될 수 있다. 시험 샘플 내의 카피수는 평행선 분석 방법에 의해 결정된다. 벡터 입자 정량을 위한 대안적인 방법은 분석 HPLC 또는 A_{260} nm에 기반한 분광광도 방법일 수 있다.

[0257] 핵산의 면역학적 유효량은 적합하게는 1 ng 내지 100 mg일 수 있다. 예를 들어, 적합한 양은 1 μ g 내지 100 mg일 수 있다. 특정 핵산(예를 들어, 벡터)의 적절한 양은 당업자에 의해 용이하게 결정될 수 있다. 핵산 성분의 예시적인 유효량은 1 ng 내지 100 μ g, 예를 들어, 1 ng 내지 1 μ g(예를 들어, 100 ng-1 μ g), 또는 1 μ g 내지 100 μ g, 예를 들어, 10 ng, 50 ng, 100 ng, 150 ng, 200 ng, 250 ng, 500 ng, 750 ng, 또는 1 μ g일 수 있다. 핵산의 유효량은 또한 1 μ g 내지 500 μ g, 예를 들어, 1 μ g 내지 200 μ g, 예를 들어, 10 내지 100 μ g, 예를 들어, 1 μ g, 2 μ g, 5 μ g, 10 μ g, 20 μ g, 50 μ g, 75 μ g, 100 μ g, 150 μ g, 또는 200 μ g을 포함할 수 있다. 대안적으로, 핵산의 예시적인 유효량은 100 μ g 내지 1 mg, 예를 들어, 100 μ g 내지 500 μ g, 예를 들어, 100 μ g, 150 μ g, 200 μ g, 250 μ g, 300 μ g, 400 μ g, 500 μ g, 600 μ g, 700 μ g, 800 μ g, 900 μ g 또는 1 mg일 수 있다.

[0258] 일반적으로, 인간 용량은 0.1 ml 내지 2 ml의 부피 내일 것이다. 따라서, 본원에 기재된 조성물은, 예를 들어, 개별적 또는 조합된 면역원성 성분 당 0.1, 0.15, 0.2, 0.5, 1.0, 1.5 또는 2.0 ml 인간 용량의 부피로 제형화될 수 있다.

- [0259] 당업자는 투여 경로 및 재조합 벡터가 사용되는 치료 또는 백신 적용에 따라 상기 용량을 조정할 수 있다. 트랜스진의 발현 수준, 또는 애주번트에 대해, 순환 항체의 수준은 투여량의 투여 빈도를 결정하기 위해 모니터링될 수 있다.
- [0260] 하나 이상의 프라이밍 및/또는 부스팅 단계가 이용되는 경우, 이러한 단계는 매시간, 매일, 매주 또는 매달, 또는 매년 투여되는 단일 용량을 포함할 수 있다. 예로서, 포유동물은 담체 중 약 10 µg 내지 약 50 µg의 플라즈미드를 함유하는 1회 또는 2회 용량을 투여받을 수 있다. 전달 양 또는 부위는 바람직하게는 포유동물의 정체 및 상태를 기초로 하여 선택된다.
- [0261] 선택된 트랜스진에 의해 인코딩되는 단백질의 치료 수준 또는 이에 대한 면역 반응의 수준은, 존재시, 부스터에 대한 필요성을 결정하기 위해 모니터링될 수 있다. 혈청에서의 CD8+ T 세포 반응, 또는 선택적으로, 항체 역가의 평가 후, 선택적 부스터 면역화가 요망될 수 있다. 선택적으로, 재조합 ChAd157 벡터는 단일 투여 또는 다양한 조합 요법, 예를 들어, 다른 활성 성분을 포함하는 요법 또는 치료 과정과 조합하여 또는 프라임-부스트 요법으로 전달될 수 있다.
- [0262] 본 발명은 이제 하기의 비제한적인 실시예에 의해 추가로 기재될 것이다.
- [0263] **실시예**
- [0264] **실시예 1: ChAd157의 단리 및 벡터 작제**
- [0265] 29개의 상이한 야생형 침팬지 아데노바이러스를 아데노바이러스 단리 및 특성화 기술을 설명하기 위한 목적으로 본원에 참조로서 포함된 문헌[Colloca *et al.* Sci Transl Med. 2012 Jan 4;4(115):115ra2 and W02010/086189]에 기술된 표준 절차를 사용하여 다른 유럽 시설에 수용된 건강한 어린 침팬지로부터 단리시켰다.
- [0266] 29개의 야생형 바이러스를 후속하여 풀링하고; 풀의 바이러스 유전체를 E. 콜리(*E. coli*) BJ5183 세포에서 BAC 서들을 사용하여 상동 재조합에 의해 클로닝하여 E1 영역의 결실을 갖는 벡터의 미니라이브리리를 생성하였다. ΔE1 벡터의 미니라이브리리를 Procell 92 세포주 내로 트랜스펙션시켰다; 구제된 벡터를 16회 계대 감염 동안 연속적으로 계대시켰다. 16회 계대배양의 바이러스 DNA를 증폭된 벡터로부터 제조하고 E. 콜리 BJ5183에서 플라즈미드 서들을 사용하여 상동 재조합에 의해 클로닝하였다. 우세한 벡터 종은 ChAd157 ΔE1 벡터로서 동정되었고 후속하여 벡터 백본의 다음 추가 변형을 포함하도록 변형되었다:
- [0267] a) ΔE1 바이러스의 E4 영역(bp 34413 내지 bp 37127)의 결실;
- [0268] b) 인간 Ad5로부터 유래된 E4orf6의 삽입.
- [0269] 1.1: ΔE1 미니라이브리리 생성
- [0270] 29개의 야생형 바이러스의 풀을 사용하여 풀링된 바이러스 유전체를 얻었다. 풀링된 바이러스 유전체를 풀링된 바이러스 DNA 및 하위그룹 C BAC 서들(#1365)(SEQ ID NO: 14)로 공동-형질전환된 E. 콜리 균주 BJ5183에서 상동 재조합에 의해 BAC 벡터 내로 클로닝하였다. 도 2의 개략도에 도시된 대로, 하위그룹 C 서들은 종 C에 속하는 ChAd의 클로닝에 전용인 BAC 벡터이므로 pIX 유전자 및 종 C ChAd 바이러스의 우측 및 좌측 말단(우측 및 좌측 ITR 포함)에서 유래된 DNA 단편을 포함한다.
- [0271] 종 C BAC 서들은 또한 좌측 말단과 pIX 유전자 사이에 삽입된 RpsL-Kana 카세트를 함유한다. 또한, ISceI 제한 부위에 측점된 Amp-LacZ-SacB 선택 카세트가 pIX 유전자와 바이러스 유전체의 우측 말단 사이에 존재한다. 특히, BAC 서들은 다음과 같은 특징부를 포함하였다: 좌측 ITR: bp 27 내지 139, hCMV(tetO) RpsL-Kana 카세트: bp 493 내지 3396, pIX 유전자: bp 3508 내지 3972, ISceI 제한 부위: bp 3990 및 7481, Amp-LacZ-SacB 선택 카세트: bp 4000 내지 7471, 우측 ITR: bp 7805 내지 7917. hCMV(tetO)는 SEQ ID NO: 37에 제공된다.
- [0272] BJ5183 세포를 정제된 바이러스 DNA의 풀 및 ISceI 제한 효소로의 분해 후 젤로부터 정제된 하위그룹 C BAC 서들 벡터로 전기천공에 의해 공동-형질전환시켰다. pIX 유전자 및 우측 ITR 서열(종 C BAC 서들 선형화 DNA의 말단에 존재함)과 풀링된 바이러스 DNA에 존재하는 동종 서열 사이에서 발생하는 상동 재조합은 BAC 서들 벡터 내에서의 상이한 바이러스 유전체 DNA의 삽입을 발생시켰다. 동시에, 바이러스 E1영역을 결실시키고 RPSL-Kana 카세트로 치환하여, BAC/Minilibrary ΔE1/TetO hCMV RpsL-Kana를 생성하였다.
- [0273] 1.2: Procell 92 세포주에서 ΔE1 미니라이브리리 증폭 및 ChAd157 ΔE1 벡터의 클로닝.

- [0274] ΔE1 미니라이브러리를 PmeI로 분해하고 Procell 92 패키징 세포주를 트랜스펙션시키는데 사용하여, 상이한 바이러스의 라이브러리를 대량으로 구제하였다. 트랜스펙션 10일 후에, 세포를 수확하고 세포 용해물을 3회 사이클의 동결(-70℃) 및 해동(+ 37℃)으로 처리하고, 2000 rpm에서 원심분리에 의해 정화시키고, 신선한 세포를 감염시키는데 사용하였다. 바이러스 증폭의 16회 연속 계대배양을 수행하여, Procell92 세포에서의 증식 효율을 위한 바이러스 종을 선택하였다. 16회 계대배양의 바이러스(들)를 2회 CsCli 구배 원심분리에 의해 정제하고 바이러스 DNA를 추출하고 E. 콜리 BJ5183 세포에서 플라스미드 서들을 사용하여 상동 재조합에 의해 클로닝하였다. 상세하게는, BJ5183 세포를 정제된 바이러스 DNA 및 하위그룹 C 플라스미드 서들(SEQ ID NO: 38)로 공동-형질전환시켰다. 도 3의 다이어그램에 도시된 대로, 하위그룹 C 플라스미드 서들은 종 C에 속하는 ChAd의 클로닝에 전용인 플라스미드 벡터이므로 종 C ChAd 바이러스의 우측 및 좌측 말단(우측 및 좌측 ITR 포함)에서 유래된 DNA 단편을 포함한다.
- [0275] 선형화된 하위그룹 C 플라스미드 서들의 말단에 존재하는 우측 및 좌측 ITR DNA 서열(PshAI/NdeI/XbaI로 분해됨) 및 바이러스 유전체 DNA 사이의 상동 재조합은 플라스미드 벡터에서의 이의 삽입을 허용하였다. 30개의 상이한 클론을 증폭시키고 제한 분석에 의해 분석하고 9개의 상이한 종을 동정하였다. 19/30개의 클론은 동일한 제한 패턴을 보였고 지배적인 종을 나타내었다; 이들 클론 중 하나를 선택하였고 pChAd157 ΔE1 TetO hCMV RpsL-Kana#1551(SEQ ID NO: 15)로서 동정하였다.
- [0276] 1.3: ChAd157 ΔE1/TetO hCMV GAG#1557의 작제
- [0277] GAG 카세트(GAG 폴리뉴클레오티드 서열 SEQ ID NO: 16)를 HCMV 프로모터 및 BGH polyA 서열(SEQ ID NO: 39) 사이에 존재하는 상동성을 이용함에 의해 E. 콜리에서 상동 재조합을 통해 선형화된 프리-아테노 역셉터 벡터 내로 클로닝하였다.
- [0278] 플라스미드 pARS CV32TetOhCMV GAG를 SpeI 및 SphI로 절단하여 tetO, HIV-GAG 및 BGH polyA 서열을 갖는 HCMV 프로모터를 함유하는 2.44 Kb 단편을 절제하였다.
- [0279] HIV-GAG 2.44Kb 단편을 HCMV 및 BGHPA의 제어 하에 RpsL-Kana 선택 카세트를 운반하는 pChAd157 ΔE1/TetO hCMV RpsL-Kana (#1551) 역셉터 벡터(SnaBI 분해됨) 내로 상동 재조합에 의해 클로닝하였다. 생성된 작제물은 pChAd157 ΔE1/TetO hCMV GAG#1557 벡터(SEQ ID NO: 17)였다.
- [0280] ChAd157 GAG를 운반하는 플라스미드의 구조는 도 4에 보고되어 있다.
- [0281] 1.4: ChAd157 ΔE1E4 Ad5E4orf6/TetO hCMV RpsL-Kana#1594의 작제.
- [0282] ChAd157 ΔE1 벡터는 백본에 다음 변형을 갖도록 후속하여 변형되었다:
- [0283] a) ΔE1 바이러스의 E4 영역(bp 34413 내지 bp 37127)의 결실;
- [0284] b) 인간 Ad5로부터 유래된 E4orf6의 삽입.
- [0285] 클로닝 및 E. 콜리에서의 상동 재조합의 여러 단계를 수반하는 전략을 사용하여 천연 E4 영역을 Ad5 E4orf6 코딩 서열로 대체함으로써 뉴클레오티드 34413 내지 37127에 걸쳐 있는 E4 영역의 결실(ΔE1 벡터 서열 좌표)을 벡터 백본에 도입하였다. E4 코딩 영역을 완전히 결실시킨 반면, E4 천연 프로모터 및 폴리아데닐화 신호는 보존되었다. 이를 위해, 하기 상술되는 바와 같이 E. 콜리 BJ5183에서의 상동 재조합에 의해 ChAd157 천연 E4 영역을 대체함으로써 Ad5orf6의 삽입을 가능하게 하는 서들 벡터를 작제하였다.
- [0286] - pARS 중C Ad5E4orf6-1의 작제:
- [0287] 올리고뉴클레오티드 5'-ATACGGACTAGTGGAGAAGTACTCGCCTACATG-3' (SEQ ID NO: 18) 및 5'-ATACGGAAGATCTAAGACTTCAGGAAATATGACTAC-3' (SEQ ID NO: 19)와 함께 주형으로서 Ad5 DNA를 이용한 PCR에 의해 Ad5orf6을 함유하는 DNA 단편을 획득하였다. PCR 단편을 BglIII 및 SpeI으로 분해하고 BglIII 및 SpeI으로 분해된 pARS 중 C RLD-EGFP 서들에 클로닝하여, 플라스미드 pARS 중 C Ad5orf6-1을 생성하였다.
- [0288] - pARS 중C Ad5E4orf6-2의 작제:
- [0289] 올리고뉴클레오티드 5'- ATTCAGTGTACAGGCGCGCCAAAGCATGACACTGATGTTTCATTTC-3' (SEQ ID NO: 20) 및 5'-ACTAGGACTAGTTATAAGCTAGAATGGGGCTTTGC-3' (SEQ ID NO: 21)와 함께 주형으로서 플라스미드 pChAd157 ΔE1/TetO hCMV RpsL-Kana(#1551)를 이용한 PCR에 의해 섬유-E4 polyA(ChAd157 ΔE1 벡터의 bp 34269 내지 bp 34412)를 함유하는 144 bp DNA 단편을 증폭시켰다. PCR 단편을 BsrGI 및 SpeI으로 분해하고, BsrGI 및 SpeI으로 분해된

pARS 하위그룹C Ad5orf6-1으로 클로닝하여, 플라스미드 pARS 중C Ad5orf6-2(SEQ ID NO: 40)를 생성하였다.

[0290] 이후 생성된 플라스미드 pARS 중C Ad5orf6-2를 사용하여 ChAd157 백본 내의 Ad5orf6으로 E4를 대체시켰다. 이를 위해, 플라스미드 pChAd157 ΔE1 TetO hCMV RpsL-Kana#1551를 PacI로 분해하고 분해된 플라스미드 pARS 중C Ad5orf6-2 BamHI/AscI와 함께 BJ5183 세포 내로 공동-형질전환시켜, pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana(#1594) 프리아데노 플라스미드(SEQ ID NO: 22)를 수득하였다.

[0291] 1.5: ChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RG#1559의 작제.

[0292] 광견병 바이러스 당단백질(RG) 발현 카세트(광견병 당단백질 폴리뉴클레오티드 서열 SEQ ID NO: 23)를 HCMV 프로모터 및 BGH polyA 서열 사이에 존재하는 상동성을 이용함에 의해 E. 콜리에서 상동 재조합을 통해 선형화된 프리-아데노 엑셉터 벡터 내로 클로닝하였다.

[0293] 플라스미드 pvjTetOhCMV-bghpolyA_RG를 SpeI 및 AsiSI로 절단하여 tetO, RG 및 BGHPolyA 서열을 갖는 HCMV 프로모터를 함유하는 2.59 Kb 단편을 절제하였다.

[0294] 생성된 RG 2.59 Kb 단편을 상동 재조합에 의해 HCMV 및 BGHPA의 제어 하에 RpsL-Kana 선택 카세트를 운반하는 pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana(#1594) 엑셉터 벡터 내로 클로닝하였다. 엑셉터 preAd 플라스미드를 제한 엔도뉴클레아제 SnaBI으로 선형화시켰다. 생성된 작제물은 pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RG#1559 벡터(SEQ ID NO: 24)였다.

[0295] ChAd157 RG를 운반하는 플라스미드의 구조는 도 6에 보고되어 있다.

[0296] 실시예 2: 벡터 생산

[0297] ChAd157의 생산성을 Procell 92 세포주에서 ChAd19 및 ChAd155와 비교하여 평가하였다.

[0298] 2.1: HIV Gag 트랜스진을 포함하는 벡터의 생산

[0299] ChAd157/GAG, ChAd19/GAG, ChAd155/GAG(HIV Gag 트랜스진을 발현시키는 ChAd157, ChAd19 및 ChAd155 벡터)를 구제하고 Procell 92에서 증폭시켰다; 용해물을 이용하여 각 벡터와 함께 단층으로 배양된 Procell 92의 1개 T25 플라스크를 감염시켰다. 300 vp/세포의 감염 다중도(MOI)가 사용되었고 ChAd19/GAG가 hCMV 프로모터에서 TetO 조작자의 삽입에 의해 매개되는 전사 조절이 결핍되었기 때문에 감염은 테트라사이클린의 존재 하에 수행되었다. 충분한 세포병변 효과가 명백할 때(ChAd157/GAG 및 ChAd155/GAG의 경우 감염 48시간 후 및 ChAd19/GAG의 경우 감염 5일 후) 감염된 세포를 수확하였다; 3회 사이클의 동결/해동(-70℃ 내지 37℃)에 의해 감염된 세포로부터 바이러스를 방출한 다음 용해물을 원심분리에 의해 정화시켰다. 정화된 용해물을 CMV 프로모터 영역에 상보적인 프라이머 및 프로브를 이용한 정량적 PCR 분석에 의해 정량하였다. 올리고뉴클레오티드 서열은 다음과 같다: CMVfor 5'-CATCTACGTATTAGTCATCGTATTACCA-3'(SEQ ID NO: 25), CMVrev 5'-GACTTGGAAATCCCCGTGAGT-3'(SEQ ID NO: 26), CMVFAM-TAMRA 프로브 5'-ACATCAATGGGCGTGGATAGCGGTT-3'(SEQ ID NO: 41)(QPCR은 ABI Prism 7900 서열 검출기 상에서 진행되었다 - Applied Biosystem).

[0300] 정화된 용해물에서 측정된 결과적 체적 역가(volumetric titer)(vp/ml) 및 세포 당 바이러스 입자에서 발현된 특이적 생산성(vp/세포)은 하기 표 1에 제공된다.

[0301] 표 1: GAG 벡터 생산성.

벡터	체적 생산성 (vp/ml)	총 vp	세포 특이적 생산성 (vp/세포)
ChAd157/GAG	4.61E+09	2.30E+10	7.68E+03
ChAd155/GAG	5.42E+09	2.71E+10	9.04E+03
ChAd19/GAG	4.80E+08	2.40E+09	8.00E+02

[0302]

[0303] 2.2: RG 트랜스진을 포함하는 벡터의 생산

[0304] 현탁액에서 배양된 Procell 92에서의 RG 백신 벡터의 생산성을 평가하기 위해 다양한 세트의 실험을 수행하였다. 실험은 5x10⁵ 개 세포/ml의 세포 밀도로 Procell 92를 감염시킴으로써 ChAd157/RG 및 ChAd155/RG를

동시에 비교하였다. 300 vp/세포의 다중 감염도(MOI)가 사용되었다. 감염된 세포를 감염 4일 후에 수확하였고; 3회 사이클의 동결/해동에 의해 감염된 세포로부터 바이러스를 방출시키고, 용해물을 원심분리에 의해 정화시켰다. 이후 정화된 용해물을 상기 보고된 대로 QPCR에 의해 정량하였다.

체적 생산성 및 세포 특이적 생산성은 하기 표 2에 제공되어 있다.

표 2: RG 벡터 생산성.

벡터	체적 생산성 (vp/ml)	총 vp	세포 특이적 생산성 (vp/세포)
ChAd157/RG	9.39E+09	4.69E+11	1.88E+04
ChAd155/RG	1.41E+10	7.04E+11	2.81E+04

실시예 3: 트랜스진 발현 수준

3.1: HIV Gag 트랜스진의 발현 수준

HeLa 세포를 HIV Gag 트랜스진을 포함하는 ChAd19, ChAd155 및 ChAd157 벡터로 감염시킴으로써 발현 수준을 동시 실험으로 비교하였다.

HeLa 세포를 35 mm 디쉬에 시딩하고 MOI = 250 vp/세포를 사용하여 ChAd19/GAG, ChAd157/GAG 및 ChAd155/GAG 정제된 바이러스로 감염시켰다. 감염된 HeLa 세포의 상청액을 감염 48시간 후에 수확하고, 상업적 ELISA 키트 (HIV-1 p24 ELISA Kit, PerkinElmer Life Science)를 이용하여 분비된 HIV GAG 단백질의 생산을 정량하였다. HIV-1 p24 항원 표준 곡선을 이용하여 제조업체의 설명서에 따라 정량을 수행하였다.

pg/ml의 GAG 단백질로 표현되는 결과가 도 7에 예시되어 있다.

3.2: RG 트랜스진의 발현 수준

ChAd155/RG 벡터와 비교하여 ChAd157/RG 벡터에 의해 제공된 광견병 당단백질 발현을 평가하기 위해 웨스턴 블롯 분석을 또한 수행하였다. 이를 위해, HeLa 세포를 35 mm 디쉬에 시딩하고 MOI = 250 vp/세포를 사용하여 ChAd157/RG 및 ChAd155/RG 정제된 바이러스로 감염시켰다. 세포 용해물을 감염 48시간 후에 수확하고 트랜스진 발현 수준을 SDS-PAGE 환원시킨 후 웨스턴 블롯 분석으로 평가하였다.

동량의 단백질 추출물을 환원성 SDS 겔에 부하시켰다; 전기영동 분리 후, 단백질을 니트로셀룰로스 막으로 옮겨 토끼 폴리클로날 항-GP(Cat. No. RBVGP11-S αDiagnostic, diluted 1:1000)로 프로빙하였다. 일차 항체와 함께 인큐베이션한 후, 막을 세척한 다음 항-토끼 호스래디쉬 퍼옥시다제(HRP) 컨쥬게이트 이차 항체와 함께 인큐베이션하였다. 마지막으로, 검정은 향상된 화학발광(ECL) 검출 시약(W3252282 PIERCE)을 사용한 화학발광에 의해 현상되었다. 웨스턴 블롯 결과가 도 8에 제시된다.

화살표로 표시된 약 57 kD의 밴드는 광견병 당단백질의 예상 중량에 상응하는 폴리클로날 항체 항-GP에 의해 나타났다.

상기 결과는 ChAd157 벡터의 발현 수준이 ChAd155에 의해 제공되는 것과 유사하게 나타난다는 것을 입증하였다.

실시예 4: 마우스 면역화 실험에 의한 면역학적 효능의 평가

4.1: HIV Gag 트랜스진을 포함하는 벡터의 면역원성

면역원성 ChAd157/GAG 벡터를 BALB/c 마우스(그룹 당 6마리)에서 ChAd155/GAG 및 ChAd19/GAG와 동시에 평가하였다. 10⁷ 개 바이러스 입자를 근내 주입하여 실험을 수행하였다. T-세포 반응을 BALB/c 마우스에서 맵핑된 GAG CD8+ T 세포 에피토프를 이용한 생체 외 인터페론-γ (IFN-γ) 효소-결합 면역스폿(ELISpot)에 의해 면역화 3주 후에 측정하였다. 수득된 결과는 도 9에 보고되어 있고, 100만 개의 비장세포 당 IFNγ 스폿 형성 세포(SFC)로 표현되었다.

각각의 도트는 단일 마우스에서의 반응을 나타내고, 선은 각각의 용량 그룹에 대한 기하평균에 해당한다. CD8 면역우세 펩티드에 대한 양성 마우스의 빈도는 x 축에 도시되어 있다.

[0322] 4.2 RG 트랜스진을 포함하는 벡터의 면역원성

[0323] ChAd157/RG 및 ChAd155/RG 벡터의 면역학적 효능을 BALB/c 마우스에서 평가하였다. 둘 모두의 벡터는 10^7 및 10^6 vp 용량으로 근내 주입되었다. 면역화된 마우스의 비장세포를 백신화 7주 후에 단리시키고 RG로부터의 펩티드 풀을 항원으로 사용하여 IFN γ ELISpot(도 10)에 의해 분석하였다.

[0324] 면역 반응의 수준은 예상대로 투여량 감소에 따라 감소하였다. 또한, ChAd155RG 벡터는 ChAd157 RG보다 높은 T 세포 반응을 유도하였지만, 유의한 차이는 없었다(도 10).

[0325] 실시예 5: 감염성의 평가

[0326] 5.1: HIV Gag 트랜스진을 포함하는 벡터의 감염성

[0327] 정제된 바이러스의 감염성은 면역세포화학 염색에 의해 감염된 세포를 가시화하기 위해 아데노바이러스 핵소 단백질에 대한 항체를 이용하는 부착성 Procell 92 세포에서 평가되었다. 핵소 단백질에 대한 항체는 아데노바이러스의 모든 혈청형을 인지한다. 이를 위해, Procell92 세포를 2×10^5 개 생육성 세포/ml의 세포 밀도로 24웰 플레이트에 시딩하고 MOI= 1 vp/세포, 0.5 vp/세포 및 0.25 vp/세포를 사용하여 ChAd157/GAG 및 ChAd155/GAG 및 ChAd19/GAG 벡터로 이중으로 감염시켰다. 감염 48시간 후에, 감염된 세포를 냉 메탄올에 의해 고정시킨 다음 항-핵소 항체로 표지하였다. 과량의 항체가 제거된다. 이후 표지된 세포를 호스레디쉬 퍼옥시다제와 퀴뉴게이션된 이차 항체와 함께 인큐베이션하고 상업적 키트 VECTOR NOVARED Substrate Kit(SK-4800)를 사용하여 검출을 수행하였다. 호스레디쉬 퍼옥시다제 효소 표지가 DAB 기질과 반응하여 암갈색 생성물이 생성될 때 검출이 이루어진다. 이후 표지된 암갈색 세포를 광학 현미경으로 정량하고 감염 역가를 계산하였다. 결과는 하기 표에 제시된다.

바이러스	Vp/ml	Ifu/ml	R (vp/ifu)
ChAd155 GAG	1.32E+11	1.58E+09	84
ChAd157 GAG	1.17E+11	1.23E+09	95
ChAd19 GAG	4.46E+10	3.86E+08	116

[0328]

[0329] 상기 결과는 ChAd155 및 ChAd157 바이러스의 감염성이 유사하고 ChAd19보다 높음을 입증하였다.

[0330] 5.2 RG 트랜스진을 포함하는 벡터의 감염성

[0331] ChAd157/RG 및 ChAd155/RG 정제된 바이러스의 감염성을 상기 보고된 대로 Hexon Immunostaining에 의해 부착성 Procell 92 세포에서 평가하였다. 결과는 하기 표에 제시된다.

바이러스	Vp/ml	Ifu/ml	R (vp/ifu)
ChAd155/RG	4.23E+11	4.06E+09	104
ChAd157/RG	1.97E+11	1.46E+09	133

[0332]

[0333] 상기 결과는 ChAd155 및 ChAd157 바이러스의 감염성이 유사함을 입증하였다.

[0334] 실시예 6: ChAd155 및 ChAd157 벡터 사이의 교차-중화의 평가

[0335] 6.1 ChAd155 및 ChAd157 벡터가 상이한 혈청형인 경우 생체 내 시험

[0336] ChAd155 및 ChAd157 벡터 사이의 교차-중화는 BALB/c 마우스(그룹 당 6마리)에서 평가되었다. 마우스를 0주 및 3주에 RG를 발현하는 10^9 vp의 ChAd155 또는 ChAd157로 사전면역화하거나 염수 완충제로 모의-백신화하였다. 3주 후, 이어서 모든 마우스를 HIV gag를 인코딩하는 10^9 Vp의 ChAd157로 1회 면역화하였다.

그룹	n	사전-면역화 2x w0 및 w3	용량 (vp)	면역화 w6	용량(vp)
1	6	PBS	-	ChAd157-GAG	10^9
2	6	ChAd155-RG	10^9	ChAd157-GAG	10^9
3	6	ChAd157-RG	10^9	ChAd157-GAG	10^9

[0337]

[0338]

사전면역화 벡터에 대한 중화 역가는 시험관 내 중화 검정에 의해 5주차(이차 주입 후 2주)에 혈청에서 측정되었다. 마지막으로, gag에 대한 T 세포 반응을 BALB/c 마우스에서 맵핑된 GAG CD8+ T 세포 에피토프를 사용하여, IFN- γ ELISpot에 의한 면역화 3주 후에 비장세포에서 검사하였다(도 12). 사전면역화에 사용된 벡터의 용량은 일부 가변성을 갖지만 두 Ad 벡터에 대해 우수한 중화 활성을 유도할 수 있었다. 항-ChAd155 중화 항체는 ChAd157에 대해 교차 반응하지 않으며 반대의 경우도 마찬가지이다(도 11). 또한, ChAd157-Gag T-세포 반응은 항-ChAd155 사전면역에 의해 영향을 받지 않았으며, 교차-중화가 관찰되지 않았음을 확인하였다(도 12).

[0339]

종합적으로, 이들 데이터는 ChAd155 및 ChAd157 바이러스가 별개의 아데노바이러스 혈청형이라는 것을 제안한다.

도면

도면1a

	10	20	30	40	50	60	70	80	
12_ChAd157	---	+	---	+	---	+	---	+	
01_ChAd3	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
02_PanAd3	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
03_ChAd17	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
04_ChAd19	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
05_ChAd24	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
06_ChAd155	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
07_ChAd11	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
08_ChAd20	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
09_ChAd31	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
10_PanAd1	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
11_PanAd2	MKRKTSDES	FNVPYDTE	GPPSVFELT	PFVSPDGFQES	PPGVLSLNLAEP	LTVTSGMLAL	KMGSGLSLDDAGNLT	80	
---	---	+	---	+	---	+	---	+	
12_ChAd157	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
01_ChAd3	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
02_PanAd3	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	159
03_ChAd17	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
04_ChAd19	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
05_ChAd24	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
06_ChAd155	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
07_ChAd11	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
08_ChAd20	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
09_ChAd31	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	160
10_PanAd1	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	159
11_PanAd2	QDITTA	PELKKTKTNLS	LETS	PLTVSTSGALT	VAAAPLAVAGTSLT	MQSEAPLT	VQDAKLTATK	GPLTVSEKGLAL	159

도면1b

	170	180	190	200	210	220	230	240																
12_ChAd157	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
01_ChAd3	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
02_PanAd3	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	239											
03_ChAd17	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
04_ChAd19	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
05_ChAd24	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
06_ChAd155	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
07_ChAd11	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
08_ChAd20	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
09_ChAd31	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	240											
10_PanAd1	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	239											
11_PanAd2	QTSAPLTAADSS	TLTVSATPP	PS	SSGSLG	DMEDP	NYT	NGKLG	HIGG	PLRWDSL	TLTVVTG	GLT	NNALQTRV	239											
	250	260	270	280	290	300	310	320																
12_ChAd157	TGALGYDT	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
01_ChAd3	TGALGYDT	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
02_PanAd3	TGALGYD	SGNLQ	LRAAGG	MR	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	319
03_ChAd17	TGALGYDT	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
04_ChAd19	TGALGYDT	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
05_ChAd24	TGALGYDT	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
06_ChAd155	TGALGYDT	SGNLQ	LRAAGG	MR	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
07_ChAd11	TGALGYD	SGNLQ	LRAAGG	MR	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
08_ChAd20	TGALGYDT	SGNLQ	LRAAGG	MR	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
09_ChAd31	TGALGYDT	SGNLQ	LRAAGG	MR	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	320
10_PanAd1	TGALGYD	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	319
11_PanAd2	TGALGYD	SGNLQ	LRAAGG	MRID	ANGQ	LIL	VA	YFFD	AQN	NLS	SLR	LGG	PPLY	NT	SH	NLD	NC	NR	GL	TT	NT	KK	LE	319

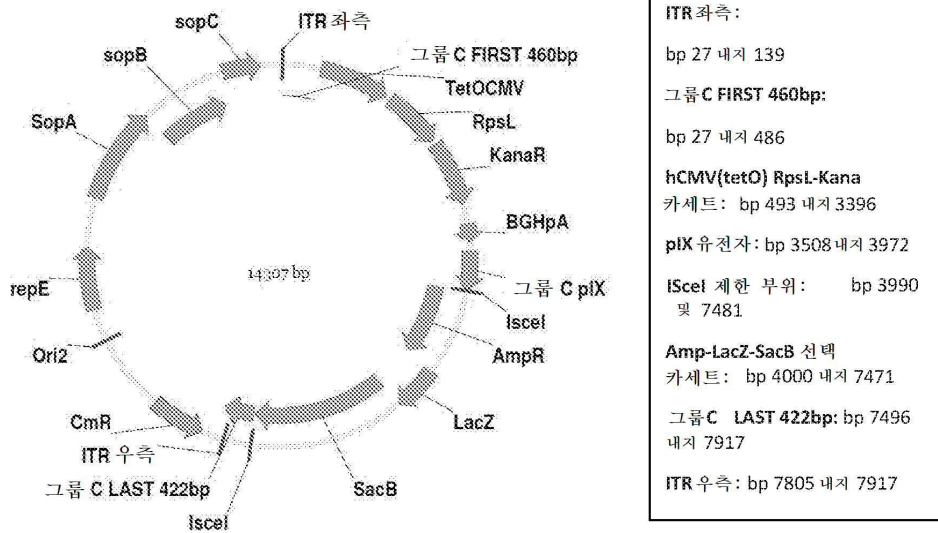
도면1c

12_ChAd157	330	340	350	360	370	380	390	400	364						
01_ChAd3				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	364						
02_PanAd3				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	363						
03_ChAd17				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	364						
04_ChAd19				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	364						
05_ChAd24				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	364						
06_ChAd155				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	400						
07_ChAd11				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	400						
08_ChAd20				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	400						
09_ChAd31				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	400						
10_PanAd1				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	399						
11_PanAd2				KISSGLDYDNGAII	KLG	TGLSFDNTGAI	TVGN	DDDKLTLM	399						
12_ChAd157	410	420	430	440	450	460	470	480							
01_ChAd3				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	444
02_PanAd3				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	444
03_ChAd17				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	442
04_ChAd19				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	444
05_ChAd24				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	444
06_ChAd155				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	480
07_ChAd11				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	480
08_ChAd20				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	480
09_ChAd31				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	479
10_PanAd1				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	480
11_PanAd2				TPDPSENCRIHS	KDK	KFTLVLTCKGSO	LASV	AVLVGGLASII	GTVA	SVIIFLRFD	NGVLI	ENSSLD	QYWN	PRNG	479

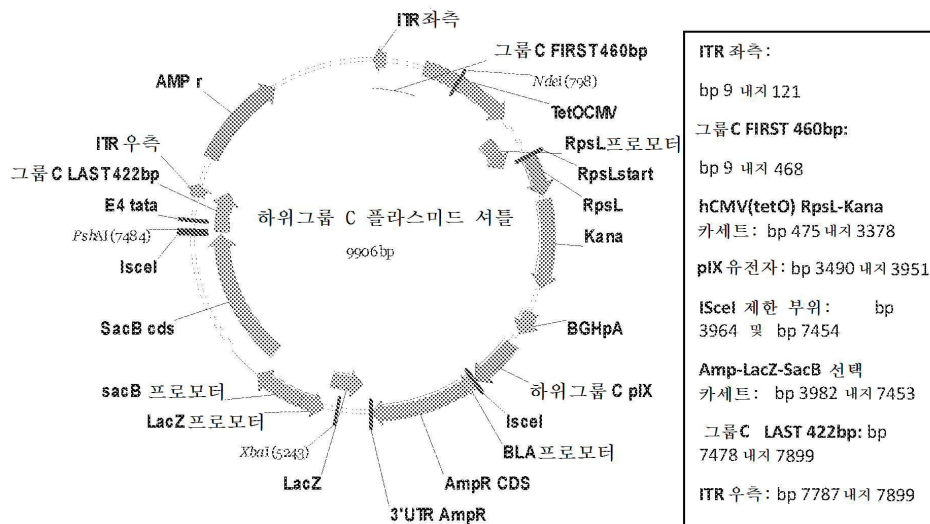
도면1d

		490	500	510	520	530	540	550	560	
12_ChAd157	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	524
01_ChAd3	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	524
02_PanAd3	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	522
03_ChAd17	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	524
04_ChAd19	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	524
05_ChAd24	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	524
06_ChAd155	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	559
07_ChAd11	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	559
08_ChAd20	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	559
09_ChAd31	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	559
10_PanAd1	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	558
11_PanAd2	STSTAFYTNVAVGEMPNI	AYPKTQSQ	TAKNIVSQVY	LN	GDKSKPMT	LTIT	LTNGTNE	STSTSQVSY	YSMSFTW	558
		570								
12_ChAd157	YAEETFTNSTFTFSYIA									543
01_ChAd3	YAEETFTNSTFTFSYIA									543
02_PanAd3	YAEETFTNSTFTFSYIA									541
03_ChAd17	YAEETFTNSTFTFSYIA									543
04_ChAd19	YAEETFTNSTFTFSYIA									543
05_ChAd24	YAEETFTNSTFTFSYIA									543
06_ChAd155	YAEETFTNSTFTFSYIA									578
07_ChAd11	YAEETFTNSTFTFSYIA									578
08_ChAd20	YAEETFTNSTFTFSYIA									578
09_ChAd31	YAEETFTNSTFTFSYIA									578
10_PanAd1	YAEETFTNSTFTFSYIA									577
11_PanAd2	YAEETFTNSTFTFSYIA									577

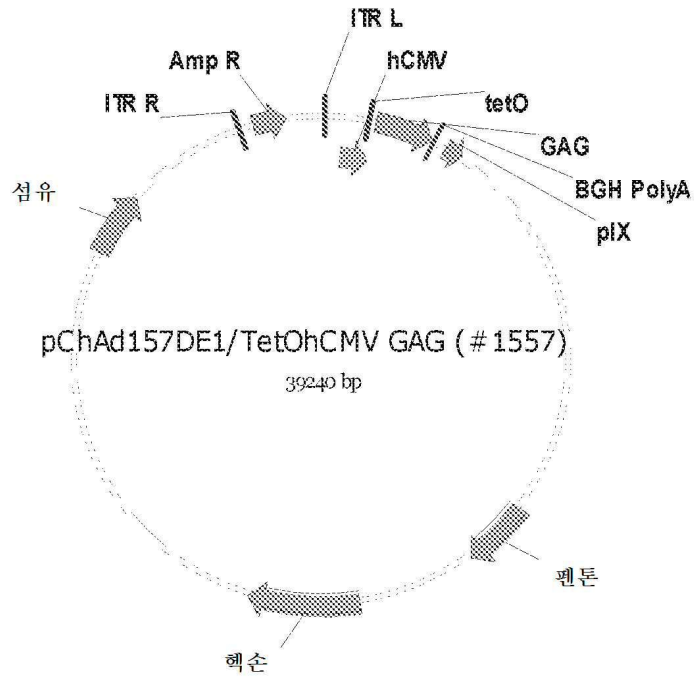
도면2



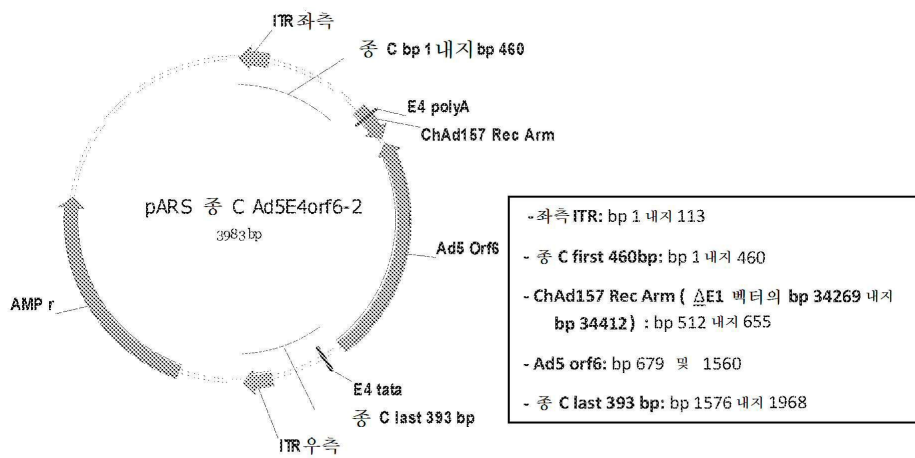
도면3



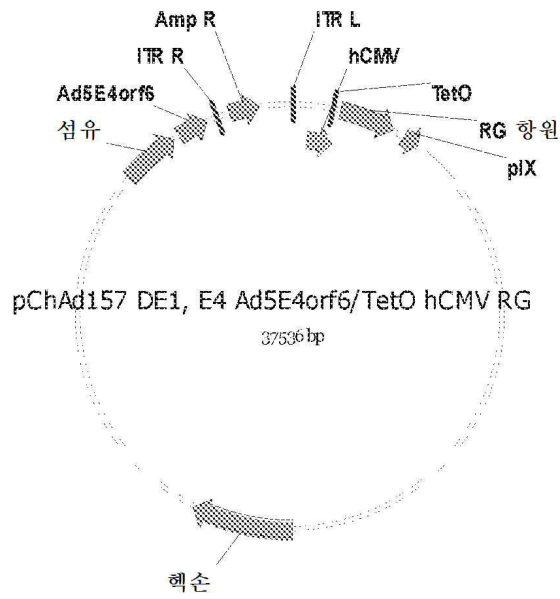
도면4



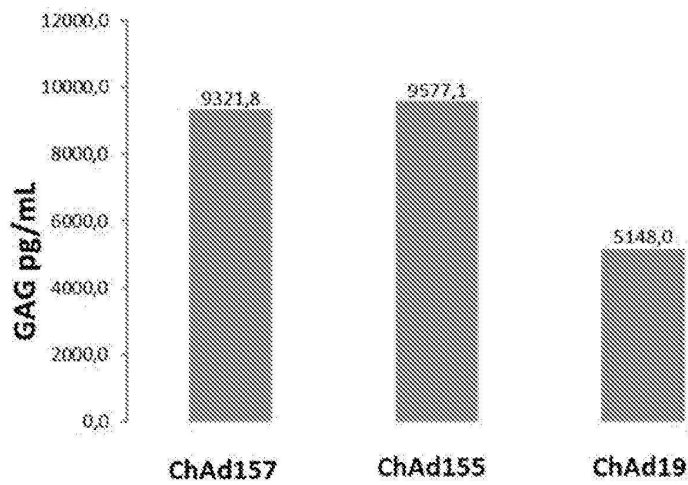
도면5



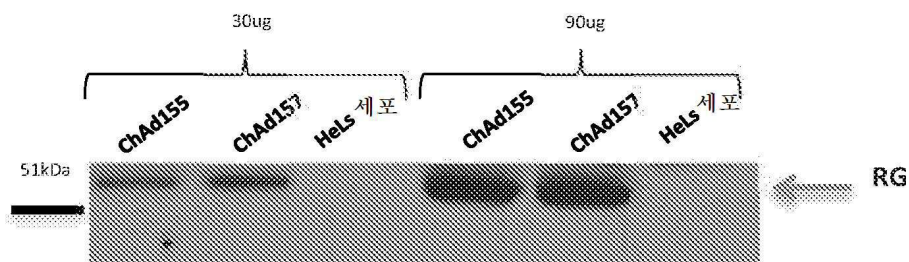
도면6



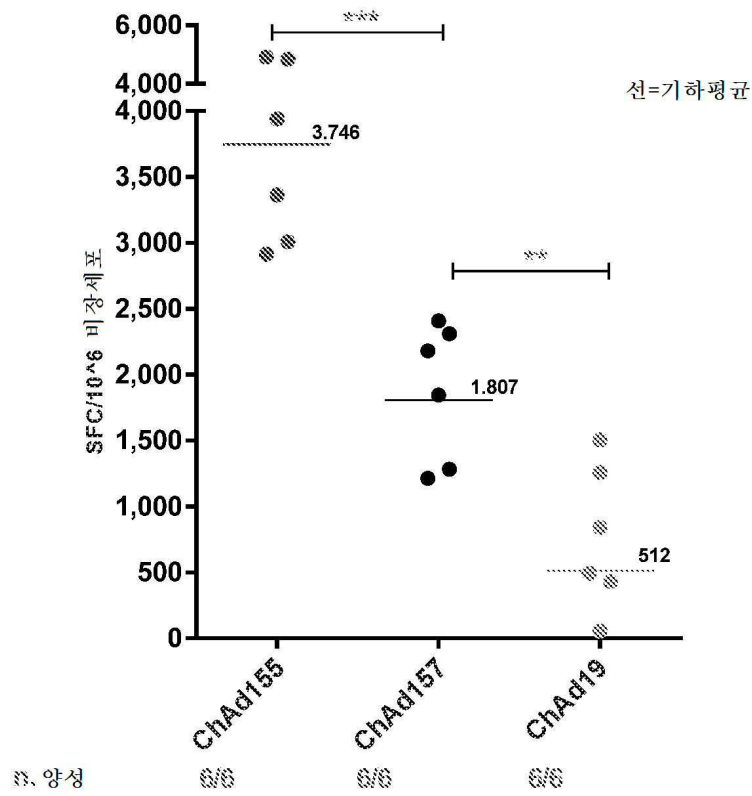
도면7



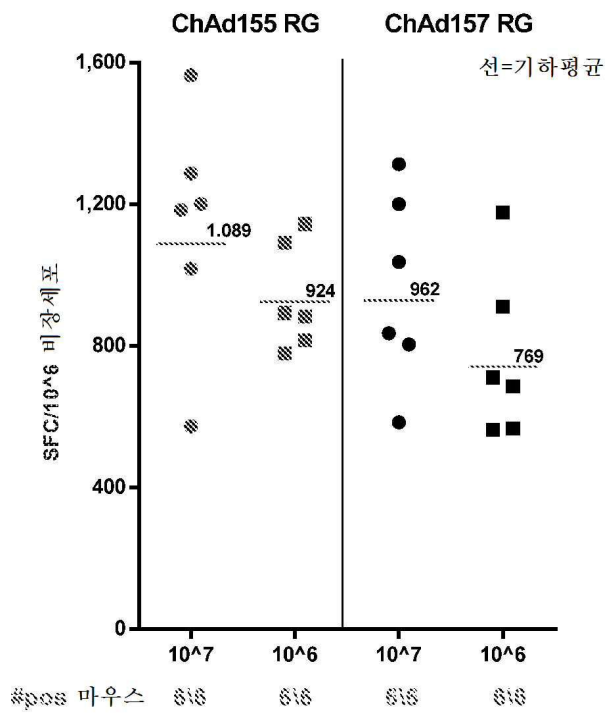
도면8



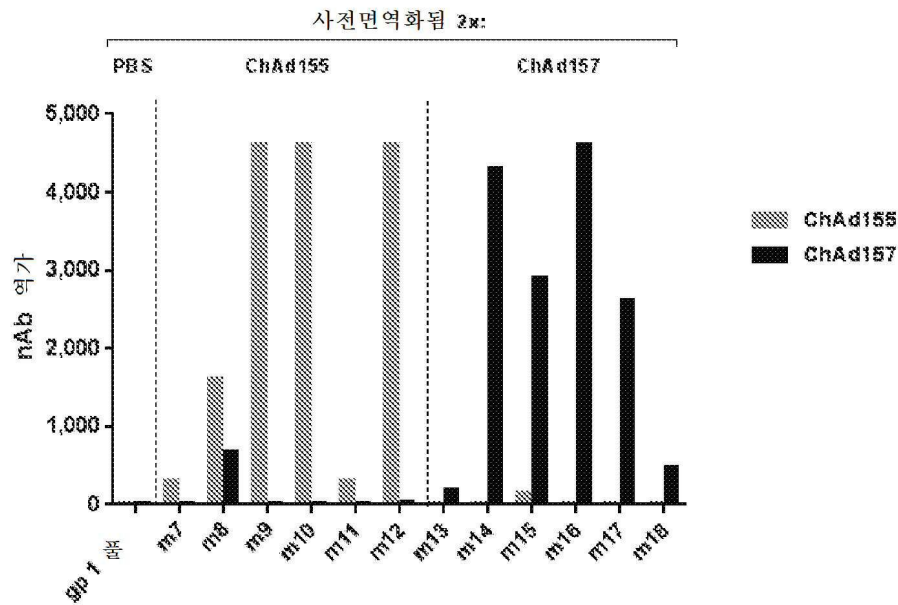
도면9



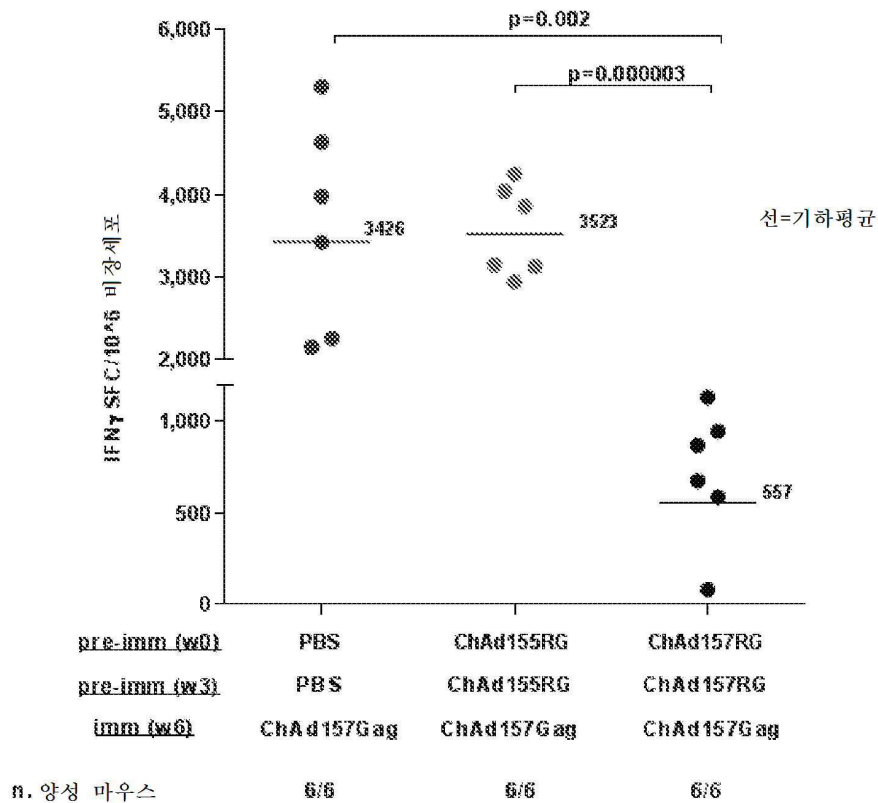
도면10



도면11



도면12



서열 목록

SEQUENCE LISTING

<110> GlaxoSmithKline Biologicals S.A.

<120> ADENOVIRUS POLYNUCLEOTIDES AND POLYPEPTIDES

<130> VU66143

<150> GB1620968.6

<151> 2016-12-09

<160> 42

<170> PatentIn version 3.5

<210> 1

<211> 543

<212> PRT

<213> Pan troglodytes

<400> 1

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu

50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser

65 70 75 80

Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn

85 90 95

Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala

100 105 110

Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu

115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr

130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu

145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175
Ser Ala Thr Pro Pro Ile Asn Val Ser Ser Gly Ser Leu Gly Leu Asp
180 185 190
Met Glu Asp Pro Met Tyr Thr His Asn Gly Lys Leu Gly Ile Arg Ile
195 200 205
Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
210 215 220
Thr Gly Asn Gly Leu Thr Val Asp Asn Asn Ala Leu Gln Thr Lys Val
225 230 235 240
Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
245 250 255
Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asn Val
260 265 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285
Gly Pro Leu Tyr Ile Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
290 295 300
Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305 310 315 320
Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335
Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
340 345 350
Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365
Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380
Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400
Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val

405 410 415
 Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
 420 425 430
 Ser Leu Asp Arg Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
 435 440 445
 Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
 450 455 460
 Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln

465 470 475 480
 Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
 485 490 495
 Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
 500 505 510
 Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
 515 520 525
 Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln

530 535 540

<210> 2

<211> 1629

<212> DNA

<213> Pan troglodytes

<400> 2

atgaagcgca caaaacgtc tgacgagagc ttcaaccccg tgtacccta tgacacggaa	60
agcggccctc cctccgtccc ttctctcacc cctcccttcg tgtctcccga tggattccaa	120
gaaagtcccc ccggggctct gtctctgaac ctggccgagc ccttggtcac ttcccacggc	180
atgctcgccc tgaaaatggg aagtggcctc tcctggacg acgctggcaa cctcacctct	240
caagatatca ccaccgctag cctccctc aaaaaacca agaccaacct cagcctagaa	300
acctcatccc ccctaactgt gagcacctca ggccgacctca ccgtagcagc cgccgctccc	360
ctggcgggtgg ccggcacctc cctcacatg caatcagagg cccccctgac agtacaggat	420
gcaaaactca ccttggtccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg	480
caaacatcgg ccccgctgac ggccgctgac agcagcacc tcaccgttag cgccacacca	540
ccaattaatg taagcagtgg aagtttaggc ttagacatgg aagaccctat gtatactcac	600

aatggaaaac tgggaataag aattgggggt ccactaagag tagtagacag cttgcataca	660
ctcactgtag ttaccggaaa tggactaact gtagataaca atgccctcca aactaaagtt	720
acgggcgccc taggttatga cacatcagga aatctacaat taagagctgc aggaggtatg	780
cgaattgacg caaatggcca acttatcctt aatgtggcat acccatttga tgctcagaac	840
aatctcagcc ttagacttgg tcagggaccc ctgtatataa acacagacca caacctggat	900
ttgaattgca acagaggctt aaccacaact accaccaaca acacaaaaaa acttgagact	960
aaaattagct caggcttaga ctatgacacc aatggtgctg tcattattaa acttggcact	1020
ggtctaagct tcgacaacac aggcgcctta actgtgggaa acactggtga tgataaactg	1080
actctgtgga cgaccccgaga cccatctcca aattgcagaa ttactcaga caaagactgc	1140
aagtttactc tagtctaac taagtgtgga agccaaatcc tggcctctgt cgccgcctta	1200
gcggtatcag gaaatctggc ttcgataaca ggcaccgttg ccagcgttac catctttctc	1260
agatttgatc agaatggagt gcttatggaa aactcctcgc tagacaggca gtactggaac	1320
ttcagaaatg gcaactcaac taacgtgcc cctacacca atgcagttgg gttcatgcca	1380
aacctcgcag catacccaca aacgcaaagc cagactgcta aaaacaacat tgtaagtcag	1440
gtttacttga atggagacaa atccaaaccc atgaccctta ccatcaccct caatggaact	1500
aatgaatcca gtgaaactag ccaggtgagt cactactcca tgtcatttac atgggcttgg	1560
gaaagtgggc aatatgccac tgaaaccttt gccaccaact ccttcacctt ttcttacatt	1620
gctgaacaa	1629
<210> 3	
<211> 593	
<212> PRT	
<213> Pan troglodytes	
<400> 3	
Met Arg Arg Ala Ala Met Tyr Gln Glu Gly Pro Pro Pro Ser Tyr Glu	
1 5 10 15	
Ser Val Val Gly Ala Ala Ala Ala Ala Pro Ser Ser Pro Phe Ala Ser	
20 25 30	
Gln Leu Leu Glu Pro Pro Tyr Val Pro Pro Arg Tyr Leu Arg Pro Thr	
35 40 45	
Gly Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Phe Asp	
50 55 60	

Thr Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Val Ala Ser
 65 70 75 80
 Leu Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Ile Gln
 85 90 95
 Asn Asn Asp Tyr Ser Pro Ser Glu Ala Ser Thr Gln Thr Ile Asn Leu
 100 105 110
 Asp Asp Arg Ser His Trp Gly Gly Asp Leu Lys Thr Ile Leu His Thr
 115 120 125
 Asn Met Pro Asn Val Asn Glu Phe Met Phe Thr Asn Lys Phe Lys Ala
 130 135 140
 Arg Val Met Val Ser Arg Ser His Thr Lys Glu Asp Arg Val Glu Leu
 145 150 155 160
 Lys Tyr Glu Trp Val Glu Phe Glu Leu Pro Glu Gly Asn Tyr Ser Glu
 165 170 175
 Thr Met Thr Ile Asp Leu Met Asn Asn Ala Ile Val Glu His Tyr Leu
 180 185 190
 Lys Val Gly Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys
 195 200 205
 Phe Asp Thr Arg Asn Phe Arg Leu Gly Leu Asp Pro Val Thr Gly Leu
 210 215 220
 Val Met Pro Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Ile
 225 230 235 240
 Leu Leu Pro Gly Cys Gly Val Asp Phe Thr Tyr Ser Arg Leu Ser Asn
 245 250 255
 Leu Leu Gly Ile Arg Lys Arg Gln Pro Phe Gln Glu Gly Phe Arg Ile
 260 265 270
 Thr Tyr Glu Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val
 275 280 285
 Glu Ala Tyr Gln Asp Ser Leu Lys Glu Asn Glu Ala Gly Gln Glu Asp
 290 295 300
 Thr Ala Pro Ala Ala Ser Ala Ala Ala Glu Gln Gly Glu Asp Ala Ala

305 310 315 320
 Asp Thr Ala Ala Ala Asp Gly Ala Glu Ala Asp Pro Ala Met Val Val
 325 330 335
 Glu Ala Ala Glu Gln Glu Glu Asp Met Asn Asp Ser Ala Val Arg Gly
 340 345 350
 Asp Thr Phe Val Thr Arg Gly Glu Glu Lys Gln Ala Glu Ala Glu Ala

 355 360 365
 Ala Ala Glu Glu Lys Gln Leu Ala Ala Ala Ala Ala Ala Ala Leu
 370 375 380
 Ala Ala Ala Glu Ala Glu Ser Glu Gly Thr Lys Pro Ala Lys Glu Pro
 385 390 395 400
 Val Ile Lys Pro Leu Thr Glu Asp Ser Lys Lys Arg Ser Tyr Asn Leu
 405 410 415
 Leu Lys Asp Ser Thr Asn Thr Ala Tyr Arg Ser Trp Tyr Leu Ala Tyr

 420 425 430
 Asn Tyr Gly Asp Pro Ser Thr Gly Val Arg Ser Trp Thr Leu Leu Cys
 435 440 445
 Thr Pro Asp Val Thr Cys Gly Ser Glu Gln Val Tyr Trp Ser Leu Pro
 450 455 460
 Asp Met Met Gln Asp Pro Val Thr Phe Arg Ser Thr Arg Gln Val Ser
 465 470 475 480
 Asn Phe Pro Val Val Gly Ala Glu Leu Leu Pro Val His Ser Lys Ser

 485 490 495
 Phe Tyr Asn Asp Gln Ala Val Tyr Ser Gln Leu Ile Arg Gln Phe Thr
 500 505 510
 Ser Leu Thr His Val Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Ala
 515 520 525
 Arg Pro Pro Ala Pro Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala
 530 535 540
 Leu Thr Asp His Gly Thr Leu Pro Leu Arg Asn Ser Ile Gly Gly Val

 545 550 555 560

Gln Arg Val Thr Val Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val
 565 570 575
 Tyr Lys Ala Leu Gly Ile Val Ser Pro Arg Val Leu Ser Ser Arg Thr
 580 585 590
 Phe

<210> 4

<211> 1779

<212> DNA

<213> Pan troglodytes

<400> 4

atgcggcgcg cgcgcatgta ccaggaggga cctcctccct cttacgagag cgtggtgggc	60
gcggcgggcg cgcgccctc ttctcccttt gcgtcgagc tgctggagcc gccgtacgtg	120
cctccgcgct acctgcggcc tacggggggg agaaacagca tccgttactc ggagctggcg	180
ccccgtttcg acaccacccg ggtgtacctg gtggacaaca agtcggcgga cgtggcctcc	240
ctgaactacc agaacgacca cagcaatttt ttgaccacgg tcatccagaa caatgactac	300
agcccgagcg agcccgacac ccagaccatc aatctggatg accggtcgca ctggggcggc	360
gacctgaaaa ccatcctgca caccaacatg cccaacgtga acgagttcat gttcaccaat	420
aagttcaagg cgcggtgat ggtgtcgcg tcgcacacca aggaagaccg ggtggagctg	480
aagtacgagt ggtggagtt cgagctgcca gagggcaact actccgagac catgaccatt	540
gacctgatga acaacgcgat cgtggagcac tatctgaaag tgggcaggca gaacggggtc	600
ctggagagcg acatcggggt caagttcgac accaggaact tccgcctggg gctggacccc	660
gtgaccgggc tggttatgcc cggggtgtac accaacgagg ctttccatcc cgacatcatc	720
ctgctgcccc gctcggggtt ggacttcact tacagccgcc tgagcaacct cctgggcac	780
cgcaagggc agcccttcca ggagggttc aggatcacct acgaggacct ggaggggggc	840
aacatccccg cgctcctga tgtggaggcc taccaggata gcttgaagga aaatgaggcg	900
ggacaggagg ataccgcccc cgccgcctcc gccgccgccc agcagggcga ggatgctgct	960
gacaccgagg ccgaggacgg ggcggaggcc gaccccgcta tgggtggtgga ggctgccgag	1020
caggaggagg acatgaatga cagtgcggtg cgcgagaca ccttcgtcac ccggggggag	1080
gaaaagcaag cggaggccga ggccgaggcc gaggaagc aactggcggc agcagcggcg	1140
gcggcgcggt tggccgaggc ggaggctgag tctgaggga ccaagccgc caaggagccc	1200
gtgattaagc cctgaccga agatagcaag aagcgcagtt acaacctgct caaggacagc	1260

accaacaccg cgtaccgcag ctggctacctg gcctacaact acggcgaccc gtcgacgggg 1320
 gtgcgtcctt ggaccctgct gtgcacgccg gacgtgacct gcggctcgga gcaggtgtac 1380

 tggtcgctgc ccgacatgat gcaagacccc gtgaccttcc gctccacgcg gcaggtcagc 1440
 aacttcccgg tgggtgggcgc cgagctgctg cccgtgcact ccaagagctt ctacaacgac 1500
 caggccgtct actcccagct catccgccag ttacacctc tgaccacagt gttcaatcgc 1560
 ttctctgaga accagattct ggcgcgcccg cccgccccca ccatcaccac cgtcagtga 1620
 aacgttctctg ctctcacaga tcacgggacg ctaccgtgc gcaacagcat cggaggagtc 1680
 cagcgagtga ccgttactga cgccagacgc cgcacctgcc cctacgttta caaggccttg 1740
 ggcatagtct cgccgcgcgt cctttccagc cgcactttt 1779

<210> 5

<211> 958

<212> PRT

<213> Pan troglodytes

<400> 5

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser

1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala

20 25 30

Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro

35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu

50 55 60

Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr

65 70 75 80

Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met

85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr

100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly

115 120 125

Ala Pro Asn Ser Cys Glu Trp Glu Gln Leu Glu Glu Ala Gln Ala Ala

130 135 140
 Val Glu Asp Glu Glu Leu Glu Asp Glu Asp Glu Glu Pro Gln Asp Glu
 145 150 155 160
 Ala Pro Val Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu Ser Gly
 165 170 175
 Glu Glu Ile Thr Lys Asn Gly Leu Gln Ile Gly Ser Asp Asn Thr Glu

 180 185 190
 Ala Gln Ser Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro Glu Pro
 195 200 205
 Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val Ala Gly
 210 215 220
 Gly Arg Val Leu Lys Lys Ser Thr Pro Met Lys Pro Cys Tyr Gly Ser
 225 230 235 240
 Tyr Ala Arg Pro Thr Asn Ser Asn Gly Gly Gln Gly Val Leu Val Ala

 245 250 255
 Asp Asp Lys Gly Val Leu Gln Ser Lys Val Glu Leu Gln Phe Phe Ser
 260 265 270
 Asn Thr Thr Thr Leu Asn Gln Arg Glu Gly Asn Asp Thr Lys Pro Lys
 275 280 285
 Val Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr His
 290 295 300
 Ile Ser Tyr Lys Pro Thr Lys Ser Asp Asp Asn Ser Lys Ile Met Leu

 305 310 315 320
 Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg Asp
 325 330 335
 Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly Val
 340 345 350
 Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln Asp
 355 360 365
 Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly Asp

 370 375 380

Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr Asp
385 390 395 400
Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu Pro
405 410 415
Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr Gln
420 425 430
Ala Ile Lys Thr Asn Gly Asn Gly Gln Glu Asn Pro Thr Trp Glu Lys
435 440 445
Asp Thr Glu Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn Asn Phe
450 455 460
Ala Met Glu Ile Asn Leu Ser Ala Asn Leu Trp Arg Asn Phe Leu Tyr
465 470 475 480
Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn Pro Ser
485 490 495
Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met Asn Lys
500 505 510
Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu Gly Ala
515 520 525
Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn His His
530 535 540
Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg
545 550 555 560
Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys
565 570 575
Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg
580 585 590
Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg
595 600 605
Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr Ala Thr
610 615 620
Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu

625 630 635 640
 Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala
 645 650 655
 Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser
 660 665 670
 Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe Thr Arg
 675 680 685
 Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr

 690 695 700
 Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu
 705 710 715 720
 Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser Val Ser
 725 730 735
 Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys
 740 745 750
 Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr

 755 760 765
 Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr
 770 775 780
 Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr Ser Phe
 785 790 795 800
 Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln Thr Lys
 805 810 815
 Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn Asn Ser

 820 825 830
 Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln Ala Tyr
 835 840 845
 Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val Asp Ser
 850 855 860
 Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg Ile Pro
 865 870 875 880

Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Ser Asp Leu Gly Gln

885 890 895
Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu

900 905 910
Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val

915 920 925
Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Thr

930 935 940
Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr

945 950 955

<210> 6

<211> 2874

<212> DNA

<213> Pan troglodytes

<400> 6

atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc	60
tcggagtacc tgagccccgg gctggtgcag ttgccccg cgaccgagag ctacttcagc	120
ctgagtaaca agtttaggaa cccacgggtg gcgcccacgc acgatgtgac caccgaccgg	180
tctcagcgcc tgacgtgctg gttcattccc gtggaccgcg aggacaccgc gtactcgtac	240
aaggcgcggt tcaccctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac	300
tttgacatcc gcgggggtgct ggaccggggc cccactttca agccttactc tggcaccgcc	360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca attagaagaa	420
gcccaggccg ctgtggaaga cgaagaatta gaagatgaag acgaggaacc acaggatgag	480
gcacctgtga aaaaaaccca tgtatacgct caggctcccc tttctggaga agaaattact	540
aaaaacggtt tgcaaatagg gtcagataac acagaagccc agtctaagcc catatatgca	600
gacacctacat tccagcctga accccaaatc ggggaatccc agtggaatga ggcagatgct	660
acagttgccg gcggtagagt gctaaagaaa tccactccca tgaagccatg ctatggttcc	720
tatgcaagac ccacaaactc caatggaggt caaggtgtgc tggtggtga tgataagggg	780
gttttcaat ctaaagtga attgcaattt ttttcaata ctactactct taatcagcgg	840
gagggtaacg atacaaaacc aaaagtgggt ctgtatagcg aagatgtgca catggaaact	900
ccagacaccc acattttcta caagcccaca aaaagcgatg acaattcaaa aatcatgctg	960

ggtcagcagt ccatgcccaa cagacctaat tacatcggct tcagagacaa ctttatcggc	1020
ctcatgtatt acaatagcac tggcaacatg ggagtgcctg caggtcaggc ctctcagttg	1080
aatgcagtgg tggacttgca agacagaaac acagaactgt cctaccagct ctctgcttgat	1140
tccatgggtg acagaaccag atacttttcc atgtggaatc aggcagtgga cagttatgac	1200
ccagatgtca gaattattga aaatcatgga actgaagacg agctcccaa ctattgtttc	1260
cctctgggcg gcataggggt aactgacact taccaggcca ttaaaaccaa tggcaatggt	1320
caagaaaacc caacctggga aaaagataca gagtttgcag accgcaatga aataggggtg	1380
ggaaacaatt tcgctatgga gatcaacctc agtccaacc tgtggagaaa ctctctgtac	1440
tccaacgtgg cgctgtacct gccagacaag cttaagtaca accctccaa tgtggacatc	1500
tctgacaacc ccaacaccta cgattacatg aacaagcgag tggaggcccc ggggctggtg	1560
gactgtctaca tcaacctggg cgcgcgctgg tcgctggact acatggacaa cgtcaacccc	1620
ttcaaccacc accgcaatgc gggcctgcgc taccgctcca tgctcctggg caacgggcgc	1680
tacgtgccct tccacatcca ggtgccccag aagtctcttg ccatcaagaa cctcctctc	1740
ctgccgggct cctacaccta cgagtggaaac ttcaggaagg atgtcaacat ggtcctccag	1800
agctctctgg gtaacgatct cagggtggac ggggccagca tcaagttcga gagcatctgc	1860
ctctacgcca cttcttccc catggccac aacacggcct ccacgctcga ggccatgctc	1920
aggaacgaca ccaacgacca gtccttcaat gactacctct ccgcccga catgctctac	1980
cccatacccg ccaacgccac caacgtcccc atctccatcc cctcgcgcaa ctgggcggcc	2040
ttccgcggtt gggccttcac ccgcctcaag accaaggaga cccctccct gggctcggga	2100
ttcgaccctt actacaccta ctggggtcc attccctacc tggacggcac ctctacctc	2160
aaccacattt tcaagaaggt ctcggtcacc ttgactcct cggtcagctg gccgggcaac	2220
gaccgtctgc tcacccccaa cgagttcgaa atcaagcgct cggtcgacgg ggagggtac	2280
aacgtggccc agtgcaacat gaccaaggac tggttcctgg tccagatgct ggccaactac	2340
aacatcggtt accagggtt ctacatcca gagagctaca aggacaggat gtactcctc	2400
ttcaggaact tccagcccat gagccggcag gtggtggacc agaccaagta caaggactac	2460
caggaggtgg gcatcatcca ccagcacaac aactcgggtt tcgtgggcta cctcgcccc	2520
accatgcgcg agggacagge ctaccccgcc aacttccct acccgctcat aggcaagacc	2580
gcggtcgaca gcatcacca gaaaaagtgc ctctgcgac gcacctctg gcgcatcccc	2640
ttctccagca acttcatgtc catgggtgcg ctctcgacc tgggccagaa ctgtctctac	2700
gccaaactcg cccacgcct cgacatgacc ttcgaggtcg acccatgga cgagccacc	2760

cttctctatg ttctgttcga agtctttgac gtggtccggg tccaccagcc gcaccgcggc 2820

gtcatcgaga ccgtgtacct gcgtacgcc ttctcgccg gcaacgccac cacc 2874

<210> 7

<211> 578

<212> PRT

<213> Pan troglodytes

<400> 7

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu

50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser

65 70 75 80

Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn

85 90 95

Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala

100 105 110

Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu

115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr

130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu

145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val

165 170 175

Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp

180 185 190

Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
195 200 205

Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
210 215 220

Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
225 230 235 240

Ser Gly Ala Leu Asn Tyr Asp Thr Ser Gly Asn Leu Glu Leu Arg Ala
245 250 255

Ala Gly Gly Met Arg Val Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
290 295 300

Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305 310 315 320

Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
325 330 335

Ile Asn Ala Gly Asp Gly Leu Gln Phe Asp Ser Gly Ser Asp Thr Asn
340 345 350

Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
355 360 365

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr

435 440 445
Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460
Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480
Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
500 505 510
Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
515 520 525
Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
530 535 540
Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
565 570 575
Gln Glu

<210> 8

<211> 1734

<212> DNA

<213> Pan troglodytes

<400> 8

atgaagcgca caaaacgtc tgacgagagc ttcaaccccg tgtacccta tgacacggaa	60
agcggccctc cctccgtccc ttctctcacc cctcccttcg tgtctcccga tggattccaa	120
gaaagtcccc ccggggtcct gtctctgaac ctggccgagc ccctgggtcac ttcccacggc	180
atgctcgccc tgaaaatggg aagtggcctc tccttgagcg acgtggcaa cctcacctct	240
caagatatca ccaccgctag ccctccctc aaaaaacca agaccaacct cagcctagaa	300
acctcatccc ccctaactgt gaggacctca ggccgacctc ccgtagcagc cgccgctccc	360
ctggcggtgg ccggcacctc cctcacatg caatcagagg cccctctgac agtacaggat	420
gcaaaactca cctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg	480

caaacatcgg ccccgctgac ggccgctgac agcagcaccc tcacagtcag tgccacacca 540
 ccccttagca caagcaatgg cagcttgggt attgacatgc aagccccat ttacaccacc 600
 aatggaaaac taggacttaa ctttggcgct cccctgcatg tggtagacag cctaaatgca 660

ctgactgtag ttactggcca aggtcttacg ataaacggaa cagccctaca aactagagtc 720
 tcaggtgccc tcaactatga cacatcagga aacctagaat tgagagctgc aggggggtatg 780
 cgagttgatg caaatggta acttatcctt gatgtagctt acccatittga tgcacaaaac 840
 aatctcagcc ttaggcttgg acagggaccc ctgtttgtta actctgcca caacttggat 900
 gttactaca acagaggcct ctacctgttc acatctggaa atacaaaaa gctagaagtt 960
 aatatcaaaa cagccaaggg tctcatittat gatgacactg ctatagcaat caatgcgggt 1020
 gatgggctac agtttgactc aggtcagat acaaatccat taaaaactaa acttggatta 1080

ggactggatt atgactccag cagagccata attgctaaac tgggaactgg cctaagcttt 1140
 gacaacacag gtgccatcac agtaggcaac aaaaatgatg acaagcttac cttgtggacc 1200
 acaccagacc catcccctaa ctgtagaatc tattcagaga aagatgctaa attcacactt 1260
 gttttgacta aatgcggcag tcaggtgttg gccagcgttt ctgtttttatc tgtaaaaggt 1320
 agccttgcgc ccatcagtgg cacagtaact agtgcctcaga ttgtcctcag atttgatgaa 1380
 aatggagtgc tactaagcaa ttcttccctt gacctcaat actggaacta cagaaaaggt 1440
 gaccttacag agggcactgc atataccaac gcagtgggat ttatgcccaa cctcacagca 1500

tacccaaaaa cacagagcca aactgctaaa agcaacattg taagtcaggt ttacttgaat 1560
 ggggacaaat ccaaaccat gacctcacc attacctca atggaactaa tgaaacagga 1620
 gatgccacag taagcactta ctccatgtca ttctcatgga actggaatgg aagtaattac 1680
 attaatgaaa cgttccaaac caactcctc accttctct acatcgcca agaa 1734

<210> 9

<211> 593

<212> PRT

<213> Pan troglodytes

<400> 9

Met Arg Arg Ala Ala Met Tyr Gln Glu Gly Pro Pro Pro Ser Tyr Glu
 1 5 10 15

Ser Val Val Gly Ala Ala Ala Ala Ala Pro Ser Ser Pro Phe Ala Ser
 20 25 30

Gln Leu Leu Glu Pro Pro Tyr Val Pro Pro Arg Tyr Leu Arg Pro Thr

35 40 45
 Gly Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Phe Asp
 50 55 60
 Thr Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Val Ala Ser
 65 70 75 80

 Leu Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Ile Gln
 85 90 95
 Asn Asn Asp Tyr Ser Pro Ser Glu Ala Ser Thr Gln Thr Ile Asn Leu
 100 105 110
 Asp Asp Arg Ser His Trp Gly Gly Asp Leu Lys Thr Ile Leu His Thr
 115 120 125
 Asn Met Pro Asn Val Asn Glu Phe Met Phe Thr Asn Lys Phe Lys Ala
 130 135 140

 Arg Val Met Val Ser Arg Ser His Thr Lys Glu Asp Arg Val Glu Leu
 145 150 155 160
 Lys Tyr Glu Trp Val Glu Phe Glu Leu Pro Glu Gly Asn Tyr Ser Glu
 165 170 175
 Thr Met Thr Ile Asp Leu Met Asn Asn Ala Ile Val Glu His Tyr Leu
 180 185 190
 Lys Val Gly Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys
 195 200 205

 Phe Asp Thr Arg Asn Phe Arg Leu Gly Leu Asp Pro Val Thr Gly Leu
 210 215 220
 Val Met Pro Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Ile
 225 230 235 240
 Leu Leu Pro Gly Cys Gly Val Asp Phe Thr Tyr Ser Arg Leu Ser Asn
 245 250 255
 Leu Leu Gly Ile Arg Lys Arg Gln Pro Phe Gln Glu Gly Phe Arg Ile
 260 265 270

 Thr Tyr Glu Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val
 275 280 285

Glu Ala Tyr Gln Asp Ser Leu Lys Glu Asn Glu Ala Gly Gln Glu Asp
 290 295 300
 Thr Ala Pro Ala Ala Ser Ala Ala Ala Glu Gln Gly Glu Asp Ala Ala
 305 310 315 320
 Asp Thr Ala Ala Ala Asp Gly Ala Glu Ala Asp Pro Ala Met Val Val
 325 330 335

 Glu Ala Pro Glu Gln Glu Glu Asp Met Asn Asp Ser Ala Val Arg Gly
 340 345 350
 Asp Thr Phe Val Thr Arg Gly Glu Glu Lys Gln Ala Glu Ala Glu Ala
 355 360 365
 Ala Ala Glu Glu Lys Gln Leu Ala Ala Ala Ala Ala Ala Ala Leu
 370 375 380
 Ala Ala Ala Glu Ala Glu Ser Glu Gly Thr Lys Pro Ala Lys Glu Pro
 385 390 395 400

 Val Ile Lys Pro Leu Thr Glu Asp Ser Lys Lys Arg Ser Tyr Asn Leu
 405 410 415
 Leu Lys Asp Ser Thr Asn Thr Ala Tyr Arg Ser Trp Tyr Leu Ala Tyr
 420 425 430
 Asn Tyr Gly Asp Pro Ser Thr Gly Val Arg Ser Trp Thr Leu Leu Cys
 435 440 445
 Thr Pro Asp Val Thr Cys Gly Ser Glu Gln Val Tyr Trp Ser Leu Pro
 450 455 460

 Asp Met Met Gln Asp Pro Val Thr Phe Arg Ser Thr Arg Gln Val Ser
 465 470 475 480
 Asn Phe Pro Val Val Gly Ala Glu Leu Leu Pro Val His Ser Lys Ser
 485 490 495
 Phe Tyr Asn Asp Gln Ala Val Tyr Ser Gln Leu Ile Arg Gln Phe Thr
 500 505 510
 Ser Leu Thr His Val Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Ala
 515 520 525

 Arg Pro Pro Ala Pro Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala

530 535 540
 Leu Thr Asp His Gly Thr Leu Pro Leu Arg Asn Ser Ile Gly Gly Val
 545 550 555 560
 Gln Arg Val Thr Val Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val
 565 570 575
 Tyr Lys Ala Leu Gly Ile Val Ser Pro Arg Val Leu Ser Ser Arg Thr
 580 585 590

Phe

<210> 10

<211> 1779

<212> DNA

<213> Pan troglodytes

<400> 10

atgcggcgcg cggcgaatga ccaggaggga cctcctccct cttacgagag cgtgggtgggc 60
 gcggcggcgg cggcgccttc ttctcccttt gcgtcgagc tgctggagcc gccgtacgtg 120
 cctccgcgct acctgcggcc tacggggggg agaaacagca tccgttactc ggagctggcg 180
 cccctgttcg acaccacccg ggtgtacctg gtggacaaca agtcggcgga cgtggcctcc 240
 ctgaactacc agaacgacca cagcaatttt ttgaccacgg tcatccagaa caatgactac 300
 agcccagcgc aggccagcac ccagaccatc aatctggatg accggtcgca ctggggcggc 360

 gacctgaaaa ccacctcgca caccaacatg cccaacgtga acgagttcat gttcaccaat 420
 aagttcaagg cgcgggtgat ggtgtcgcg tcgcacacca aggaagaccg ggtggagctg 480
 aagtagcagt gggtggagtt cgagctgcca gagggcaact actccgagac catgaccatt 540
 gacctgatga acaacgcgat cgtggagcac tatctgaaag tgggcaggca gaacggggtc 600
 ctggagagcg acatcggggt caagttcgac accaggaact tccgcctggg gctggacccc 660
 gtgaccgggc tggttatgcc cggggtgtac accaacgagg ccttccatcc cgacatcatc 720
 ctgctgcccc gctgcggggt ggacttcact tacagccgcc tgagcaacct cctgggcatac 780

 cgcaagcggc agcccttcca ggagggttc aggatcacct acgaggacct ggaggggggc 840
 aacatccccg cgctcctcga tgtggaggcc taccaggata gcttgaagga aatgaggcg 900
 ggacaggagg ataccgcccc gccgcctcc gccgccgcg agcagggcga ggatgctgct 960
 gacaccgcgg ccgcggacgg ggcagaggcc gacccccta tgggtgtgga ggctcccag 1020
 caggaggagg acatgaatga cagtgcggtg cgcggagaca ccttcgtcac ccggggggag 1080

gaaaagcaag cggaggccga ggccgcgcc gaggaagc aactggcggc agcagcggcg 1140
gcggcggcgt tggccgcgcc ggagctgag tctgagggga ccaagcccgc caaggagccc 1200

gtgattaagc cctgaccga agatagcaag aagcgcagtt acaacctgct caaggacagc 1260
accaacaccg cgtaccgcag ctggtacctg gcctacaact acggcgaccc gtcgacgggg 1320
gtgcgtcctt ggacctgct gtgcacgccg gacgtgacct gcggctcgga gcaggtgtac 1380
tggtcgctgc ccgacatgat gcaagacccc gtgaccttcc gctccacgcg gcaggtcagc 1440
aacttcccgg tggtagggcg cgagctgctg cccgtgcact ccaagagctt ctacaacgac 1500
caggccgtct actcccagct catccgccag ttacaccttc tgaccacgt gttcaatcgc 1560
tttctgaga accagattct ggccgccccg cccgccccca ccatcaccac cgtcagttaa 1620

aacgttctct ctctcacaga tcacgggacg ctaccgtgc gcaacagcat cggaggagtc 1680
cagcgagtga ccgttactga cgccagacgc cgcacctgcc cctacgttta caaggccttg 1740
ggcatagtct cgccgcgcgt cctttccagc cgcactttt 1779

<210> 11

<211> 960

<212> PRT

<213> Pan troglodytes

<400> 11

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser

1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala

20 25 30

Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro

35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu

50 55 60

Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr

65 70 75 80

Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met

85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr

100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
 115 120 125
 Ala Pro Asn Ser Cys Glu Trp Glu Gln Glu Glu Thr Gln Ala Val Glu
 130 135 140
 Glu Ala Ala Glu Glu Glu Glu Glu Asp Ala Asp Gly Gln Ala Glu Glu
 145 150 155 160

 Glu Gln Ala Ala Thr Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu
 165 170 175
 Ser Gly Glu Lys Ile Ser Lys Asp Gly Leu Gln Ile Gly Thr Asp Ala
 180 185 190
 Thr Ala Thr Glu Gln Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro
 195 200 205
 Glu Pro Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val
 210 215 220

 Ala Gly Gly Arg Val Leu Lys Lys Ser Thr Pro Met Lys Pro Cys Tyr
 225 230 235 240
 Gly Ser Tyr Ala Arg Pro Thr Asn Ala Asn Gly Gly Gln Gly Val Leu
 245 250 255
 Thr Ala Asn Ala Gln Gly Gln Leu Glu Ser Gln Val Glu Met Gln Phe
 260 265 270
 Phe Ser Thr Ser Glu Asn Ala Arg Asn Glu Ala Asn Asn Ile Gln Pro
 275 280 285

 Lys Leu Val Leu Tyr Ser Glu Asp Val His Met Glu Thr Pro Asp Thr
 290 295 300
 His Leu Ser Tyr Lys Pro Ala Lys Ser Asp Asp Asn Ser Lys Ile Met
 305 310 315 320
 Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile Gly Phe Arg
 325 330 335
 Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly Asn Met Gly
 340 345 350

 Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val Asp Leu Gln

355 360 365
 Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp Ser Met Gly
 370 375 380
 Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
 385 390 395 400
 Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu
 405 410 415

 Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr
 420 425 430
 Gln Ala Val Lys Thr Asn Asn Gly Asn Asn Gly Gly Gln Val Thr Trp
 435 440 445
 Thr Lys Asp Glu Thr Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn
 450 455 460
 Asn Phe Ala Met Glu Ile Asn Leu Ser Ala Asn Leu Trp Arg Asn Phe
 465 470 475 480

 Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn
 485 490 495
 Pro Ser Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met
 500 505 510
 Asn Lys Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu
 515 520 525
 Gly Ala Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn
 530 535 540

 His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn
 545 550 555 560
 Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala
 565 570 575
 Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn
 580 585 590
 Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp
 595 600 605

Leu Arg Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr
 610 615 620
 Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala
 625 630 635 640
 Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser
 645 650 655
 Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro
 660 665 670

 Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe
 675 680 685
 Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp
 690 695 700
 Pro Tyr Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe
 705 710 715 720
 Tyr Leu Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser
 725 730 735

 Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu
 740 745 750
 Ile Lys Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn
 755 760 765
 Met Thr Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile
 770 775 780
 Gly Tyr Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr
 785 790 795 800

 Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln
 805 810 815
 Thr Lys Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn
 820 825 830
 Asn Ser Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln
 835 840 845
 Ala Tyr Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val

850	855	860	
Asp Ser Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg			
865	870	875	880
Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Ser Asp Leu			
	885	890	895
Gly Gln Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr			
	900	905	910
Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe			
	915	920	925
Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile			
	930	935	940
Glu Thr Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr			
945	950	955	960
<210> 12			
<211> 2880			
<212> DNA			
<213> Pan troglodytes			
<400> 12			
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc			60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc			120
ctgagtaaca agtttaggaa cccacgggtg gcgcccacgc acgatgtgac caccgaccgg			180
tctcagcgcc tgacgtgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac			240
aaggcgcggt tcacctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac			300
tttgacatcc gcggggtgct ggaccggggt cccactttca agccctactc tggcaccgcc			360
tacaactccc tgcccccaa gggcgctccc aactcctgcg agtgggagca agaggaaact			420
caggcagttg aagaagcagc agaagaggaa gaagaagatg ctgacggtca agctgaggaa			480
gagcaagcag ctacaaaaa gactcatgta tatgtcagg ctcccctttc tggcgaaaaa			540
attagtaaag atggtctgca aataggaacg gacgtacag ctacagaaca aaaacctatt			600
tatgcagacc ctacattcca gcccgaaacc caaatcgggg agtcccagtg gaatgaggca			660
gatgctacag tcgccggcgg tagagtgcta aagaaatcta ctcccatgaa accatgctat			720
ggttcctatg caagaccac aaatgcta atggaggtcagg gtgtactaac ggcaaatgcc			780

cagggacagc tagaatctca gggtgaaatg caattctttt caacttctga aaacgcccgt 840
aacgaggcta acaacattca gcccaaattg gtgctgtata gtgaggatgt gcacatggag 900
accccgata cgcaccttcc ttacaagccc gcaaaaagcg atgacaattc aaaaatcatg 960
ctgggtcagc agtccatgcc caacagacct aattacatcg gcttcagaga caactttatc 1020

ggcctcatgt attacaatag cactggcaac atgggagtgc ttgcaggtca ggcctctcag 1080
ttgaatgcag tgggtgactt gcaagacaga aacacagaac tgtcctacca gctcttgctt 1140
gattccatgg gtgacagaac cagatacttt tccatgtgga atcaggcagt ggacagtat 1200
gaccagatg ttagaattat tgaatatcat ggaactgaag acgagctccc caactattgt 1260
ttccctctgg gtggcatagg ggtaactgac acttaccagg ctgttaaaac caacaatggc 1320
aataacgggg gccagggtgac ttggacaaaa gatgaaactt ttgcagatcg caatgaaata 1380
gggggtgggaa acaatttcgc tatggagatc aacctcagt ccaacctgtg gagaaacttc 1440

ctgtactcca acgtggcgct gtacctacca gacaagctta agtacaaccc ctccaatgtg 1500
gacatctctg acaaccccaa cacctacgat tacatgaaca agcgagtggg ggccccgggg 1560
ctggtggact gctacatcaa cctgggcgcg cgtgggtcgc tggactacat ggacaacgtc 1620
aacccttca accaccaccg caatgcgggc ctgcgtacc gctccatgct cctgggcaac 1680
gggcgtacg tgccttcca catccagggt ccccaagaat tctttgcat caagaacctc 1740
ctcctctgc cgggctcta cacctacgag tggaaactca ggaaggatgt caacatggtc 1800
ctccagagct ctctgggtaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc 1860

atctgcctct acgccactt ctccccatg gccacaaca cggcctccac gctcgaggcc 1920
atgctcagga acgacaccaa cgaccagtcc ttcaatgact acctctccgc cgccaacatg 1980
ctctaccca taccgcca cccaccaac gtccccatct ccatccctc gcgcaactgg 2040
gcggccttcc ggggtgggc cttcacccgc ctcaagacca aggagacccc ctccctgggc 2100
tcgggattcg acccctacta cacctactcg ggctccattc cctacctgga cggcaccttc 2160
tacctcaacc aactttcaa gaaggtctcg gtcaccttcg actcctcggt cagctggccg 2220
ggcaacgacc gtctgtcac cccaacgag ttcgagatca agcgtcgggt cgacggggag 2280

ggctacaacg tggcccagtg caacatgacc aaggactggg tcctgggtcca gatgctggcc 2340
aactacaaca tcggctacca gggcttctac atcccagaga gctacaagga caggatgtac 2400
tccttcttca ggaacttcca gcccatgagc cggcaggtgg tggaccagac caagtacaag 2460
gactaccagg aggtgggcat catccaccag cacaacaact cgggcttcgt gggctacctc 2520
gccccacca tgcgcgaggg acaggcctac cccgccaact tccctatcc gctcatagge 2580
aagaccgagg tcgacagcat caccagaaa aagttctct gcgaccgcac cctctggcgc 2640

atcccccttct ccagcaactt catgtccatg ggtgcgctct cggacctggg ccagaacttg 2700

ctctacgcca actccgcca cgccctcgac atgaccttcg aggtcgacce catggacgag 2760

cccacccttc tctatgttct gttcgaagtc tttgacgtgg tccgggtcca ccagccgcac 2820

cgcggcgtca tcgagaccgt gtacctgcgt acgcccttct cggccggcaa cgccaccacc 2880

<210> 13

<211> 37830

<212> DNA

<213> Pan troglodytes

<400> 13

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg 60

cggggcggga ggcggtccg ggggcgggcc ggcgggcggg gcggtgtggc ggaagtggac 120

tttctaagtg tggcggatgt gacttgctag tgccgggcgc ggtaaaagtg acgttttccg 180

tgcgcgacaa cgcccacggg aagtacatt tttcccgcg tttttaccgg atgtttagt 240

gaatttgggc gtaaccaagt aagatttggc cattttcgcg ggaaaactga aacggggaag 300

tgaaatctga ttaatttcgc gttagtcata ccgcgtaata tttgtcgagg gccgaggac 360

tttggccgat tacgtggagg actcggccag gtgttttttg aggtgaattt ccgcgttccg 420

ggtcaaagtc tccgttttat tattatagtc agctgacgcg gagtgtattt atacctctg 480

atctcgtcaa gtggccactc ttgagtcca gcgagtagag ttttctctc tgccgctctc 540

cgtccgctc cgctcggctc tgacaccggg gaaaaaatga gacatttcac ctacgatggc 600

ggtgtgctca ccggccagct ggctgctgaa gtcttgaca cctgatcga ggaggtattg 660

gccgataatt atctccctc gactcctttt gagccaccta cacttcacga actctacgat 720

ctggatgtgg tggggcccag cgatccgaac gacgagcggg tttccagttt tttccagag 780

tccatgttgt tggccagcca ggagggggtc gaacttgaga cccctcctcc gatcgtggat 840

tcccccgatc cgccgcagct gactaggcag ccgagcgct gtgcgggacc tgagactatg 900

ccccagctgc tacctgaggt gatcgatctc acctgtaatg agtctggtt tccaccagc 960

gaggatgagg acgaagaggg tgagcagttt gtgttagatt ctgtggaaca acccgggcga 1020

ggatgcaggt cttgtcaata tcaccgaaa aacacaggag actcccagat tatgtgttct 1080

ctgtgttata tgaagatgac ctgtatgttt atttacagta agtttatcat ctgtgggcag 1140

gtgggctata gtgtgggtgg tggcttttgg ggggtttttt aatatatgtc aggggttatg 1200

ctgaagactt ttttattgtg atttttaag gtccagtgtc tgagcccag caagaacctg 1260

aaccggagcc tgagccttct cgccccagga gaaagcctgt aatcttaact agaccagcg 1320

caccggtagc gagaggcctc agcagcgcg agaccaccga ctccggtgct tcctcatcac	1380
ccccggagat tcacccctg gtgcccctgt gtcccgttaa gcccgttgcc gtgagagtca	1440
gtggcgcgcg gtctgctgtg gagtgcattg aggacttgct ttttgattca caggaacctt	1500
tggacttgag ctigaaacgc cccaggcatt aaacctggtc acctggactg aatgagttga	1560
cgcctatgtt tgcttttgaa tgacttaatg tgtatagata ataaagagtg agataatgtt	1620
ttaattgcat ggtgtgttta acttggcgcg agtctgctgg gtatataagc ttccctgggc	1680
taaacttggc tacacttgac ctcatggagg cctgggagtg tttggagaac tttgccggag	1740
ttcgtgcctt gctggacgag agctctaaca atacctcttg gtggtggagg tatttgtggg	1800
gctctcccca gggcaagtta gttttagaa tcaaggagga ttacaagtgg gaatttgaag	1860
agcttttgaa atcctgtggt gagctattgg attctttgaa tctaggccac caggctctct	1920
tccaggagaa ggtcatcagg actttggatt ttccacacc ggggcgcatt gcagccgcgg	1980
ttgcttttct agcttttttg aaggatagat ggagcgaaga gaccacttg agttcgggct	2040
acgtcctgga ttttctggcc atgcaactgt ggagagcatg gatcagacac aagaacaggc	2100
tgcaactgtt gtcttcctgc cgcccgttgc tgattccggc ggaggagcaa caggccgggt	2160
cagaggaccg gggccgtcgg gatccggagg agagggcacc gagccgggc gagaggagcg	2220
cgctgaacct gggaaaccgg ctgagcggcc atccacatcg ggagtgaatg tcgggcaggt	2280
ggtggatctt ttccagaac tgcggcgat tttgactatt agggaggatg ggcaatttgt	2340
taagggtctt aagagggaga ggggggcttc tgagcataac gaggaggcca gtaatttagc	2400
ttttagcttg atgaccagac accgtccaga gtgcatcact ttccagcaga ttaaggacaa	2460
ttgtgccaat gagttagatc tgttgggtca gaagtatagc atagagcagc tgaccactta	2520
ctggctgcag ccgggtgatg atctggagga agctattagg gtgtatgcta aggtggccct	2580
gcggcccgat tgcaagtaca agctcaaggg gctggtgaat atcaggaatt gttgctacat	2640
ttctggcaac ggggcggagg tggagataga gaccgaagac aggttggctt tcagatgcag	2700
catgatgaat atgtggccgg gggctgctgg catggacggg gtggtgatta tgaatgtgag	2760
gttcacgggg cccaacttta acggcacggg gtttttgggg aacaccaacc tggctctgca	2820
cggggtgagc ttctatgggt ttaacaacac ctgtgtggag gcctggaccg atgtgaaggt	2880
ccgcggttgc gccttttatg gatgttggaa ggccatagtg agccgcccta agagcaggag	2940
ttccattaag aaatgcttgt ttgagaggig caccttgggg atcctggccg agggcaactg	3000
cagggtgcgc cacaatgtgg cctccgagtg cggttgcttc atgctagtca agagcgtggc	3060

ggtaatcaag cataatatgg tgtgcggcaa cagcgaggac aaggcctcac agatgctgac 3120

ctgcacggat ggcaactgcc acttgctgaa gaccatccat gtaaccagcc acagccggaa 3180

ggcctggccc gtgttcgagc acaacttgct gaccgctgc tccttgcatc tgggcaacag 3240

gcggggggtg ttcttgccct atcaatgcaa ctttagtcac accaagatct tgctagagcc 3300

cgagagcatg tccaagtgta acttgaacgg ggtgtttgac atgacatga agatctggaa 3360

ggtgctgagg tacgacgaga ccaggtcccg gtgcagaccc tgcgagtgcg ggggcaagca 3420

tatgaggaac cagcccgtga tgctggatgt gaccgaggag ctgaggacag accacttggt 3480

tctggcctgc accagggccg agtttggttc tagcgatgaa gacacagatt gaggtgggtg 3540

agtgggcgtg gcctggggtg gtcatgaaaa tatataagtt gggggtctta gggtctcttt 3600

atttgtgttg cagagaccgc cggagccatg agcgggagca gcagcagcag cagtagcagc 3660

agcgcttgg atggcagcat cgtgagccct tatttgacga cgcggtgcc cactgggcc 3720

ggggtgcgtc agaattgat gggctccagc atcgacggcc gaccgtcct gcccgaaat 3780

tccgccacgc tgacctatgc gaccgtcgcg gggacgccgt tggacgccac cgcgccgcc 3840

gccgccaccg cagccgcctc ggccgtgcgc agcctggcca cggactttgc attcctggga 3900

ccactggcga caggggctac ttctcgggcc gctgctgccg ccgttcgcga tgacaagctg 3960

accgcccctgc tggcgcagtt ggatgcgctt actcgggaac tgggtgacct ttctcagcag 4020

gtcatggccc tgcgccagca ggtctcctcc ctgcaagctg gcgggaatgc ttctcccaca 4080

aatgccgttt aagataaata aaaccagact ctgtttggat taaagaaaag tagcaagtgc 4140

attgctctct ttatttcata atttccgcg cgcgataggc cctagaccag cgttctcggt 4200

cgttgagggt gcggtgtatc ttctccagga cgtggtagag gtggctctgg acgttgagat 4260

acatgggcat gagccccgcc cgggggtgga ggtagcacca ctgcagagct tcatgctccg 4320

gggtggtgtt gtagatgac cagtcgtagc aggagcgtg ggcatggtgc ctaaaaatgt 4380

ccttcagcag caggccgatg gccaggggga ggcccttgggt gtaagtgttt aaaaacggt 4440

taagtggga aggtgcatt cggggagaga tgatgtgcat cttggactgt attttagat 4500

tggcgatgtt tccgccaga tcccttctgg gattcatgtt gtgcaggacc accagtacag 4560

tgtatccgtt gcacttgggg aatttgtcat gcagcttaga gggaaaagcg tggaagaact 4620

tggagacgcc ttttggcct cccagatatt ccatgcattc gtccatgatg atggcaatgg 4680

gccccggga ggcagcttg gcaaagatat ttctggggtc gctgacgtcg tagttgtgtt 4740

ccagggtgag gtcgtcatag gccattttta caaagcgcg gcggagggtg cccgactggg 4800

ggatgatggt cccctctggc cctggggcgt agttgccctc gcagatctgc atttcccagg 4860
 ccttaatctc ggagggggga atcatatcca cctgcggggc gatgaagaaa acggtttccg 4920
 gagccgggga gattaactgg gatgagagca ggtttctaag cagctgtgat tttccacaac 4980
 cgggtgggccc ataaataaca cctataaccg gttgcagctg gtagtttaga gagctgcagc 5040
 tgccgtcgtc ccggaggagg ggggccacct cgttgagcat gtcctgacg cgcatttct 5100
 ccccagccag atccgccaga aggcgctcgc cggccaggga cagcagctct tgcaaggaag 5160
 caaagttttt cagcggcttg aggcgcgtccg ccgtgggcat gtttttcagg gtctggctca 5220

 gcagctccag gcggtccag agctcgggtga cgtgctctac ggcatctcta tccagcatat 5280
 ctctctgttt cgcgggttgg ggcgactttc gctgtagggc accaagcggg ggtcgtccag 5340
 cggggccaga gtcatgtcct tccatgggag cagggtcctc gtcagggtgg tctgggtcac 5400
 ggtgaagggg tgcgtccggg gctgagcgtc tgccaagggt cgcttgaggc tggttctgct 5460
 ggtgctgaag cgctgccggt cttcgccctg cgcgtcggcc aggtagcatt tgaccatggt 5520
 gtcatagtcc agccccctcg cggcgtgtcc cttggcgcgc agcttgccct tggaggtggc 5580
 gccgcacgag gggcagagca ggctcttgag cgcgtagagc ttggggggcga ggaagaccga 5640

 ttcgggggag taggcgtccg cgccgcagac cccgcacacg gtctcgact ccaccagcca 5700
 ggtgagctcg gggcgcgcgg ggtcaaaaac caggtttccc ccatgctttt tgatgcgttt 5760
 cttacctcgg gtctccatga ggtggtgtcc ccgctcgggtg acgaagaggc tgtccgtgtc 5820
 tccgtagacc gacttgaggg gtcttttctc caggggggtc cctcggctct cctcgtagag 5880
 gaactcggac cactctgaga cgaaggcccg cgtccaggcc aggacgaagg aggtatgtg 5940
 ggaggggtag cggctgttgt ccaactagggg gtccaccttc tccaagggtg gaagacacat 6000
 gtcgccttcc tcggcgtcca ggaagggtgat tggctttag gttaggcca cgtgaccggg 6060

 ggttctgtac ggggggggtat aaaagggggg gggggcgcgc tcgtcgtcac tctcttccgc 6120
 atcgtgtctc gcgagggcca gctgttgggg ttagtattcc ctctcgaagg cgggcatgac 6180
 ctccgcgtg aggttgtcag ttccaaaaa cgaggaggat ttgatgttca cctgtcccga 6240
 ggtgatacct ttgagggtac ccgcgtccat ctggtcagaa aacacgatct ttttattgtc 6300
 cagcttgggt gcgaacgacc cgtagagggc gttggagagc agcttggcga tggagcgag 6360
 ggtctggttc ttgtccctgt cggcgcgtc cttggccgcg atgttagct gcacgtactc 6420
 gcgcgcgacg cagcgcact cggggaagac ggtggtgcgc tcgtcgggca ccaggcgac 6480

 gcgccagccg cggttgtgca ggggtaccag gtccacgtg gtggcgacct cgccgcgag 6540
 gcgctcgttg gtccagcaga gacggccgc cttgcgcgag cagaaggggg gcagggggtc 6600
 gagctgggtc tcgtccgggg ggtccgcgtc cacggtgaaa accccggggc gcaggcgcg 6660

gtcgaagtag tctatcttgc aaccttgcat gtccagcgcc tgctgccagt cgcgggcggc 6720

gagcgcgcg tctaggggt tgagcggcgg gccccagggc atgggggtggg tgagtgcgga 6780

ggcgtacatg ccgcagatgt catagacgta gaggggctcc cgcaggaccc cgatgtaggt 6840

ggggtagcag cggccgcccgc ggatgctggc gcgcacgtag tcatacagct cgtgcgaggg 6900

ggcgaggagg tcggggccca ggttgggtgcg ggcggggcgc tccgcgcgga agacgatctg 6960

cctgaagatg gcatgcgagt tggaagagat ggtggggcgc tggaagacgt tgaagctggc 7020

gtcctgcagg ccgacggcgt cgcgcacgaa ggaggcgtag gagtgcgcga gcttgtgtac 7080

cagctcggcg gtgacctgca cgtcgagcgc gcagtagtcg agggctctgc ggatgatgtc 7140

atatttagcc tgccccctct tttccacag ctgcgggttg aggacaaact cttcgcggtc 7200

tttccagtac tcttgatcg ggaaaccgtc cggttccgaa cggtaagagc ctagcatgta 7260

gaactggttg acggcctggt aggcgcagca gcccttctcc acggggaggg ctaggcctg 7320

cgcggccttg cggagcgagg tgtgggtcag ggcaagggtg tccctgacca tgactttgag 7380

gtactggtgc ttgaagtcgg agtcgtcgca gccccccgc tcccagagcg agaagtcggt 7440

gcgcttcttg gagcgggggt tgggcagagc gaagggtgaca tcgttgaaga ggattttgcc 7500

cgcgcggggc atgaagtgc gggatgatgcg gaagggcccc ggcaattcag agcggttgtt 7560

gatgacctgg gcggcgagca cgatctctgc gaagccgttg atgttgtggc ccacgatgta 7620

gagttccagg aagcggggcc ggccctttac ggtgggcagc ttcttttagct cttcgtaggt 7680

gagtcctcgc ggcgaggcga ggccgtgctc ggccagggcc cagtcgcga ggtgcgggtt 7740

gtctctgagg aaggacttcc agaggtcgcg ggccaggagg gtctgcaggc ggtctctgaa 7800

ggtcctgaac tggcggccca cggccatttt ttcggggttg atgcagtaga aggtgagggg 7860

gtcttctgc cagcggctcc agtcgagctg cagggcgagg tcgcgcgcgg cggtgaccag 7920

gcgctcgtcg cccccgaatt tcatgaccag catgaaggc acgagctgct ttccgaaggc 7980

ccccatcaa gttaggtct ctacatcgta ggtgacaaag aggcgctccg tgcgaggatg 8040

cgagccgac gggaagaact ggatctcccg ccaccagttg gaggagtggc tgttgatgtg 8100

gtggaagtag aagtcccgtc gccgggcccga acactcgtgc tggcttttgt aaaagcgagc 8160

gcagtactgg cagcgtgca cgggctgtac ctcatgcacg agatgcacct ttgccccgcg 8220

cacgaggaag ccgaggggaa atctgagccc ccgcctggc tcgcggcatg gctggttctc 8280

ttctactttg gatcggtgc cgtctccgtc tggctcctcg aggggtgtta cggtaggagc 8340

gaccaccacg ccgcgcgagc cgcaggcca gatatcggcg cgcggcggtc ggagtttgat 8400

gacgacatcg cgcagctggg agctgtccat ggtctggagc tcccgcggcg gcggcaggtc 8460

agccgggagt tcttgagggt tcacctgca ggtcgggcc agggcgcggg gcaggcttag 8520

gtggtacctg atctctaggg gcgtgttggg ggcggcgtcg atggcttgca ggagcccgca	8580
gccccggggg gcgacgacgg tccccgcgg ggtggtggtg gtggtggcgg tgcagctcag	8640
aagcgggtgcc gcgggcgggc ccccgaggt agggggggct ccggtcccgc gggcaggggc	8700
ggcagcggca cgtcggcgtg gagcgcgggc aggagtgggt gctgtgccg gaggttgctg	8760
gcgaaggcga cgacgcggcg gttgatctcc tggatctggc gcctctcgt gaagacgacg	8820
ggccccgtga gcttgaacct gaaagagagt tcgacagaat caatctcgtg gtcattgacc	8880
gcggcctggc gcaggatctc ctgcacgtct cccgagttgt cttggtaggc gatctcggcc	8940
atgaactgct cgatctcttc ctctggagg tctccgcgtc cggcgcgttc cacggtggcc	9000
gccaggtcgt tggagatgcg ccccatgagc tgcgagaagg cgttgagtcc gccctcgttc	9060
cagactcggc tgtagaccac gccccctgg tcatcgcggg cgcgcatgac cacctgcgcg	9120
aggttgagct ccacgtgccg gcggaagacg gcgtagtgc gcagacgtg gaagaggtag	9180
ttgagggtgg tggcgggtg ctcgccacg aagaagtca tgaccagcg gcgcaacgtg	9240
gattcgttga tgtccccaa ggctccagc cgttccatgg cctcgtagaa gtccacggcg	9300
aagttgaaaa actgggagtt gcgcgccgac acggtcaact cctcctccag aagacggatg	9360
agctcggcga cgggtgtcgcg caccctcgcg tcgaaggcta tggggatctc ttcctccgt	9420
agcatcacca cctcctctc ttcctctct tctggcactt ccatgatggc ttcctcctct	9480
tcgggggggtg gcggcggcgg cgggtggggga gggggcgctc tgcgccggcg gcggcgcacc	9540
gggagcgggt ccacgaagcg gcgcatcatc tcccccggc ggcggcgcgt ggtctcggtg	9600
acggcgcggc cgttctcccc ggggcgcagt tgggaagacg cgcgggacat ctggtgctgg	9660
ggcgggtggc cgtgaggcag cgagacggcg ctgacgatgc atctcaaca ttgctgcgta	9720
ggtacccgc cgaggacgt gagggagtcc atatccaccg gatccgaaaa cttttcgagg	9780
aaggcgtcta accagtcgca gtcgaaggt aggctgagca ccgtggcggg cggcgggggg	9840
tggggggagt gtctggcgga ggtgctgctg atgatgtaat tgaagtaggc ggacttgaca	9900
cggcggtagg tcgacaggag caccatgtcc ttgggtccgg cctgctggat gcggaggcgg	9960
tcggctatgc cccaggcttc gttctggcat cggcgcagggt cctttagta gtcttgcatg	10020
agcctttcca ccggcacctc ttctcttcc tcttctgctt cttccatgtc tgcttcggcc	10080
ctggggcggc gccgcgcgcc cctgcccccc atgcgcgtga ccccgaacc cctgagcgggt	10140
tggagcaggg ccaggctggc gacgacgcgc tcggccagga tggcctgctg cacctgcgtg	10200
agggtggttt ggaagtcac caagtccacg aagcgggtgg aggcgcccggt gttgatggtg	10260

taggtgcagt tggccatgac ggaccagttg acggtctggt ggcccggttg cgacatctcg	10320
gtgtacctga gtgcgagta ggcgcgaggag tcgaagacgt agtcgttgca agtccgcacc	10380
aggtactggt agcccaccag gaagtgcggc ggcggtctggc ggtagagggg ccagcgcagg	10440
gtggcggggg ctccgggggc caggctcttc agcatgaggc ggtggtaggc gtagatgtac	10500
ctggacatcc aggtgatacc cgcgcgggtg gtggaggcgc gcgggaagtc gcgcacccgg	10560
ttccagatgt tgcgcagggg cagaaagtgc tccatggtag gcgtgctctg tccagtcaga	10620
cgcgcgcagt cgttgatact ctagaccagg gaaaacgaaa gccggtcagc gggcactctt	10680
ccgtggtctg gtgaatagat cgcaagggtg tcatggcgga gggcctcggg tcgagccccg	10740
ggtccgggccc ggacgggtccg ccatgatcca cgcggttacc gcccgcgtgt cgaaccagg	10800
tgtgcgacgt cagacaacgg tggagtgttc cttttggcgt ttttctggcc gggcgccggc	10860
gccgcgtaag agactaagcc gcgaaagcga aagcagtaag tggctcgtc cccgtagccg	10920
gagggatcct tgctaagggt tgcgttgcgg cgaaccccg ttcgaatccc gtactcgggc	10980
cggccggacc cgcggctaag gtgttgatt ggctccccc tcgtataaag accccgcttg	11040
cggattgact ccggacacgg ggacgagccc cttttatitt tgctttcccc agatgcatcc	11100
ggtgctgcgg cagatgcgcc ccccggcca gcagcagcaa caacaccagc aagagcggca	11160
gcaacagcag cgggagtcac gcaggggccc ctacccacc ctgcgcgggc cggccacctc	11220
ggcgtccgcg gccgtgtctg gcgcctgcgg cggcgcgagg gggcggcgtg acgacccga	11280
ggagccccg cggcgacagg ccagacacta cctggacctg gaggagggcg agggcctggc	11340
gcggctgggg gcgccgtctc ccgagcgcca ccccggggtg cagctgaagc gcgactcgcg	11400
cgaggcgtag gtgcctcggc agaacctgtt cagggaccgc gcgggcgagg agcccagga	11460
gatgcgggac aggaggttca gcgcagggcg ggagctgcgg caggggctga accgcgagcg	11520
gtgctgcgc gaggaggact ttgagcccga cgcgcggacg gggatcagcc ccgcgcgcgc	11580
gcacgtggcg gccgccgacc tggtagcggc gtacgagcag acggtgaacc aggagatcaa	11640
cttccaaaag agtttcaaca accacgtgcg cagctggtg gcgcgcgagg aggtgacct	11700
cgggctgatg cacctgtggg actttgtaag cgcgtggtg cagaaccca acagcaagcc	11760
tctgacggcg cagctgttcc tgatagtga gcacagcagg gacaacgagg cgtttaggga	11820
cgcgtgctg aacatcaccg agcccagggt tcggtggctg ctggacctga ttaacatcct	11880
gcagagcata gtggtgcagg agcgcagcct gagcctggcc gacaaggtgg cggccatcaa	11940
ctactcgatg ctgagcctgg gcaagtttta cgcgcgcaag atctaccaga cgccgtacgt	12000
gcccatagac aaggaggtag agatcgacgg tttttacatg cgcatggcgc tgaagtgct	12060
caccctgagc gacgacctgg gcgtgtaccg caacgagcgc atccacaagg ccgtgagcgt	12120

gagccggcgg cgcgagctga gcgaccgcga gctgatgcac agcctgcagc gggcgctggc	12180
gggcgccggc agcggcgaca gggaggcgga gtcctacttc gatgcggggg cggacctgcg	12240
ctgggcgccc agccggcggg ccctggaggc cgcgggggtc cgcgaggact atgacagga	12300
cggcgaggag gatgaggagt acgagctaga ggaggcgag tacctggact aaaccgcggg	12360
tggtgtttcc ggtagatgca agaccgaac gtggtggacc cggcgctgcg ggcggctctg	12420
cagagccagc cgtccggcct taactcctca gacgactggc gacaggtcat ggaccgcatc	12480
atgtcgtga cggcgcgtaa ccgggacgcg ttccggcagc agccgcaggc caacaggctc	12540
tccgccatcc tggaggcggt ggtgcctgcg cgtcgaacc ccacgcacga gaaggtgctg	12600
gccatagtga acgcgctggc cgagaacagg gccatccgcc cggacgaggc cgggctggtg	12660
tacgacgcgc tgctgcagcg cgtggcccg tacaacagcg gcaacgtgca gaccaacctg	12720
gaccggctgg tgggggacgt gcgcgaggcg gtggcgagc gcgagcgcg ggatcggcag	12780
ggcaacctgg gctccatggt ggcgtgaat gccttctga gcacgcagcc ggccaacctg	12840
ccgcgggggc aggaagacta caccaacttt gtgagcgcg tgcggctgat ggtgaccgag	12900
acccccaga gcgaggtgta ccagtcgggc ccggactact tcttcagac cagcagacag	12960
ggcctgcaga cggtagaacct gagccaggct ttcaagaacc tgcgggggct gtggggcgtg	13020
aaggcgccca ccggcgaccg ggcgacggtg tccagcctgc tgacgcccac ctgcgcctg	13080
ctgctgctgc tgatcgccc gttcagggc agcggcagcg tgtccggga cacctacctg	13140
gggcacctgc tgacctgta ccgcgaggcc atcgggcagg cgcaggtgga cgagcacacc	13200
ttccaggaga tcaccagcgt gagccgcgcg ctggggcagg aggacacgag cagcctggag	13260
gcgactctga actacctgt gaccaaccgg cggcagaaga ttccctcgct gcacagcctg	13320
acctccgagg aggagcgcat cttgcgtac gtgcagcaga gcgtgagcct gaacctgatg	13380
cgcgacgggg tgacgccag cgtggcgctg gacatgaccg cgcgcaacat ggaaccgggc	13440
atgtacgcc cgacccggcc ttacatcaac gcctgatgg actacctgca tcgcgcggcg	13500
gccgtgaacc ccgagtactt taccaacgcc atcctgaacc cgcactggct cccgccgcc	13560
gggtttctaca ggggggctt cgaggtccc gagaccaacg atggcttct gtgggacgac	13620
atggacgaca gcgtgttctc cccgcggccg caggcgctgg cggaagcgtc cctgctgcgt	13680
cccaagaagg aggaggagga ggaggcgagt cgcgcgcgc gcagcagcg cgtggcttct	13740
ctgtccgagc tggggcgggc agccgccgcg cccccgggt cctgggcgg cagccccctt	13800
ccgagcctgg tggggtctct gcacagcgag cgcaccaccc gccctcggt gctgggcgag	13860
gacgagtacc tgaataactc cctgctgcag ccggtgcggg agaaaaacct gcctcccgc	13920
ttcccaaca acgggataga gagcctggtg gacaagatga gcagatggaa gacctatgcg	13980

caggagcaca gggacgcgcc tgcgtccgg ccgccacgc ggcgccagcg ccacgaccgg	14040
cagcgggggc tggigtggga tgacgaggac tccgcggacg atagcagcgt gctggacctg	14100
ggagggagcg gcaaccctgt cgcgccctg cgcgcccgcc tggggaggat gttttaaaaa	14160
aaaaaaaaa aagcaagaag catgatgcaa aaattaata aaactcacca aggccatggc	14220
gaccgagcgt tggtttcttg tgttcccttc agtatgcggc gcgcggcgat gtaccaggag	14280
ggacctcttc cctcttacga gagcgtgggt ggcgcggcgg cggcggcgcc ctcttctccc	14340
tttgcgtcgc agctgctgga gccgccgtac gtgcctccgc gctacctgcg gcctacgggg	14400
gggagaaaca gcatccgtta ctcgagctg gcgccctgt tcgacaccac ccgggtgtac	14460
ctggtggaca acaagtgcgc ggacgtggcc tcctgaact accagaacga ccacagcaat	14520
tttttgacca cggtcaccca gaacaatgac tacagcccga gcgaggccag caccagacc	14580
atcaatctgg atgaccgtc gcactggggc ggcgacctga aaaccatcct gcacaccaac	14640
atgcccaacg tgaacgagtt catgttcacc aataagtca aggcgcgggt gatggtgtcg	14700
cgctcgaca ccaaggaaga ccgggtggag ctgaagtacg agtgggtgga gttcgagctg	14760
ccaggaggca actactccga gaccatgacc attgacctga tgaacaacgc gatcgtggag	14820
cactatctga aagtgggcag gcagaacggg gtcttgaga gcgacatcgg ggtcaagttc	14880
gacaccagga acttccgct ggggctggac cccgtgaccg ggctggttat gcccggggtg	14940
tacaccaacg aggccttcca tcccgacatc atcctgctgc ccggtgcgg ggtggacttc	15000
acttacgcc gcctgagcaa cctcctgggc atccgcaagc ggcagccctt ccaggagggc	15060
ttcaggatca cctacgagga cctggagggg ggcaacatcc ccgcgtcct cgatgtggag	15120
gcctaccagg atagcttgaa ggaaaatgag gcgggacagg aggataccgc ccccgccgc	15180
tccgccgcc ccgagcagg cgaggatgct gctgacaccg cggccgcgga cggggcagag	15240
gccgaccccg ctatggtggt ggaggctccc gagcaggagg aggacatgaa tgacagtgcg	15300
gtgcgcggag acaccttctg caccggggg gaggaagaagc aagcggaggc cgaggccgcg	15360
gccgaggaaa agcaactggc ggcagcagcg gcggcgcgcg cgttgccgc gccggaggct	15420
gagtctgagg ggaccaagcc cgccaaggag cccgtgatta agccctgac cgaagatagc	15480
aagaagcgca gttacaacct gctcaaggac agcaccaaca ccggtaccg cagctggtac	15540
ctggcctaca actacggcga cccgtcgacg ggggtgcgct cctggaccct gctgtgcacg	15600
ccggactgta cctgcggctc ggagcagggt tactggtcgc tgcccacat gatgcaagac	15660
cccgtgacct tccgtccac gcggcaggct agcaacttc cggtggtggg cgccgagctg	15720

ctgcccgctgc actccaagag cttctacaac gaccaggccg tctactccca gctcatccgc	15780
cagttacact ctctgacca cgtgttcaat cgctttcttg agaaccagat tctggcgcg	15840
ccgccgccc ccaccatcac caccgtcagt gaaaacgttc ctgctctcac agatcacggg	15900
acgtaccgc tgcgaacag catcggagga gtccagcgag tgaccgttac tgacgccaga	15960
cgcgcacct gccctacgt ttacaaggcc ttgggcatag tctcgcgcg cgtcctttcc	16020
agccgcactt tttagcaac accaccatca tgtccatcct gatctaccc agcaataact	16080
ccggctgggg actgctgcgc gcgccagca agatgttcgg aggggcgagg aagcgttccg	16140
agcagcacc cgtgcgcgtg cgcgggcaact tccgcgccc ctggggagcg cacaaacgcg	16200
gccgcgcggg gcgcaccacc gtggacgacg ccatcgactc ggtggtggag caggcgcgca	16260
actacaggcc cgcggtctct accgtggacg cgcccatcca gaccgtggtg cggggcgcg	16320
ggcggtagc caagctgaag agcccccga agcgcgtggc ccgccccac cgcgccgac	16380
ccggggccgc cgccaaacgc gccgccgagg cctgtcttcg ccgggccaag cgcacgggc	16440
gccgcgccg catgagggcc gcgcgccgt ttgccgccg catcacgcc gccaccatgg	16500
cccccgtag ccgaagacgc gcggccgccc ccgccgccg gccatcagt gacatggcca	16560
gcaggcgccg gggcaacgtg tactgggtgc gcgactcggg gaccggcacg cgcgtgccc	16620
tgcgttccg cccccgcgg acttgagatg atgtgaaaa acaacactga gtctctgct	16680
gttgtgtga tccagcggc ggcggcgcg gcagcgtcat gtccaagcg aaaatcaaag	16740
aagagatgct ccaggctcgt gcgcgggaga tctatgggcc ccgaagaag gaagagcagg	16800
attcgaagcc ccgaagata aagcgggtca aaaagaaaa gaaagatgat gacgatgcc	16860
atggggaggt ggagtctctg gcgccacgg gccccaggcg ccggtgcag tggaaaggcc	16920
ggcgcgtaaa gcgcgtcctg gcccccgca ccgcggtggt cttcacgcc gccgagcgt	16980
ccaccggac tttaagcgc gtctatgacg aggtgtacgg cgacgaagac ctgctggagc	17040
aggccaacga gcgcttcgga gagtttgctt acgggaagcg tcagcggcg ctggggaagg	17100
aggacctgct ggcgctgcc ctggaccagg gcaacccac cccagtctg aagcccgtga	17160
ccctgcagca ggtgctgcc agcagcgac cctccaggc gaagcgggt ctgaagcgcg	17220
agggcggcga cctggcgccc accgtgcagc tcatggtgcc caagcggcag aggtggagg	17280
atgtgctgga gaaaatgaa gtagacccg gtctgcagcc ggacatcagg gtccgcccc	17340
tcaagcaggt ggcgcgggc ctcggcgtgc agaccgtgga cgtggtcatc cccaccgca	17400
actccccgc gccgccacc actaccgtg cctccacgga catggagaca cagaccgatc	17460
ccgccgagc gcgagccga gccgccgcg cgacctctc ggcgagggtg cagacggacc	17520
cctggctgcc gccggcgatg tcagctccc gcgcgcgtc cgggcgcagg aagtacggcg	17580

ccgccaacgc gtcctgccc gagtacgcct tgcaccttc catcgcgccc acccccggct	17640
accgaggcta tacctaccgc ccgcgaagag ccaagggttc caccgcccgt ccccgccgac	17700
gcgcccgcgc caccaccgcg cgccgcccgc gcagacgcca gcccgcactg gctccagtct	17760
ccgtgaggaa agtggcgcgcg gacggacaca cctgggtgct gcccagggcg cgctaccacc	17820
ccagcatcgt ttaaaagcct gttgtgggttc ttgcagatat ggccctcact tgccgcctcc	17880
gtttcccggt gccgggatac cgaggaggaa gatcgcgccg caggagggtt ctggccggcc	17940
gcggccttag cgagggcagc cgccgcgcgcg accggcgggcg acgcgccacc agccgacgca	18000
tgcgcggcgg ggtgctgccc ctgttaatcc cctgatcgc cgcgcgatc ggccgctg	18060
ccgggatcgc ctccgtggcc ttgcaagcgt cccagaggca ttgacagact tgcaaaactg	18120
caaatatgga aaaaaaac ccaataaaaa agtctagact ctacgctcg cttggtcctg	18180
tgactatfff gtagaatgga agacatcaac ttgctgcgc tggccccgcg tcacggctcg	18240
cgccccgttc tgggacactg gaacgatac ggccaccagca acatgagcgg tggcgcttc	18300
agttggggct ctctgtggag cggcattaaa agtatcgggt ctgccgttaa aaattacggc	18360
tccccggcct ggaacagcag caccggccag atgttgagag acaagttgaa agagcagaac	18420
ttccagcaga aggtggtgga gggcctggcc tccggcatca accgggtggt ggacctggcc	18480
aaccaggccg tgcagaataa gatcaacagc agactggacc cccggccgcg ggtggaggag	18540
gtgccccgg cgctggagac ggtgtcccc gatggcgctg gcgagaagcg cccgcccgc	18600
gataggaag agaccactct ggtcacgcag accgatgagc cgccccgta tgaggaggcc	18660
ctgaagcaag gtctgcccac caccggccc atcgcgcca tggccaccgg ggtggtgggc	18720
cgccacacc ccgccacgt ggacttgcct ccgcccgccg atgtgccga gcagcagaag	18780
gcggcacagc cgggcccgc cgcgaccgc tccgttctt ccgccgttc tctgcgccgc	18840
gcggccagcg gccccgcgg ggggtcgcg aggcacggca actggcagag cagctgaac	18900
agcatcgtgg gtctgggggt gcggtccgtg aagcgccgc gatgctactg aatagcttag	18960
ctaacgtgtt gtatgtgtgt atgcgcccta tctgccgc agaggagctg ctgagtcgcc	19020
gccgttcgcg cgcccaccac caccgcaact ccgccctca agatggcgac cccatcgatg	19080
atgccgcagt ggtcgtacat gcacatctcg ggccaggacg cctcgagta cctgagcccc	19140
gggctggtgc agttcgccg gcgcccgag agctacttca gcctgagtaa caagtttagg	19200
aaccaccagg tggcgccac gcacgatgtg accaccgacc ggtctcagcg cctgacgctg	19260
cggttcattc ccgtggaccg cgaggacacc gcgtactcgt acaaggcgcg gttcacctg	19320
gccgtggcg acaaccgcgt gctggacatg gcctccacct actttgacat ccgcccgggtg	19380
ctggaccggg gtccacttt caagccctac tctggcaccg cctacaactc cctggcccc	19440

aaggcgctc ccaactcctg cgagtgggag caagaggaaa ctcaggcagt tgaagaagca	19500
gcagaagagg aagaagaaga tgctgacggt caagctgagg aagagcaagc agctaccaaa	19560
aagactcatg tatatgctca ggctccctt tctggcgaaa aaattagtaa agatggtctg	19620
caaataggaa cggacgctac agctacagaa caaaaaccta tttatgcaga ccctacattc	19680
cagcccgaac cccaatcgg ggagtcccag tggaatgagg cagatgctac agtcgccggc	19740
ggtagagtgc taaagaaatc tactcccatg aaaccatgct atggttccta tgcaagacct	19800
acaaatgcta atggaggta ggggtgacta acggcaaatg cccagggaca gctagaatct	19860
caggttgaaa tgcaattctt ttcaacttct gaaaacgccc gtaacgaggc taacaacatt	19920
cagcccaaat tgggtgctga tagtgaggat gtgcacatgg agaccccga tacgcacctt	19980
tcttacaagc ccgcaaaaag cgatgacaat tcaaaaatca tgctgggtca gcagtccatg	20040
ccaacagac ctaattacat cggcttcaga gacaacttta tcggcctcat gtattacaat	20100
agcactggca acatgggagt gcttcagggt caggcctctc agttgaatgc agtgggtggac	20160
ttgcaagaca gaaacacaga actgtcctac cagctcttgc ttgattccat gggtagacaga	20220
accagatact ttcccatgtg gaatcaggca gtggacagtt atgaccaga tgttagaatt	20280
attgaaaatc atggaactga agacgagctc cccaactatt gtttcctct gggtagcata	20340
ggggtaaactg acaattacca ggctgttaaa accaacaatg gcaataacgg gggccagggtg	20400
acttgacaa aagatgaac ttttcagat cgcaatgaaa taggggtggg aaacaatttc	20460
gctatggaga tcaacctcag tgccaacctg tggagaaact tcctgtactc caactggcg	20520
ctgtacctac cagacaagct taagtacaac cctccaatg tggacatctc tgacaacccc	20580
aacacctacg attacatgaa caagcgagt gtggccccgg ggctgggtgga ctgtacatc	20640
aacctggcg cgcgctggtc gctggactac atggacaacg tcaaccctt caaccaccac	20700
cgcaatgcgg gcctgcgcta ccgctccatg ctctgggca acggcgcta cgtgccttc	20760
cacatccagg tgccccagaa gttctttgcc atcaagaacc tcctcctct gccgggctcc	20820
tacacctacg agtggaaact caggaaggat gtcaacatgg tcctccagag ctctctgggt	20880
aacgatctca gggtaggacg ggccagcatc aagttcgaga gcatctgcct ctacgccacc	20940
ttcttccca tggccacaa cacggcctcc acgctcgagg ccatgctcag gaacgacacc	21000
aacgaccagt cttcaatga ctacctctcc gccgccaaca tgctctaccc cataccgcc	21060
aacgccacca acgtcccat ctccatcccc tcgcgcaact gggcggcctt ccgcggtgg	21120
gccttcacc gcctcaagac caaggagacc cctccctgg gctcgggatt cgaccctac	21180

tacacctact cgggctccat tccctacctg gacggcacct tctacctcaa ccacactttc	21240
aagaaggtct cggtcacctt cgactcctcg gtcagctggc cgggcaacga ccgtctgctc	21300
acccccaacg agttcgagat caagcgctcg gtcgacgggg agggctacaa cgtggcccag	21360
tgcaacatga ccaaggactg gttcctggtc cagatgctgg ccaactacaa catcggtac	21420
cagggttctt acatcccaga gagctacaag gacaggatgt actccttctt caggaacttc	21480
cagcccatga gccggcaggt ggtggaccag accaagtaca aggactacca ggaggtgggc	21540
atcatccacc agcacaacaa ctggggttc gtgggtacc tcgccccac catgcgcgag	21600
ggacaggcct accccgcaa cttcccctat ccgtcatag gcaagaccgc ggtcgacagc	21660
atcaccaga aaaagtctt ctgcgaccgc accctctggc gcatcccctt ctccagcaac	21720
ttcatgtcca tgggtgcgt ctggacctg ggccagaact tgctctacgc caactccgcc	21780
cacgcccctg acatgacctt cgaggtcgac cccatggacg agcccacct tctctatgtt	21840
ctgttcgaag tctttgactg ggtccgggtc caccagccgc accgcggcgt catcgagacc	21900
gtgtacctgc gtacgccctt ctggccggc aacgccacca cctaaagaag caagccgcag	21960
tcctgcgcgc ctgcatgccg tcgggttcca ccgagcaaga gctcagggcc atcgtcagag	22020
acctgggatg cgggccctat ttttgggca ccttcgacaa gcgttccct ggctttgtct	22080
ccccacacaa gctggcctgc gccatcgta acacggccgg ccgcgagacc gggggcgtgc	22140
actggctggc cttcgcctgg aaccgcgct ccaaaacatg cttctcttt gaccttctg	22200
gcttttcgga ccagcggctc aagcaatct acgagttcga gtacgaggc ttgctgcgtc	22260
gcagcgcct cgcctctcg cccgaccgt gcgtcacct cgaaaagtcc acccagaccg	22320
tgaggggcc cgactcggc gcctgcggtc tcttctgtg catgtttctg cagcctttg	22380
tgactggcc tcagagtcac atggaccgca accccacct gaacttctg acgggggtgc	22440
ccaactccat gctccagagc cccaggtcg agcccacct gcgcgcaac caggagcagc	22500
tctacagctt cctggagcgc cactgcctt acttcgccg ccacagcgca cagatcagga	22560
gggccacctc cttctgccac ttgcaagaga tgcaagaagg gtaataacga tgtacacact	22620
ttttttctca ataatggca tctttttatt tatacaagct ctttgggta ttcatttccc	22680
accaccacc gccgttgtcg ccatctggct ctatttagaa atcgaaaggg ttctgccggg	22740
agtcgccgt cgccacgggc agggacacgt tgcgatactg gtacgggtg cccacttga	22800
actcgggcac caccaggcga ggcagctcg ggaagttttc gctccacagg ctgcgggtca	22860
gcaccagcgc gttcatcagg tcggcgccg agatcttgaa gtcgcagttg gggccgccgc	22920
cctgcgcgcg cgagttgcgg tacaccgggt tgcagcactg gaacaccaac agcgcgggt	22980
gcttcacgct ggccagcacg ctgcggtcgg agatcagctc ggctccagg tcctccgct	23040

tgctcagcgc gaacggggtc atcttgggca cttgccgccc caggaagggc gcgtgccccg	23100
gtttcagatt gcagtcgagc cgcagcggga tcagcaggtg cccgtgcccc gactcggcgt	23160
tggggtacag cgcgcgcatg aaggcctgca tctggcggaa ggccatctgg gccttggcgc	23220
cctccgagaa gaacatgccg caggacttgc ccgagaactg gtttgcgggg cagctggcgt	23280
cgtgcaggca gcagcgcgcg tcggtgttgg cgatctgcac cacgttgcgc ccccaccggt	23340
tcttcacgat cttggccttg gacgattgct ccttcagcgc gcgtgcccc ttctcgttgg	23400
tcacatccat ctcgatcaca tgttccttgi tcacatgct gctgccgtgc agacacttca	23460
gctcgccctc cgtctcggtg cagcgggtgct gccacagcgc gcagcccgtg ggctcgaag	23520
actttaggt caccctccgc aaggactgca ggtaccctg caaaaagcgg cccatcatgg	23580
tcacgaaggt ctgttctg ctgaaggta gctgcagccc gcggtgctcc tcgttcagcc	23640
aggtcttgca cagggccgc agcgcctcca cctggtcggg cagcatcttg aagttcacct	23700
tcagtcatt ctccacgtgg tacttgtcca tcagcgtgcg cgccgcctcc atgcccttct	23760
cccaggccga caccagcggc aggtcacgg ggttcttcac catcacgtg gccgccgct	23820
ccgccgctt ttcgcttcc gcccgctgt tctcttctc ttcctctct tctcgcgc	23880
cgccactcg cagccccgc accacggggt cgtcttctg caggcgtgc accttgcgt	23940
tgcgttgcg cccctgcttg atgcgcacgg gcgggttgc gaagcccacc atcaccagcg	24000
cggcctctt ttgctcgtc tcgctgtcca gaatgacct cggggagggg gggttggta	24060
tcctcagta cgaggcacgc ttcttttct tcttgggggc gttcgccagc tccgcggtg	24120
cggccgctgc cgaggtcgaa ggccgagggc tgggcgtgcg cggcaccagc gcgtcctgcg	24180
agccgtctc gtctctctg gactcgagac ggaggcgggc ccgcttctc gggggcgcgc	24240
ggggcggcgg aggcggcggc ggcgacggag acggggacga gacatcgtcc aggggtgggtg	24300
gacggcgggc cgcgcgcgt ccgcgctcgg gggtgggtc gcgtgggtcc tcttcccgac	24360
tggccatctc cactgctcc ttctctata ggcagaaaga gatcatggag tctctcatgc	24420
gagtcgagaa ggaggaggac agcctaaccg cccctctga gccctccacc accgccgcca	24480
ccaccgcaa tgccgccgcg gacgacgcgc ccaccgagac caccgccagt accacctec	24540
ccagcgacgc acccccgtc gagaatgaag tgctgatcga gcaggaccg ggttttgtga	24600
gcggagagga ggaatgagtg gatgagaagg agaaggagga ggtcgccgc tcagtgccaa	24660
aagaggataa aaagcaagac caggacgacg cagataagga tgagacagca gtcgggcggg	24720
ggaacggaag ccatgatgct gatgacggct acctagacgt gggagacgac gtgctgctta	24780
agcactgca ccgccagtgc gtcacgtct gcgacgcgt gcaggagcgc tgcgaagtgc	24840
ccctggacgt ggcgagggtc agccgcgcct acgagggca cctcttcgcg ccgcacgtgc	24900

ccccaagcg ccgggagaac ggcacctgcg agcccaaccc gcgtctcaac ttctacccgg	24960
tcttcgcggt acccgaggtg ctggccacct accacatctt tttccaaaac tgcaagatcc	25020
ccctctcttg ccgcgccaac cgcacccgcg ccgacaaaac cctgaccctg cggcagggcg	25080
cccacatacc tgatatcgcc tctctggagg aagtgcctaa gatcttcgag ggtctcggtc	25140
gcgacgagaa acgggcggcg aacgtcttgc acggagacag cgaacacgag agtcactcgg	25200
gggtgctggt ggagctcgag ggcgacaacg cgcgcctggc cgtactcaag cgcagcatag	25260
aggtcaccca ctttgcctac ccggcgctca acctgcccc caaggtcatg agtgtggtca	25320
tgggcgagct catcatgcgc cgcgcccagc cctggccgc ggatgcaaac ttgcaagagt	25380
cctccgagga aggcctgccc gcggtcagcg acgagcagct ggcgcgctgg ctggagaccc	25440
gcgacccgcg gcagctggag gagcggcgca agctcatgat ggccgcggtg ctggtcaccg	25500
tggagctcga gtgtctcgag cgcttcttcg cggacccga gatgcagcgc aagctcgagg	25560
agacctgca ctacacctc cgccagggt acgtgcgcca ggctgcaag atctccaacg	25620
tggagctctg caacctggtc tctacctgg gcacctgca cgagaaccgc ctcgggcaga	25680
acgtctgca ctccacctc aaaggggagg cgcgccgga ctacatccgc gactgcgcct	25740
acctcttct ctgctacacc tggcagacgg ccatgggggt ctggcagcag tgcctggagg	25800
agcgcaacct caaggagctg gaaaagctcc tcaagcgac cctcaggac ctctggacgg	25860
gcttcaacga gcgtcgggt gccgccgcgc tggcggacat catctttccc gagcgctgc	25920
tcaagacct gcagcagggc ctgcccact tcaccagcca gagcatgctg cagaacttca	25980
ggactttcat cctggagcgc tcgggcatcc tgccggccac ttgctgcgcg ctgcccagcg	26040
acttctgtgc catcaagtac agggagtgcc cgcgccgct ctggggccac tgctacctct	26100
tccagctggc caactacct gcctaccact cggacctcat ggaagacgtg agcggcgagg	26160
gcctgtctga gtgccaactgc cgctgcaacc tetgcaagcc ccaccgtct ctagtctga	26220
accgcagct gctcagcgag agtcagatta tcggtacctt cgagctgcag ggtccctcgc	26280
ctgacgagaa gtcccggtc ccagggtga aactactcc ggggctgtgg acttccgct	26340
acctacgcaa atttgtacct gaggactacc acgcccacga gatcaggttc tacgaagacc	26400
aatccgccc gcccaaggcg gagctcaccg cctgcgtcat caccagggg cacatcctgg	26460
gccaattgca agccatcaac aaagcccgcc gagagtctt gctgaaaaag ggtcgggggg	26520
tgtacctgga cccccagtc ggcgaggagc taaaccgct acccccgcc cgccccagc	26580
agcgggacct tgcttccag gatggcacc agaaaagaag agcagccgc gccgccgcg	26640

cagccataca tgcttctgga ggaagaggag gaggactggg acagtcaggc agaggaggtt	26700
tcggacgagg agcaggagga gatgatggaa gactgggagg aggacagcag cctagacgag	26760
gaagcttcag aggccgaaga ggtggcagac gcaacaccat cgccctcggt cgcagcccc	26820
tcgccggggc ccctgaaatc ctccgaaccc agcaccagcg ctataacctc cgctcctccg	26880
gcgccggcgc caccgcccc cagacccaac cgtagatggg acaccacagg aaccggggtc	26940
ggtaagtcca agtgcgccg gccgccaccg cagcagcagc agcagcagcg ccagggtac	27000
cgctcgtggc gcgggcacaa gaacgccata gtcgcctgct tgcaagactg cgggggcaac	27060
atctctttcg ccgcccgtt cctgtattc caccacgggg tcgcctttcc ccgaatgtc	27120
ctgcattact accgtcatct ctacagcccc tactgcagcg gcgaccaga ggccggcagcg	27180
gcagccacag cggcgaccac cacctaggaa gatatcctcc gcgggcaaga cagcggcagc	27240
agcggccagg agaccgcgg cagcagcggc gggagcgggt ggcgccactg gcctctcgcc	27300
caacgaaccc ctctgaccc gggagctcag acacaggatc ttccccactt tgtatgcat	27360
cttccaacag agcagaggcc aggagcagga gctgaaaata aaaaacagat ctctgcgctc	27420
cctcacccgc agctgtctgt atcacaaaag cgaagatcag cttcggcgca cgctggagga	27480
cgccggaggca ctcttcagca aatactgcgc gctcactctt aaagactagc tccgcgccct	27540
tctcgaattt aggcgggaga aaactacgtc atcgccggcc gccgccagc ccgccagcc	27600
gagatgagca aagagattcc cagccatac atgtggagct accagccgca gatgggactc	27660
gcggcgggag cggcccagga ctactccacc cgcataaact acatgagcgc gggacccac	27720
atgatctcac aggtcaacgg gatccgcgcc cagcgaaccc aaatactgct ggaacaggcg	27780
gccatcacgg ccacgcccc ccataatctc aacccccgaa attggcccg cgcctcgtg	27840
taccagaaa cccctccgc caccaccgta ctactccgc gtgacgccca ggccgaagtc	27900
cagatgacta actcaggggc gcagctcgcg ggcggctttc gtcacggggc gcggccgctc	27960
cgaccaggta taagacacct gatgatcaga ggccgaggta tccagctcaa cgacgagtcg	28020
gtgagctctt cgctcggtct ccgtccggac ggaactttcc agctcgccgg atccggccgc	28080
tcttcgttca cgccccgcca ggcgtacctg actctgcaga cctcgtctc ggagccccgc	28140
tccggcggca tcggaacct ccagttcgtg gaggagtctg tgccctcggt ctacttcaac	28200
cccttctcgg gacctccgg acgtacccc gaccagtca ttccgaactt tgacgcggtg	28260
aaggactcgg cggacggcta cgactgaatg tcagggtctg aggcagagca gcttcgcctg	28320
agacacctcg agcactgccg ccgccacaag tgcttcgcc gcggttctgg tgagtctgc	28380
tactttcagc taccgagga gcataccgag gggccggcgc acggcgtccg cctgaccacc	28440
cagggcgagg ttacctgtc cctcatccgg gagtttacc tccgtcccct gctagtggag	28500

cgggagcggg gtccctgtgt cctaactatc gctgcaact gccctaacc tggattacat	28560
caagatcttt gctgtcatct ctgtgctgag ttttaataaac gctgagatca gaattctactg	28620
gggctcctgt cgccatcctg tgaacgccac cgtcttcacc caccgccacc aggcccaggc	28680
gaacctcacc tgcgggtctgc atcggagggc caagaagtac ctcacctggt acttcaacgg	28740
cacccccctt gtggtttaca acagcttcga cggggacgga gtctccctga aagaccagct	28800
ctccggtctc agctactcca tccacaagaa caccaccctc caactcttcc ctccctacct	28860
gccgggaacc tacgagtgcg tcaccggccg ctgcaccac ctcaccgcc tgatcgtaaa	28920
ccagagcttt ccgggaacag ataactccct ctccccaga acaggaggtg agctcaggaa	28980
actccccggg gaccaggcg gagacgtacc ttcgaccctt gtggggttag gatTTTTTat	29040
taccgggttg ctggctcttt taatcaaagt ttcttgaga ttgttcttt cttctactg	29100
gtatgaacac ctcaacctcc aataactcta cctttcttc ggaatcaggt gacttctctg	29160
aaatcgggct tgggtgtctg ctactctgt tgatttttt cttatcata ctcagccttc	29220
tgtgcctcag gctcgccgcc tgctgcgcac acatctatat ctactgctgg ttgctcaagt	29280
gcagggtcgc ccaccaaga tgaacaggta catggtccta tcgacctag gcctgtctggc	29340
cctggcggcc tgcagcgccg ccaaaaaaga gattaccttt gaggagcccg cttgcaatgt	29400
aactttcaag cccgagggtg accaatgcac caccctcgtc aaatgcgtta ccaatcatga	29460
gaggctgcgc atcgactaca aaaacaaaac tggccagttt gcggtctata gtgtgtttac	29520
gccccgagac ccctctaact actctgtcac cgtcttcag ggcggacagt ctaagatatt	29580
caattacact ttcccttttt atgagttatg cgatgcggtc atgtacatgt caaaacagta	29640
caacctgttg cctccctctc ccagggcgtg tgtggaaaat actgggtctt actgctgtat	29700
ggctttcgca atcactacgc tcgctcta at ctgcacggtg ctatacata aattcaggca	29760
gaggcgaatc ttatcgatg aaaagaaaat gccttgatcg ctaacaccgg ctttctatct	29820
gcagaatgaa tgcaatcacc tcctactaa tcaccaccac cctccttgcg attgcccatg	29880
ggttgacacg aatcgaagtg ccagtggggt ccaatgtcac catggtgggc cccgccggca	29940
attccacct catgtgggaa aaatttgtcc gcaatcaatg ggttcatttc tgctctaacc	30000
gaatcagtat caagcccaga gccatctgcg atgggcaaaa tctaactctg atcaatgtgc	30060
aatgatgga tctgtgggtac tattacgggc agcggggaga aatcattaat tactggcgac	30120
cccacaagga ctacatgctg catgtagtgc aggcacttcc cactaccacc cccactacca	30180
cctctccac caccaccacc actactacta ctactactac tactactact actaccacta	30240
ccgtgcccc ccataccgcg aaaagcacca tgattagcac aaagccccct cgtgtctact	30300
cccacgccgg cgggcccac gtgtgcgacct cagaaaccac cgagctttgc ttctgccaat	30360

gcactaacgc cagcgctcat gaactgttcg acctggagaa tgaggatgtc cagcagagct	30420
ccgcttgcct gaccaggag gctgtggagc ccgttgcct gaagcagatc ggtgattcaa	30480
taattgactc ttcttctttt gccactcccg aataccctcc cgattctact ttccacatca	30540
cgggtaccaa agaccctaac ctctctttct acctgatgct gctgctctgt atctctgtgg	30600
tctcttccgc gctgatgta ctggggatgt tctgctgcct gatctgccgc agaaagagaa	30660
aagctcgctc tcagggccaa ccactgatgc ccttcccta cccccggat ttgcagata	30720
acaagatatg agctcgctgc tgacactaac cgctttacta gcctgcgctc taacccttgt	30780
cgcttgcgac tcgagattcc acaatgtcac agctgtggca ggagaaaatg ttactttcaa	30840
ctccacggcc gataccagt ggtcgtggag tggtcagggt agctacttaa ctatctgcaa	30900
tagctccact tccccggca tatcccaac caagtaccaa tgcaatgcca gcctgttcac	30960
cctcatcaac gcttccacc tggacaatgg actctatgta ggctatgtac ctttgggtgg	31020
gcaaggaaag acccacgctt acaacctgga agttcgccag ccagaaacca ctaccaagc	31080
tttctccacc accaccacca ccaccacat caccagcagc agcagcagca gcagccacag	31140
cagcagcagc agattattga ctttggtttt ggccagctca tctgccgcta ccagggccat	31200
ctacagctct gtgccccaaa ccactcagat ccaccgcca gaaacgacca cggccaccac	31260
cctacacacc tcacgcgac agatgccgac caacatcacc cccttggctc ttcaaatggg	31320
acttacaagc cccactcaa aaccagtgga tgcggccgag gtctccgcc tcgtcaatga	31380
ctgggcgggg ctgggaatgt ggtgttctgc cataggcatg atggcgctct gcctgttct	31440
gctctggctc atctgtgcc tccaccgcag gcgagccaga cccccatct atagacccat	31500
cattgtcctg aaccccgata atgatgggat ccatagattg gatggcctga aaaacctact	31560
tttttcttt acagtatgat aaattgagac atgcctcgca ttttcttgta catgttcctt	31620
ctccacctt ttctggggtg ttctacgtg gccgtgtgt ctcacctgga ggtagactgc	31680
ctctcacctc tcaactgtca cctgctttac ggatttgtca ccctcactct catctgcagc	31740
ctaatacagc taatcatgc ctccatccag tgcattgatt acatctgtgt gcgcctcgca	31800
tacttcagac accaccgca gtaccgagac aggaacattg cccaacttct aagactgctc	31860
taatcatgca taagactgtg atctgccttc tgatcctctg catcctgccc accctcacct	31920
cctgccagta caccacaaaa tctccgcgca aaagacatgc ctctgccgc ttacccaac	31980
tgtggaatat acccaaatgc tacaacgaaa agagcgagct ctccgaagct tggctgtatg	32040
gggtcatctg tgtcttagtt ttctgcagca ctgtctttgc cctcataatc taccctact	32100

ttgatttggg atggaacgcg atcgatgcca tgaattaccc cacctttccc gcacccgaga	32160
taattccact gcgacaagtt gtacccgttg tegttaatca acgccccca tcccctacgc	32220
ccactgaaat cagctacttt aacctaacag gcggagatga ctgacgcctt agatctagaa	32280
atggacggca tcagtaccga gcagcgtctc ctgagagggc gcaggcaggc ggctgagcaa	32340
gagcgctca atcaggagct ccgagatctc gttaacctgc accagtcaa aagaggcatc	32400
ttttgtctgg taaagcaggc caaagtcacc tacgagaaga ccggcaacag ccaccgctc	32460
agttacaaat tgcccacca gcgccagaag ctggtgctca tggagggtga gaatcccatc	32520
accgtcacc agcactcggg agagaccgag ggggtgtctg actccccctg tcgggggtcca	32580
gaagacctct gcacccctgg aaagacctg tgcggtctca gagatttagt cccctttaac	32640
taatcaaaca ctggaatcaa taaaagaat cacttactta aaatcagaca gcaggctctt	32700
gtccagttta ttcagcaga cctccttccc ctctcccaa ctctggtact ccaaagcct	32760
tctggcggca aacttctcc acacctgaa gggaatgtca gattcttgct cctgtccctc	32820
cgcaccact atcttcatgt tgttcagat gaagcgcacc aaaacgtctg acgagagctt	32880
caacccctg tacccctatg acacggaaag cggccctccc tccgtccctt tctcaccac	32940
tcccttcgtg tctccgatg gattccaaga aagtcccccc ggggtcctgt ctctgaacct	33000
ggccgagccc ctggteactt cccacggcat gctcgccctg aaaatgggaa gtggcctctc	33060
cctggacgac gtggcaacc tcacctctca agatatacacc accgctagcc ctccctcaa	33120
aaaaaccaag accaacctca gcctagaaac ctcatcccc ctaactgtga gcacctcagg	33180
cgcctcacc gtagcagccg ccgtccctt ggcggtggcc ggcacctccc tcaccatgca	33240
atcagaggcc cccctgacag tacaggatgc aaaactcacc ctggccacca aaggccccct	33300
gacctgtct gaaggcaaac tggccttgca aacatggcc ccgtgacgg ccgtgacag	33360
cagcacctc acagtcagt ccacaccacc ccttagcaca agcaatggca gcttgggtat	33420
tgacatgcaa gccccattt acaccacaa tggaaaacta ggacttaact ttggcgctcc	33480
cctgcatgtg gtagacagcc taaatgact gactgtagt actggccaag gtcttacgat	33540
aaacggaaca gcctacaaa ctagagtctc aggtgcctc aactatgaca catcaggaaa	33600
cctagaattg agagctgcag ggggtatgcg agttgatgca aatggtaac ttatccttga	33660
tgtagcttac ccatttgatg caaaaacaa tctcagcctt aggtctggac agggaccctt	33720
gtttgttaac tctgccaca acttgatgt taactacaac agaggcctt accgtttcac	33780
atctggaaat accaaaaagc tagaagttaa tatcaaaaca gccaaagggtc tcatttatga	33840
tgacactgct atagcaatca atgcgggtga tgggtacag ttgactcag gctcagatac	33900
aatccatta aaaactaac ttggattagg actggattat gactccagca gagccataat	33960

tgctaaactg ggaactggcc taagctttga caacacaggt gccatcacag taggcaacaa	34020
aatgatgac aagcttacct tgtggaccac accagacca tcccctaact gtagaatcta	34080
ttcagagaaa gatgctaaat tcacacttgt ttgactaaa tgcggcagtc aggtgttggc	34140
cagcgtttct gttttatctg taaaaggtag ccttgcgccc atcagtgga cagtaactag	34200
tgctcagatt gtcctcagat ttgatgaaaa tggagtctta ctaagcaatt cttcccttga	34260
ccctcaatac tggaactaca gaaaagggtga ctttacagag ggcaactgcat ataccaacgc	34320
agtgggattt atgccaaccc tcacagcata cccaaaaaca cagagccaaa ctgctaaaag	34380
caacatttga agtcaggttt acttgaatgg ggacaaatcc aaacccatga ccctcaccat	34440
tacctcaat ggaactaatg aaacaggaga tgccacagta agcacttact ccatgtcatt	34500
ctcatggaac tggaatggaa gtaattacat taatgaaacg ttccaaacca actccttcac	34560
cttctcctac atcgccaag aataaaaagc atgacgtgtg tgatttgatt caatgtgttt	34620
ctgttttatt ttcaagcaca acaaatcat tcaagtcatt cttccatctt agcttaatag	34680
acacagtagc ttaatagacc cagtagtgca aagccccatt ctagcttata gatcagacag	34740
tgataattaa ccaccaccac caccatacct ttgattcag gaaatcatga tcatcacagg	34800
atcctagtct tcaggccgcc ccctccctcc caagacacag aatacacagt cctctcccc	34860
cgactggctt taaataacac catctggttg gtcacagaca tgttcttagg ggttatattc	34920
cacacggtct cctgcgcgc cagcgctcg tcggtgatgt tgataaaact tcccggcagc	34980
tcgctcaagt tcacgtcgct gtccagcggc tgaacctccg gctgacgcga taactgtgcg	35040
accggtgctt ggacgaacgg aggcgcgcc tacaaggggg tagagtcata atcctcggtc	35100
aggatagggc ggtgatgcag cagcagcgag cgaacatct gctgccgcc cggtccgtc	35160
cggcaggaac acaacacgcc ggtggtctcc tccgcgataa tccgcaccgc ccgcagcatc	35220
agcttctctg ttctccgcgc gcagcacctc acccttatct cgtcaaact ggcgcagtag	35280
gtacagcaca gcaccacgat gttattcatg atcccacagt gcaggcgct gtatccaaag	35340
ctcatgccgg gaaccaccgc cccacgtgg ccatcgtacc acaagcgac gtaaatcaag	35400
tgtegacccc tcatgaacgc gctggacaca aacattactt ccttgggcat gttgtaattc	35460
accacctccc ggtaccagat aaacctctgg ttgaacaggg caccttcac caccatcctg	35520
aaccaagagg ccagaacctg cccaccggt atgcaactga gggaaccgg gttggaacaa	35580
tgacaatgca gactccaagg ctctgaaccg tggatcatcc ggctgctgaa ggcatcgatg	35640
ttggcacaac acagacacac gtgcatgcac ttctcatga ttagcagctc ttccctcgtc	35700
aggatcatat cccaaggaat aacctattct tgaatcaacg taaaaccac acagcaggga	35760
aggcctcgca cataactcac gttgtgcatg gtcagcgtgt tgcatccgg aaacagcggg	35820

tgatcctcca gtatcgaggc gcgggtctcc ttctcacagg gaggtaaagg gtccctgctg	35880
tacggactgc gccgggacga ccgagatcgt gttgagcgta gtgtcatgga aaagggaacg	35940
ccggacgtgg tcatacttct tgaagcagaa ccaggttcgc gcgtggcagg cctccttgcg	36000
tctgcggtct cgcggtctag ctcgctccgt gtgatatgtg tagtacagcc actcccgcag	36060
agcgtcgagg cgcaccctgg ctcccgatc tatgtagact ccgtcttgca ccgcggccct	36120
gataatatcc accaccgtag aataagcaac acccagccaa gcaatacact cgctctgcga	36180
gcggcagaca ggaggagcgg gcagagatgg gagaacctg ataaaaact ttttttaaag	36240
aatatatttc aattcttcga aagtaagatc tatcaagtgg cagcgtctcc ctccactggc	36300
gcggtcaaac tciacggcca aagcacagac aacggcattt ctaagatgtt ccttaatggc	36360
gtccaaaaga cacaccgtc tcaagttgca gtaactatg aatgaaaacc catccggctg	36420
attttccaat atagacgcgc cggcagcgtc caccaaaacc agataatttt ctctctcca	36480
gcggtttacg atctgtctaa gcaaatccct tatatcaagt ccgaccatgc caaaaatctg	36540
ctcaagagcg cctccacct tcatgtacaa gcagcgcac atgattgcaa aaattcaggt	36600
tcttcagaga cctgtataag attcaaatg ggaacattaa caaaaattcc tctgtcgcgc	36660
agatcccttc gcagggcaag ctgaacataa tcagacaggt ccgaacggac cagtgaggcc	36720
aaatccccac caggaaccag atccagagac cctatactga ttatgacgcg catactcggg	36780
gctatgctga ccagcgtagc gccgatgtag gcgtgctgca tgggcggcga gataaaatgc	36840
aaagtgtcgg ttaaaaaatc aggcgaagcc tcgcgcaaaa aagctaacac atcataatca	36900
tgctcatgca ggtagttgca ggtaagctca ggaacaaaa cggaataaca cacgattttc	36960
ctctcaaaca tgacttcgcg gatactgcgt aaaacaaaa attataaata aaaaattaat	37020
taaataactt aaacattgga agcctgtctc acaacaggaa aaaccacttt aatcaacata	37080
agacgggcca cgggcatgcc ggcatagccg taaaaaaatt ggtccccgtg attaacaagt	37140
accacagaca gtcccccgt catgtcgggg gtcatcatgt gagactctgt atacactct	37200
ggattgtgaa catcagacaa acaaagaaat cgagccacgt agcccgagg tataatcacc	37260
cgcaggcgga ggiacagcaa aacgaccccc ataggaggaa tcacaaaatt agtaggagaa	37320
aaaaatacat aaacaccaga aaaaccctgt tgctgaggca aaatagcgcc ctcccgatcc	37380
aaaacaacat aaagcgttc cacaggagca gccataacaa agacccgagt ctaccagta	37440
aaagaaaaa gatctctcaa cgcagcacca gcaccaacac ttcgcagtgt aaaaggccaa	37500
gtgccgagag agtatatata ggaataaaaa gtgacgtaaa cgggcaaagt ccaaaaaacg	37560

cccagaaaaa ccgcacgcga acctacgccc cgaaacgaaa gccaaaaaac actagacact 37620
cccttcgggc gtcaacttcc gctttccac gctacgtcac ttcccccggt caaacaact 37680
acatatcccg aacttccaag tgcacagcc caaaacaccg cctacacctc cccgcccgc 37740
ggcccgcccc cggaccgcc tcccgccccg cgccgcccac ctattatca tattggcttc 37800
aatccaaat aaggtatatt attgatgatg 37830
<210> 14
<211> 15422
<212> DNA
<213> Pan troglodytes
<400> 14
tgcccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 60

cggggagagg cggtttgcgt attgggcgct agatctccca tcacatatac ctgccgttca 120
ctattattta gtgaaatgag atattatgat attttctgaa ttgtgattaa aaaggcaact 180
ttatgccat gcaacagaaa ctataaaaaa tacagagaat gaaaagaaac agatagattt 240
tttagttctt taggcccgtg gtctgcaaat ccttttatga ttttctatca aacaaaagag 300
gaaaatagac cagttgcaat ccaaacgaga gtctaataga atgaggctga aaagtaaact 360
gcgcgggttt gttactgata aagcaggcaa gacctaaaat gtgtaaaggg caaagtgtat 420
actttggcgt cacccttac atattttagg tcttttttta ttgtgcgtaa ctaacttgc 480

atcttcaaac aggagggtcg gaagaagcag accgctaaca cagtacataa aaaaggagac 540
atgaacgatg aacatcaaaa agtttgcaaa acaagcaaca gtattaacct ttactaccgc 600
actgtcggca ggagcgcaa ctcaagcgtt tgcgaaagaa acgaaccaa agccatataa 660
ggaaacatac ggcatttccc atattacacg ccatgatatg ctgcaaatcc ctgaacagca 720
aaaaaatgaa aaatatcaag ttcttgaatt cgattcgtcc acaattaaaa atatctcttc 780
tgcaaaaggc ctggacgttt gggacagctg gccattacaa aacgtgacg gcaactgtcg 840
aaactatcac ggctaccaca tcgtctttgc attagccgga gacctaataa atgcggatga 900

cacatcgatt tacatgttct atcaaaaagt cggcgaaact tctattgaca gctggaaaaa 960
cgctggccgc gtctttaag acagcgacaa attcgatgca aatgattcta tcctaaaaga 1020
ccaaacacaa gaatggtcag gttcagccac attacatct gacggaaaaa tccgtttatt 1080
ctacactgat ttctccggtg aacattacgg caaacaacac ctgacaactg cacaagttaa 1140
cgtatcagca tcagacagct ctttgaacat caacgggtga gaggattata aatcaatctt 1200
tgacggtgac ggaaaaacgt atcaaatgt acagcagttc atcgatgaag gcaactacag 1260

ctcaggcgac aaccatacgc tgagagatcc tcactacgta gaagataaag gccacaaata	1320
cttagtattt gaagcaaaca ctggaactga agatggctac caaggcgaag aatctttatt	1380
taacaaagca tactatggca aaagcacatc attcttccgt caagaaagtc aaaaacttct	1440
gcaaagcgat aaaaaacgca cggctgagtt agcaaacggc gctctcggta tgattgagct	1500
aaacgatgat tacacactga aaaaagtgat gaaaccgtg attgcatcta acacagtaac	1560
agatgaaatt gaacgcgcga acgtctttta aatgaacggc aaatggtacc tgttactga	1620
ctccccgga tcaaaaatga cgattgacgg cattacgtct aacgatattt acatgcttgg	1680
ttatgtttct aattctttta ctggcccata caagccgtg aacaaaactg gccttgtgtt	1740
aaaaatggat ctigatccta acgatgtaac ctttacttac tcacacttcg ctgtacctca	1800
agcgaaagga aacaatgtcg tgattacaag ctatatgaca aacagaggat tctacgcaga	1860
caaacaatca acgtttgcgc caagcttctt gctgaacatc aaaggcaaga aaacatctgt	1920
tgtcaaagac agcatccttg aacaaggaca attaacagtt aacaataat agggataaca	1980
gggtaatgct agaagaccgc agtcttacca gtaaaagaaa aaagatctct caacgcagca	2040
ccagcaccaa cacttcgcag tgtaaaaggc caagtgccga gagagtatat ataggaataa	2100
aaagtgacgt aaacgggcaa agtcacaaaa acgccagaa aaaccgcacg cgaacctacg	2160
ccccgaaacg aaagccaaaa aacactagac actcccttcc ggctgaact tccgctttcc	2220
cacgtacgt cacttgcccc agtcaaaca actacatat ccgaacttcc aagtcgccac	2280
gccccaaaca cgcctacac ctccccgcc gccggccgc ccccaaacc gcctccgcc	2340
ccgcgccccg cccgcgcgcg cccatctcat tatcatattg gcttcaatcc aaaataaggt	2400
atattattga tgatggttta aacggatcct ctagagtcga cctgcaggca tgcaagcttg	2460
agtattctat agtgcacct aaatagcttg gcgtaatcat ggtcatagct gtttctgtg	2520
tgaaattgtt atccgtcac aattccacac aacatacgag ccggaagcat aaagtgtaaa	2580
gcctggggtg cctaagtgt gagtaactc acattaattg cgttcgctc actgcccgct	2640
ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgaaccctt	2700
gcggccgcc gggccgtcga ccaatttca tgtttgacag cttatcatcg aatttctgcc	2760
attcatccgc ttattatcac ttattcaggc gtagcaacca ggctttaaag ggcaccaata	2820
actgccttaa aaaaattacg ccccgccctg ccactcatcg cagtactgtt gtaattcatt	2880
aagcattctg ccgacatgga agccatcaca aacggcatga tgaacctgaa tcgccagcgg	2940
catcagcacc ttgtcgcctt gcgtataata ttgcccctg gtgaaaacgg gggcgaagaa	3000

gttgtccata ttggccacgt ttaaatcaaa actggtgaaa ctcacccagg gattggctga 3060

gacgaaaaac atattctcaa taaacccttt agggaaatag gccaggtttt caccgtaaca 3120

cgccacatct tgcgaatata tgtgtagaaa ctgccggaaa tcgtcgtggt attcactcca 3180

gagcgaatgaa aacgtttcag ttgtctcatg gaaaacgggtg taacaagggt gaacactatc 3240

ccatacacc agctcacctg ctttcattgc catacggaat tccggatgag cattcatcag 3300

gcgggcaaga atgtgaataa aggccggata aaacttgtgc ttatTTTTct ttacggtctt 3360

taaaaaggcc gtaatatcca gctgaacggt ctggttatag gtacattgag caactgactg 3420

aaatgcctca aaatgttctt tacgatgcca ttgggatata tcaacggtgg tatatccagt 3480

gatttttttc tccattttag ctctcttagc tctgaaaat ctcgataact caaaaaatac 3540

gcccggtagt gatcttattt cattatgggtg aaagtTggaa cctcttacgt gccgatcaac 3600

gtctcatttt cgccaaaagt tggcccagggt ctccccggtg tcaacaggga caccaggatt 3660

tatttattct gcgaagtgat ctccgtcac aggtatttat tcgcgataag ctcatggagc 3720

ggcgtaaccg tcgcacagga aggacagaga aagcgcggat ctgggaagtg acggacagaa 3780

cggtcaggac ctggattggg gaggcgggtg ccgccgctgc tgctgacggt gtgacgttct 3840

ctgttccggt cacaccacat acgttccgcc attcctatgc gatgcacatg ctgtatgccg 3900

gtataccgct gaaagtcttg caaagcctga tgggacataa gtccatcagt tcaacggaag 3960

tctacacgaa ggtttttgcg ctggatgtgg ctgcccgga ccgggtgcag ttTgcgatgc 4020

cggagtctga tcgggttgcg atgctgaaac aattatcctg agaataaatg ccttggcctt 4080

tatatggaaa tgtggaactg agtggatatg ctgtttttgt ctgttaaaca gagaagctgg 4140

ctgttatcca ctgagaagcg aacgaaacag tcgggaaaat ctccattat cgtagagatc 4200

cgcattatta atctcaggag cctgtgtagc gtttatagga agtagtgttc tgtcatgatg 4260

cctgcaagcg gtaacgaaaa cgatttgaat atgccttcag gaacaataga aatcttcgtg 4320

cgggtgttacg ttgaagtTga gcggattatg tcagcaatgg acagaacaac ctaatgaaca 4380

cagaaccatg atgtggtctg tccttttaca gccagtagtg ctcgccgcag tcgagcgaca 4440

gggcgaagcc ctcgagtgag cgaggaagca ccagggaaca gcacttatat attctgctta 4500

cacacgatgc ctgaaaaaac ttcccttggg gttatccact tatccacggg gatattttta 4560

taattatttt ttttatagtt tttagatctt cttttttaga gcgccttgta ggcctttatc 4620

catgctggtt ctagagaagg tgttgtgaca aattgcctt tcagtgtgac aaatcacctt 4680

caaatgacag tctgtctgtg gacaaattgc ccttaaccct gtgacaaatt gccctcagaa 4740

gaagctgttt ttTcacaaag ttatccctgc ttattgactc ttttttattt agtgtgacaa 4800

tctaaaaact tgtcacactt cacatggatc tgtcatggcg gaaacagcgg ttatcaatca 4860

caagaaacgt aaaaatagcc cggaatcgt ccagtcaaac gacctcactg aggcggcata	4920
tagtctctcc cgggatcaaa aacgtatgct gtatctgttc gttgaccaga tcagaaaatc	4980
tgatggcacc ctacaggaac atgacggtat ctgcgagatc catgttgcta aatatgctga	5040
aatatcggga ttgacctctg cggaagccag taaggatata cggcaggcat tgaagagttt	5100
cgcggggaag gaagtggttt tttatcgccc tgaagaggat gccggcgatg aaaaaggcta	5160
tgaatctttt ccttggttta tcaaactgac gcacagtcca tccagagggc ttacagtgt	5220
acatatcaac ccatatctca ttcccttctt tatcggttta cagaaccggt ttacgcagtt	5280
tcggcttagt gaaacaaaag aaatcaccaa tccgtatgcc atgcgtttat acgaatccct	5340
gtgtcagtat cgtaagccgg atggctcagg catcgtctct ctgaaaatcg actggatcat	5400
agagcgttac cagctgcctc aaagttacca gcgtatgcct gacttcgcc gccgcttctt	5460
gcaggctctgt gttaatgaga tcaacagcag aactccaatg cgctctcat acattgagaa	5520
aaagaaaggc cgccagacga ctcatatcgt attttcttc cgcatatca ctccatgac	5580
gacaggatag tctgagggtt atctgtcaca gatttgaggg tggttcgtca catttgttct	5640
gacctactga gggtaatgtg tcacagtttt gctgtttctt tcagcctgca tggattttct	5700
catacttttt gaactgtaat ttttaaggaa gccaaatttg agggcagttt gtcacagttg	5760
atctcttctt ctttcccttc gtcatgtgac ctgatatcgg gggttagttc gtcacattg	5820
atgaggggtg attatcacag tttattactc tgaattggct atccgcgtgt gtacctctac	5880
ctggagtttt tcccacgggt gatatctctt ctgctgctga gcgtaagagc tatctgacag	5940
aacagtcttt ctttgccttc tcgccagttc gctcgtatg ctcggttaca cggctcggc	6000
gagcgttagt gataataagt gactgaggta tgtgtcttc ttatctctt ttgtagtgtt	6060
gctcttattt taaacaactt tgcggttttt tgatgacttt gcgattttgt tgttgctttg	6120
cagtaaatg caagatttaa taaaaaacg caaagcaatg attaaaggat gttcagaatg	6180
aaactcatgg aaacacttaa ccagtgcata aacgctggtc atgaaatgac gaaggctatc	6240
gccattgcac agtttaatga tgacagcccg gaagcgagga aaataaccg gcgctggaga	6300
ataggatgaag cagcggattt agttgggtt tcttctcagg ctatcagaga tgccgagaaa	6360
gcaggcgac taccgaccc ggatatgaa attcgaggac gggttagca acgtgttggt	6420
tatacaattg aacaattaa tcatatgcgt gatgtgttg gtacgcgatt gcgacgtgt	6480
gaagacgtat ttccaccggt gatcgggtt gctgccata aagtggtgt ttacaaaacc	6540
tcagtttctg ttcatcttgc tcaggatctg gctctgaagg ggctacgtgt ttgtctctg	6600
gaaggtaacg acccccagg aacagcctca atgtatcacg gatgggtacc agatcttcat	6660
attcatgcag aagacactct cctgccttc tatcttggg aaaaggacga gtgcattat	6720

gcaataaagc ccacttgctg gccggggcctt gacattattc cttcctgtct ggctctgcac 6780

cgatttgaag ctgagttaat gggcaaattt gatgaaggta aactgcccac cgatccacac 6840

ctgatgctcc gactggccat tgaaactgtt gctcatgact atgatgtcat agttattgac 6900

agcgcgccta acctgggtat cggcacgatt aatgtcgtat gtgctgctga tgtgctgatt 6960

gttcccacgc ctgctgagtt gtttgactac acctccgcac tgcagttttt cgatatgctt 7020

cgtgatctgc tcaagaacgt tgatcttaaa gggttcgagc ctgatgtacg tattttgctt 7080

accaaataca gcaatagtaa tggctctcag tccccgtgga tggaggagca aattcgggat 7140

gcctggggaa gcatggttct aaaaaatgtt gtacgtgaaa cggatgaagt tggtaaaggt 7200

cagatccgga tgagaactgt ttttgaacag gccattgac aacgctcttc aactggtgcc 7260

tggagaaatg ctctttctat ttgggaacct gtctgcaatg aaattttcga tcgtctgatt 7320

aaaccacgct gggagattag ataataagc gtgcgcctgt tattccaaaa catacgtca 7380

atactcaacc ggttgaagat acttcgttat cgacaccagc tgccccgatg gtggattcgt 7440

taattgcgcg cgtaggagta atggctcgcg gtaatgccat tactttgcct gtatgtggtc 7500

gggatgtgaa gtttactctt gaagtgtctc ggggtgatag tgttgagaag acctctcggg 7560

tatggtcagg taatgaacgt gaccaggagc tgcttactga ggacgcactg gatgatctca 7620

tcccttcttt tctactgact ggtcaacaga caccggcggt cggtcgaaga gtatctggtg 7680

tcatagaaat tgccgatggg agtcgccgtc gtaaagctgc tgcacttacc gaaagtatt 7740

atcgtgttct ggttggcgag ctggatgatg agcagatggc tgcattatcc agattgggta 7800

acgattatcg cccaacaagt gcttatgaac gtggtcagcg ttatgcaagc cgattgcaga 7860

atgaatttgc tggaatatt tctgcgctgg ctgatgcgga aaatatttca cgtaagatta 7920

ttaccgctg tatcaacacc gccaaattgc cttaatcagt tgttgccttt ttttctcacc 7980

ccgtgaact atctgccgg tcaggtgatg cacttcaaaa agcctttaca gataaagagg 8040

aattacttaa gcagcaggca tctaaccctc atgagcagaa aaaagctggg gtgatatttg 8100

aagctgaaga agttatcact cttttaactt ctgtgcttaa aacgtcatct gcatcaagaa 8160

ctagtttaag ctcacgacat cagtttgctc ctggagcgac agtattgtat aaggcgata 8220

aaatggtgct taacctggac aggtctcgtg ttccaactga gtgtatagag aaaattgagg 8280

ccattcttaa ggaacttgaa aagccagcac cctgatgcga ccacgtttta gtctacgttt 8340

atctgtcttt acttaatgtc ctttggtaca ggccagaaag cataactggc ctgaatattc 8400

tctctgggcc cactgttcca cttgtatcgt cggtctgata atcagactgg gaccacggtc 8460

ccactcgtat cgtcggctctg attattagtc tgggaccacg gtccactcg tatcgtcgg	8520
ctgattatta gtctgggacc acggccccac tcgtatcgtc ggtctgataa tcagactggg	8580
accacggtec cactcgtatc gtcggctctga ttattagtct gggaccatgg tccactcgt	8640
atcgtcggtc tgattattag tctgggacca cggccccact cgtatcgtcg gtctgattat	8700
tagtctggaa ccacggcccc actcgtatcg tcggctctgat tattagtctg ggaccacgg	8760
cccactcgta tcgtcggctc gattattagt ctgggaccac gatccactc gtgttgcgg	8820
tctgattatc ggctcgggac cacggcccc cttgtattgt cgatcagact atcagcgtga	8880
gactacgatt ccatcaatgc ctgtcaaggg caagtattga catgctcgtc taacctgtag	8940
aacggagtaa cctcggctgtg cggttgtatg cctgctgtgg attgctcgtc tgtcctgctt	9000
atccacaaca ttttgcgcac ggttatgtgg acaaaatacc tggttacca gcccggtccg	9060
gcacgttaac cgggctgcat ccgatgcaag tgtgtcgtc tcgacgagct cgcgagctcg	9120
gacatgaggt tgccccgtat tcagtgtcgc tgatttgtat tgtctgaagt tgtttttacg	9180
ttaagttgat gcagatcaat taatacgata cctgcgtcat aattgattat ttgacgtgg	9240
ttgatggcct ccacgcacgt tgtgatatgt agatgataat cattatcact ttacgggtcc	9300
tttccggtga tccgacaggt tacggggcgg cgacctcgcg ggttttcgt atttatgaaa	9360
attttccggt ttaaggcgtt tccgttcttc ttcgtcataa cttaatgttt ttatttaaaa	9420
taccctctga aaagaaagga aacgacaggt gctgaaagcg agctttttgg cctctgtcgt	9480
ttcctttctc tgtttttgtc cgtggaatga acaatggaag tccgagctca tcgctaataa	9540
cttcgtatag catacattat acgaagtat attcgatgcg gccgaagggt gtccgctca	9600
gcgggtgttg gcgggtgtcg gggctggcct aactatgcgg catcagagca gattgtactg	9660
agagtgcacc atatcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcac	9720
aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcgtt gcgggcctct	9780
tcgtatttac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg	9840
ccagggtttt ccagtcacg acgttgtaaa acgacggcca gtgaattgta atacgactca	9900
ctatagggcg aattcgagct cggatcccg ggatcctcgt ttaaacatc atcaataata	9960
taccttattt tggattgaag ccaatatgat aatgagatgg gcggcgcggg gcggggcgcg	10020
gggcgggagg cgggtttggg ggcgggcccg cggcgggggc ggtgtggcgg aagtggactt	10080
tgtaagtgtg gcggatgtga cttgctagt cgggcgcggg taaaagtac gttttccgtg	10140
cgcgacaacg cccccgggaa gtgacatttt tcccgcggtt tttaccgat gttgtagtga	10200
atttggcggt aaccaagtaa gatttggcca ttttcgctgg aaaactgaaa cggggaagt	10260
aatctgatt aattttgcgt tagtcatacc gcgtaatat tgtctagggc cgagggactt	10320

tggccgatta cgiggaggac tcgcccaggt gttttttgag gtgaatttcc gcgtttccggg	10380
tcaaagtctg cgttttatta ttataggata tccattgca tacgttgtat ccatacata	10440
atatgtacat ttatattggc tcatgtccaa cattaccgcc atgttgacat tgattattga	10500
ctagttatta atagtaatca attacgggggt cattagtica tagcccatat atggagtcc	10560
gcgttacata acttacggta aatggcccg cttggctgacc gccaacgac ccccgcccat	10620
tgacgtcaat aatgacgtat gtteccatag taacgccaat agggactttc cattgacgtc	10680
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	10740
caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgccagtc	10800
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta	10860
ccatggtgat gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg	10920
gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac caaatcaac	10980
gggactttcc aaaatgtcgt aacaactccg ccccatgac gcaaatgggc ggtaggcgtg	11040
tacggtgga ggtctatata agcagagctc tccctatcag tgatagagat ctccctatca	11100
tgatagaga tcgtcgacga gctcgttttag tgaaccgtca gatcgctgg agacgccatc	11160
cacgctgttt tgacctcat agaagacacc gggaccgatc cagcctccgc ggccgggaac	11220
ggtgcattgg aacgcggatt ccccggtcca agagttagat cttccgttta tctaggtacc	11280
gggccccccc tcgaggtcga cggtatcgat aagcttcacg ctgccgcaag cactcagggc	11340
gcaagggtg ctaaaggaag cggaacacgt agaaagccag tccgcagaaa cgggtgctgac	11400
cccggatgaa tgtcagctac tgggctatct ggacaaggga aaacgcaagc gcaaagagaa	11460
agcaggtagc ttgcagtggg cttacatggc gatagctaga ctgggcggtt ttatggacag	11520
caagcgaacc ggaattgcca gctggggcgc cctctggtaa ggttgggaag ccctgcaaag	11580
taaactggat ggctttcttg ccgccaagga tctgatggcg caggggatca agatctaacc	11640
aggagctatt taatggcaac agttaaccag ctggtacgca aaccacgtgc tcgcaaagtt	11700
gcgaaaagca acgtgcctgc gctggaagca tgcccgaata aacgtggcgt atgtactcgt	11760
gtatatacta ccactcctaa aaaaccgaac tccgcgtgc gtaaagtatg ccgtgttcgt	11820
ctgactaacg gtttcgaagt gacttcctac atcgggtggtg aaggtcaca cctgcaggag	11880
cactccgtga tctgatccg tggcggtcgt gttaaagacc tcccgggtgt tcgttaccac	11940
accgtacgtg gtgcgcttga ctgctccggc gttaaagacc gtaagcaggc tcgttccaag	12000
tatggcgtga agcgtcctaa ggcttaatgg tagatctgat caagagacag gatgacggtc	12060
gtttcgcgtg cttgaacaag atggattgca cgcaggttct ccggccgctt ggggtggagag	12120
gctattcggc tatgactggg cacaacagac aatcggctgc tctgatgccg ccgtgttcgg	12180

gctgtcagcg caggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa 12240

tgaactgcag gacgaggcag cgcggtatc gtggctggcc acgacggcg ttccttgccg 12300

agctgtgctc gacgttgtca ctgaagcggg aagggaactgg ctgctattgg gcgaagtgcc 12360

ggggcaggat ctctgtcat ctacattgc tctgcccag aaagtatcca tcatggctga 12420

tgcaatgcgg cggtgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 12480

acatcgcatc gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct 12540

ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat 12600

gcccgcggc gaggatctcg tcgtgacca tggcgatgcc tgcttgccga atatcatggt 12660

ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 12720

tcaggacata gcgttgcta cccgtgatat tgctgaagag cttggcggcg aatgggctga 12780

ccgttctctc gtgttttacg gtatcgccgc tcccattcg cagcgcatcg ctttctatcg 12840

ctttcttgac gagtcttctc gagcgggact ctggggttcg aaatgaccga ccaagcgacg 12900

cccaacctgc catcacgaga ttctgattcc accgccct tctatgaaag gttgggcttc 12960

ggaatcgttt tccgggacgc cggtcggatg atcctccagc gcggggatct catgctggag 13020

ttcttcgcc accccgggct cgatccctc ggggggaatc agaattcagt cgacagcggc 13080

cgcatctgc tgtccttct agttgccagc catctgtgt ttgcccctcc cccgtgcctt 13140

ccttgacctt ggaagggtgcc actccactg tcttttcta ataaatgag gaaattgcat 13200

cgattgtctc gattagggtg cattctattc tggggggtgg ggtggggcag gacagcaagg 13260

gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggccgatc 13320

agcgatcgct gaggtgggtg agtgggcgtg gcctggggtg gtcataaaaa tatataagtt 13380

gggggtctta gggtctcttt atttgtgtt cagagaccgc cggagccatg agcgggagca 13440

gcagcagcag cagtagcagc agcgcttgg atggcagcat cgtgagccct tatattgacga 13500

cgcggtatgcc ccaactgggcc ggggtgcgtc agaattgat gggctccagc atcgacggcc 13560

gacctgtctt gcccgaat tccgccagc tgacctatgc gacctcgcg gggacgccgt 13620

tggacccac cgccgccgcc gccgccaccg cagccgcctc ggccgtgcgc agcctggcca 13680

cggactttgc attcctggga ccaactggcg caggggctac ttctcgggcc gctgctgccg 13740

ccgttcgca tgacaagctg accgccctgc tggcgagtt ggtatgcctt actcggaac 13800

tgggtgacct ttctcagcag gtcatggccc tgcgccagca ggtctctcc ctgcaagctg 13860

gcgggaatgc ttctccaca aatgccgttt aaggcgccgc ctagggataa cagggtata 13920

cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 13980
 cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg 14040
 tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc 14100
 tggatgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 14160
 atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga 14220
 gcacttttaa agttctgcta tgtggcgagg tattatcccg tattgacgcc gggcaagagc 14280
 aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag 14340

 aaaagcatct tacggatggc atgacagtaa gagaattatg cagtctgcc ataaccatga 14400
 gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg 14460
 cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga 14520
 atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt 14580
 tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 14640
 ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggg 14700
 ttattgctga taaatctgga gccggtgagc gtgggtctcg cggatcatt gcagcactgg 14760

 ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 14820
 tggatgaacg aaatagacag atcgtgaga taggtgcctc actgattaag catttgtaac 14880
 tgtcagacca agtttactca tatatacttt agattgattt aaaatacgtat tatatgtatt 14940
 agtcacgct attaccatgg ttaatgcgcc gctacagggc gcgtccattc gccattcagg 15000
 ctgcgcaact gttgggaagg gcgatcggg cgggcctctt cgctattacg ccagctggcg 15060
 aaagggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc ccagtcacga 15120
 cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga attgggccct 15180

 ctagatgcat gctcgagcgg ccgccagtgt gatggatgc tgcagaattc cagcacactg 15240
 gcggccgtta ctagtggatc cgagctcgg accaagcttg gcgtaatcat ggtcatagct 15300
 gtttcctgtg tgaattgtt atccgctcac aattccacac aacatacag ccggaagcat 15360
 aaagtgtaaa gcctgggggtg cctaatagag gagctaactc acattaattg cgttgcgctc 15420
 ac 15422

 <210> 15
 <211> 39535
 <212> DNA
 <213> Pan troglodytes
 <220><221> misc_feature
 <222> (24391)..(24393)

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

<400> 15

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcggggcgg gaggcgggtt tgggggcggg ccggcgggcg gggcgggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttccgc ggtttttacc	240
ggatgttgtg gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcggc aggtgttttt tgaggatgaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt	480
gtatccatat cataatatgt acatttataat tggctcatgt ccaacattac cgccatgttg	540
acattgatta ttgactagt attaatagta atcaattacg gggtcattag ttcatagccc	600
atataatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccga	660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caataggagc	720
ttccattga cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca	780
agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggccgcctg	840
gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca tctacgtatt	900
agtcacgct attaccatgg tgatgcgggt ttggcagtac atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg	1020
gcacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat	1080
gggcggtagg cgtgtacgtt gggagggtcta tataagcaga gctctcccta tcagtatag	1140
agatctccct atcagtata gagatcgtcg acgagctcgt ttagtgaacc gtcagatcg	1200
ctggagacgc catccacgt gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggccgg gaacggtgca ttggaacgcg gattccccgt gccaaagatg agatcttccg	1320
tttatctagg taccgggccc cccctcgagg tcgacgtat cgataagctt cacgctgccg	1380
caagcactca gggcgcaagg gctgctaaag gaagcggaac acgtagaaag ccagtccgca	1440
gaaacggtgc tgaccccgga tgaatgtcag ctactgggct atctggacaa gggaaaacgc	1500
aagcgcaaag agaaagcagg tagcttgacg tgggcttaca tggcgatagc tagactgggc	1560
ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgcctctg gtaaggttgg	1620

gaagccctgc aaagtaaact ggatggcttt cttgccgcca aggatctgat ggcgagggg	1680
atcaagatct aaccaggagc tatttaatgg caacagttaa ccagctggta cgcaaaccac	1740
gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgctgga agcatgccg caaaaacgtg	1800
gcgtatgtac tcgtgtatat actaccactc ctaaaaaacc gaactccgcg ctgcgtaaag	1860
tatgccgtgt tcgtctgact aacggtttcg aagtgacttc ctacatcggg ggtgaaggtc	1920
acaacctgca ggagcactcc gtgatcctga tccgtggcgg tcgtgttaaa gacctccgg	1980
gtgttcgtta ccacaccgta cgtgggtgcg ttgactgctc cggcggttaa gaccgtaagc	2040
aggctcgttc caagtatggc gtgaagcgtc ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggicgtttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat	2220
gccgcctgt tccgctgtc agcgcagggg cggccggttc tttttgtcaa gaccgacctg	2280
tccggtgccc tgaatgaact gcaggacgag gcagcgcggc tatcgtggct ggccacgacg	2340
ggcgttccct gcgcagctgt gctcgactgt gtcactgaag cgggaaggga ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac ctgccattc	2520
gaccaccaag cgaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc	2580
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgccga cggcgaggat ctgctcgtga cccatggcga tgctctgtg	2700
ccgaatatca tgggtgaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccctg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc	2880
atgccttct atgccttct tgacgagttc ttctgagcgg gactctgggg ttcgaaatga	2940
ccgaccaagc gacgccaac ctgccatcac gagatttca ttccaccgc gccttctatg	3000
aaaggttggg cttcggaatc gttttccggg acgccgctg gatgatectc cagcgcgggg	3060
atctcatgct ggagtcttc gccaccccg ggtcgtatcc cctcggggg aatcagaatt	3120
cagtcgacag cggcccgat ctgctgtgcc ttctagtgc cagccatctg ttgtttgcc	3180
ctccccgtg ctttcttga ccttgaagg tgcactccc actgtcttt ctaataaaa	3240
tgaggaaatt gcatcgatt gtctgagtag gtgtcattct attctgggg gtgggtggg	3300
gcaggacagc aagggggagg attgggaaga caatagcagg catgctggg atgcgtggg	3360

ctctatggcc gatcagcgat cgctgagggtg ggtgagtggtg cgtggcctgg ggtggatcatg 3420
 aaaatatata agttgggggt cttaggggtct ctttatttgt gttgcagaga ccgccggagc 3480
 catgagcggg agcagcagca gcagcagtag cagcagcgcc ttggatggca gcatcgtgag 3540
 cccttatttg acgacgcgga tgccccactg ggccgggggtg cgtcagaatg tgatgggctc 3600
 cagcatcgac ggccgacccg tcctgcccgc aaattccgcc acgtgacct atgcgacct 3660
 cgccggggacg ccgttggacg ccaccgccgc cgccgccgcc accgcagccg cctcgccgt 3720
 gcgcagcctg gccacggact ttgcattcct gggaccactg gcgacagggg ctacttctcg 3780

ggccgctgct gccgccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc 3840
 gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcc agcaggtctc 3900
 ctccctgcaa gctggcggga atgcttctcc cacaaatgcc gtttaagata aataaaacca 3960
 gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttatct cataattttc 4020
 cgcgcgcat aggccctaga ccagcggtct cggtcgttga gggcgcggtg tatcttctcc 4080
 aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccggggg 4140
 tggaggtagc accactgcag agcttcatgc tccgggggtg tgtttagat gatccagtcg 4200

tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg 4260
 gggaggccct tgggtgaagt gtttcaaaaa cggttaagtt gggaagggtg cattcgggga 4320
 gagatgatgt gcatcttga ctgtattttt agattggcga tgtttccgcc cagatccctt 4380
 ctgggattca tgttgtgcag gaccaccagt acagtgtatc cggtgcactt ggggaatttg 4440
 tcatgcagct tagagggaag agcgtggaag aacttggaga cgcccttgtg gcctcccaga 4500
 ttttccatgc attcgtccat gatgatggca atgggcccgc gggaggcagc ttgggcaaag 4560
 atatttctgg ggtcgctgac gtcgtagtgt tgttccaggg tgaggtcgtc ataggccatt 4620

tttacaagc gccggcggag ggtgcccgc tgggggatga tggtcacctc tggccccggg 4680
 gcgtagttgc cctgcagat ctgcatttcc caggccttaa tctcggaggg gggaatcata 4740
 tccacctgcg gggcgtatga gaaaacggtt tccggagccg gggagattaa ctgggatgag 4800
 agcaggtttc taagcagctg tgattttcca caaccgttg gccataaat aacacctata 4860
 accggttga gctggtagtt tagagagctg cagctgccgt cgtcccggag gaggggggccc 4920
 acctcgttga gcatgtccct gacgcgcatg ttctccccga ccagatccgc cagaaggcgc 4980
 tcgccccca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttgaggccg 5040

tccgccgtgg gcatgttttt cagggtcttg ctacagcagct ccaggcggtc ccagagctcg 5100
 gtgacgtgct ctacggcatc tctatccagc atatctcttc gtttcgcggg ttggggcgac 5160
 tttcgctgta gggcaccaag cgggtggtcgt ccagcggggc cagagtcatg tccttccatg 5220

ggcgcaggggt cctcgtcagg gtggctctggg tcacggtgaa ggggtgcgct ccgggctgag 5280
cgcttgccaa ggtgcgcttg aggtcgttgc tgcctgtgct gaagcgtgc cggctcttcgc 5340
cctgcgcgtc ggccaggtag catttgacca tgggtgtcata gtccagcccc tccgcggcgt 5400
gtcccttggc gcgcagcttg cccttggagg tggcgccgca cgaggggcag agcaggctct 5460

tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgccgc 5520
agaccccgca cacggtctcg cactccacca gccaggtagag ctcggggcgc gccgggtcaa 5580
aaaccaggtt tccccatgc tttttgatgc gtttcttacc tcgggtctcc atgaggtggt 5640
gtccccgtc ggtgacgaag aggtcgtccg tgtctccgta gaccgacttg aggggtcttt 5700
tctccagggg ggtccctcgg tcttctcgt agaggaaactc ggaccactct gagacgaagg 5760
cccgcgtcca ggccaggacg aaggaggcta tgtgggaggg gtagcggctc ttgtccacta 5820
gggggtccac ctctccaag gtgtgaagac acatgtcgcc ttctcggcg tccaggaagg 5880

tgattggctt gtaggtgtag gccacgtgac cgggggttcc tgacgggggg gtataaaagg 5940
gggtgggggc gcgtcgtcg tcaactctctt ccgcctcgt gtctgcgagg gccagctgct 6000
ggggtgagta ttccctctcg aaggcgggca tgacctcgc gctgaggttg tcagtttcca 6060
aaaacagga ggatttgatg ttacactgtc ccgaggtagt acctttgagg gtaccgcgt 6120
ccatctggtc agaaaacacg atctttttat tgtccagctt ggtggcgaac gaccgtaga 6180
gggcgttggg gagcagcttg gcgatggagc gcagggtctg gttcttgtcc ctgtcggcgc 6240
gtccttggc cgcatgttg agctgcacgt actcgcgcgc gacgcagcgc cactcgggga 6300

agacggtggt gcgtcgtcg ggcaccaggc gcacgcgcca gccgcggttg tgcagggtga 6360
ccagggtccac gctggtggcg acctcgccgc gcaggcgtc gttggtccag cagagacggc 6420
cgcccttgcg cgagcagaag gggggcaggg ggtcgagctg ggtctcgtcc ggggggtccg 6480
cgtccacggt gaaaaccccg gggcgcaggc gcgcgtcgaa gtagtctatc ttgcaacctt 6540
gcatgtccag cgcttctgc cagtgcggg cggcgagcgc gcgtcgtag gggttgagcg 6600
gcgggcccc gggcatggg tgggtgagtg cggaggcgta catgccgag atgtcataga 6660
cgtagagggg ctcccgagg accccgatgt aggtggggta gcagcgccg ccgcggatgc 6720

tggcgcgcac gtagtcatac agctcgtgcg agggggcgag gaggtcggg cccaggttgg 6780
tgcgggcggg gcgtccgtg cggaagacga tctgcctgaa gatggcatgc gatttgaag 6840
agatggtggg gcgttggaag acgttgaagc tggcgtcctg caggccgacg gcgtcgcga 6900
cgaaggaggc gtaggagtcg cgcagcttgt gtaccagctc ggcggtgacc tgcacgtcga 6960
gcgcgcagta gtcgagggtc tcgcggatga tgtcatattt agcctgcccc ttctttttcc 7020
acagctcgcg gttgaggaca aactcttcgc ggtctttcca gtactcttgg atcgggaaac 7080

cgtcgggttc cgaacggtaa gagcctagca ttagaactg gttgacggcc tggtaggcgc 7140

agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg 7200

tcaggcgaa ggtgtccctg accatgactt tgaggtagtg gtgcttgaag tcggagtcgt 7260

cgcagccgcc ccgtcccag agcgagaagt cggtcgcctt cttggagcgg gggttgggca 7320

gagcgaaggt gacatcgttg aagaggattt tgcccgcgcg gggcatgaag ttgcgggtga 7380

tgcggaaggg ccccgccact tcagagcggg tgttgatgac ctggcgggcg agcacgatct 7440

cgtcgaagcc gttgatgttg tggccacga ttagagttc caggaagcgg ggccggccct 7500

ttacggtggg cagcttcttt agctcttcgt aggtgagctc ctggggcgag gcgaggccgt 7560

gctcgccag ggcccagtc gcgaggtgcg ggttgtctct gaggaaggac tcccagaggt 7620

cgcgggccag gaggtctgc aggcgggtccc tgaaggtcct gaactggcgg cccacggcca 7680

ttttttcggg ggtgatcgag tagaaggtga gggggtcttg ctgccagcgg tcccagtcga 7740

gctgcagggc gaggtcgcgc gcggcggtga ccaggcgtc gtcgccccg aatttcatga 7800

ccagcatgaa gggcacgagc tgctttccga agcccccat ccaagttag gtctctacat 7860

cgtaggtgac aaagaggcgc tccgtgcgag gatgcgagcc gatcggaag aactggatct 7920

ccgcccacca gtggaggag tggctgttga tgtggtggaa gtagaagtc cgtcgccggg 7980

ccgaacactc gtgctggtt ttgtaaaagc gagcgcagta ctggcagcgc tgcacgggct 8040

gtacctcatg cagagatgc acctttcgcc cgcgcacgag gaagccgagg ggaaatctga 8100

gccccccgcc tggtcgcgg catgctggt gctcttctac tttggatgcg tgtccgtctc 8160

cgtctggctc ctgaggggt gttacggtgg agcggaccac cacgccgcgc gagccgcagg 8220

tccagatcgc ggcgcgccgc ggtcggagtt tgatgacgac atcgcgcagc tgggagctgt 8280

ccatggtctg gagtccccc ggccggcgga ggtcagccgg gatttcttgc aggttcacct 8340

cgcagagtcg ggccagggcg cggggcaggt ctaggtggta cctgatctct aggggcgtgt 8400

tggtagggcg gtcgatggct tgcaggagcc gcgagccccg gggggcgacg acggtgcccc 8460

gcggggtggt ggtggtggtg gcggtgcagc tcagaagcgg tgcgcgggc gggccccgg 8520

aggtaggggg ggctccggtc ccgcgggcag gggcggcagc ggcacgtcgg cgtggagcgc 8580

gggcaggagt tgggtctgtg cccgaggtt gctggcgaag gcgacgacgc ggcggttgat 8640

ctctggatc tggcgcctc gcgtgaagac gacggggccg gtgagcttga acctgaaaga 8700

gagttcgaca gaatcaatct cgggtgcatt gaccgggcc tggcgagga tctctgcac 8760

gtctcccag ttgtcttggg aggcgatctc ggccatgaac tgctcgatct ctctctctg 8820

gaggtctccg cgtccggcgc gttccacggt ggccgccagg tcgttgaga tgcgccccat	8880
gagctgcgag aaggcgttga gtccgccctc gttccagact cggctgtaga ccacgcccc	8940
ctggtcatcg cgggcgcgca tgaccacctg cgcgaggttg agtccacgt gccgcgcgaa	9000
gacggcgtag ttgcgcagac gctggaagag gtagttgagg gtggtggcgg tgtgctcggc	9060
cacgaagaag ttcatgacct agcggcgcaa cgtggattcg ttgatgtccc ccaaggcctc	9120
cagccgttcc atggcctcgt agaagtccac ggcaagtgtg aaaaactggg agttgcgcgc	9180
cgacacggtc aactcctcct ccagaagacg gatgagctcg gcgacggtgt cgcgcacctc	9240
gcgctcgaag gctatgggga tctcttctc cgctagcacc accacctcct cctcttctc	9300
ctcttctggc acttccatga tggcttctc ctcttcgggg ggtggcggcg gcggcggtgg	9360
gggagggggc gctctgcgcc ggccggcgcg caccgggagg cggccacga agcgcgcgat	9420
catctcccc cgccggcggc gcatggtctc ggtgacggcg cggccgttct cccggggcg	9480
cagttggaag acgcccgcgg acatctggtg ctggggcggg tggccgtgag gcagcgagac	9540
ggcgtgacg atgcatctca acaattgctg cgtaggtacg ccgccgaggg acctgaggga	9600
gtccatatec accggatccg aaaacctttc gaggaaggcg tctaaccagt cgcagtcgca	9660
aggtaggctg agcaccgtgg cgggcggcgg ggggtggggg gagtgtctgg cggaggtgct	9720
gctgatgatg taattgaagt aggcggactt gacacggcgg atggtcgaca ggagcaccat	9780
gtccttgggt cggccctgct ggatgcggag gcggtcggct atgcccagg ctctgttctg	9840
gcatcggcgc aggtccttgt agtagtcttg catgagcctt tccaccgca cctcttctc	9900
ttctcttct gcttcttcca tgtctgttc ggccctgggg cggcgccgcg ccccttgc	9960
ccccatgcgc gtgacccga acccctgag cgttggagc agggccaggt cggcgacgac	10020
gcgctcgcc aggatggcct gctgcacctg cgtgagggtg gtttgaagt catccaagtc	10080
cacgaagcgg tggtaggcgc ccgtgtgat ggtgtaggtg cagttggcca tgacggacca	10140
gttgacggtc tggtagcccg gttgcgacat ctcggtgtac ctgagtcgcg agtaggcgcg	10200
ggagtgaag acgtagtctg tgcaagtcg caccaggtac tggtagcca ccaggaagtg	10260
cggcgccggc tggcggtaga ggggccagcg cagggtggcg ggggtccgg gggccagtc	10320
ttccagcatg aggcggtggt aggcgtagat gtacctggac atccaggtga taccgcggc	10380
ggtggtggag gcgcgcggga agtcgcgcac ccggttccag atgttgcga ggggcagaaa	10440
gtgctccatg gtaggcgtgc tctgtccagt cagacgcgcg cagtcgttga tactctagac	10500
cagggaaac gaaagccggt cagcgggcac tcttccgtgg tctggtgaat agatcgcaag	10560
ggtatcatgg cggagggcct cggttcgagc ccgggtccg ggccggacgg tccgccatga	10620
tccacgcggt taccgcccgc gtgtcgaacc caggtgtgcg acgtcagaca acggtggagt	10680

gttcccttttg gcgtttttct ggccggggcgc cggcgccgcg taagagacta agccgcgaaa	10740
gcgaaagcag taagtggctc gctccccgta gccggaggga tccttgctaa gggttgctt	10800
gcggcgaacc ccggttcgaa tcccgtaactc ggcccgcccg gacccgcgcg taaggtgttg	10860
gattggcctc cccctcgtat aaagaccccg cttgcggatt gactccggac acggggacga	10920
gcccccttta tttttgcttt cccagatgc atccggtgtt gcgacagatg cccccccgc	10980
cccagcagca gcaacaacac cagcaagagc gccagcaaca gcagcgggag tcatgcaggg	11040
ccccctcacc caccctcggc ggcccggcca cctcggcgtc cgcggccgtg tctggcgctt	11100
gcggcggcgg cggcgggggg ccggctgacg accccgagga gccccgcgg cgcagggcca	11160
gacactacct ggacctggag gagggcgagg gcctggcgcg gctggggcg ccgtctcccg	11220
agcgccacc gcgggtgcag ctaaagcgcg actcgcgcga ggctacgtg cctcggcaga	11280
acctgttcag ggaccgcgcg ggcgaggagc ccgaggagat gcgggacagg aggttcagcg	11340
cggggcggga gctgcggcag gggctgaacc gcgagcggct gctgcgcgag gaggactttg	11400
agcccgcgcg gcggacgggg atcagccccg cgcgcgcgca cgtggcggcc gccgacctgg	11460
tgacggcgta cgagcagacg gtgaaccagg agatcaactt ccaaaagagt ttcaacaacc	11520
acgtgcgcac gctggtggcg cgcgaggagg tgaccatcgg gctgatgcac ctgtgggact	11580
ttgtgagcgc gctggtgcag aacccaata gcaagcctct gacggcgcag ctgttcctga	11640
tagtgacga cagcaggagc aacgaggcgt ttaggagcgc gctgctgaac atcaccgagc	11700
ccgagggccg gtggctgctg gacctgatta acatcctgca gagcatagt gtgcaggagc	11760
gcagcctgag cctggccgac aaggtggcgg ccatcaacta ctcgatgctg agcctgggca	11820
agttttacgc gcgcaagatc taccagacgc cgtacgtgcc catagacaag gaggtgaaga	11880
tcgacggttt ttacatgcgc atggcgctga aggtgctcac cctaagcgac gacctgggcg	11940
tgtaccgcaa cgagcgcac cacaaggccg tgagcgtgag ccggcggcgc gagctgagcg	12000
accgcgagct gatgcatagc ctgcagcggg cgtggcggg gcccggcagc ggcgacaggg	12060
aggcggagtc ctacttcgat gcggggggcg acctgcgctg ggccccagc cggcggggcc	12120
tggaggccgc ggggggtccg gaggactatg acgaggacgg cgaggaggat gaggagtacg	12180
agctagagga gggcgagtac ctggactaaa ccgcgggtgg tgtttccgt agatgcaaga	12240
cccgaacgtg gtggaccgg cgctcggggc ggctctgcag agccagccgt ccggccttaa	12300
ctcctcagac gactggcgac aggtcatgga ccgcatcatg tcgctgacgg cgcgtaaccc	12360
ggacgcgttc cggcagcagc cgcaggccaa caggctctcc gccatcctgg aggcgttgtt	12420
gcctgcgcgc tcgaacccca cgcacagaaa ggtgctggcc atagtgaac cgctggccga	12480
gaacagggcc atccgcccgg acgaggccgg gctggtgtac gacgcgctgc tgcagcgcgt	12540

ggccccctac aacagcggca acgtgcagac caacctggac cggctggtgg gggacgtgcg	12600
cgaggcgggtg ggcgcagcgcg agcgcgcgga tcggcagggc aacctgggct ccatggtggc	12660
gctgaatgcc ttcttgagca cgcagccggc caacgtgccg cgggggcagg aagactacac	12720
caactttgtg agcgcgctgc ggctgatggt gaccgagacc cccagagcg aggtgtacca	12780
gtcgggtccg gactacttct tccagaccag cagacagggc ctgcagacgg tgaacctgag	12840
ccaggctttc aagaacctgc gggggctgtg gggcgtgaag gcgcccacg gcgaccgggc	12900
gacggtgtcc agcctgtga cgcccaactc gcgcctgtg ctgctgtga tcgcgccgtt	12960
cacggacagc ggcagcgtgt cccgggacac ctacctgggg cacctgtga ccctgtaccg	13020
cgaggccatc gggcaggcgc aggtggacga gcacacctc caggagatca ccagcgttag	13080
ccgcgcgctg gggcaggagg acacgagcag cctggaggcg actctgaact acctgtgac	13140
caaccggcgg cagaagattc cctcgtgca cagcctgacc tccgaggagg agcgcattt	13200
gcgctacgtg cagcagagcg tgagcctgaa cctgatgcgc gacggggtga cggccagtgt	13260
ggcgtggac atgaccgcgc gcaacatgga accgggcatg tacgccgcgc accggcctta	13320
catcaaccgc ctgatggact acctgcatcg cgcggcggcc gtgaaccccg agtactttac	13380
caacgccatc ctgaacccgc actggctccc gccgccggg tttacagcg ggggcttcga	13440
ggtcccggag gccaacgatg gtttctgtg ggacgacatg gacgacagcg tgtttcccc	13500
gcggccgcag gcgtggcgg aagcgtccct gctgcgtccc aagaaggagg aggaggcgag	13560
tcgccccgc ggcagcagcg gcgtggcttc tctgtccgag ctgggggcgg cagccgccgc	13620
gcgccccggg tccttggcg gcagccctt tccgagcctg gtggggtctc tgcacagcga	13680
gcgcaccacc cgccctcggc tgctgggcga ggacgagtac ctgaataact ccctgtgca	13740
gccggtgcgg gagaaaaacc tgcctccgc cttccccaac aacgggatag agagcctggt	13800
ggacaagatg agcagatgga agacctatgc gcaggagcac agggacgcgc ccgcgtccg	13860
gccgcccacg cggcgcacg gccacgaccg gcagcggggg ctggtgtggg atgacgagga	13920
ctccgcggac gatagcagcg tgctggacct gggaggagc ggcaaccgt tcgcgcacct	13980
gcgccccgc ctggggagga tgttttaaaa aaaaaaaaaa gcaagaagca tgatgcaaaa	14040
attaaataaa actcaccaag gccatggcga ccgagcgttg gtttcttgtg ttcccttcag	14100
tatgcggcgc gcggcgtgt accaggaggg acctctctcc tcttacgaga gcgtggtggg	14160
cgcggcggcg gcggcgcct cttctccctt tgcgtcgcag ctgctggagc cgccgtacgt	14220
gcctccgcgc tacctgcggc ctacgggggg gagaaacagc atccgttact cggagctggc	14280

gccccgtttc gacaccaccc ggggtgtacct ggtggacaac aagtcggcgg acgtggcctc	14340
cctgaactac cagaacgacc acagcaattt tttagaccag gtcattccaga acaatgacta	14400
cagccccgagc gagggccagca cccagaccat caatctggat gaccggctcg actggggcgg	14460
cgacctgaaa accatcctgc acaccaacat gcccacgtg aacgagttca tgttcaccaa	14520
taagttcaag gcgcgggtga tgggtgcgcg ctgcacacc aaggaagacc ggggtggagct	14580
gaagtacgag tgggtggagt tcgagctgcc agagggaac tactccgaga ccatgacat	14640
tgacctgatg aacaacgcga tcgtggagca ctatctgaaa gtgggcaggc agaacggggt	14700
cctggagagc gacatcgggg tcaagttcga caccaggaac ttccgctgg ggctggaccc	14760
cgtgaccggg ctggttatgc ccggggtgta caccaacgag gccttccatc ccgacatcat	14820
cctgctgcc ggctgcgggg tggacttcac ttacagccgc ctgagcaacc tcctgggcat	14880
ccgaagcgg cagcccttc aggagggtt caggatcacc tacgaggacc tggagggggg	14940
caacatcccc gcgtcctcg atgtggaggc ctaccaggat agcttgaagg aaaatgaggc	15000
gggacaggag gataccgcc ccgccctc cgccgccgc gagcaggcg aggatgctgc	15060
tgacaccgc gccgcggagc gggcggaggc cgacccgct atggtggtgg aggctgccga	15120
gcaggaggag gacatgaatg acagtgcggt gcgcggagac accttcgtca cccgggggga	15180
ggaaaagcaa gcggaggccg aggccgcggc cgaggaaaag caactggcgg cagcagcggc	15240
ggcggcggcg ttggccgagg cggaggctga gtctgagggg accaagccc ccaaggagcc	15300
cgtgattaag cccctgacc aagatagcaa gaagcgagt tacaacctgc tcaaggacag	15360
caccaacacc gcgtaccga gctgtacct ggcctacaac tacggcgacc cgtcgacggg	15420
ggtgcgctcc tggacctgc tgtgcacgc ggacgtgacc tgcggctcgg agcaggtgta	15480
ctggtcgctg cccgacatga tgcaagacc cgtgaccttc cgctccacgc ggcaggtcag	15540
caacttccc gtggtgggag ccgagctgct gccctgcac tccaagagct tctacaacga	15600
ccaggccgtc tactcccagc tcatccgcca gttacctct ctgaccacg tgttcaatcg	15660
ctttctgag aaccagattc tggcgcgcc gccgcccc accatcacca ccgtcagtga	15720
aaacgttct gctctcacag atcacgggac gctaccgtg cgcaacagca tcggaggagt	15780
ccagcgagt accgttactg acgccagac ccgcacctgc ccctacgttt acaaggcctt	15840
gggcatagtc tcgccgcg tctttccag ccgcacttt tgagcaacac caccatcatg	15900
tccatcctga tctacccag caataactcc ggtggggac tgctgcgcgc gccagcaag	15960
atgttcggag gggcgaggaa gcgttccgag cagcaccg tgcgctgcg cgggcacttc	16020
cgccccctt ggggagcgca caaacgcggc cgcgcggggc gcaccaccgt ggacgacgc	16080
atcgactcgg tgggtggagca ggcgcgcaac tacaggcccg cggtctctac cgtggacgcg	16140

gccatccaga ccgtggtgcg gggcgcgcg cggtacgcca agctgaagag cgcgcggaag	16200
cgcgtggccc gccgccaccg ccgccgaccc gggggccgccc ccaaaccgcg cgcgcggccc	16260
ctgcttcgcc gggccaagcg caccggcgcg cgcgccgcca tgaggcgcg cgccgcttg	16320
gccgcccggca tcaccgccc caccatggcc ccccgtaacc gaagacgcg gcccgccc	16380
gccgcccgcg ccatcagtga catggccagc aggcgcccgg gcaacgtgta ctgggtgcg	16440
gactcggtag cggcacgcg cgtgcccgtg cgcttcgcc ccccgcgac ttgagatgat	16500
gtgaaaaaac aacactgagt ctctgctgt tgtgtgtatc ccagcgcgcg cgccgcgcg	16560
cgcagcgtca tgtccaagcg caaaatcaaa gaagagatgc tccaggtcgt cgcgccggag	16620
atctatgggc cccgaagaa ggaagagcag gattcgaag cccgcaagat aaagcgggtc	16680
aaaaagaaaa agaaagatga tggcgatgcc gatggggagg tggagtctct gcgcccacg	16740
gcgcccaggc gcccggtgca gtggaaggcg cgcgcgtaaa agcgcgctct gcgcccggc	16800
accgcggtgg tcttcacgcc cggcgagcg tccaccgga cttcaagcg cgtctatgac	16860
gaggtgtacg gcgacgaaga cctgctggag caggccaacg agcgcttcgg agagtttct	16920
tacgggaagc gtcagcgccc gctggggaag gaggacctgc tggcgctgcc gctggaccag	16980
ggcaacccca ccccgagct gaagcccggt accctgcagc aggtgctgcc gagcagcgca	17040
ccctccgagg cgaagcgggg tctgaagcg gagggcgcg acctggcgcc caccgtgcag	17100
ctcatggtgc ccaagcggca gaggtggag gatgtgctgg agaaaatgaa agtagacccc	17160
ggtctgcagc cggacatcag ggtccgtccc atcaagcagg tggcgccggg cctcgcgctg	17220
cagaccgtgg acgtggtcat cccaccggc aactccccg ccgccaccac cactaccgt	17280
gcctccacgg acatggagac acagaccgat cccgcccag ccgcccacc cgcgcgccc	17340
gcgacctct cggcgagggt gcagacggac ccttggtgc cgccggcgat gtcagctccc	17400
cgcgcgctc gggcgcgag gaagtacggc gcgccaacg cgctcctgcc cgagtacgc	17460
ttgcatcctt ccatcgccc caccgccgc taccgaggt atacctacc cccggaaga	17520
gccaagggtt ccaccgccc tccccgccga cgcgccgccc ccaccaccg ccgcccgcg	17580
cgcagacgc agccgcact ggctccagtc tccgtgagga gaggcgcg cgacggacac	17640
acctggtgc tgcccaggc gcgtaccac ccagcatcg tttaaaagc tgttgtgtt	17700
cttgagata tgccctcac ttgccgctc cgtttccgg tgccgggata ccgaggagga	17760
agatcgccc gcaggagggg tctggccgac cgcgccctga gcggaggcag ccgcccgcg	17820
caccgcccgc gacgcccac cagccgacgc atgcgcccgc gggtgctgcc cctgttaac	17880
ccctgatcg ccgcccgat cggcgccgtg cccgggatcg cctccgtggc ctgacggcg	17940
tcccagaggc attgacagac ttgcaactt gcaaatatgg aaaaaacc caataaaaa	18000

gtctagactc tcacgctcgc ttggctcctgt gactatittg tagaatggaa gacatcaact	18060
ttgcgtcgtc gggcccgctg cacggtcgc gcccgttcct gggacactgg aacgatatcg	18120
gcaccagcaa catgagcggg gggccttca gttggggctc tctgtggagc ggcattaaaa	18180
glatcgggtc tgccgttaaa aattacggct cccgggctg gaacagcagc acgggccaga	18240
tgttgagaga caagttaaa gagcagaact tccagcagaa ggtgtggagc ggcctggcct	18300
ccggcatcaa cgggggtgtg gacctggcca accaggccgt gcagaataag atcaacagca	18360
gactggacct cggcccgccg gtggaggagg tgccgccggc gctggagacg gtgtccccg	18420
atgggcgtgg cgagaagcgc ccggggcccg atagggaaga gaccactctg gtcacgcaga	18480
ccgatgagcc gccccgtat gaggaggccc tgaagcaagg tctgccacc acgcggccca	18540
tcgcgcccat ggccaccggg gtggtgggcc gccacacccc cgccacgtg gacttgctc	18600
cgcccgccga tgtccgcag cagcagcaga aggcggcaca gccgggcccg cccgtgaccg	18660
cctccgttc ctcgccggt cctctgcgc gcccgccag cggcccccgc gggggggtcg	18720
cgaggcacgg caactggcag agcacgtga acagcatctg gggtctgggg gtgcgtccg	18780
tgaagcgccg ccgatgtac tgaatagctt agctaactgt ttgtatgtgt gtatgcgcc	18840
tatgtcgccg ccagaggagc tgctgagtcg ccgccgttcg cgcgccacc accaccaccg	18900
ccactccgc cctcaagatg gcgaccccat cgatgatgcc gcagtggtcg tacatgcaca	18960
tctcgggcca ggacgcctcg gactacctga gccccgggt ggtgcagttc gcccgcgcca	19020
ccgagagcta cttcagcctg agtaacaagt ttaggaaccc cacggtggcg cccacgcacg	19080
atgtgaccac cgaccggtct cagcgctga cgtgcggtt cattcccgtg gaccgcgagg	19140
acaccgcta ctctacaag gcgcggttca ccttgccgt gggcgacaac cgcgtgctgg	19200
acatggcctc cacctacttt gacatccgcg gggtgctgga ccggggcccc actttcaagc	19260
cttactctgg caccgcctac aactccctgg ccccaaggc cgtcccaac tctgctgagt	19320
gggagcaatt agaagaagcc caggccgtg tggaagacga agaattagaa gatgaagacg	19380
aggaaccaca ggatgaggca cctgtgaaaa aaacccatgt atacgtcag gtcgcccttt	19440
ctggagaaga aattactaaa aacggtttgc aaatagggtc agataacaca gaagcccagt	19500
ctaagcccat atatgcagat cctacattcc agcctgaacc ccaaatcggg gaatcccagt	19560
ggaatgaggc agatgctaca gttgccggcg gtagagtgt aaagaaatcc actccatga	19620
agccatgcta tggttcctat gcaagacca caaactcaa tggagggtcaa ggtgtgctgg	19680
tggtgatga taagggggtt cttcaatcta aagttgaatt gcaatTTTT tcaaatacta	19740

ctactcttaa tcagcgggag ggtaacgata caaaacaaaa agtgggtgctg tatagcgaag	19800
atgtgcatat ggaaactcca gacaccaca tttcttaca gcccacaaaa agcgaatgaca	19860
attcaaaaat catgctgggt cagcagtcga tgcccaacag acctaattac atcggtctca	19920
gagacaactt tatcggcctc atgtattaca atagcactgg caacatggga gtgcttgag	19980
gtcaggcctc tcagttgaat gcagtgggtg acttgcaaga cagaaacaca gaactgtcct	20040
accagctctt gcttgattcc atgggtgaca gaaccagata cttttccatg tggaatcagg	20100
cagtggacag ttatgacca gatgtcagaa ttattgaaaa tcatggaact gaagacgagc	20160
tcccaacta ttgtttccct ctgggcggca taggggtaac tgacacttac caggccatta	20220
aaaccaatgg caatggtcaa gaaaacccaa cctgggaaaa agatacagag ttgacagacc	20280
gcaatgaaat aggggtggga aacaatttcg ctatggagat caacctcagt gccaacctgt	20340
ggagaaaact cctgtactcc aacgtggcgc tgtacctgcc agacaagctt aagtacaacc	20400
cctccaatgt ggacatctct gacaaccca acacctacga ttacatgaac aagcgagtgg	20460
tggccccggg gctggtggac tgctacatca acctgggcgc gcgtggtcg ctggactaca	20520
tggacaacgt caacccttc aaccaccacc gcaatgcggg cctgcgtac cgctccatgc	20580
tcctgggcaa cgggcgctac gtgcccttc acatccaggt gcccagaag ttctttgcca	20640
tcaagaacct cctcctcctg ccgggctcct acacctacga gtggaacttc aggaaggatg	20700
tcaaatggt cctccagagc tctctgggta acgatctcag ggtggacggg gccagcatca	20760
agttcgagag catctgcctc tacgccact tcttcccat ggcccacaac acggcctcca	20820
cgctcgaggc catgctcagg aacgacacca acgaccagtc cttcaatgac tacctctccg	20880
ccgccaacat gctctacccc ataccgcca acgccacaa cgtccccatc tccatccct	20940
cgcgcaactg ggcggccttc cgcgctggg ccttcaccg cctcaagacc aaggagacc	21000
cctccctggg ctcggtatc gaccctact acacctactc gggtccatt ccctacctgg	21060
acggcacctt ctacctaac cacactttca agaaggtctc ggtcaccttc gactcctcgg	21120
tcagctggcc gggcaacgac cgtctgctca ccccaacga gttcgaaatc aagcgctcgg	21180
tcgacgggga gggtacaaac gtggcccagt gcaacatgac caaggactgg ttcctggtec	21240
agatgctggc caactacaac atcggtacc agggcttcta catcccagag agctacaagg	21300
acaggatgta ctctttcttc aggaacttcc agcccatgag ccggcaggtg gtggaccaga	21360
ccaagtacaa ggactaccag gaggtgggca tcatccacca gcacaacaac tcgggcttcg	21420
tgggtacct cgccccacc atgcgcgagg gacaggccta ccccgccaac ttccctacc	21480
cgctcatagg caagaccgag gtcgacagca tcaccagaa aaagtctctc tgcgatcgca	21540
ccctctggcg catcccttc tccagcaact tcatgtccat ggggtgcgctc tcggacctgg	21600

gccagaactt gctctacgcc aactccgccc acgccctcga catgaccttc gaggtcgacc 21660
ccatggacga gcccaccctt ctctatgttc tgttcgaagt ctttgacgtg gtccgggtcc 21720
accagccgca ccgcgcgctc atcgagaccg tgtacctgcg tacgcccttc tcggccggca 21780
acgccaccac ctaaagaagc aagccgcagt catcgccgcc tgcattgccgt cgggttcac 21840

cgagcaagag ctgagggcca tctcagaga cctgggatgc gggccctatt ttttgggcac 21900
tttcgacaag cgtttccctg gctttgtctc cccacacaag ctggcctgcg ccatcgtcaa 21960
cacggccggc cgcgagaccg ggggctgca ctggctggcc ttcgcctgga acccgcgctc 22020
caaaacatgc ttctctttg acccttcgg cttttcggac cagcggtca agcaaacta 22080
cgagttcgag tacgagggt tgcctgctcg cagcgccatc gcctcctcgc ccgaccgtg 22140
cgtcacctc gaaaagtcca cccagaccgt gcaggggccc gactcggccg cctgcggtct 22200
cttctgctgc atgtttctgc acgctttgt gactggcct cagagtccta tggaccgaa 22260

ccccaccatg aacttgctga cgggggtgcc caactccatg ctccagagcc cccaggtcga 22320
gcccaccctg cgccgaacc aggagcagct ctacagcttc ctggagcgcc actcgccta 22380
cttcgcccgc cacagcgac agatcaggag ggccacctcc ttctgccact tgcaagagat 22440
gcaagaaggg taataacgat gtacacactt ttttctcaat aaatggcatt tttttatita 22500
tacaagctct ctgggttatt catttccac caccaccacc acccgccgtt gtcgccatct 22560
ggctctatct agaaatcgaa agggttctgc cgggagtcgc cgtgcgccac gggcagggac 22620
acgttcgat actggtagcg ggtgccccac ttgaactcgg gcaccaccag gcgaggcagc 22680

tcggggaagt tttcgtcca caggtcgcgg gtcagcacca gcgcgttcac caggtcgggc 22740
gccgagatct tgaagtcga gttggggccg ccgccctgcg cgcgcgagtt gcggtacacc 22800
gggttcgac actggaacac caacagcgcc gggtgcttca cgctagccag cacgtcgcgg 22860
tcggagatca gtcggcgctc caggtcctcc gcgttgctca gcgcgaacgg ggtcatcttg 22920
ggcacttgcc tccccaggaa gggcgcgctgc ccgggtttcg agttgcagtc gcagcgagc 22980
gggatcagca ggtgcccatg cccggactcg gcgttgggg acagcgcgcg catgaaggcc 23040
tgcatctggc ggaaggccat ctgggccttg gcgccctccg agaagaacat gccgcaggac 23100

ttgcccgaga actggtttgc ggggcagctg gcgtcgtgca ggcagcagcg cgcgtcggtg 23160
ttggcgatct gcaccaggt gcgccccac cgtttcttca cgatcttgge ctggacgat 23220
tgctccttca gcgcgcgctg cccgttctcg ctggtcacat ccatctcgat cacatgttcc 23280
ttgttcacca tgcgtgctgc gtgcaggcac ttcagctcgc cctccgtctc ggtgcagcgg 23340
tgctgccaca gcgcgcagc cgtgggctcg aaagacttgt aggtcacctc cgcaaggac 23400
tgcaggtacc cctgcaaaaa gcggcccatc atggtcacga aggtcttgtt gctgctgaag 23460

gtcagctgca gcccgcggtg ctctctgttc agccaggtct tgcacacggc cgccagcgcc	23520
tccacctggt cgggcagcat cttgaagtgc accttcagct cattctccac gtggtacttg	23580
tccatcagcg tgcgcgcgcg ctccatgcc ttctcccagg ccgacaccag cggcaggctc	23640
acggggttct tcaccatcac cgtggccgcc gcctccgccg cgctttcgct ttccgccccg	23700
ctgttctctt cctcttcttc ctcttctctg ccgccgccca ctgcagccc ccgcaccacg	23760
gggtcgtctt cctgcaggcg ctgcaccttg cgcttgccgt tgcgcccctg cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcac agcgcggcct cttcttgctc gtctcgtctg	23880
tccagaatga cctccgggga ggggggggtg gtcacctca gtaccgaggc acgttcttt	23940
ttcttcttg gggcgcttcg cagctccgcg gctgcggccg ctgccgaggt cgaaggccga	24000
gggctgggcg tgcgcggcac cagcgcgtcc tgcgagccgt cctcgtctc ctcgactcg	24060
agacggaggc gggccccgtt ctctgggggc gcgcggggcg gcggaggcgg cggcggcgac	24120
ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgcc gcgtccgcgc	24180
tcgggggttg tctcgcgtg gtctcttcc cgaactggcca tctccactg ctcttctcc	24240
tataggcaga aagagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accgccccct ctgagccctc caccaccgcc gccaccaccg ccaatgccgc cgcggacgac	24360
gcgcccaccg agaccaccgc cagtaccacc nnctcccca gcgacgcacc cccgctcgag	24420
aatgaagtgc tgatcgagca ggaccgggt tttgtgagcg gagaggagga tgaggaggat	24480
gagaaggaga aggaggaggt cgccgcctca gtgccaaaag aggataaaaa gcaagaccag	24540
gacgacgagc ataaggatga gacagcagtc gggcggggga acggaagcca tgatgctgat	24600
gacggctacc tagacgtggg agacgacgtg ctgcttaagc acctgcaccg ccagtgcgtc	24660
atcgtctgcg acgcgtgca ggagcgggtg gaagtgcgcc tggacgtggc ggaggtcagc	24720
cgcgcctacg agcggcacct ctctcgccg cagtgcccc ccaagcgccg ggagaacggc	24780
acctgcgagc ccaaccgcg tctcaacttc taccgggtct tcgcggtacc cgagggtctg	24840
gccacctacc acatcttttt ccaaaactgc aagatcccc tctctgcg cgctaaccgc	24900
accgcgcgcg aaaaaacct gacctgcgg caggcgcccc acatacctga tattgcctct	24960
ctggaggaag tgcccaagat ctctgagggt ctcggtcgcg acgagaaacg ggcggcgaac	25020
gctctgcacg gagacagcga aaacgagagt cactcggggg tgctggtgga gctcgagggc	25080
gacaacgcgc gcctggccgt actcaagcgc agcatagagg tcaccactt tgcctaccgc	25140
gcgtcaacc tgcccccaa ggtcatgagt gtggtcatgg gcgagctcat catgcgccgc	25200

gcccagcccc tggccgcgga tgcaaacttg caagagtcct cagaggaagg cctgcccgcg	25260
gtcagcgacg agcagctggc gcgttggttg gagaccgcg accccgcgca gctggaggag	25320
cggcgcaagc tcatgatggc cgcgttgctg gtcaccgtgg agctcgagtg tctgcagcgc	25380
ttcttcgagg accccgagat gcagcgcaag ctgcaggaga cctgcaacta caccttccgc	25440
cagggtctacg tgcgccaggc ctgcaagatc tccaacgtgg agctctgcaa cctggtctcc	25500
tacctgggca tctgcacga gaaccgcctc gggcagaacg tctgcaactc caccctcaaa	25560
ggggaggcgc gccgcgacta catccgcgac tgcgcctacc tcttcctctg ctacacctgg	25620
cagacggcca tgggggtctg gcagcagtg cctggaggagc gcaacctcaa ggagctggaa	25680
aagctcctca agcgcaccct caggggacctc tggacgggct tcaacgagcg ctcggtggcc	25740
gccgcgttgg cggacatcat cttccccgag cgcctgtca agacctgca gcagggcctg	25800
cccgaactca ccagccagag catgtctgag aacttcagga ctttcatcct ggagcgctcg	25860
ggcatcctgc cggccacttg ctgcgcgttg cccagcgact tcgtgcccac caagtacagg	25920
gagtgtccgc cgcgctctg gggccactgc tacctcttcc agctggccaa ctacctgcc	25980
taccactcgg acctcatgga agacgtgagc ggcgagggcc tgctcgagtg ccactgccgc	26040
tgcaacctct gcacgcccc cgcctctcta gtctgcaacc cgcagctgct cagcgagagt	26100
cagattatcg gtaccttga gctgcagggt ccctcgctg acgagaagtc cgcggctccg	26160
gggctgaaac tactccggg gctgtggact tccgcctacc tacgcaaatt tgtacctgag	26220
gactaccacg cccacgagat caggttctac gaagaccaat cccgcccgc caaggcgag	26280
ctcaccgct gcgtcatcac ccaggggcac atcctgggccc aattgcaagc catcaacaaa	26340
gcccccgag agttcttct gaaaaagggt cggggggtgt acctggacc ccagtccgc	26400
gaggagctaa acccgctacc cccgcccgcg cccagcagc gggaccttgc ttcccaggat	26460
ggcaccaga aagaagcagc agcccccgc gcagccatac atgcttcttg aggaagagga	26520
ggaggactgg gacagtcagg cagaggagat gatggaagac tgggaggagg acagcgcct	26580
agacgaggaa gtttcagagg ccgaagaggt ggcagacgca acaccatcac cctcggtcgc	26640
agccccctcg cgggggcccc tgaaatctc cgaaccagc accagcgta taacctccgc	26700
tcctccggcg ccggcgccac ccgcccgcag acccaaccgt agatgggaca ccacaggaac	26760
cggggtcgtt aagtccaagt gcccgcgccc gccaccgcag cagcagcagc agcagcgcca	26820
gggctaccgc tcgtggcgcg ggcacaagaa cgccatagtc gcctgcttgc aagactgcgg	26880
gggcaacatc tctttgccc gccgttctt gctattccac cacggggtcg cctttccccg	26940
caatgtcctg cattactacc gtcattctta cagcccctac tgcagcgcg acccagaggc	27000
ggcagcgga gccacagcgg cgaccaccac ctaggagat atcctccgcg ggcaagacag	27060

cggcagcagc ggccaggaga cccgcggcag cagcggcggg agcgggtgggc gcaactgcgcc	27120
tctcgcccaa cgaaccctc tcgaccggg agctcagaca caggatcttc cccactttgt	27180
atgccatctt ccaacagagc agaggccagg agcaggagct gaaaaataaa aacagatctc	27240
tgcgctccct caccgcagc tgtctgtatc acaaaagcga agatcagctt cggcgcacgc	27300
tggaggacgc ggaggcactc ttcagcaaat actgcgcgt cactcttaaa gactagctcc	27360
gcgccttct cgaatttagg cgggagaaaa ctacgtcatc gccggccgcc gccagcccg	27420
cccagccgag atgagcaaag agattccac gccatacatg tggagctacc agccgcagat	27480
gggactcgcg gcgggagcgg cccaggacta ctccaccgc atgaactaca tgagcgcggg	27540
accccatatg atctcacagg tcaacgggat ccgcgccag cgaaacaaa tactgctgga	27600
acaggcggcc atcacgccca cgcgccca taatctcaac ccccgaaatt ggcccgccgc	27660
cctagtgtac caggaaacc cctccgccac caccgtacta cttccgctg acgccaggc	27720
cgaagtccag atgactaact caggggcgca gctcgcgggc ggctttcgtc acggggcgcg	27780
gccgctccga ccaggtataa gacacctgat gatcagaggc cgaggtatcc agctcaacga	27840
cgagtcggtg agctcttcgc tcggtctcgc tccggacgga actttccagc tcgccggatc	27900
cggctcgtct tcgttcacgc cccgccaggc gtacctgact ctgcagacct cgtcctcgga	27960
gccccgtcc ggccgcatcg gaacctcca gtctgtggag gatttcgtgc cctcggctta	28020
cttcaacccc ttctcgggac ctcccggacg ctaccccgac cagttcattc cgaactttga	28080
cgcggtgaag gactcggcgg acggctacga ctgaatgtca ggtgccgagg cagagcagct	28140
tcgcctgaga cacctcagc actgccgccg ccacaagtgc ttcccccgcg gttccggtga	28200
gttctgtac tticagctac ccgaggagca taccgagggg ccggcgcacg gcgtccgcct	28260
gaccaccag ggcgaggtta cctgttcct catccgggag ttaccctcc gtccctgct	28320
agtggagcgg gagcggggtc cctgtgtcct aactatgcc tgcaactgcc ctaacctgg	28380
attacatcaa gatctttgtc gtcatctctg tgctgagttt aataaacgt gagatcagaa	28440
tctactgggg ctctgtcgc catctgtga acgccaccgt cttaccacac cccgaccagg	28500
cccaggcgaa ctcacctgc ggtctgcac ggaggtecaa gaagtacac accgtgtact	28560
tcaacggcac cccctttgtg gtttacaaca gcttcgacgg ggacggagtc tccctgaaag	28620
accagctctc cgtctcagc tactccatcc acaagaacac caccctcaa ctcttcctc	28680
cctacctgcc gggaacctac gattgcgtca ccggccgctg caccacctc accgcctga	28740
tcgtaaacca gagctttccg ggaacagata actccctctt cccagaaca ggaggtgagc	28800
tcaggaaact ccccggggac caggcgggag acgtacctc gacctttgtg gggttaggat	28860
tttttattac cgggttgctg gctcttttaa tcaaagcttc cttgagattt gttcttctc	28920

tctacgtgta tgaacacctc agcctccaat aactctaccc tttcttcggg atcaggtgac	28980
ttttctgaaa tcgggcttgg tgtgctgctt actctgttga ttttttcct tatcatactc	29040
agccttctgt gcctcaggct cgccgcctgc tgcgcacaca tctatatcta ctgctggttg	29100
ctcaagtga ggggtcgcca cccaagatga acaggtacat ggtcctatcg atcctaggcc	29160
tgttgccct ggccgcctgc agccgcgcca aaaaagagat tacctttgag gagcccgctt	29220
gcaatgtaac tttaagccc gaggtgacc aatgcaccac cctcgtcaaa tgcgttacca	29280
atcatgagaa gctgcgcac gactacaaaa aaaaaactgg ccggtttgcg gtctatagt	29340
tgtttacgcc cggagacccc tctaactact ctgtcacctt cttccagggc ggacagtcta	29400
agatatcaaa ttacactttc cttttttaig agttgtgcga tgcggtcatg tacatgtcaa	29460
aacagtacaa cctgtggcct cctctcccc aggcgtgtgt ggaaaatact gggctctact	29520
gctgtatggc tttggcaatc actacgctcg ctctaactcg cacggtgcta tatataaaat	29580
tcaggcagag gcgaatcttt atcgatgaaa agaaaatgcc ttgatcgcta acaccggctt	29640
tctatctgca gaatgaatgc aatcacctcc ctactaatca ccaccacct ccttgcgatt	29700
gccccgggt tgacacgaat cgaagtgcga gtgggggtcca atgtcacctat ggtgggcccc	29760
gccggcaatt ccacctcat gtgggaaaaa tttgtccgca atcaatgggt tcatttctgc	29820
tctaaccgaa tcagtatcaa gccagagcc atctgcgatg ggcaaaatct aactctgatc	29880
aatgtgcaaa tgatggatgc tgggtactat tacgggcagc ggggagaaat cattaattac	29940
tggcgacccc acaaggacta catgctgcat gtagtcgagg cacttccac taccacccc	30000
actaccacct ctcccaccac cactaccact actactacta ctactaccac taccgtgcc	30060
cgccatacc gcaaaagcac catgattagc aaaaagcccc ctctgtctca ctcccagcc	30120
ggcgggcccc tcggtgcgac ctacagaaacc accgagcttt gcttctgcca atgactaac	30180
gccagcgctc atgaactgtt cgacctggag aatgaggatg cccagcagag ctccgttgc	30240
ctgaccagg aggtgttgga gcccgttgcc ctgaagcaga tcggtgattc aataattgac	30300
tcttcttctt ttgccactcc cgaataacct ccgattcta cttccacat caccgggtacc	30360
aaagacccta acctctcttt ctacctgatg ctgctgctct gtatctctgt ggtctcttcc	30420
gcgctgatgt tactggggat gttctgctgc ctgatctgcc gcagaaagag aaaagctcgc	30480
tctcagggcc aaccactgat gcccttcccc taccceccgg attttgcaga taacaagata	30540
tgagctcgct gcigacacta accgctttac tagcctgcgc tctaaccctt gtcgcttgcg	30600
actcgagatt ccacaatgtc acagctgtgg caggagaaaa tgttactttc aactccacgg	30660

ccgataccca gtggctgtgg agtggctcag gtagctactt aactatctgc aatagctcca	30720
cttccccag catatccca accaagtagc aatgcaatgc cagcctgttc accctcatca	30780
acgcttccac cctggacaat ggactctatg taggctatgt acccttttgt gggcaaggaa	30840
agacccacgc ttacaacctg gaagtctgcc agcccagaac cactacccaa gctwcymcca	30900
ycaccagcac cagcagcagc agccacagca gcagcagcag attattgact ttggttttgg	30960
ccagctcatc tgcgctacc caggccatct acagctctgt gcccgaacc actcagaccc	31020
accgcccaga aacgaccacc gccaccacc tacacacctc cagcgatcag atgccgacca	31080
acatcacccc cttggctctt caaatgggac ttacaagccc cactccaaaa ccagtggatg	31140
cggccgaggt ctccgccctc gtcaatgact gggcggggct gggaatgtgg tggttcgcca	31200
taggcatgat ggcgctctgc ctgcttctgc tctggctcat ctgctgcctc caccgcaggc	31260
gagccagacc ccccatctat agaccatca ttgtcctgaa ccccgataat gatgggatcc	31320
atagattgga tggcctgaaa aacctacttt tttcttttac agtatgataa attgagacat	31380
gcctcgcatt ttctgtaca tgttcttctt cccacctttt ctggggtgtt ctacgtggc	31440
cgtgtgtctt cacttgagg tagactgcct ctacccttc actgtctacc tgctttacgg	31500
attggtcacc ctactctca tctgcagcct aatcacagta atcatgcct tcatccagt	31560
cattgattac atctgtgtgc gcctcgata cttcagacac caccgcagt accgagacag	31620
gaacattgcc caacttctaa gactgctcta atcatgata agactgtgat ctgccttctg	31680
atcctctgca tctgcccac cctcacctcc tgccagtaca ccacaaaac tccgcgcaa	31740
agacatgcct cctgccgtt cacccaactg tggaatatac ccaaatgcta caacgaaaag	31800
agcgagctct ccgaagcttg gctgtatggg gtcatctgtg tcttagtttt ctgcagcact	31860
gtctttgcc tcatgatcta cccctacttt gatttgggat ggaacgcgat ccatgccatg	31920
aattacccca ctttccgc acccgagata attccactgc gacaagttgt accggtgtc	31980
gttaatcaac gcccccatc ccctacgcc actgaaatca gctacttta ctaacaggc	32040
ggagatgact gacgccctag atctagaaat ggacggcatc agtaccgagc agcgtctcct	32100
agagaggcgc aggcaggcgg ctgagcaaga ggcctcaat caggagctcc gagatctcgt	32160
taacctgcac cagtcaaaa gaggcacttt ttgtctggta aagcaggcca aagtcaccta	32220
cgagaagacc ggcaacagcc accgcctcag ttacaaattg cccaccagc gccagaagct	32280
gggtctcatg gtgggtgaga atcccatcac cgtcaccag cactcggtag agaccgagg	32340
gigtctgcac tccccctgc ggggtccaga agacctctgc accctggtaa agacctgtg	32400
cggctcaga gatttagtcc cttttaacta atcaaacact ggaatcaata aaaagaatca	32460
cttacttaaa atcagacagc aggtctctgt ccagtttatt cagcagcacc tccttccct	32520

cctcccaact ctggtactcc aaacgccttc tggcggcaaa cttctccac accctgaagg	32580
gaatgtcaga ttcttgctcc tgctccctccg caccactat cttcatgttg ttgcagatga	32640
agcgcaccaa aacgtctgac gagagcttca accccgtgta cccctatgac acggaaagcg	32700
gccctccctc cgtcccttcc ctcacccctc ccttcgtgtc tcccgatgga ttccaagaaa	32760
gtccccccgg ggtcctgtct ctgaacctgg ccgagccctt ggtcacttcc caggcatgc	32820
tcgccctgaa aatgggaagt ggcctctccc tggacgacgc tggcaacctc acctctcaag	32880
atatcaccac cgctagccct cccctcaaaa aaaccaagac caacctcagc ctagaacct	32940
catccccctt aactgtgagc acctcaggcg ccttcaccgt agcagccgcc gctccccctg	33000
cgggtggccgg cacttccctc accatgcaat cagaggcccc cctgacagta caggatgcaa	33060
aactcacctt ggccacaaa ggccccctga ccgtgtctga aggcaactg gccttgcaaa	33120
catcgcccc gctgacggcc gctgacagca gcacctcac cgtttagcgc acaccaccaa	33180
ttaatgtaag cagtgggaagt ttaggcttag acatggaaga ccctatgtat actcacaatg	33240
gaaaactggg aataagaatt ggggttccac taagagtagt agacagcttg catacactca	33300
ctgtagttag cggaatgga ctaactgtag ataacaatgc cctccaaact aaagttacgg	33360
gcgccttagg ttatgacaca tcaggaaatc tacaattaag agctgcagga ggtatgcgaa	33420
ttgacgcaaa tggccaactt atccttaatg tggcataccc atttgatgct cagaacaatc	33480
tcagccttag acttggtcag ggaccctgt atataaacac agaccacaac ctggatttga	33540
attgcaacag aggtctaacc acaactacca ccaacaacac aaaaaactt gagactaaaa	33600
ttagctcagg cttagactat gacaccaatg gtgctgtcat tattaaactt ggcaactggtc	33660
taagcttcca caacacaggc gccctaactg tgggaaacac tggatgatgat aaactgactc	33720
tgtggacgac cccagaccca tctccaaatt gcagaattca ctgagacaaa gactgcaagt	33780
ttactctagt cctaactaag tgtggaagcc aaatcctggc ctctgtcgcc gccctagcgg	33840
tatcaggaaa tctggcttcg ataacaggca ccgttgccag cgttaccatc tttctcagat	33900
ttgatcagaa tggagtgtt atggaaaact cctcgctaga caggcagtac tggaaacttca	33960
gaaatggcaa ctcaactaac gctgccccct acaccaatgc agttgggttc atgccaacc	34020
tcgcagcata ccccaaacg caaagccaga ctgctaaaaa caacattgta agtcaggttt	34080
acttgaatgg agacaaatcc aaacccatga cccttaccat caccctcaat ggaactaatg	34140
aatccagtga aactagccag gtgagtcact actccatgtc atttaccatg gcttgggaaa	34200
gtgggcaata tgccactgaa acctttgcca ccaactcctt cacttttct tacattgtctg	34260
aacaataaaa agcatgacac tgatgttcat ttctgattct tattttatta ttttcaaca	34320
caacaaaatc attcaagtca ttcttccatc ttagcttaat agacacagta gcttaataga	34380

cccagtagtg caaagcccca ttctagctta tagatcagac agtgataatt aaccaccacc	34440
accaccatac cttttgattc aggaaatcat gatcatcaca ggatcctagt cttcaggccg	34500
ccccctccct cccaagacac agaatacaca gtcctctccc cccgactggc tttaaataac	34560
accatctggt tggtcacaga catgttctta ggggttatat tccacacggt ctctgccgc	34620
gccaggcgct cgtcggtgat gttgataaac tctcccggca gctcgctcaa gttcacgtcg	34680
ctgtccagcg gctgaacctc cggtgacgc gataactgtg cgaccggctg ctggacaaac	34740
ggaggccgcg cctacaaggg ggtagagtca taatctcgg tcaggatagg gcggtgatgc	34800
agcagcagcg agcgaaacat ctgctgccgc cgccgctccg tccggcagga aaacaacacg	34860
ccggtggtct cctccgcgat aatccgcacc gcccgagca tcagcttctt cgttctccgc	34920
gcgcagcacc tcacctgat ctgctcaag tcggcgagc aggtacagca cagcaccacg	34980
atgttattca tgatccaca gtgcaggcg ctgtatccaa agtcatgcc gggaaccacc	35040
gccccacgt ggccatcgt ccacaagcg acgtaaatta agtgcgacc cctcatgaac	35100
gtgctggaca caaacattac ttcttgggc atgttgtaat tcaccacctc ccggtaccag	35160
ataaacctct ggtaaacag ggcaccttc accacatcc tgaaccaaga ggccagaacc	35220
tgcccaccgg ctatgcactg cagggaaccc gggttggaac aatgacaatg cagactcaa	35280
ggctcgtaac cgtggatcat ccggtcgtg aaggcatcga tgttggcaca acacagacac	35340
acgtgcatgc actttctcat gattagcagc tcttcctcg tcaggatcat atccaagga	35400
ataaccatt ctggaatcaa cgtaaaaccc acacagcagg gaaggcctcg cacataactc	35460
acgttgtgca tggtcagcg gttgcattcc ggaacacg cgatgacctc cagtatcgag	35520
gcgcgggtct cttctcaca gggaggtaaa gggtccctgc tgtacggact gcgccgggac	35580
gaccgagatc gtgttgagcg tagtgtcatg gaaaaggga cgccggacgt ggtcatactt	35640
cttgaagcag aaccaggctc gcgcgtggca ggcctccttg cgtctgcgtt ctgcgcgtct	35700
agctcgtcc gtgtgatagt ttagtacag cactcccgc agagcgtcga ggcgcaccct	35760
ggcttccgga tctatgtaga ctccgtcttg caccgggcc ctgataatat ccaccaccgt	35820
agaataagca acaccagcc aagcaataca ctcgctctgc gagcggcaga caggaggagc	35880
gggcagagat gggaagacca tgataaaaaa ctttttttaa agaattttt ccaattcttc	35940
gaaagtgaaga tctatcaagt ggcagcgctc cctccactg gcgcgggtcaa actctacggc	36000
caaagcacag acaacggcat ttctaagatg ttcttaatg gcgtccaaa gacacaccgc	36060
tctcaagtcg cagttaaacta tgaatgaaaa ccatccggc tgattttcca atatagacgc	36120

gccggcggcg tccaccaaac ccagataatt tttttctctc cagcggttta gaatctgtct	36180
aagcaaatcc cttatatcaa gtccggccat gccaaaaatc tgctcaagag cgccctccac	36240
cttcattgacc aagcagcgca tcatgattgc aaaaattcag gttcttcaga gacctgtata	36300
agattcaaaa tgggaacatt aacaaaaatt cctctgtcgc gcagatccct tcgcagggca	36360
agctgaacat aatcagacag gtctgaacgg accagtgagg ccaaatcccc accaggaacc	36420
agatccagag accctatact gattatgacg cgcatactcg gggctatgct gaccagcgta	36480
gcgccgatgt aggcgtgctg catgggcggc gagataaaat gcaaagtgtt ggttaaaaaa	36540
tcaggcaaag cctcgcgcaa aaaagctaac acatcataat catgctcatg caggtagttg	36600
caggtaagct caggaaccaa aacggaataa cacacgattt tcctctcaaa catgacttcg	36660
cggatactgc gtaaaacaaa aattataaat aaaaaattaa ttaacttaaa cattggaagc	36720
ctgtctcaca acaggaaaaa ccactttaat caacataaga cgggccacgg gcatgccggc	36780
atagccgtaa aaaaattggt ccccgtaggt aacaagtacc acagacagct ccccggtcat	36840
gtcgggggtc atcatgtgag actctgtata cacgtctgga ttgtgaacat cagacaaaca	36900
aagaaatcga gccacgtagc ccggagggtat aatcacccgc aggcggaggt acagcaaac	36960
gacccccata ggaggaatca caaaattagt aggagaaaaa aatacataaa caccagaaaa	37020
accctgttgc tgaggcaaaa tagcgcctc ccatccaaa acaacataaa gcgcttccac	37080
aggagcagcc ataacaaga cccgagttt accagtaaaa agaaaaaaga tctctcaacg	37140
cagcaccagc accaacactt cgcagtgtaa aaggccaagt gccgagagag tatatatagg	37200
aataaaaagt gacgtaaacg ggcaaagtcc aaaaaacgcc cagaaaaacc gcacgcgaac	37260
ctacgccccg aaacgaaagc caaaaaacac tagacactcc cttccggcgt caacttcgc	37320
tttccacgc tagtctactt gcccagtc aacaaactac atatccgaa cttccaagtc	37380
gccacgccc aaacaccgcc tacacctccc cgcccgccgg cccgcccga aacccgcctc	37440
ccgccccgcg ccccgccccg cgccgcccac ctcatatca tattggcttc aatccaaaat	37500
aaggtatatt attgatgatg gtttaaacgg atccaattct tgaagacgaa agggcctcgt	37560
gatacgcta ttttatagg ttaatgtcat gataataatg gtttcttaga cgtcagggtg	37620
cacttttcgg ggaatgtgc gcggaacccc tatttgitta ttttctaaa tacattcaaa	37680
tatgtatccg ctcatgagac aataacctg ataatgctt caataatatt gaaaaggaa	37740
gagtatgagt attcaacatt tccgtgtcgc ccttattccc tttttgcgg cattttcct	37800
tcctgttttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg	37860
tgacagagtg ggttacatcg aactggatct caacacgggt aagatcctg agagttttcg	37920
ccccgaagaa cgttttccaa tgatgagcac ttttaagtt ctgctatgtg gcgcggtatt	37980

atcccggtgtt gacgccgggc aagagcaact cggtcgccgc atacactatt ctacagaatga 38040

cttggttagt tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga 38100

attatgcagt gctgccataa ccatgagtga taacactgcg gccaaattac ttctgacaac 38160

gatcggagga ccgaaggagc taaccgcttt ttgcacaac atgggggatc atgtaactcg 38220

ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 38280

gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 38340

agcttcccg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct 38400

gcgctcgcc cttccggctg gctggtttat tgcgtataaa tctggagccg gtgagcgtgg 38460

gtctcgggt atcattgcag cactggggcc agatggtaag ccttcccgta tcgtagtat 38520

ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 38580

tgccctactg attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat 38640

tgatttaaaa ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatcccttaa 38700

cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 38760

gatacttttt ttctgcgct aatctgctgc ttgcaaacia aaaaaccacc gctaccagcg 38820

gtggtttgtt tgccgagta agagctacca actctttttc cgaaggtaac tggcttcagc 38880

agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag 38940

aactctgtag caccgctac atacctcgct ctgctaatec tgttaccagt ggctgctgcc 39000

agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 39060

cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac 39120

accgaactga gatacttaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga 39180

aaggcggaca ggtatccgtt aagcggcagg gtcggaacag gagagcgac gagggagctt 39240

ccagggggaa acgcctgcta tctttatagt cctgtcgggt ttgccacct ctgacttgag 39300

cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg 39360

gcctttttac ggttcttgcc cttttgctgg ccttgaagct gtcctgatg gtcgtcatct 39420

acctgcctgg acagcatggc ctgcaacgcg ggcatccga tgccgccgga agcgagaaga 39480

atcataatgg ggaaggccat ccagcctcgc gtcgcagatc cgaattcggt taaac 39535

<210> 16

<211> 1500

<212> DNA

<213> Human immunodeficiency virus

<400> 16

atgggtgcta gggcttctgt gctgtctggt ggtgagctgg acaagtggga gaagatcagg 60
ctgaggcctg gtggcaagaa gaagtacaag ctaaagcaca ttgtgtgggc ctccaggag 120
ctggagaggt ttctgtgaa ccctggcctg ctggagacct ctgaggggtg caggcagatc 180
ctgggccagc tccagccctc cctgcaaaca ggctctgagg agctgaggtc cctgtacaac 240
acagtggcta ccctgtactg tgtgcaccag aagattgatg tgaaggacac caaggaggcc 300

ctggagaaga ttgaggagga gcagaacaag tccaagaaga aggccagca ggctgctgct 360
ggcacaggca actccagcca ggtgtcccag aactaccca ttgtgcagaa cctccagggc 420
cagatggtgc accaggccat ctcccccg accctgaatg cctgggtgaa ggtggtggag 480
gagaaggcct tctccctga ggtgatcccc atgttctctg ccctgtctga ggggtgccacc 540
ccccaggacc tgaacacat gctgaacaca gtggggggcc atcaggctgc catgcagatg 600
ctgaaggaga ccatcaatga ggaggtgct gagtgggaca ggctgcatcc tgtgcacgt 660
ggccccattg cccccggcca gatgaggag cccaggggct ctgacattgc tggcaccacc 720

tccaccctcc aggagcagat tggctggatg accaacaacc ccccatccc tgtgggggaa 780
atctacaaga ggtggatcat cctgggcctg aacaagattg tgaggatgta ctccccacc 840
tccatcctgg acatcaggca gggccccaag gagccctca gggactatgt ggacaggttc 900
tacaagacc tgagggtga gcaggcctcc caggaggtga agaactggat gacagagacc 960
ctgctggtgc agaatgcaa ccctgactgc aagaccatcc tgaaggccct gggccctgct 1020
gccaccctgg aggagatgat gacagcctgc cagggggtgg ggggccctgg tcacaaggcc 1080
agggtgctgg ctgaggccat gtcccagtg accaactccg ccaccatcat gatgcagagg 1140

ggcaacttca ggaaccagag gaagacagt aagtgttca actgtggcaa ggtgggccac 1200
attgccaaga actgtaggcc cccaggaag aagggtgct ggaagtgtgg caaggagggc 1260
caccagatga aggactgcaa tgagaggcag gccaacttcc tgggcaaat ctggccctcc 1320
cacaaggga ggcttgcaa ctctctccag tccaggcctg agcccacagc cctcccag 1380
gagtccttca ggtttggga ggagaagacc acccccagcc agaagcagga gccattgac 1440
aaggagctgt accccctggc ctccctgagg tcctgtttg gcaacgacc ctctcccag 1500

<210> 17

<211> 39240

<212> DNA

<213> Pan troglodytes

<220><221> misc_feature

<222> (24096)..(24098)

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

<400> 17

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcgggcggg gaggcgggtt tggggcgggg ccggcgggcg gggcgggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgccccg ggaagtgaca tttttccgc ggtttttacc	240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggcgg attacgtgga ggactcggc aggtgttttt tgaggagaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt	480
gtatccatat cataatatgt acatttataat tggctcatgt ccaacattac cgccatgttg	540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc	600
atataatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccac	660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caataggagc	720
tttccattga cgicaatggg tggagtattt acggtaaaact gcccacttgg cagtacatca	780
agtgtatcat atccaagta cggcccctat tgacgtcaat gacggtaaat ggcccgcctg	840
gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca tctacgtatt	900
agtcacgcgt attaccatgg tgatgcggtt ttggcagtag atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg	1020
gcacaaaaat caacgggact ttcaaaaatg tcgtaacaac tccgccccat tgacgcaaat	1080
gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctcccta tcagtatag	1140
agatctccct atcagtata gagatcgtcg acgagctcgt ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggccgg gaacggtgca ttggaacgcg gattccccgt gccaagagtg agatctacca	1320
tgggtgctag ggcttctgtg ctgtctggtg gtgagctgga caagtgggag aagatcaggc	1380
tgaggcctgg tggcaagaag aagtacaagc taaagcacat tgtgtgggcc tccagggagc	1440
tggagaggtt tgcgtgaac cctggcctgc tggagacctc tgagggtgc aggcagatcc	1500
tgggccagct ccagccctcc ctgcaaacag gctctgagga gctgaggtcc ctgtacaaca	1560
cagtggctac cctgtactgt gtgcaccaga agattgatgt gaaggacacc aaggaggccc	1620
tggagaagat tgaggaggag cagaacaagt ccaagaagaa ggcccagcag gctgctgctg	1680
gcacaggcaa ctccagccag gtgtcccaga actaccccat tgtgcagaac ctccagggcc	1740

agatggtgca ccaggccatc tccccccgga ccttgaatgc ctgggtgaag gtggtggagg	1800
agaaggcctt ctccctgag gtgatccca tgttctctgc cctgtctgag ggtgccaccc	1860
cccaggacct gaacaccatg ctgaacacag tggggggcca tcaggctgcc atgcagatgc	1920
tgaaggagac catcaatgag gaggtgctg agtgggacag gctgcatcct gtgcacgctg	1980
gccccattgc ccccgccag atgagggagc ccaggggctc tgacattgct ggcaccacct	2040
ccaccctcca ggagcagatt ggctggatga ccaacaaccc ccccatccct gtgggggaaa	2100
tctacaagag gtggatcatc ctgggcctga acaagattgt gaggatgtac tccccacct	2160
ccatcctgga catcaggcag ggccccaagg agcccttcag ggactatgtg gacaggttct	2220
acaagaccct gagggctgag caggcctccc aggaggtgaa gaactggatg acagagaccc	2280
tgctggtgca gaatgccaac cctgactgca agaccatcct gaaggccctg ggccctgctg	2340
ccaccctgga ggagatgatg acagcctgcc aggggggtggg gggccctggt cacaaggcca	2400
gggtgctggc tgaggccatg tcccagggtga ccaactccgc caccatcatg atgcagaggg	2460
gcaacttcag gaaccagagg aagacagtga agtgcttcaa ctgtggcaag gtgggccaca	2520
ttgccaagaa ctgtagggcc cccaggaaga agggctgctg gaagtgtggc aaggagggcc	2580
accagatgaa ggactgcaat gagaggcagg ccaacttcct gggcaaaatc tggccctccc	2640
acaagggcag gcctggcaac ttctccagt ccaggcctga gcccacagcc cctcccagg	2700
agtccttcag gtttggggag gagaagacca ccccagcca gaagcaggag cccattgaca	2760
aggagctgta cccctggcc tcctgaggt cctgtttgg caacgacccc tcctccagt	2820
aaaataaagc ccgggcagat ctgatctgct gtgccttcta gttgccagcc atctgttgtt	2880
tgccctccc ccgtgccttc cttgacctg gaaggtgcca ctccactgt cctttcctaa	2940
taaaatgagg aaattgcatc gcattgtctg agtaggtgtc attctattct ggggggtggg	3000
gtggggcagc acagcaaggg ggaggattgg gaagacaata gcaggcatgc tggggatgcg	3060
gtgggctcta tggccgatca gcgatcgctg aggtgggtga gtgggcgtgg cctggggtgg	3120
tcatgaaaat atataagttg ggggtcttag ggtctcttta tttgtgttgc agagaccgcc	3180
ggagccatga gcgggagcag cagcagcagc agtagcagca gcgccttgga tggcagcatc	3240
gtgagccctt atttgacgac gcggatgccc cactgggccg gggtgcgtca gaatgtgatg	3300
ggctccagca tcgacggccg accgtctctg ccgcgaaatt ccgccacgt gacctatgcg	3360
accgtcgcg ggacgccgtt ggacgccacc gccgccgcc cgcgccaccg agccgcctcg	3420
gccgtgcgca gcctggccac ggactttgca ttcttgggac cactggcgac aggggctact	3480
tctcgggccg ctgtgcccgc cgttcgcgat gacaagctga ccgcctgct ggcgagttg	3540
gatgcgtta ctcggaact gggtagacct tctcagcagg tcatggccct gcgccagcag	3600

gtctctctccc tgcaagctgg cggaatgct tctcccacaa atgccgttta agataaataa	3660
aaccagactc tgtttgatt aaagaaaagt agcaagtga ttgctctctt tatttcataa	3720
ttttccgcgc gcgataggcc ctagaccagc gttctcggtc gttgagggtg cgggtgtatct	3780
tctccaggac gtggtagagg tggctctgga cgttgagata catgggcatg agcccgctccc	3840
gggggtggag gtagcaccac tgcagagctt catgctccgg ggtggtgttg tagatgatcc	3900
agtcgtagca ggagcgctgg gcatggtgcc taaaaatgtc cttcagcagc aggccgatgg	3960
ccagggggag gcccttggtg taagtgttta caaacgggtt aagttgggaa ggggtgcattc	4020
ggggagagat gatgtgcatc ttggactgta tttttagatt ggcgatgttt ccgccagat	4080
cccttctggg attcatgttg tgcaggacca ccagtacagt gtatccggtg cacttgggga	4140
atttgtcatg cagcttagag ggaaaagcgt ggaagaactt ggagacgccc ttgtggcctc	4200
ccagattttc catgcattcg tccatgatga tggcaatggg ccgcggggag gcagcttggg	4260
caaagatatt tctggggctg ctgacgtcgt agttgtgttc cagggtgagg tcgtcatagg	4320
ccatttttac aaagcgcggg cggaggggtgc ccgactgggg gatgatggtc ccctctggcc	4380
ccggggcgta gttgcctcg cagatctgca tttcccaggc cttaatctcg gaggggggaa	4440
tcataccac ctgcggggcg atgaagaaaa cggtttccgg agccggggag attaactggg	4500
atgagagcag gtttctaagc agctgtgatt ttccacaacc ggtgggcca taaataacac	4560
ctataaccgg ttgcagctgg tagtttagag agctgcagct gccgtcgtcc cggaggaggg	4620
gggccacctc gttgagcatg tccctgacgc gcatgttctc ccgaccaga tccgccagaa	4680
ggcgctcgcc gccagggac agcagctctt gcaaggaagc aaagttttcc agcggttga	4740
ggccgtccgc cgtgggcatg ttttcagggt tctggctcag cagctccagg cgggtccaga	4800
gctcggtagc gtgctctacg gcatctctat ccagcatatc tctcgtttc gcgggttggg	4860
gcgactttcg ctgtaggga ccaagcgggt gtcgtccagc ggggccagag tcatgtcctt	4920
ccatgggcgc agggctctcg tcagggtgtt ctgggtcacg gtgaagggtt gcgtccggg	4980
ctgagcgctt gccaaagtgc gcttgaggct ggttctgctg gtgctgaagc gctgccggtc	5040
ttcgccctgc gcgtcggcca ggtagcattt gaccatggtg tcatagtcca gccctccgc	5100
ggcgtgtccc ttggcgcgca gcttgcctt ggaggtggcg ccgcacagg ggcagagcag	5160
gctcttgagc gcgtagagct tggggggcag gaagaccgat tcgggggagt aggcgtccgc	5220
gccgcagacc ccgcacacgg tctcgcactc caccagccag gtgagctcgg ggcgcgccgg	5280
gtcaaaaacc aggtttcccc catgcttttt gatgcgtttc ttacctcggg tctccatgag	5340

gtggtgtccc cgctcgggtga cgaagaggct gtccgtgtct ccgtagaccg acttgagggg 5400

tcttttctcc aggggggtcc ctccgtcttc ctctagagg aactcggacc actctgagac 5460

gaaggcccg cgtccaggcca ggacgaagga ggctatgtgg gaggggtagc ggtcgttgtc 5520

cactaggggg tccaccttct ccaagggtgtg aagacacatg tcgccttctt cggcgtccag 5580

gaaggtgatt ggctttagg tgtaggccac gtgaccgggg gttcctgacg ggggggtata 5640

aaaggggtg ggggcgcgt ctctgtcact ctcttcgca tcgtgtctg cgaggccag 5700

ctgtgggggt gatlattccc tctgaaggc ggcatgacc tccgcgtga ggttgtcagt 5760

ttccaaaaac gaggaggatt tgatgttcac ctgtcccag gtgatactt tgagggtacc 5820

cgctccatc tggtcagaaa acacgatctt ttattgtcc agcttgggtg cgaacgaccc 5880

gtagaggcg ttggagagca gcttggcgt ggagcgcagg gtctggttct tgtccctgtc 5940

ggcgcgtcc ttggccgca tgttagctg cactactcg cgcgcgacgc agcgcactc 6000

ggggaagacg gtgtgctc ctctgggcac caggcgcacg cgcagccgc ggttgtgcag 6060

ggtgaccagg tccacgttg ttggcacctc gccgcgcagg cgctcgttg tccagcagag 6120

acggccgccc ttgcgcgagc agaagggggg cagggggtcg agctgggtct cgtccggggg 6180

gtccgcgtcc acggtgaaaa ccccggggcg caggcgcgcg tcgaagtagt ctatcttgca 6240

accttgcatg tccagcgct gctgccagt cggggcggcg agcgcgcgt cgtaggggtt 6300

gagcggcggg cccagggca tggggtgggt gattgcggag gcgtacatgc cgcagatgtc 6360

atagacgtag aggggtccc gcaggacccc gatgtagggt gggtagcagc ggcccccgcg 6420

gatgctggcg cgcacgtagt catacagctc gtgcgagggg gcgaggaggt cggggcccag 6480

gttgggtcgg gcggggcgt ccgtgcgga gacgatctgc ctgaagatgg catgcgagtt 6540

ggaagagatg gtggggcgt ggaagacgtt gaagctggcg tcctgcagc cgacggcgtc 6600

gcgcacaag gaggcgtagg agtcgcgcag cttgtgtacc agctcggcgg tgacctgcac 6660

gtcagcgcg cagtagtca gggctctcg gatgatgtca tattagcct gccccctctt 6720

tttccacagc tcgcggtga ggacaaactc ttgcggtct ttccagtact cttggatcgg 6780

gaaaccgtcc ggttccgaac ggtaagagcc tagcatgtag aactggttga cggcctggta 6840

ggcgcagcag ccttttcca cggggagggc gtaggcctgc gcggccttgc ggagcgaggt 6900

gtgggtcagg gcgaagggt ccttgacat gactttgagg tactggtgt tgaagtcgga 6960

gtcgtgcag ccgccccgt cccagagcga gaagtcggtg cgcttcttg agcgggggtt 7020

gggcagagcg aaggtgacat cgttgaagag gattttgcc gcgcggggca tgaagttgcg 7080

ggtgatgcgg aaggccccg gcacttcaga gcggttggtg atgacctggg cggcgagcac 7140
 gatctcgtcg aagccgttga tgtttgtggc cacgatgtag agttccagga agcggggccg 7200
 gccctttacg gtgggcagct tctttagctc ttcgtagggt agctcctcgg gcgagcgag 7260
 gccgtgctcg gccagggccc agtccgcgag gtgcgggttg tctctgagga aggactccca 7320
 gaggtcgcgg gccaggaggg tctgcaggcg gtcctgaag gtccgaact ggcggccac 7380
 ggccatTTTT tcgggggtga tgcagtagaa ggtgaggggg tcttctgcc agcggccca 7440
 gtcgagctgc agggcgaggt cgcgcgcggc ggtgaccagg cgctcgtcgc cccgaattt 7500

 catgaccagc atgaaggga cgagctgctt tccgaaggcc cccatccaag ttaggtctc 7560
 tacatcgtag gtgacaaaga ggcgtccgt gcgaggatgc gagccgatcg ggaagaactg 7620
 gatctccgc caccagtgg aggagtggct gttgatgtgg tggaagtaga agtccgtcg 7680
 ccgggccgaa cactcgtgct ggcttttga aaagcgagcg cagtactggc agcgtgcac 7740
 gggctgtacc tcatgcacga gatgcacctt tcgccgcgc acgaggaagc cgaggggaaa 7800
 tctgagcccc ccgctggct cgcggcatgg ctggtgctct tctactttgg atcgtgtcc 7860
 gtctcgtct ggctcctga ggggtgttac ggtggagcgg accaccacgc cgcgcgagcc 7920

 gcaggtccag atacggcgc gcggcggtcg gagtttgatg acgacatcgc gcagctggga 7980
 gctgtccatg gtctggagct cccgcggcgg cggcaggtca gccgggagtt ctgacagtt 8040
 cacctcgag agtcgggcca gggcgcgggg caggtctagg tggtaacctga tctctagggg 8100
 cgtgttggtg gcggcgtcga tggcttcag gagcccgag ccccgggggg cgacacggt 8160
 gcccccggg gtggtggtg tgggtggcgt gcagctcaga agcggtgccg cggcggggc 8220
 cccggaggta gggggggctc cggtcccgc ggaggggagc gcagcggcac gtcggcgtgg 8280
 agcgcgggca ggagtgggt ctgtccccg aggttgctgg cgaaggcgac gacgcggcgg 8340

 ttgatctcct ggatctggcg cctctgcgtg aagacgacgg gcccggtgag ctgaaacctg 8400
 aaagagagtt cgacagaatc aatctcgtg tcattgaccg cggcctggcg caggatctcc 8460
 tgcacgtctc ccgagttgtc ttggtaggcg atctcgcca tgaactgctc gatctctcc 8520
 tcctggaggt ctccgcgtcc ggcgcttcc acggtggcgg ccaggtcgtt ggagatgcgc 8580
 cccatgagct gcgagaaggc gttgagtcg cctcgttcc agactcggct gtagaccacg 8640
 cccccctggt catcggggc gcgcatgacc acctgcgca ggttgagctc cacgtgccgc 8700
 gcgaagacgg cgtagttgcg cagacgtgg aagaggtagt tgagggtgtt ggcggtgtgc 8760

 tcggccacga agaagttcat gaccagcgg cgcaacgtgg attcgttgat gtccccaaag 8820
 gcctccagcc gttccatggc ctctagaag tccacggcga agttgaaaaa ctgggagttg 8880
 cgcgccgaca cgtcaactc ctctccaga agacggatga gctcggcgac ggtgtcgcgc 8940

acctcgcgct cgaaggctat ggggatctct tcctccgcta gcatcaccac ctctctctct	9000
tcctctctct ctggcacttc catgatggct tcctctctct cgggggggtgg cggcggcggc	9060
ggtgggggag ggggcgctct gcgccggcgg cggcgccaccg ggaggcggtc cacgaagcgc	9120
gcgatcatct ccccgcgggc gcggcgcatg gtctcggtag cggcgcgggc gtctctccgg	9180
gggcgcagtt ggaagacgcc gccggacatc tggtagctggg gcgggtggcc gtgaggcagc	9240
gagacggcgc tgacgatgca tctcaacaat tgctgcgtag gtacgccgcc gagggacctg	9300
agggagtcca tatccaccgg atccgaaaac ctttcgagga aggcgtctaa ccagtgcag	9360
tcgaaggta ggctgagcac cgtggcgggc ggccgggggt ggggggagtg tctggcggag	9420
gtgctgctga tgatgtaatt gaagtaggcg gacttgacac ggcggtatgt cgacaggagc	9480
accatgtcct tgggtccggc ctgctggatg cggaggcggg cggctatgcc ccaggcttcg	9540
ttctggcatc ggcgagggtc cttgtagtag tcttgcata gcctttccac cggcacctct	9600
tcctctctct ctctgcttc tccatgtct gcttcggccc tggggcggcg ccgcgcccc	9660
ctgccccca tgcgcgtgac cccgaacccc ctgagcgggt ggagcagggc caggctggcg	9720
acgacgcgct cggccaggat ggcctgctgc acctgcgtga gggtagtttg gaagtcatcc	9780
aagtccacga agcgggtgta ggcccccgtg ttgatggtgt aggtgcagtt ggccatgacg	9840
gaccagttag cggctctggt gcccggttgc gacatctcgg tgtacctgag tcgcgagtag	9900
gcgcgggagt cgaagacgta gtcgttgcaa gtccgcacca ggtactggtg gccaccagg	9960
aagtgcggcg gcggctggcg gtagaggggc cagcgagggg tggcgggggc tccggggggc	10020
aggctctcca gcatgaggcg gtggtaggcg tagatgtacc tggacatcca ggtgataccc	10080
gcggcgggtg tggaggcgcg cgggaagtcg cgacccgggt tccagatgtt gcgcaggggc	10140
agaaagtgt ccatggtagg cgtgtctgtt ccagtcagac gcgcgcagtc gttgatactc	10200
tagaccaggg aaaacgaaag ccggtcagcg ggcactcttc cgtggtctgg tgaatagatc	10260
gcaagggtat catggcggag ggcctcgggt cgagccccgg gtccgggccg gacggtccgc	10320
catgatccac gcggttaccg cccgcgtgtc gaaccagggt gtgcgacgtc agacaacggt	10380
ggagtgttcc ttttggcgtt tttctggccg ggccggggcg ccgcgtaaga gactaagcgc	10440
cgaagcgaa agcagtaagt ggctcgtcc ccgtagccgg agggatcctt gctaagggtt	10500
gcgttggcg gaaccccggt tcgaatccg tactcgggcc ggccggaccg gcggctaagg	10560
tggttgattg gcctccccct cgtataaaga ccccgcttgc ggattgactc cggacacggg	10620
gacgagcccc ttttatTTTT gctttccca gatgcatccg gtgttgcgac agatgcgcc	10680
cccgccccag cagcagcaac aacaccagca agagcggcag caacagcagc gggagtcatg	10740
cagggcccc tcacccacc tcggcggccc ggccacctcg gcgtccgcgg ccgtgtctgg	10800

cgcttgcggc ggccggcggcg gggggccggc tgacgacccc gaggagcccc cgccggcgag	10860
ggccagacac tacctggacc tggaggaggc cgagggcctg gcgcggctgg gggcgccgtc	10920
tcccggagcg caccgcggg tgcagctaaa gcgcgactcg cgcgaggcgt acgtgcctcg	10980
gcagaacctg ttcaggacc gcgcggcgga ggagcccag gagatgcggg acaggaggtt	11040
cagcgcgggg cgggagctgc ggcaggggct gaaccgcgag cggctgctgc gcgaggagga	11100
ctttgagccc gacgcgcgga cggggatcag cccgcgcgc gcgcacgtgg cggccgccga	11160
cctggtgacg gcgtacgagc agacggtgaa ccaggagatc aacttccaaa agagtttcaa	11220
caaccacgtg cgcacgtgg tggcgcgca ggaggtgacc atcgggctga tgcacctgtg	11280
ggactttgtg agcgcgctgg tgcagaacct caatagcaag cctctgacgg cgcagctgtt	11340
cctgatagtg cagcacagca gggacaacga ggcgtttagg gacgcgctgc tgaacatcac	11400
cgagcccag ggccggtggc tgctggacct gattaacatc ctgcagagca tagtggtgca	11460
ggagcgcagc ctgagcctgg ccgacaaggt ggcggccatc aactactcga tgctgagcct	11520
gggcaagttt tacgcgcgca agatctacca gacgccgtac gtgcccatac acaaggaggt	11580
gaagatgcac ggtttttaca tgccgatggc gctgaagggt ctcaccctaa gcgacgacct	11640
gggcgtgtac cgcaacgagc gcatccacaa ggccgtgagc gtgagccggc ggcgcgagct	11700
gagcgaccgc gagctgatgc atagcctgca gcgggcgctg gcgggcgcgc gcagcggcga	11760
cagggaggcg gagtctact tcgatgcggg ggcggacctg cgctggcgcc ccagccggcg	11820
ggccctggag gccgcggggg tccgcgagga ctatgacgag gacggcgagg aggatgagga	11880
gtacgagcta gaggaggcg agtacctgga ctaaaccgcg ggtggtgttt ccggtagatg	11940
caagaccga acgtggtgga cccgcgcgtg cggcgggctc tgacagagca gccgtccggc	12000
cttaactcct cagacgactg gcgacaggtc atggaccgca tcatgtcgct gacggcgct	12060
aaccggagc cgttccggca gcagccgag gccaacaggc tctccgcat cctggaggcg	12120
gtggtgcctg cgcgctcga cccacgcac gagaagggtg tggccatagt gaacgcgctg	12180
gccgagaaca gggccatccg cccggacgag gccgggctgg tgtacgacgc gctgctgcag	12240
cgctggccc gctacaacag cggcaacgtg cagaccaacc tggaccggt ggtgggggac	12300
gtgcgcgagg cgtggcgca gcgcgagcgc gcggatcggc agggcaacct gggctccatg	12360
gtggcgctga atgccttct gagcacgag ccggccaacg tgcccgggg gcaggaagac	12420
tacaccaact ttgtgagcg gctgcggctg atggtgaccg agacccccca gagcgaggtg	12480
taccagtcgg gtccgacta cttcttcag accagcagac agggcctgca gacggtgaac	12540

ctgagccagg ctttcaagaa cctgcggggg ctgtggggcg tgaaggcgcc caccggcgac	12600
cgggcgacgg tgtccagcct gctgacgccc aactcgcgcc tgctgctgct gctgatcgcg	12660
ccgttcacgg acagcggcag cgtgtcccgg gacacctacc tggggcacct gctgaccctg	12720
taccgcgagg ccatcgggca ggcgaggtg gacgagcaca ctttccagga gatcaccagc	12780
gttagccgcg cgctggggca ggaggacacg agcagcctgg aggcgactct gaactacctg	12840
ctgaccaacc ggcggcagaa gattccctcg ctgcacagcc tgacctccga ggaggagcgc	12900
atcttgcgct acgtgcagca gagcgtgagc ctgaacctga tgcgcgacgg ggtgacgccc	12960
agtgtggcgc tggacatgac cgcgcgcaac atggaaccgg gcatgtacgc cgcgaccgg	13020
ccttacatca accgcctgat ggactacctg catcgcgagg cgcccgtaaa ccccagtagc	13080
tttaccacg ccatcctgaa cccgactgg ctcccgcgc cggggttcta cagcgggggc	13140
ttcgaggtcc cggaggccaa cgatggcttc ctgtgggacg acatggacga cagcgtgttc	13200
tcccccggc cgaggcgct ggcggaagcg tccctgctgc gtccaagaa ggaggaggag	13260
gcgagtcgcc gccgcggcag cagcggcgtg gcttctctgt ccgagctggg ggcggcagcc	13320
gccgcgcgcc cgggtccct gggcggcagc cctttccga gcctggtggg gtctctgcac	13380
agcgagcgca ccacccgccc tcggctgctg ggcgaggacg agtacctgaa taactccctg	13440
ctgcagccgg tgcgggagaa aaacctgcct ccgccttcc ccaacaacgg gatagagagc	13500
ctggtggaca agatgagcag atggaagacc tatgcgcagg agcacaggga cgcgcccgcg	13560
ctccggccgc ccacgcggcg ccagcgccac gaccggcagc gggggctggt gtgggatgac	13620
gaggactccg cggacgatag cagcgtgctg gacctgggag ggagcggcaa cccgttcgcg	13680
cacctgcgcc cccgcctggg gaggatgttt taataaaaaa aaaaagcaag aagcatgatg	13740
caaaaattaa ataaactca ccaaggccat ggcgaccgag cgttggtttc ttgtgttccc	13800
ttcagtatgc ggcgcgcggc gatgtaccag gagggacctc ctcctctta cgagagcgtg	13860
gtgggcgcgg cggcggcggc gccctcttct cctttgctg cgagctgct ggagccgccg	13920
tacgtgcctc cgcgctacct gcggcctacg ggggggagaa acagcatccg ttactcgag	13980
ctggcgcgcc tgttcgacac caccgggtg tacttggtgg acaacaagtc ggcgacgtg	14040
gcctccctga actaccagaa cgaccacagc aattttttga ccacggtcat ccagaacaat	14100
gactacagcc cgagcgaggc cagcaccag accatcaatc tggatgaccg gtcgactgg	14160
ggcggcgacc tgaataacat cctgcacacc aacatgccca acgtgaacga gttcatgttc	14220
accaataagt tcaaggcgcg ggtgatggtg tcgcgctcgc acaccaagga agaccgggtg	14280
gagctgaagt acgagtgggt ggagttcgag ctgccagagg gcaactactc cgagaccatg	14340
accattgacc tgatgaacaa cgcgatcgtg gagcactatc tgaaagtggg caggcagaac	14400

ggggtcctgg agagcgacat cggggtcaag ttcgacacca ggaacttccg cctggggctg	14460
gacccctga ccgggtggt tatgcccggg gtgtacacca acgaggcctt ccatcccgac	14520
atcatcctgc tgcccggctg cggggtggac ttcaactaca gccgcctgag caacctcctg	14580
ggcatccgca agcggcagcc cttccaggag ggcttcagga tcacctacga ggacctggag	14640
gggggcaaca tccccgcgt cctcgatgtg gaggcctacc aggatagctt gaaggaaaat	14700
gaggcgggac aggaggat acgccccgcc gcctccgccg ccgcccagca gggcgaggat	14760
gctgctgaca ccgcgccgc ggacggggcg gagccgacc ccgctatggt ggtggaggct	14820
gccgagcagg aggaggacat gaatgacagt gcggtgcgcg gagacacctt cgtcacccgg	14880
ggggaggaaa agcaagcgga ggccgaggcc gcggccgagg aaaagcaact ggccgagca	14940
gcggcggcgg cgcggttggc cgccgaggag gctgagtctg aggggaccaa gcccaccaag	15000
gagccctga ttaagccct gaccgaagat agcaagaagc gcagttacaa cctgctcaag	15060
gacagcacca acaccgcgt ccgcagctgg tacctggcct acaactacgg cgaccgtcg	15120
acgggggtgc gtcctggac cctgctgtgc acgccgacg tgacctgcgg ctccgagcag	15180
gtgtactggt cgctgccga catgatgcaa gacccgtga ccttcgcgc cacgcggcag	15240
gtcagcaact tcccggtggt gggcgccgag ctgctgccg tgactccaa gagcttctac	15300
aacgaccagg ccgtctact ccagctcat cgccagtca cctctctgac ccacgtgttc	15360
aatcgctttc ctgagaacca gattctggcg cgcccgccg cccccacat caccacgtc	15420
agtgaacacg ttctgtctt cacagatcac gggacgtac cgctgcgcaa cagcatcgga	15480
ggagtccagc gagtgacct tactgacgcc agacgccga cctgccccta cgtttacaag	15540
gccttgggca tagtctgcc gcgcgtcctt tcagccgca ctttttgagc aacaccacca	15600
tcatgtccat cctgatctca ccagcaata actccgctg gggactgtg cgcgcgcca	15660
gcaagatgtt cggaggggag aggaagcgtt ccgagcagca cccgtgcgc gtgcgcgggc	15720
acttcgcgc cccctgggga gcgcacaaac gcggccgcgc gggcgccacc accgtggacg	15780
acgccatga ctcggtggtg gagcaggcgc gcaactacag gcccgcggtc tctaccgtg	15840
acgcggccat ccagaccgtg gtgcggggcg cgcggcggta cgccaagctg aagagccgcc	15900
ggaagcgcgt ggcccgcgc caccgccgc gacccggggc cgccgcaaa cgcgcgcgc	15960
cggccctgct tcgcccggc aagcgacgg gccgcgcgc cgccatagg gccgcgcgc	16020
gcttggcgc cggcatcacc gccgccacca tggcccccg taccgaaga cgcgcggccg	16080
ccgcgcgcgc cgccgccatc agtgacatgg ccagcaggcg ccggggcaac gtgtactggg	16140
tgcgcgactc ggtgaccggc acgcgcgtgc ccgtgcgtt cgcccccg cggacttgag	16200
atgatgtgaa aaaacaacac tgagtctcct gctgttgtgt gtatcccagc ggccggcg	16260

gcgcgcgcag cgtcatgtcc aagcgcaaaa tcaaagaaga gatgctccag gtcgtcgcgc 16320

cggagatcta tgggcccccg aagaaggaag agcaggattc gaagccccgc aagataaagc 16380

gggtcaaaaa gaaaaagaaa gatgatggcg atgccgatgg ggaggtggag ttcctgcgcg 16440

ccacggcgcc caggcgcccc gtgcagtga agggccggcg cgtaaagcgc gtcctgcgcc 16500

ccggcaccgc ggtggtcttc acgcccggcg agcgtccac ccggactttc aagcgcgtct 16560

atgacgaggt gtacggcgac gaagacctgc tggagcaggc caacgagcgc ttcggagagt 16620

ttgcttacgg gaagcgtcag cggccgctgg ggaaggagga cctgctggcg ctgccgctgg 16680

accaggggcaa cccaccccc agtctgaagc ccgtgacct gcagcaggtg ctgccgagca 16740

gcgcaccctc cgaggcgaag cggggtctga agcgcgaggg cggcgacctg gcgcccaccg 16800

tgcagctcat ggtgcccaag cggcagaggc tggaggatgt gctggagaaa atgaaagtag 16860

accccggtct gcagccggac atcagggtcc gtcccatcaa gcaggtggcg cggggcctcg 16920

gcgtgcagac cgtggacgtg gtcatccca ccggcaactc ccccgccgc accaccacta 16980

ccgtgcctc caccgacatg gagacacaga ccgatccgc gcgagccgc gccaccgccg 17040

ccgccgcgac ctctcggcg gagggtcaga cggaccctg gctgccgcg gcgatgtcag 17100

ctccccgcgc gcgtcgcggg cgcaggaagt acggcgccgc caacgcgctc ctgcccgagt 17160

acgccttgca tccttcac cgcgccacc ccggtaccg aggtatacc taccgccgc 17220

gaagagccaa ggtttccacc cgccgtccc gccgacgcgc cgccgccacc acccgccgc 17280

gccgccgcag accccagccc gcactggctc cagtctccgt gaggagagt gcgcgcgacg 17340

gacacacctt ggtgctgccc agggcgcgct accaccccag catcgtttaa aagcctgttg 17400

tggttcttgc agatatggcc ctacttgcc gcctccgttt cccggtgccg ggataccgag 17460

gaggaagatc gcgccgagg aggggtctgg ccggccgcgg cctgagcgga ggcagccgc 17520

gcgcgcaccg gcggcgacgc gccaccagcc gacgcatgcg cggcgggggtg ctgccctgt 17580

taatccccct gatcgccgcg gcgatcggcg ccgtgcccg gatcgctcc gtggccttgc 17640

aggcgtccca gaggcattga cagacttgca aacttgcaaa tatggaaaa aacccaata 17700

aaaaagtcta gactctcacg ctgccttgg cctgtgacta tttttagaa tggagacat 17760

caactttgcg tcgctggccc cgcgtcacgg ctgcgcccc ttcctgggac actggaacga 17820

tatcggcacc agcaacatga gcggtggcg cttcagttgg ggctctctgt ggagcggcat 17880

taaaagtatc ggtctgcgg ttaaaaattt cggctcccgg gcctggaaca gcagcacggg 17940

ccagatgttg agagacaagt tgaaagagca gaacttccag cagaaggtgg tggaggcct 18000

ggcctccggc atcaacgggg tggaggacct ggccaaccag gccgtgcaga ataagatcaa	18060
cagcagactg gacccccggc cgccggtgga ggaggtgccg ccggcgctgg agacggtgtc	18120
ccccgatggg cgtggcgaga agcgccccgc gcccgatagg gaagagacca ctctggtcac	18180
gcagaccgat gagccgcccc cgtatgagga ggccctgaag caaggtctgc ccaccacgcg	18240
gccccatcgc ccatggcca ccggggtggt gggccgccac acccccgcca cgtggactt	18300
gcctccggcc gccgatgtgc cgcagcagca gcagaaggcg gcacagccgg gcccgccgt	18360
gaccgcctcc cgttcctccg ccggtcctct gcgccgcgcg gccagcggcc cccgcggggg	18420
ggtcgcgagg cacggcaact ggcagagcac gctgaacagc atcgtgggtc tgggggtgcg	18480
gtccgtgaag cgccgccgat gctactgaat agcttagcta acgtgttgta tgtgtgtatg	18540
cgccctatgt cgccgccaga ggagctgctg agtcgccgcg gttcgcgcgc ccaccaccac	18600
caccgccact cgcccccta agatggcgac cccatcgatg atgccgcagt ggtcgtacat	18660
gcacatctcg ggccaggacg cctcgagta cctgagcccc gggctggtgc agttcgcccc	18720
cgccaccgag agctacttca gcctgagtaa caagtttagg aacccacagg tggcgccac	18780
gcacgatgtg accaccgacc ggtctcagcg cctgacgtg cggttcattc ccgtggaccg	18840
cgaggacacc gcgtactcgt acaaggcgcg gttcacctg gccgtggcg acaaccggt	18900
gctggacatg gcctccact actttgacat ccgcggggtg ctggaccggg gccccacttt	18960
caagccttac tctggcaccg cctacaactc cctggcccc aaggcgctc ccaactcctg	19020
cgagtgggag caattagaag aagcccaggc cgctgtggaa gacgaagaat tagaagatga	19080
agacaggaa ccacaggatg aggcacctgt gaaaaaac catgtatacg ctcaggctcc	19140
cctttctgga gaagaaatta ctaaaacgg ttgcaaata gggtcagata acacagaagc	19200
ccagtctaag ccatatatg cagatcctac attccagcct gaaccccaa tcggggaatc	19260
ccagtggaat gaggcagatg ctacagttgc cgccggtaga gtgctaaaga aatccactcc	19320
catgaagcca tgctatggtt cctatgcaag acccacaac tccaatggag gtcaaggtgt	19380
gctggtggct gatgataagg gggttcttca atctaaagtt gaattgcaat ttttttcaa	19440
tactactact cttaatcagc gggagggtaa cgatacaaaa ccaaaagtgg tgctgtatag	19500
cgaagatgtg cacatgaaa ctccagacac ccacatttct tacaagccca caaaaagcga	19560
tgacaattca aaaaatcgc tgggtcagca gtccatgcc aacagacctt attacatcgg	19620
cttcagagac aactttatcg gcctcatgta ttacaatagc actggcaaca tgggagtgt	19680
tgcaggtcag gcctctcagt tgaatgcagt ggtggacttg caagacagaa acacagaact	19740
gtcctaccag ctcttgcttg attccatggg tgacagaacc agatactttt ccatgtggaa	19800
tcaggcagtg gacagttatg acccagatgt cagaattatt gaaaatcatg gaactgaaga	19860

cgagctcccc aactattgtt tccctctggg cggcataggg gtaactgaca cttaccaggc	19920
cattaaaacc aatggcaatg gtcaagaaaa cccaacctgg gaaaaagata cagagtttgc	19980
agaccgcaat gaaatagggg tgggaacaa tttcgctatg gagatcaacc tcagtgccaa	20040
cctgtggaga aacttcctgt actccaacgt ggcgctgtac ctgccagaca agcttaagta	20100
caacccctcc aatgtggaca tctctgacaa cccaacacc tacgattaca tgaacaagcg	20160
agtgggtggc cgggggctgg tggactgcta catcaacctg ggcgcgcgct ggtcgctgga	20220
ctacatggac aacgtcaacc ctttaacca ccaccgcaat gcgggcctgc gctaccgctc	20280
catgctcctg ggcaacgggc gctacgtgcc ctccacatc caggtgcccc agaagtcttt	20340
tgccatcaag aacctctcc tctgccggg ctctacacc tacgagtgga acttcaggaa	20400
ggatgtcaac atggctctcc agagctctct gggtaacgat ctcagggtgg acggggccag	20460
catcaagttc gagagcatct gcctctacgc caccttcttc cccatggccc acaacacggc	20520
ctccacgctc gaggccatgc tcaggaacga caccaacgac cagtccttca atgactacct	20580
ctccgccgc aacatgtctt accccatacc cgccaacgcc accaacgtcc ccatctccat	20640
cccctcgcg aactggcgcg ctttcgcgcg ctgggccttc acccgctca agaccaagga	20700
gacccctcc ctgggctcgg gattcgacct ctactacacc tactcgggct ccattcccta	20760
cctggacggc accttctacc tcaaccacac tttaagaag gtctcgggtca ctttcgactc	20820
ctcggtcagc tggccgggca acgaccgtct gtcaccccc aacgagttcg aaatcaagcg	20880
ctcggtcgac ggggagggct acaacgtggc ccagtgcaac atgaccaagg actggttctt	20940
ggtccagatg ctggccaact acaacatcgg ctaccagggc ttctacatcc cagagagcta	21000
caaggacagg atgtactcct tcttcaggaa ctccagccc atgagccggc aggtggtgga	21060
ccagaccaag tacaaggact accaggaggt ggcatcacc caccagcaca acaactcggg	21120
cttcgtgggc tacctcgccc ccacatgcg cgagggacag gcctaccccg ccaacttccc	21180
ctaccgctc ataggcaaga ccgcggtcga cagcatcacc cagaaaaagt tctctcgca	21240
tcgcacctc tggcgcatcc ctttctccag caacttcatt tccatgggtg cgctctcgga	21300
cctgggccag aacttgctct acgccaactc cgcccacgcc ctgcacatga ctttcgaggt	21360
cgaccccatg gacgagccca ccttctctta tgttctgttc gaagtctttg acgtggtccg	21420
ggtccaccag ccgcaccgcg gcgtcatcga gacctgtac ctgcgtacgc ctttctcggc	21480
cggcaacgcc accacctaaa gaagcaagcc gcagtcacgc ccgcctgcat gccgtcgggt	21540
tccaccgagc aagagctcag ggccatcgtc agagacctgg gatgcgggcc ctattttttg	21600
ggcactttcg acaagcgctt ccctggcttt gtctccccc acaagctggc ctgcgccatc	21660
gtcaacacgg ccggccgcga gaccgggggc gtgcactggc tggccttcgc ctggaacccg	21720

cgctccaaaa catgcttcct ctttgacccc ttcggctttt cggaccagcg gctcaagcaa 21780

atctacagat tcgagtacga gggcttgctg cgtcgcagcg ccatcgcctc ctgccccgac 21840

cgctgcgtca cctcgaataa gtccaccag accgtgcagg ggcccgaact ggccgcctgc 21900

ggtctcttct gctgcatgtt tctgcacgcc tttgtgcact ggcctcagag tcccatggac 21960

cgcaacccca ccatgaactt gctgacgggg gtgccaact ccatgtcca gagccccag 22020

gtcgagccca cctgcgccg caaccaggag cagctctaca gtttcttga gcgcactcg 22080

ccctacttcc gccgccacag cgcacagatc aggagggcc cctccttctg ccacttgcaa 22140

gagatgcaag aagggaata acgatgtaca cacttttttc tcaataaatg gcattttttt 22200

attatacaa gcctctctgg gtattcattt cccaccacca ccaccaccg cgtttgtcgc 22260

catctggctc tatttagaaa tcgaaagggt tctgccggga gtgccgtgc gccacgggca 22320

gggacacgtt gcgatactg tagcgggtgc cccacttga ctcgggcacc accaggcgag 22380

gcagctcggg gaagttttcg ctccacaggc tgcgggtcag caccagcgcg ttcacaggt 22440

cgggcgccga gatcttgaag tcgcagttgg ggccgccgcc ctgcgcgcgc gaggttcggt 22500

acaccgggtt gcagcactgg aacaccaaca gcgccgggtg cttcacgcta gccagcacgc 22560

tgcggtcgga gatcagctcg gcgtccaggt cctccgcgtt gctcagcgcg aacgggggtca 22620

tcttgggcac ttgcctcccc aggaaggcg cgtgccccgg tttcagttg cagtgcagc 22680

gcagcgggat cagcaggtgc ccatgcccgg actcggcggtt ggggtacagc gcgcgcatga 22740

aggctgcat ctggcggaag gccatctggg ccttggcgcc ctccgagaag aacatgccgc 22800

aggacttgc cgagaactgg tttgcggggc agctggcgtc gtgcaggcag cagcgcgct 22860

cgggtgttggc gatctgcacc acgttgcgcc cccaccggtt cttcacgac ttggccttgg 22920

acgattgctc cttcagcgcg cgctgccccg tctcgtggt cacatccatc tcgatcatc 22980

gttcttgtt caccatgctg ctgccgtgca ggcacttcag ctgcacctcc gtctcggtgc 23040

agcgggtgctg ccacagcgcg cagcccgtgg gctcgaaaga cttglaggtc acctccgca 23100

aggactgcag gtacccctgc aaaaagcggc ccatcatggt cacgaagtc ttgttgctgc 23160

tgaaggtcag ctgcagcccc cgggtctcct cgttcagcca ggtcttgac acggccgcca 23220

gcgcctccac ctggtcgggc agcatcttga agttcacctt cagctcattc tccacgtggt 23280

acttgcctat cagcgtgcgc gccgcctcca tgccttctc ccaggccgac accagcgga 23340

ggctcacggg gtctttcacc atcacgtgg ccgccgctc cgccgcgctt tcgctttccg 23400

ccccgtgtt ctcttctct tctctctct cctgcgccg gccactcgc agccccgca 23460

ccacggggtc gtcttctgc aggcgctgca ccttgcgctt gccgttgcgc ccctgcttga	23520
tgcgcacggg cgggttgctg aagcccacca tcaccagcgc gccctcttct tgctcgtcct	23580
cgctgtccag aatgacctcc ggggaggggg ggttggtcat cctcagtacc gaggcacgct	23640
tctttttctt cctgggggctg ttgccagct ccgcggctgc ggccgctgcc gaggtcgaag	23700
gccgagggct gggcgtgcgc ggcaccagcg cgtcctgcga gccgtcctcg tctcctcgg	23760
actcgagacg gaggcgggccc cgcttcttcg ggggcgcgcg gggcggcgga ggcggcgcg	23820
gcgacggaga cggggacgag acatcgcca gggtgggtgg acggcgggcc gcgccgcgtc	23880
cgcgctcggg ggtggtctcg cgctggctct ctccccact ggccatctcc cactgctcct	23940
tctcctatag gcagaaagag atcatggagt ctctcatgcg agtcgagaag gaggaggaca	24000
gcctaaccgc cccctctgag ccctccacca ccgccccac caccgccaat gccgccgcgg	24060
acgacgcgcc caccgagacc accgccagta ccacnnnct ccccgagcgc gcacccccgc	24120
tcgagaatga agtgctgac gagcaggacc cgggttttgt gagcggagag gaggatgagg	24180
tggatgagaa ggagaaggag gaggtcgccg cctcagtgcc aaaagaggat aaaaagcaag	24240
accaggacga cgcagataag gatgagacag cagtcgggcg ggggaacgga agccatgatg	24300
ctgatgacgg ctacctagac gtgggagacg acgtgctgct taagcacctg caccgccagt	24360
gcgtcatcgt ctgcgacgcg ctgcaggagc ggtgcgaagt gccctggac gtggcggagg	24420
tcagccgcgc ctacgagcgg cactcttctg cgccgcacgt gcccccaag cgccgggaga	24480
acggcacctg cgagcccaac ccgcgtctca acttctaccc ggtcttcgcg gtaccgagg	24540
tgttgccac ctaccacatc tttttccaaa actgcaagat cccctctcc tgccgcgcta	24600
accgcacccg cgccgacaaa accctgaccc tgcggcaggg cgccacata cctgatattg	24660
cctctctgga ggaagtgcc aagatcttcg aggtctctcg tcgcgacgag aaacgggcgg	24720
cgaacgtctt gcacggagac agcgaacacg agagtcactc gggggtgctg gtggagctcg	24780
agggcgacaa cgcgcgcttg gccgtactca agcgcagcat agaggtcacc cactttgcct	24840
acccgcgct caacctgcc cccaaggta tgagtgtgt catgggcgag ctcatcatgc	24900
gccgcgccc gccctggcc gcggatgcaa acttgcaaga gtctcagag gaaggcctgc	24960
ccgcggtcag cgacgagcag ctggcgcgct ggctggagac ccgcgacccc gcgcagctgg	25020
aggagcggcg caagctcatg atggccgcgg tgctggtcac cgtggagctc gagtgtctgc	25080
agcgcttctt cgcgacccc gagatgcagc gcaagctcga ggagaccctg cactacacct	25140
tccgccaggg ctacgtgcgc caggcctgca agatctccaa cgtggagctc tgcaacctgg	25200
tctcctacct gggcatcctg cagcagaacc gcctcgggca gaacgtcctg cactccacc	25260
tcaaagggga ggcgcgccg gactacatcc gcgactgcgc ctacctctt ctctgtaca	25320

cctggcagac ggccatgggg gtctggcagc agtgcctgga ggagcgcaac ctcaaggagc	25380
tggaaaagct cctcaagcgc accctcaggg acctctggac gggcttcaac gagcgctcgg	25440
tggccgccgc gctggcggac atcatcttcc ccgagcgctt gctcaagacc ctgcagcagg	25500
gcctgcccga cttcaccagc cagagcatgc tgcagaactt caggactttc atcctggagc	25560
gctcgggcat cctgccggcc acttctgctcgc cgtgcccag cgacttcgtg cccatcaagt	25620
acagggagtg cccgcgcgcg ctctggggcc actgctacct cttccagctg gccaactacc	25680
tgcctacca ctcggaacctc atggaagacg tgagcggcga gggcctgctc gagtgccact	25740
gccgtgcaa cctctgcacg cccaccgct ctctagtctg caaccgcag ctgctcagcg	25800
agagtcatat tatcggtacc ttccagctgc agggctccctc gcctgacgag aagtccgcgg	25860
ctccggggct gaaactcact ccggggctgt ggacttccgc ctacctacgc aaatttgtac	25920
ctgaggacta ccacgccac gagatcaggt tctacgaaga ccaatcccgc ccgccaagg	25980
cggagctcac cgctgcgtc atcaccaggg ggacacatcct gggccaattg caagccatca	26040
acaaagcccg ccgagagttc ttgctgaaaa agggctcgggg ggtgtacctg gacccccagt	26100
ccggcgagga gctaaacccg ctacccccgc cgccgcccga gcagcgggac cttgcttccc	26160
aggatggcac ccagaaagaa gcagcagccg ccgccgcagc catactgct tctggaggaa	26220
gaggaggagg actgggacag tcaggcagag gagatgatgg aagactggga ggaggacagc	26280
agcctagacg aggaagcttc agagccgaa gaggtggcag acgcaacacc atcacctcgc	26340
gtcgcagccc cctcgccggg gccctgaaa tcctccgaac ccagcaccag cgctataacc	26400
tccgtctctc cggcgccggc gccacccgcc cgcagaccca accgtagatg ggacaccaca	26460
ggaaccgggg tcggtaatgc caagtgcccg ccgccgccac cgcagcagca gcagcagcag	26520
cgccagggct accgtctgtg gcgcgggcac aagaacgcca tagtcgcctg cttgcaagac	26580
tgcgggggca acatctcttt cgcccgccgc ttctctctat tccaccacgg ggtcgccttt	26640
ccccgcaatg tcttgcatca ctaccgtcat ctctacagcc cctactgcag cggcgaccca	26700
gaggcggcag cggcagccac agcggcgacc accacctagg aagatatcct ccgcgggcaa	26760
gacagcggca gcagcggcca ggagacccgc ggcagcagcg gcgggagcgg tgggcgcact	26820
gcgcctctcg cccaacgaac ccctctgcac ccgggagctc agacacagga tcttccccac	26880
tttgtatgcc atcttccaac agagcagagg ccaggagcag gagctgaaaa taaaaaacag	26940
atctctgcgc tccctcaccg gcagctgtct gtatcacaaa agcgaagatc agcttcggcg	27000
cacgtggag gacgcggagg cactcttcag caaatactgc gcgtcactc ttaaagacta	27060
gtcccgccg cttctcgaat ttagcgggga gaaaactacg tcctcgccgg ccgccgccc	27120
gccccccag ccgagatgag caaagagatt cccacgcat acatgtggag ctaccagccg	27180

cagatgggac tcgcggcggg agcggcccag gactactcca cccgcatgaa ctacatgagc	27240
gcgggacccc acatgatctc acaggtcaac gggatccgcg cccagcgaaa ccaaatactg	27300
ctggaacagg cgcccatcac cgccacgccc cgccataatc tcaacccccg aaattggccc	27360
gccgccctag tgiaccagga aacccccctc gccaccaccg tactacttcc gcgtgacgcc	27420
caggccgaag tccagatgac taactcaggg gcgcagctcg cgggcggctt tcgtcacggg	27480
gcgcggccgc tccgaccagg tataagacac ctgatgatca gaggccgagg tatccagctc	27540
aacgacgagt cggtagactc ttgcctcggc ctccgtccgg acggaacttt ccagctcgcc	27600
ggatccggtc gctcttcgtt cagccccgc caggcgtacc tgactctgca gacctcgtec	27660
tcggagcccc gctccggcgg catcggaacc ctccagtctg tggaggagtt cgtgccctcg	27720
gtctacttca accccttctc gggacctccc ggacgctacc ccgaccagtt cattccgaac	27780
tttgacgagg tgaaggactc ggcggacggc tacgactgaa tgtcaggtgc cgaggcagag	27840
cagcttcgcc tgagacacct cgagcactgc cgccgccaca agtgcttcgc ccgcggttcc	27900
ggtgagttct gctactttca gctaccgag gagcataccg aggggccggc gcacggcgtc	27960
cgctgacca ccaggggcga gggttacctgt tcctcatcc gggagtacac cctccgtccc	28020
ctgctagtgg agcgggagcg gggtcctgt gtccctaacta tcgcctgcaa ctgccctaac	28080
cctggattac atcaagatct ttgctgtcat ctctgtctg agtttaataa acgtgagat	28140
cagaatctac tggggctcct gtcgcatcc tgtgaacgcc accgtcttca cccacccga	28200
ccaggcccag gcgaacctca cctgcggtct gcatcggagg tccaagaagt acctcacctg	28260
gtacttcaac ggcacccctt ttgtggttta caacagcttc gacggggacg gagtctccct	28320
gaaagaccag ctctccggtc tcagctactic catccacaag aacaccaccc tccaactctt	28380
ccctccctac ctgccgggaa cctacgagtg cgtcacggc cgctgcaccc acctcacccg	28440
cctgatcgta aaccagagct ttccgggaac agataactcc ctcttccca gaacaggagg	28500
tgagctcagg aaactccccg gggaccaggg cgagacgta ccttcgaccc ttgtggggtt	28560
aggatttttt attaccgggt tgcctggctct tttaatcaaa gcttccttga gatttgttct	28620
ttcctttctac gtgtatgaac acctcagcct ccaataactc taccctttct tcgggatcag	28680
gtgacttttc tgaatcggg cttggtgtgc tgcttactct gttgattttt ttccttatca	28740
tactcagcct tctgtgcctc aggtcgccg cctgtgcgc acacatctat atctactgct	28800
ggttgctcaa gtgcagggtt cgccacccaa gatgaacagg tacatggtcc tatcgatcct	28860
aggcctgctg gccctggcgg cctgcagcgc cgccaaaaaa gagattacct ttgaggagcc	28920

cgcttgcaat gtaactttca agcccgaggg tgaccaatgc accaccctcg tcaaatgcgt	28980
taccaatcat gagaagctgc gcatcgacta caaaaacaaa actggccggt ttgcggtcta	29040
tagtgtgttt acgcccggag acccctctaa ctactctgtc accgtcttcc agggcggaca	29100
gtctaagata ttcaattaca ctttcccttt ttatgagttg tgcgatgcgg tcatgtacat	29160
gtcaaaacag tacaacctgt ggcctccctc tccccaggcg tgtgtggaaa atactgggtc	29220
ttactgctgt atggctttgg caatcactac gctcgctcta atctgcacgg tgctatatat	29280
aaaattcagg cagaggcgaa tctttatcga tgaaaagaaa atgccttgat cgtaacacc	29340
ggctttctat ctgcagaatg aatgcaatca cctccctact aatcaccacc accctccttg	29400
cgattgccca tgggttgaca cgaatcgaag tgccagtggg gtccaatgtc accatggtgg	29460
gccccgccgg caattccacc ctcatgtggg aaaaatttgt ccgcaatcaa tgggttcatt	29520
tctgtcttaa ccgaatcagt atcaagccca gagccatctg cgatgggcaa aatctaactc	29580
tgatcaatgt gcaaatgatg gatgctgggt actattacgg gcagcgggga gaaatcatta	29640
attactggcg accccacaag gactacatgc tgcattagtg cgaggcactt cccactacca	29700
ccccactac cacctctccc accaccacta ccactactac tactactact accactaccg	29760
ctgcccgcca taccgcaaa agcaccatga ttagcacaaa gcccctcgt gctcactccc	29820
acgccggcgg gcccatcggg ggcacctcag aaaccaccga gctttgcttc tgccaatgca	29880
ctaaccggag cgctcatgaa ctgttcgacc tggagaatga ggatgcccgag cagagctccg	29940
cttgctgac ccaggaggct gtggagcccc ttgccctgaa gcagatcggg gattcaataa	30000
ttgactcttc ttcttttgcc actcccgaat accctcccga ttctacttcc cacatcacgg	30060
gtaccaaaga ccctaacctc tctttctacc tgatgtgct gctctgtatc tctgtggtct	30120
cttcgcgct gatgttactg gggatgttct gctgcctgat ctgccgcaga aagagaaaag	30180
ctcgctctca gggccaacca ctgatgcctt tcccctaccc cccggatttt gcagataaca	30240
agatatgagc tcgtgtctga cactaaccgc ttactagcc tgcgtcttaa cccttgcgc	30300
ttgcgactcg agattccaca atgtcacagc tgtggcagga gaaaaatgta ctttcaactc	30360
cacggccgat acccagtggg cgtggagtgg ctgaggtagc tacttaacta tetgcaatag	30420
ctccacttcc ccagcatat cccaaccaa gtaccaatgc aatgccagcc tgttcaccct	30480
catcaacgt tccacctgg acaatggact ctatgtaggc tatgtaccct ttggtgggca	30540
aggaaagacc cagcttaca acctggaagt tcgccagccc agaaccacta cccaagctwc	30600
ymccaycacc agcaccagca gcagcagcca cagcagcagc agcagattat tgactttggt	30660
tttggccagc tcatctgccg ctaccaggc catctacagc tctgtgccg aaaccactca	30720
gaccaccgc ccagaaacga ccaccgccac caccctacac acctccagcg atcagatgcc	30780

gaccaacatc acccccttgg ctcttcaaat gggacttaca agccccactc caaaaccagt	30840
ggatgcggcc gaggtctccg ccctcgtcaa tgactgggcg gggctgggaa tgtggtggtt	30900
cgccataggc atgatggcgc tctgcctgct tctgctctgg ctcatctgct gcctccaccg	30960
caggcgagcc agacccccca tctatagacc catcattgtc ctgaaccccg ataatgatgg	31020
gatccataga ttggatggcc tgaaaaacct acttttttct ttacagtat gataaattga	31080
gacatgcctc gcattttctt gtacatgttc cttctccac cttttctggg gtgttctacg	31140
ctggccgctg tgctcacct ggaggtagac tgctctcac ctttactgt ctacctgctt	31200
tacggattgg tcacctcac tctcatctgc agcctaata cagtaatcat cgccttcac	31260
cagtgcattg attacatctg tgtgcgcctc gcatacttca gacaccacc gcagtaccga	31320
gacaggaaca ttgcccaact tctaagactg ctctaatacat gcataagact gtgatctgcc	31380
ttctgacct ctgcatcctg cccacctca cctcctgcca gtacaccaca aaatctccgc	31440
gcaaaagaca tgctcctgc cgcttccacc aactgtggaa tatacccaaa tgctacaacg	31500
aaaagagcga gctctccgaa gcttggctgt atggggctcat ctgtgtctta gttttctgca	31560
gcactgtctt tgccctcatg atctacccct actttgatit gggatggaac gcgatcgatg	31620
ccatgaatta cccaccttt cccgcaccg agataattcc actgcgacaa gttgtaccg	31680
ttgtcgttaa tcaacgcccc ccatcccta cgcctactga aatcagctac tttaacctaa	31740
caggcggaga tgactgacgc cctagatcta gaaatggacg gcatcagtac cgagcagcgt	31800
ctcctagaga ggcgcaggca ggcggtgag caagagcgcc tcaatcagga gctccgagat	31860
ctcgttaacc tgcaccagtg caaaagaggc atcttttgtc tggtaaagca ggccaaagtc	31920
acctacgaga agaccggcaa cagccaccgc ctcatgtaca aattgcccac ccagcgccag	31980
aagctggtgc tcatggtggg tgagaatccc atcaccgtca cccagcactc ggtagagacc	32040
gaggggtgtc tgcactcccc ctgtcggggt ccagaagacc tctgcacct ggtaaagacc	32100
ctgtgcggtc tcagagattt agtccccctt aactaatcaa aacttggaat caataaaaag	32160
aatcacttac ttaaaatcag acagcaggtc tctgtccagt ttattcagca gcacctcctt	32220
ccccctctcc caactctggt actccaaacg ctttctggcg gcaaacttcc tccacacct	32280
gaagggaatg tcagattctt gctcctgtcc ctccgcacc actatcttca tgttgttgca	32340
gatgaagcgc accaaaacgt ctgacgagag ctcaacccc gtgtaccct atgacacgga	32400
aagcgccct cctccgtcc ctttctcac cctcccttc gtgtctccg atggattcca	32460
agaaagtccc cccgggttcc tgtctctgaa cctggccgag cccctggtca cttcccacgg	32520
catgctgcc ctgaaaatgg gaagtggcct ctccctggac gacgttgga acctcacctc	32580
tcaagatatc accaccgcta gccctccct caaaaaaac aagaccaacc tcagcctaga	32640

aacctcatcc cccctaactg tgagcacctc aggcgcctc accgtagcag cgcgcgtcc	32700
cctggcggtg gccggcacct cctcaccat gcaatcagag gccccctga cagtacagga	32760
tgcaaaactc accctggcca ccaaaggccc cctgaccgtg tctgaaggca aactggcctt	32820
gcaaacatcg gccccgtga cggccgctga cagcagcacc ctaccgtta gcgccacacc	32880
accaattaat gtaagcagtg gaagttagg cttagacatg gaagacccta tgtatactca	32940
caatggaaaa ctgggaataa gaattggggg tccactaaga gtagtagaca gcttgcatat	33000
actcactgta gttaccggaa atggactaac ttagataac aatgccctcc aaactaaagt	33060
tacgggcgcc ctaggttatg acacatcagg aaatctacaa ttaagagctg caggaggtat	33120
gcgaattgac gcaaattggcc aactiatcct taatgtggca taccatttg atgctcagaa	33180
caatctcagc cttagacttg gtcagggacc cctgtatata aacacagacc acaacctgga	33240
tttgaattgc aacagaggtc taaccacaac taccaccaac aacacaaaaa aacttgagac	33300
taaaattagc tcaggcttag actatgacac caatggtgct gtcattatta aacttggcac	33360
tggcttaagc ttcgacaaca caggcgccct aactgtggga aacactggtg atgataaact	33420
gactctgtgg acgacccag acccatctcc aaattgcaga attcactcag acaaagactg	33480
caagtttact ctatgcctaa ctaagtgtgg aagccaaatc ctggcctctg tcgccgccct	33540
agcggtatca ggaatctgg cttcgataac aggcaccgtt gccagcgtta ccattttct	33600
cagatttgat cagaatggag tgcttatgga aaactcctcg ctagacaggc agtactggaa	33660
cttcagaaat ggcaactcaa ctaacgtgc ccctacacc aatgcagtg ggttcatgcc	33720
aaacctcgca gcatacccca aaacgcaaag ccagactgct aaaaaaaca ttgtaagtca	33780
ggtttacttg aatggagaca aatccaaacc catgacctt accatcacc tcaatggaac	33840
taatgaatcc agtgaaacta gccaggtgag tcactactcc atgtcattta catgggcttg	33900
ggaaagtggg caatatgcca ctgaaacctt tgccaccaac tcttcacct tttcttat	33960
tgctgaacaa taaaagcat gacactgatg ttcatttctg attcttattt tattattttc	34020
aaacacaaca aaatcattca agtcattctt ccattctagc ttaatagaca cagtagctta	34080
atagaccag tagtgcaaag cccattctta gcttatagat cagacagtga taattaacca	34140
ccaccaccac catacctttt gattcaggaa atcatgatca tcacaggatc ctagtcttca	34200
ggcgcgcgcc tccctcccaa gacacagaat acacagtct ctcgcccgga ctggctttaa	34260
ataacaccat ctggttggtc acagacatgt tcttaggggt tatattccac acggtctcct	34320
gccgcgccag gcgctcgtc gtgatgttga taaactctcc cggcagctcg ctcaagttca	34380

cgctgctgtc cagcggctga acctccggct gacgcgataa ctgtgcgacc ggctgctgga	34440
caaacggagg ccgcgcctac aagggggttag agtcataatc ctcggtcagg ataggcggt	34500
gatgcagcag cagcgagcga aacatctgct gccgccgccg ctccgtccgg caggaaaaca	34560
acacgccggt ggctctctcc gcgataatcc gcaccgcccg cagcatcagc ttcctcgttc	34620
tccgcgcgca gcacctcacc ctgatctcgc tcaagtcggc gcagtaggta cagcacagca	34680
ccacgatgtt attcatgac ccacagtga gggcgctgta tccaaagctc atgccgggaa	34740
ccaccgcccc cactggcca tcgtaccaca agcgcacgta aattaaagtgt cgaccctca	34800
tgaactgtct ggacacaaac attacttctt tgggcatgtt gtaattcacc acctcccgt	34860
accagataaa cctctggta aacaggcac cttccaccac catcctgaac caagaggcca	34920
gaacctgccc accggctatg cactgcaggg aaccgggtt ggaacaatga caatgcagac	34980
tccaaggctc gtaaccgtgg atcatccggc tgcgaaggc atcgatgttg gcacaacaca	35040
gacacacgtg catgcacttt ctcatgatta gcagctcttc cctcgtcagg atcatatccc	35100
aaggaataac ccattcttga atcaacgtaa aaccacaca gcagggaagg cctcgcat	35160
aactcacgtt gtgcatggc agcgtgttgc attccgaaa cagcgatga tcctccagta	35220
tcgaggcgcg ggctctcttc tcacaggag gtaaagggtc cctgctgtac ggactgcgcc	35280
gggacgaccg agatcgtgtt gagcgtagt tcatggaaaa gggaacgccg gacgtggtca	35340
tacttcttga agcagaacca ggcttcgcgcg tggcaggcct ccttgcgtct gcggtctcgc	35400
cgtctagctc gctccgtgtg atagtgttag tacagccact cccgcagagc gtcgaggcgc	35460
accctggctt ccgatctat gtagactccg tcttgaccg cgccctgat aatatccacc	35520
accgtagaat aagcaacacc cagccaagca atacactgc tctgcgagcg gcagacagga	35580
ggagcgggca gagatgggag aaccatgata aaaaactttt tttaaagaat atttccaat	35640
tcttcgaaag taagatctat caagtggcag cgtcccctc cactggcgcg gtcaaaactct	35700
acggccaaag cacagacaac ggcatcttca agatgttctt taatggcgct caaaagacac	35760
accgtctca agtcgcagta aactatgaat gaaaacccat ccggtgatt ttccaatata	35820
gacgcgccgg cgcgctccac caaaccaga taattttctt ctctccagcg gtttagaatc	35880
tgtctaagca aatcccttat atcaagtcg gccatgcaa aaatctgctc aagagcgccc	35940
tccaccttca tgaccaagca gcgcatcatg attgcaaaaa ttcaggttct tcagagacct	36000
gtataagatt caaaatggga acattaacaa aaattctct gtcgcgcaga tcccttcgca	36060
gggcaagctg aacataatca gacaggtctg aacggaccag tgaggccaaa tccccaccag	36120
gaaccagatc cagagacctt atactgatta tgacgcgcat actcggggct atgctgacca	36180
gcgtagcgcc gatgtaggcg tgctgcatgg gcggcgagat aaaatgcaa gtgctggtta	36240

aaaaatcagg caaagcctcg cgcaaaaaag ctaacacatc ataatcatgc tcatgcaggt	36300
agttgcaggt aagctcagga accaaaacgg aataacacac gattttcctc tcaaactga	36360
cttcgcggat actgcgtaaa acaaaaatta taaataaaaa attaattaac ttaaacttg	36420
gaagcctgtc tcacaacagg aaaaaccact ttaatcaaca taagacgggc cacgggcatg	36480
ccggcatagc cgtaaaaaaa ttggccccg tgattaacaa gtaccacaga cagctccccg	36540
gtcatgtcgg gggtcatcat gtgagactct gtatacacgt ctggattgtg aacatcagac	36600
aaacaaagaa atcgagccac gtagcccgga ggtataatca cccgcaggcg gaggtacagc	36660
aaaacgaccc ccataggagg aatcacaaaa ttagtaggag aaaaaaatac ataaacacca	36720
gaaaaaccct gttgctgagg caaaatagcg cctccccgat ccaaacaac ataaagcgt	36780
tccacaggag cagccataac aaagaccga gtcttaccag taaaaagaaa aaagatctct	36840
caacgcagca ccagcaccaa cacttcgcag tgtaaaaggc caagtgccga gagagtatat	36900
ataggaataa aaagtgcgt aaacgggcaa agtccaaaa acgcccagaa aaaccgcacg	36960
cgaacctacg ccccgaaacg aaagccaaaa aacactagac actcccttcc ggcgtcaact	37020
tccgctttcc cagctacgt cacttgcccc agtcaaaaa actacatac ccgaacttcc	37080
aagtcgccac gcccaaaaa cgcctacac ctccccgcc gccggcccg ccccaaaccc	37140
gcctccccgc ccgcgccccg ccccgcccg cccatctcat tatcatattg gcttcaatcc	37200
aaaataaggt atattattga tgatggttta aacggatcca attcttgaag acgaaagggc	37260
ctcgtgatac gcctattttt ataggttaat gtcataata taatggtttc ttagacgtca	37320
ggtaggcactt ttccgggaaa tgtgcgcgga acccctatit gtttatitit ctaaatacat	37380
tcaaataatgt atccgctcat gagacaataa cctgataaa tgcttcaata atattgaaaa	37440
aggaagagta tgagtattca acatttccgt gtgcgctta ttcccttttt tgcggcattt	37500
tgctttcttg tttttgctca ccagaaaacg ctggtgaaag taaaagatgc tgaagatcag	37560
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtlaagat ccttgagagt	37620
tttcgccccg aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg	37680
gtattatccc gtgttgacgc cgggcaagag caactcggtc gccgcataca ctattctcag	37740
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta	37800
agagaattat gcagtgtcgc cataacctg agtgataaca ctgcggccaa cttaactctg	37860
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta	37920
actcgcttg atcgttgga accggagctg aatgaagcca taccaaagca cgagcgtgac	37980
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt	38040
actctagctt cccggcaaca attaataagac tggatggagg cggataaagt tgcaggacca	38100

cttctgcgct cggcccttcc ggctggctgg ttatttgctg ataatctgg agccggtgag 38160

cgtaggtctc gcgtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta 38220

gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgtgag 38280

atagggtcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atataactt 38340

tagattgatt taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc 38400

cttaacgtga gttttcgttc cactgagcgt cagaccccggt agaaaagatc aaaggatctt 38460

cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac 38520

cagcgggtggg ttgtttgccg gatcaagagc taccaactct ttttcgaag gtaactggct 38580

tcagcagagc gcagatacca aatactgtcc ttctagtgtg gccgtagtta ggccaccact 38640

tcaagaactc ttagcaccg cctacatacc tcgtctgtct aatcctgtta ccagtggctg 38700

ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata 38760

aggcgagcgc gtcgggctga acgggggggtt cgtgcacaca gccagcttg gagcgacga 38820

cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag 38880

ggagaaaaggc ggacaggtat ccggttaagcgc gcagggtcgg aacagagagc cgcacgaggg 38940

agcttccagg gggaaacgcc tggatcttt atagtctgt cggttttcgc cacctctgac 39000

ttgagcgtcg attttttga tgctcgtcag gggggcggag cctatggaaa aacgccagca 39060

acgggcctt ttiacggttc ctggcctttt gctggccttg aagctgtccc tgatggctgt 39120

catctacctg cctggacagc atggcctgca acgcgggcat cccgatgccg ccggaagcga 39180

gaagaatcat aatggggaag gccatccagc ctgcgtcgc agatccgaat tcgttttaac 39240

<210> 18

<211> 33

<212> DNA

<213> Pan troglodytes

<400> 18

atacggacta gtggagaagt actgcctac atg 33

<210> 19

<211> 36

<212> DNA

<213> Pan troglodytes

<400> 19

atacgaaga tctaagactt caggaaatat gactac 36

<210> 20
 <211> 45
 <212> DNA
 <213> Pan troglodytes
 <400> 20
 attcagtgtgta caggcgcgcc aaagcatgac actgatgttc atttc 45
 <210> 21
 <211> 35
 <212> DNA
 <213> Pan troglodytes
 <400> 21
 actaggacta gttataagct agaatggggc tttgc 35
 <210> 22
 <211> 37740
 <212> DNA
 <213> Pan troglodytes
 <220><221> misc_feature
 <222> (24391)..(24393)
 <223>
 > n is a, c, g, or t
 <400> 22
 catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg 60
 cggggcgggg cgcggggcgg gaggcgggtt tgggggcggg ccggcgggcg gggcgggtgtg 120
 gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag 180
 tgacgttttc cgtgcgcgac aacgccccg ggaagtgaca tttttccgc ggtttttacc 240
 ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact 300
 gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta 360
 gggccgaggg actttggcgg attacgtgga ggactcgccc aggtgttttt tgaggatgaat 420
 ttcgcgttc cgggtcaaag tctgcgtttt attattatag gataatccat tgcatacgtt 480
 glatccatat cataatatgt acatttataat tggctcatgt ccaacattac cgccatgttg 540
 acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcataagccc 600
 atatatggag ttcgcgtta cataacttac ggtaaattggc ccgcctggct gaccgcccac 660
 cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 720
 ttccattga cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca 780

agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggccccgctg 840

 gcattatgcc cagtaacatga ccttatggga ctttctact tggcagtaca tctacgtatt 900
 agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg 960
 gtttgactca cggggatttc caagtcacca ccccatggac gtcaatggga gtttgttttg 1020
 gcacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat 1080
 gggcggtagg cgtgtacggt gggagggtcta tataagcaga gctctcccta tcagtgatag 1140
 agatctccct atcagtata gagatcgctg acgagctcgt ttagtgaacc gtcagatcg 1200
 ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct 1260

 ccgcggccgg gaacggtgca ttggaacgcg gattccccgt gccaaagatg agatcttccg 1320
 tttatctagg taccgggccc cccctcgagg tcgacggtat cgataagctt cacgctgccg 1380
 caagcactca gggcgcaagg gctgctaaag gaagcggaac acgtagaaag ccagtccgca 1440
 gaaacggtgc tgaccccgga tgaatgtcag ctactgggct atctggaca gggaaaacgc 1500
 aagcgcaaag agaaagcagg tagcttgtag tgggcttaca tggcgatagc tagactgggc 1560
 ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgccctctg gtaaggttgg 1620
 gaagccctgc aaagtaact ggatggcttt ctgcccga aggatctgat ggcgcagggg 1680

 atcaagatct aaccaggagc tatttaatgg caacagttaa ccagctggta cgcaaaccac 1740
 gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgctgga agcatgccg caaaaacgtg 1800
 gcgtatgtac tcgtgtatat actaccactc ctaaaaaacc gaactccgcg ctgcgtaaag 1860
 tatgccgtgt tcgtctgact aacggtttcg aagtgacttc ctacatcggt ggtgaaggtc 1920
 acaacctgca ggagcactcc gtgatcctga tccgtggcgg tcgtgtttaa gacctccgg 1980
 gtgttcgtta ccacaccgta cgtgggtgcgc ttgactgctc cggcgtttaa gaccgtaagc 2040
 aggctcgctc caagtatggc gtgaagcgtc ctaaggetta atggtagatc tgatcaagag 2100

 acaggatgac ggtcgtttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc 2160
 gcttgggtgg agaggctatt cggtatgac tgggcacaac agacaatcgg ctgctctgat 2220
 gccgccgtgt tccggtgtc agcgcagggg cgcccgttc tttttgtcaa gaccgacctg 2280
 tccggtgccc tgaatgaact gcaggacgag gcagcgcggc tatcgtggct gccacgacg 2340
 ggcgttccct gcgcagctgt gctcgactgt gtcactgaag cgggaaggga ctggctgcta 2400
 ttggcggaag tgccggggca ggatctcctg tcactcacc ttgctcctgc cgagaaagta 2460
 tccatcatgg ctgatgcaat gcggcggtg catacgttg atccggtac ctgccattc 2520

gaccaccaag cgaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc 2580
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gtctgccagg 2640
ctcaaggcgc gcatgccga cggcgaggat ctctcgtga cccatggcga tgcctgcttg 2700
ccgaatatca tgggtgaaaa tggccgcttt tctggattca tgcactgtgg ccggctgggt 2760
gtggcggacc gctatcagga catagcgttg gctaccctg atattgctga agagcttgcc 2820
ggcgaatggg ctgaccgtt cctcgtgctt tacggtatcg ccgtcccgga ttcgcagcgc 2880
atcgcttct atcgcttct tgacgagttc tctgagcgg gactctgggg ttcgaaatga 2940

ccgaccaagc gacgccaac ctgccatcac gagatttga ttcaccgcc gccttctatg 3000
aaaggttggg cticggaatc gttttccggg acgccggctg gatgacctc cagcgcgggg 3060
atctcatgct ggagtcttc gccaccccg ggctcgatcc cctcgggggg aatcagaatt 3120
cagtgcagag cggccgcgat ctgctgtgcc tctagtgtc cagccatctg ttgtttgcc 3180
ctccccgtg ccttcctga ccctggaagg tgccactccc actgtcctt cctaataaaa 3240
tgaggaaatt gcatgcatt gtctgagtag gtgtcattct attctggggg gtgggtggg 3300
gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcggtggg 3360

ctctatggcc gatcagcga cgctgaggtg ggtgagtggg cgtggcctgg ggtggcatg 3420
aaaatatata agttgggggt cttaggtct ctttattgt gttgcagaga ccgccggagc 3480
catgagcggg agcagcagca gcagcagtag cagcagcgc ttggatggca gcatcgtgag 3540
cccttatttg acgacgcga tgccccactg ggccgggggtg cgtcagaatg tgatgggctc 3600
cagcatcgac ggccgacccg tctgcccgc aaattccgc acgtgacct atgcgacct 3660
cgccgggacg ccgttggacg ccaccgccgc cgccgccgc accgcagccg cctcgccgt 3720
gcgcagcctg gccacggact ttgcattctt gggaccactg gcgacagggg ctacttctcg 3780

ggccgctgct gccgcgctt gcgatgaaa gctgaccgcc ctgctggcgc agttggatgc 3840
gcttactcgg gaactgggtg acccttctca gcaggtcatg gccctgcgc agcaggtctc 3900
ctccctgcaa gctggcggga atgttctcc cacaaatgcc gtttaagata aataaaacca 3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttattt cataatttc 4020
cgcgcgcat aggccctaga ccagcgttct cggtcgtga gggcgcgtg tatcttctcc 4080
aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtccggggg 4140
tggaggtagc accactgcag agcttcatgc tccgggggtg tgtttagat gatccagtcg 4200

tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg 4260
gggaggccct tgggtgaagt gtttacaaaa cggttaagtt gggaagggtg cattcgggga 4320
gagatgatgt gcatcttga ctgtattttt agattggcga tgtttccgc cagatccctt 4380

ctgggattca tgttgtgcag gaccaccagt acagtgtatc cgggtgcactt ggggaatttg 4440

tcatgcagct tagagggaaa agcgtggaag aacttgaga cgcccttgtg gcctcccaga 4500

ttttccatgc attcgtccat gatgatggca atgggcccgc gggaggcagc ttgggcaaag 4560

atatttctgg ggtcgtgac gtcgtagtig tgttccaggg tgaggtcgtc ataggccatt 4620

tttacaagc gcgggcggag ggtgcccac tgggggatga tggccccctc tggccccggg 4680

gcgtagtigc cctcgcagat ctgcatttcc caggccttaa tctcggaggg gggaatcata 4740

tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctgggatgag 4800

agcaggtttc taagcagctg tgattttcca caaccgttg gccataaat aacacctata 4860

accggttgca gctggtagt tagagagctg cagctgccgt cgtcccggag gaggggggcc 4920

acctcgttga gcatgtccct gacgcgatg ttctcccga ccagatccgc cagaaggcgc 4980

tcgccccca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttagggccg 5040

tccgcgtgg gcatgtttt cagggtctgg ctacagcagc ccaggcggtc ccagagctcg 5100

gtgacgtgct ctacggcatc tctatccagc atatctctc gtttcgcggg ttggggcgac 5160

tttcgtgta gggcaccaag cggtggtcgt ccagcggggc cagagtcatg tctttccatg 5220

ggcgcagggt cctcgtcagg gtggtctggg tcacggtgaa ggggtgcgct ccgggctgag 5280

cgcttgccaa ggtgcgcttg aggttggttc tgctggtgct gaagcgctgc cggcttctgc 5340

cctgcgcgtc ggccaggtag catttgacca tgggtgcata gtccagcccc tccggggcgt 5400

gtcccttggc gcgcagcttg cccttggagg tggcgccga cgaggggcag agcaggctct 5460

tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgccgc 5520

agaccccgca cacggtctcg cactccacca gccaggtgag ctcggggcgc gccgggtcaa 5580

aaaccaggtt tccccatgc tttttgatgc gtttcttacc tcgggtctcc atgaggtggt 5640

gtccccgctc ggtgacgaag aggtgtccg tgtctcgtga gaccgacttg aggggtcttt 5700

tctccagggg ggtccctcgg tcttctcgt agaggaactc ggaccactct gagacgaagg 5760

cccgcgtcca ggccaggacg aaggaggcta tgtgggaggg gtagcggtcg ttgtccacta 5820

gggggtccac ctctccaag gtgtgaagac acatgtcgcc ttctcggcg tccaggaagg 5880

tgattggctt gtaggtgtag gccacgtgac cgggggttcc tgacgggggg gtataaaagg 5940

gggtgggggc gcgtcgtcg tcaactctt ccgcatcgct gtctcgagg gccagctgct 6000

ggggtgagta ttccctctcg aaggcgggca tgacctcgc gctgaggttg tcagtttcca 6060

aaaacgagga ggatttgatg ttacactgtc ccgaggtgat accttgagg gtaccgcgt 6120

ccatctggtc agaaaacacg atctttttat tgtccagctt ggtggcgaac gaccgtaga 6180

gggcgttgga gagcagcttg gcgatggagc gcagggtctg gttcttctcc ctgtcggcgc 6240

gctccttggc cgcatgttg agctgcacgt actcgcgcgc gacgcagcgc cactcgggga 6300

agacggtggt gcgtcgtcg ggcaccaggc gcacgcgcca gccgcggttg tgcagggtga 6360

ccagggtccac gctggtggcg acctcgccgc gcaggcgctc gttggtccag cagagacggc 6420

cgccttgcg cgagcagaag gggggcaggg ggtcgagctg ggtctcgtcc ggggggtccg 6480

cgtccacggt gaaaaccccg gggcgaggc gcgcgtcgaa gtagtctatc ttgcaacctt 6540

gcatgtccag cgctgtgc cagtcgaggc cggcgagcgc gcgtcgtag gggttgagcg 6600

gcgggcccc gggcatgggg tgggtgagtg cggaggcgta catccgcag atgtcataga 6660

cgtagagggg ctcccgagg accccgatgt aggtggggta gcagcgccg ccgcggatgc 6720

tggcgcgcac gtagtcatac agctcgtcg agggggcgag gaggtcgggg ccaggttgg 6780

tgcgggcggg gcgtccgtg cggaagacga tctgcctgaa gatggcatgc gaggttgaag 6840

agatggtggg gcgttgaag acgttgaagc tggcgtctg caggccgacg gcgtcgcga 6900

cgaaggaggc gtaggagtcg cgcagcttgi gtaccagctc ggcggtgacc tgcacgtcga 6960

gcgcgcagta gtcgagggtc tcgcggatga tgcataattt agcctgcccc ttctttttcc 7020

acagctcgcg gttgaggaca aactcttcgc ggtctttcca gtactcttgg atcgggaaac 7080

cgtccggttc cgaacggtaa gagcctagca ttagaactg gttgacggcc tggtaggcgc 7140

agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg 7200

tcaggcgcaa ggtgtccctg accatgactt tgaggtactg gtgcttgaag tcggagtcgt 7260

cgcagccgcc ccgtcccag agcgagaagt cggtcgcctt cttggagcgg gggttgggca 7320

gagcgaaggt gacatcgttg aagaggattt tgcccgcgcg gggcatgaag ttgcgggtga 7380

tgcggaaggg ccccgccact tcagagcggg tgtgatgac ctgggcggcg agcacgatct 7440

cgtcgaagcc gttgatgttg tggccacga ttagagttc caggaagcgg ggcggccct 7500

ttacggtggg cagcttcttt agctcttcgt aggtgagctc ctcgggcgag gcgaggccgt 7560

gctcgccag ggcccagtc gcgaggtgcg ggttgtctct gaggaaggac tcccagaggt 7620

cgcgggccag gaggtctgc aggcgggtccc tgaaggtcct gaactggcgg cccacggcca 7680

ttttttcggg ggtgatgcag tagaaggatga ggggtcttg ctgccagcgg tcccagtcga 7740

gctgcagggc gaggtcgcgc gcggcggtga ccaggcgctc gtcgccccg aatttcatga 7800

ccagcatgaa gggcacgagc tgctttccga agcccccat ccaagtgtag gtctctacat 7860

cgtaggtgac aaagaggcgc tccgtgcgag gatgcgagcc gatcggaag aactggatct 7920

cccgccacca gttggaggag tggctgttga tgtgtggaa gtagaagtcc cgtcgccggg 7980

ccgaacactc gtgctggctt ttgtaaaagc gagcgagta ctggcagcgc tgcacgggct 8040
 giacctcatg cagagatgc acctttcgcc cgcgcacgag gaagccgagg ggaatctga 8100
 gcccccgcc tggctcgagg catggctggt gctcttctac tttggatgag tgtccgtctc 8160
 cgtctggctc ctgaggggtt gttacgggtg agcggaccac cacgccgcgc gagccgcagg 8220
 tccagatgc ggccgcggc ggtcggagtt tgatgacgac atcgcgagc tgggagctgt 8280
 ccatggtctg gagctccgc ggccggcgca ggtcagccgg gagttcttgc aggttcacct 8340
 cgcagagtcg ggccagggcg cggggcagggt ctagggtgta cctgatctct aggggcgtgt 8400

 tggtaggggc gtcgatggct tgcaggagcc cgcagccccg gggggcgagc acggtgcccc 8460
 gcgggggtgtt ggtggtggtg gcggtgcagc tcagaagcgg tgccgcgggc gggccccgg 8520
 aggtaggggg ggtccggtc ccgcgggcag ggccggcagc ggcacgtcgg cgtggagcgc 8580
 gggcaggagt tggtagctg cccggaggtt gctggcgaag gcgacgagc ggcggttgat 8640
 ctcttgatc tggcgctct gcgtgaagac gacgggccc gtgagcttga acctgaaaga 8700
 gaggctgaca gaatcaatct cgggtgcatt gaccgcggcc tggcgagga tctctgcac 8760
 gtctcccgag ttgtcttgg aggcgatctc ggccatgaac tgctgatct ctctctctg 8820

 gaggtctccg cgtccggcg gttccacggt ggccgccagg tcgttgaga tgcgccccat 8880
 gagctgcgag aaggcgttga gtcgccctc gttccagact cggctgtaga ccacgcccc 8940
 ctggtcatcg cgggcgcgca tgaccacctg cgcgaggttg agctccacgt gccgcgcgaa 9000
 gacggcgtag ttgcgcagac gctggaagag gtagttgagg gtggtggcgg tgtgctcggc 9060
 cacgaagaag ttcatgacc agcggcgcaa cgtggattcg ttgatgtccc ccaaggcctc 9120
 cagccgttcc atggcctct agaagtccac ggcaagtgt aaaaactggg agttgcgcgc 9180
 cgacacggtc aactctctc ccagaagacg gatgagctcg gcgacggtgt cgcgcacctc 9240

 gcgctcgaag gctatgggga tctcttctc cgtagcacc accacctct ctcttctc 9300
 ctcttctggc acttccatga tggcttctc ctcttcgggg ggtggcggcg gcggcggtgg 9360
 gggagggggc gctctgcgc ggccggcgcg caccgggagg cggctccagc agcgcgcgat 9420
 catctccccg cggcgggcg gcatggtctc ggtgacggcg cggccgttct cccggggcg 9480
 cagttggaag acgcccggc acatctggtg ctggggcggg tggccgtgag gcagcgagac 9540
 ggcgctgac atgcatctc acaattgctg cgtaggtacg ccgcccaggg acctgaggga 9600
 gtccatatcc accggatccg aaaaccttc gaggaaggcg tctaaccagt cgcagtcgca 9660

 aggtaggctg agcaccgtgg cgggcggcgg ggggtggggg gagggtctgg cggaggtgct 9720
 gctgatgatg taattgaagt aggcggactt gacacggcgg atggtcgaca ggagcaccat 9780
 gtccctgggt ccggcctgct ggatgcggag gcggtcggct atgccccagg ctctgtctg 9840

gcatcggcgc aggtccttgt agtagtcttg catgagcctt tccaccggca cctcttctcc	9900
ttctctttct gcttcttcca tgtctgcttc ggccctgggg cggcgccgcg cccccctgcc	9960
ccccatgcgc gtgaccccca accccctgag cggttggagc agggccaggt cggcgacgac	10020
gcgctcggcc aggatggcct gctgcacctg cgtgagggtg gtttggaaagt catccaagtc	10080
cacgaagcgg tggtaggcgc ccgtgttgat ggtgtaggtg cagttggcca tgacggacca	10140
gttgacggtc tggtggcccg gttgcgacat ctccgtgtac ctgagtcgcg agtaggcgcg	10200
ggagtcgaag acgtagtcgt tgcaagtcg caccaggtac tggtagccca ccaggaagtg	10260
cggcgccggc tgccggtaga ggggccagcg cagggtggcg ggggctccgg gggccaggtc	10320
ttccagcatg aggcggtggt aggcgtagat gtacctggac atccaggtga taccgcggc	10380
ggtggtggag gcgcgcggga agtcgcgcac ccggttccag atgttgcga ggggcagaaa	10440
gtgctccatg gtaggcgtgc tctgtccagt cagacgcgcg cagtcgttga tactctagac	10500
cagggaaaac gaaagccggt cagcgggcac tcttccgtgg tctggtgaat agatcgcaag	10560
ggatatcatgg cggagggcct cgggttcgagc cccgggtccg ggccggacgg tccgccatga	10620
tccacgcggt taccgccgcg gtgtcgaacc caggtgtgcg acgtcagaca acggtggagt	10680
gttctttttg gcgtttttct ggccggggcg cggcgccgcg taagagacta agccgcgaaa	10740
gcgaaagcag taagtggctc gtcctccgta gccggaggga tccttgctaa gggttgcgtt	10800
gcggcgaacc ccggttcgaa tcccgctact gggccggccg gaccgcggcg taagggtttg	10860
gattggcctc cccctcgtat aaagaccccg ctgccggatt gactccggac acggggacga	10920
gccccittta tttttgcttt ccccatatgc atccggtgtt gcgacagatg cccccccgc	10980
cccagcagca gcaacaacac cagcaagagc ggagcgaaca gcagcgggag tcatgcaggg	11040
ccccctacc caccctcgcc ggcccggcca cctcggcgtc cgcggccgtg tctggcgctt	11100
gcggcggcgg cggcgggggg ccggtgacg accccagga gcccccgcgg cgcagggcca	11160
gacactacct ggacctggag gagggcgagg gcctggcgcg gctggggcg cgtctctccg	11220
agcgcacccc gcgggtgcag ctaaagcgcg actcgcgcga ggctacgtg cctcggcaga	11280
acctgttcag ggaccgcgcg ggcgaggagc ccgaggagat gcgggacagg aggttcagcg	11340
cggggcggga gctgcggcag gggctgaacc gcgagcggct gctgcgcgag gaggactttg	11400
agcccagcgc gcggacgggg atcagccccg cgcgcgcgca cgtggcggcc gccgacctgg	11460
tgacggcgta cgagcagacg gtgaaccagg agatcaactt ccaaagagt ttcaacaacc	11520
acgtgcgcac gctggtggcg cgcgaggagg tgaccatcgg gctgatgcac ctgtgggact	11580
ttgtgagcgc gctggtgcag aaccccaata gcaagcctct gacggcgag ctgttctga	11640
tagtgacga cagcagggac aacgaggcgt ttagggacgc gctgctgaac atcaccgagc	11700

ccgagggccg gtggctgctg gacctgatta acatcctgca gagcatagtg gtgcaggagc	11760
gcagcctgag cctggccgac aaggtggcgg ccatcaacta ctcgatgctg agcctgggca	11820
agttttacgc gcgcaagatc taccagacgc cgtacgtgcc catagacaag gaggtgaaga	11880
tcgacggttt ttacatgcgc atggcgctga aggtgctcac cctaagcgac gacctgggcg	11940
tgtaccgcaa cgagcgcatc cacaaggccg tgagcgtgag ccggcggcgc gagctgagcg	12000
accgcgagct gatgcatagc ctgcagcggg cgtggcgggg cgccggcagc ggcgacaggg	12060
aggcggagtc ctacttcgat gcggggggcg accctgcgtg ggcgcccagc cggcggggccc	12120
tggaggccgc gggggtccgc gaggactatg acgaggacgg cgaggaggat gaggagtacg	12180
agctagagga gggcgagtac ctggactaaa ccgcgggtgg tgtttccggt agatgcaaga	12240
cccgaacgtg gtggaccgga cgctgcgggc ggctctgcag agccagccgt ccggccttaa	12300
ctcctcagac gactggcgac aggtcatgga ccgcatcatg tcgtgacgg cgcgtaacct	12360
ggacgcgttc cggcagcagc cgcaggccaa caggctctcc gccatcctgg aggcggtggt	12420
gcctgcgcgc tcgaacccca cgcacgagaa ggtgctggcc atagtgaacg cgctggccga	12480
gaacagggcc atccgccgg acgaggccgg gctggtgtac gacgcgtgc tgcagcgct	12540
ggcccgtac aacagcggca acgtgcagac caacctggac cggctggtgg gggacgtgcg	12600
cgaggcggtg gcgcagcgcg agcgcgcgga tcggcagggc aacctgggct ccatggtggc	12660
gctgaatgcc ttctgagca cgcagccggc caacgtgccg cgggggcagg aagactacac	12720
caactttgtg agcgcgtgc ggctgatggt gaccgagacc cccagagcg aggtgtacca	12780
gtcgggtccg gactacttct tccagaccag cagacagggc ctgcagacgg tgaacctgag	12840
ccaggctttc aagaacctgc gggggctgtg gggcgtgaag gcgcccaccg gcgaccgggc	12900
gacggtgtcc agcctgtga cgcceaactc gcgcctgctg ctgctgctga tcgcgccgtt	12960
cacggacagc ggcagcgtgt cccgggacac ctacctgggg cacctgctga ccctgtaccg	13020
cgaggccatc gggcaggcgc aggtggacga gcacaccttc caggagatca ccagcgttag	13080
ccgcgcgctg gggcaggagg acacgagcag cctggaggcg actctgaact acctgctgac	13140
caaccggcgg cagaagattc cctcgtgca cagcctgacc tccgaggagg agcgcatttt	13200
gcgctacgtg cagcagagcg tgagcctgaa cctgatgcgc gacggggtga cgcccagtgt	13260
ggcgctggac atgaccgcgc gcaacatgga accgggcatg tacgccgcgc accggcctta	13320
catcaaccgc ctgatggact acctgcatcg cgcggcggcc gtgaaccccg agtactttac	13380
caacgccatc ctgaaccgc actggctccc gccgccggg ttctacagcg ggggcttcga	13440

gggtcccgag gccaacgatg gcttctctgt ggacgacatg gacgacagcg tgttctcccc	13500
gcggcccgag gcgtggcg aagcgtccct gctgcgtccc aagaaggagg aggaggcgag	13560
tcgccccgc ggcagcagcg gcgtggcttc tctgtccgag ctggggcgcg cagccgccgc	13620
gcgccccggg tccttggcg gcagccctt tccgagcctg gtggggtctc tgcacagcga	13680
gcgcaccacc cgcctcgcg tgcctggcg ggacgagtac ctgaataact ccctgctgca	13740
gccggtgagg gagaaaaacc tgcctccgc cttccccaac aacgggatag agagcctggt	13800
ggacaagatg agcagatgga agacctatgc gcaggagcac agggacgcgc ccgcgtccg	13860
gccccccacg cgccgccagc gccacgaccg gcagcggggg ctggtgtggg atgacgagga	13920
ctccgcggac gatagcagcg tgctggacct gggaggagc ggcaaccctg tcgcgcacct	13980
gcgccccgc ctggggagga tgttttaaaa aaaaaaaaaa gcaagaagca tgatgcaaaa	14040
attaaataaa actcaccaag gccatggcg ccgagcgttg gtttcttgtg ttccttcag	14100
tatgcggcg gcggcgatgt accaggagg acctcctccc tcttacgaga gcgtggtggg	14160
cgcgcgcg gcggcgacct cttctccctt tgcgtcgag ctgctggagc cgccgtacgt	14220
gcctccgcgc tacctcgcg ctacgggggg gagaaacagc atccgttact cggagctggc	14280
gccccgtgtc gacaccacc ggggtgtacct ggtggacaac aagtcggcg acgtggcctc	14340
cctgaactac cagaacgacc acagcaattt ttgaccacg gtcattcaga acaatgacta	14400
cagcccgagc gaggccagca ccagacct caatctggat gaccggtcgc actggggcg	14460
cgacctgaaa accatcctgc acaccaacat gccaacgtg aacgagttca tgttcacaa	14520
taagttaag gcgcggtga tgggtgcgc ctcgcacacc aaggaagacc ggggtgagct	14580
gaagtacgag tgggtggagt tcgagctgcc agagggaac tactccgaga ccatgacct	14640
tgacctgatg aacaacgca tcgtggagca ctatctgaaa gtgggcaggc agaacggggt	14700
cctggagagc gacatcgggg tcaagttcga caccaggaac ttccgcctgg ggctggacct	14760
cgtgaccggg ctggttatgc ccgggtgta caccaacgag gccttccatc ccgacctat	14820
cctgctgccc ggtgcgggg tggacttcac ttacagccgc ctgagcaacc tcttggcat	14880
ccgcaagcgg cagcccttcc aggagggtt caggatcacc tacaggacc tggagggggg	14940
caacatcccc gcgtcctcg atgtggaggc ctaccaggat agcttgaagg aaaatgaggc	15000
gggacaggag gataccgcc ccgccgctc cgcccgcc gagcaggcg aggatgctgc	15060
tgacaccgcg gcccgggagc gggcgaggc cgacccgct atggtggtg aggctgccga	15120
gcaggaggag gacatgaatg acagtgcgt gcgcggagac accttcgtca cccgggggga	15180
ggaaaagcaa gcggaggccg aggcccggc cgaggaaaag caactggcg cagcagcggc	15240
ggcgcgcg ttggcccg cgagggtga gtctgaggg accaagccc ccaaggagcc	15300

cgtgattaag cccctgaccg aagatagcaa gaagcgagct tacaacctgc tcaaggacag	15360
caccaacacc gcgtaccgca gctggtacct ggcttacaac tacggcgacc cgtcgacggg	15420
ggtgcgctcc tggacctgc tgtgcacgcc ggacgtgacc tgcggctcgg agcaggtgta	15480
ctggtcgctg cccgacatga tgcaagaccc cgtgaccttc cgctccacgc ggcaggtcag	15540
caacttcccg gtggtgggcg ccgagctgct gcccggtcac tccaagagct tctacaacga	15600
ccaggccgtc tactcccgag tcatccgcca gttcacctct ctgacctcag tgttcaatcg	15660
ctttctctag aaccagattc tggcgcgccc gcccgcccc accatcacca ccgtcagtga	15720
aaacgttctt gctctcacag atcacgggac gctaccgtg cgcaacagca tcggaggagt	15780
ccagcgagtg accgttactg acgccagacg ccgcacctgc ccctacgttt acaaggcctt	15840
gggcatagtc tcgccgcgcg tcttttccag ccgcactttt tgagcaacac caccatcatg	15900
tccatcctga tctcaccag caataactcc ggtgggggac tgctgcgcgc gccagcaag	15960
atgttcggag gggcgaggaa gcgttccgag cagcaccccg tgcgcgtgcg cgggcacttc	16020
cgcgccccct ggggagcgca caaacgcggc cgcgcggggc gcaccaccgt ggacgacgcc	16080
atcgactcgg tgggtggagca ggcgcgcaac tacaggcccc cggtctctac cgtggacgcg	16140
gccatccaga ccgtggtgcg gggcgcgagg cggtacgcca agctgaagag ccgccggaag	16200
cgcgtggccc gccgccaccg ccgccgaccc ggggccgccc ccaaacgcgc cgcccgggcc	16260
ctgcttcgcc gggccaagcg cacgggcccgc cgcgcgcgca tgagggccgc gcgcgccttg	16320
gccgccggca tcaccgccgc caccatggcc ccccgtaacc gaagacgcgc ggccgccgcc	16380
gccgccgccc ccattcagtga catggccagc aggcgcgggg gcaacgtgta ctgggtgcgc	16440
gactcggtag ccggcacgcg cgtgcccgtg cgcttccgcc ccccgcgac ttgagatgat	16500
gtgaaaaaac aactactagt ctctgctgt tgtgtgtatc ccagcggcgg cggcgcgcg	16560
cgcagcgtca tgtccaagcg caaaatcaaa gaagagatgc tccaggtcgt cgcgccggag	16620
atctatgggc cccgaagaa ggaagagcag gattcgaag cccgcaagat aaagcgggtc	16680
aaaaagaaaa agaaagatga tggcgatgcc gatggggagg tggagtctct gcgcgccacg	16740
gcgccaggc gcccggtgca gtggaagggc cggcgcgtaa agcgcgtcct gcgccccggc	16800
accgcggtgg tcttcacgcc cggcgagcgc tccaccgga ctttcaagcg cgtctatgac	16860
gagggtgtag gcgacgaaga cctgctggag caggccaacg agcgcttcgg agagtttgct	16920
tacgggaagc gtcagcgccc gctggggaag gaggacctgc tggcgctgcc gctggaccag	16980
ggcaacccca ccccgagtct gaagcccgctg acctgcagc aggtgctgcc gagcagcgca	17040
ccctccgagg cgaagcgggg tctgaagcgc gagggcggcg acctggcgcc caccgtgcag	17100
ctcatggtgc ccaagcgga gaggtggag gatgtgctgg agaaaatgaa agtagacccc	17160

ggctctgcagc cggacatcag ggtccgtccc atcaagcagg tggcgccggg cctcggcgtg	17220
cagaccgtgg acgtgggtcat cccacccggc aactcccccg ccgccaccac cactaccgt	17280
gcctccacgg acatggagac acagaccgat ccgcccgcag ccgcccgcac cgcgccgcc	17340
gcgacctcct cggcggaggt gcagacggac ccttggtgc cgccggcgat gtcagctccc	17400
cgcgcgcgtc gggggcgagc gaagtacggc gccgccaacg cgctcctgcc cgagtacgcc	17460
ttgcatcctt ccatcgcgcc cacccccggc taccgaggct atacctaccg cccggaaga	17520
gccaaagggtt ccaccgccg tccccgccga cgcgccgccg ccaccaccg cgcgccgcc	17580
cgcagacgcc agcccgact ggctccagtc tccgtgagga gaggggcg cgacggacac	17640
accctggtgc tgcccagggc gcgtaccac cccagcatcg tttaaaagcc tgttgtgtt	17700
cttcagata tggccctcac ttgccgctc cgtttcccg tgccgggata ccgaggagga	17760
agatcgccc gcaggagggg tctggccggc cgcggcctga gcggaggcag ccgccgcgcg	17820
caccggcggc gacgcgccac cagccgacgc atgcgcggcg ggggtctgcc cctgttaac	17880
cccctgatcg ccgcccgat cggcgccgtg cccgggatcg cctccgtgcc ctgcaggcg	17940
tcccagagge attgacagac ttgcaaatct gcaaatatgg aaaaaaaccc caataaaaaa	18000
gtctagactc tcacgtcgc ttggtcctgt gactattttg tagaatggaa gacatcaact	18060
ttgcgtcgt ggccccgct caccgctgc gccggttctt gggacactgg aacgatatcg	18120
gcaccagcaa catgagcgtt ggcccttca gttggggctc tctgtggagc ggcattaaaa	18180
gtatcgggtc tgccgttaaa aattacggt cccgggctg gaacagcagc acgggccaga	18240
tgttgagaga caagttaaa gagcagaact tccagcagaa ggtggtggag ggcctggcct	18300
ccggcatcaa cggggtggtg gacctggcca accaggccgt gcagaataag atcaacagca	18360
gactggacc cggccgccg gtggaggagg tgccgccggc gctggagacg gtgtccccg	18420
atgggcgtgg cgagaagcg ccgcccgcg atagggaaga gaccactctg gtcacgcaga	18480
ccgatgagcc gccccglat gaggaggccc tgaagcaagg tctgccacc acgcggccca	18540
tcgcgcccac ggccaccggg gtgggtgggc gccacacccc cgccacgtg gacttgctc	18600
cgcccgcca tgtccgcag cagcagcaga aggcggcaca gccgggccc cccgtgaccg	18660
cctccgttc ctcccggt cctctgcgc gcgcggccag cggccccgc ggggggtcg	18720
cgaggcacgg caactggag agcacgtga acagcatgt ggggtctggg gtgcgtccg	18780
tgaagcgcc ccgatgtac tgaatagctt agctaactg ttgtatgtgt gtatgcgcc	18840
tatgtcgcc ccagaggagc tgctgagtc cgccgttcg cgcgccacc accaccaccg	18900

ccactccgcc cctcaagatg gcgaccccat cgatgatgcc gcagtggctg tacatgcaca	18960
tctcgggcca ggacgcctcg gagtacctga gccccgggct ggtgcagttc gcccgcgcca	19020
ccgagagcta cttcagcctg agtaacaagt ttaggaaccc cacggtggcg cccacgcacg	19080
atgtgaccac cgaccggctc cagcgccctga cgctgcgggt cattcccgtg gaccgcgagg	19140
acaccgcgta ctctacaag gcgcgggttca ccttggccgt gggcgacaac cgcgtgctgg	19200
acatggcctc cacctaactt gacatccgcg ggggtgctgga cgggggcccc actttcaagc	19260
cttactctgg caccgcctac aactccctgg cccccaaggg cgctcccaac tcctgcgagt	19320
gggagcaatt agaagaagcc caggccgctg tggaagacga agaattagaa gatgaagacg	19380
aggaaccaca ggatgaggca cctgtgaaaa aaacccatgt atacgctcag gctccccctt	19440
ctggagaaga aattactaaa aacggtttgc aaatagggtc agataacaca gaagcccagt	19500
ctaagcccat atatgcagat cctacattcc agcctgaacc ccaaactcggg gaatcccagt	19560
ggaatgaggc agatgctaca gttgccggcg gtagagtgt aaagaaatcc actcccatga	19620
agccatgcta tggttcctat gcaagaccca caaactccaa tggagggtcaa ggtgtgctgg	19680
tggctgatga taaggggggt cttcaatcta aagttgaatt gcaatTTTT tcaaatacta	19740
ctactcttaa tcagcgggag ggtaacgata caaaacaaaa agtgggtgctg tatagcgaag	19800
atgtgcacat ggaaactcca gacaccaca tttcttaca gccacaaaa agcgtatgaca	19860
attcaaaaat catgctgggt cagcagteca tgcccaacag acctaatat atcggttca	19920
gagacaactt tatcgccctc atgtattaca atagcactgg caacatggga gtgcttcag	19980
gtcaggcctc tcagttgaat gcagtgggtg acttgcaaga cagaaacaca gaactgtcct	20040
accagctctt gcttgattcc atgggtgaca gaaccagata cttttccatg tggaatcagg	20100
cagtggacag ttatgaccca gatgtcagaa ttattgaaaa tcatggaact gaagacgagc	20160
tccccaaacta ttgtttccct ctgggcggca taggggtaac tgacacttac caggccatta	20220
aaaccaatgg caatggtcaa gaaaaccaa cctgggaaaa agatacagag tttgcagacc	20280
gcaatgaaat aggggtggga aacaatttcg ctatggagat caacctcagt gccaacctgt	20340
ggagaaaactt cctgtactcc aacgtggcgc tgtacctgcc agacaagctt aagtacaacc	20400
cctccaatgt ggacatctct gacaaccca acacctacga ttacatgaac aagcgagtgg	20460
tggccccggg gctggtggac tgetacatca acctgggcgc gcgtgggtcg ctggactaca	20520
tggacaacgt caacccttc aaccaccacc gcaatgcggg cctgcgtac cgctccatgc	20580
tcctgggcaa cgggcgttac gtgcccttc acatccaggt gcccagaag ttctttgcca	20640
tcaagaacct cctcctctg cgggctcct acacctacga gtggaacttc aggaaggatg	20700
tcaacatggt cctccagagc tctctgggt acgatctcag ggtggacggg gccagcatca	20760

agttcgagag catctgcctc tacgccacct tcttcccat ggcccaaac acggcctcca	20820
cgctcgaggc catgctcagg aacgacacca acgaccagtc cttcaatgac tacctctccg	20880
ccgccaacat gctctacccc ataccgcga acgccaccaa cgtcccatc tccatccct	20940
cgcgcaactg ggcggccttc cgcggtggg ccttaccgc cctcaagacc aaggagacc	21000
cctccctggg ctcggtatc gaccctact acacctactc gggtccatt ccctacctgg	21060
acggcacctt ctacctcaac cacactttca agaaggctc ggtcacctc gactcctcg	21120
tcagctggcc gggcaacgac cgtctgctca ccccaacga gttcgaaac aagcgctcg	21180
tcgacgggga gggtacaac gtggccagc gcaacatgac caaggactgg ttcctggtec	21240
agatgctggc caactacaac atcggtacc agggcttcta catcccagag agctacaagg	21300
acaggatgta ctcttcttc aggaacttc agcccatgag ccggcaggtg gtggaccaga	21360
ccaagtaca ggactaccag gaggtgggca tcatcacca gcacaacaac tcgggcttcg	21420
tgggtacct cgccccacc atgcgcgagg gacaggccta cccgccaac tccccctacc	21480
cgctcatagg caagaccgcg gtcgacagca taccacagaa aaagtctc tgcgatcgca	21540
ccctctggcg catcccttc tccagcaact tcatgtccat gggtgcgctc tcggacctgg	21600
gccagaactt gctctacgc aactcgcgc acgcccctga catgacctc gaggtcgacc	21660
ccatggacga gccaccctt ctctatgtc tgttcgaagt ctttgacgtg gtccgggtcc	21720
accagccgca ccgcccgcg atcgagaccg tgtacctgc tacgcccctc tcggccgga	21780
acgccaccac ctaaagaagc aagccgcagc catcgccgc tgcattgccg cgggttcac	21840
cgagcaagag ctcagggcca tcgtcagaga cctgggatgc gggccctatt ttttggcac	21900
tttcgacaag cgcttccctg gctttgtct cccacacaag ctggcctgcg ccatcgtcaa	21960
cacggccggc cgcgagaccg gggcgctgca ctggctggcc ttcgcctgga acccgcgctc	22020
caaaacatgc ttctctttg accccttcgg cttttcggac cagcggtca agcaaatcta	22080
cgagttcgag tacgagggtc tgctgcgtcg cagcgccatc gcctcctcgc ccgaccgtg	22140
cgtcacctc gaaaagtcca ccagaccgt gcaggggccc gactcgccg cctgcggtct	22200
cttctgctgc atgtttctgc acgcctttgt gcactggcct cagagtcca tggaccgaa	22260
ccccaccatg aacttgctga cgggggtgcc caactccatg ctccagagcc ccaggtcga	22320
gcccacctg cgcgcgaacc aggagcagct ctacagctc ctggagcgcc actcgcccta	22380
cttcgccgc cacagcgac agatcaggag ggccacctc ttctgccact tgcaagagat	22440
gcaagaagg taataacgat gtacacactt ttttctcaat aaatggcatt tttttattta	22500
tacaagctct ctgggtatt catttccac caccaccac acccgccgtt gtcgccatct	22560
ggctctatct agaaatcgaa agggttctgc cgggagtcgc cgtgcgccac gggcaggac	22620

acgttgcgat actggtagcg ggtgccccac ttgaactcgg gcaccaccag gcgaggcagc	22680
tctgggaagt tttcgtcca caggctgcgg gtcagcacca gcgcgttcat caggctgggc	22740
gccgagatct tgaagtcgca gttggggccg ccgccctgcg cgcgcgagtt gcggtacacc	22800
gggttgcagc actggaacac caacagcgcc gggtgcttca cgctagccag cacgctgcgg	22860
tctggagatca gctcggcgtc caggctcctcc gcgttgcctca gcgcgaacgg ggtcatcttg	22920
ggcacttgcc tccccaggaa gggcgcgtgc cccggtttcg agttgcagtc gcagcgcagc	22980
gggatcagca ggtgcccatg cccggactcg gcgttggggg acagcgcgcg catgaaggcc	23040
tgcactctggc ggaaggccat ctgggccttg gcgccctccg agaagaacat gccgcaggac	23100
ttgcccgaga actggtttgc ggggcagctg gcgtcgtgca ggcagcagcg cgcgtcgggtg	23160
ttggcgatct gcaccacgtt gcgccccac cggttcttca cgatcttggc cttggacgat	23220
tgctccttca gcgcgcgtg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttaccac tgctgctgcc gtgcaggcac ttcagctcgc cctccgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtgggctcg aaagacttgt aggtcacctc cgcgaaggac	23400
tgcaggtacc cctgcaaaaa gcggcccatc atgggtcacga aggtcttgtt gctgctgaag	23460
gtcagctgca gcccgcggtg ctctctgttc agccaggtct tgcacacggc cgccagcgcc	23520
tccacctggt cgggcagcat cttgaagttc accttcagct cattctccac gtggtacttg	23580
tccatcagcg tgcgcgccgc ctccatgcc tttctccagg ccgacaccag cggcaggctc	23640
acggggttct tcaccatcac cgtggccgcc gcctccgccg cgctttcgct ttccgccccg	23700
ctgttctctt cctcttctc ctcttctctg ccgccgccca ctgcagccc ccgcaccacg	23760
gggtcgtctt cctgcaggcg ctgcaccttg cgcttgccgt tgcgccctg cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcac agcgcggcct cttcttgctc gtctctgctg	23880
tccagaatga cctccgggga gggggggttg gtcatctca gtaccgaggc acgttctttt	23940
ttcttcttg gggcgttcgc cagctccgcg gctgcggccg ctgccgaggt cgaaggccga	24000
gggctgggcg tgcgcggcac cagcgcgtcc tgcgagccgt cctcgtctc ctcggactcg	24060
agacggaggc gggccccgtt cttcgggggc gcgcggggcg gcggaggcgg cggcggcgac	24120
ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgcc gcgtccgcgc	24180
tctggggttg tctcgcgtg gtctcttcc cgaactggcca tctccactg ctcttctcc	24240
tataggcaga aagagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accgccccct ctgagccctc caccaccgcc gccaccaccg ccaatgccgc cgcggacgac	24360

gcgcccaccg agaccaccgc cagtaccacc nnnctcccca gcgacgcacc cccgctcgag	24420
aatgaagtgc tgatcgagca ggacccgggt tttgtgagcg gagaggagga tgaggtggat	24480
gagaaggaga aggaggaggt cgccgcctca gtgccaaaag aggataaaaa gcaagaccag	24540
gacgacgcag ataaggatga gacagcagtc gggcggggga acggaagcca tgatgctgat	24600
gacggctacc tagacgtggg agacgacgtg ctgcttaagc acctgcaccg ccagtgcgtc	24660
atcgtctgcg acgcgtgca ggagcgggtg gaagtgcccc tggacgtggc ggaggtcagc	24720
cgcgcctacg agcggcacct cttcgcgccg cacgtgcccc ccaagcgccg ggagaacggc	24780
acctgcgagc ccaaccgcg tctcaacttc taccgggtct tcgcggtacc cgaggtgctg	24840
gccacctacc acatcttttt caaaaactgc aagatcccc tctcctgccg cgctaaccgc	24900
acccgcgcg acaaaacct gacctgcgg cagggcgccc acatacctga tattgcctct	24960
ctggaggag tgcccaagat cttcgagggt ctcggtcgcg acgagaaaac ggcggcgaac	25020
gctctgcacg gagacagcga aaacgagagt cactcggggg tgctgggtgga gctcgagggc	25080
gacaacgcgc gcctggccgt actcaagcgc agcatagagg tcaccactt tgctaccgc	25140
gcgtcaacc tgcccccaa ggtcatgagt gtggtcatgg gcgagctcat catgcgccgc	25200
gcccagcccc tgcccgcgga tgcaacttg caagagtcct cagaggaagg cctgcccgcg	25260
gtcagcgacg agcagctggc gcgctggctg gagaccgcg accccgcgca gctggaggag	25320
cggcgcaagc tcatgatggc cgcggtgctg gtcaccgtgg agctcgagt tctgcagcgc	25380
ttcttcgcgg accccgagat gcagcgcaag ctcgaggaga cctgcacta caccttcgc	25440
cagggtacg tgcgccagc ctgcaagatc tccaacgtgg agctctgcaa cctggtctcc	25500
tacctgggca tctgcacga gaaccgcctc gggcagaacg tctgcactc caccctcaa	25560
ggggaggcgc gccgcgacta catccgcgac tgcgcctacc tcttcctctg ctacacctgg	25620
cagacggcca tgggggtctg gcagcagtc ctggaggagc gcaacctcaa ggagctggaa	25680
aagctctca agcgcacct cagggaacctc tggacgggt tcaacgagcg ctcggtggcc	25740
gccgcgtgg cggaatcat cttccccgag cgctgtctca agaccctgca gcagggcctg	25800
cccgaactca ccagccagag catgtctgag aacttcagga cttcatcct ggagcgtcg	25860
ggcatcctgc cggccacttg ctgcgcgtg ccagcgcact tcgtgcccac caagtacagg	25920
gagtgcgccg cggcgtctg gggccactgc tacctcttc agctggccaa ctacctgcc	25980
tacctcgg acctcatgga agacgtgagc ggcgagggcc tgctcgagt cactgccgc	26040
tgcaacctct gcacgcccc cgcctctcta gtctgcaacc gcgagctgct cagcgagagt	26100
cagattatcg gtaccttga gctgcagggt cctcgcctg acgagaagtc cgcggtccg	26160
gggctgaaac tactccggg gctgtggact tccgcctacc tacgcaaatt tgtacctgag	26220

gactaccacg cccacgagat caggttctac gaagaccaat cccgcccgcc caaggcggag	26280
ctcaccgcct gcgtcatcac ccaggggcac atcctgggcc aattgcaagc catcaacaaa	26340
gccccccgag agttcttctg gaaaaagggt cgggggggtgt acctggaccc ccagtccggc	26400
gaggagctaa acccgctacc cccgcccgcg cccagcagc gggaccttgc ttcccaggat	26460
ggcaccacaga aagaagcagc agcccccgcg gcagccatac atgcttctgg aggaagagga	26520
ggaggactgg gacagtcagg cagaggagat gatggaagac tgggaggagg acagcagcct	26580
agacaggaa gcttcagagg ccgaagaggt ggcagacgca acaccatcac cctcggtcgc	26640
agccccctcg cgggggcccc tgaaatctc cgaaccacgc accagcgtta taacctccgc	26700
tcctccggcg ccggcgccac ccgcccgcag acccaaccgt agatgggaca ccacaggaac	26760
cggggtcggt aagtccaagt gccccccgcc gccaccgcag cagcagcagc agcagcgcca	26820
gggctaccgc tcgtggcgcg ggcacaagaa cgccatagtc gcctgcttgc aagactgcgg	26880
gggcaacatc tctttcgccc gccgcttct gctattccac cacggggtcg cttttccccg	26940
caatgtctg cattactacc gtcatctcta cagcccctac tgcagcggcg acccagaggc	27000
ggcagcggca gccacagcgg cgaccaccac ctaggaagat atcctccgcg ggcaagacag	27060
cggcagcagc ggccaggaga cccgcggcag cagcggcggg agcggtgggc gactgcgcc	27120
tctcgcccaa cgaaccctc tcgaccggg agctcagaca caggatcttc cccactttgt	27180
atgccatctt ccaacagagc agaggccagg agcaggagct gaaaataaaa aacagatctc	27240
tgcgtccct caccgcagc tgtctgtatc acaaaagcga agatcagctt cggcgcacgc	27300
tggaggacgc ggaggcactc ttacagcaat actgcgcgt cactcttaaa gactagctcc	27360
gcgcccttct cgaatttagg cgggagaaaa ctacgtcatc gccggccgcc gccagccccg	27420
cccagccgag atagcaaa agattccac gccatacatg tggagctacc agccgcagat	27480
gggactcgcg ggggagcgg cccaggacta ctccaccgc atgaactaca tgagcgggg	27540
acccacatg atctcacagg tcaacgggat ccgcgccag cgaaacaaa tactgtgga	27600
acaggcggcc ataccgcca cgccccgcca taatctcaac ccccgaaatt ggcccccg	27660
cctagtgtac caggaaacc cctccgccac caccgtacta cttccgcgtg acgcccaggc	27720
cgaagtccag atgactaact caggggcgca gctcgcgggc ggctttcgtc acggggcgcg	27780
gccgtccga ccaggtataa gacacctgat gatcagaggc cgaggatatc agctcaacga	27840
cgagtccgtg agctcttcgc tcggctctcg tccggacgga actttccagc tcgccggatc	27900
cggctcgtct tcgttcacgc cccgccaggc gtacctgact ctgcagacct cgtcctcgga	27960
gccccgtcc ggcgcatcg gaacctcca gtctgtggag gatttcgtgc ctcgggtcta	28020
cttcaacccc ttctcgggac ctcccgacg ctaccccgac cagttcattc cgaactttga	28080

cgcggtgaag gactcggcgg acggctacga ctgaatgtca ggtgccgagg cagagcagct	28140
tgcctgaga cacctcgagc actgccgccg ccacaagtgc ttgcccgcg gttccggtga	28200
gttctgctac tticagctac ccgaggagca taccgagggg ccggcgcaag gcgtccgcct	28260
gaccaccag ggcgagggtta cctgttcct catccgggag ttaccctcc gtccctgct	28320
agtggagcgg gagcggggtc cctgtgtcct aactatcgcc tgcaactgcc ctaaccttg	28380
attacatcaa gatctttgct gtcatctctg tctgagttt aataaacgt gagatcagaa	28440
tctactgggg ctctgtcgc catcctgtga acgccaccgt cttacccac ccgaccagg	28500
cccaggcgaa cctcacctgc ggtctgcac ggaggtccaa gaagtacctc acctggtact	28560
tcaacggcac cccctttgtg gtttacaaca gcttcgacgg ggacggagtc tccctgaaag	28620
accagctctc cggctctcagc tactccatcc acaagaacac caccctcaa ctcttcctc	28680
cctacctgcc gggaacctac gagtgcgtca ccggccgctg caccacctc accgcctga	28740
tcgtaaacca gagctttccg ggaacagata actccctctt cccagaaca ggaggtgagc	28800
tcaggaaact ccccggggac caggcgagg acgtacctc gaccttgtg gggttaggat	28860
ttttattac cgggttgcgt gctcttttaa tcaaagctc cttgagattt gttcttctc	28920
tctacgtgta tgaacacctc agcctccaat aactctacc tttcttcggg atcaggtgac	28980
ttttctgaaa tcgggcttgg tgtgtgctt actctgtga ttttttctc taccatactc	29040
agccttctgt gcctcaggct cgccgcctgc tgcgcacaca tctatatcta ctgctggttg	29100
ctcaagtga ggggtcgcca cccaagatga acaggtacat ggtctatcg atcctaggcc	29160
tgttgccct ggcggcctgc agcgccgcca aaaaagagat tacctttgag gagcccgctt	29220
gcaatgtaac tticaagccc gaggtgacc aatgcaccac cctcgtcaaa tgcgttacca	29280
atcatgagaa gtgcgcac gactacaaaa aaaaaactgg ccggtttgcg gtctatagt	29340
tgtttacgcc cggagacccc tctaactact ctgtcaccgt cttccagggc ggacagtcta	29400
agatatcaa ttacacttc ctttttatg agttgtgca tgcggtcatg tacatgtcaa	29460
aacagtacaa cctgtggcct cctctcccc aggcgtgtgt ggaaaatact gggctctact	29520
gctgtatggc ttggcaatc actacgctcg ctctaactcg cacggtgcta tatataaat	29580
tcaggcagag gcgaatcttt atcgatgaaa agaaaatgcc ttgatcgcta acaccggctt	29640
tctatctgca gaatgaatgc aatcacctcc ctactaatca ccaccacct ccttgcgatt	29700
gcccattgggt tgacacgaat cgaagtcca gtgggtcca atgtcacat ggtgggcccc	29760
gccggcaatt ccacctcat gtgggaaaaa ttgtccgca atcaatgggt tcatttctgc	29820

tctaaccgaa tcagtatcaa gcccagagcc atctgcatg ggcaaaatct aactctgac	29880
aatgtgcaaa tgatggatgc tgggtactat tacgggcagc ggggagaaat cattaattac	29940
tggcgacccc acaaggacta catgctgcat gtagtcgagg cacttccac taccaccccc	30000
actaccacct ctcccaccac cactaccact actactacta ctactaccac taccgctgcc	30060
cgccatcccc gcaaaagcac catgattagc acaaagcccc ctctgctca ctcccacgcc	30120
ggcgggcccc tcggtgcgac ctacagaaacc accgagcttt gcttctgcc atgcactaac	30180
gccagcgctc atgaactgtt cgacctggag aatgaggatg cccagcagag ctccgcttgc	30240
ctgacccagg aggtctgtga gcccgttgcc ctgaagcaga tcggtgatc aataattgac	30300
tcttctctt ttgccactcc cgaataccct cccgattcta ctttccacat cacgggtacc	30360
aaagacccta acctctctt ctacctgatg ctgctgctct gtatctctgt ggtctcttcc	30420
gcgctgatgt tactggggat gttctgctgc ctgatctgcc gcagaaagag aaaagctcgc	30480
tctcagggcc aaccactgat gcccttcccc taccctccgg attttgaga taacaagata	30540
tgagctcgt gctgacacta accgctttac tagcctgcgc tctaaccctt gtcgcttgcg	30600
actcgagatt ccacaatgtc acagctgtgg caggagaaaa tgttacttcc aactccacgg	30660
ccgatacca gtggtcgtgg agtggctcag gtagctactt aactatctgc aatagctcca	30720
cttccccag catatcccc accaagtagc aatgcaatgc cagcctgttc accctcatca	30780
acgtttccac cctggacaat ggactctatg taggctatgt acccttttgt gggcaaggaa	30840
agaccacgc ttacaacctg gaagttcgcc agcccagaac cactaccaa gctwcmcca	30900
ycaccagcac cagcagcagc agccacagca gcagcagcag attattgact ttggttttgg	30960
ccagctcatc tgccgttacc caggccatct acagctctgt gcccgaaacc actcagaccc	31020
accgcccaga aacgaccacc gccaccacc tacacacctc cagcgatcag atgccgacca	31080
acatcacccc ctggtctctt caaatgggac ttacaagccc cactccaaaa ccagtggatg	31140
cggccgaggt ctccgcctc gtcaatgact gggcggggct gggaatgtgg tggttcgcca	31200
taggcatgat ggcgtctgc ctgcttctgc tctggctcat ctgctgcctc caccgcaggc	31260
gagccagacc ccccatctat agaccatca ttgtctgaa ccccgataat gatgggatec	31320
atagattgga tggcctgaaa aacctacttt tttcttttac agtatgataa attgagacat	31380
gcctcgatt tttctgtaca tgttcttct cccacctttt ctggggtgtt ctacgtggc	31440
cgctgtgtct cacttggagg tagactgcct ctacccttc actgtctacc tgctttacgg	31500
attggtcacc ctactctca tctgcagcct aatcacagta atcatgcct tcatccagtg	31560
cattgattac atctgtgtgc gcctcgata ctccagacac caccgcagc accgagacag	31620
gaacattgcc caacttctaa gactgctcta atcatgcata agactgtgat ctgccttctg	31680

atcctctgca tcttgcacac cctcacctcc tgccagtaca ccacaaaatc tccgcgcaaa	31740
agacatgcct cctgccgctt cacccaactg tggaatatac ccaaatgcta caacgaaaag	31800
agcgagctct ccgaagcttg gctgtatggg gtcactctgtg tcttagtttt ctgcagcact	31860
gtctttgccc tcatgatcta cccctacttt gatttgggat ggaacgcgat cgatgccatg	31920
aattacccca cttttccgc acccgagata attccactgc gacaagtgtg acccgttgtc	31980
gttaatcaac gcccccatc ccctacgcc actgaaatca gctacttta ctaacaggc	32040
ggagatgact gacgccctag atctagaaat ggacggcatc agtaccgagc agcgtctcct	32100
agagaggcgc aggcaggcgg ctgagcaaga gcgcctcaat caggagctcc gagatctcgt	32160
taacctgcac cagtgcacaa gaggcatttt ttgtctggta aagcaggcca aagtcaccta	32220
cgagaagacc ggcaacagcc accgcctcag ttacaaattg cccaccagc gccagaagct	32280
ggtgctcatg gtgggtgaga atcccatcac cgtcaccag cactcggtag agaccgaggg	32340
gtgtctgcac tccccctgc ggggtccaga agacctctgc accctggtaa agacctgtg	32400
cggctctcaga gatttagtcc cttttaacta atcaaacact ggaatcaata aaaagaatca	32460
cttacttaaa atcagacagc aggtctctgt ccagttttatt cagcagcacc tctttccct	32520
cctcccaact ctggtactcc aaacgccttc tggcggcaaa cttctctcac accctgaagg	32580
gaatgtcaga ttcttgctcc tgctccctcg caccactat cttcatgttg ttgcagatga	32640
agcgaccaa aacgtctgac gagagcttca acccgtgta ccctatgac acggaaagcg	32700
gccctccctc cgtcccttc ctcacccctc ccttcgtgtc tccgatgga ttccaagaaa	32760
gtcccccggt ggtcctgtct ctgaacctgg ccgagccct ggtcacttcc cacggcatgc	32820
tcgccctgaa aatgggaagt ggcctctccc tggacgacgc tggcaacctc acctctcaag	32880
atatcaccac cgctagccct cccctcaaaa aaaccaagac caacctcagc ctagaaacct	32940
catccccct aactgtgagc acctcaggcg cctcaccgt agcagcggcc gctccctgg	33000
cgggtggccgg cactccctc accatgcaat cagaggcccc cctgacagta caggatgcaa	33060
aactcaccct ggccacaaa ggccccctga ccgtgtctga aggcaaatg gccttgcaaa	33120
catcgcccc gctgacggcc gctgacagca gcacctcac cgttagcgcc acaccacaa	33180
ttaatgtaag cagtgggaagt ttaggcttag acatggaaga ccctatgtat actcacaatg	33240
gaaaactggg aataagaatt ggggtccac taagagtagt agacagcttg cataactca	33300
ctgtagttag cggaatgga ctaactgtag ataacaatgc cctccaaact aaagttacgg	33360
gcgcctagg ttaagacaca tcaggaaatc tacaattaag agctgcagga ggtatgcgaa	33420
ttgacgcaaa tggccaactt atccttaatg tggcataccc atttgatgct cagaacaatc	33480
tcagccttag acttggtcag ggaccctgt atataaacac agaccacaac ctggatttga	33540

attgcaacag aggtctaacc acaactacca ccaacaacac aaaaaaactt gagactaaaa	33600
ttagctcagg cttagactat gacaccaatg gtgctgtcat tattaaactt ggcactggtc	33660
taagcttcga caacacaggc gccctaactg tgggaaacac tggatgatgat aaactgactc	33720
tgtggacgac ccagaccca tctccaaatt gcagaattca ctgagacaaa gactgcaagt	33780
ttactctagt cctaactaag tgttgaagcc aaatcctggc ctctgtcgcc gccctagcgg	33840
tatcaggaaa tctggcttcg ataacaggca cgtttgccag cgttaccatc tttctcagat	33900
ttgatcagaa tggagtgttt atggaaaact cctcgctaga caggcagtac tggaaacttca	33960
gaaatggcaa ctcaactaac gctgccccct acaccaatgc agttgggttc atgccaaacc	34020
tcgcagcata ccccaaacg caaagccaga ctgctaaaaa caacattgta agtcaggttt	34080
acttgaatgg agacaaatcc aaacccatga cccttaccat caccctcaat ggaactaatg	34140
aatccagtga aactagccag gtgagtcact actccatgtc atttcatggt gcttgggaaa	34200
gtgggcaata tgccactgaa acccttgcca ccaactcctt caccttttct tacattgctg	34260
aacaataaaa agcatgacac tgatgttcat ttctgattct tattttatta ttttcaaaca	34320
caacaaaatc attcaagta ttttccatc ttagcttaat agacacagta gcttaataga	34380
cccagtagtg caaagcccca ttctagctta taactagtgg agaagtactc gcctacatgg	34440
gggtagagtc ataatcgtgc atcaggatag ggcggtgggtg ctgcagcagc gcgcgaataa	34500
actgctgccg ccgccgtcc gtcctgcagg aatacaacat ggacgtggtc tcctcagcga	34560
tgattcgac cggccgcagc ataaggcgcc ttgtcctccg ggacacagcag cgcaccctga	34620
tctacttaa atcagcagc taactgcagc acagcaccac aatatgttc aaaatccac	34680
agtgaaggc gcigtatcca aagctcatgg cggggaccac agaaccacg tggccatcat	34740
accacaagcg caggtagatt aagtggcgac cctcataaa cacgtggac ataaacatta	34800
cctcttttgg catgttgtaa ttcaccact cccgtacca tataaacctc tgattaaaca	34860
tggcgccatc caccaccatc ctaaaccagc tggccaaaac ctgcccgcg gctatacact	34920
gcaggaacc gggactggaa caatgacagt ggagagccca ggactcgtta ccatggatca	34980
tcatgctcgt catgatatca atgttggcac aacacaggca cacgtgcata cacttcctca	35040
ggattacaag ctctcccgc gttagaacca tatcccaggg aacaacccat tcctgaatca	35100
gcgtaaatcc cacactgcag ggaagacctc gcacgtaact cacgttgtgc attgtcaaag	35160
tgttacattc gggcagcagc ggatgatcct ccagtatggt agcgcgggtt tctgtctcaa	35220
aaggaggtag acgatcccta ctgtacggag tgcgccgaga caaccgagat cgtgttggtc	35280

gtagtgtcat gccaaatgga acgccggacg tagtcatatt tcctgaagtc ttagatctct	35340
caacgcagca ccagcaccaa cacttcgcag tgtaaaaggc caagtgccga gagagtatat	35400
ataggaataa aaagtgcgt aaacgggcaa agtccaaaaa acgccagaa aaaccgcacg	35460
cgaacctacg ccccgaaacg aaagccaaaa aacactagac actcccttcc ggcgtcaact	35520
tccgctttcc cagctacgt cacttgcccc agtcaaaaaa actacatatc ccgaacttcc	35580
aagtcgccac gccccaaaaa ccgcctacac ctccccgcc gccggccgc ccccaaaccc	35640
gcctccgcgc ccgcgccccg ccccgccgcg cccatctcat tatcatattg gcttcaatcc	35700
aaaataaggt atattattga tgatggttta aacggatcca attctgaag acgaaagggc	35760
ctcgtgatac gcctatTTTT ataggTTaat gtcagtataa taatggttcc ttagacgtca	35820
ggtggcactt ttcggggaaa tgtgcgcgga acccctatTT gtttattttt ctaaatacat	35880
tcaaataatgt atccgctcat gagacaataa ccttgataaa tgcttcaata atattgaaaa	35940
aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt tgcggcattt	36000
tgccttctcg tttttgctca ccagaaaacg ctggtgaaag taaaagatgc tgaagatcag	36060
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaatat ccttgagagt	36120
tttcgccccg aagaacgttt tccaatgatg agcactttta aagtctctgt atgtggcgcg	36180
gtattatccc gtgttgacgc cgggcaagag caactcggtc gccgcataca ctattctcag	36240
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta	36300
agagaattat gcagtgtcgc cataaccatg agtgataaca ctgcggccaa ctacttctg	36360
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta	36420
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac	36480
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt	36540
actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca	36600
cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag	36660
cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta	36720
gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgtgag	36780
ataggTgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atataactt	36840
tagattgatt taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaaacc	36900
cttaacgtga gttttcgttc cactgagcgt cagaccccg agaaaagatc aaaggatctt	36960
cttgagatcc ttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgtac	37020
cagcgttggg ttgtttgccg gatcaagacg taccaactct ttttccgaag gtaactggct	37080
tcagcagagc gcagatacca aatactgtcc ttctagtgtg gccgtagtta ggccaccact	37140

tcaagaactc tgiagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg 37200
ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata 37260
aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca gcccagcttg gagcgaacga 37320
cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag 37380

ggagaaaggc ggacaggtat ccggaagcg gcagggtcgg aacaggagag cgcacgaggg 37440
agcttccagg gggaaacgcc tggatatctt atagtctgt cggttttcgc cacctctgac 37500
ttgagcgtcg attttttga tgctcgtcag gggggcggag cctatggaaa aacgccagca 37560
acgcggcctt ttacggctc ctggcctttt gctggccttg aagctgtccc tgatggtcgt 37620
catctacctg cctggacagc atggcctgca acgcgggcat cccgatgccg ccggaagcga 37680
gaagaatcat aatggggaag gccatccagc ctgcgctcgc agatccgaat tcgtttaaac 37740

<210> 23

<211> 1572

<212> DNA

<213> Rabies virus

<400> 23

atggtgcctc aagccctgct gttcgtgcc ctgctggtct tctccctctg ctttggcaag 60
ttccccatct acaccatccc tgacaagctc ggcccttgg ccccatga catacatcac 120
ctcagctgcc ccaacaacct ggtgtgtggag gatgagggt gcacaaacct gagcggcttc 180
tcctatatgg aactcaaggt gggctatatc tccgcatca aggtcaatgg attcacatgc 240
accggcgtcg tgacagagcg tgaaacatac accaactttg tgggctacgt caccaccaca 300
ttcaagagga agcacttcag gccaccctc gacgcttgca gggctgccta caattggaag 360
atggctggcg accccaggta tgaggagtcc ctgcacaatc cctaccccga ctaccattgg 420

ctcaggacag tcaagaccac caaggagtcc ctggtcatta tctcccttag cgtggccgac 480
ctagaccctg atgacaaaag cctgcactcc aggtctctcc ctacggcaa atgctccggc 540
attacagtga gtccacctc ctgcagcaca aaccacgact acaccatctg gatgcctgag 600
aatcctaggc tcggcacctc ctgtgacata ttacaaata gcaggggcaa gagggcttcc 660
aaaggcagca aaacctgcgg ctttgtcgac gaaagaggcc tgtacaagtc cctcaagggc 720
gcttgtaaac tcaagctgtg cggagtgtg ggactcagac tcatggacgg cacatgggtg 780
gccatgcaga ccagcgatga gaccaagtgg tgcccccccg atcagctggt gaatctgcac 840

gacttcaggt ccgacgaaat tgagcacctc gtggtcgagg agctggtgaa gaagagagaa 900
gagtgcctgg atgctctgga gtccatcatg accacaaat ccgtgtcctt cagaaggctg 960

agccacctca ggaagctggt ccccggtttt ggcaaggcct acacaatttt caataagaca 1020
ctgatggagg ccgatgtca ctacaaatcc gtgaggacct ggaacgagat catccccctcc 1080
aaaggtgcc tgagggtggg aggaagatgc cccccccacg tcaacggcgt cttcttcaac 1140
ggcattatcc tcggaccga tggccatgtc ctgatccctg aaatgcaaag ctccctgctg 1200
cagcagcaca tggaactcct ggagagctcc gtcacccccc tgatgcaccc tctcgctgac 1260

cccagcacgg tgtttaaaga cggcgatgag gccgaggact tcgtggaagt gcatctgcct 1320
gatgtgcata agcaagtcag cggcgctgat ctgggcctgc ctaattgggg caagtatgtc 1380
ctgctctccg ccggagctct gattgccctg atgtgatca tcttctgat gacctgctgc 1440
agaagagtca acagacctga gagcacccaa agatccctcg gcggaaccgg aaggaaggtc 1500
agcgtgacca gccagtccgg caaagtgatt tcctcctggg agagctataa aagcggcgga 1560
gagaccaggc tg 1572

<210> 24

<211> 37536

<212> DNA

<213> Pan troglodytes

<220><221> misc_feature

<222> (24187)..(24189)

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

<400> 24

catcatcaat aatatacttt attttggatt gaagccaata tgataatgag atgggcggcg 60
cggggcgggg cgcgggcggg gaggcgggtt tggggcgggg ccggcgggcg gggcggtgtg 120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag 180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttccgc ggtttttacc 240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact 300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatitgtcta 360

gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggatgaat 420
ttccgcgttc cgggtcaaa gctgcgtttt attattatag gatatcccat tgcatacgtt 480
glatccatat cataatatgt acatttatat tggctcatgt ccaacattac cgccatgttg 540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagcc 600
atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccc 660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caataggac 720

tttccattga cgtaaatggg tggagtattt acggtaaact gccacttgg cagtacatca	780
agtgatcat atgccaagta cccccctat tgacgtcaat gacggtaaat ggcccgcctg	840
gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca tctacgtatt	900
agtcacgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgtttg	1020
gaacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat	1080
gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctcccta tcagtatag	1140
agatctccct atcagtata gagatcgtcg acgagctcgt ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggcccg gaacggtgca ttggaacgcg gattccccgt gccaagagtg agatcttccg	1320
tttatctagg taccagatat cgccaccatg gtgcctcaag cctgctgtt cgtgccctg	1380
ctggtcttct cctctgctt ttgcaagttc ccatctaca ccatccctga caagctcggc	1440
ccctggtccc ccatgacat acatcacctc agctgcccc acaacctggt ggtggaggat	1500
gagggtgca caaacctgag cggtctctcc tatatggaac tcaaggtggg ctatatctcc	1560
gccatcaagg tcaatggatt cacatgcacc ggcgtcgtga cagaggctga aacatacacc	1620
aactttgtgg gctacgtcac caccacattc aagaggaagc acttcaggcc caccctgac	1680
gcttgagggt ctgcctacaa ttggaagatg gctggcgacc ccaggtatga ggagtcctg	1740
cacaatccct accccgacta ccattggctc aggacagtca agaccacaa ggagtcctg	1800
gtcattatct ccctagcgt ggccgaccta gaccgctatg acaaaagcct gcactccagg	1860
gtcttcccta gcggcaaatg ctccggcatt acagttagct ccactactg cagcacaaac	1920
cagactaca ccatctggat gcctgagaat cctaggtcgt gcacctctg tgacatattt	1980
acaaatagca ggggcaagag ggcttccaaa ggcagcaaaa cctgcggctt tgtcgacgaa	2040
agaggcctgt acaagtcct caaggcgct tglaaactca agctgtcgg agtgctggga	2100
ctcagactca tggacggcac atgggtggcc atgcagacca gcgatgagac caagtgtgc	2160
cccccgatc agctggtgaa tctgcacgac ttcaggtcg acgaaattga gcacctcgtg	2220
gtcgaggagc tggtaagaa gagagaagag tgcttgatg ctctggagtc catcatgacc	2280
accaaatecg tgtccttcag aaggctgagc cacctcagga agctgggtccc cggtttggc	2340
aaggcctaca caattttcaa taagacactg atggaggccg atgctcacta caaatccgtg	2400
aggacctgga acgagatcat cccctccaaa ggctgcctga gggtgggagg aagatgccac	2460

ccccacgtca acggcgtctt cttcaacggc attatcctcg gacccgatgg ccatgtcctg	2520
atccctgaaa tgcaagctc cctgtctcag cagcacatgg aactcctgga gagctccgtc	2580
atccccctga tgcacctct cgtgacccc agcaccgtgt ttaaagacgg cgatgaggcc	2640
gaggacttcg tggaagtga tctgcctgat gtgcataagc aagtcagcgg cgtcgatctg	2700
ggcctgccta attggggcaa gtatgtcctg ctctccgccg gagctctgat tgccctgatg	2760
ctgatcatct tctgatgac ctgtgcaga agagtcaaca gacctgagag cacccaaaga	2820
tccctcggcg gaaccggaag gaaggtcagc gtgaccagcc agtcggcaa agtgatttcc	2880
tcctgggaga gctataaaag cggcggagag accaggctgt gatgagcggc cgcgatctgc	2940
tgtgccttct agttgccagc catctgttgt ttgccctcc cccgtgcctt ccttgacct	3000
ggaaggtgcc actccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct	3060
gagtagtgtt cattctattc tggggggtgg ggtggggcag gacagcaagg gggaggattg	3120
ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggccgatc agcgatcgct	3180
gaggtgggtg agtgggcgtg gcctgggggtg gtcatgaaaa tatataagtt gggggtctta	3240
gggtctcttt atttgtgtg cagagaccgc cggagccatg agcgggagca gcagcagcag	3300
cagtagcagc agcgccttgg atggcagcat cgtgagccct tatttgacga cgcggatgcc	3360
ccactgggcc ggggtgcgtc agaattgat gggctccagc atcgacggcc gaccttct	3420
gcccgcaaat tccgccagc tgacctatgc gaccgtcgcg gggacgccgt tggacgccac	3480
cgccgccgcc gccgccaccg cagccgcctc ggccgtgcgc agcctggcca cggactttgc	3540
attcctggga ccaactggca caggggctac ttctcgggcc gctgctgccg ccgttcgca	3600
tgacaagctg accgccctgc tggcgcagtt ggatgcgctt actcgggaac tgggtgacct	3660
ttctcagcag gtcatggccc tgcgccagca ggtctcctcc ctgcaagctg gcgggaatgc	3720
ttctcccaca aatgcctgtt aagataaata aaaccagact ctgtttggat taaagaaaag	3780
tagcaagtgc attgtctctt ttatttcata attttccgcg cgcgataggc cctagaccag	3840
cgttctcgtt cgttgagggt gcggtgtatc ttctccagga cgtggtagag gtggctctgg	3900
acgttgagat acatgggcat gagcccgctc cgggggtgga ggtagcacca ctgcagagct	3960
tcatgtctcg gggtaggtgt gtagatgatc cagtcgtagc aggagcgtg ggcatgggtc	4020
ctaaaaatgt ccttcagcag caggccgatg gccaggggga ggcccttggg gtaagtgttt	4080
acaaaacggt taagtggga aggggtgcatt cggggagaga tgatgtgcat cttggactgt	4140
atttttagat tggcgtgtt tccgccaga tccttcttgg gattcatgtt gtgcaggacc	4200
accagtacag tgtatccgtt gcacttgggg aatttgtcat gcagcttaga gggaaaagcg	4260
tggaagaact tggagacgcc cttgtggcct ccagatttt ccatgcattc gtccatgatg	4320

atggcaatgg gcccgcgga ggcagcttgg gcaaagatat ttctggggtc gctgacgtcg	4380
tagttgtgtt ccagggtgag gtcgtcatag gccattttta caaagcgcg gcgagggtg	4440
cccactggg ggatgatgtt cccctctggc cccggggcgt agttgccctc gcagatctgc	4500
atttcccagg ccttaatctc ggagggggga atcatacca cctgcggggc gatgaagaaa	4560
acggtttccg gagccggga gattaactgg gatgagagca ggtttctaag cagctgtgat	4620
tttccacaac cggtagggcc ataaataaca cctataaccg gttgcagctg gtagtttaga	4680
gagctgcagc tgccgtcgtc ccggaggagg ggggccacct cgttgagcat gtccctgacg	4740
cgcattgtct ccccgaccag atccgccaga aggcgctcgc cggccaggga cagcagctct	4800
tgcaaggaag caaagttttt cagcggcttg aggccgtccg ccgtgggcat gtttttcagg	4860
gtctggctca gcagctccag gcggtcccag agctcgggtga cgtgctctac ggcatctcta	4920
tccagcatat ctctctgttt cgcgggttgg ggcgactttc gctgtagggc accaagcgg	4980
ggctgtccag cggggccaga gtcatgtcct tccatggcg cagggtcctc gtcagggtgg	5040
tctgggtcac ggtgaagggg tgcgtccgg gctgagcgt tgccaaggtg cgcttgaggc	5100
tggttctgct ggtgctgaag cgctgccgtt ctctgccctg cgcgtcgcc aggtagcatt	5160
tgaccatggt gtcatagtc agccctccg cggcgtgtcc cttggcgcgc agcttgccct	5220
tggaggtggc gccgcacgag gggcagagca ggctcttgag cgcgtagagc ttgggggcga	5280
ggaagaccga ttccggggag taggcgtccg cgcgcagac cccgcacacg gtctgcact	5340
ccaccagcca ggtgagctcg gggcgcgcg ggtcaaaaac caggtttccc ccatgctttt	5400
tgatgcgttt cttaacctcg gtctccatga ggtggtgtcc ccgtcgggtg acgaagaggc	5460
gtccgtgtc tccgtagacc gacttagagg gtcttttctc caggggggtc cctcgtctt	5520
cctcgtagag gaactcggac cactctgaga cgaaggcccg cgtccaggcc aggacgaagg	5580
aggctatgtg ggaggggtag cggctgttgt ccactagggg gtcaccttc tccaaggtgt	5640
gaagacacat gtgccttcc tcggcgtcca ggaaggtgat tggctttag gtgtaggcca	5700
cgtgaccggg ggttctgac ggggggtat aaaaggggtt gggggcgcgc tcgtcgtcac	5760
tctcttccgc atcgtgtct gcgaggcca gctgctgggg tgagtattcc ctctgaagg	5820
cgggcatgac ctccgcgtg aggttgtcag ttccaaaaa cgaggaggat ttgatgttca	5880
cctgtccga ggtgatacct ttgagggtac ccgcgtccat ctggtcagaa aacacgatct	5940
ttttattgtc cagcttggtg gcgaacgacc cgtagagggc gttggagagc agcttgcgca	6000
tggagcgcag ggtctggttc ttgtccctgt cggcgcgtc cttggccgcg atgttgagct	6060
gcacgtactc gcgcgcgacg cagcgcact cggggaagac ggtggtgcgc tcgtcgggca	6120
ccaggcgcac gcgccagccg cggttgtgca gggtgaccag gtccacgtg gtggcgacct	6180

cgccgcgcag gcgctcgttg gtccagcaga gacggccgcc cttgcgcgag cagaaggggg 6240

gcagggggtc gagctgggtc tcgtccgggg ggtccgcgtc cacggtgaaa accccggggc 6300

gcaggcgcg c gtcgaagtag tctatcttgc aaccttgcac gtccagcgcc tgctgccagt 6360

cgcgggcggc gagcgcgcg tcgtaggggt tgagcggcgg gcccagggc atggggtggg 6420

tgagtgcgga ggctacatg ccgcagatgt catagacgta gaggggctcc cgcaggaccc 6480

cgatgtaggt ggggtagcag cggccgccgc ggatgctggc gcgcacgtag tcatacagct 6540

cgtgcgaggg ggcgaggagg tcggggccca ggttgggtgc ggcggggcgc tccgtgcgga 6600

agacgatctg cctgaagatg gcatgcgagt tggaagagat ggtggggcgc tggaagacgt 6660

tgaagctggc gtctgcagg ccgacggcgt cgcgcacgaa ggaggcgtag gagtgcgcga 6720

gcttgtgtac cagctcggcg gtgacctgca cgtcgagcgc gcagtagtcg agggctctgc 6780

ggatgatgtc atatitgacc tgcccccttct tttccacag ctgcgggttg aggacaaact 6840

cttcgcggtc ttccagtlac tcttggatcg ggaaaccgtc cggttccgaa cggttaagagc 6900

ctagcatgta gaactgggtg acggcctggg aggcgcagca gcccttctcc acggggaggg 6960

cgtaggcctg cgcggccttg cggagcgagg tgtgggtcag ggcaaggtg tccctgacca 7020

tgactttgag gtactgggtg ttgaagtcgg agtcgtcgca gccgccccgc tcccagagcg 7080

agaagtcggg gcgcttcttg gagcgggggt tgggcagagc gaaggtgaca tcgttgaaga 7140

ggattttgcc cgcgcggggc atgaagttgc gggatgatgc gaaggggccc ggcacttcag 7200

agcgggtgtt gatgacctgg gcggcgagca cgatctcgtc gaagccgttg atgttggc 7260

ccacgatgta gattccagg aagcggggcc ggccctttac ggtgggcagc ttcttttagct 7320

cttcgtaggt gagctcctcg ggcgaggcga ggccgtgctc ggccagggcc cagtccgcga 7380

ggtgcgggtt gtctctgagg aaggactccc agaggtcgcg ggccaggagg gtctgcagc 7440

ggtccctgaa ggtcctgaac tggcggccca cgccatttt ttcgggggtg atgcagtaga 7500

aggtgagggg gtcttgctgc cagcgggtccc agtcgagctg cagggcgagg tcgcgcgcgg 7560

cgggtgaccag gcgctcgtcg cccccgaatt tcatgaccag catgaaggc acgagctgct 7620

ttccgaaggc ccccatccaa gtgtaggtct ctacatcgta ggtgacaaaaggcgcctccg 7680

tgcgaggatg cgagccgac gggaagaact ggatctccc ccaccagttag gaggagtggc 7740

tgttgatgtg gtggaagtag aagtccgtc gccgggccga acactcgtgc tggcttttgt 7800

aaaagcgagc gcagtactgg cagcgtgca cgggctgtac ctcatgcacg agatgcacct 7860

ttcggccgcg cagaggaag ccgaggggaa atctgagccc cccgcctggc tcgcggcatg 7920

gctgggtgctc ttctactttg gatgcgtgtc cgtctccgtc tggctcctcg aggggtgtta 7980
cgggtggagcg gaccaccacg ccgcgcgagc cgcaggtcca gatacggcg cgcggcggtc 8040
ggagtttgat gacgacatcg cgcagctggg agctgtccat ggtctggagc tcccgcggcg 8100
gcggcaggtc agccgggagt tcttcaggt tcacctcgca ggtcgggcc agggcgcggg 8160
gcaggtctag gtgtacctg atctctaggg gcgtgttggt ggcgcgctcg atggcttgca 8220
ggagcccga gccccgggg gcgacgacgg tcccccggg ggtggtggtg gtggtggcg 8280
tgcagctcag aagcggtgcc gcggcggggc ccccgagggt agggggggct ccggtcccgc 8340

gggcaggggc ggcagcgga cgtcggcgtg gagcgcgggc aggagtgtgt gctgtgccc 8400
gaggttgctg gcgaaggcga cgaccgggc gttgatctcc tggatctggc gcctctgct 8460
gaagacgacg gggccggtga gcttgaacct gaaagagagt tcgacagaat caatctcggt 8520
gtcattgacc gcggcctggc gcagatctc ctgcacgtct cccagattgt ctggtaggc 8580
gatctcgcc atgaactgct cgatctcttc ctctggagg tctccgctc cggcgcttc 8640
cacggtggcc gccaggtcgt tggagatgcg ccccatgagc tgcgagaagg cgttgagtc 8700
gccctcgttc cagactcggc tgtagaccac gccccctgg tcatcgcggg cgcgcatgac 8760

cacctgcgcg aggttgagct ccacgtgccg gcgaagacg gcgtagtgc gcagacgtg 8820
gaagaggtag ttgagggtgg tggcgggtg ctcgccacg aagaagttca tgaccagcg 8880
gcgaacgtg gattcgtga tgtecccaaa ggcctccagc cgttccatgg cctcgtagaa 8940
gtccacggcg aagttgaaa actgggagtt gcgcgccgac acggtcaact cctcctcag 9000
aagacgatg agctcgcgga cgggtgcgcg cacctgcgc tcgaaggcta tggggtctc 9060
ttctccgt agcatcacca cctctcttc ttctctct tctggcactt ccatgatggc 9120
ttctctct tgggggggtg gcggcgcgcg cgggtgggga gggggcgctc tgcgccggcg 9180

gcggcgacc gggaggcggc ccacgaagcg cgcgatcatc tcccgcggc ggcggcgcat 9240
ggtctcgggt acggcgcggc cgttctccc gggcgcgagt tggaagacg gcccgacat 9300
ctggtgctgg gcgggtggc cgtgaggcag cgagacggcg ctgacgatgc atctcaaaa 9360
ttgctgcgta ggtacgccg cgagggaact gagggagtcc atatccaccg gatccgaaaa 9420
cctttcgagg aaggcgtcta accagtcgca gtcgaagggt aggctgagca ccgtggcggg 9480
cggcgggggg tggggggagt gtctggcgga ggtgctgctg atgatgta tgaagtaggc 9540
ggacttgaca cggcggtagg tcgacaggag caccatgtcc ttgggtccg cctgctggat 9600

gcggaggcgg tcggctatgc ccaggcttc gttctggcat cggcgaggt ccttgtagta 9660
gtcttgcatg agcctttcca ccggcacctc ttctcttcc tcttctgctt ctccatgtc 9720
tgcttcggcc ctggggcggc gccgcgccc cctgcccc atgcgctga cccgaaccc 9780

cctgagcggg tggagcaggg ccaggtcggc gacgacgcgc tcggccagga tggcctgctg	9840
cacctgctg aggggtggtt ggaagtcac caagtcacg aagcgggtgt aggcgcccgt	9900
gttgatggtg taggtgcagt tggccatgac ggaccagtgt acggtctggt gggccggttg	9960
cgacatctcg gtgtacctga gtcgcgagta ggcgcgggag tcgaagacgt agtcgttgca	10020
agtccgcacc aggtactggt agcccaccag gaagtgcggc ggcggctggc ggtagagggg	10080
ccagcgcagg gtggcggggg ctccgggggc caggtcttcc agcatgaggc ggtggtaggc	10140
glagatgtac ctggacatcc aggtgatacc cgcggcgggtg gtggaggcgc gcgggaagtc	10200
gcgcaccceg ttccagatgt tgcgcagggg cagaaaagtc tccatggtag gcgtgctctg	10260
tccagtcaga cgcgcgcagt cgttgatact ctagaccagg gaaaacgaaa gccggtcagc	10320
gggcactctt ccgtggtctg gtgaatagat cgcaagggtg tcatggcgga gggcctcggg	10380
tcgagccccg ggtccgggcc ggacgggtccg ccatgatcca cgcggttacc gcccgcgtgt	10440
cgaaccagg tgtgcgacgt cagacaacgg tggagtgttc cttttggcgt ttttctggcc	10500
gggcgcgggc gccgcgtaag agactaagcc gcgaaagcga aagcagtaag tggctcgctc	10560
cccgtagccg gagggatcct tgctaagggt tgcgttgagg cgaaccccg ttcgaatccc	10620
glactcgggc cgcccgacc cgcgcctaag gtgttggtt ggcctcccc tcgtataaag	10680
accccgcttg cggattgact ccggacacgg ggacgagccc cttttatttt tgctttcccc	10740
agatgcattc ggtgttgca cagatgcgc ccccgcccca gcagcagca caacaccagc	10800
aagagcggca gcaacagcag cgggagtcac gcaggggccc ctcaccacc ctcggcggcc	10860
cggccacctc ggcgtccgag gccgtgtctg gcgcctgcgg cggcggcgcc gggggccgg	10920
ctgacgacc cgaggagccc ccgcggcgca ggccagaca ctacctggac ctggaggagg	10980
gcgagggcct ggcgcggctg ggggcgccgt ctcccagcgc ccaccgcgg gtgcagctaa	11040
agcgcgactc gcgcgaggcg tacgtgcctc ggcagaaact gttcaggagc gcgcggggcg	11100
aggagcccga ggagatgcgg gacaggaggt tcagcgcggg gcgggagctg cggcaggggc	11160
tgaaccgca gggctgctg cgcgaggagg actttgagcc cgacgcgcgg acggggatca	11220
gccccgcgcg cgcgcacgtg gcggccgccg acctggtgac ggcgtacgag cagacggtga	11280
accaggagat caacttcaa aagagtttca acaaccacgt gcgcacgtg gtggcgcgcg	11340
aggaggtgac catcgggctg atgcacctgt gggactttgt gagcgcgtg gtgcagaacc	11400
ccaatagcaa gcctctgacg gcgcagctgt tcctgatagt gcagcacagc agggacaacg	11460
aggcggttag ggacgcgtg ctgaacatca ccgagcccga gggccggttg ctgctggacc	11520
tgattaacat cctgcagagc atagtgggtg aggagcgcag cctgagcctg gccgacaagg	11580
tggcgcccat caactactcg atgctgagcc tgggcaagtt ttacgcgcgc aagatctacc	11640

agacgccgta cgtgcccata gacaaggagg tgaagatcga cggtttttac atgcgcatgg	11700
cgctgaaggt gctcaccta agcgacgacc tgggcgtgta ccgcaacgag cgcattccaca	11760
aggccgtgag cgtgagccgg cggcgcgagc tgagcgaccg cgagctgatg catagcctgc	11820
agcgggcgct ggcggggccc ggcagcggcg acagggaggc ggagtcctac ttcgatgcgg	11880
gggcggaacct gcgtgggcg cccagccggc ggcccctgga ggccgcgggg gtccgcgagg	11940
actatgacga ggacggcgag gaggatgagg agtacgagct agaggagggc gagtacctgg	12000
actaaaccgc gggtaggtgt tccgtagat gcaagaccg aacgtggtgg acccgcgct	12060
gcgggcggt ctgcagagcc agccgtccgg ccttaactcc tcagacgact ggcgacaggt	12120
catggaccgc atcatgtcg tgacggcgcg taaccggac gcgttccggc agcagccgca	12180
ggccaacagg ctctcccca tcttgaggc ggtggtgcct gcgcgctcga accccacgca	12240
cgagaaggtg ctggccatag tgaacgcgt ggccgagaac agggccatcc gcccgacga	12300
ggccgggctg gtgtacgac cgctgctgca gcgcgtggcc cgctacaaca gcggcaact	12360
gcagaccaac ctggaccggc tggtagggga cgtgcgcgag gcgtggcg gcgcgagcg	12420
cgcgatcgg cagggaacc tgggtccat ggtggcgctg aatgccttc tgagcacga	12480
gccggccaac gtgccgggg ggcaggaaga ctacaccaac tttgtgagcg cgctgcggct	12540
gatggtgacc gagaccccc agagcgaggt gtaccagtc ggtccggact acttcttcca	12600
gaccagcaga cagggcctgc agacggtgaa cctgagccag gctttcaaga acctgcgggg	12660
gctgtggggc gtgaaggcgc ccaccggcga cggggcagc gtgtccagcc tgctgacgc	12720
caactcggc ctgctgtgc tgctgatgc gccgttcacg gacagcggca gcgtgtcccg	12780
ggacacctac ctggggcacc tgctgacct gtaccgcgag gccatcgggc aggcgaggt	12840
ggacgagcac accttcagg agatcaccag cgttagccgc gcgtggggc aggaggacac	12900
gagcgcctg gaggcgactc tgaactacct gctgaccaac cggcggcaga agattccctc	12960
gctgcacagc ctgacctccg aggaggagcg catcttgcgc tacgtgcagc agagcgtgag	13020
cctgaacctg atgcgcgac gggtagcgc cagtgtggcg ctggacatga ccgcgcgcaa	13080
catggaaccg ggcatgtacg ccgcgcaccg gccttacatc aaccgcctga tggactacct	13140
gcatcgcgcg cggccgtga accccgagta cttaccaac gccatcctga accgcactg	13200
gtcccgccg cccgggttct acagcggggg ctctgaggtc ccggaggcca acgatggctt	13260
cctgtgggac gacatggacg acagcgtgtt ccccccg cgcagggcgc tggcggaagc	13320
gtccctgctg cgtcccaaga aggaggagga ggcgagtcgc cggcgcgga gcagcggcgt	13380

ggctttctctg tccgagctgg gggcggcagc cgccgcgcgc cccgggtccc tggcggcag	13440
cccctttccg agcctggtgg ggtctctgca cagcgagcgc accacccgcc ctcggtgct	13500
gggcgaggac gagtacctga ataactccct gctgcagccg gtgcgggaga aaaacctgcc	13560
tcccgccttc cccaacaacg ggatagagag cctggtggac aagatgagca gatggaagac	13620
ctatgcgcag gaggcacagg acgcgccccg gctccggccg cccacgcggc gccagcgcca	13680
cgaccggcag cgggggctgg tgtgggatga cgaggactcc gcggacgata gcagcgtgct	13740
ggacctggga gggagcgga acccgttcgc gcacctgcgc ccccgctgg ggaggatgtt	13800
ttaaaaaaa aaaaaagcaa gaagcatgat gcaaaaatta aataaaactc accaaggcca	13860
tggcgaccga gcgttggttt cttgtgttcc cttcagtatg cggcgcgcgg cgatgtacca	13920
ggaggggacct cctccctctt acgagagcgt ggtgggcgcg gcggcggcgg cgcctcttc	13980
tccctttgcg tcgcagctgc tggagccgcc gtactgcct ccgcgctacc tgcggcctac	14040
gggggggaga aacagcatcc gttactcgga gctggcgccc ctgttcgaca ccaccgggt	14100
gtacctggtg gacaacaagt cggcggacgt ggctccctg aactaccaga acgaccacag	14160
caattttttg accacggtea tccagaacaa tgactacagc ccgagcgagg ccagcaccca	14220
gaccatcaat ctggatgacc ggtcgcactg gggcggcgac ctgaaaacca tctgcacac	14280
caacatgccc aacgtgaacg agttcatgtt caccaataag ttcaaggcgc gggatgatgt	14340
gtcgcgctcg cacaccaagg aagaccgggt ggagctgaag tacgagtggg tggagttcga	14400
gctgccagag ggcaactact ccgagacat gaccattgac ctgatgaaca acgcgatcgt	14460
ggagcactat ctgaaagtgg gcaggcagaa cggggtcctg gagagcgaca tcggggtcaa	14520
gttcgacacc aggaacttcc gcctggggct ggaccccggt accgggctgg ttatgccgg	14580
ggtgtacacc aacgaggcct tccatcccga catcatcctg ctgcccggt gcggggtgga	14640
cttcacttac agccgcctga gcaacctcct gggcatccgc aagcggcagc cttccagga	14700
gggcttcagg atcacctacg aggacctgga ggggggcaac atccccgcgc tctcgtatgt	14760
ggaggcctac caggatagct tgaaggaaaa tgaggcggga caggaggata ccgccccgc	14820
cgctccgcc gccgcgagc agggcgagga tgtgtctgac accgcggcgc cggacggggc	14880
ggaggccgac cccgctatgg tggaggaggc tgccgagcag gaggaggaca tgaatgacag	14940
tgcggtgcgc ggagacacct tcgtcaccgc gggggaggaa aagcaagcgg aggccgaggc	15000
cgcgcccgag gaaaagcaac tggcggcagc agcggcggcg gcggcgttgg ccgcggcgga	15060
ggctgagtct gaggggacca agcccgcaa ggagcccggtg attaaagccc tgaccgaaga	15120
tagcaagaag cgcatgtaca acctgctcaa ggacagcacc aacaccggt accgcagctg	15180
gtacctggcc tacaactacg gcgacccgtc gacgggggtg cgctcctgga ccctgctgtg	15240

cacgccggac gtgacctgcg gctcggagca ggtgtactgg tcgctgcccc acatgatgca	15300
agaccccggtg accttccgct ccacgcggca ggtcagcaac ttcccgggtg tgggcgccga	15360
gctgctgccc gtgcactcca agagcttcta caacgaccag gccgtctact cccagctcat	15420
ccgccagtgc acctctctga cccacgtgtt caatcgcttt cctgagaacc agattctggc	15480
gcgccccccc gccccacca tcaccaccgt cagtgaatac gttcctgctc tcacagatca	15540
cgggacgcta ccgctgcgca acagcatcgg aggagtccag cgagtgcacg ttactgacgc	15600
cagacgccgc acctgcccc acgtttacaa ggcccttgggc atagtctcgc cgcgcgtcct	15660
ttccagccgc actttttgag caacaccacc atcatgtcca tctgatctc acccagcaat	15720
aactccggct ggggactgct gcgcgcgccc agcaagatgt tcggaggggc gaggaagcgt	15780
tccgagcagc accccgtgcg cgtgcgcggg cacttccgcg ccccttgggg agcgcacaaa	15840
cgcggccgcg cggggcgcac caccgtggac gacgccatcg actcgggtgt ggagcaggcg	15900
cgcaactaca ggcccgcggt ctctaccgtg gacgcggcca tccagaccgt ggtgcggggc	15960
gcgcggcggt acgccaagct gaagagccgc cggaagcgcg tggcccgccg ccaccgccgc	16020
cgaccggggg ccgccccaa acgcgcgccc gcggccctgc ttgcggggc caagcgcacg	16080
ggccgcgccc ccgccatgag ggccgcgcgc cgcttggccg ccggcatcac cggccacc	16140
atggcccccc gtaccgaag acgcgcggcc gccgcgccg ccgccccat cagtgcacatg	16200
gccagcaggc gccggggcaa cgtgtactgg gtgcgcgact cggtgaccgg cacgcgcgtg	16260
cccgtgcgct tccgcccccc gcggacttga gatgatgtga aaaaacaaca ctgagtctcc	16320
tgctgttgtg tgtatccag cggcggcggc ggcgcgcgca gcgtcatgtc caagcgcaaa	16380
atcaaagaag agatgtcca ggtcgtcgcg ccggagatct atgggcccc gaagaaggaa	16440
gagcaggatt cgaagcccc caagataaag cgggtcaaaa agaaaaagaa agatgatggc	16500
gatgccgatg gggagggtga gttcctgccc gccacggcgc ccaggcgccc ggtgcagtgg	16560
aaggcgccgc gcgtaaagcg cgtcctgccc cccggcaccg cgggtggtctt cacgcccggc	16620
gagcgtcca cccggacttt caagcgctc tatgacgagg tgtacggcga cgaagacctg	16680
ctggagcagg ccaacgagcg cttcggagag ttgtcttacg ggaagcgtca gcggccgctg	16740
gggaaggagg acctgctggc gctgccgtg gaccaggga accccacccc cagtctgaag	16800
cccgtgacct tgcagcaggt gctgccgagc agcgaccct cccaggcgaa gcggggtctg	16860
aagcgcgagg gcggcgacct ggccccacc gtgcagctca tggcgccaa gcggcagagg	16920
ctggaggatg tgctggagaa aatgaaagta gaccccggtc tgcagccgga catcagggtc	16980
cgtcccatca agcagggtgc gccgggcctc ggctgtcaga ccgtggacgt ggtcatcccc	17040
accggcaact ccccgccgc caccaccact accgtgcct ccacggacat ggagacacag	17100

accgatcccg ccgcagccgc cgccaccgcc gccgccgcga cctcctcggc ggaggtgcag	17160
acggaccctt ggctgccgcc ggcgatgtca gctccccgcg cgcgtcgcgg gcgcaggaag	17220
tacggcgccg ccaacgcgct cctgcccagag tacgccttgc atccttccat cgcgcccacc	17280
cccggctacc gaggtatata ctaccgcccgc cgaagagcca aggggtccac ccgccgtccc	17340
cgccgacgcg ccgccgccac caccgcccgc cgccgccgca gacgccagcc cgcactggct	17400
ccagtctccg tgaggagagt ggccgcgcgc ggacacaccc tgggtctgcc cagggcgcg	17460
taccaccca gcatcgttta aaagcctgtt gtggttcttg cagatatggc ctcacttgc	17520
cgctccgtt tcccggtgcc gggataccga ggaggaagat cgccgccgag gagggtctg	17580
gccggccgcg gcctgagcgg aggcagccgc cgccgcgacc ggccggcgacg cgccaccagc	17640
cgacgcagtc gggcggggt gctgcccctg ttaatcccc tgatcgccgc ggcgatcgcc	17700
gccgtgccc ggatgcctc cgtggccttg caggcgctcc agaggcattg acagacttgc	17760
aaacttgcaa atatggaata aaacccaat aaaaaagtct agactctcac gctcgttgg	17820
tcctgtgact atttgtaga atggaagaca tcaacttgc gtcgtggcc ccgcgtcacg	17880
gtcgcgccc gttcctggga cactggaacg atatcgccac cagcaacatg agcgggtggc	17940
ccttcagtgt gggctctctg tggagcggca ttaaaagtat cgggtctgcc gttaaaaatt	18000
acggctcccg ggcttgaac agcagcacgg gccagatgtt gagagacaag ttgaaagagc	18060
agaacttcca gcagaagggt gtggagggcc tggcctccgg catcaacggg gtggtggacc	18120
tggccaacca ggccgtgacg aataagatca acagcagact ggacccccgg ccgccgtgg	18180
aggaggtgcc gccggcgctg gagacggtgt ccccgatgg gcgtggcgag aagcgccgc	18240
ggcccgatag ggaagagacc actctggtca cgagaccga tgagccgcc ccgtatgagg	18300
aggccctgaa gcaaggtctg cccaccacgc ggcccatgc gccatggcc accggggtg	18360
tgggcgcga cccccgcc acgtggact tgcctcgcc cgccgatgtg ccgcagcagc	18420
agcagaaggc ggcacagccg ggccgcccgc tgaccgctc ccgttctcc gccgtctc	18480
tgcccgccgc ggccagcgcc ccccgccggg gggtcgcgag gcacggcaac tggcagagca	18540
cgctgaacag catcgtgggt ctgggggtgc ggtccgtgaa gcgccgccga tgctactgaa	18600
tagcttagct aacgtgttgt atgtgtgtat gcgccctatg tcgccgccag aggagctgt	18660
gagtcgcgcg cgttcgcgcg cccaccacca ccaccgccac tccgccctc aagatggcga	18720
ccccatgat gatccgcag tggctgtaca tgcacatctc gggccaggac gcctcgaggt	18780
acctgagccc cgggctggtg cagttcgccc gcgccaccga gagctacttc agcctgagta	18840

acaagtttag gaaccccacg gtggcgccca cgcacgatgt gaccaccgac cggcttcagc	18900
gcctgacgct gcggttcatt cccgtggacc gcgaggacac cgcgtactcg tacaaggcgc	18960
ggttcaccct ggccgtgggc gacaaccgcg tgctggacat ggccctccacc tactttgaca	19020
tccgcggggt gctggaccgg ggccccactt tcaagcctta ctctggcacc gcctacaact	19080
ccctggcccc caagggcgct cccaactcct gcgagtggga gcaattagaa gaagcccagg	19140
ccgctgtgga agacgaagaa ttagaagatg aagacgagga accacaggat gaggcacctg	19200
tgaaaaaaac ccatgtatac gctcaggctc ccttttctgg agaagaaatt actaaaaacg	19260
gtttgcaaat agggtcagat aacacagaag cccagtctaa gcccatatat gcagatccta	19320
cattccagcc tgaaccccaa atcggggaat cccagtggaa tgaggcagat gctacagttg	19380
ccggcggtag agtgctaaag aaatccactc ccatgaagcc atgctatggt tcctatgcaa	19440
gaccacaaaa ctccaatgga ggtcaaggtg tgctggtggc tgatgataag ggggttcttc	19500
aatctaaagt tgaattgcaa tttttttcaa atactactac tcttaatcag cgggagggtg	19560
acgatacaaa accaaaagtg gtgctgtata gcgaagatgt gcacatggaa actccagaca	19620
cccacatttc ttacaagccc acaaaaagcg atgacaattc aaaaatcatg ctgggtcagc	19680
agtccatgcc caacagacct aattacatcg gcttcagaga caactttatc ggccctcatgt	19740
attacaatag cactggcaac atgggagtgc ttgcaggtca ggccctctcag ttgaatgcag	19800
tggtggactt gcaagacaga aacacagaac tgcctacca gctcttgctt gattccatgg	19860
gtgacagaac cagatacttt tccatgtgga atcaggcagt ggacagttat gaccagatg	19920
tcagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt ttccctctgg	19980
gcggcatagg ggtaactgac acttaccagg ccattaaaac caatggcaat ggtcaagaaa	20040
acccaacctg ggaaaaagat acagagtttg cagaccgcaa tgaaataggg gtgggaaca	20100
atttcgctat ggagatcaac ctacgtgcca acctgtggag aaacttctg tactccaacg	20160
tggcgctgta ctgccagac aagcttaagt acaaccctc caatgtggac atctctgaca	20220
acccaacac ctacgattac atgaacaagc gagtgggtggc cccggggctg gtggactgct	20280
acatcaacct gggcgcgcgc tggtcgctgg actacatgga caacgtcaac ccttcaacc	20340
accaccgcaa tgcgggcctg cgctaccgct ccatgctcct gggcaacggg cgctacgtgc	20400
ccttccacat ccaggtgccc cagaagttct ttgccatcaa gaacctctc ctctgcccg	20460
gtcctacac ctacgagtgg aacttcagga aggatgtcaa catggtctc cagagctctc	20520
tgggtaacga tctcagggtg gacggggcca gcatcaagtt cgagagcatc tgcctctacg	20580
ccaccttctt ccccatggcc cacaacacgg cctccacgct cgaggccatg ctcaggaacg	20640
acaccaacga ccagtcttc aatgactacc tctccgcgc caacatgctc taccatac	20700

ccgccaacgc caccaacgtc cccatctcca tcccctcgcg caactgggcg gccttccgcg	20760
gctgggcctt caccgcctc aagaccaagg agacccctc cctgggctcg ggattcgacc	20820
cctactacac ctactcgggc tccattccct acctggacgg caccttctac ctcaaccaca	20880
ctttcaagaa ggctcgggtc accttcgact cctcggtcag ctggccgggc aacgaccgtc	20940
tgctcacccc caacgagttc gaaatcaagc gctcggtcga cggggagggc tacaacgtgg	21000
cccagtgcaa catgaccaag gactgggttc tggccagat gctggccaac tacaacatcg	21060
gctaccaggg ctctacatc ccagagagct acaaggacag gatglactcc ttcttcagga	21120
acttcagcc catgagccgg cagggtggtgg accagaccaa gtacaaggac taccaggagg	21180
tgggcatcat ccaccagcac aacaactcgg gcttcgtggg ctacctcgcc cccaccatgc	21240
gcgagggaca ggcctacccc gccaaattcc cctaccgct cataggcaag accgcggtcg	21300
acagcatcac ccagaaaaag ttctctcgcg atcgaccct ctggcgcatc cccttctcca	21360
gcaacttcat gtccatgggt gcgctctcgg acctgggcca gaacttgctc tacgccaact	21420
ccgccacgc cctcgacatg accttcgagg tcgaccccat ggacgagccc acccttctct	21480
atgttctgtt cgaagtcttt gacgtggtcc gggccacca gccgcaccgc ggcgtcatcg	21540
agaccgtgta ctgcgtacg ccttctcgg ccggcaacgc caccaccta agaagcaagc	21600
cgcagtcac gccgcctgca tgccgtcggg ttccaccgag caagagctca gggccatcgt	21660
cagagacctg ggatgcgggc cctatTTTTT gggcacttcc gacaagcgt tccctggctt	21720
tgttcccca cacaagctgg cctgcgccat cgtcaacac gccggccgcg agaccggggg	21780
cgtgcaactgg ctggccttcg cctggaaccc gcgctccaaa acatgcttcc tctttgacct	21840
cttcggcttt tcggaccagc ggctcaagca aatctacgag ttcgagtacg agggcttgct	21900
gcgtcgcagc gccatgcct cctcggccga ccgctgcgtc accctcgaaa agtccacca	21960
gaccgtgcag gggcccgact cggccgcctg cggctctctc tgctgcatgt ttctgcacgc	22020
ctttgtgcac tggcctcaga gtcccatgga ccgcaacccc accatgaact tgctgacggg	22080
ggtgcccac tccatgtcc agagcccca ggtcgagccc accctgcgcc gcaaccagga	22140
gcagctctac agcttctcgg agcgccactc gcctacttc cgccgccaca gcgcacagat	22200
caggagggcc acctcttct gccacttgca agagatgcaa gaaggtaat aacgatgtac	22260
acactTTTTT ctcaataaat ggcatTTTTT tatttataca agctctctgg ggtattcatt	22320
tcccaccacc accaccacc gccgttgtcg ccatctggct ctatttagaa atcgaaaggg	22380
ttctgccggg agtcgccgtg cgccacgggc agggacacgt tgcgatactg gtagcgggtg	22440
ccccacttga actcgggcac caccaggcga ggcagctcgg ggaagttttc gctccacagg	22500
ctgcgggtca gcaccagcgc gttcatcagg tcgggcgccg agatcttgaa gtcgcagttg	22560

gggccgccgc cctgcgcgcg cgagttgcgg tacaccgggt tgcagcactg gaacaccaac 22620

agcgccgggt gcttcacgct agccagcacg ctgcggtcgg agatcagctc ggcgtccagg 22680

tcctcgcgt tgcacagcgc gaacggggtc atcttgggca cttgcctccc caggaagggc 22740

gcgtgccccg gtttcgagtt gcagtcgcag cgcagcggga tcagcaggtg cccatgcccc 22800

gactcggcgt tggggtacag cgcgcgcgat aaggcctgca tctggcggaa ggccatctgg 22860

gccttggcgc cctccgagaa gaacatgccg caggacttgc ccgagaactg gtttgcgggg 22920

cagctggcgt cgtgcaggca gcagcgcgcg tcggtgttgg cgatctgcac cacgttgcgc 22980

ccccaccgt tcttcacgat cttggccttg gacgattgct ccttcagcgc gcgtgcccc 23040

ttctcgttg tcacatccat ctcgatcaca tgttcttgt tcacatgct gctgccgtgc 23100

aggcacttca gctcgcctc cgtctcgggt cagcgggtgct gccacagcgc gcagcccgtg 23160

ggctcgaaag actttaggt cacctccgcg aaggactgca ggtaccctg caaaaagcgg 23220

cccatcatgg tcacgaaggt cttgttgcg ctgaaggtca gctgcagccc gcggtgctcc 23280

tcgttcagcc aggtcttga cacggccgc agcgcctcca cctggtcggg cagcatcttg 23340

aagttcacct tcagtcatt ctccacgtgg tacttgtcca tcagcgtgcg cgccgcctcc 23400

atgcccttct cccaggccga caccagcggc aggcctcacgg ggttcttcac catcacgtg 23460

gccgcgcgt cgcgcgcgt ttcgtttcc gcccgcgtgt tctcttctc ttctcctct 23520

tcctcgcgc cgcacctcg cagcccccgc accacgggggt cgtcttctg caggcgtgc 23580

accttgcgt tgcggttgc cccctgctt atgcgcacgg gcgggttgc gaagcccacc 23640

atcaccagcg cgccctctc ttgctcgtc tcgctgtcca gaatgacctc cggggagggg 23700

gggttggta tcctcaglac cgaggcacgc tttttttct tcctgggggc gttcgccagc 23760

tcccggtg cgcgcgtcg cgaggtcgaa ggccgagggc tgggcgtgcg cggcaccagc 23820

gcgtcctgc agcgtcctc gtctcctcg gactcgagac ggaggcgggc ccgtttctc 23880

gggggcgcgc ggggcggcgg aggcggcggc ggcgacggag acggggacga gacatcgtcc 23940

aggggtgggtg gacggcgggc cgcgcgcgt ccgcgctcgg ggggtggtctc gcgtgggtcc 24000

tttcccgac tggccatctc ccactgctc ttctcctata ggcagaaaga gatcatggag 24060

tctctcatgc gatlcgagaa ggaggaggac agcctaaccg cccctctga gccctccacc 24120

accgcgcga ccaccgcaa tgccgcgcg gacgacgcgc ccaccgagac caccgccagt 24180

accacnnc tccccagca cgcacccccg ctcgagaatg aagtgtgat cgagcaggac 24240

ccgggttttg tgagcggaga ggaggatgag gtggatgaga aggagaagga ggaggtcgcc 24300

gcctcagtgc caaaagagga taaaaagcaa gaccaggacg acgcagataa ggatgagaca	24360
gcagtcgggc gggggaacgg aagccatgat gctgatgacg gctacctaga cgtgggagac	24420
gacgtgctgc ttaagcacct gcaccgccag tgcgtcatcg tctgcgacgc gctgcaggag	24480
cgggtgcgaag tgcccttggc cgtggcggag gtcagccgcg cctacgagcg gcacctcttc	24540
gcgccgcacg tgcccccaa gcgccgggag aacggcacct gcgagcccaa cccgcgtctc	24600
aacttctacc cggctcttcg ggtacccgag gtgctggcca cctaccacat ctttttcaa	24660
aactgcaaga tcccccttc ctgccgcgct aaccgcaccc gcgccgacaa aacctgacc	24720
ctgccgcagg gcgccacat acctgatatt gcctctctgg aggaagtgcc caagatcttc	24780
gagggtctcg gtcgcgacga gaaacgggcg gcgaacgctc tgcacggaga cagcgaac	24840
gagagtcact cgggggtgct ggtggagctc gagggcgaca acgcgcgcct ggccgtactc	24900
aagcgcagca tagaggtcac ccactttgcc taccggcgcc tcaacctgcc cccaaggtc	24960
atgagtgtgg tcatggcgca gctcatcatg cgccgcgccc agccctggc gcgggatgca	25020
aacttgcaag agtcctcaga ggaaggcctg cccgcggtca gcgacgagca gctggcgcg	25080
tggctggaga cccgcgaccc cgcgcgctg gaggagcggc gcaagctcat gatggccg	25140
gtgctggtca ccgtggagct cgagtgtctg cagcgcttct tcgcggaccc cgagatgcag	25200
cgcaagctcg aggagacct gcactacacc ttccgccagg gctacgtgcg ccaggcctgc	25260
aagatctcca acgtggagct ctgcaacctg gtctcctacc tgggcatcct gcacgagaac	25320
cgctcgggc agaactcct gcactccacc ctcaaagggg aggcgcgcgc cgactacatc	25380
cgcgactgcg cctacctctt cctctgtac acctggcaga cggccatggg ggtctggcag	25440
cagtgcctgg aggagcgcaa cctcaaggag ctggaaaagc tcctcaagcg caccctcagg	25500
gacctctgga cgggcttcaa cgagcgctcg gtggccgccc cgctggcgga catcatcttc	25560
cccagcgccc tgetcaagac cctgcagcag ggctgcccc acttcaccag ccagagcatg	25620
ctgcagaact tcaggacttt catcctggag cgctcgggca tcctgccggc cacttgctgc	25680
gcgtgcccc gcgacttctg gcccatcaag tacaggaggt gcccgccgc gctctggggc	25740
cactgtacc ttttccagct ggccaactac ctgcctacc actcgacct catggaagac	25800
gtgagcggcg agggcctgct cgagtggcac tgccgtgca acctctgac gccccaccgc	25860
tctctagtct gcaaccgca gctgtcagc gagagtcaga ttatcggtac cttcgagctg	25920
cagggtccct cgctgacga gaagtccgcg gctccggggc tgaaactcac tccggggctg	25980
tggacttccg cctacctacg caaatgtgta cctgaggact accacgcca cgagatcagg	26040
ttctacgaag accaatcccg cccgccaag gcggagctca ccgctgcgt catcaccag	26100
gggcacatcc tgggccaatt gcaagccatc aacaaagccc gccgagagtt cttgtgaaa	26160

aagggtcggg ggggtgtacct ggacccccag tccggcgagg agctaaaccc gctacccccg	26220
ccgccgcccc agcagcggga cttgtcttcc caggatggca cccagaaaga agcagcagcc	26280
gccgccgcag ccatacatgc ttctggagga agaggaggag gactgggaca gtcaggcaga	26340
ggagatgatg gaagactggg aggaggacag cagcctagac gaggaagctt cagaggccga	26400
agaggaggca gacgcaacac catcaccttc ggtcgcagcc ccctcgccgg ggcctctgaa	26460
atcctccgaa cccagcacca gcgtataac ctccgtctct ccggcgccgg cgccacccgc	26520
ccgcagaccc aaccgtagat gggacaccac aggaaccggg gtcggtaagt ccaagtcccc	26580
gccgccgcca ccgcagcagc agcagcagca gcgccagggc taccgtctgt ggcgcgggca	26640
caagaacgcc atagtgcct gcttgaaga ctgcgggggc aacatctctt tcgccgcccg	26700
cttctgcta ttccaccagc gggctgcctt tccccgaat gtcctgcatt actaccgtca	26760
tctctacagc ccctactgca gcggcgaccc agaggcggca gcggcagcca cagcgcgac	26820
caccacctag gaagatatcc tccgcgggca agacagcggc agcagcggcc aggagaccg	26880
cggcagcagc ggcgggagcg gtggcgcac tgcgcctctc gcccaacgaa cccctctga	26940
cccgggagct cagacacagg atcttcccca ctttgtatgc catcttcaa cagagcagag	27000
gccaggagca ggagctgaaa ataaaaaca gatctctgcg ctccctcacc cgcagctgtc	27060
tgtatcacia aagcgaagat cagcttcggc gcacgtgga ggacgcggag gcactcttca	27120
gcaaatactg cgcgctcact cttaaagact agctccgcgc cttctcgaa tttagcggg	27180
agaaaactac gtcatcgccg gccgcgccc agcccgccca gccgagatga gcaaagagat	27240
tcccacgcca tacatgtgga gctaccagcc gcagatggga ctgcggcgg gagcgccca	27300
ggactactcc acccgcatga actacatgag cgcgggaccc cacatgatct cacaggtcaa	27360
cgggatccgc gccagcgaa accaaatact gctggaacag gcggccatca ccgccacgc	27420
ccgccataat ctcaaccccc gaaattggcc cgcgcctta gtgtaccagg aaacccctc	27480
cgccaccacc gtactacttc cgcgtgacgc ccaggccgaa gtccagatga ctaactcagg	27540
ggcgagctc gcgggcggct ttctgtcacgg gcgcgcggcg ctccgaccag gtataagaca	27600
cctgatgata agaggccgag gtatccagct caacgacgag tcggtgagct cttcgtcgg	27660
tctccgtccg gacggaactt tccagctcgc cggatccggc cgtcttctgt tcacgccccg	27720
ccaggcgta cgtactcgc agacctcgc ctccgagccc cgtccggcg gcatcggaac	27780
cctccagttc gtggaggagt tcgtgccctc ggtctacttc aaccccttct cgggacctcc	27840
cggacgctac cccgaccagt tcattccgaa ctttgacgcg gtgaaggact cggcgagcgg	27900
ctacgactga atgtcaggtg ccgaggcaga gcagcttcgc ctgagacacc tcgagcactg	27960
ccgccgccac aagtgcttcg cccgcggttc cggtgagttc tgctacttcc agctaccga	28020

ggagcataacc gaggggcccgg cgcacggcgt ccgcctgacc acccagggcg aggttacctg	28080
ttccctcadc cgggagttca cctccgtcc cctgctagtg gagcgggagc ggggtccctg	28140
tgctctaact atcgctgca actgcctaa cctggatta catcaagac tttgtgtca	28200
tctctgtgct gagtttaata aacgtgaga tcagaatcta ctggggctcc tgtcgccatc	28260
ctgtgaacgc caccgtcttc acccaccg accaggccca ggcaaacctc acctgcggtc	28320
tgcatcggag gtccaagaag tacctcacct ggtacttcaa cggcaccccc tttgtggttt	28380
acaacagctt cgacggggac ggagtctccc tgaaagacca gctctccgt ctcagctact	28440
ccatccacaa gaacaccacc ctccaactct tccctcccta cctgccggga acctacgagt	28500
gcgtcaccgg ccgtgcacc cacctaccc gcctgatcgt aaaccagagc tttccgggaa	28560
cagataactc cctcttccc agaacaggag gtgagctcag gaaactccc ggggaccagg	28620
gcggagacgt accttcgacc cttgtgggt taggattttt tattaccggg ttgttggtc	28680
ttttaatcaa agcttccttg agatttggtc tttccttcta cgtglatgaa cacctcagcc	28740
tccaataact ctaccttcc ttcgggatca ggtgactttt ctgaaatcgg gcttggtgtg	28800
ctgcttactc tgttgatttt tttccttacc atactcagcc tttgtgcct caggctcgcc	28860
gcctgtctcg cacacatcta tatctactgc tggttgctca agtgcagggg tcgccacca	28920
agatgaacag gtacatggtc ctatgatcc taggcctgct ggccctggcg gcctgcagcg	28980
ccgcaaaaa agagattacc tttgaggagc ccgcttgcaa tgtaactttc aagcccagg	29040
gtgaccaatg caccaccctc gtcaaatcg ttaccaatca tgagaagctg cgcacgact	29100
acaaaaacaa aactggccgg tttgcggtct atagtgtgt tacgcccga gacctctta	29160
actactctgt caccgtcttc caggcggac agtctaagat attcaattac actttccctt	29220
tttatgagtt gtgcgatcg gtcatgtaca tgcacaaaca gtacaacctg tggcctccct	29280
ctccccaggc gtgtgtggaa aatactgggt ctactgctg tatggctttg gcaatcacta	29340
cgctcgtctt aatctgcag gtgctatata taaaattcag gcagaggcga atctttatcg	29400
atgaaaaaa aatgccttga tcgctaacac cgcttttcta tctgcagaat gaatgcaatc	29460
acctccctac taatcaccac caccctcctt gcgattgccc atgggttgac acgaatcgaa	29520
gtgccagtgg ggtccaatgt caccatggtg ggccccgcc gcaattccac cctcatgtgg	29580
gaaaaatttg tccgaatca atgggttcat tttgtctcta accgaatcag tatcaagccc	29640
agagccatct gcgatgggca aaatctaact ctgatcaatg tgcaaatgat ggatgctggg	29700
tactattacg ggcagcgggg agaaatcatt aattactggc gacccacaa ggactacatg	29760

ctgcatgtag tcgaggcact tcccactacc acccccacta ccacctctcc caccaccact	29820
accactacta ctactactac taccactacc gctgcccgcc ataccgcaa aagcaccatg	29880
attagcacia agccccctcg tgctcactcc cacgccggcg ggcccatcgg tgcgacctca	29940
gaaaccaccg agctttgtct ctgccaatgc actaacgcca gcgctcatga actgttcgac	30000
ctggagaatg aggatgccca gcagagctcc gcttgccctga cccaggaggc tgtggagccc	30060
gttgccctga agcagatcgg tgattcaata attgactctt cttcttttgc cactcccgaa	30120
taccctcccg attctacttt ccacatcacg ggtaccaaag accctaacct ctctttctac	30180
ctgatgctgc tgctctgtat ctctgtggtc tcttccgcgc tgatgttact ggggatgttc	30240
tgctgcctga tctgccgcag aaagagaaaa gctcgctctc agggccaacc actgatgccc	30300
ttccctacc ccccgattt tgcagataac aagatatgag ctgctgctg acactaacg	30360
ctttactagc ctgctctcta acccttgtcg ctgcgactc gagattccac aatgtcacag	30420
ctgtggcagg agaaaatgtt actttcaact ccacggccga taccagtggt tcgtggagtg	30480
gctcaggtag ctacttaact atctgcaata gctccacttc cccagcata tcccaacca	30540
agtaccaatg caatgccage ctgttcaccc tcatcaacgc ttccacctg gacaatggac	30600
tctatgtagg ctatgtaccc tttggtgggc aaggaaagac ccacgcttac aacctggaag	30660
ttgccagcc cagaaccact acccaagctw cymccaycac cagcaccage agcagcagcc	30720
acagcagcag cagcagatta ttgactttgg ttttggccag ctcatctgcc gctaccagg	30780
ccatctacag ctctgtgccc gaaaccactc agaccaccg cccagaaacg accaccgcca	30840
ccacctaca cacctccage gatcagatgc cgaccaacat caccctctg gctcttcaaa	30900
tgggacttac aagccccact ccaaaccag tggatgcggc cgaggtctcc gccctcgtca	30960
atgactgggc ggggctggga atgtggtggt tcgcatagg catgatggcg ctctgcctgc	31020
ttctgctctg gctcatctgc tgccctccacc gcaggcgagc cagaccccc atctatagac	31080
ccatcattgt ccigaacccc gataatgatg ggatccatag attggatggc ctgaaaaacc	31140
tacttttttc ttttacagta tgataaattg agacatgcct cgcattttct tgtacatgtt	31200
ccttctccca cttttctgg ggtgttctac gctggccgct gtgtctcacc tggaggtaga	31260
ctgcctctca ccttccactg tctacctgct ttacggattg gtcacctca ctctcatctg	31320
cagcctaate acagtaata tcgcttcat ccagtgcatt gattacatct gtgtgcgct	31380
cgcatacttc agacaccacc cgcagtaccg agacaggaac attgccaac ttctaagact	31440
gctctaata tgataagac tgtgatctgc cttctgatcc tctgcatcct gccaccctc	31500
acctctgcc agtacaccac aaaatctccg cgcaaaagac atgcctctg ccgcttcacc	31560
caactgtgga atataccaa atgctacaac gaaaagagcg agctctccga agcttggtg	31620

tatggggtca tctgtgtctt agttttctgc agcactgtct ttgccctcat gatctacccc	31680
tactttgatt tgggatggaa cgcgatcgat gccatgaatt accccacctt tccgcaccc	31740
gagataattc cactgcgaca agttgtaccc gttgtcgta atcaacgccc cccatcccct	31800
acgccactg aaatcagcta ctttaacctt acaggcggag atgactgacg ccctagatct	31860
agaaatggac ggcatcagta ccgagcagcg tctcctagag aggcgcaggc aggcggctga	31920
gcaagagcgc ctcaatcagg agctccgaga tctcgtaac ctgcaccagt gcaaaagagg	31980
catcttttgt ctggtaaagc aggccaaagt cacctacgag aagaccggca acagccaccg	32040
cctcagttac aaattgccca cccagcgcca gaagctggtg ctcatggtgg gtgagaatcc	32100
catcaccgtc acccagcact cggtagagac cgagggtgt ctgcactccc cctgtcgggg	32160
tccagaagac ctctgcaccc tggtaaagac cctgtcgggt ctgagagatt tagtcccctt	32220
taactaatca aacactggaa tcaataaaaa gaatcactta cttaaaatca gacagcaggt	32280
ctctgtccag tttattcagc agcactcct tcccctctc ccaactctgg tactccaaac	32340
gccttctggc ggcaaaactt ctccacacc tgaagggaat gtcagattct tgctcctgtc	32400
cctccgacc cactatctt atgttgttgc agatgaagcg caccaaaacg tetgacgaga	32460
gcttcaacc cgigtacccc tatgacacgg aaagcggccc tccctccgtc ctttctctca	32520
ccccctcctt cgtgtctccc gatggattcc aaaaaagtc ccccggtgtc ctgtctctga	32580
acctggccga gccctgggtc acttcccacg gcatgctgc cctgaaaatg ggaagtggcc	32640
tctccctgga cgacgtggc aacctcacct ctcaagatat caccacgct agccctcccc	32700
tcaaaaaaac caagaccaac ctgagcctag aaacctcatc cccctaact gtgagcacct	32760
caggcgccct caccgtagca gccgcgctc ccttggcggg gcccggcacc tccctcacca	32820
tgcaatcaga ggccccctg acagtacagg atgcaaaact caccctggcc accaaaggcc	32880
ccctgaccgt gtctgaagc aaactggcct tgcaaacatc ggccccgtg acggccgctg	32940
acagcagcac cctcaccgtt agcgccacac caccaattaa tgtaagcagt ggaagttag	33000
gcttagacat ggaagacct atgtatact acaatggaaa actgggaata agaattgggg	33060
gtccactaag agtagtagac agcttgcata cactcactgt agttaccgga aatggactaa	33120
ctgtagataa caatgcctc caaactaaag ttacgggcgc cctaggttat gacacatcag	33180
gaaatctaca attaagagct gcaggaggta tgcaattga cgcaaatggc caacttatcc	33240
ttaatgtggc ataccattt gatgtcaga acaatctcag ccttagactt ggtcaggac	33300
ccctgtatat aaacacagac cacaacctgg atttgaattg caacagaggt ctaaccacaa	33360
ctaccaccaa caacacaaaa aaacttgaga ctaaaattag ctgagctta gactatgaca	33420
ccaatggtgc tgcattatt aaacttggca ctggtctaag cttcgacaac acaggcggcc	33480

taactgtggg aaacactggg gatgataaac tgactctgtg gacgaccca gacccatctc	33540
caaattgcag aattcactca gacaaagact gcaagtttac tctagtccta actaagtgtg	33600
gaagccaaat cctggcctct gtgcgcgcc tagcggatc aggaaatctg gcttcgataa	33660
caggcaccgt tgccagcgtt accatctttc tcagatttga tcagaatgga gtgcttatgg	33720
aaaactctc gctagacagg cagtactgga acttcagaaa tggcaactca actaacgtg	33780
ccccctacac caatgcagtt gggttcatgc caaacctgc agcatacccc aaaacgcaa	33840
gccagactgc taaaaacaac attgtaagtc aggtttactt gaatggagac aaatccaaac	33900
ccatgacct taccatcacc ctcaatggaa ctaatgaatc cagtgaact agccaggtga	33960
gtcactactc caigtcatth acatgggctt gggaaagtgg gcaatatgcc actgaaacct	34020
ttgccaccaa ctcttcacc ttttcttaca ttgctgaaca ataaaaagca tgacactgat	34080
gttcatttct gattcttatt ttattatth caaacacaac aaaatcttc aagtcattct	34140
tccatcttag cttaatagac acagtagctt aatagacca gtagtgcaa gcccattct	34200
agcttataac tagtggagaa gtactgcct acatgggggt agagtcataa tcgtgcatca	34260
ggatagggcg gtggtgtgc agcagcgcg gaataaactg ctgcgcgcg cgtccgtcc	34320
tgcaggaata caacatggca gtggtctcct cagcgtgat tcgcaccgcc cgcagcataa	34380
ggcgccttgt cctccgggca cagcagcga cctgatctc acttaaatca gcacagtaac	34440
tgcagcacag caccacaata ttgttcaaaa tcccacagtg caaggcgtg tatccaaagc	34500
tcatggcggg gaccacagaa cccactggc catcatacca caagcgcagg tagattaagt	34560
ggcgacctt cataaacag ctggacataa acattacctc ttttggcatg ttgtaattca	34620
ccacctccg gtacatatata aacctctgat taaacatggc gccatccacc accatcctaa	34680
accagctggc caaaacctgc ccgccggcta tacactgcag ggaaccggga ctggaacaat	34740
gacagtggag agcccaggac tcgtaacct ggatcatcat gctcgtcatg atatcaatgt	34800
tggcacaaca caggcacag tgcatacact tcctcaggat tacaagctcc tcccgcgtta	34860
gaaccatatc ccagggaaca accatttct gaatcagct aaatcccaca ctgcaggga	34920
gacctgcac gtaactcac ttgtgattg tcaaagtgtt acattcgggc agcagcggat	34980
gacctccag tatggtagcg cgggtttctg tctcaaaagg aggtagacga tccctactgt	35040
acggagtgcg ccgagacaac cgagatctg ttggtcgtg tgtcatgcca aatggaacgc	35100
cggacgtagt catatttct gaagcttag atctctcaac gcagcaccag caccaacact	35160
tcgcagtgt aaaggccaag tgccgagaga gtatatatag gaataaaaag tgacgtaaac	35220

gggcaaagtc caaaaaacgc ccagaaaaac cgcacgcgaa cctacgcccc gaaacgaaag	35280
ccaaaaaaca ctagacactc ctttccggcg tcaacttccg ctttcccacg ctacgtcact	35340
tgccccagtc aaacaaacta catatcccga acttccaagt cgccacgccc aaaacaccgc	35400
ctacacctcc ccgcccccg gcccccccc aaaccgcct cccgccccgc gcccccccc	35460
gcgcgcccc tctcattatc atattggctt caatccaaaa taaggatat tattgatgat	35520
ggtttaaacg gatccaattc ttgaagacga aagggcctcg tgatagcct atttttatag	35580
gttaatgtca tgataataat ggtttcttag acgtcagggtg gcacttttcg gggaaatgtg	35640
cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga	35700
caataaccct gataaatgct tcaataatat tgaanaagga agagtatgag tattcaacat	35760
ttccgtgtcg cccttatcc ctttttgcg gcattttgcc ttcctgtttt tgctcaccca	35820
gaaacgctgg tgaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc	35880
gaactggatc tcaacagcgg taagatcctt gagagttttc gcccgaaga acgttttcca	35940
atgatgagca cttttaagt tctgctatgt ggcgcggtat tatcccggtg tgacgccggg	36000
caagagcaac tcggtcgccg catacactat tctcagaatg acttggttga gtactcacca	36060
gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata	36120
accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag	36180
ctaaccgctt ttttgacaa catgggggat catgtaactc gccttgatcg ttgggaaccg	36240
gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca	36300
acaacgttgc gcaaacattt aactggcgaa ctacttactc tagcttcccg gcaacaatta	36360
atagactgga tggaggcgga taaagttgca ggaccattc tgcgctcggc ctttccggct	36420
ggctggttta ttgctgataa atctggagcc ggtgagcgtg ggtctcgcg tatcattgca	36480
gcactggggc cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag	36540
gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat	36600
tggttaactgt cagaccaagt ttactcatat atactttaga ttgatttaaa aggatctagg	36660
tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt tcgttccact	36720
gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg	36780
taatctgctg cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc	36840
aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgag ataccaaata	36900
ctgtccttct agttagccg tagttaggcc accacttcaa gaactctgta gcaccgcta	36960
catactcgc tctgctaate ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc	37020
ttaccgggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg	37080

ggggttcgtg cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac 37140
 agcgtgagct atgagaaaagc gccacgcttc ccgaaggag aaaggcggac aggtatccgg 37200
 taagcggcag ggtcggaaaca ggagagcgca cgaggagct tccaggggga aacgcctggt 37260
 atctttatag tcctgtcggg ttctgccacc tctgacttga gcgtcgattt ttgtgatgct 37320

cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg 37380
 ccttttgctg gccttgaagc tgteccctgat ggtcgtcatc tacctgcctg gacagcatgg 37440
 cctgcaacgc gggcatcccg atgccgccgg aagcgagaag aatcataatg gggaaggcca 37500
 tccagcctcg cgtcgcagat ccgaattcgt ttaaac 37536

<210> 25

<211> 29

<212> DNA

<213> Human cytomegalovirus

<400> 25

catctacgta ttagtcacg ctattacca 29

<210> 26

<211> 21

<212> DNA

<213> Human cytomegalovirus

<400> 26

gacttggaaa tccccgtgag t 21

<210> 27

<211> 543

<212> PRT

<213> Pan troglodytes

<400> 27

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu

50	55	60	
Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser			
65	70	75	80
Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn			
	85	90	95
Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala			
	100	105	110
Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu			
	115	120	125
Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr			
	130	135	140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu			
145	150	155	160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val			
	165	170	175
Ser Ala Thr Pro Pro Ile Asn Val Ser Ser Gly Ser Leu Gly Leu Asp			
	180	185	190
Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile			
	195	200	205
Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val			
	210	215	220
Thr Gly Asn Gly Leu Thr Val Asp Asn Asn Ala Leu Gln Thr Arg Val			
225	230	235	240
Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala			
	245	250	255
Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asn Val			
	260	265	270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln			
	275	280	285
Gly Pro Leu Tyr Ile Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn			
	290	295	300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
 305 310 315 320
 Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
 325 330 335
 Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
 340 345 350
 Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
 355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
 370 375 380
 Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
 385 390 395 400
 Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val
 405 410 415
 Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
 420 425 430

Ser Leu Asp Arg Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
 435 440 445
 Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
 450 455 460
 Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
 465 470 475 480
 Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
 485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
 500 505 510
 Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
 515 520 525
 Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
 530 535 540

<210> 28

<211> 542

<212> PRT

<213> Pan paniscus

<400> 28

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu

50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser

65 70 75 80

Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn

85 90 95

Leu Ser Leu Gln Thr Ser Ala Pro Leu Thr Val Ser Ser Gly Ser Leu

100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu Thr

115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Gly Leu

130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Thr Leu Gln

145 150 155 160

Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val Gly

165 170 175

Thr Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp Met

180 185 190

Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile Gly

195 200 205

Gly Pro Leu Gln Val Val Asp Ser Leu His Thr Leu Thr Val Val Thr

210 215 220

Gly Asn Gly Ile Thr Val Ala Asn Asn Ala Leu Gln Thr Lys Val Ala
 225 230 235 240
 Gly Ala Leu Gly Tyr Asp Ser Ser Gly Asn Leu Glu Leu Arg Ala Ala
 245 250 255
 Gly Gly Met Arg Ile Asn Thr Gly Gly Gln Leu Ile Leu Asp Val Ala
 260 265 270
 Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln Gly
 275 280 285
 Pro Leu Tyr Val Asn Thr Asn His Asn Leu Asp Leu Asn Cys Asn Arg
 290 295 300
 Gly Leu Thr Thr Thr Thr Ser Ser Asn Thr Thr Lys Leu Glu Thr Lys
 305 310 315 320
 Ile Asp Ser Gly Leu Asp Tyr Asn Ala Asn Gly Ala Ile Ile Ala Lys
 325 330 335
 Leu Gly Thr Gly Leu Thr Phe Asp Asn Thr Gly Ala Ile Thr Val Gly
 340 345 350
 Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro Ser
 355 360 365
 Pro Asn Cys Arg Ile His Ala Asp Lys Asp Cys Lys Phe Thr Leu Val
 370 375 380
 Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu Ala
 385 390 395 400
 Val Ser Gly Asn Leu Ser Ser Met Thr Gly Thr Val Ser Ser Val Thr
 405 410 415
 Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser Ser
 420 425 430
 Leu Asp Lys Glu Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn Ala
 435 440 445
 Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ser Ala Tyr
 450 455 460
 Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Glu Val

465 470 475 480
 Tyr Leu His Gly Asp Lys Ser Lys Pro Met Ile Leu Thr Ile Thr Leu
 485 490 495
 Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr Ser
 500 505 510
 Met Ser Phe Thr Trp Ser Trp Asp Ser Gly Lys Tyr Ala Thr Glu Thr

 515 520 525
 Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
 530 535 540
 <210> 29
 <211> 543
 <212> PRT
 <213> Pan troglodytes
 <400> 29
 Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
 20 25 30

 Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45
 Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asp Ile Thr Ser Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

 Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110
 Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125
 Thr Met Gln Ser Glu Ala Pro Leu Ala Val Gln Asp Ala Lys Leu Thr
 130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175

Ser Ser Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp
180 185 190

Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
195 200 205

Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
210 215 220

Thr Gly Asn Gly Leu Thr Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225 230 235 240

Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
245 250 255

Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
290 295 300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305 310 315 320

Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335

Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
340 345 350

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu

385 390 395 400
 Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val
 405 410 415

 Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
 420 425 430
 Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
 435 440 445
 Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
 450 455 460
 Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
 465 470 475 480

 Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
 485 490 495
 Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
 500 505 510
 Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
 515 520 525
 Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
 530 535 540

 <210> 30
 <211> 543
 <212> PRT
 <213> Pan troglodytes
 <400> 30

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

50 55 60
 Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95
 Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110
 Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu

 115 120 125
 Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
 130 135 140
 Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
 145 150 155 160
 Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
 165 170 175
 Ser Ala Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp

 180 185 190
 Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
 195 200 205
 Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
 210 215 220
 Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
 225 230 235 240
 Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala

 245 250 255
 Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
 260 265 270
 Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
 275 280 285
 Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
 290 295 300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr

305 310 315 320

Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile

325 330 335

Lys Leu Gly Thr Gly Val Ser Phe Asp Ser Thr Gly Ala Leu Ser Val

340 345 350

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro

355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu

370 375 380

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu

385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val

405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser

420 425 430

Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn

435 440 445

Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala

450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln

465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr

485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr

500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu

515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln

530 535 540

<210> 31

<211> 543

<212> PRT

<213> Pan troglodytes

<400> 31

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu

50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser

65 70 75 80

Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn

85 90 95

Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala

100 105 110

Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu

115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr

130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu

145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val

165 170 175

Ser Ala Thr Pro Pro Ile Asn Val Ser Ser Gly Ser Leu Gly Leu Asp

180 185 190

Met Glu Asn Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile

195 200 205

Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val

210 215 220
 Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
 225 230 235 240
 Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
 245 250 255
 Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
 260 265 270

 Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
 275 280 285
 Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
 290 295 300
 Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
 305 310 315 320
 Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
 325 330 335

 Lys Leu Gly Thr Gly Val Ser Phe Asp Ser Thr Gly Ala Leu Ser Val
 340 345 350
 Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
 355 360 365
 Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
 370 375 380
 Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
 385 390 395 400

 Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
 405 410 415
 Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
 420 425 430
 Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
 435 440 445
 Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
 450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
 465 470 475 480
 Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile Leu Thr Ile Thr
 485 490 495
 Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
 500 505 510
 Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
 515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
 530 535 540

<210> 32

<211> 578

<212> PRT

<213> Pan troglodytes

<400> 32

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

35 40 45
 Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95
 Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala

100 105 110
 Leu Thr Leu Ala Ala Ala Val Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125
 Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr

130 135 140
 Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
 145 150 155 160
 Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Ile

 165 170 175
 Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
 180 185 190
 Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
 195 200 205
 Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
 210 215 220
 Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val

 225 230 235 240
 Ser Gly Ala Leu Asn Tyr Asp Ser Ser Gly Asn Leu Glu Leu Arg Ala
 245 250 255
 Ala Gly Gly Met Arg Val Asp Ala Asn Gly Lys Leu Ile Leu Asp Val
 260 265 270
 Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
 275 280 285
 Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn

 290 295 300
 Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
 305 310 315 320
 Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
 325 330 335
 Ile Asn Pro Gly Asp Gly Leu Glu Phe Gly Ser Gly Ser Asp Thr Asn
 340 345 350
 Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Ser Arg

 355 360 365
 Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
 370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
 385 390 395 400
 Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
 405 410 415
 Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
 420 425 430
 Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
 435 440 445
 Val Thr Ser Ala Gln Ile Ile Leu Arg Phe Asp Glu Asn Gly Val Leu
 450 455 460
 Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
 465 470 475 480
 Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
 485 490 495
 Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
 500 505 510
 Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile
 515 520 525
 Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
 530 535 540
 Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
 545 550 555 560
 Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
 565 570 575
 Gln Glu

<210> 33

<211> 578

<212> PRT

<213> Pan troglodytes

<400> 33

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro

1 5 10 15
 Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
 20 25 30

 Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45
 Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

 Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110
 Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125
 Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
 130 135 140
 Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
 145 150 155 160

 Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
 165 170 175
 Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
 180 185 190
 Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
 195 200 205
 Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
 210 215 220

 Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
 225 230 235 240
 Ser Gly Ala Leu Asn Tyr Asp Thr Ser Gly Asn Leu Glu Leu Arg Ala
 245 250 255

Ala Gly Gly Met Arg Val Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
290 295 300

Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305 310 315 320

Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
325 330 335

Ile Asn Ala Gly Asp Gly Leu Gln Phe Asp Ser Gly Ser Asp Thr Asn
340 345 350

Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
355 360 365

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
435 440 445

Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn

500 505 510
 Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
 515 520 525
 Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
 530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
 545 550 555 560
 Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
 565 570 575
 Gln Glu

<210> 34

<211> 578

<212> PRT

<213> Pan troglodytes

<400> 34

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20 25 30
 Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45
 Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn

85 90 95
 Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110
 Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
130 135 140
Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145 150 155 160
Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175
Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
180 185 190
Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
195 200 205
Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
210 215 220
Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
225 230 235 240
Ser Gly Ala Leu Asn Tyr Asp Thr Ser Gly Asn Leu Glu Leu Arg Ala
245 250 255
Ala Gly Gly Met Arg Val Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270
Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285
Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
290 295 300
Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305 310 315 320
Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
325 330 335
Ile Asn Ala Gly Asp Gly Leu Gln Phe Asp Ser Gly Ser Asp Thr Asn
340 345 350
Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
355 360 365
Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly

370 375 380
 Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
 385 390 395 400
 Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala

 405 410 415
 Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
 420 425 430
 Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
 435 440 445
 Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
 450 455 460
 Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly

 465 470 475 480
 Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
 485 490 495
 Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
 500 505 510
 Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
 515 520 525
 Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val

 530 535 540
 Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
 545 550 555 560
 Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
 565 570 575
 Gln Glu

<210> 35

<211> 577

<212> PRT

<213> Pan paniscus

<400> 35

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Gln Thr Ser Ala Pro Leu Thr Val Ser Ser Gly Ser Leu
100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Gly Leu
130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Thr Leu Gln
145 150 155 160

Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val Ser
165 170 175

Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Ser Ile Asp Met
180 185 190

Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Ala Leu Asn Ile Gly
195 200 205

Ala Pro Leu His Val Val Asp Thr Leu Asn Ala Leu Thr Val Val Thr
210 215 220

Gly Gln Gly Leu Thr Ile Asn Gly Arg Ala Leu Gln Thr Arg Val Thr
225 230 235 240

Gly Ala Leu Ser Tyr Asp Thr Glu Gly Asn Ile Gln Leu Gln Ala Gly

245 250 255
 Gly Gly Met Arg Ile Asp Asn Asn Gly Gln Leu Ile Leu Asn Val Ala
 260 265 270

 Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln Gly
 275 280 285
 Pro Leu Ile Val Asn Ser Ala His Asn Leu Asp Leu Asn Leu Asn Arg
 290 295 300
 Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val Asn
 305 310 315 320
 Ile Lys Thr Ala Lys Gly Leu Phe Tyr Asp Gly Thr Ala Ile Ala Ile
 325 330 335

 Asn Ala Gly Asp Gly Leu Gln Phe Gly Ser Gly Ser Asp Thr Asn Pro
 340 345 350
 Leu Gln Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Asn Lys Ala
 355 360 365
 Ile Ile Thr Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala
 370 375 380
 Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr Thr
 385 390 395 400

 Pro Asp Pro Ser Pro Asn Cys Arg Ile Asn Ser Glu Lys Asp Ala Lys
 405 410 415
 Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser Val
 420 425 430
 Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr Val
 435 440 445
 Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu Leu
 450 455 460

 Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly Asp
 465 470 475 480
 Ser Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro Asn
 485 490 495

Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn Ile
 500 505 510
 Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys Pro Met Thr Leu
 515 520 525

Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val Ser
 530 535 540
 Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr Ile
 545 550 555 560
 Asn Asp Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Gln
 565 570 575
 Glu

<210> 36

<211> 577

<212> PRT

<213> Pan paniscus

<400> 36

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro

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

65 70 75 80
 Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95
 Leu Ser Leu Gln Thr Ser Ala Pro Leu Thr Val Ser Ser Gly Ser Leu
 100 105 110
 Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu Thr

115 120 125
 Met Gln Ser Gln Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Gly Leu

 130 135 140
 Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Thr Leu Gln
 145 150 155 160
 Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val Ser
 165 170 175
 Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Ser Ile Asp Met
 180 185 190
 Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Ala Leu Asn Ile Gly

 195 200 205
 Ala Pro Leu His Val Val Asp Thr Leu Asn Ala Leu Thr Val Val Thr
 210 215 220
 Gly Gln Gly Leu Thr Ile Asn Gly Arg Ala Leu Gln Thr Arg Val Thr
 225 230 235 240
 Gly Ala Leu Ser Tyr Asp Thr Glu Gly Asn Ile Gln Leu Gln Ala Gly
 245 250 255
 Gly Gly Met Arg Ile Asp Asn Asn Gly Gln Leu Ile Leu Asn Val Ala

 260 265 270
 Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln Gly
 275 280 285
 Pro Leu Ile Val Asn Ser Ala His Asn Leu Asp Leu Asn Leu Asn Arg
 290 295 300
 Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val Asn
 305 310 315 320
 Ile Lys Thr Ala Lys Gly Leu Phe Tyr Asp Gly Thr Ala Ile Ala Ile

 325 330 335
 Asn Ala Gly Asp Gly Leu Gln Phe Gly Ser Gly Ser Asp Thr Asn Pro
 340 345 350
 Leu Gln Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Asn Lys Ala
 355 360 365

Ile Ile Thr Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala

370 375 380

Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr Thr

385 390 395 400

Pro Asp Pro Ser Pro Asn Cys Arg Ile Asn Ser Glu Lys Asp Ala Lys

405 410 415

Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser Val

420 425 430

Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr Val

435 440 445

Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu Leu

450 455 460

Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly Asp

465 470 475 480

Ser Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro Asn

485 490 495

Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn Ile

500 505 510

Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys Pro Met Thr Leu

515 520 525

Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val Ser

530 535 540

Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr Ile

545 550 555 560

Asn Asp Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Gln

565 570 575

Glu

<210> 37

<211> 845

<212> DNA

<213> Human cytomegalovirus

<400> 37

ccattgcata cgttgatcc atatcataat atgtacattt atattggctc atgtccaaca 60
 ttaccgccat gttgacattg attattgact agttattaat agtaatcaat tacgggggtca 120
 ttagttcata gcccatafat ggagttccgc gttacataac ttacggtaaa tggcccgctt 180
 ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta 240
 acgccaatag ggactttcca ttgacgtcaa tgggtggagt attiacggta aactgcccac 300
 ttggcagtac atcaagtgt tcatatgcc agtacgcccc ctattgacgt caatgacggt 360
 aaatggcccc cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag 420

tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca gtacatcaat 480
 gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat 540
 gggagtttgt tttggcaca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc 600
 ccattgacgc aaatgggcgg taggcgtgtg cgggtggagg tctatataag cagagctctc 660
 cctatcagtg atagagatct ccctatcagt gatagagatc gtcgacgagc tcgttttagtg 720
 aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag aagacaccgg 780
 gaccgatcca gcctccgcgg ccgggaacgg tgcattggaa cgcggattcc ccgtgccaag 840

agtga 845

<210> 38

<211> 15422

<212> DNA

<213> Pan troglodytes

<400> 38

tgcccgtttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 60
 cggggagagg cggtttgcgt attgggcgct agatctccca tcacatatac ctgccgttca 120
 ctattattta gtgaaatgag atattatgat attttctgaa ttgtgattaa aaaggcaact 180
 ttatgcccac gcaacagaaa ctataaaaaa tacagagaat gaaaagaaac agatagattt 240
 tttagttctt taggcccgtg gtctgcaaat ccttttatga ttttctatca aacaaaagag 300

gaaaatagac cagttgcaat ccaaacgaga gtctaataga atgaggctga aaagtaaact 360
 gcgcgggttt gttactgata aagcaggcaa gacctaaat gtgtaaaggg caaagtgtat 420
 actttggcgt cacccttac atattttagg tcttttttta ttgtgcgtaa ctaacttgcc 480
 atcttcaaac aggagggtcg gaagaagcag accgctaaca cagtacataa aaaaggagac 540
 atgaacgatg aacatcaaaa agtttgcaaa acaagcaaca gtattaacct ttactaccgc 600

actgctggca ggaggcgcaa ctcaagcgtt tgcgaaagaa acgaacaaaa agccatataa	660
ggaaacatac ggcatttccc atattacacg ccatgatatg ctgcaaatcc ctgaacagca	720
aaaaaatgaa aaatatcaag ttcttgaatt cgattcgtcc acaattaaaa atatctcttc	780
tgcaaaaggc ctggacgttt gggacagctg gccattacaa aacgctgacg gcactgtcgc	840
aaactatcac ggctaccaca tcgtctttgc attagccgga gatcctaaaa atgcggatga	900
cacatcgatt tacatgttct atcaaaaagt cggcgaaact tctattgaca gctggaaaaa	960
cgctggccgc gtctttaaag acagcgacaa attcgatgca aatgattcta tcctaaaaga	1020
ccaaacacaa gaatggtcag gttcagccac atttacatct gacggaaaaa tccgtttatt	1080
ctacactgat ttctccggtg aacattacgg caaacaacaa ctgacaactg cacaagttaa	1140
cgatcagca tcagacagct ctttgaacat caacggtgta gaggattata aatcaatctt	1200
tgacggtgac ggaaaaacgt atcaaaatgt acagcagttc atcgatgaag gcaactacag	1260
ctcaggcgac aaccatacgc tgagagatcc tcactacgta gaagataaag gccacaaata	1320
cttagtattt gaagcaaaca ctggaactga agatggctac caaggcgaag aatctttatt	1380
taacaaagca tactatggca aaagcacatc attcttccgt caagaaagtc aaaaacttct	1440
gcaaagcgat aaaaaacgca cggctgagtt agcaaacggc gctctcggtg tgattgagct	1500
aaacgatgat tacacactga aaaaagtgat gaaaccgctg attgcatcta acacagtaac	1560
agatgaaatt gaacgcgcga acgtctttta aatgaacggc aaatggtacc tgttactga	1620
ctcccgcgga tcaaaaatga cgattgacgg cattacgtct aacgatattt acatgcttgg	1680
ttatgtttct aattctttta ctggcccata caagccgctg aacaaaactg gccttgtgtt	1740
aaaaatggat ctigatccta acgatgtaac ctttacttac tcacacttcg ctgtacctca	1800
agcgaagga aacaatgctg tgattacaag ctatatgaca aacagaggat tctacgcaga	1860
caaacaatca acgtttgcgc caagcttccct gctgaacatc aaaggcaaga aaacatctgt	1920
tgtcaaagac agcatccttg aacaaggaca attaacagtt aacaataat agggataaca	1980
gggtaatgct agaagaccgc agtcttacca gtaaaagaaa aaagatctct caacgcagca	2040
ccagcaccaa cacttcgcag tgtaaaaggc caagtgccga gagagtatat ataggaataa	2100
aaagtgacgt aaacgggcaa agtccaaaaa acgccagaa aaaccgcacg cgaacctacg	2160
ccccgaacg aaagccaaaa aacactagac actcccttcc ggctcaact tccgttttcc	2220
cacgctacgt cacttgcccc agtcaaaaaa actacatata ccgaacttcc aagtcgccac	2280
gccccaaaca ccgcctacac ctccccgcc gccggccgc ccccaaacc gcctccgcc	2340

ccgcgccccg ccccgcgccg cccatctcat tatcatattg gcttcaatcc aaaataaggt 2400

atattattga tgatggttta aacggatcct ctagagtcga cctgcaggca tgcaagcttg 2460

agtattctat agtgtcacct aaatagcttg gcgtaatcat ggtcatagct gtttctgtg 2520

tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa 2580

gcctgggggtg cctaagtgtg gagctaactc acattaattg cgttgcgtc actgcccgt 2640

ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgaaccctt 2700

gcggccgccg gggccgtcga ccaattctca tgtttgacag cttatcatcg aatttctgcc 2760

attcatccgc ttattatcac ttattcaggc gtagcaacca ggcgtttaag ggcaccaata 2820

actgccttaa aaaaattacg ccccgccctg ccactcatcg cagtaactgtt gtaattcatt 2880

aagcattctg ccgacatgga agccatcaca aacggcatga tgaacctgaa tcgccagcgg 2940

catcagcacc ttgtgcctt gcgtataata tttgccatg gtgaaaacgg gggcgaagaa 3000

gttgtccata ttggccacgt ttaaatcaaa actggtgaaa ctaccccagg gattggctga 3060

gacgaaaaac atatttcaa taaaccctt agggaaatag gccaggtttt caccgtaaca 3120

cgccacatct tgcgaatata tgtgtagaaa ctgccgaaa tcgtcgtggt attactcca 3180

gagcgatgaa aacgtttcag ttgtctcatg gaaaacggtg taacaagggt gaacactatc 3240

ccatacacc agctcacctg ctttcattgc catacggaat tccgatgag cattcatcag 3300

gcgggcaaga atgtgaataa aggccggata aaacttgtgc ttatttttct ttacggtctt 3360

taaaaaggcc gtaatatcca gctgaacggt ctggttatag gtacattgag caactgactg 3420

aatgcctca aatgttctt tacgatgcca ttgggatata tcaacggtgg tatatccagt 3480

gatttttttc tccattttag ctctcttagc tctgaaaat ctcgataact caaaaaatac 3540

gcccgttagt gatcttattt cattatggtg aaagttggaa cctcttacgt gccgatcaac 3600

gtctcatttt cgccaaaagt tggcccaggg ctccccgta tcaacaggga caccaggatt 3660

tatttattct gcgaagtgat ctccgtcac aggtatttat tcgcgataag ctcatggagc 3720

ggcgtaaccg tcgcacagga aggacagaga aagcgcggat ctgggaagtg acggacagaa 3780

cggtcaggac ctggattggg gaggcggttg ccgccgtgc tgctgacggt gtgacgttct 3840

ctgttccgtt cacaccacat acgttccgcc attcctatgc gatgcacatg ctgtatgccg 3900

gtataccgtt gaaagtctg caaagcctga tgggacataa gtccatcagt tcaacggaag 3960

tctacacgaa ggtttttgct ctggatgtgg ctccccgca ccgggtgcag ttgctgatgc 4020

cggagtctga tgcggttgct atgctgaaac aattatcctg agaataaatg ccttggcctt 4080

tatatggaaa tgtggaactg agtggatatg ctgtttttgt ctgttaaaca gagaagctgg	4140
ctgttatcca ctgagaagcg aacgaacag tcgggaaaat ctcccattat cgtagagatc	4200
cgcattatta atctcaggag cctgtgtagc gtttatagga agtagtggtc tgtcatgatg	4260
cctgcaagcg gtaacgaaaa cgatttgaat atgccttcag gaacaataga aatcttcgtg	4320
cgggtgttac ttgaagtga gcggttatg tcagcaatgg acagaacaac ctaatgaaca	4380
cagaaccatg atgtggtctg tccttttaca gccagtagtg ctgcgccgag tcgagcgaca	4440
gggcgaagcc ctcgagttag cgaggaagca ccagggaaca gcacttatat attctgctta	4500
cacacgatgc ctgaaaaaac ttcccttggg gttatccact tatccacggg gatattttta	4560
taattatfff ttttatagtt ttttagatctt cttttttaga gcgccttgta ggcctttatc	4620
catgctgggt ctagagaagg tgttgtgaca aattgccctt tcagtgtagc aaatcacctt	4680
caaatgacag tctgtctgt gacaaattgc ccttaaccct gtgacaaatt gccctcagaa	4740
gaagctgttt tttcacaaag ttatccctgc ttattgactc tttttatfff agtgtgacaa	4800
tctaaaaact tgtcacactt cacatggatc tgcctggcg gaaacagcgg ttatcaatca	4860
caagaaacgt aaaaatagcc cgcgaaatct ccagtcaaac gacctactg aggcggcata	4920
tagtctctcc cgggatcaaa aacgtatgct gtatctgttc gttgaccaga tcagaaaatc	4980
tgatggcacc ctacaggaac atgacggtat ctgcgagatc catgttgcta aatatgctga	5040
aatattcgga ttgacctctg cggaagccag taaggatata cggcaggcat tgaagagttt	5100
cgcggggaag gaagtgtttt tttatcgccc tgaagaggat gccggcgtg aaaaaggcta	5160
tgaatctfff ccttggttta tcaaactgac gcacagtcca tccagagggc ttacagtggt	5220
acatatcaac ccatatctca ttcccttctt tatcggttta cagaaccggt ttacgcagtt	5280
tcggcttagt gaaacaaaag aaatcaccaa tccgtatgcc atgcgtttat acgaatccct	5340
gtgtcagtat cgtaagccgg atggtcagg catcgtctct ctgaaaatcg actggatcat	5400
agagcgttac cagctgcctc aaagtacca gcgtatgcct gacttccgc gccgcttctt	5460
gcaggtctgt gttaatgaga tcaacagcag aactccaatg cgcctctcat acattgagaa	5520
aaagaaaggc cgccagacga ctcatatcgt attttcttc cgcgatatca ctccatgac	5580
gacaggatag tctgagggtt atctgtcaca gatttgaggg tggttcgtca catttgttct	5640
gacctactga gggtaatttg tcacagtttt gctgtttcct tcagcctgca tggattttct	5700
catacttttt gaactgtaat ttttaaggaa gccaaatttg agggcagttt gtcacagttg	5760
atttcttct ctttcccttc gtcagtgtac ctgatatcgg gggttagttc gtcacattg	5820
atgagggttg attatcacag tttattactc tgaattggct atccgcgtgt gtacctctac	5880
ctggagtttt tcccacggtg gatatttctt cttgcgtga gcgtaagagc tatctgacag	5940

aacagtcttt ctttgcttcc tcgccagttc gctcgctatg ctcggttaca cggctgcggc	6000
gagcgctagt gataataagt gactgaggta tgtgctcttc ttatctcctt ttgtagtgtt	6060
gctcttattt taaacaactt tgcggttttt tgatgacttt gcgattttgt tgttgctttg	6120
cagtaaattg caagatttaa taaaaaacg caaagcaatg attaaaggat gttcagaatg	6180
aaactcatgg aaacacttaa ccagtgcata aacgctggtc atgaaatgac gaaggctatc	6240
gccattgcac agtttaatga tgacagcccg gaagcgagga aaataaccg gcgctggaga	6300
ataggtgaag cagcggattt agttgggggt tcttctcagg ctatcagaga tgccgagaaa	6360
gcaggcgac taccgcacc ggatatgaa attcgaggac gggttgagca acgtgttggt	6420
tatacaattg aacaaattaa tcatatgcgt gatgtgttg gtacgcgatt gcgacgtgct	6480
gaagacgtat ttccaccgt gatcgggggt gctgccata aaggtagcgt ttacaaaacc	6540
tcagtttctg ttcatcttgc tcaggatctg gctctgaagg ggctacgtgt tttgctcgtg	6600
gaaggtaacg acccccaggg aacagcctca atgtatcacg gatgggtacc agatcttcat	6660
attcatgcag aagacactct cctgcctttc tatcttgggg aaaaggacga tgtcacttat	6720
gcaataaagc ccacttgcgt gccggggctt gacattattc cttctgtct ggctctgcac	6780
cgtattgaaa ctgagttaat gggcaaattt gatgaaggta aactgccac cgatccacac	6840
ctgatgctcc gactggccat tgaaactgtt gctcatgact atgatgtcat agttattgac	6900
agcgcgccta acctgggtat cggcacgatt aatgtcgtat gtgctgctga tgtgctgatt	6960
gttcccacgc ctgctgagtt gtttgactac acctccgcac tgcagttttt cgatatgctt	7020
cgtgatctgc tcaagaacgt tgatcttaaa gggttcgagc ctgatgtacg tattttgctt	7080
accaaataca gcaatagtaa tggctctcag tccccgtga tggaggagca aattcgggat	7140
gcctggggaa gcatggttct aaaaaatgtt gtacgtgaaa cggatgaagt tggtaaagg	7200
cagatccgga tgagaactgt ttttgaacag gccattgac aacgtcttc aactggtgcc	7260
tggagaaatg ctctttctat ttgggaacct gtctgcaatg aaattttcga tcgtctgatt	7320
aaaccacgct gggagattag ataataaagc gtgcgcctgt tattccaaaa catacgtca	7380
atactcaacc ggttgaagat acttcgttat cgacaccage tgccccgatg gtggattcgt	7440
taattgcgcg cgtaggagta atggctcgcg gtaatgcat tactttgcct gtatgtggc	7500
gggatgtgaa gtttactctt gaagtgtctc ggggtgatag tgttgagaag acctctcggg	7560
tatggtcagg taatgaacgt gaccaggagc tgcttactga ggacgactg gatgatctca	7620
tcccttcttt tctactgact ggtcaacaga caccggcgtt cggtcgaaga gtatctggtg	7680
tcatagaaat tgccgatggg agtcgccgtc gtaaagctgc tgcacttacc gaaagtgatt	7740
atcgtgttct ggttggcgag ctggatgatg agcagatggc tgcattatcc agattgggta	7800

acgattatcg cccaacaagt gcttatgaac gtggtcagcg ttatgcaagc cgattgcaga 7860

atgaatttgc tggaatatatt tctgcgctgg ctgatgcgga aaatatttca cgtaagatta 7920

ttaccgctg tatcaacacc gccaaattgc ctaaatacagt tgttgctctt ttttctcacc 7980

ccggtgaact atctgcccg tccagtgatg cacttcaaaa agcctttaca gataaagagg 8040

aattacttaa gcagcaggca tctaaccttc atgagcagaa aaaagctggg gtgatatttg 8100

aagctgaaga agttatcact cttttaactt ctgtgcttaa aacgtcatct gcatcaagaa 8160

ctagtttaag ctcacgacat cagtttgctc ctggagcgac agtattgtat aaggcgata 8220

aaatggctgt taacctggac aggtctcgtg ttccaactga gtgtatagag aaaattgagg 8280

ccattcttaa ggaacttgaa aagccagcac cctgatgcga ccacgtttta gtctacgttt 8340

atctgtcttt acttaatgtc ctttgttaca ggccagaaag cataactggc ctgaatatc 8400

tctctgggcc cactgttcca ctgtatcgt cggcttgata atcagactgg gaccacggtc 8460

ccactcgtat cgtcggctcg attattagtc tgggaccacg gtccactcg tatcgtcgg 8520

ctgattatta gtctgggacc acgggccac tcgtatcgtc ggtctgataa tcagactggg 8580

accacggtec cactcgtatc gtcggtctga ttattagctt gggaccatgg tccactcgt 8640

atcgtcggtc tgattattag tctgggacca cgggccact cgtatcgtcg gtctgattat 8700

tagtctggaa ccacggtecc actcgtatcg tcggtctgat tattagctg ggaccacgg 8760

cccactcgtat tcgtcggctt gattattagt ctgggaccac gatccactc gtgttgcgg 8820

tctgattatc ggtctgggac caggttcca ctgtattgt cgtacagact atcagcgtga 8880

gactacgatt ccatcaatgc ctgtcaaggg caagtattga catgtcgtcg taacctgtag 8940

aacggagtaa cctcggctgt cggttgtatg cctgctgtgg attgctgctg tgtcctgctt 9000

atccacaaca ttttgcgcac ggttatgtgg acaaaatacc tggttacca ggccgtgccg 9060

gcacgttaac cgggctgcat ccgatgcaag tgtgtcgtg tcgacgagct cgcgagctcg 9120

gacatgaggt tgccccgtat tcagtgtcgc tgatttgtat tgtctgaagt tgtttttacg 9180

ttaagtgtat gcagatcaat taatacgata cctgcgtcat aattgattat ttgacgtggt 9240

ttgatggcct ccacgcagct tgtgatatgt agatgataat cattatcact ttacgggtcc 9300

tttccggtga tccgacaggt tacggggcgg cgacctcgcg ggttttcgt atttatgaaa 9360

attttccggt ttaaggcgtt tccgttcttc ttgcataa cttaatgttt ttatttaaaa 9420

taccctctga aaagaaagga aacgacaggt gctgaaagcg agctttttgg cctctgtcgt 9480

ttcctttctc tgtttttgtc cgtggaatga acaatggaag tccgagctca tcgctaataa 9540

cttcgtatag catacattat acgaagttat attcgaatgcg gccgcaaggg gttcgcgtca	9600
gcgggtgttg gcgggtgtcg gggctggctt aactatgcgg catcagagca gattgtactg	9660
agagtgcacc atatgcgggtg tgaaataaccg cacagatgcg taaggagaaa ataccgcac	9720
aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggg gcgggcctct	9780
tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg	9840
ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattgta atacgactca	9900
ctatagggcg aattcgagct cggtaaccgg ggatcctcgt ttaaacatc atcaataata	9960
tacctatttt tggattgaag ccaatatgat aatgagatgg gcggcgcggg gcggggcgcg	10020
gggcgggagg cgggtttggg ggccggccgg cggcgccggc ggtgtggcgg aagtggactt	10080
tgtaagtgtg gcggatgtga cttgctagt cgggcgcggg taaaagtac gttttccgtg	10140
cgcgacaacg cccccgggaa gtgacatttt tcccgcggtt ttaccggat gttgtagtga	10200
atttgggcgt aaccaagtaa gatttggcca ttttcgcggg aaaactgaaa cggggaagtg	10260
aaatctgatt aattttgcgt tagtcatacc gcgtaaatatt tgtctagggc cgagggactt	10320
tggccgatta cgtggaggac tcgcccaggt gttttttgag gtgaatttcc gcgttccggg	10380
tcaaagtctg cgttttatta ttataggata tccattgca tacgttgtat ccatacata	10440
atatgtacat ttatattggc tcatgtccaa cattaccgcc atgttgacat tgattattga	10500
ctagttatta atagtaatca attacggggt cattagtca tagcccatat atggagtcc	10560
gcgttacata acttacgga aatggccgc ctggctgacc gccaacgac ccccgccat	10620
tgacgtcaat aatgacgtat gttcccatag taacccaat agggactttc cattgacgtc	10680
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	10740
caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt	10800
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta	10860
ccatggtgat gcggttttgg cagtacatca atggcggtgg atagcggttt gactcacggg	10920
gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac	10980
gggactttcc aaaatgtcgt aacaactccg cccattgac gcaaatgggc ggtaggcgtg	11040
tacggtggga ggtctatata agcagagctc tcctatcag tgatagagat ctccctatca	11100
tgatagaga tcgtcgaca gctcgttttag tgaaccgtca gatcgctgg agacgccatc	11160
cacgtgtttt tgacctccat agaagacacc gggaccgatc cagcctccgc ggccgggaac	11220
ggtgcattgg aacgcggatt ccccggtcca agagttagat cttccgttta tctaggtacc	11280
gggccccccc tcgaggtcga cggatcgtat aagcttcacg ctgccgcaag cactcagggc	11340
gcaagggtg ctaaaggaag cggaacacgt agaaagccag tccgcagaaa cgggtgctgac	11400

cccggatgaa tgcagctac tgggctatct ggacaaggga aaacgcaagc gcaaagagaa	11460
agcaggtagc ttgcagtggc cttacatggc gatagctaga ctgggcggtt ttatggacag	11520
caagcgaacc ggaattgcc a gctggggcgc cctctggtaa ggttgggaag ccctgcaaag	11580
taaactggat ggctttcttg ccgccaagga tctgatggcg caggggatca agatctaacc	11640
aggagctatt taatggcaac agttaaccag ctggtacgca aaccacgtgc tcgcaaagtt	11700
gcgaaaagca acgtgcctgc gctggaagca tgcccgc aaa aacgtggcgt atgtactcgt	11760
gtatatacta ccactcctaa aaaaccgaac tccgcgtgc gtaaagtatg ccgtgttcgt	11820
ctgactaacg gtttcgaagt gacttcctac atcgggtgtg aaggtcaca cctgcaggag	11880
cactccgtga tctgatccg tggcggctgt gttaaagacc tcccgggtgt tcgttaccac	11940
accgtacgtg gtgcgcttga ctgctccggc gttaaagacc gtaagcaggc tcgttccaag	12000
tatggcgtga agcgtcctaa ggcttaatgg tagatctgat caagagacag gatgacggtc	12060
gtttcgcatt ctgaacaag atggattgca cgcaggttct ccggccgctt gggtagagag	12120
gctattcggc tatgactggg cacaacagac aatcggtgc tctgatgcc ccgtgttcg	12180
gctgtcagcg cagggggcgc cggttctttt tgtcaagacc gacctgtccg gtgcctgaa	12240
tgaactgcag gacgaggcag cgcgctatc gtggctggcc acgacgggcg ttccttgcgc	12300
agctgtgtc gacgttgtca ctgaagcggg aagggaactgg ctgctattgg gcgaagtgc	12360
ggggcaggat ctctgtcat ctacacctgc tctgcccag aaagtatcca tcatggctga	12420
tgcaatgcgg cgctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa	12480
acatcgcac gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct	12540
ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcac	12600
gccccagcgc gaggatctgc tcgtgacca tggcgatgcc tgcttgccga atatcatggt	12660
ggaaaatggc cgctttcttg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta	12720
tcaggacata gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga	12780
ccgttctctc gtgctttacg gtatcgccgc tcccattcg cagcgcatcg cttctatcg	12840
ccttcttgac gagtcttct gagcgggact ctggggttcg aaatgaccga ccaagcgacg	12900
cccaacctgc catcacgaga ttctgattcc accgccct tctatgaaag gttgggcttc	12960
ggaatcgttt tccgggacgc cggtggatg atctccagc gcggggatct catgctggag	13020
ttcttcgccc acccgggct cgatccctc ggggggaatc agaattcagt cgacagcggc	13080
cgcatctgc tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt	13140
ccttgacctt ggaaggtgcc actccactg tcttttcta ataaaatgag gaaattgcat	13200
cgcatgtgtc gagtaggtgt cattctattc tgggggtgg ggtggggcag gacagcaagg	13260

gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggccgatc	13320
agcgatcgct gaggtgggtg agtgggcgtg gcctggggtg gtcatgaaaa tatataagtt	13380
gggggtctta gggtctcttt atttgtgttg cagagaccgc cggagccatg agcgggagca	13440
gcagcagcag cagtagcagc agcgcccttg atggcagcat cgtgagccct tatttgacga	13500
cgcgatgcc ccaactgggc ggggtgcgtc agaattgat gggctccagc atcgacggcc	13560
gaccgcctct gcccgcaaat tccgccagc tgacctatgc gaccgtcgcg gggacgccgt	13620
tggacgccac cgccgccgc gccgccaccg cagccgcctc ggccgtgcgc agcctggcca	13680
cggactttgc attcctggga ccaactggca caggggctac ttctcgggcc gctgctgccg	13740
ccgttcgcga tgacaagctg accgccctgc tggcgcagtt ggatgcgctt actcggaac	13800
tgggtgacct ttctcagcag gtcattggccc tgcgccagca ggtctctctc ctgcaagctg	13860
gcgggaatgc ttctccaca aatgccgttt aaggcgcgcg ctagggataa cagggttaata	13920
cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac	13980
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gattattcaa catttccgtg	14040
tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgctcac ccagaaacgc	14100
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg	14160
atctcaacag cggtaaagatc cttgagagtt ttgcceccga agaacgtttt ccaatgatga	14220
gcacttttaa agttctgcta tgtggcgcggt tattatcccg tattgacgcc gggcaagagc	14280
aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag	14340
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgcctgc ataaccatga	14400
gtgataaacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg	14460
cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga	14520
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt	14580
tgcgcaaact attaaactggc gaactactta ctctagcttc ccgcaacaa ttaatagact	14640
ggatggaggc ggataaagtt gcaggaccac ttctgcgtc ggcccttccg gctggctggt	14700
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg	14760
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta	14820
tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac	14880
tgtcagacca agtttactca tataacttt agattgattt aaaatacgtata tatatgtatt	14940
agtcacgct attaccatgg ttaatgcgcc gctacagggc gcgtccattc gccattcagg	15000

ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg ccagctggcg 15060
aaaggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc ccagtcacga 15120
cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga attgggcct 15180
ctagatgcat gctcgagcgg ccgccagtgt gatggatata tgcagaattc cagcacactg 15240
gcggccgtta ctagtggatc cgagctcggt accaagcttg gcgtaatcat ggcatagct 15300
gtttcctgtg tgaattgtt atccgtcac aattccacac aacatacag ccggaagcat 15360
aaagtgtaaa gcctgggggtg cctaagtgt gagctaactc acattaattg cgttgcgctc 15420

ac 15422

<210> 39

<211> 225

<212> DNA

<213> Bos taurus

<400> 39

ctgtgccttc tagttgccag ccattctgtt tttgcccctc ccccgctcct tccttgacct 60
tggaagggtc cactccact gtcctttcct aataaaatga ggaaattgca tcgcatgtc 120
tgagtaggtg tcattctatt ctgggggggtg ggggtgggca ggacagcaag ggggaggatt 180
gggaagacaa tagcaggcat gctggggatg cgtgggctc tatgg 225

<210> 40

<211> 3983

<212> DNA

<213> Pan troglodytes

<400> 40

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg 60
cggggcgggg cgcggggcgg gaggcgggtt tggggcgggg ccggcgggcg gggcggtgtg 120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag 180
tgacgttttc cgtgcgcgac aacgccccg ggaagtgaca tttttccgc ggtttttacc 240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact 300
gaaacgggga agtgaaatct gattaatatt gcgttagtca taccgcgtaa tatgttcta 360
gggccgaggg actttggcgg attacgtgga ggactcgccc aggtgttttt tgaggatgaat 420

ttccgcgttc cgggtcaaag tctccgtttt attattatag gatatccat tgcatacgtt 480
glatccatat cataatatgt acaggcgcgc caaagcatga cactgatgtt cttttctgat 540
tcttatttta ttattttcaa acacaacaaa atcattcaag tcattcttcc atcttagctt 600

aatagacaca gtagcttaat agaccagta gtgcaaagcc ccattctagc ttataactag 660
tggagaagta ctgcctaca tgggggtaga gtcataatcg tgcacagga tagggcggtg 720
gtgctgcagc agcgcgcaa taaactgtg ccgcccgccg tccgtcctgc aggaatacaa 780
catggcagtg gtctcctcag cgatgattcg caccgcccgcc agcataaggc gccttgtcct 840

ccgggcacag cagcgacccc tgatctcact taaatcagca cagtaactgc agcacagcac 900
cacaatattg ttcaaaatcc cacagtgcga ggcgctgtat ccaaagctca tggcggggac 960
cacagaaccc acgtggccat cataccacaa gcgcaggtag attaagtggc gaccctcat 1020
aaacacgtg gaacataaca ttacctctt tggcatgttg taattcacca cctcccgta 1080
ccatataaac ctctgattaa acatggcgcc atccaccacc atcctaaacc agctggccaa 1140
aacctgcccg ccggctatac actgcaggga accgggactg gaacaatgac agtggagagc 1200
ccaggactcg taacatgga tcatcatgct cgtcatgata tcaatgttgg cacaacacag 1260

gcacacgtgc atacacttcc tcaggattac aagctcctcc cgcgtagaa ccataccca 1320
gggaacaacc cattctgaa tcagcgtaaa tcccacactg cagggaagac ctgcacgta 1380
actcagttg tgcattgtca aagtgttaca ttgggcagc agcgatgat cctccagtat 1440
ggtagcgcg gtttctgtct caaaaggagg tagacgatcc ctactgtacg gagtgcgccg 1500
agacaaccga gatcgtgtg gtcgtagtgt catgccaaat ggaacgccg acgtatcat 1560
atctctgaa gtcttagatc tctcaacgca gcaccagcac caacacttcg cagtgtaaaa 1620
ggccaagtgc cgagagagta tatataggaa taaaagtga cgtaaacggg caaagtcca 1680

aaaacgccc gaaaaaccg acgcgaacct acgccccgaa acgaaagcca aaaaacacta 1740
gacactccct tccggcgta acttccgctt tcccacgta cgtcacttgc cccagtcaaa 1800
caaaactacat atcccgaact tccaagtcg cagccccaaa acaccgcta cacctccccg 1860
cccgcggcc cgccccaaa cccgcctccc gcccccgcc ccgccccg cgcctcatct 1920
cattatcata ttggcttcaa tccaaaataa ggtatattat tgatgatggt ttaaacgat 1980
ccaattcttg aagacgaaag ggcctcgtga tacgcctatt tttataggtt aatgtcatga 2040
taataatggt ttcttagacg tcaggtagga ctttcgggg aaatgtgcgc ggaacccta 2100

tttgtttatt ttctaaata cattcaaata tgcattcgt catgagacaa taacctgat 2160
aatgtctca ataatttga aaaaggaaga gtatgagtat tcaacattc cgtgtcgccc 2220
ttattccctt ttttgcgca ttttgcctt ctgttttgc tcaccagaa acgtggtga 2280
aagtaaaaga tgctgaagat cagttgggtg cagcagtggt ttacatgaa ctggatctca 2340
acagcggtga gatccttgag agttttcgcc ccgaagaacg tttccaatg atgagcatt 2400
ttaaagtctt gctatgtggc gcggtattat ccggtgtga cgccgggcaa gagcaactcg 2460

gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc 2520

atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata 2580

acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt 2640

tgcaacaat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag 2700

ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca acgttgcgca 2760

aactattaac tggcgaacta cttactctag cttcccggca acaattaata gactggatgg 2820

aggcgataa agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg 2880

ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag 2940

atggtaagcc ctcccgtatc gtagttatct acacgacggg gagicaggca actatggatg 3000

aacgaaatag acagatcgct gagatagggt cctcactgat taagcattgg taactgtcag 3060

accaagttaa ctcatatata ctttagattg atttaaaagg atctaggtga agatcctttt 3120

tgataatctc atgacaaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc 3180

cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt 3240

gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccggatcaag agctaccaac 3300

tccttttccg aaggtaactg gcttcagcag agcgcagata ccaataactg tccttctagt 3360

gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgtct 3420

gctaactctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga 3480

ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac 3540

acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg 3600

agaaagcgcc acgcttcccg aaggagaaaa ggcgacagg tatccggtaa gcggcagggt 3660

cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggatc tttatagtcc 3720

tgtegggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg 3780

gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttccctggcct tttgctggcc 3840

ttgaagctgt cctgatggt cgtcatctac ctgcctggac agcatggcct gcaacgcggg 3900

catccgatg ccgccggaag cgagaagaat cataatgggg aaggccatcc agcctcgcgt 3960

cgcagatccg aattcgttta aac 3983

<210> 41

<211> 25

<212> DNA

<213> Human cytomegalovirus

<400> 41

acatcaatgg gcgtggatag cggtt	25
<210> 42	
<211> 1187	
<212> DNA	
<213> Human cytomegalovirus	
<400> 42	
ccattgcata cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca	60
ttaccgccat gttgacattg attattgact agttattaat agtaatcaat tacgggggtca	120
ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgctt	180
ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta	240
acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggt aactgcccac	300
ttggcagtac atcaagtgt tcatatgcca agtacgcccc ctattgacgt caatgacggt	360
aaatggcccc cctggcatta tgcccagtac atgacctat gggactttcc tacttggcag	420
tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca gtacatcaat	480
gggcgtggat agcggtttga ctacgggga ttccaagtc tccaccccat tgacgtcaat	540
gggagtttgt ttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc	600
ccattgacgc aaatggcgcg taggcgtgta cgggtggagg tctatataag gcgaagcgct	660
ccctatcagt gatagagatc tccctatcag tgatagagat cgtcgacgag ctgcgcgcgg	720
gcgggagtcg ctgcgcgctg ccttcgcccc gtgccccgct ccgccgccgc ctgcgcgccg	780
ccgccccggc tctgactgac cgcgttacta aaacaggtaa gtccggcctc cgcgccgggt	840
tttggcgctt cccgcgggcg cccccctctt cacggcgagc gctgccacgt cagacgaagg	900
gcgcagcgag cgtcctgac ctccgccccg gacgtcagg acagcgcccc gctgctcata	960
agactcggcc ttagaacccc agtatcagca gaaggacatt ttaggacggg acttgggtga	1020
ctctagggca ctggttttct ttccagagag cggaacaggc gaggaaaagt agtcccttct	1080
cggcgattct gcggagggat ctccgtgggg cggatgaacgc cgatgatgcc tctactaacc	1140
atgttcatgt tttctttttt tttctacagg tctgggtga cgaacag	1187