



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년03월05일
(11) 등록번호 10-2775366
(24) 등록일자 2025년02월26일

(51) 국제특허분류(Int. Cl.)
C12N 15/85 (2006.01) A01K 67/0278 (2024.01)
C12N 5/0735 (2010.01) C12N 9/64 (2006.01)
(52) CPC특허분류
C12N 15/8509 (2013.01)
A01K 67/0278 (2024.01)
(21) 출원번호 10-2023-7003004(분할)
(22) 출원일자(국제) 2017년02월27일
심사청구일자 2023년01월26일
(85) 번역문제출일자 2023년01월26일
(65) 공개번호 10-2023-0021759
(43) 공개일자 2023년02월14일
(62) 원출원 특허 10-2018-7026552
원출원일자(국제) 2017년02월27일
심사청구일자 2022년02월10일
(86) 국제출원번호 PCT/US2017/019574
(87) 국제공개번호 WO 2017/151453
국제공개일자 2017년09월08일
(30) 우선권주장
62/301,023 2016년02월29일 미국(US)
(56) 선행기술조사문헌
W02015171861 A1
(뒷면에 계속)
전체 청구항 수 : 총 12 항

(73) 특허권자
리제너론 파마슈티칼스 인코포레이티드
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
(72) 발명자
퍼셀 은감보 리사
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
무지카 알렉산더 오
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
탕 야준
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
(74) 대리인
특허법인와이에스장

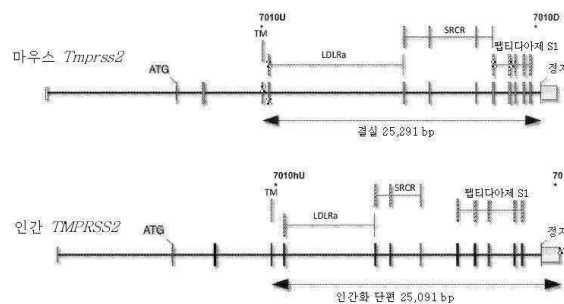
심사관 : 박영관

(54) 발명의 명칭 인간화 TMPRSS 유전자를 갖는 설치류

(57) 요약

유전자 변형된 마우스 및 랫트와 같은 설치류, 및 이를 만들고 사용하기 위한 방법 및 조성물이 제공된다. 설치류는 내인성 설치류 *Tmprss2*, *Tmprss4*, 또는 *Tmprss11d* 유전자와 같은 적어도 하나의 내인성 설치류 *Tmprss* 유전자의 인간화를 포함한다.

대표도



(52) CPC특허분류

C12N 5/0606 (2013.01)
C12N 9/6424 (2013.01)
A01K 2207/15 (2013.01)
A01K 2217/072 (2013.01)
A01K 2227/105 (2013.01)
A01K 2267/0337 (2013.01)
C12N 2510/00 (2013.01)

(56) 선행기술조사문헌

KR1020150008383 A
 WO2015077071 A1
 KR101693243 B1
 J Virol. vol.84 no.11 pp.5605-5614 2010.
 N Kuhn, Studies on the host response to
 influenza A virus infections in mouse
 knock-out mutants, 박사학위 University of
 Veterinary Medicine Hannover(2015.12.31)
 Yu Sun, "Characterization of the TMPRSS2
 Protease as a Modulator of Prostate Cancer
 Metastasis"(2009.03.31.)
 JP2017536126 A

명세서

청구범위

청구항 1

5' 뉴클레오티드 서열 및 3' 뉴클레오티드 서열이 측면에 위치하는, 인간 *TMPRSS* 단백질의 세포외 도메인 (ectodomain)을 암호화하는 핵산을 포함하는 핵산 벡터로서,

상기 5' 뉴클레오티드 서열 및 3' 뉴클레오티드 서열은 동족 설치류 *Tmprss* 단백질의 세포외 도메인을 암호화하는 설치류 *Tmprss* 유전자좌에서 설치류 게놈 DNA의 측면에 위치하는 게놈 DNA 서열에 대해 상동성이고, 설치류 게놈 DNA와 인간 *TMPRSS* 단백질의 세포외 도메인을 암호화하는 핵산의 대체 및 상동성 재조합을 매개하여 인간화 *Tmprss* 유전자를 형성할 수 있고; 및

설치류는 마우스 또는 랫트인, 벡터.

청구항 2

제1 항에 있어서, 인간 *TMPRSS* 단백질은 인간 *TMPRSS2* 단백질이고, 동족 설치류 *Tmprss* 단백질은 설치류 *Tmprss2* 단백질이고, 인간화 *Tmprss* 유전자는 인간화 *Tmprss2* 유전자인 것을 특징으로 하는 벡터.

청구항 3

제2 항에 있어서, 핵산은 인간 *TMPRSS2* 유전자의 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 포함하는 것을 특징으로 하는 벡터.

청구항 4

제2 항에 있어서, 인간화 *Tmprss2* 유전자는

- (i) 내인성 설치류 *Tmprss2* 유전자의 코딩 엑손 1 내지 2,
- (ii) 내인성 설치류 *Tmprss2* 유전자의 코딩 엑손 3의 5' 부분 및 인간 *TMPRSS2* 유전자의 코딩 엑손 3의 3' 부분을 포함하는 엑손 3, 및
- (iii) 인간 *TMPRSS2* 유전자의 코딩 엑손 4 내지 코딩 엑손 13을 포함하는 것을 특징으로 하는 벡터.

청구항 5

제1 항에 있어서, 인간 *TMPRSS* 단백질은 인간 *TMPRSS4* 단백질이고, 동족 설치류 *Tmprss* 단백질은 설치류 *Tmprss4* 단백질이고, 인간화 *Tmprss* 유전자는 인간화 *Tmprss4* 유전자인 것을 특징으로 하는 벡터.

청구항 6

제5 항에 있어서, 핵산은 인간 *TMPRSS4* 유전자의 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 포함하는 것을 특징으로 하는 벡터.

청구항 7

제5 항에 있어서, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자의 코딩 엑손 1 내지 코딩 엑손 3, 및 인간 *TMPRSS4* 유전자의 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 포함하는 것을 특징으로 하는 벡터.

청구항 8

제1 항에 있어서, 인간 *TMPRSS* 단백질은 인간 *TMPRSS11D* 단백질이고, 동족 설치류 *Tmprss* 단백질은 설치류 *Tmprss11d* 단백질이고, 인간화 *Tmprss* 유전자는 인간화 *Tmprss11d* 유전자인 것을 특징으로 하는 벡터.

청구항 9

제8 항에 있어서, 핵산은 인간 *TMPRSS11D* 유전자의 코딩 엑손 3 내지 코딩 엑손 10 내의 정지 코돈을 포함하는

것을 특징으로 하는 벡터.

청구항 10

제8 항에 있어서, 인간화 *Tmprss11d* 유전자는 내인성 설치류 *Tmprss11d* 유전자의 코딩 엑손 1 내지 2, 및 인간 *TMPRSS11D* 유전자의 코딩 엑손 3 내지 코딩 엑손 10을 포함하는 것을 특징으로 하는 벡터.

청구항 11

제1 항 내지 제10 항 중 어느 한 항에 있어서, 설치류는 마우스인 것을 특징으로 하는 벡터.

청구항 12

제1 항 내지 제10 항 중 어느 한 항에 있어서, 설치류는 랫트인 것을 특징으로 하는 벡터.

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

발명의 설명

기술 분야

[0001]

관련 출원에 대한 상호 참조

[0002]

본 출원은 그 내용 전체가 참조로서 본원에 통합되는 2016년 2월 29일에 출원된 미국 가출원 번호 제62/301,023호의 우선권의 이익을 주장한다.

[0003]

참조로서 서열 목록 포함

[0004]

2017년 2월 13일에 생성되고, EFS-Web을 통해 미국 특허 상표청에 제출된 33093_10234US01_SequenceListing.txt로 명명된 275 KB의 ASCII 텍스트 파일 내 서열 목록이 참조로써 본원에 통합된다.

배경 기술

[0005]

II형 막관통 세린 프로테아제는 N 말단 막관통 도메인을 특징으로 하는 프로테아제 계열이다(Bugge 외의 *J. Biol. Chem.* 284 (35): 23177-23181, 2009; Hooper 외의 *J. Biol. Chem.* 272(2): 857-860, 2001). 본 계열의 모든 구성원은 단쇄 지모겐(single-chain zymogen)으로서 발현되고 고도로 보존된 R/(IV)VGG 모티프 내에서의 절단에 의해 단백질 분해 방식으로 활성화된다. 본 계열의 일 구성원으로, 막관통 프로테아제인 세린 4형

(TMPRSS4)은 상피를 가로지르는 나트륨과 물의 흐름을 조절하는 상피 나트륨 채널(ENaC)를 활성화시키는 것으로 나타났다(Guipponi 외의 2002 *Hum. Mol. Genet.* 11:2829; Vuagniaux 외의 2002 *J. Gen. Physiol.* 120:191). TMPRSS4의 단백질 분해 활성자는 알려져 있지 않지만; 현재까지 이용 가능한 데이터는 단백질이 자체 활성화됨을 암시한다. 활성화될 때, TMPRSS4의 촉매 도메인은 이황화 결합을 통해 단백질의 N 말단에 결합된 채로 유지된다. TMPRSS4, TMPRSS2 및 TMPRSS11D (또는 인간 기도 트립신 유사 프로테아제; *Human Airway Trypsin-like protease*; "HAT")는 바이러스 생명 주기의 제1 필수 단계인 인플루엔자 A 혈구 응집소(hemagglutinin; HA)를 시험관 내에서 절단하는 것으로 나타났다. 이러한 절단은, 단백질이 전구체 단백질(HA0)로서 합성되고 활성을 위해 HA1 및 HA2로의 절단을 필요로 하므로 HA의 활성화에 필수적이다. Caco-2 세포에서 *TMPRSS4*의 RNAi 녹다운은 바이러스의 확산을 감소시켰다. 또한, TMPRSS4는 인플루엔자로 감염된 마우스의 폐에서 강하게 상향 조절되는 것으로 나타났다(Böttcher 외의 2006 *J. Virol.* 80:9896; Böttcher 외의 2009 *Vaccine* 27: 6324; Böttcher-Friebershäuser 외의 2010 *J. Virol.* 84: 5604; Bertam 외의 2010 *J. Virol.* 84:10016; Bertam 외의 2010 *J. Virol.* 84:10016; Böttcher-Friebershäuser 외의 2011 *J. Virol.* 85: 1554; Bahgat 외의 2011 *J. Virol.* 85: 8:27).

[0006] 바이러스 감염 및 다른 질병의 치료 및 예방을 위해 인간 II형 막관통 세린 프로테아제를 특이적으로 표적화하는 항체를 포함하는 화합물을 식별하고 시험하기 위해 생체 내 시스템, 예를 들어 감염의 설치류 모델을 개발하는 것이 필요하다.

발명의 내용

[0007] 본 발명은 새로운 치료제를 확인하고 개발하기 위해 생체 내 시스템을 제공하기 위해 설치류 동물을 조작하는 것이 바람직하다는 인식을 포함한다. 예를 들어, 본 발명은 인간화 *Tmprss* 유전자를 갖는 설치류가 바이러스 감염의 치료 및 예방을 위한 치료제의 식별과 개발에 사용하기에 바람직하다는 인식을 포함한다.

[0008] 일 양태에서, 본 발명은 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열 및 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss* 유전자를 함유하는 게놈을 가진 설치류를 제공하되, 상기 인간화 *Tmprss* 유전자는 내인성 설치류 *Tmprss* 유전자의 5' 조절 서열(들), 예컨대 프로모터 및/또는 증강자(들)에 의해 조절된다.

[0009] 일부 구현예에서, 본원에 개시된 설치류 내의 인간화 *Tmprss* 유전자는 인간 TMPRSS 단백질의 세포외 도메인(ectodomain)과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유하는 인간화 *Tmprss* 단백질을 암호화한다. 일부 구현예에서, 인간화 *Tmprss* 단백질은 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포질 및 막관통 부분을 함유한다.

[0010] 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열 및 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss* 유전자를 함유하되, 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열은 동족 인간 *TMPRSS* 유전자에 의해 암호화된 인간 TMPRSS 단백질의 세포외 도메인과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 폴리펩티드를 암호화한다. 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열 및 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss* 유전자를 함유하되, 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss* 유전자에 의해 암호화된 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 폴리펩티드를 암호화한다.

[0011] 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss* 유전자의 인접 게놈 서열을 동족 인간 *TMPRSS* 유전자의 인접 게놈 서열과 대체함으로써 생성된 내인성 설치류 *Tmprss* 유전자좌에 위치한 인간화 *Tmprss* 유전자를 함유한다. 특정 구현예에서, 삽입되는 동족 인간 *TMPRSS* 유전자의 인접 게놈 서열은 인간 *TMPRSS* 유전자에 의해 암호화된 인간 TMPRSS 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화하는 엑손 서열을 포함한다. 일부 구현예에서, 동족 인간 *TMPRSS* 유전자의 인접 게놈 서열은 또한 동족 인간 *TMPRSS* 유전자의 3' UTR을 포함한다.

[0012] 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss* 유전자좌에서 인간화 *Tmprss* 유전자에 대해 이형접합체이다. 다른 구현예에서, 설치류는 내인성 설치류 *Tmprss* 유전자좌에서 인간화 *Tmprss* 유전자에 대해 동

형접합체이다.

- [0013] 또 다른 구현예에서, 설치류는 둘 이상의 인간화 *Tmprss* 유전자(예를 들어, 인간화 *Tmprss2*, 인간화 *Tmprss4*, 및 인간화 *Tmprss11d* 유전자 중 둘 이상)를 상이한 내인성 설치류 *Tmprss* 유전자좌에 (각각의 내인성 설치류 *Tmprss* 유전자좌는 각각 동족 인간 *TMPRSS* 유전자로 인간화됨) 함유한다.
- [0014] 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss2* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss2* 유전자를 함유하되, 인간화 *Tmprss2* 유전자는 내인성 설치류 *Tmprss2* 유전자의 프로모터에 의해 조절된다.
- [0015] 일부 구현예에서, 인간화 *Tmprss2* 유전자는 인간화에 사용된 인간 *TMPRSS2* 유전자에 의해 암호화된 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유하는 인간화 *Tmprss2* 단백질을 암호화한다. 인간 *TMPRSS2* 단백질은, 일부 구현예에서, 서열 번호 4에 명시된 바와 같은 아미노산 서열과 적어도 85% 동일한 (예: 적어도 90%, 95%, 98%, 99% 또는 100% 동일한) 아미노산 서열을 함유한다. 일부 구현예에서, 인간화 *Tmprss2* 단백질은 예를 들어 서열 번호 4에 명시된 바와 같은 인간 *TMPRSS2* 단백질의 잔기 W106 내지 G492 또는 C 말단 387 아미노산으로 이루어진 아미노산 서열과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유한다. 일부 구현예에서, 인간화 *Tmprss2* 유전자는 인간화되는 내인성 설치류 *Tmprss2* 유전자에 의해 암호화된 설치류 *Tmprss2* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포질 및 막관통 부분을 추가로 함유하는 인간화 *Tmprss2* 단백질을 암호화한다. 예시적인 내인성 설치류 *Tmprss2* 단백질은 서열 번호 2에 명시된다.
- [0016] 일부 구현예에서, 설치류는 내인성 설치류 *Tmprss2* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss2* 유전자를 함유하되, 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS2* 유전자에 의해 암호화된 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 암호화한다. 특정 구현예에서, 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS2* 유전자의 코딩 엑손 13에 코딩 엑손 4 내지 정지 코돈을 함유하는 인간 *TMPRSS2* 유전자의 인접 게놈 서열이다. 특정 구현예에서, 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자의 3' UTR을 추가로 함유한다. 일부 구현예에서, 인간화 *Tmprss2* 유전자에 포함된 내인성 설치류 *Tmprss2* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss2* 유전자에 의해 암호화된 내인성 설치류 *Tmprss2* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포질 및 막관통 부분을 암호화한다.
- [0017] 특정 구현예에서, 인간화 *Tmprss2* 유전자는 내인성 설치류 *Tmprss2* 유전자의 코딩 엑손 1 내지 2 및 인간 *TMPRSS2* 유전자의 코딩 엑손 4 내지 코딩 엑손 13을 포함하되, 인간화 *Tmprss2* 유전자는 내인성 설치류 *Tmprss2* 유전자에 의해 암호화된 설치류 *Tmprss2* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 포함하는 인간화 *Tmprss2* 단백질, 및 인간 *TMPRSS2* 유전자에 의해 암호화된 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화한다. 인간화 *Tmprss2* 유전자는 엑손 3을 일부 구현예에서는 인간 *TMPRSS2* 유전자의 코딩 엑손 3에 함유하고, 다른 구현예에서는 내인성 설치류 *Tmprss2* 유전자의 코딩 엑손 3에 함유한다. 일부 구현예에서, 인간화 *Tmprss2* 유전자는 내인성 설치류 *Tmprss2* 유전자의 코딩 엑손 3의 5' 부분 및 인간 *TMPRSS2* 유전자의 코딩 엑손 3의 3' 부분을 포함하는 엑손 3을 함유한다.
- [0018] 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss4* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS4* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss4* 유전자를 함유하되, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자의 프로모터에 의해 조절된다.
- [0019] 일부 구현예에서, 인간화 *Tmprss4* 유전자는 인간화에 사용된 인간 *TMPRSS4* 유전자에 의해 암호화된 인간 *TMPRSS4* 단백질의 세포외 도메인과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유하는 인간화 *Tmprss4* 단백질을 암호화한다. 인간 *TMPRSS4* 단백질은, 일부 구현예에서, 서열 번호 11에 명시된 바와 같은 아미노산 서열과 적어도 85% 동일한 (예: 적어도 90%, 95%, 98%, 99% 또는 100% 동일한) 아미노산 서열을 함유한다. 일부 구현예에서, 인간화 *Tmprss4* 단백질은 서열 번호 11에 명시된 바와 같은 인간 *TMPRSS4* 단백질의 잔기 K54 내지 L437 또는 C 말단 384 아미노산으로 이루어진 아미노산 서열과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유한다. 일부 구현예에서, 인간화 *Tmprss4* 유전자는 인간화되는 내인성 설치류 *Tmprss4* 유전자에 의해 암호화된 설치류 *Tmprss4* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는

100% 동일한) 세포질 및 막관통 부분을 추가로 함유하는 인간화 *Tmprss4* 단백질을 암호화한다. 예시적인 내인성 설치류 *Tmprss4* 단백질은 서열 번호 9에 명시된다.

- [0020] 일부 구현예에서, 설치류는 내인성 설치류 *Tmprss4* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS4* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss4* 유전자를 함유하되, 인간 *TMPRSS4* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS4* 유전자에 의해 암호화된 인간 *TMPRSS4* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화한다. 특정 구현예에서, 인간 *TMPRSS4* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS4* 유전자의 코딩 엑손 13에 코딩 엑손 4 내지 정지 코돈을 함유하는 인접 게놈 서열이다. 일부 구현예에서, 인간화 *Tmprss4* 유전자에 포함된 내인성 설치류 *Tmprss4* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss4* 유전자에 의해 암호화된 내인성 설치류 *Tmprss4* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 암호화한다.
- [0021] 특정 구현예에서, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자의 코딩 엑손 1 내지 3 및 인간 *TMPRSS4* 유전자의 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유한다.
- [0022] 일부 구현예에서, 본원에 개시된 설치류는 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss11d* 유전자를 함유하되, 인간화 *Tmprss11d* 유전자는 내인성 설치류 *Tmprss11d* 유전자의 프로모터에 의해 조절된다.
- [0023] 일부 구현예에서, 인간화 *Tmprss11d* 유전자는 인간화에 사용된 인간 *TMPRSS11D* 유전자에 의해 암호화된 인간 *TMPRSS11D* 단백질의 세포외 도메인과 실질적으로 동일한 (예: 서열이 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유하는 인간화 *Tmprss11d* 단백질을 암호화한다. 인간 *TMPRSS11D* 단백질은, 일부 구현예에서, 서열 번호 18에 명시된 바와 같은 아미노산 서열과 적어도 85% 동일한 (예: 적어도 90%, 95%, 98%, 99% 또는 100% 동일한) 아미노산 서열을 함유한다. 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 예를 들어 서열 번호 18에 명시된 바와 같은 인간 *TMPRSS11D* 단백질의 잔기 A42 내지 I418 또는 C 말단 377 아미노산으로 이루어진 아미노산 서열과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유한다. 일부 구현예에서, 인간화 *Tmprss11d* 유전자는 인간화되는 내인성 설치류 *Tmprss11d* 유전자에 의해 암호화된 설치류 *Tmprss11d* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 (예: 적어도 85%, 90%, 95%, 98%, 99% 또는 100% 동일한) 세포질 및 막관통 부분을 추가로 함유하는 인간화 *Tmprss11d* 단백질을 암호화한다. 예시적인 내인성 설치류 *Tmprss11d* 단백질은 서열 번호 16에 명시된다.
- [0024] 일부 구현예에서, 설치류는 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열을 포함하는 인간화 *Tmprss11d* 유전자를 함유하되, 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS11D* 유전자에 의해 암호화된 인간 *TMPRSS11D* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화한다. 특정 구현예에서, 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS11D* 유전자의 코딩 엑손 3에 내지 코딩 엑손 10 내의 정지 코돈을 함유하는 인접 게놈 서열이다. 특정 구현예에서, 인간 *TMPRSS11D* 유전자의 인접 게놈 서열은 인간 *TMPRSS11D* 유전자의 3' UTR을 추가로 함유한다. 일부 구현예에서, 인간화 *Tmprss11d* 유전자에 포함된 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss11d* 유전자에 의해 암호화된 내인성 설치류 *Tmprss11d* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 암호화한다.
- [0025] 특정 구현예에서, 인간화 *Tmprss11d* 유전자는 내인성 설치류 *Tmprss11d* 유전자의 코딩 엑손 1 내지 2 및 인간 *TMPRSS11D* 유전자의 코딩 엑손 3 내지 코딩 엑손 13을 함유한다.
- [0026] 또 다른 양태에서, 본 발명은 본원에 기술된 바와 같은 인간화 *Tmprss* 유전자를 함유하는 게놈을 가진 단리된 설치류 세포 또는 조직을 제공한다. 특정 구현예에서, 인간화 *Tmprss* 유전자는 인간화 *Tmprss2* 유전자, 인간화 *Tmprss4* 유전자, 및 인간화 *Tmprss11d* 유전자로 이루어지는 군으로부터 선택된다.
- [0027] 또 다른 양태에서, 본 발명은 본원에 기술된 바와 같은 인간화 *Tmprss* 유전자를 함유하는 게놈을 가진 설치류 배주 줄기 세포를 제공한다. 특정 구현예에서, 인간화 *Tmprss* 유전자는 인간화 *Tmprss2* 유전자, 인간화 *Tmprss4* 유전자, 및 인간화 *Tmprss11d* 유전자로 이루어지는 군으로부터 선택된다.
- [0028] 또 다른 양태에서, 본원에 기술된 설치류 배아 줄기 세포로부터 생성된 설치류 배아가 또한 제공된다.
- [0029] 일 양태에서, 본 발명은 설치류 내에서 내인성 *Tmprss* 유전자의 인간화에 사용하기에 적합한 핵산 벡터를 제공한다. 일부 구현예에서, 핵산 벡터는 5' 상동 아암 및 3' 상동 아암이 측면에 위치한 인간 *Tmprss* 핵산 서열(예: 인간 *TMPRSS* 단백질의 세포외 도메인을 암호화하는 인간 게놈 DNA)을 포함한다. 5' 및 3' 상동 아암은

인간 *Tmprss* 핵산 서열에 대해 5' 및 3'에 각각 위치하는 핵산 서열이며, 동족 설치류 *Tmprss* 단백질의 세포외 도메인을 암호화하는 설치류 게놈 DNA의 측면에 위치하는 설치류의 내인성 *Tmprss* 유전자좌에서 게놈 DNA 서열에 대해 상동성이다. 따라서, 5' 및 3' 상동 아암은 동족 설치류 *Tmprss* 단백질의 세포외 도메인을 암호화하는 설치류 게놈 DNA의 상동성 재조합 및 인간 *Tmprss* 핵산 서열과의 대체를 매개하여 본원에 기술된 바와 같은 인간화 *Tmprss* 유전자를 형성할 수 있다.

[0030] 추가적인 양태에서, 본 발명은 인간화 *Tmprss* 유전자를 함유하는 게놈을 가진 설치류를 제공하는 방법에 관한 것이다. 상기 방법은 내인성 설치류 *Tmprss* 유전자의 게놈 서열을 동족 인간 *TMPRSS* 유전자의 게놈 서열로 대체시켜 인간화 *Tmprss* 유전자를 형성하기 위해 설치류의 게놈을 변형시키는 것을 포함한다.

[0031] 일부 구현예에서, 본 발명은 인간화 *Tmprss* 유전자를 갖는 설치류(예: 마우스 또는 랫트)를 만드는 방법으로서, (a) 설치류 배아 줄기 세포의 내인성 설치류 *Tmprss* 유전자좌에 동족 *TMPRSS* 유전자의 뉴클레오타이드 서열을 함유하는 게놈 단편을 삽입하여 (본원에 기술된 것들과 같은) 인간화 *Tmprss* 유전자를 형성하는 단계; (b) (a) 단계의 인간화 *Tmprss* 유전자를 포함하는 설치류 배아 줄기 세포를 수득하는 단계; 및 (c) (b) 단계의 설치류 배아 줄기 세포를 이용하여 설치류를 생성하는 단계를 포함하는 방법을 제공한다.

[0032] 일부 구현예에서, 인간화 *Tmprss* 유전자는 인간화 *Tmprss2* 유전자, 인간화 *Tmprss4* 유전자, 및 인간화 *Tmprss11d* 유전자로 이루어지는 군으로부터 선택된다. 다양한 구현예에서, 인간화 *Tmprss* 유전자는 인간화에 사용된 인간 *TMPRSS* 유전자에 의해 암호화된 인간 *TMPRSS* 단백질의 세포외 도메인과 실질적으로 동일한 (예: 서열이 적어도 90%, 95%, 98%, 99% 또는 100% 동일한) 세포외 도메인을 함유하는 인간화 *Tmprss* 단백질을 암호화한다. 특정 구현예에서, 인간화 *Tmprss* 단백질은 인간 *TMPRSS2* 단백질, 인간 *TMPRSS4* 단백질, 및 인간 *TMPRSS11D* 단백질로 이루어지는 군으로부터 선택된 인간 *TMPRSS* 단백질의 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss* 단백질은 인간화되는 내인성 설치류 *Tmprss* 유전자에 의해 암호화된 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 추가로 함유한다.

[0033] 또 다른 양태에서, 본 발명은 인플루엔자 바이러스 감염에 있어서 화합물(예: 인간 *TMPRSS* 단백질을 특이적으로 표적화하는 후보 억제제)의 치료 효능을 평가하기 위해 본원에 기술된 설치류를 사용하는 방법을 제공한다. 상기 방법은, 본원에 기술된 설치류를 제공하는 단계; 설치류에 인플루엔자 바이러스와 후보 화합물을 투여하는 단계; 및 설치류에서 인플루엔자 바이러스 감염의 존재 및 중증도를 모니터링하여 후보 약물의 치료 효능을 알아내는 단계를 포함할 수 있다.

[0034] 일부 구현예에서, 인플루엔자 바이러스는 화합물 이전에 설치류에 투여된다. 일부 구현예에서, 인플루엔자 바이러스는 화합물 이후에 설치류에 투여된다.

[0035] 일부 구현예에서, 후보 화합물은 인간 *TMPRSS* 단백질에 대해 특이적인 항체 또는 이의 항원 결합 단편이다. 특정 구현예에서, 후보 화합물은 인간 *TMPRSS2* 단백질, 인간 *TMPRSS4* 단백질, 및 인간 *TMPRSS11D* 단백질로 이루어지는 군으로부터 선택된 인간 *TMPRSS* 단백질에 대해 특이적인 항체 또는 이의 항원 결합 단편이다.

[0036] 본 발명의 다른 특징, 목적 및 장점은 다음의 상세한 설명에서 명백해진다. 그러나, 상세한 설명은 본 발명의 특정한 구현예를 나타내지만, 단지 예시로서 제공될 뿐 제한하고자 함이 아닌 것으로 이해해야 한다. 본 발명의 범주 내 다양한 변형예들 및 수정예들이 상세한 설명으로부터 당업자에게 명백해질 것이다.

도면의 간단한 설명

[0037] 다음의 도면들로 구성된 본원에 포함된 도면은 단지 예시적인 목적이며 제한하고자 하는 것이 아니다.

도 1a 내지 1d. 마우스 *Tmprss2*의 인간화를 위한 예시적 전략

도 1a는 마우스 *Tmprss2* 및 인간 *TMPRSS2* 유전자의 게놈 조직의 도면을 도시하며, 축척에 비례하지는 않는다. 엑손은 게놈 서열 전체에 걸쳐 삽입된 가는 막대로 표시되며, 두 가지 유전자 모두에 대한 첫 번째 코딩 엑손은 엑손 위에 시작 코돈 "ATG"로 표시되고, 마지막 코딩 엑손은 엑손 위에 "정지" 코돈으로 표시되어 있다. 결실할 약 25,291 bp의 마우스 게놈 단편과 삽입할 약 25,091 bp의 인간 게놈 단편이 표시되어 있다. 실시예 1에 기재된 검정에 사용된 프로브의 위치가 표시되어 있다. TM: 막관통 도메인; SRCR: 스캐빈저 수용체 시스템-리치 유사 도메인; LDLRa: 저밀도 리포단백질 수용체 클래스 A.

도 1b는 내인성 마우스 *Tmprss2* 유전자의 인간화를 위한 예시적인 변형된 BAC 벡터를 접합 서열(서열 번호 22, 23 및 24)과 함께 도시하며, 축척에 비례하지는 않는다.

도 1c는 네오마이신 카세트가 결실된 후의 인간화 *Tmprss2* 대립유전자를 접합 서열(서열 번호 22 및 25)과 함께 도시하며, 축척에 비례하지는 않는다.

도 1d는 인간 *TMPRSS2* 단백질(서열 번호 4), 마우스 *Tmprss2* 단백질(서열 번호 2), 및 인간화 *Tmprss2* 단백질("7010 돌연변이 단백질")(서열 번호 7)의 서열 배열을 제시한다.

도 2a 내지 2d 마우스 *Tmprss4*의 인간화를 위한 예시적 전략

도 2a는 마우스 *Tmprss4* 및 인간 *TMPRSS4* 유전자의 게놈 조직의 도면을 도시하며, 축척에 비례하지는 않는다. 엑손은 게놈 서열 전체에 걸쳐 삽입된 가는 막대로 표시되며, 두 가지 유전자 모두에 대한 첫 번째 엑손(및 첫 번째 코딩 엑손)은 엑손 위에 시작 코돈 "ATG"로 표시되고, 마지막 코딩 엑손은 엑손 위에 "정지" 코돈으로 표시되어 있다. 결실할 약 11,074 bp의 마우스 게놈 단편과 삽입할 약 14,963 bp의 인간 게놈 단편이 표시되어 있다. 실시예 2에 기재된 검정에 사용된 프로브의 위치가 표시되어 있다. TM: 막관통 도메인; SRCR: 스캐빈저 수용체 시스테인-리치 유사 도메인; LDLRa: 저밀도 리포단백질 수용체 클래스 A.

도 2b는 내인성 마우스 *Tmprss4* 유전자의 인간화를 위한 예시적인 변형된 BAC 벡터를 접합 서열(서열 번호 38, 39 및 40)과 함께 도시하며, 축척에 비례하지는 않는다.

도 2c는 네오마이신 카세트가 결실된 후의 인간화 *Tmprss4* 대립유전자를 접합 서열(서열 번호 41 및 40)과 함께 도시하며, 축척에 비례하지는 않는다.

도 2d는 인간 *TMPRSS4* 단백질(서열 번호 11), 마우스 *Tmprss4* 단백질(서열 번호 9), 및 인간화 *Tmprss4* 단백질("7224 돌연변이 단백질")(서열 번호 14)의 서열 배열을 제시한다.

도 3a 내지 3d 마우스 *Tmprss11d*의 인간화를 위한 예시적 전략

도 3a는 마우스 *Tmprss11d* 및 인간 *TMPRSS11D* 유전자의 게놈 조직의 도면을 도시하며, 축척에 비례하지는 않는다. 엑손은 게놈 서열 전체에 걸쳐 삽입된 가는 막대로 표시되며, 두 가지 유전자 모두에 대한 첫 번째 엑손(및 첫 번째 코딩 엑손)은 엑손 위에 시작 코돈 "ATG"로 표시되고, 마지막 코딩 엑손은 엑손 위에 "정지" 코돈으로 표시되어 있다. 결실할 약 35,667 bp의 마우스 게놈 단편과 삽입할 약 33,927 bp의 인간 게놈 단편이 표시되어 있다. 실시예 3에 기재된 검정에 사용된 프로브의 위치가 표시되어 있다. TM: 막관통 도메인; SEA: 성게 정자 단백질, 엔테로키나아제(enterokinase) 및 아그린(agrins)에서 발견된 도메인.

도 3b는 내인성 마우스 *Tmprss11d* 유전자의 인간화를 위한 예시적인 변형된 BAC 벡터를 접합 서열(서열 번호 57, 58 및 59)과 함께 도시하며, 축척에 비례하지는 않는다.

도 3c는 네오마이신 카세트가 결실된 후의 인간화 *Tmprss11* 대립유전자를 접합 서열(서열 번호 57 및 60)과 함께 도시하며, 축척에 비례하지는 않는다.

도 3d는 인간 *TMPRSS11D* 단백질(서열 번호 18), 마우스 *Tmprss11d* 단백질(서열 번호 16), 및 인간화 *Tmprss11d* 단백질("7226 돌연변이 단백질")(서열 번호 21)의 서열 배열을 제시한다.

도 4는, MAID7225 HumInTMPRSS4 마우스가 고 용량의 중증 인플루엔자 A H1N1이나 중증 마우스 적응형 H3N2의 투여에 대해 민감성에 있어 차이가 없음을 나타내는 실험 결과를 도시한다. A/Puerto Rico/08/1934(H1N1)(연회색 원, 점선)를 투여한 MAID7225 HumIn TMPRSS4 마우스는 야생형 마우스(연회색 사각형, 점선)와 비교하여 유사한 생존율을 보였다. 마찬가지로, A/Aichi/02/1968-X31(H3N2)(진회색 삼각형, 점선)을 투여한 MAID7225 HumIn TMPRSS4 마우스는 야생형 마우스(연회색 역삼각형, 파선)와 비교하여 유사한 생존율을 보였다. 0일 차에 1150 PFU의 A/Puerto Rico/08/1934(H1N1) 또는 10,000 PFU의 A/Aichi/02/1968-X31(H3N2)로 감염시켰다. 감염되지 않는 음성 대조 MAID7225 HumIn TMPRSS4 및 야생형 마우스(검은 다이아몬드, 실선)를 대조군에 포함시켰다.

발명을 실시하기 위한 구체적인 내용

[0038]

본 발명은 II 형 막관통 세린 프로테아제(또는 *transmembrane protease/serine*을 지칭하는 "*Tmprss*")를 암호화하는 인간화 유전자를 갖는 유전자 변형된 설치류(예: 마우스 및 랫트)에 관한 것이다. 유전자 변형된 설치류는, 인플루엔자 바이러스 감염과 같은 질병의 치료 및 예방을 위해 인간 *TMPRSS* 분자를 특이적으로 표적화하는 후보 화합물에 대한 스크리닝에 사용하기에 적합하다. 따라서, 본 발명은 인간화 *Tmprss* 유전자를 갖는 유전자 변형된 설치류, 유전자 변형된 설치류로부터 분리된 세포 및 조직, 유전자 변형된 설치류를 만드는 방법 및 조성물, 및 치료 화합물의 스크리닝 및 시험을 위한 유전자 변형된 설치류의 용도를 제공한다. 본 발명의 다

양한 구현예가 아래에 더 기술된다.

II 형 막관통 세린 프로테아제("Tmprss")

비인간 분자의 경우 "Tmprss"로 지칭되고 인간 분자의 경우 "TMPRSS"("transmembrane protease/serine")로도 본원에서 지칭되는 II 형 막관통 세린 프로테아제는 N 말단 막관통 도메인 및 C 말단 세포외 세린 프로테아제 도메인을 특징으로 하는 단백질 집단이다. 상기 집단에서 적어도 18개의 구성원을 식별하였는데, 이들은 4개의 하위 집단으로 그룹화된다(전술한 Bugge 외의 (2009) 내용 참조). 모든 구성원은 (i) 짧은 N 말단 세포질 도메인, (ii) 막관통 도메인, 및 (iii) 프로테아제 도메인 및 막관통 도메인을 프로테아제 도메인과 연결시키는 줄기 영역을 함유하는 세포외 도메인 등을 포함하여, 집단을 정의하는 몇 개의 공통적인 구조적 특징을 공유한다. 줄기 영역은 다음의 6가지 상이한 유형의 모듈 구조 도메인의 조합을 함유한다: SEA (sea urchin sperm protein/enteropeptidase/agrin) 도메인, 그룹 A 스캐빈저 수용체 도메인 (group A scavenger receptor domain), LDLA (low-density lipoprotein receptor class A) 도메인, CUB (Cls/Clr urchin embryonic growth factor, bone morphogenetic protein-1) 도메인, MAM (meprip/A5 antigen/receptor protein phosphatase mu) 도메인, 및 프리즐 도메인(frizzled domain). Bugge 외의 전술한 (2009) 리뷰 참조. 예를 들어, 둘 다 헵신(hepsin)/TMPRSS 하위 집단에 속하는 TMPRSS2 및 TMPRSS4는 줄기 영역에서 단일 LDLA 도메인이 선행하는 그룹 A 스캐빈저 수용체 도메인을 갖는다. HAT/DESC 하위 집단에 속하고 human airway trypsin-like protease를 지칭하는 "HAT"로도 알려진 TMPRSS11D는 단일 SEA 도메인을 갖는다. 전술한 Bugge 외의 (2009)의 도 1 참조.

II 형 막관통 세린 프로테아제는 초기에 비활성 효소 전구체로서 생성되는데, 이는 프로테아제 도메인 전의 컨센서스 활성화 모티프에서 염기성 아미노산 잔기를 따르는 절단에 의한 활성화를 필요로 한다. 활성화된 프로테아제의 일부는 프로도메인(prodomain)과 프로테아제 도메인 사이에서의 이황화 결합의 결과로서 막 결합을 유지한다. 세포외 도메인은 세포 국지화, 활성화, 억제 및/또는 이들 프로테아제의 기질 특이성에 대해 중요한 것으로 여겨진다(전술한 Bugge 외의 (2009); Szabo 외의 *Int. J. Biochem. Cell Biol.* 40: 1297-1316 (2008) 참조).

II 형 막관통 세린 프로테아제의 구성원에 대한 다양한 생화학적 및 병태생리학적 정보가 문서화되었다. TMPRSS2, TMPRSS4 및 TMPRSS11D는 바이러스 생명 주기의 제1 필수 단계인 인플루엔자 A 혈구 응집소(hemagglutinin; HA)를 시험관 내에서 절단하는 것으로 나타났다. 본원에 기술된 인간화 *Tmprss* 유전자를 갖는 유전자 변형된 설치류 동물은 TMPRSS 분자의 생물학적 기능에 대한 철저한 이해는 물론이고 인간 TMPRSS 분자를 특이적으로 표적화하는 치료 화합물의 스크리닝을 가능하게 하는 유용한 생체 내 시스템을 제공한다.

마우스, 인간 및 인간화 *Tmprss* 핵산 및 단백질 서열을 포함하는 예시적인 *Tmprss* 서열이 본 출원에 제공되며 다음의 표에 요약되어 있다. 실시예 섹션에서의 분석에 사용된 프라이머와 프로브 서열 및 예시적인 인간화 *Tmprss* 대립유전자의 삽입 접합 서열 또한 표에 포함된다.

[0044] 서열의 요약 설명

서열 번호	설명	특징
1	생쥐 <i>Tmprss2</i> , mRNA, NM_015775.2	길이: 3175 bp CDS: 231~1703 엑손: 1~177; 178~245 (제 2 엑손, 및 제 1 코딩 엑손); 246~465; 466~552; 553~672; 673~799; 800~910; 911~954; 955~1123; 1124~1299; 1300~1395; 1396~1538; 1539~1691; 1692~3161.
2	생쥐 <i>Tmprss2</i> , 단백질	길이: 490 aa
3	인류 <i>TMPRSS2</i> , 전사 변이체 2, mRNA, NM_005656.3	길이: 3212 bp CDS: 135~1613 엑손: 1~78; 79~149 (제 2 엑손, 및 제 1 코딩 엑손); 150~372; 373~459; 460~579; 580~706; 707~817; 818~861; 862~1033; 1034~1209; 1210~1305; 1306~1448; 1449~1601; 1602~3204.
4	인류 <i>TMPRSS2</i> , 전사 변이체 2, 단백질	길이: 492 aa 세포외 도메인: W106 에서 시작함.
5	인간화 <i>Tmprss2</i> 계통 단편	길이: 27,947 bp 1~84: 마우스 서열 85~25175: 인간 서열 (총 25091 bp) 25176~27866: <i>Xhd</i> -LoxP-카세트- loxP- <i>ICeUI</i> - <i>NheI</i> (총 2691 bp) 27867~27947: 마우스 서열

[0045]

6	카세트가 결실된 인간화 <i>Tmprss2</i> 게놈 단편	길이: 25,333 bp 1~84: 마우스 서열 85~25175: 인간 서열 (총 25091 bp) 25176~25252: <i>XhoI</i> - <i>loxP</i> - <i>ICeUI</i> - <i>NheI</i> (77 bp) 25253~25333: 마우스 서열
7	인간화 <i>Tmprss2</i> 단백질	길이: 491 aa
8	생쥐 <i>Tmprss4</i> , mRNA, NM_145403.2	길이: 2267 bp CDS: 289~1596 엑손: 1~291 (제 1 엑손 및 제 1 코딩 엑손); 292~325; 326~439; 440~592; 593~722; 723~824; 825~865; 866~1025; 1026~1192; 1193~1291; 1292~1434; 1435~1584; 1585~2267.
9	생쥐 <i>Tmprss4</i> , 단백질	길이: 435 aa
10	인류 <i>TMPRSS4</i> , 전사 변이체 4, mRNA, NM_001173551.1	길이: 3543 bp CDS: 292~1599 엑손: 1~294 (제 1 엑손 및 제 1 코딩 엑손); 295~328; 329~442; 443~595; 596~725; 726~827; 828~868; 869~1028; 1029~1195; 1196~1294; 1295~1437; 1438~1587; 1588~3529.
11	인류 <i>TMPRSS4</i> , 전사 변이체 4, 단백질	길이: 437 aa 세포외 도메인: K54 에서 시작함.

[0046]

12	카세트를 함유하는 인간화 <i>Tmprss4</i> 게놈 단편	길이: 20,078 bp 1~18: 마우스 서열 19~5014: SalI/XhoI-LoxP- <i>hUbi</i> - EM7-Neo-Pm- <i>Cre</i> -loxP- <i>ICeuI</i> -NheI (총 4996 bp) 5015~19977: 인간 서열 (총 14,963 bp) 19978~20078: 마우스 서열
13	카세트가 결실된 인간화 <i>Tmprss4</i> 게놈 단편	길이: 15159 bp 1~18: 마우스 서열 19~95: SalI/XhoI-LoxP- <i>ICeuI</i> -NheI (총 77 bp) 96~15058: 인간 서열 (총 14,963 bp) 15059~15159: 마우스 서열
14	인간화 <i>Tmprss4</i> 단백질	길이: 435 aa
15	생쥐 <i>Tmprss11d</i> , mRNA, NM_145561.2	길이: 2046 bp CDS: 36~1289 엑손: 1~43 (제 1 엑손 및 제 1 코딩 엑손), 44~165, 166~284; 285~352; 353~507; 508~546; 547~724; 725~984; 985~1127; 1128~2046.
16	생쥐 <i>Tmprss11d</i> , 단백질	길이: 417 aa
17	인류 <i>TMPRSS11D</i> , mRNA, NM_004262.2	길이: 2800 bp CDS: 66~1322 엑손: 1~73 (제 1 엑손 및 제 1 코딩 엑손); 74~195; 196~314; 315~382; 383~540; 541~579; 580~757; 758~1017; 1018~1160; 1161~2783.
18	인류 <i>TMPRSS11D</i> , 단백질	길이: 418 aa 세포외 도메인: A42 에서 시작함.

[0047]

19	카세트를 함유하는 인간화 <i>Tmprss11d</i> 게놈 단편	길이: 38,992 1~19: 마우스 서열 20~33,946: 인간 서열 (총 33,927 bp) 33,947~38,942: XhoI-LoxP- <i>hUbi</i> -EM7-Neo-Pm- <i>Cre</i> -loxP- <i>ICeuI</i> -NheI (총 4,996 bp) 38,943~38,992: 마우스 서열
20	카세트가 결실된 인간화 <i>Tmprss11d</i> 게놈 단편	길이: 34,073 bp 1~19: 마우스 서열 20~33,946: 인간 서열 (총 33,927 bp) 33,947~34,023: XhoI-LoxP- <i>ICeuI</i> -NheI (77 bp) 34,024~34,073: 마우스 서열
21	인간화 <i>Tmprss11d</i> 단백질	418 aa
22	<i>Tmprss2</i> 인간화를 위한 5' 마우스/5' 인간 접합 서열	5' 마우스//5' 인간
23	<i>Tmprss2</i> 인간화를 위한 3' 마우스/카세트 접합 서열	인간//XhoI//loxP 카세트
24	<i>Tmprss2</i> 인간화를 위한 카세트/3' 마우스 접합 서열	카세트 (loxP)/ICEUI/NheI//마우스
25	<i>Tmprss2</i> 인간화를 위한 3' 인간/loxP/3' 마우스 접합	3' 인간//XhoI/(loxP)/ICEUI/NheI//3' 마우스
26-37	<i>Tmprss2</i> 인간화를 위한 대립유전자 상실 및 대립유전자 획득 검증을 위한 프라이머 및 프로브	표 1
38	<i>Tmprss4</i> 인간화를 위한 5' 마우스/카세트 접합 서열	5' 마우스//SalI-XhoI/(loxP) 카세트
39	<i>Tmprss4</i> 인간화를 위한 카세트/5' 인간 접합 서열	카세트 (loxP)/ICEUI/NheI//5' 인간

[0048]

40	<i>Tmprss4</i> 인간화를 위한 3' 인간/3' 마우스 접합 서열	3' 인간/3' 마우스
41	<i>Tmprss4</i> 인간화를 위한 5' 마우스/loxP/5' 인간 접합	5' 마우스//SalI/XhoI//(<i>loxP</i>) /ICEUI//NheI//5' 인간
42-56	<i>Tmprss4</i> 인간화를 위한 대립유전자 상실 및 대립유전자 획득 검증을 위한 프라이머 및 프로브	표 2
57	<i>Tmprss11d</i> 인간화를 위한 5' 마우스/5' 인간 접합 서열	5' 마우스//5' 인간
58	<i>Tmprss11d</i> 인간화를 위한 3' 인간/카세트 접합 서열	3' 인간//XhoI//(<i>loxP</i>) 카세트
59	<i>Tmprss11d</i> 인간화를 위한 카세트/3' 마우스 접합 서열	카세트 (<i>loxP</i>)/ICEUI//NheI//3' 마우스
60	<i>Tmprss11d</i> 인간화를 위한 3' 인간/loxP/3' 마우스 접합	3' 인간//XhoI//(<i>loxP</i>)/ICEUI//NheI//3' 마우스
61-72	<i>Tmprss11d</i> 인간화를 위한 대립유전자 상실 및 대립유전자 획득 검증을 위한 프라이머 및 프로브	표 3

[0049]

[0050]

인간화 *Tmprss* 설치류 동물

[0051]

일 양태에서, 본 발명은 인간화 *Tmprss* 단백질을 암호화하는 인간화 *Tmprss* 유전자를 생식계열에 함유하는 설치류 동물을 제공한다.

[0052]

"인간화"라는 용어가 핵산 또는 단백질의 맥락에서 사용될 때, 이는 자연에서 설치류 동물에서 발견되는 특정한 유전자 또는 단백질의 구조와 실질적으로 또는 동일하게 상응하는 부분을 포함하는 구조(즉, 뉴클레오티드 또는 아미노산 서열)를 가진 핵산 또는 단백질을 포함하고, 관련된 설치류 유전자 또는 단백질에서 발견되는 것과 상이한 대신에 상응하는 인간 유전자 또는 단백질에서 발견되는 구조와 밀접하게 또는 동일하게 상응하는 부분도 포함한다. 인간화 유전자를 함유하거나 인간화 단백질을 발현하는 설치류는 "인간화" 설치류이다.

[0053]

일부 구현예에서, 본 발명의 설치류는 마우스, 랫트, 및 햄스터로부터 선택된다. 일부 구현예에서, 본 발명의 설치류는 쥐상목 상과(superfamily Muroidea)로부터 선택된다. 일부 구현예에서, 본 발명의 유전자 변형된 설치류는 칼로미스쿠스과(Calomyscidae)(예를 들어, 마우스 유사 햄스터), 비단털쥐과(Cricetidae)(예를 들어, 햄스터, 미국 랫트 및 마우스, 들쥐), 쥐과(Muridae)(트루 마우스 및 랫트, 게르빌루스쥐, 아프리카가시쥐, 크레스티드 랫트), 네소미스과(Nesomyidae)(클라이밍 마우스, 락 마우스(rock mice), 흰꼬리 랫트, 말라가시 랫트 및 마우스), 가시겨울잠쥐과(Platacanthomyidae)(예를 들어, 가시겨울잠쥐), 및 소경쥐과(Spalacidae)(예를 들어, 두더쥐, 대나무쥐, 및 동북)로부터 선택되는 과로부터 유래한다. 일부 특정 구현예에서, 본 발명의 유전자 변형된 설치류는 트루 마우스 또는 랫트(패밀리 쥐과), 게르빌루스쥐, 아프리카가시쥐, 크레스티드 랫트로부터 선택된다. 일부 특정 구현예에서, 본 발명의 유전자 변형된 마우스는 쥐과(family Muridae)의 구성원으로부터 선택된다.

[0054]

일부 구현예에서, 본원에 기술된 설치류는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 및 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 포함하는 게놈 내에 인간화 *Tmprss* 유전자를 함유하되, 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS* 유전자의 뉴클레오티드 서열은 인간화 *Tmprss* 유전자가 *Tmprss* 단백질을 암호화하고, 프로모터 및/또는 증강자(들)와 같은 내인성 설치류 *Tmprss* 유전자의 5' 조절 요소(들)에 의해 조절되도록 서로 작동 가능하게 연결된다.

- [0055] 본 발명은 특히 대등한 (like-for-like) 인간화에 관한 것인데, 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열이 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열에 작동 가능하게 연결되어 인간화 유전자를 형성한다. 예를 들어, 일부 구현예에서, 내인성 설치류 *Tmprss2* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열에 작동 가능하게 연결되어 인간화 *Tmprss2* 유전자를 형성한다. 다른 구현예에서, 내인성 설치류 *Tmprss4* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS4* 유전자의 뉴클레오티드 서열에 작동 가능하게 연결되어 인간화 *Tmprss4* 유전자를 형성한다. 또 다른 구현예에서, 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열에 작동 가능하게 연결되어 인간화 *Tmprss11d* 유전자를 형성한다.
- [0056] 일부 구현예에서, 본 발명의 유전자 변형된 설치류는 그의 게놈에 인간화 *Tmprss* 유전자를 함유하되, 인간화 *Tmprss* 유전자는 인간 *TMPRSS* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 함유하는 인간화 *Tmprss* 단백질을 암호화한다. 용어 "세포외 도메인(ectodomain)"은 세포막의 외부로 연장되는 막관통 단백질의 부분, 즉 막관통 단백질의 세포외 부분을 지칭한다. *TMPRSS* 분자의 세포외 도메인은 프로테아제 도메인 및 막관통 도메인을 프로테아제 도메인과 연결시키는 줄기 영역을 포함한다. 세포외 도메인 또는 "인간 *TMPRSS* 단백질의 세포외 도메인과 실질적으로 동일한" 폴리펩티드라는 말은, 일부 구현예에서 인간 *TMPRSS* 단백질의 세포외 도메인과 서열이 적어도 85%, 90%, 95%, 95%, 99% 또는 100% 동일한 폴리펩티드; 일부 구현예에서, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 이하의 아미노산(들)만큼 인간 *TMPRSS* 단백질의 세포외 도메인과 상이한 폴리펩티드; 일부 구현예에서, 예를 들어 세포외 도메인의 N 말단 또는 C 말단에서 아미노산을 결여하거나 추가적인 아미노산을 가짐으로써 세포외 도메인의 N 말단 또는 C 말단에서 인간 *TMPRSS* 단백질의 세포외 도메인과 상이한 폴리펩티드; 및 일부 구현예에서, 인간 *TMPRSS* 단백질의 실질적으로 세포외 도메인인 폴리펩티드를 의미한다. 인간 *TMPRSS* 단백질의 "실질적으로 세포외 도메인"이라는 말은, 세포외 도메인과 동일하거나, N 말단 또는 C 말단에서 1~5개(즉, 1, 2, 3, 4, 또는 5개)의 아미노산의 결여하거나 추가적인 1~5개의 아미노산을 가짐으로써 세포외 도메인과 상이한 폴리펩티드를 의미한다.
- [0057] 일부 구현예에서, 인간화 *Tmprss* 유전자는 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 추가로 함유하는 인간화 *Tmprss* 단백질을 암호화한다. 세포질 및 막관통 부분 또는 "내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한" 폴리펩티드라는 말은, 일부 구현예에서, 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 서열이 적어도 85%, 90%, 95%, 95%, 99% 또는 100% 동일한 폴리펩티드; 일부 구현예에서, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 이하의 아미노산(들)만큼 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 상이한 폴리펩티드; 일부 구현예에서, 막관통 도메인의 C 말단에서 아미노산을 결여하거나 추가적인 아미노산을 가짐으로써 C 말단에서만 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 상이한 폴리펩티드; 및 일부 구현예에서, 내인성 설치류 *Tmprss* 단백질의 세포질 도메인 및 실질적으로 막관통 도메인으로 이루어진 폴리펩티드를 의미한다. 내인성 설치류 *Tmprss* 단백질의 "실질적으로 막관통 도메인"이라는 말은, 막관통 도메인과 동일하거나, C 말단에서 1~5개의 아미노산을 결여하거나 추가적인 1~5개의 아미노산을 가짐으로써 막관통 도메인과 상이한 폴리펩티드를 의미한다.
- [0058] 일부 구현예에서, 유전자 변형된 설치류에서의 인간화 *Tmprss* 유전자는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열 및 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 포함하되, 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS* 유전자에 의해 암호화된 인간 *TMPRSS* 단백질의 세포외 도메인과 실질적으로 동일한 폴리펩티드를 암호화한다. 특정 구현예에서, 인간화 *Tmprss* 유전자에서의 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열은 인간 *TMPRSS* 유전자에 의해 암호화된 인간 *TMPRSS* 단백질의 세포외 도메인을 암호화한다.
- [0059] 일부 구현예에서, 유전자 변형된 설치류의 게놈 내의 인간화 *Tmprss* 유전자는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열 및 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 포함하되, 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열은 설치류 *Tmprss* 유전자에 의해 암호화된 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 폴리펩티드를 암호화한다. 특정 구현예에서, 인간화 *Tmprss* 유전자 내에 존재하는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss* 유전자에 의해 암호화된 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 도메인을 암호화한다.
- [0060] 일부 구현예에서, 인간화 *Tmprss* 유전자는 내인성 설치류 *Tmprss* 유전자좌에서 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열을 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열과 대체함으로써 생성된다.
- [0061] 일부 구현예에서, 내인성 설치류 *Tmprss* 유전자좌에서의 설치류 *Tmprss* 유전자의 인접 게놈 서열은 인간화 *Tmprss* 유전자를 형성하기 위해 동족 인간 *TMPRSS* 유전자의 인접 게놈 서열과 대체되었다.
- [0062] 특정 구현예에서, 내인성 설치류 *Tmprss* 유전자에 삽입된 인간 *TMPRSS* 유전자의 인접 게놈 서열은 인간 *TMPRSS*

유전자에 의해 암호화된 인간 TMPRSS 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화하는 인간 TMPRSS 유전자의 엑손을 전체적으로 또는 부분적으로 포함한다.

[0063] 특정 구현예에서, 인간화 대체 이후 내인성 설치류 *Tmprss* 유전자좌에 남아있고 삽입된 인접 인간 TMPRSS 게놈 서열에 작동 가능하게 연결된 내인성 설치류 *Tmprss* 유전자의 게놈 서열은 내인성 설치류 *Tmprss* 유전자에 의해 암호화된 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 암호화한다.

[0064] 내인성 *Tmprss* 단백질 및 인간 TMPRSS 단백질이 막관통 도메인과 세포외 도메인 사이의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 인간 TMPRSS 단백질의 세포외 도메인을 정교하게 암호화하는 인간 TMPRSS 게놈 서열을 삽입하는 것은 필요하지 않을 수 있다. 생성된 인간화 *Tmprss* 유전자에 의해 암호화된 인간화 *Tmprss* 단백질이 인간 TMPRSS 단백질의 세포외 도메인과 동일한 세포외 도메인 및 내인성 설치류 *Tmprss* 단백질의 막관통 도메인과 동일한 막관통 도메인을 포함하도록, 내인성 설치류 *Tmprss* 단백질의 세포질 도메인과 실질적으로 막관통 도메인을 암호화하는 내인성 설치류 *Tmprss* 유전자의 게놈 서열에 작동 가능하게 연결된 인간 TMPRSS 단백질의 실질적으로 세포외 도메인을 암호화하는 인간 TMPRSS 유전자의 약간 길거나 짧은 게놈 서열을 삽입하는 것이 가능하다.

[0065] 일부 구현예에서, 인간화 *Tmprss* 유전자에 포함된 인간 TMPRSS 유전자의 뉴클레오티드 서열은 인간 TMPRSS 유전자의 3' 미번역 영역("UTR")을 또한 포함한다. 특정 구현예에서, 인간화 *Tmprss* 유전자는 또한, 인간 TMPRSS 유전자의 3' UTR에 추가하여 인간 TMPRSS 유전자좌로부터 추가적인 인간 게놈 서열을 인간 TMPRSS 3' UTR 다음에 포함한다. 추가적인 인간 게놈 서열은 인간 TMPRSS 유전자의 3' UTR의 바로 하류의 인간 TMPRSS 유전자좌에서 발견되는 적어도 10~200 bp, 예컨대 50 bp, 75 bp, 100 bp, 125 bp, 150 bp, 175 bp, 200 bp 또는 그 이상으로 이루어질 수 있다. 다른 구현예에서, 인간화 *Tmprss* 유전자 내에 존재하는 인간 TMPRSS 유전자의 뉴클레오티드 서열은 인간 3' UTR을 포함하지 않고; 대신에, 내인성 설치류 *Tmprss* 유전자의 3' UTR이 포함되고 인간화 *Tmprss* 유전자의 정지 코돈의 바로 뒤에 위치한다. 예를 들어, 인간화 *Tmprss* 유전자는 내인성 설치류 *Tmprss* 단백질의 세포질 및 막관통 도메인을 암호화하는 엑손 서열을 함유하는 내인성 설치류 *Tmprss* 유전자의 뉴클레오티드 서열을 포함할 수 있는데, 인간 TMPRSS 단백질의 세포외 도메인 내지 정지 코돈을 암호화하는 엑손을 함유하는 인간 TMPRSS 유전자의 뉴클레오티드 서열이 이를 따르고, 내인성 설치류 *Tmprss* 유전자의 3' UTR은 정지 코돈의 바로 뒤에 위치한다.

[0066] 일부 구현예에서, 인간화 *Tmprss* 유전자는 설치류 내에서 암호화된 인간화 *Tmprss* 단백질을 발현시킨다. 일부 구현예에서, 인간화 *Tmprss* 단백질은 대조군 설치류(예: 인간화 *Tmprss* 유전자가 없는 설치류)에서의 대응 설치류 *Tmprss* 단백질과 필적하거나 실질적으로 동일한 패턴으로 발현된다. 일부 구현예에서, 인간화 *Tmprss* 단백질은 대조군 설치류(예: 인간화 *Tmprss* 유전자가 없는 설치류)에서의 대응 설치류 *Tmprss* 단백질과 필적하거나 실질적으로 동일한 레벨로 발현된다. 특정 구현예에서, 인간화 *Tmprss* 단백질은 세포 표면에서 발현되고 검출된다. 특정 구현예에서, 인간화 *Tmprss* 단백질 또는 가용성 형태(예: 제거된 세포외 도메인 형태)는 설치류의 혈청에서, 예컨대 대조군 설치류에서의 대응 설치류 *Tmprss* 단백질 또는 이의 가용성 형태에 필적하거나 실질적으로 동일한 레벨로 발현되고 검출된다. 인간화 설치류에서의 인간화 유전자 또는 단백질을 대조군 설치류의 내인성 설치류 유전자 또는 단백질과 비교하는 맥락에서, 용어 "필적하는(comparable)"은 비교되는 분자 또는 레벨이 서로 동일하지 않을 수 있지만, 관찰된 차이점 또는 유사성에 기초하여 결론이 합리적으로 도출될 수 있도록 이들 사이의 비교를 허용하기에 충분히 유사함을 의미하며; 발현 레벨을 지칭할 때의 용어 "실질적으로 동일한(substantially the same)"은 비교되는 레벨이 서로 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 또는 1%보다 더 크게 상이하지 않음을 의미한다.

[0067] 일부 구현예에서, 본 발명은 본원에 기술된 비인간 바와 같은 설치류 동물로부터 단리된 세포 또는 조직을 제공한다. 일부 구현예에서, 세포는 수상세포, 림프구(예컨대, B 또는 T 세포), 대식세포 및 단핵구로부터 선택된다. 일부 구현예에서, 조직은 지방, 방광, 뇌, 가슴, 골수, 눈, 심장, 장, 신장, 간, 폐, 림프절, 근육, 췌장, 혈장, 혈청, 피부, 비장, 위, 흉선, 고환, 난자, 및 이의 조합으로부터 선택된다.

[0068] 일부 구현예에서, 본 발명은 본원에 기술된 바와 같은 인간화 *Tmprss* 유전자를 함유하는 게놈을 가진 설치류 배아 줄기 세포를 제공한다. 일부 구현예에서, 설치류 배아 줄기 세포는 마우스 배아 줄기 세포이다. 일부 구현예에서, 설치류 배아 줄기 세포는 랫트 배아 줄기 세포이다. 인간화 *Tmprss* 유전자를 게놈에 함유하는 설치류 배아 줄기 세포는 아래에 후술하는 바와 같이 인간화 설치류를 만드는 데 사용될 수 있다.

[0069] 일부 구현예에서, 본원에서 제공된 설치류는 그의 게놈 내의 인간화 *Tmprss* 유전자에 대해 이형접합체이다. 다

른 구현예에서, 본원에서 제공된 설치류는 그의 게놈 내의 인간화 *Tmprss* 유전자에 대해 동형접합체이다.

- [0070] 특정 구현예에서, 설치류는 그의 게놈 내에 다수의, 즉 둘 이상의 인간화 *Tmprss* 유전자를 포함한다. 즉, 설치류 내의 둘 이상의 상이한 내인성 *Tmprss* 유전자좌가 동족 인간 *TMPRSS* 유전자의 뉴클레오티드 서열을 사용하여 인간화되었다. 예를 들어, 설치류가 다음으로부터 선택된 둘 이상의 유전자좌에서 인간화되었다: *Tmprss2*, *Tmprss4*, 및 *Tmprss11d*.
- [0071] (마우스와 같은) 예시적인 인간화 *Tmprss2* 설치류, (마우스와 같은) 인간화 *Tmprss4* 설치류, 및 (마우스와 같은) 인간화 *Tmprss11d* 설치류가 아래에 더 기술된다.
- [0072] 인간화 *Tmprss2* 설치류
- [0073] 일부 구현예에서, 본 발명은 내인성 설치류 *Tmprss2* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열을 포함하고, 프로모터 및/또는 증강자(들)과 같은 내인성 설치류 *Tmprss2* 유전자의 5' 조절 요소(들)에 의해 조절되는 인간화 *Tmprss2* 유전자를 함유하는 게놈을 가진 설치류를 제공한다. 예시적인 설치류는 마우스와 랫트를 포함한다.
- [0074] 일부 구현예에서, 인간화 *Tmprss2* 유전자는 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 함유하는 인간화 *Tmprss2* 단백질을 암호화한다.
- [0075] 특정 구현예에서, 인간 *TMPRSS2* 단백질은 서열 번호 4에 명시된 바와 같은 아미노산 서열과 적어도 85%, 90%, 95%, 98%, 99% 또는 100%의 동일성을 갖는 아미노산 서열을 갖는다.
- [0076] 일부 구현예에서, 인간화 *Tmprss2* 단백질은 인간 *TMPRSS2* 단백질의 C 말단 387 아미노산, 예를 들어, 인간 *TMPRSS2* 단백질의 아미노산 106 내지 492를 함유한다. 일부 구현예에서, 인간화 *Tmprss2* 단백질은 서열 번호 4의 W106 내지 G492로 이루어진 아미노산 서열과 실질적으로 동일한 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss2* 단백질은, 서열 번호 4의 W106 내지 G492로 이루어진 아미노산 서열과 적어도 85%, 90%, 95%, 98%, 99% 또는 100%의 동일성을 갖는 세포외 도메인; 서열 번호 4의 W106 내지 G492로 이루어진 아미노산 서열과 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 이하의 아미노산(들)만큼 상이한 세포외 도메인; 또는, 예를 들어, N 말단 또는 C 말단에서 1~5개의 아미노산을 결여하거나 추가적인 1~5개의 아미노산을 가짐으로써 서열 번호 4의 W106 내지 G492로 이루어진 아미노산 서열과 N 말단 또는 C 말단에서만 상이한 세포외 도메인을 함유한다.
- [0077] 일부 구현예에서, 인간화 *Tmprss2* 단백질은 내인성 설치류 *Tmprss2* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 추가로 함유한다. 일부 구현예에서, 인간화 *Tmprss2* 단백질은 내인성 설치류 *Tmprss2* 단백질의 막관통 도메인 및 세포질 도메인을 추가로 포함한다.
- [0078] 특정 구현예에서, 인간화 *Tmprss2* 단백질은 내인성 설치류 *Tmprss2* 단백질의 막관통 도메인과 세포질 도메인, 및 인간 *TMPRSS2* 단백질의 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss2* 유전자는 서열 번호 7에 명시된 바와 같은 아미노산 서열을 갖는 인간화 *Tmprss2* 단백질을 암호화한다.
- [0079] 일부 구현예에서, 인간화 *Tmprss2* 유전자는 내인성 설치류 *Tmprss2* 유전자좌에서 내인성 설치류 *Tmprss2* 유전자의 뉴클레오티드 서열을 인간 *TMPRSS2* 유전자의 뉴클레오티드 서열과 대체함으로써 생성된다.
- [0080] 일부 구현예에서, 내인성 설치류 *Tmprss2* 유전자좌에서의 설치류 *Tmprss2* 유전자의 인접 게놈 서열은 인간화 *Tmprss2* 유전자를 형성하기 위해 인간 *TMPRSS2* 유전자의 인접 게놈 서열과 대체되었다.
- [0081] 특정 구현예에서, 내인성 설치류 *Tmprss2* 유전자에 삽입된 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자에 의해 암호화된 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화하는 인간 *TMPRSS2* 유전자의 엑손 서열, 즉 엑손을 전체적으로 또는 부분적으로 포함한다. 내인성 *Tmprss2* 단백질 및 인간 *TMPRSS2* 단백질이 막관통 도메인 및 세포외 도메인의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 인간 *TMPRSS2* 단백질의 세포외 도메인을 정교하게 암호화하는 인간 *TMPRSS2* 게놈 서열을 삽입하는 것은 필요하지 않을 수 있고, 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 갖는 인간화 *Tmprss2* 단백질을 만들기 위해 인간 *TMPRSS2* 단백질의 실질적으로 세포외 도메인을 암호화하는 약간 길거나 짧은 인간 *TMPRSS2* 게놈 서열을 사용하는 것이 가능하다.
- [0082] 특정 구현예에서, 내인성 설치류 *Tmprss2* 유전자에 삽입되는 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자의 적어도 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유한다.

- [0083] 특정 구현예에서, 내인성 설치류 *Imprss2* 유전자에 삽입되는 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자의 인트론 3 및 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유한다. 특정 구현예에서, 내인성 설치류 *Imprss2* 유전자에 삽입되는 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자의 코딩 엑손 3의 3' 부분, 인트론 3, 및 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유한다. 특정 구현예에서, 인간화에 포함된 인간 *TMPRSS2* 유전자의 코딩 엑손 3의 3' 부분은 5~10 염기쌍의 길이, 즉 코딩 엑손 3의 3' 말단의 약 5, 6, 7, 8, 9 또는 10 염기쌍의 길이이다.
- [0084] 일부 구현예에서, 내인성 설치류 *Imprss2* 유전자에 삽입되는 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자의 3' UTR을 또한 함유한다. 특정 구현예에서, 인간화를 위해 인간 *TMPRSS2* 유전자의 전체 코딩 엑손 13은 인간 *TMPRSS2* 유전자의 3' UTR을 포함하는 인접 인간 *TMPRSS2* 게놈 서열에 포함된다. 특정 구현예에서, 인간 *TMPRSS2* 유전자의 인접 게놈 서열은 인간 *TMPRSS2* 유전자의 3' UTR의 하류에서 추가적인 인간 게놈 서열을 포함한다. 추가적인 인간 게놈 서열은 인간 *TMPRSS2* 유전자좌에서 인간 *TMPRSS2* 유전자의 3' UTR의 바로 하류에서 발견되는 적어도 10~200 bp, 또는 적어도 10, 20, 30, 40, 50, 75, 100, 125, 150, 175, 또는 200 bp의 서열일 수 있다.
- [0085] 일부 구현예에서, 인간화 *Imprss2* 유전자좌에 남아 있는 내인성 설치류 *Imprss2* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Imprss2* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 암호화한다. 내인성 *Imprss2* 단백질 및 인간 *TMPRSS2* 단백질이 막관통 도메인 및 세포외 도메인의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 내인성 설치류 *Imprss2* 단백질의 막관통 도메인을 정교하게 암호화하는 내인성 설치류 *Imprss2* 게놈 서열을 유지시키는 것은 필요하지 않을 수 있고, 내인성 설치류 *Imprss2* 단백질과 동일한 막관통 도메인을 갖는 인간화 *Imprss2* 단백질을 암호화하기 위해 인간화 대체에서 내인성 설치류 *Imprss2* 단백질의 실질적으로 막관통 도메인을 암호화하는 약간 길거나 짧은 설치류 *Imprss2* 게놈 서열을 유지시키는 것이 가능하다. 일부 구현예에서, 인간화 *Imprss2* 유전자좌에 남아 있는 내인성 설치류 *Imprss2* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Imprss2* 유전자의 엑손 1~2 및 코딩 엑손 3의 5' 부분을 포함하되, 코딩 엑손 3의 5' 부분은 코돈 엑손 3의 실질적인 부분, 즉, 코딩 엑손 3의 3' 말단에서 5~10 염기쌍을 제외한 전체 코딩 엑손 3이다.
- [0086] 특정 구현예에서, 인간화 *Imprss2* 유전자는 내인성 설치류 *Imprss2* 유전자의 코딩 엑손 1 내지 2 및 코딩 엑손 3의 5' 부분, 및 인간 *TMPRSS2* 유전자의 코딩 엑손 3의 3' 부분 및 코딩 엑손 4 내지 코딩 엑손 13을 포함하되, 인간화 *Imprss2* 유전자는 설치류 *Imprss2* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 함유하는 인간화 *Imprss2* 단백질, 및 인간 *TMPRSS2* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화한다. 특정 구현예에서, 인간화 *Imprss2* 유전자는 내인성 설치류 *Imprss2* 유전자에 의해 암호화된 설치류 *Imprss2* 단백질의 세포질 도메인 및 막관통 도메인, 및 인간 *TMPRSS2* 유전자에 의해 암호화된 인간 *TMPRSS2* 단백질의 세포외 도메인을 함유하는 인간화 *Imprss2* 단백질을 암호화한다. 특정 구현예에서, 인간화 *Imprss2* 유전자는 서열 번호 7에 명시된 바와 같은 아미노산 서열을 갖는 인간화 *Imprss2* 단백질을 암호화한다.
- [0087] 일부 구현예에서, 인간화에 사용된 인간 *TMPRSS2* 유전자 및 설치류 *Imprss2* 유전자의 엑손과 인트론은 서열 번호 1, 3 및 5 내지 6에서 발견되는 것들이다.
- [0088] 일부 구현예에서, 인간화 *Imprss2* 유전자는 설치류에서 암호화된 인간화 *Imprss2* 단백질의 발현을 초래한다. 일부 구현예에서, 인간화 *Imprss2* 단백질은 대조군 설치류(예: 인간화 *Imprss2* 유전자가 없는 설치류)에서의 대응 설치류 *Imprss2* 단백질과 필적하거나 실질적으로 동일한 패턴으로 발현된다. 일부 구현예에서, 인간화 *Imprss2* 단백질은 대조군 설치류(예: 인간화 *Imprss2* 유전자가 없는 설치류)에서의 대응 설치류 *Imprss2* 단백질과 필적하거나 실질적으로 동일한 레벨로 발현된다. 특정 구현예에서, 인간화 *Imprss2* 단백질은 세포 표면에서 발현되고 검출된다. 특정 구현예에서, 인간화 *Imprss2* 단백질 또는 가용성 형태(예: 제거된 세포외 도메인 형태)는 설치류의 혈청에서, 예컨대 대조군 설치류에서의 대응 설치류 *Imprss2* 단백질 또는 이의 가용성 형태에 필적하거나 실질적으로 동일한 레벨로 발현되고 검출된다.
- [0089] 인간화 *Imprss4* 설치류
- [0090] 일부 구현예에서, 본 발명은 내인성 설치류 *Imprss4* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS4* 유전자의 뉴클레오티드 서열을 포함하고, 프로모터 및/또는 증강자(들)과 같은 내인성 설치류 *Imprss4* 유전자의 5' 조절 요소(들)에 의해 조절되는 인간화 *Imprss4* 유전자를 함유하는 게놈을 가진 설치류를 제공한다. 예시적인 설치류는

마우스와 랫트를 포함한다.

- [0091] 일부 구현예에서, 인간화 *Tmprss4* 유전자는 인간 TMPRSS4 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 함유하는 인간화 *Tmprss4* 단백질을 암호화한다. 특정 구현예에서, 인간 TMPRSS4 단백질은 서열 번호 11에 명시된 바와 같은 아미노산 서열과 적어도 85%, 90%, 95%, 98%, 99% 또는 100%의 동일성을 갖는 아미노산 서열을 갖는다.
- [0092] 일부 구현예에서, 인간화 *Tmprss4* 단백질은 인간 TMPRSS4 단백질의 C 말단 384 아미노산, 예를 들어, 인간 TMPRSS4 단백질의 아미노산 54 내지 437을 함유한다. 일부 구현예에서, 인간화 *Tmprss4* 단백질은 서열 번호 11의 K54 내지 L437로 이루어진 아미노산 서열과 실질적으로 동일한 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss4* 단백질은, 서열 번호 11의 K54 내지 L437로 이루어진 아미노산 서열과 적어도 85%, 90%, 95%, 98%, 99% 또는 100%의 동일성을 갖는 세포외 도메인; 서열 번호 11의 K54 내지 L437로 이루어진 아미노산 서열과 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 이하의 아미노산(들)만큼 상이한 세포외 도메인; 또는, 예를 들어, N 말단 또는 C 말단에서 1~5개의 아미노산을 결여하거나 추가적인 1~5개의 아미노산을 가짐으로써 서열 번호 11의 K54 내지 L437로 이루어진 아미노산 서열과 N 말단 또는 C 말단에서만 상이한 세포외 도메인을 함유한다.
- [0093] 일부 구현예에서, 인간화 *Tmprss4* 단백질은 내인성 설치류 *Tmprss4* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 추가로 함유한다. 일부 구현예에서, 인간화 *Tmprss4* 단백질은 내인성 설치류 *Tmprss4* 단백질의 막관통 도메인 및 세포질 도메인을 추가로 포함한다.
- [0094] 특정 구현예에서, 인간화 *Tmprss4* 단백질은 내인성 설치류 *Tmprss4* 단백질의 막관통 도메인과 세포질 도메인, 및 인간 TMPRSS4 단백질의 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss4* 유전자는 서열 번호 14에 명시된 바와 같은 아미노산 서열을 갖는 인간화 *Tmprss4* 단백질을 암호화한다.
- [0095] 일부 구현예에서, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자좌에서 내인성 설치류 *Tmprss4* 유전자의 뉴클레오티드 서열을 인간 TMPRSS4 유전자의 뉴클레오티드 서열과 대체함으로써 생성된다.
- [0096] 일부 구현예에서, 내인성 설치류 *Tmprss4* 유전자좌에서의 설치류 *Tmprss4* 유전자의 인접 게놈 서열은 인간화 *Tmprss4* 유전자를 형성하기 위해 인간 TMPRSS4 유전자의 인접 게놈 서열과 대체되었다.
- [0097] 특정 구현예에서, 내인성 설치류 *Tmprss4* 유전자에 삽입된 인간 TMPRSS4 유전자의 인접 게놈 서열은 인간 TMPRSS4 유전자에 의해 암호화된 인간 TMPRSS4 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화하는 인간 TMPRSS4 유전자의 엑손 서열, 즉 엑손을 전체적으로 또는 부분적으로 포함한다. 내인성 *Tmprss4* 단백질 및 인간 TMPRSS4 단백질이 막관통 도메인 및 세포외 도메인의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 인간 TMPRSS4 단백질의 세포외 도메인을 정교하게 암호화하는 인간 TMPRSS4 게놈 서열을 삽입하는 것은 필요하지 않을 수 있고, 인간 TMPRSS4 단백질의 세포외 도메인과 동일한 세포외 도메인을 갖는 인간화 *Tmprss4* 단백질을 만들기 위해 인간 TMPRSS4 단백질의 실질적으로 세포외 도메인을 암호화하는 약간 길거나 짧은 인간 TMPRSS4 게놈 서열을 사용하는 것이 가능하다.
- [0098] 특정 구현예에서, 내인성 설치류 *Tmprss4* 유전자에 삽입되는 인간 TMPRSS4 유전자의 인접 게놈 서열은 인간 TMPRSS4 유전자의 적어도 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유한다.
- [0099] 특정 구현예에서, 내인성 설치류 *Tmprss4* 유전자에 삽입되는 인간 TMPRSS4 유전자의 인접 게놈 서열은 인간 TMPRSS4 유전자의 인트론 3의 3' 부분, 및 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유한다. 특정 구현예에서, 인간화에 포함된 인간 TMPRSS4 유전자의 코딩 엑손 3의 3' 부분은 140~160 염기쌍의 길이, 즉 인트론 3의 3' 말단의 약 140, 145, 150, 155, 160 염기쌍의 길이이다.
- [0100] 일부 구현예에서, 내인성 설치류 *Tmprss4* 유전자에 삽입되는 인간 TMPRSS4 유전자의 인접 게놈 서열은 인간 TMPRSS4 유전자의 3' UTR을 함유한다. 특정 구현예에서, 내인성 설치류 *Tmprss4* 유전자에 삽입되는 인간 TMPRSS4 유전자의 인접 게놈 서열은 인간 TMPRSS4 유전자의 3' UTR을 포함하지 않고, 내인성 설치류 *Tmprss4* 유전자의 3' UTR은 인간화 *Tmprss4* 유전자 내에서 정지 코돈의 바로 뒤에 위치한다.
- [0101] 일부 구현예에서, 인간화 *Tmprss4* 유전자좌에 남아 있는 내인성 설치류 *Tmprss4* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss4* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 암호화한다. 내인성 *Tmprss4* 단백질 및 인간 TMPRSS4 단백질이 막관통 도메인 및 세포외 도메인의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 내인성 설치류 *Tmprss4* 단백질의 막관통 도메인을 정교하게 암호화하는 내인성 설치류 *Tmprss4* 게놈 서열을 유지시키는 것은 필요하지 않을 수 있고, 내인성 설치류 *Tmprss4* 단백질과 동일

한 막관통 도메인을 갖는 인간화 *Tmprss4* 단백질을 암호화하기 위해 인간화 대체에서 내인성 설치류 *Tmprss4* 단백질의 실질적으로 막관통 도메인을 암호화하는 약간 길거나 짧은 설치류 *Tmprss4* 게놈 서열을 유지시키는 것이 가능하다.

[0102] 특정 구현예에서, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자의 코딩 엑손 1 내지 3 및 인간 *TMPRSS4* 유전자의 코딩 엑손 4 내지 코딩 엑손 13의 정지 코돈을 함유한다. 특정 구현예에서, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자의 코딩 엑손 1 내지 3 및 인트론 3의 5' 부분, 및 인간 *TMPRSS4* 유전자의 인트론 3의 3' 부분 및 코딩 엑손 4 내지 코딩 엑손 13의 정지 코돈을 함유한다. 특정 구현예에서, 인간화 *Tmprss4* 유전자는 내인성 설치류 *Tmprss4* 유전자에 의해 암호화된 설치류 *Tmprss4* 단백질의 세포질 도메인 및 막관통 도메인, 및 인간 *TMPRSS4* 유전자에 의해 암호화된 인간 *TMPRSS4* 단백질의 세포외 도메인을 함유하는 인간화 *Tmprss4* 단백질을 암호화한다. 특정 구현예에서, 인간화 *Tmprss4* 유전자는 서열 번호 14에 명시된 바와 같은 아미노산 서열을 갖는 인간화 *Tmprss4* 단백질을 암호화한다.

[0103] 일부 구현예에서, 인간화에 사용된 인간 *TMPRSS4* 유전자 및 설치류 *Tmprss4* 유전자의 엑손과 인트론은 서열 번호 8, 10 및 12 내지 13에서 발견되는 것들이다.

[0104] 일부 구현예에서, 인간화 *Tmprss4* 유전자는 설치류에서 암호화된 인간화 *Tmprss4* 단백질의 발현을 초래한다. 일부 구현예에서, 인간화 *Tmprss4* 단백질은 대조군 설치류(예: 인간화 *Tmprss4* 단백질을 암호화하는 인간화 *Tmprss4* 유전자가 없는 설치류)에서의 대응 설치류 *Tmprss4* 단백질과 필적하거나 실질적으로 동일한 패턴으로 발현된다. 일부 구현예에서, 인간화 *Tmprss4* 단백질은 대조군 설치류(예: 인간화 *Tmprss4* 단백질을 암호화하는 인간화 *Tmprss4* 유전자가 없는 설치류)에서의 대응 설치류 *Tmprss4* 단백질과 필적하거나 실질적으로 동일한 레벨로 발현된다. 특정 구현예에서, 인간화 *Tmprss4* 단백질은 세포 표면에서 발현되고 검출된다. 특정 구현예에서, 인간화 *Tmprss4* 단백질 또는 가용성 형태(예: 제거된 세포외 도메인 형태)는 설치류의 혈청에서, 예컨대 대조군 설치류에서의 대응 설치류 *Tmprss4* 단백질 또는 이의 가용성 형태에 필적하거나 실질적으로 동일한 레벨로 발현되고 검출된다.

[0105] 인간화 *Tmprss11d* 설치류

[0106] 일부 구현예에서, 본 발명은 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열 및 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열을 포함하고, 프로모터 및/또는 증강자(들)과 같은 내인성 설치류 *Tmprss11d* 유전자의 5' 조절 요소(들)에 의해 조절되는 인간화 *Tmprss11d* 유전자를 함유하는 게놈을 가진 설치류를 제공한다. 예시적인 설치류는 마우스와 랫트를 포함한다.

[0107] 일부 구현예에서, 인간화 *Tmprss11d* 유전자는 인간 *TMPRSS11D* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 함유하는 인간화 *Tmprss11d* 단백질을 암호화한다.

[0108] 특정 구현예에서, 인간 *TMPRSS11D* 단백질은 서열 번호 18에 명시된 바와 같은 아미노산 서열과 적어도 85%, 90%, 95%, 98%, 99% 또는 100%의 동일성을 갖는 아미노산 서열을 갖는다.

[0109] 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 인간 *TMPRSS11D* 단백질의 C 말단 377 아미노산, 예를 들어, 인간 *TMPRSS11D* 단백질의 아미노산 42 내지 418을 함유한다. 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 서열 번호 18의 A42 내지 I418로 이루어진 아미노산 서열과 실질적으로 동일한 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss11d* 단백질은, 서열 번호 18의 A42 내지 I418로 이루어진 아미노산 서열과 적어도 85%, 90%, 95%, 98%, 99% 또는 100%의 동일성을 갖는 세포외 도메인; 서열 번호 18의 A42 내지 I418로 이루어진 아미노산 서열과 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개 이하의 아미노산(들)만큼 상이한 세포외 도메인; 또는, 예를 들어, N 말단 또는 C 말단에서 1~5개의 아미노산을 결여하거나 추가적인 1~5개의 아미노산을 가짐으로써 서열 번호 18의 A42 내지 I418로 이루어진 아미노산 서열과 N 말단 또는 C 말단에서만 상이한 세포외 도메인을 함유한다.

[0110] 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 내인성 설치류 *Tmprss11d* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 추가로 함유한다. 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 내인성 설치류 *Tmprss11d* 단백질의 막관통 도메인 및 세포질 도메인을 포함한다.

[0111] 특정 구현예에서, 인간화 *Tmprss11d* 단백질은 내인성 설치류 *Tmprss11d* 단백질의 막관통 도메인과 세포질 도메인, 및 인간 *TMPRSS11D* 단백질의 세포외 도메인을 함유한다. 특정 구현예에서, 인간화 *Tmprss11d* 유전자는 서열 번호 21에 명시된 바와 같은 아미노산 서열을 갖는 인간화 *Tmprss11d* 단백질을 암호화한다.

- [0112] 일부 구현예에서, 인간화 *Tmprss11d* 유전자는 내인성 설치류 *Tmprss11d* 유전자좌에서 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열을 인간 *TMPRSS11D* 유전자의 뉴클레오티드 서열과 대체함으로써 생성된다.
- [0113] 일부 구현예에서, 내인성 설치류 *Tmprss11d* 유전자좌에서의 설치류 *Tmprss11d* 유전자의 인접 게놈 서열은 인간화 *Tmprss11d* 유전자를 형성하기 위해 인간 *TMPRSS11D* 유전자의 인접 게놈 서열과 대체되었다. 특정 구현예에서, 내인성 설치류 *Tmprss11d* 유전자에 삽입된 인간 *TMPRSS11D* 유전자의 인접 게놈 서열은 인간 *TMPRSS11D* 유전자에 의해 암호화된 인간 *TMPRSS11D* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 암호화하는 인간 *TMPRSS11D* 유전자의 엑손 서열, 즉 엑손을 전체적으로 또는 부분적으로 포함한다. 내인성 *Tmprss11d* 단백질 및 인간 *TMPRSS11D* 단백질이 막관통 도메인 및 세포외 도메인의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 인간 *TMPRSS11D* 단백질의 세포외 도메인을 정교하게 암호화하는 인간 *TMPRSS11D* 게놈 서열을 삽입하는 것은 필요하지 않을 수 있고, 인간 *TMPRSS11D* 단백질의 세포외 도메인과 실질적으로 동일한 세포외 도메인을 갖는 인간화 *Tmprss11d* 단백질을 만들기 위해 인간 *TMPRSS11D* 단백질의 실질적으로 세포외 도메인을 암호화하는 약간 길거나 짧은 인간 *TMPRSS11D* 게놈 서열을 사용하는 것이 가능하다.
- [0114] 특정 구현예에서, 내인성 설치류 *Tmprss11d* 유전자에 삽입되는 인간 *TMPRSS11D* 유전자의 인접 게놈 서열은 인간 *TMPRSS11D* 유전자의 적어도 코딩 엑손 3 내지 코딩 엑손 10 내의 정지 코돈을 함유한다.
- [0115] 특정 구현예에서, 내인성 설치류 *Tmprss11d* 유전자에 삽입되는 인간 *TMPRSS11D* 유전자의 인접 게놈 서열은 인간 *TMPRSS11D* 유전자의 인트론 2의 3' 부분 및 코딩 엑손 3 내지 코딩 엑손 10 내의 정지 코돈을 함유한다. 특정 구현예에서, 인간화에 포함된 인간 *TMPRSS2* 유전자의 인트론 2의 3' 부분은 약 444 염기쌍의 길이이다.
- [0116] 일부 구현예에서, 내인성 설치류 *Tmprss11d* 유전자에 삽입되는 인간 *TMPRSS11D* 유전자의 인접 게놈 서열은 인간 *TMPRSS11D* 유전자의 3' UTR을 함유한다. 특정 구현예에서, 인간화를 위해, 인간 *TMPRSS11D* 유전자의 전체 코딩 엑손 10은 인간 *TMPRSS11D* 유전자의 3' UTR을 포함하는 인접 인간 *TMPRSS11D* 게놈 서열에 포함된다. 특정 구현예에서, 인간 *TMPRSS11D* 유전자의 인접 게놈 서열은 인간 *TMPRSS11D* 유전자의 3' UTR의 하류에서 추가적인 인간 게놈 서열을 포함한다. 추가적인 인간 게놈 서열은 인간 *TMPRSS11D* 유전자좌에서 인간 *TMPRSS11D* 유전자의 3' UTR의 바로 하류에서 발견되는 10~200 bp, 50~200 bp, 또는 약 150, 160, 170, 180 bp의 서열일 수 있다.
- [0117] 일부 구현예에서, 인간화 *Tmprss11d* 유전자좌에 남아 있는 내인성 설치류 *Tmprss11d* 유전자의 뉴클레오티드 서열은 내인성 설치류 *Tmprss11d* 유전자에 의해 암호화된 내인성 설치류 *Tmprss11d* 단백질의 세포질 및 막관통 부분과 실질적으로 동일한 세포질 및 막관통 부분을 암호화한다. 내인성 *Tmprss11d* 단백질 및 인간 *TMPRSS11D* 단백질이 막관통 도메인 및 세포외 도메인의 접합부 근처에서 공통 아미노산을 공유하는 상황에서, 내인성 설치류 *Tmprss11d* 단백질의 막관통 도메인을 정교하게 암호화하는 내인성 설치류 *Tmprss11d* 게놈 서열을 유지시키는 것은 필요하지 않을 수 있고, 내인성 설치류 *Tmprss11d* 단백질과 동일한 막관통 도메인을 갖는 인간화 *Tmprss11d* 단백질을 암호화하기 위해 인간화 대체에서 내인성 설치류 *Tmprss11d* 단백질의 실질적으로 막관통 도메인을 암호화하는 약간 길거나 짧은 설치류 *Tmprss11d* 게놈 서열을 유지시키는 것이 가능하다.
- [0118] 특정 구현예에서, 인간화 *Tmprss11d* 유전자는 내인성 설치류 *Tmprss11d* 유전자의 코딩 엑손 1 내지 2 및 인간 *TMPRSS11D* 유전자의 코딩 엑손 3 내지 코딩 엑손 10을 함유한다. 특정 구현예에서, 인간화 *Tmprss11d* 유전자는 내인성 설치류 *Tmprss11d* 유전자에 의해 암호화된 설치류 *Tmprss11d* 단백질의 세포질 도메인 및 막관통 도메인, 및 인간 *TMPRSS11D* 유전자에 의해 암호화된 인간 *TMPRSS11D* 단백질의 세포외 도메인을 함유하는 인간화 *Tmprss11d* 단백질을 암호화한다. 특정 구현예에서, 인간화 *Tmprss11d* 유전자는 서열 번호 21에 명시된 바와 같은 아미노산 서열을 갖는 인간화 *Tmprss11d* 단백질을 암호화한다.
- [0119] 일부 구현예에서, 인간화에 사용된 인간 *TMPRSS11D* 유전자 및 설치류 *Tmprss11d* 유전자의 엑손과 인트론은 서열 번호 15, 17 및 19 내지 20에서 발견되는 것들이다.
- [0120] 일부 구현예에서, 인간화 *Tmprss11d* 유전자는 설치류 내에서 암호화된 인간화 *Tmprss11d* 단백질의 발현을 초래한다. 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 대조군 설치류(예: 인간화 *Tmprss11d* 단백질을 암호화하는 인간화 *Tmprss11d* 유전자가 없는 설치류)에서의 대응 설치류 *Tmprss11d* 단백질과 필적하거나 실질적으로 동일한 패턴으로 발현된다. 일부 구현예에서, 인간화 *Tmprss11d* 단백질은 대조군 설치류(예: 인간화 *Tmprss11d* 단백질을 암호화하는 인간화 *Tmprss11d* 유전자가 없는 설치류)에서의 대응 설치류 *Tmprss11d* 단백질과 필적하거나 실질적으로 동일한 레벨로 발현된다. 특정 구현예에서, 인간화 *Tmprss11d* 단백질은 세포 표면에서 발현되고 검출된다. 특정 구현예에서, 인간화 *Tmprss11d* 단백질 또는 가용성 형태(예: 제거된 세포외 도메인 형태)는 설치류의 혈청에서, 예컨대 대조군 설치류에서의 대응 설치류 *Tmprss11d* 단백질 또는 이의 가용성 형태에 필적하거나

실질적으로 동일한 레벨로 발현되고 검출된다.

[0121] **인간화 *Tmprss* 설치류 동물을 만드는 방법**

[0122] 본 개시의 추가적인 양태는 전술한 인간화 *Tmprss* 설치류를 만드는 방법을 비롯하여 인간화 *Tmprss* 설치류를 만드는 데 사용하기에 적합한 핵산 벡터 및 비인간 배아 줄기 세포에 관한 것이다.

[0123] 본원에서 제공된 설치류는 당업계에 공지된 방법을 사용하여 만들 수 있다. 예시적인 구현예에서, 설치류 *Tmprss* 유전자를 보유하는 박테리아 인공 염색체(BAC) 클론은 박테리아 상동성 재조합 및 VELOCIGENE® 기술을 사용하여 변형될 수 있다(예: 미국 특허 제6,586,251호 및 Valenzuela 외의 (2003), High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, *Nature Biotech.* 21(6):652-659 참조). 결과적으로, 설치류 *Tmprss* 뉴클레오티드 서열이 원 BAC 클론으로부터 결실되었고, 인간 *Tmprss* 뉴클레오티드 서열이 삽입되어, 5' 및 3' 설치류 상동 아암이 측면에 위치한 인간화 *Tmprss* 유전자를 갖는 변형된 BAC 클론이 생성되었다. 변형된 BAC 클론은, 일단 선형화되면, 예컨대 전기 천공에 의해 설치류 배아 줄기(ES) 내에 도입될 수 있다. 마우스 ES 세포와 랫트 ES 세포 모두는 당업계에 기술되어 있다. 참조: 예를 들어, 미국 특허번호 제7,576,259호, 제7,659,442호, 제7,294,754호, 및 미국 특허 공개번호 제2008-0078000 A1호(이들 모두는 본원에 참조로서 통합됨)는 마우스 ES 세포 및 유전자 변형된 마우스를 만드는 VELOCIMOUSE® 방법을 기술하고 있고; 미국 특허 공개번호 제2014/0235933 A1호, 제2014/0310828 A1호, Tong 외의 (2010) *Nature* 467:211-215, 및 Tong 외의 (2011) *Nat Protoc.* 6(6): doi:10.1038/nprot.2011.338(이들 모두는 본원에 참조로서 통합됨)은 랫트 ES 세포 및 유전자 변형된 랫트를 만드는 방법을 기술하고 있다.

[0124] 게놈에 통합된 인간화 *Tmprss* 유전자를 갖는 ES 세포가 선택될 수 있다. 일부 구현예에서, 내인성 설치류 *Tmprss* 유전자좌에 통합된 인간화 *Tmprss*를 갖는 ES 세포는 설치류 대립유전자의 결실 및/또는 인간 대립유전자의 증가 분석에 기초하여 선택될 수 있다. 그런 다음, 선택된 ES 세포는 VELOCIMOUSE 방법(예: 미국 특허 번호 제7,576,259호, 제7,659,442호, 제7,294,754호, 및 미국 특허 공개번호 제2008-0078000 A1호 참조), 또는 미국 특허 공개번호 제2014/0235933 A1호 및 제2014/0310828 A1호에 기술된 방법을 사용함으로써 상실배 전시기(pre-morula stage)의 배아(예: 8 세포기 배아)에 주입하기 위한 공여자 ES 세포로서 사용된다. 공여자 ES 세포를 포함하는 배아는 배반포기까지 배양된 다음, 대리모에 이식되어 공여자 ES 세포로부터 완전히 유래된 F0 설치류를 생산한다. 인간화 *Tmprss* 유전자를 가진 설치류 새끼는 설치류 대립유전자의 상실 및/또는 인간 대립유전자의 증가 분석을 사용하여, 꼬리 단편으로부터 단리된 DNA의 유전자형 분석에 의해 확인될 수 있다.

[0125] 인간화 *Tmprss* 유전자에 대해 이형접합체인 설치류는 생성된 동형접합체 설치류와 교배시킬 수 있다. 하나의 인간화 *Tmprss* 유전자를 함유하는 설치류는 또 다른 인간화 *Tmprss* 유전자를 함유하는 설치류와 교배되어 다수의 인간화 *Tmprss* 유전자를 함유하는 설치류를 만들 수 있다. 예를 들어, 인간화 *Tmprss2* 유전자를 함유하는 설치류는 또 다른 인간화 *Tmprss4* 유전자를 함유하는 설치류와 교배되어 인간화 *Tmprss2* 유전자와 인간화 *Tmprss4* 유전자를 함유하는 설치류를 만들 수 있다.

[0126] **인간화 *Tmprss* 유전자를 갖는 설치류를 사용하는 방법**

[0127] 본원에 개시된 설치류는 인간 TMPRSS 단백질을 특이적으로 표적화하는 화합물의 식별 및 시험을 위한 인간화 *Tmprss* 단백질을 발현하는 생물학적 물질(예: 세포)의 유용한 생체 내 시스템 및 공급원을 제공한다.

[0128] 일 양태에서, 본원에 개시된 설치류는 인플루엔자 바이러스 감염을 치료 및/또는 예방하기 위한 인간 TMPRSS 단백질 억제제와 같은 후보 화합물의 능력을 알아내는 데 사용된다.

[0129] 일부 구현예에서, 본원에 기술된 인간화 *Tmprss* 유전자를 함유하고 인간화 *Tmprss* 단백질을 발현하는 설치류는 인플루엔자 바이러스 감염 실험에 앞서 후보 화합물과 함께 투여된다. 화합물의 예방적 효능은 설치류가 대조 설치류(들)와 비교하여 인플루엔자 바이러스 감염에 대한 더 적은 증상을 나타내고/내거나 덜 위중한 증상을 나타내고/내거나 개선된 생존력을 나타내는지 여부를 알아 냄으로써 평가할 수 있다.

[0130] 다른 구현예에서, 인간화 *Tmprss* 유전자를 함유하고 인간 TMPRSS 단백질의 세포의 도메인을 포함하는 인간화 *Tmprss* 단백질을 발현하는 설치류에게 실험적 인플루엔자 바이러스 감염 후 상기 인간 TMPRSS 단백질의 후보 억제제를 투여한다. 후보 억제제의 치료 효능은 설치류가 대조 설치류(들)와 비교하여 인플루엔자 바이러스 감염에 대한 더 적은 증상을 나타내고/내거나 덜 위중한 증상을 나타내고/내거나 개선된 생존력을 나타내는지 여부를 알아 냄으로써 평가할 수 있다.

[0131] 적합한 대조 설치류는, 예를 들어, 실험적 감염없이 인간화 *Tmprss* 유전자를 함유하는 설치류; 실험적 감염으로

인간화 *Tmprss* 유전자를 함유하고, 임의의 화합물이 없는 설치류; 및 실험적 감염으로 인간화 *Tmprss* 유전자를 함유하고, 치료 효과가 있는 것으로 알려진 화합물을 함유하는 설치류를 포함한다.

[0132] 본 발명의 방법에서 평가할 수 있는 화합물은, 후보 TMPRSS 억제제, 예를 들어, 소분자 프로테아제 억제제, 핵산계 억제제(예: siRNA, 리보자임, 안티센스 작제물 등), 항원 결합 단백질(예: 항체 또는 이의 항원 결합 단편), 또는 차단 펩티드/펩티드 억제제를 포함한다. TMPRSS 억제제는 헤마글루티닌(hemagglutinin) 전구체 단백질(HA0)을 단백질 분해에 의해 HA1 및 HA2 서브유닛으로 절단하는 TMPRSS 단백질의 능력을 억제 또는 감소시킴으로써 기능할 수 있다.

[0133] 일부 구현예에서, 후보 억제제는 항체 또는 이의 항원 결합 단편이다. 단클론 항체와 다클론 항체 모두가 본 발명의 목적에 적합하다. 특정 구현예에서, 항체는 TMPRSS 단백질에 특이적으로 결합하여 상기 TMPRSS 단백질의 프로테아제 활성을 억제하고, 다른 TMPRSS 단백질의 프로테아제 활성은 실질적으로 억제하지 않는다. 예를 들어, 항TMPRSS2 항체 억제제는 TMPRSS2 단백질에 특이적으로 결합하여 TMPRSS2 단백질의 프로테아제 활성을 억제하는데, TMPRSS4 또는 TMPRSS11D의 단백질 분해 활성에 대해서는 영향을 미치지 않거나, 동일하거나 실질적으로 동일한 실험 조건 하에서 시험된 비억제성 대조 분자에 비해 TMPRSS4 또는 TMPRSS11D의 단백질 분해 활성을 25% 이하(예: 20%, 15%, 10%, 5% 이하) 만큼 감소시킨다.

[0134] 일부 구현예에서, 억제제는 항TMPRSS2 항체 또는 이의 항원 결합 단편이다. 일부 구현예에서, 억제제는 항TMPRSS4 항체 또는 이의 항원 결합 단편이다. 일부 구현예에서, 억제제는 항TMPRSS11D 항체 또는 이의 항원 결합 단편이다.

[0135] 실험적 인플루엔자 바이러스 감염은 알려진 프로토콜을 따라 유도되고 모니터링될 수 있다. 예를 들어, 미국 특허 공개번호 제2013/0273070 A1호 참조. 예를 들어, 설치류 동물에게 인플루엔자 바이러스가 비강내 투여될 수 있다. 감염된 동물을 평가하여 감염의 증상과 중증도를 알아낼 수 있다. 예를 들어, 동물의 (1) 체중 변화 및 생존, (2) 유동 세포 계측법을 통한 세포 변화, (3) 면역 화학, 폐 전체의 PAS 및 H&E 염색, 및 (4) 혈청 내 사이토카인 레벨에 대해 분석할 수 있다. 바이러스에 걸리기 쉬운 것으로 알려진 대조 동물은 감염되지 않은 동물과 비교하여, 수지상 세포의 빈도, 인플루엔자 양성 폐포 대식세포, 폐에서의 호중구 세포 또는 상피 세포, 및 IFN γ 레벨에서 상당한 증가를 보였다.

[0136] 실시예

[0137] 다음의 실시예는 당업자에게 본 발명의 방법 및 조성물을 어떻게 제조하고 사용하는지를 설명하기 위해 제공되며, 발명자가 그들의 발명으로 간주한 것의 범위를 제한하고자 하는 것이 아니다. 다르게 명시되지 않는 한, 온도는 섭씨로 표시되고, 압력은 대기압이거나 대기압에 가깝다.

[0138] 실시예 1. 내인성 *Tmprss2* 유전자의 인간화

[0139] 본 실시예는 설치류(예를 들어, 마우스)에서 *Tmprss2*를 암호화하는 내인성 유전자를 인간화하는 예시적인 방법을 예시한다. 본 실시예에 기술된 방법은 임의의 인간 서열, 또는 원하는 인간 서열(또는 서열 단편)의 조합을 사용하여 설치류의 내인성 *Tmprss2* 유전자를 인간화하는데 사용될 수 있다.

[0140] 내인성 *Tmprss2* 유전자의 인간화를 위한 표적화 벡터를 박테리아 인공 염색체(BAC) 클론 및 VELOCIGENE® 기술을 사용하여 작제하였다(예를 들어, 본원에 참조로서 포함된, 미국 특허 번호 제6,586,251호 및 Valenzuela 외의 (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, *Nature Biotech.* 21(6):652-659 참조).

[0141] 요약하자면, 마우스 *Tmprss2* 유전자를 함유하는 마우스 박테리아 인공 염색체(BAC) 클론 bMQ-264A15를 사용하여 다음과 같이 변형시켰다. 5' 마우스 상동 뉴클레오티드 서열, (인간 *TMPRSS2* 유전자의 코딩 엑손 3의 마지막 7 bp, 인트론 3, 및 코딩 엑손 4 내지 코딩 엑손 13(코딩 엑손 13의 일부인 3' UTR을 포함함)을 함유하는) 약 25,091 bp의 인간 *TMPRSS2* 게놈 DNA, 약 2,691 bp의 자체 결실 네오마이신 카세트, 및 3' 마우스 상동 서열을 포함하도록 DNA 단편을 생성하였다. 이러한 DNA 단편을 사용하여 박테리아 세포에서 상동성 재조합을 통해 BAC 클론 bMQ-264A15를 변형시켰다. 결과적으로, BAC 클론에서 세포의 도메인을 암호화하는 (약 25,291 bp의) 마우스 *Tmprss2* 게놈 단편을 약 25,091 bp의 인간 *TMPRSS2* 게놈 단편으로 대체하였고, 약 2,691 bp의 자체 결실 네오마이신 카세트가 그 뒤에 이어졌다. 구체적으로, 대체된 마우스 *Tmprss2* 게놈 단편은 마우스 *Tmprss2* 유전자의 코딩 엑손 3의 마지막 7 bp, 인트론 3, 및 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 포함하였다(도 1a 내지 1b). 삽입된 인간 *TMPRSS2* 게놈 단편은 (인간 *TMPRSS2*의 3' UTR을 포함하는) 인간 *TMPRSS2* 유전자의 코딩 엑손 3의 마지막 7 bp, 인트론 3, 및 코딩 엑손 4 내지 코딩 엑손 13, 및 인간 *TMPRSS2*의 3' UTR 하류의

131 bp의 인간 3' 게놈 서열을 포함하였다(도 1a 내지 1b). 생성된 변형 BAC 클론은, 5'에서 3'까지, (i) 마우스 *Tmprss2* 5' UTR, 마우스 *Tmprss2* 엑손 1(비코딩), 코딩 엑손 1 내지 3(코딩 엑손 3의 마지막 7 bp 제외)를 포함하는 약 12 kb의 마우스 게놈 DNA를 함유하는 5' 마우스 상동 아암; (ii) 인간 코딩 엑손 3의 마지막 7 bp, 인트론 3, 인간 코딩 엑손 4 내지 13(인간 *TMPRSS2*의 3' UTR 포함), 및 인간 3' 게놈 서열을 포함하는 약 25,091 bp의 인간 *TMPRSS2* 게놈 단편; (iii) 약 2,691 bp의 자체 결실 네오마이신 카세트, 그 뒤에 이어지는 (iv) 마우스 *Tmprss2* 3' UTR 및 원래 BAC 클론에 남아 있는 마우스 게놈 DNA를 함유하는 45 kb의 3' 마우스 상동 아암을 포함하였다. 도 1a 내지 1b 참조. 접합 서열 또한 도 1b의 하단에 명시되어 있다. 인간 *TMPRSS2* 게놈 단편 및 네오마이신 카세트를 비롯하여 상류 및 하류 삽입 접합부를 함유하는 변형된 BAC 클론의 부분은 서열 번호 5에 명시되어 있다. 인간화 *Tmprss2* 유전자에 의해 암호화된 단백질의 아미노산 서열은 서열 번호 7에 명시되어 있다. 이러한 인간화 *Tmprss2* 단백질("7010 돌연변이 단백질"), 마우스 *Tmprss2* 단백질(서열 번호 2), 및 인간 *TMPRSS2* 단백질(서열 번호 4)의 배열은 도 1d에 제공된다.

[0142] 전술한 바와 같이, 인간화 *Tmprss2* 유전자를 함유하는 변형된 BAC 클론을 사용하여 마우스 배아 줄기(ES) 세포를 전기 천공하여 인간화 *Tmprss2* 유전자를 포함하는 변형된 ES 세포를 생성시켰다. 인간화 *Tmprss2* 유전자를 함유하는 양성으로 표적화된 ES 세포를 인간 *TMPRSS2* 서열(예를 들어, 인간 *TMPRSS2*의 코딩 엑손 4 내지 13)의 존재를 검출하고 마우스 *Tmprss2* 서열의 상실 및/또는 보유(예를 들어, 마우스 *Tmprss2*의 코딩 엑손 4 내지 13의 상실)를 확인한 검정(전술한 Valenzuela 문헌 참조)에 의해 확인하였다. 표 1에는 전술한 바와 같이(도 1a 내지 1b) 내인성 *Tmprss2* 유전자의 인간화를 확인하기 위해 사용한 프라이머 및 프로브가 명시되어 있다. 정확히 표적화된 ES 세포 클론이 선택되었으면, 예를 들어 전기 천공에 의해 Cre 재조합 효소를 도입함으로써 네오마이신 선택 카세트를 절제할 수 있다. 대안적으로, 네오마이신 선택 카세트는 ES 클론으로부터 생성된 자손을 Cre 재조합 효소를 발현하는 결실자(deletor) 설치류 균주와 교배시킴으로써 제거할 수 있다. 카세트의 결실 후의 인간화 *Tmprss2* 유전자좌는 도 1c에 도시되어 있으며, 접합 서열은 도 1c의 하단에 도시되어 있다.

[0143] 선택된 ES 세포 클론(카세트 유무를 불문함)은 인간화 *Tmprss2* 대립유전자를 게놈에 함유하는 한 배의 새끼를 생성하기 위해 VELOCIMOUSE® 방법을 사용하여 암컷 마우스를 이식하는 데 사용하였다(예: 미국 특허 번호 제 7,294,754호 및 Poueymirou 외의 F0 generation mice that are essentially fully derived from the donor gene-targeted ES cells allowing immediate phenotypic analyses, 2007, *Nature Biotech.* 25(1):91-99 참조). 인간화 *Tmprss2* 대립유전자를 갖는 마우스는 인간 *TMPRSS2* 유전자 서열의 존재를 검출하는 대립유전자 검정(전술한 Valenzuela 외의 문헌 참조)의 변형을 사용하여 꼬리 조각으로부터 단리된 DNA의 유전형을 분석함으로써 재확인하고 동정할 수 있다. 새끼의 유전형을 확인하고, 특성화를 위해 인간화 *Tmprss2* 유전자좌에 대해 이형접합성인 동물 코호트를 선택한다. 인간화 *Tmprss2* 유전자좌에 대해 동형접합성인 동물은 이형접합성 동물의 교배에 의해 만들어진다.

표 1

명칭	프라이머	서열(5'~3')	서열 번호
7010U	정방향	GCCGTGACTGTGACCTTCTC	(서열 번호 26)
	프로브 (BHQ)	TGGAGGAGCCACCTGATGCCTC	(서열 번호 27)
	역방향	GCCTTGCCCTCAATGGAAAC	(서열 번호 28)
7010D	정방향	GGTTGCACAGCAAGGAAGAAG	(서열 번호 29)
	프로브 (BHQ)	CCAGGAGTTCCTGTGAGCCTACCC	(서열 번호 30)
	역방향	TGGAATGGAAGGAGCTGGAG	(서열 번호 31)
7010hU	정방향	GTCCACCTCCTGCAACTG	(서열 번호 32)
	프로브 (BHQ)	TGAGCCTTCCCATCAGCCTGGG	(서열 번호 33)
	역방향	CCACAATGGCACATGGGTCTG	(서열 번호 34)
7010hTD	정방향	GGTGCTTGCTCCCCAAGA	(서열 번호 35)
	프로브 (BHQ)	CCTAAAAGGTGTGTGAATGG	(서열 번호 36)
	역방향	GGCAATAAAGAAGGAAGACGTTTT	(서열 번호 37)

[0144]

[0145] 실시예 2. 내인성 *Tmprss4* 유전자의 인간화

[0146] 본 실시예는 설치류(예를 들어, 마우스)에서 *Tmprss4*를 암호화하는 내인성 유전자를 인간화하는 예시적인 방법을 예시한다. 본 실시예에 기술된 방법은 원하는 대로 임의의 인간 서열, 또는 인간 서열(또는 서열 단편)의 조합을 사용하여 설치류의 내인성 *Tmprss4* 유전자를 인간화하는데 사용될 수 있다.

[0147] 박테리아 인공 염색체(BAC) 클론 및 VELOCIGENE® 기술을 사용하여 내인성 *Tmprss4* 유전자의 인간화를 위한 표적화 벡터를 작제하였다(예: 미국 특허 번호 제6,586,251호 및 전술한 Valenzuela 외의 (2003) 문헌 참조).

[0148] 요약하자면, 마우스 *Tmprss4* 유전자를 함유하는 마우스 박테리아 인공 염색체(BAC) 클론 RP23-71M15를 사용하고 다음과 같이 변형시켰다. 5' 마우스 상동 뉴클레오티드 서열, 약 4,996 bp의 자체 결실 네오마이신 카세트, (인간 *TMPRSS4* 유전자의 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 함유하는) 약 14,963 bp의 게놈 DNA, 및 3' 마우스 상동 서열을 포함하도록 DNA 단편을 생성하였다. 이러한 DNA 단편을 사용하여 박테리아 세포에서 상동성 재조합을 통해 BAC 클론 RP23-71M15를 변형시켰다. 결과적으로, BAC 클론에서 세포의 도메인을 암호화하는 (약 11,074 bp의) 마우스 게놈 단편을 약 4,996 bp의 자체 결실 네오마이신 카세트로 대체하였고, 약 14,963 bp의 인간 게놈 DNA가 그 뒤에 이어졌다. 구체적으로, 결실되고 대체된 마우스 게놈 단편은 마우스 인트론 3의 3' 130 bp, 마우스 *Tmprss4* 유전자의 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 포함하였다(도 2a 내지 2b). 삽입된 인간 게놈 단편은 약 150 bp의 인간 *TMPRSS4* 인트론 3의 3' 부분, 및 인간 *TMPRSS4* 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈을 포함하였다(도 2a 내지 2b). 생성된 변형된 BAC 클론은, 5'에서 3'까지, (마우스 *Tmprss4* 5' UTR, 마우스 *Tmprss4* 코딩 엑손 1 내지 3, (130 bp의 3'이 없는) 마우스 *Tmprss4* 인트론 3의 일부를 포함하는) 약 44.8 kb의 마우스 게놈 DNA, 약 4,996 bp의 자체 결실 네오마이신 카세트, 약 150 bp의 인간 *TMPRSS4* 인트론의 3' 부분, 인간 *TMPRSS4* 코딩 엑손 4 내지 코딩 엑손 13 내의 정지 코돈, 및 바로 뒤에 이어지는 마우스 *Tmprss4* 3' UTR 및 원래 BAC 클론에 남아 있는 마우스 게놈 DNA(합이 약 118 kb인 3' 마우스 상동 아암)를 포함하였다. 도 2a 내지 2b 참조. 접합 서열 또한 도 2b의 하단에 명시되어 있다. 네오마이신 카세트 및 인간 *TMPRSS4* 게놈 단편을 비롯하여 상류 및 하류 삽입 접합부를 함유하는 변형된 BAC 클론의 부분은 서열 번호 12에 명시되어 있다. 인간화 *Tmprss4* 유전자에 의해 암호화된 단백질의 아미노산 서열은 서열 번호 14에 명시되어 있다. 이러한 인간화 *Tmprss4* 단백질("7224 돌연변이 단백질"), 마우스 *Tmprss4* 단백질(서열 번호 9), 및 인간 *TMPRSS4* 단백질(서열 번호 11)의 배열은 도 2d에 제공된다.

[0149] 전술한 바와 같이, 인간화 *Tmprss4* 유전자를 함유하는 변형된 BAC 클론을 사용하여 마우스 배아 줄기(ES) 세포를 전기 천공하여 인간화 *Tmprss4* 유전자를 포함하는 변형된 ES 세포를 생성시켰다. 인간화 *Tmprss4* 유전자를 함유하는 양성으로 표적화된 ES 세포를, 인간 *TMPRSS4* 서열(예를 들어, 인간 *TMPRSS4*의 코딩 엑손 4 내지 13)의 존재를 검출하고 마우스 *Tmprss4* 서열의 상실 및/또는 보유(예를 들어, 마우스 *Tmprss4*의 코딩 엑손 4 내지 13의 상실)를 확인한 검정(전술한 Valenzuela 문헌 참조)에 의해 확인하였다. 표 2에는 전술한 바와 같이 (도 2a 내지 2b) 내인성 *Tmprss4* 유전자의 인간화를 확인하기 위해 사용한 프라이머 및 프로브가 명시되어 있다. 정확히 표적화된 ES 세포 클론이 선택되었으면, 예를 들어 전기 천공에 의해 Cre 재조합 효소를 도입함으로써 네오마이신 선택 카세트를 절제할 수 있다. 대안적으로, 네오마이신 선택 카세트는 ES 클론으로부터 생성된 자손을 Cre 재조합 효소를 발현하는 결실자(deletor) 설치류 균주와 교배시킴으로써 제거할 수 있다. 카세트의 결실 후의 인간화 *Tmprss4* 유전자좌는 도 2c에 도시되어 있으며, 접합 서열은 도 2c의 하단에 도시되어 있다.

[0150] 선택된 ES 세포 클론(카세트 유무를 불문함)은 인간화 *Tmprss4* 대립유전자를 게놈에 함유하는 한 배의 새끼를 생성하기 위해 VELOCIMOUSE® 방법을 사용하여 암컷 마우스를 이식하는 데 사용하였다(예: 미국 특허 번호 제 7,294,754호 및 전술한 Poueymirou 외의 (2007) 문헌 참조). 인간화 *Tmprss4* 대립유전자를 갖는 마우스는 인간 *TMPRSS4* 유전자 서열의 존재를 검출한 대립유전자 검정(전술한 Valenzuela 외의 문헌 참조)의 변형을 사용하여 꼬리 조각으로부터 단리된 DNA의 유전형을 분석함으로써 재확인하고 동정하였다. 새끼의 유전형을 확인하였고, 특성화를 위해 인간화 *Tmprss4* 유전자좌에 대해 이형접합성인 동물 코호트를 선택하였다. 인간화 *Tmprss4* 유전자좌에 대해 동형접합성인 동물을 이형접합성 동물의 교배에 의해 만들었다.

표 2

명칭	프라이머	서열(5'~3')	서열 번호
7224mTU	정방향	GAGCAGGGCCATGACACAT	(서열 번호 42)
	프로브 (BHQ)	ACCATTAGATCCCAGCACTGGACA	(서열 번호 43)
	역방향	AAACCCCTTCCCGAGAGAGAA	(서열 번호 44)
7224mTU2	정방향	GAGGAACACTGTGTCAAGGACTT	(서열 번호 45)
	프로브 (BHQ)	CCTGAAAAGCCCGAGTGGCAG	(서열 번호 46)
	역방향	GGGCAGAGACCACATCTGA	(서열 번호 47)
7224mTD	정방향	GGAAGCCCTCTCTCGATACTTG	(서열 번호 48)
	프로브 (BHQ)	TTCTACCCTGAGGGCATGCAGC	(서열 번호 49)
	역방향	TGGGATGTAGAAGGTTGTCAGA	(서열 번호 50)
7224hTU	정방향	CTGAGCCTGGAACACACATG	(서열 번호 51)
	프로브 (BHQ)	TCTGAGAGCCCAGCACTATCGCC	(서열 번호 52)
	역방향	GCTGAGGGTCAGGCTTGAG	(서열 번호 53)
7224hTD	정방향	TCTGCAGGGTAGGGAGAGAAG	(서열 번호 54)
	프로브 (BHQ)	TGTTTCAGAAAAGGAAGACTCACGTTC	(서열 번호 55)
	역방향	GAGACCGATGAAGAGAAAGTCAGA	(서열 번호 56)

[0151]

[0152]

실시예 3. 내인성 *Tmprss11d* 유전자의 인간화

[0153]

본 실시예는 설치류(예를 들어, 마우스)에서 *Tmprss11d*를 암호화하는 내인성 유전자를 인간화하는 예시적인 방법을 예시한다. 본 실시예에 기술된 방법은 원하는 대로 임의의 인간 서열, 또는 인간 서열(또는 서열 단편)의 조합을 사용하여 설치류의 내인성 *Tmprss11d* 유전자를 인간화하는데 사용될 수 있다.

[0154]

박테리아 인공 염색체(BAC) 클론 및 VELOCIGENE® 기술을 사용하여 내인성 *Tmprss11d* 유전자의 인간화를 위한 표적화 벡터를 작제하였다(예: 미국 특허 번호 제6,586,251호 및 전술한 Valenzuela 외의 (2003) 문헌 참조).

[0155]

요약하자면, 마우스 *Tmprss11d* 유전자를 함유하는 마우스 박테리아 인공 염색체(BAC) 클론 RP23-95N22를 사용하고 다음과 같이 변형시켰다. 5' 마우스 상동 뉴클레오타이드 서열, (인간 *TMPRSS11D* 유전자의 인트론 2의 3' 말단에서의 444 bp, 및 코딩 엑손 3 내지 코딩 엑손 10(코딩 엑손 10의 일부인 3' UTR을 포함함)을 함유하는) 약 33,927 bp의 인간 *TMPRSS11D* 게놈 DNA, 약 4,996 bp의 자체 결실 네오마이신 카세트, 및 3' 마우스 상동 서열을 포함하도록 DNA 단편을 생성하였다. 이러한 DNA 단편을 사용하여 박테리아 세포에서 상동성 재조합을 통해 BAC 클론 RP23-95N22를 변형시켰다. 결과적으로, BAC 클론에서 세포의 도메인을 암호화하는 (약 35,667 bp의) 마우스 *Tmprss11d* 게놈 단편을 약 33,927 bp의 인간 *TMPRSS11D* 게놈 단편으로 대체하였고, 약 4,996 bp의 자체 결실 네오마이신 카세트가 그 뒤에 이어졌다. 구체적으로, 대체된 마우스 *Tmprss11d* 게놈 단편은 마우스 *Tmprss11d* 유전자의 인트론 2의 3' 부분, 및 코딩 엑손 3 내지 코딩 엑손 10 내의 정지 코돈을 포함하였다(도 3a 내지 3b). 삽입된 인간 *TMPRSS11D* 게놈 단편은 (인간 *TMPRSS11D*의 3' UTR을 포함하는) 인간 *TMPRSS11D* 유전자의 인트론 2의 3' 말단에서의 444 bp, 및 코딩 엑손 3 내지 코딩 엑손 10, 및 인간 *TMPRSS11D*의 3' UTR 하류의 약 172 bp의 인간 3' 게놈 서열을 포함하였다(도 3a 내지 3b). 생성된 변형된 BAC 클론은, 5'에서 3'까지, (i) 마우스 *Tmprss11d* 5' UTR, 마우스 *Tmprss11d* 코딩 엑손 1 내지 2 및 인트론 2의 5' 부분을 포함하는 약 143 kb의 마우스 게놈 DNA를 함유하는 5' 마우스 상동 아암; (ii) 인간 *TMPRSS11D*의 인트론 2의 3' 부분 및 코딩 엑손 3 내지 10(3' UTR 포함함), 및 인간 3' 게놈 서열을 포함하는 인간 *TMPRSS11D* 게놈 단편; (iii) 약 4,996 bp의 자체 결실 네오마이신 카세트, 그 뒤에 이어지는 (iv) 마우스 *Tmprss11d* 3' UTR 및 원래 BAC 클론에 남아 있는 마우스 게놈 DNA를 함유하는 10 kb의 3' 마우스 상동 아암을 포함하였다. 도 3a 내지 3b 참조. 접합 서열 또한 도 3b의 하단에 명시되어 있다. 인간 *TMPRSS11D* 게놈 단편 및 네오마이신 카세트를 비롯하여 상류 및 하류 삽입 접합부를 함유하는 변형된 BAC 클론의 부분은 서열 번호 19에 명시되어 있다. 인간화 *Tmprss11d*

유전자에 의해 암호화된 단백질의 아미노산 서열은 서열 번호 21에 명시되어 있다. 이러한 인간화 *Tmprss11d* 단백질("7226 돌연변이 단백질"), 마우스 *Tmprss11d* 단백질(서열 번호 16), 및 인간 *TMPRSS11D* 단백질(서열 번호 18)의 배열은 도 3d에 제공된다.

[0156] 전술한 바와 같이, 인간화 *Tmprss11d* 유전자를 함유하는 변형된 BAC 클론을 사용하여 마우스 배아 줄기(ES) 세포를 전기 천공하여 인간화 *Tmprss11d* 유전자를 포함하는 변형된 ES 세포를 생성시켰다. 인간화 *Tmprss11d* 유전자를 함유하는 양성으로 표적화된 ES 세포는, 인간 *TMPRSS11D* 서열(예를 들어, 인간 *TMPRSS11D*의 코딩 엑손 3 내지 10)의 존재를 검출하고 마우스 *Tmprss11d* 서열의 상실 및/또는 보유(예를 들어, 마우스 *Tmprss11d*의 코딩 엑손 3 내지 10의 상실)를 확인하는 검정(전술한 Valenzuela 문헌 참조)에 의해 동정하였다. 표 3에는 전술한 바와 같이(도 3a 내지 3b) 내인성 *Tmprss11d* 유전자의 인간화를 확인하기 위해 사용한 프라이머 및 프로브가 명시되어 있다. 정확히 표적화된 ES 세포 클론이 선택되었으면, 예를 들어 전기 천공에 의해 Cre 재조합 효소를 도입함으로써 네오마이신 선택 카세트를 절제할 수 있다. 대안적으로, 네오마이신 선택 카세트는 ES 클론으로부터 생성된 자손을 Cre 재조합 효소를 발현하는 결실자(deletor) 설치류 군주와 교배시킴으로써 제거할 수 있다. 카세트의 결실 후의 인간화 *Tmprss11d* 유전자좌는 도 3c에 도시되어 있으며, 접합 서열은 도 3c의 하단에 도시되어 있다.

[0157] 선택된 ES 세포 클론(카세트 유무를 불문함)은 인간화 *Tmprss11d* 대립유전자를 게놈에 함유하는 한 배의 새끼를 생성하기 위해 VELOCIMOUSE® 방법을 사용하여 암컷 마우스를 이식하는 데 사용하였다(예: 미국 특허 번호 제 7,294,754호 및 전술한 Poueymirou 외의 (2007) 문헌 참조). 인간화 *Tmprss11d* 대립유전자를 갖는 마우스는 인간 *TMPRSS11D* 유전자 서열의 존재를 검출하는 대립유전자 검정(전술한 Valenzuela 외의 문헌 참조)의 변형을 사용하여 꼬리 조각으로부터 단리된 DNA의 유전형을 분석함으로써 재확인하고 동정하였다. 새끼의 유전형을 확인하였고, 특성화를 위해 인간화 *Tmprss11d* 유전자좌에 대해 이형접합성인 동물 코호트를 선택하였다. 인간화 *Tmprss11d* 유전자좌에 대해 동형접합성인 동물은 이형접합성 동물의 교배에 의해 만들어진다.

표 3

명칭	프라이머	서열(5'~3')	서열 번호
7226mTU	정방향	TCCTCTCCAGACAAGAAAGCT	(서열 번호 61)
	프로브 (BHQ)	TCATAGCAGCTTTCAAATCCTAAACGT TGA	(서열 번호 62)
	역방향	TCGTGTGTAGCTGGTGAGTT	(서열 번호 63)
7226mTD	정방향	CATGCGATCACAGGAGGAGATC	(서열 번호 64)
	프로브 (BHQ)	AATTGGGCCCCGAAGCCAGATGC	(서열 번호 65)
	역방향	CGGAAGGCTTCTGTGACTTC	(서열 번호 66)
7226hTU	정방향	GTCTCCCACTTCTGACATAATGAAC	(서열 번호 67)
	프로브 (BHQ)	CCCAGTGTTAACCCTACATCTGGTTCC	(서열 번호 68)
	역방향	TGGGAAGAGACTCTTGGACA	(서열 번호 69)
7226hTD	정방향	ATGAGCTCCTAGTACAGCTAAAGTT	(서열 번호 70)
	프로브 (MGB)	ATGCATGATCATCTATGCGTCAGAGC	(서열 번호 71)
	역방향	TGCCCAGATGCAGGGAGTTAG	(서열 번호 72)

[0158]

[0159] 실시예 4. MAID7225 HumIn 대 야생형 *Tmprss4* 마우스에서 1군 및 2군 인플루엔자 A 바이러스의 평가

[0160] 감염의 동물 모델로서 인간화 *Tmprss* 설치류를 사용의 유효성을 입증하기 위해, 중증 인플루엔자에 감염된 인플루엔자 A 1군 및 2군 모델에서 MAID7225 HumIn *TMPRSS4* 마우스의 생존율 대 야생형(WT) 한 배 새끼의 생존율을 평가하는 실험을 수행하였다.

[0161] MAID7225 HumIn *TMPRSS4* 마우스는 그의 게놈 내의 인간화 *Tmprss4* 유전자에 대해 동형접합체이고, 실시예 2에 기술된 바와 같이 생성되었다. 이들 연구에 사용된 바이러스 군주에는 역사적인 A/Puerto Rico/08/1934 (H1N1)

인플루엔자 A 바이러스 1군 분리 균주 및 실험실 내 마우스 적응형 A/Aichi/02/1968 (HA, NA) X-31 (H3N2) 인플루엔자 A 바이러스 2군 분리 균주가 포함되었다. 모든 실험은 6~8주령 암컷 및 수컷 MAID7225 HumIn TMPRSS4 마우스 또는 WT 한 배 새끼에서 수행하였다. 마우스에 1,150 플라크 형성 단위(PFU)의 A/Puerto Rico/08/1934(H1N1) 또는 10,000 PFU의 A/Aichi/02/1968-X31(H3N2)을 투여하였다. 이들 생존 모델에서, 마우스에 감염 후(p.i.) 0일 차에 비강 내(IN) 투여하였다. 감염 후 14일 차까지 매일 마우스의 몸무게를 재고 관찰하였고, 시작 몸무게의 20% 줄었을 때 희생시켰다. 결과는 생존 백분율로서 보고된다(표 4).

표 4

군 ID	군당 마우스 수	생존율(생존한 마우스의 수/군의 총 마우스 수)
미감염 대조군(2 HumIn, 2 WT 마우스)	4	100 (4/4)
WT TMPRSS4; H1_PR34 감염	10	20 (2/10)
HumIn TMPRSS4; H1_PR34 감염	8	25 (2/8)
WT TMPRSS4; H3_X31 감염	9	11.1 (1/9)
HumIn TMPRSS4; H3_X31 감염	8	25 (2/8)

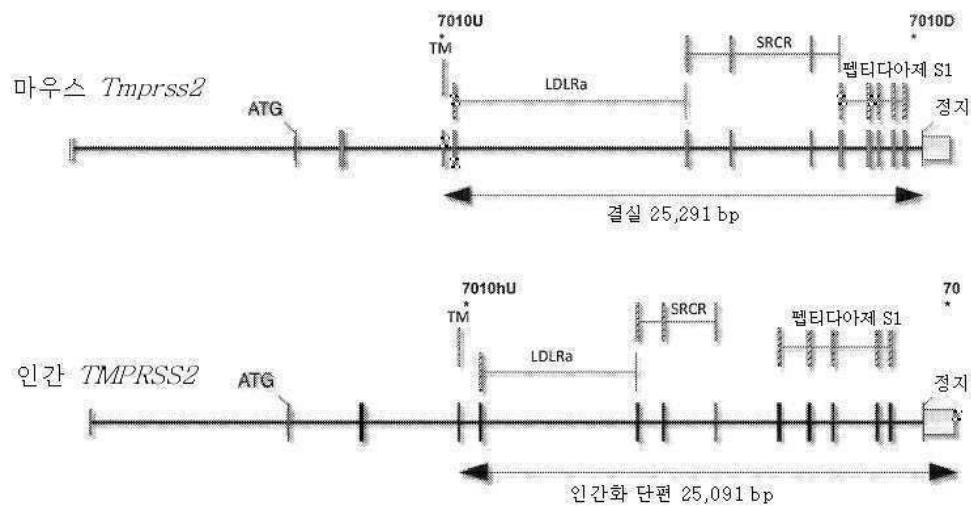
[0162]

[0163] MAID7225 HumIn TMPRSS4 마우스의 생존을 중증 인플루엔자 A 1군 바이러스[A/Puerto Rico/08/1934 (H1N1)] 및 중증 마우스 적응형 인플루엔자 A 2군 바이러스[A/Aichi/02/1968-X31 (H3N2)] 모두를 투여한 후의 WT 한 배 새끼와 비교하였다(도 4). MAID7225 HumIn TMPRSS4 마우스의 생존은 H1N1을 투여하거나 (각각 25%; n=8 및 20%; n=10) H3N2를 투여한 (각각, 25%; n=8 and 11.1%; n=9) 야생형 마우스와 차이가 없었다.

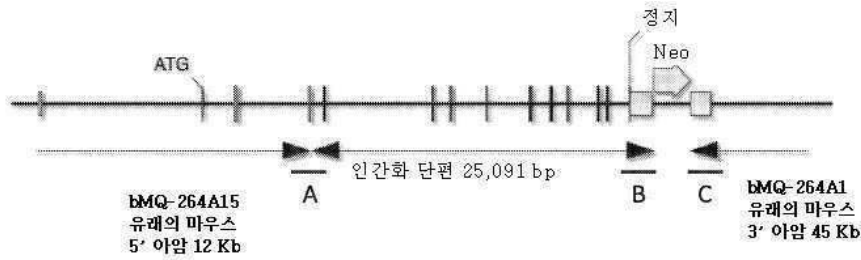
[0164] 간행물, 웹사이트 및 다른 참조 자료는 본 발명의 배경기술을 설명하고, 그의 시행이 참조로서 본원에 포함되는 것을 고려하여 추가적인 상세한 설명을 제공하기 위해 언급되었다.

도면

도면1a

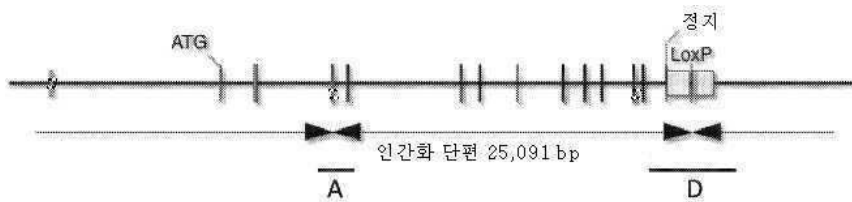


도면1b



A. 5' 마우스 // 5' 인간 (서열 번호 22)	AGCACCCCTC TCTTCCGCAG AGTCTAAGAA ATCGCTGTGT TTAGCCCTCG CCCTGGGCAC TGTCTCAGG GGAGCTGCTG TGGCTGCTGT CTTGCTTTGG // AAGTTCAGTA AGTGCAGGGA GCCTCGATCC CACCATGTGC TCCTGCAGTC CCCAGTGCTC TGAGCCAGAC CCTGCTCTCT GGGCTATTGA GACCTCTGGA GGGCCTCCGT GAGGTTCTCT TCTTACATAA CGAGGCTGTC TCTCTCCCT TCTCTTG
B. 인간 // <u>XhoI</u> // (<u>loxP</u>) 카세트 (서열 번호 23)	GGTCAGAGGA CCAAAGGTGA GGCAAGGCCA GACTTGGTGC TCCTGTGGTT// CTCGAG // <u>ATAACTTCG TATAATGTAT GCTATACGAA GTTAT</u> ATGCATGGCC TCCGCGCCGG GTTTTGGGCG CTCCCGCGGG CGCCCCCTC CTCACGGCGA GCGCTGCCAC GTCAGACGAA GGCGCGAGCG AGCGTCCTGA
C. 카세트 (<u>loxP</u>) /CEU// <u>NheI</u> // 마우스 (서열 번호 24)	ATTGTTTTGC CAAGTTCTAA TTCCATCAGA CCTCGACCTG CAGCCCCCTAG <u>ATAACTTCGT ATATGTATG CTATACGAAG TTAT</u> // GCTAGTAACATAACGGTCCTAAGGTAGCGA // GCTAGC // TCCACGTGGC TTTGTCCAG ACTTCCTTTG TCTTCAACAA CCTTCTGCAA

도면1c

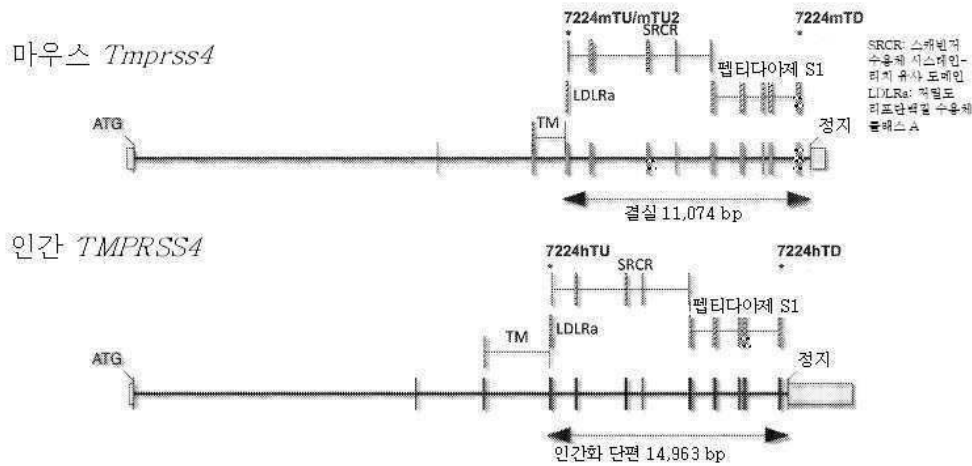


A. 5' 마우스 // 5' 인간 (서열 번호 22)	AGCACCCCTC TCTTCCGCAG AGTCTAAGAA ATCGCTGTGT TTAGCCCTCG CCCTGGGCAC TGTCTCAGG GGAGCTGCTG TGGCTGCTGT CTTGCTTTGG // AAGTTCAGTA AGTGCAGGGA GCCTCGATCC CACCATGTGC TCCTGCAGTC CCCAGTGCTC TGAGCCAGAC CCTGCTCTCT GGGCTATTGA GACCTCTGGA GGGCCTCCGT GAGGTTCTCT TCTTACATAA CGAGGCTGTC TCTCTCCCT TCTCTTG
D. 인간 // <u>XhoI</u> // (<u>loxP</u>) /CEU// <u>NheI</u> // 마우스 (서열 번호 25)	GGTCAGAGGA CCAAAGGTGA GGCAAGGCCA GACTTGGTGC TCCTGTGGTT// CTCGAG // <u>ATAACTTCG TATAATGTAT GCTATACGAA GTTAT</u> /GCTAGTAACATAACGGTCCTAAGGTAGCGA // GCTAGC // TCCACGTGGC TTTGTCCAG ACTTCCTTTG TCTTCAACAA CCTTCTGCAA

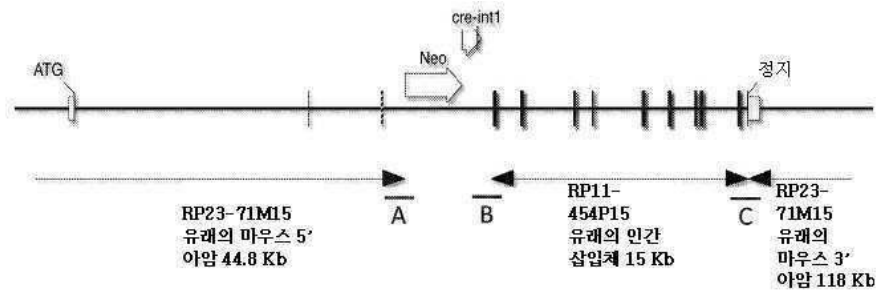
도면1d

Tmprss2 단백질 배열	
hTMPrSS2	10 20 30 40 50 60 MALNSGSPFAIGPYENHGYQFENPYFAQPTVVFTVYEVHFAQYYPSVPQYAPRVLTTQA
mTmprss2	MALNSGSPGIGPCYENHGYQSEHICPPRPVFAFNGYNLYPAQYYPSVPQYAPRITTTQA
7010 돌연변이 단백질	MALNSGSPGIGPCYENHGYQSEHICPPRPVFAFNGYNLYPAQYYPSVPQYAPRITTTQA ***** * * * * * , * , * * , * , ***** , **
hTMPrSS2	70 80 90 100 110 120 SNPVVCTQPKSPSGTCTSKTKKALCITLTGLTFLVGAALAAAGLLWKFMGSKCSNSGIEQ
mTmprss2	STSVIHTHPKS-SGALCTSKSKSLCLALALGTVLTGAAVAALLWRFWDNCSSTSEMEQ
7010 돌연변이 단백질	STSVIHTHPKS-SGALCTSKSKSLCLALALGTVLTGAAVAALLWRFWDNCSSTSEMEQ * , *
hTMPrSS2	130 140 150 160 170 180 DSSGTCINPSNWC DGVS HC PGGEDENRCVRLYGFNFILOVYSSQRRKSWHPVCQDWNENY
mTmprss2	GSSGTCISSSLWC DGVAHC PNGEDENRCVRLYGQSFILQVYSSQKAWYPVCQDDWSES
7010 돌연변이 단백질	DSSGTCINPSNWC DGVS HC PGGEDENRCVRLYGFNFILOVYSSQRRKSWHPVCQDWNENY ***** * * * * * , * , * * * * , * , * * * * , * , * * * * , *
hTMPrSS2	190 200 210 220 230 240 GRAACRDMGYKNNFYSSQGI VDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLH
mTmprss2	GRAACRDMGYKNNFYSSQGI PDQSGATSFMKLNVSSGNVDLYKKLYHSDC SSRMVVSLH
7010 돌연변이 단백질	GRAACRDMGYKNNFYSSQGI VDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLH ***** , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , *
hTMPrSS2	250 260 270 280 290 300 CIACGVNLSSRQSRIVGGESALPGANFWQVSLHVQNVHVC GGSII TPENIVTAAHCVEK
mTmprss2	CIECGVRS-VKRQSRIVGGLNASPGDWFQVSLHVQGVHVC GGSII TPENIVTAAHCVEE
7010 돌연변이 단백질	CIACGVNLSSRQSRIVGGESALPGANFWQVSLHVQNVHVC GGSII TPENIVTAAHCVEK ** *** * * * * * , * , * * * * , * , * * * * , * , * * * * , *
Tmprss2 단백질 배열	
hTMPrSS2	310 320 330 340 350 360 PLNNPWHHTAFAGILRQSFMYGAGYQVEKVISHPNYDSKTKNNDIALMKLQKPLTFNDL
mTmprss2	PLSSPYWTAFAGILRQSLMFYGSRRHQVEKVI SHPNYDSKTKNNDIALMKLQTPAFNDL
7010 돌연변이 단백질	PLNNPWHHTAFAGILRQSFMYGAGYQVEKVISHPNYDSKTKNNDIALMKLQKPLTFNDL ***** , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , *
hTMPrSS2	370 380 390 400 410 420 VKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCNSRYVDNLI
mTmprss2	VKPVCLPNPGMMLDLQECWISGWGATYEGKTS DVLNAAVPLIEPSKCNKYIYNNLI
7010 돌연변이 단백질	VKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCNSRYVDNLI ***** , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , *
hTMPrSS2	430 440 450 460 470 480 TPAMICAGFLQGNVDS CQGD SGGPLVTSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVF
mTmprss2	TPAMICAGFLQGSVDSCQGD SGGPLVTLKNGIWWLIGDTSWGS GCAKALRPGVYGNVTVF
7010 돌연변이 단백질	TPAMICAGFLQGNVDS CQGD SGGPLVTSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVF ***** , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , * , *
hTMPrSS2	490 TDWIYRQMRADG
mTmprss2	TDWIYQMRANS
7010 돌연변이 단백질	TDWIYRQMRADG *****
실선: TM 실선 박스: LDLRa 점선 박스: SRCR 도메인 점선: 펩시다아제 S1	

도면2a

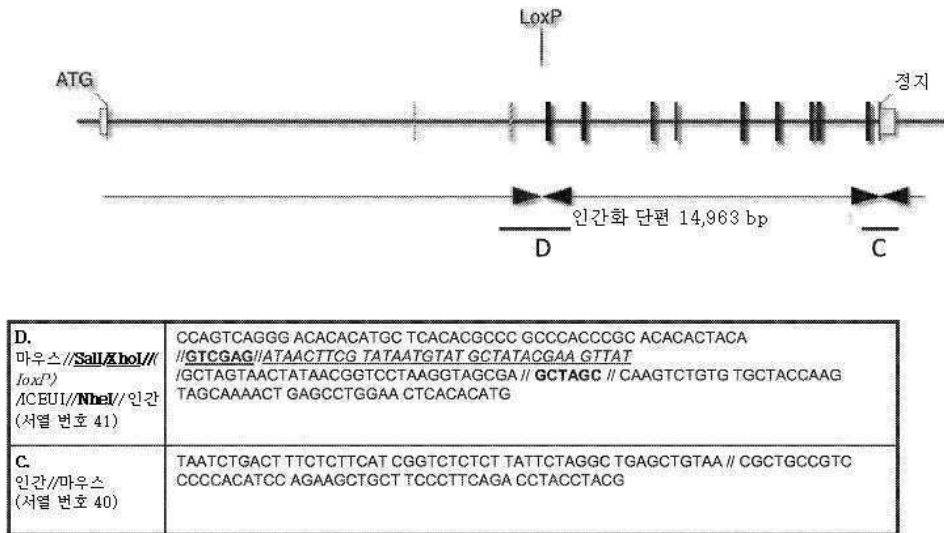


도면2b



A. 마우스// <i>SalI</i> - <i>XhoI</i> //(<i>loxP</i>) 카세트 (서열 번호 38)	CCAGTCAGGG ACACACATGC TCACACGCC GCCACCCGC ACACACTACA //GTCGAG //ATAACTTCGT ATAATGTATG CTATACGAAG TTAT ATGCATGGCC TCCGCGCCGG GTTTTGGCGC
B. 카세트 (<i>loxP</i>)// <i>ICE1</i> // <i>NheI</i> //인간 (서열 번호 39)	ATTCTAGTTG TGGTTTGTCC AAACATCA ATGTATCTTA TCATGTCTGG AATAAGTTCGTATAATGTATGCTATACGAAGTTAT/GCTAG / TAACTATAACGGTCTAAGGTAGCGA //GCTAGC// CAAGTCTGTG TGCTACCAAG TAGCAAACT GAGCCTGGAA CTCACACATG CGTGTCTGAG AGCCCAGCAG TATCGC
C. 인간/마우스 (서열 번호 40)	TAATCTGACT TTCTCTTCAT CGGTCTCTCT TATTCTAGGC TGAGCTGTAA // CGCTGCCGTC CCCCACATCC AGAAGTGCT TCCCTTCAGA CCTACCTACG

도면2c



도면2d

Tmprss4 단백질 배열	
hTMPRSS4	MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPILIIALLSLASIIIVVLIKVILDKY
mTmprss4	MESDSCQPLNNRDIVPFRKPRRPQETFKKVGIPILIIAVLLSLIALVIVALLIKVILDKY
7224 돌연변이 단백질	MESDSCQPLNNRDIVPFRKPRRPQETFKKVGIPILIIAVLLSLIALVIVALLIKVILDKY

hTMPRSS4	YFLCGQPLHFIERKQLCDGELDCPLGEDEEHCVKSFPEGPAAVAVRLSKDRSTLQVLDSAT
mTmprss4	YFICGSPLTFIQRGQLCDGHLDCASGEDEEHCVKDFPEKPGVAVRLSKDRSTLQVLDAAT
7224 돌연변이 단백질	YFLCGQPLHFIERKQLCDGELDCPLGEDEEHCVKSFPEGPAAVAVRLSKDRSTLQVLDSAT

hTMPRSS4	GNWFSACFDNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELRMRNS
mTmprss4	GTWASVCFDNFTEALAKTACRQMGYSQPAFRAVEIRPDQNLPAQVTGNSQELQVQNGS
7224 돌연변이 단백질	GNWFSACFDNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELRMRNS

hTMPRSS4	GPCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWFWQVSIQYDRQHVCVGGSIIDPHWV
mTmprss4	RSCLSGSLVSLRCLDCGKSLKTPRVVGVEAPVDSWFWQVSIQYNKQHVCGGSIIDPHWI
7224 돌연변이 단백질	GPCLSGSLVSLHCLACGKSLKTPRVVGVEASVDSWFWQVSIQYDRQHVCVGGSIIDPHWV

hTMPRSS4	LTAAHCFRKHTDFVFNWKVRAGSDKLGSPFLAVAKIIIEFNPMYPKDNDIALMKLQFPL
mTmprss4	LTAAHCFRKLYLDVSSWKVRAGSNILGNSPFLPAKIFIAEPNPLYPKERDIALVQLQMP
7224 돌연변이 단백질	LTAAHCFRKHTDFVFNWKVRAGSDKLGSPFLAVAKIIIEFNPMYPKDNDIALMKLQFPL

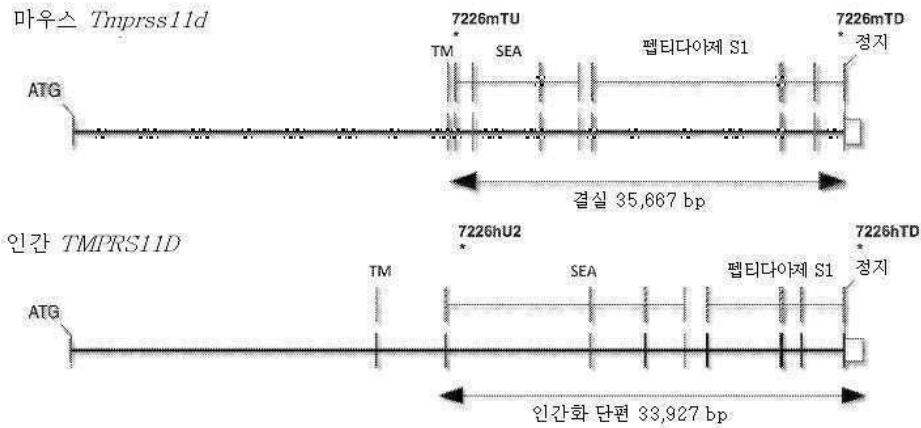
Tmprss4 단백질 배열	
hTMPRSS4	TFSGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGGKMSDILLQASVQVIDSTRCNAED
mTmprss4	TFSGSVRPICLPFSDEVLPATPVWVIGWFTENGKMSDMLLQASVQVIDSTRCNAED
7224 돌연변이 단백질	TFSGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGGKMSDILLQASVQVIDSTRCNAED

hTMPRSS4	AYQGEVTEKMMKAGIPEGGVDTCCQDGGGGLMYQSDQWHVVGIVSWGYGCGGPSTPGVYT
mTmprss4	AYEGEVTAEMLCAGTFQGGKDTCCQDGGGGLMYHSDKQVVGIVSWGHGCGGPSTPGVYT
7224 돌연변이 단백질	AYQGEVTEKMMKAGIPEGGVDTCCQDGGGGLMYQSDQWHVVGIVSWGYGCGGPSTPGVYT

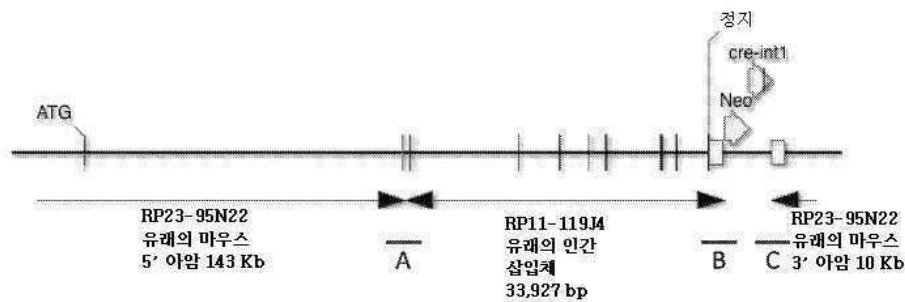
hTMPRSS4	KVSAYLNWIYNVWKAEL
mTmprss4	KVTAYLNWIYNVRKSEM
7224 돌연변이 단백질	KVSAYLNWIYNVWKAEL

실선: TM
실선 박스: LDLRa
점선 박스: SRCR 도메인
점선: 펩시다아제 S1

도면3a

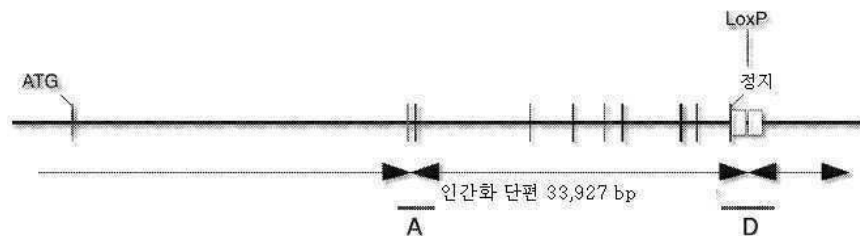


도면3b



A 마우스 // 인간 (서열 번호 57)	GACCATTTTAAGGTTTTGCT TGGTTGTTTT GGAGGGAGGG TGGTGCTTTG // CTAATGGTGA ATTACTAACT CCTCAATAAA GAATATTATT TGAATAAATT
B 인간 // XhoI / (loxP) 카세트 (서열 번호 58)	GCTGCCTTTT AATCCAGCGC TATAATTGAG GCAAGCGTCC AGCTTGACAC // CTCGAG // ATAACCTTCG TATAATGTAT GCTATACGAA GTTAT ATGCATGGCC TCCGCGCCGG GTTTTGGCGC CTCCCGCGGG CGCCCCCTC CTCACGGCGA GCGCTGCCAC GTCAGACGAA GGGCGCAGCG AGCGTCCTGA
C 카세트 (loxP) // ICEU // Nhe // 마우스 (서열 번호 59)	ATTGTTTTGC CAAGTTCTAA TTCCATCAGA CCTCGACCTG CAGCCCCTAG ATAACCTTCGT ATAATGTATG CTATACGAAG TTAT / GCTAGTAACATAACGGTCCTAAGGTAGCGA // GCTAGC //TGCAACCGAG GAAAAACGT GCCATGAGGT CTCTGTATCC AAGTGTGACT

도면3c



A 마우스 // 인간 (서열 번호 57)	GACCATTTTA AGGTTTTGCT TGGTTGTTTT GGAGGGAGGG TGGTGCTTTG // CTAATGGTGA ATTACTAACT CCTCAATAAA GAATATTATT TGAATAAATT
D 인간 // XhoI // (loxP) // ICEU // Nhe // 마우스 (서열 번호 60)	CCAGTCAGGG ACACACATGC TCACACGCCG GCGCACCCGC ACACACTACA // CTCGAG // ATAACCTTCG TATAATGTAT GCTATACGAA GTTAT /GCTAGTAACATAACGGTCCTAAGGTAGCGA // GCTAGC // TGCAACCGAG GAAAAACGT GCCATGAGGT CTCTGTATCC AAGTGTGACT

도면 3d

Tmprss11d 단백질 배열

10 20 30 40 50 60
 hTMPRSS11D MYRPARTVTSRFLNPYVVCIFVAGVVLAVTIALLVFLAFD KSYFYRRSSFQLLNVE
 mTmprss11d MYRPRLMSPSRFFTPFAVAFVVIITVGLLAMMAGLLIHFLAFD KAYFYHSSSFQLNVE
 7226 돌연변이 MYRPRLMSPSRFFTPFAVAFVVIITVGLLAMMAGLLIHFLAFD KSYFYRRSSFQLLNVE
 단백질 *****

hTmPRSS11D YNSQLNSPATQEYRTLSGRIESLITKTFKESNLRNQIFRAHVAKLRQDGSQVRADVVMKIF
mTmPRSS11D YTEALNSPATHEYRTLSEIEAMITDFPRGSSLKSEFIRTHVVKLRKEGTGQVQVADVVMKIF
722 돌연변이 YNSQLNSPATQEYRTLSGRIESLITKTFKESNLRNQIFRAHVAKLRQDGSQVRADVVMKIF
단백질

hTMPRSS11D QFTRRNNNGASMKSRIESVLRQMLNNSGNLEINPSTETISLTDDQAAANWNLNECGAGPDIL
mTmprss11d RSSKRNRRNKVMKTRIQSVLRR-LSSSGNLEIAPSNETISLTDDQNTENVLTQECGARGPDIL
7226 돌연변이 QFTRRNNNGASMKSRIESVLRQMLNNSGNLEINPSTETISLTDDQAAANWNLNECGAGPDIL
다백전

190 200 210 220 230 240
 hTMRSS11D TLSEQRILGGTEAEGSPWPQVSLRLNNAHHCGLSINNMMWILTAACHCFRSNSNPRDWIA
 mTmrss11d TLSEERIIGGMQAEFGDPWPQVSLQLNVVHHCGGALISNNMWLTAACHCFKSYNPQYWTAA
 7226 돌연변이 TLSEQRILGGTWPQVSLRLNNAHHCGLSINNMMWILTAACHCFRSNSNPRDWIA
 단백질 *****

250 260 270 280 290 300

hTmPrSS11D TSGISTTFPKLRMVRNILLHNHYKSATHENDIALVRLNSVTFTKDIHSVCLPAATQNI

TmPrss11d TFGVSTMSFRLRVRAILAHGDYSSVTRNDIAVQLDRSVAFSRNIHRVCLPAATQNI

7226 돌연변이 TSGISTTFPKLRMVRNILLHNHYKSATHENDIALVRLNSVTFTKDIHSVCLPAATQNI

단백질

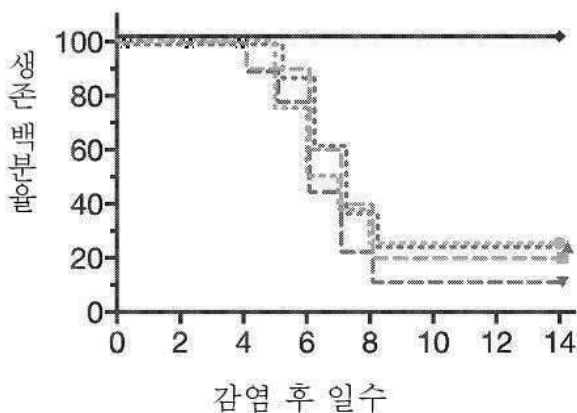
Tmprss11d 단백질 배열

310 320 330 340 350 360
hTMPRSS11D PPGSTAYVTGWGAQ^YAGHTVPELRQ^GQVRIISNDVCNAPHSYNGAILSGMLCAGVPGQG
mTmprss11d IPGSVAYVTGWGSLTYGGNAVTNLRQGEVRIISSEECNTPAGYSGSVLPQMLCAGMRSGA
7226 돌연변이 PPGSTAYVTGWGAQ^YAGHTVPELRQ^GQVRIISNDVCNAPHSYNGAILSGMLCAGVPGQG
다배정 *****

370 380 390 400 410
hTMPRSS11D VDACQGDSGGFLVQEDSRRLWFIVGIVSWGDCGLPDKPGVYTRVTAYLWDWIRQQTGI
mTmprss11d VDACQGDSGGFLVQEDSRRLWFVVGIVSWGQCGLPNKPGVYTRVTAYRNWIRQQTGI
726 돌연변이
VDACQGDSGGFLVQEDSRRLWFIVGIVSWGDCGLPDKPGVYTRVTAYLWDWIRQQTGI
다배치

실선: TM
실선 박스: SEA 도메인
점선: 펄시다아제 S1

도면4



서 열 목 록

SEQUENCE LISTING

<110> REGENERON PHARMACEUTICALS, INC.

<120> RODENTS HAVING A HUMANIZED TMPRSS GENE

<130> 33093PCT (10234W001)

<150> 62/301,023

<151> 2016-02-29

<160> 72

<170> PatentIn version 3.5

<210> 1

<211> 3175

<212> DNA

<213> Mus musculus

<400> 1

```
gcctttcctg gccgttcct ctttctggcc gaggtgcctg cgttagggg tgtcacctg      60
gctcccgga cgcgcctcc ggagatttaa gcgagaactg gagtaggtcg tgtacttga      120
gcggacgagg aagccaagag ctccgacaga ggcggagagg ggcgggaagc gcaacaggtc      180

acctggagga agccccatac tgacctctc atgtgtctga cacaggcagg atggcattga      240
actcagggtc acctccagga atcggacctt gctatgagaa ccacgggtat cagtctgagc      300
acatctgtcc tccgagacca ccagtggctc ccaatggcta caacttgtat ccagcccagt      360
actaccctc tccagtgcct cagtatgtc cgaggattac aacgcaagcc tcaacatctg      420
tcatccacac acatccaag tcctcaggag cactgtgcac ctcaaagtct aagaaatcgc      480
tgtgtttagc cctcgccctg ggcactgtcc tcacgggagc tgctgtggct gctgtcttgc      540
tttgagggtt ctgggacagc aactgttcta cgtctgagat ggagtgtggg tcttcaggca      600

catgcatcag ctcttctctc tgggtgtgacg gggtagcaca ttgtcccaac ggagaagatg      660
agaaccgttg tgctctctc tacggacaaa gcttcacct ccaggtttac tcatctcaga      720
ggaaagcctg gtatcccggtg tgccaggatg attggagtga gagctacggg agagcagcat      780
glaaagacat gggatacaag aacaattttt attctagcca agggatacca gaccagagcg      840
gggcaacgag ctttatgaag ctgaatgtga gctcaggcaa cgttgacctc tataaaaaac      900
tctaccacag tgactcatgt tcatcccga tgggtgtttc tttgcgtgt atagaatgcg      960
gggttcgctc agtgaacgc cagagcagga ttgtgggtgg attgaatgcc tcaccaggag      1020

actggccctg gcaggtcagc ctgcacgtcc aaggcgtcca cgtctgcgga ggctccatca      1080
tcacccccga gtggattgtg acggccgccc actgtgtgga agaaccctc agcagcccga      1140
ggtactggac ggcatttgcg ggaattctga gacagtctct catgttctat ggaagtagac      1200
```

accaggtaga aaaagtaatt tcccatccaa attacgactc taagaccaag aataacgaca 1260
 ttgctctcat gaagctgcag acacctttgg cttttaatga tctagtgaag ccagtgtgtc 1320
 tgccgaacct aggcatgatg ctagacctag accaggaatg ctggatttcg ggggtggggg 1380
 ccacctatga gaaagggaag acctcggacg tgttgaatgc tgccatggta cccttgatcg 1440

agccctccaa atgtaatagt aaatacatat acaacaacct aatcacacca gccatgatct 1500
 gtgccggctt cctccagggg tctgtcgact cttgccaggg agacagtgga gggccgctgg 1560
 ttactttgaa gaatgggatc tgggtggctga ttggggacac gagctggggc tcgggctgtg 1620
 ccaaggcact cagacctgga gtatacggga acgtgacggg atttacagat tggatctacc 1680
 agcaaatgag ggcgaacagc taatccacgt ggctttgtcc cagacttcct ttgtcttcaa 1740
 caaccttctg caagaaaacc aagggcctga attttaactt cctgtgcaca atgtaccttt 1800
 tgagatgatt cgaagggcct ttcactttta ttaaacagtg acttgtttga ctgtgtctcc 1860

tggtcctgtg agggcttcag tgccccacc cttggccact tctgcagctc ccaccagaat 1920
 ggatgaccag attctgttgg gtttgggcac atagggccaa aggcagagga ggggtggcact 1980
 ctcatgttgg aacttctttt gggctcatgc tcaggccttt tttggatcac taaggactat 2040
 gacctctgag taacctgatg acctgagaaa gagtaaggag gccaggcagg gccttgggcc 2100
 caggaaacagg taccttgaga gtgagagcta ccattgcct gtggcctaaa tctgctgtgc 2160
 aggttgggct ggctacactg tcatgatttc attaacagcc tgggtgaaca tggctgggag 2220
 taaagggctt gctctcctgc atgttgacat gacggccctt tccaagggtg atggaggctt 2280

tcccaagcta agggcctagg cagatctctc agagcaagaa gctaataccg gcatgtccct 2340
 tgggtgagct ctacatgggt ttattcagtc tggttcttgg ctccccacta ctgtttctct 2400
 cagcctctca gagcctgaaa cttacctctt agctttggct acaggcatgg cctagtacct 2460
 gatggagcct gtatagctca gctaataaaa tggaggetca ggtccatcag aatcaggac 2520
 ttgtgatctc agtcaccttg cttctgggtt gtgtttcttc tcttactacc tactgcacc 2580
 tggacactag agtggatgaa tgtctggagt tcacctgcat ttggactgtg tgattgtgcc 2640
 tcagacacta gacctcttcc agatggtttag gttgttctgt agactggcaa tgagattaga 2700

agttcctagc ttcagataaa gatgaaagag aggagatcat tgtcttctgt cttcttctgg 2760
 ccctgggttt ataccaggaa agccatgcca gaattaccaa atatgaagta tgaatgtctt 2820
 acccaggtg aggcctctgcc tccttctctc tgccctgggtt ttcagaaggc agtgaatggg 2880
 tcataactgg gactccatct ttgctgggga aagtctccca cctagggaat ggttaccact 2940
 ccatgtaaag aaaactccct catgcgtcct ctgggacctt cttagatgct gtaaggtacc 3000
 tacatacaga ctaaattgtc aagcaccttg aagtgtgaga acctgtcccc tccttagctc 3060

tccttgcttt tgctgttggg tgggtatttc ctgctttgtg tctgttctga gctgtgagat 3120

tccactgtga aatatatgaa taaagtatat aattctttta aaaaaaaaaa aaaaa 3175

<210> 2

<211> 490

<212> PRT

<213> Mus musculus

<400> 2

Met Ala Leu Asn Ser Gly Ser Pro Pro Gly Ile Gly Pro Cys Tyr Glu

1 5 10 15

Asn His Gly Tyr Gln Ser Glu His Ile Cys Pro Pro Arg Pro Pro Val

20 25 30

Ala Pro Asn Gly Tyr Asn Leu Tyr Pro Ala Gln Tyr Tyr Pro Ser Pro

35 40 45

Val Pro Gln Tyr Ala Pro Arg Ile Thr Thr Gln Ala Ser Thr Ser Val

50 55 60

Ile His Thr His Pro Lys Ser Ser Gly Ala Leu Cys Thr Ser Lys Ser

65 70 75 80

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

85 90 95

Ala Ala Val Ala Ala Val Leu Leu Trp Arg Phe Trp Asp Ser Asn Cys

100 105 110

Ser Thr Ser Glu Met Glu Cys Gly Ser Ser Gly Thr Cys Ile Ser Ser

115 120 125

Ser Leu Trp Cys Asp Gly Val Ala His Cys Pro Asn Gly Glu Asp Glu

130 135 140

Asn Arg Cys Val Arg Leu Tyr Gly Gln Ser Phe Ile Leu Gln Val Tyr

145 150 155 160

Ser Ser Gln Arg Lys Ala Trp Tyr Pro Val Cys Gln Asp Asp Trp Ser

165 170 175

Glu Ser Tyr Gly Arg Ala Ala Cys Lys Asp Met Gly Tyr Lys Asn Asn

180 185 190

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

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

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

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

 Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Leu Lys Asn

435 440 445
 Gly Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly Cys Ala
 450 455 460
 Lys Ala Leu Arg Pro Gly Val Tyr Gly Asn Val Thr Val Phe Thr Asp
 465 470 475 480
 Trp Ile Tyr Gln Gln Met Arg Ala Asn Ser
 485 490
 <210> 3
 <211> 3212
 <212> DNA
 <213> Homo sapiens
 <400> 3
 gagtaggcgc gagctaagca ggaggcggag gcggaggcgg agggcgaggg gcggggagcg 60
 ccgcctggag cgcggcaggt catattgaac attccagata cctatcatta ctgatgctg 120
 ttgataacag caagatggct ttgaactcag ggtcaccacc agctattgga ctttactatg 180
 aaaacatgg ataccaaccg gaaaaccct atccgcaca gccactgtg gtccccactg 240
 tctacaggt gcatcggct cagtactacc cgtccccgt gcccagtag ccccagagg 300
 tcctgacgca ggcttccaac cccgtcgtct gcacgcagcc caaatccca tccgggacag 360
 tgtgcacctc aaagactaag aaagactgt gcatcacctt gaccctgggg accttcctcg 420
 tgggagctgc gctggccgct ggcctactct ggaagtcat gggcagcaag tgctccaact 480
 ctgggataga gtgcgactcc tcaggtacct gcatcaacc ctctaactgg tgtgatggcg 540
 tgtcacactg ccccgccggg gaggacgaga atcggtgtgt tcgcctctac ggaccaaact 600
 tcatccttca ggigtactca tctcagagga agtcttgga cctgtgtgc caagacgact 660
 ggaacgagaa ctacggcgcg gcggcctgca gggacatggg ctataagaat aatttttact 720
 ctaccaagg aatagtggat gacagcggat ccaccagctt tatgaaactg aacacaagtg 780
 ccggcaatgt cgatatctat aaaaaactgt accacagtga tgcctgttct tcaaaagcag 840
 tggtttcttt acgtgtata gcctgcgggg tcaacttgaa ctcaagccgc cagagcagga 900
 ttgtggcgcg cgagagcgcg ctcccggggg cctggccctg gcaggtcagc ctgcacgtcc 960
 agaacgtcca cgtgtgcgga ggctccatca tcaccccgga gtggatcgtg acagccgccc 1020
 actgcgtgga aaaacctctt aacaatccat ggcatcggac ggcatcgtg gggattttga 1080
 gacaatcttt catgttctat ggagccggat accaagtaga aaaagtatt tctcatccaa 1140
 attatgactc caagaccaag aacaatgaca ttgcgtgat gaagctgcag aagcctctga 1200

ctttcaacga cctagtgaac ccagtgtgtc tgcccaaccc aggcatgatg ctgcagccag 1260

aacagctctg ctggatttcc ggggtggggg ccaccgagga gaaagggaag acctcagaag 1320

tgctgaacgc tgccaagggtg cttctcattg agacacagag atgcaacagc agatatgtct 1380

atgacaacct gatcacacca gccatgatct gtgccggctt cctgcagggg aacgtcgatt 1440

cttgccaggg tgacagtga gggcctctgg tcacttcgaa gaacaatatc tgggtggctga 1500

taggggatac aagctggggg tctggctgtg ccaaagctta cagaccagga gtgtacggga 1560

atgtgatggt attcacggac tggatttacc gacaaatgag ggacagcggc taatccacat 1620

ggtcttcgtc cttgacgtcg ttttacaaga aaacaatggg gctgggtttg cttccccgtg 1680

catgatttac tcttagagat gattcagagg tcacttcatt tttattaaac agtgaacttg 1740

tctggctttg gcactctctg ccattctgtg caggctgcag tggctccct gccagcctg 1800

ctctccctaa ccccttgtcc gcaaggggtg atggccggct ggttgtgggc actggcggtc 1860

aagtgtggag gagaggggtg gaggtgccc cattgagatc ttcctgctga gtcctttcca 1920

ggggccaatt ttggatgagc atggagctgt cacctctcag ctgctggatg acttgagatg 1980

aaaaaggaga gacatggaaa gggagacagc cagggtggcag ctgcagcggc tgccctctgg 2040

ggccacttgg tagtgtcccc agcctacctc tccacaaggg gattttgctg atgggttctt 2100

agagccttag cagccctgga tgggtggccag aaataaagg accagccctt catgggtggt 2160

gacgtggtag tcacttgtaa ggggaacaga aacatttttg ttcctatggg gtgagaatat 2220

agacagtgcc cttggtgcga gggaagcaat tgaaaaggaa cttgccctga gcactcctgg 2280

tgcaggtctc cacctgcaca ttgggtgggg ctccctgggag ggagactcag ccttctctct 2340

cactctccct gaccctgtc ctagcaccct ggagagtga catgccctt ggtcctggca 2400

gggcgccaag tctggcaca tgttgccctc ttcaggcctg ctatgactg gaaattgagg 2460

tccatggggg aaatcaagga tgctcagttt aaggtacact gtttccatgt tatgtttcta 2520

cacattgcta cctcagtgt cctggaaact tagcttttga tgtctccaag tagtcacct 2580

tcatttaact ctttgaact gtatcatctt tgccaagtaa gagggtggc ctatttcagc 2640

tgctttgaca aaatgactgg ctccgactt aacgttctat aaatgaatgt gctgaagcaa 2700

agtgcccatg gtggcggcga agaagagaaa gatgtgttt gtttggact cctgtgtgtc 2760

ccttccaatg ctgtgggttt ccaaccaggg gaagggtccc ttttgattg ccaagtcca 2820

taacatgag cactactcta ccatgttct gcctcctggc caagcaggct ggtttgcaag 2880

aatgaaatga atgattctac agctaggact taaccttgaa atggaaagtc atgcaatccc 2940

atttcagga tctgtctgtg cacatgcctc tctagagagc agcattccca gggaccttgg 3000
 aaacagttag cactgtaagg tgcttgctcc ccaagacaca tcctaaaagg tgttgtaatg 3060
 gtgaaaacgt cttccttctt tattgccctt tcttatttat gtgaacaact gtttgtcttt 3120
 ttttgtatct tttttaact gtaaagttca attgtgaaaa tgaatatcat gcaaataaat 3180
 tatgcaattt ttttttcaaa gtaaaaaaaaa aa 3212

<210> 4

<211> 492

<212> PRT

<213> Homo sapiens

<400> 4

Met Ala Leu Asn Ser Gly Ser Pro Pro Ala Ile Gly Pro Tyr Tyr Glu

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

65 70 75 80
 Thr Lys Lys Ala Leu Cys Ile Thr Leu Thr Leu Gly Thr Phe Leu Val
 85 90 95
 Gly Ala Ala Leu Ala Ala Gly Leu Leu Trp Lys Phe Met Gly Ser Lys
 100 105 110
 Cys Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn
 115 120 125
 Pro Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp

130 135 140
 Glu Asn Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Val
 145 150 155 160
 Tyr Ser Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp
 165 170 175

Asn Glu Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn
 180 185 190
 Asn Phe Tyr Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser
 195 200 205
 Phe Met Lys Leu Asn Thr Ser Ala Gly Asn Val Asp Ile Tyr Lys Lys
 210 215 220
 Leu Tyr His Ser Asp Ala Cys Ser Ser Lys Ala Val Val Ser Leu Arg
 225 230 235 240
 Cys Ile Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile
 245 250 255
 Val Gly Gly Glu Ser Ala Leu Pro Gly Ala Trp Pro Trp Gln Val Ser
 260 265 270
 Leu His Val Gln Asn Val His Val Cys Gly Gly Ser Ile Ile Thr Pro
 275 280 285
 Glu Trp Ile Val Thr Ala Ala His Cys Val Glu Lys Pro Leu Asn Asn
 290 295 300
 Pro Trp His Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Phe Met
 305 310 315 320
 Phe Tyr Gly Ala Gly Tyr Gln Val Glu Lys Val Ile Ser His Pro Asn
 325 330 335
 Tyr Asp Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln
 340 345 350
 Lys Pro Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn
 355 360 365
 Pro Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp
 370 375 380
 Gly Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala
 385 390 395 400
 Lys Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr
 405 410 415
 Asp Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly

420 425 430
Asn Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser
435 440 445
Lys Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly

450 455 460
Cys Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe
465 470 475 480
Thr Asp Trp Ile Tyr Arg Gln Met Arg Ala Asp Gly
485 490

<210> 5

<211> 27947

<212> DNA

<213> Artificial Sequence

<220><223> Recombinant polynucleotide

<400> 5

gcagagtcta agaaatcgct gtgtttagcc ctgcacctgg gcactgtcct cacgggagct 60

gtgtggctg ctgtcttctt ttggaagttc agtaagtga gggagcctcg atcccacat 120

gtgctcctgc agtccccagt gctctgagcc agaccctgct ctctgggcta ttgagacctc 180

tggaggccct ccgtgaggtt cctctcttac ataacgaggc tgtctctctt cccttctctt 240

gtttagctat gagattgaca catcatgggg aaagcattta gaatgtacc agtgctttgg 300

ggtgcttggg gccaccacgc actgtgagca caggttcttc taccttgggg ccacaccag 360

ttacctgtat ctactgcac agcagtggct gttggggacc agggccaccc ctccatgtcc 420

cacctcctgc aactgcagcc tgagccttcc catcagcctg ggggtggtga gacccatgtg 480

ccattgtgga tccttcaagt tacctgtgtg gcagagagga cgtgtgagt cgtccaac 540

ccaaacactg agagggtcct tccattgcc ccacggaag taagggtccc cagtgtctaat 600

tccacttata ctgtctggtg gcaaggacac ttctctcct tattaagtg ggggattggc 660

tgggtgaggt ggctcacgcc tgttatccca gcactttaag aggccaagc aggtggacca 720

cctgaggta ggagttttag accacaagc tggccaacat gttgaaact catctctact 780

aaaaatacaa aaattagtca ggcgtggtgg cgtgcacctg taatcccagc tacttaggag 840

gctggggcag gaggatcact tgaaccagg agttggaggt tgcagtgagc caagattgtg 900

cccctgcact ccagcctggg tgacagaatg agacttcac tcaaaaacaa aacaaaacaa 960

aacacagtgg ggccaggagt tggaggctgc agcgagctac agtaatgcca cgggtttcct	1020
cactccatga ggctcattgc gtttctcagc ctgaaggga cctctcttct gttttctctg	1080
caagtgggca gcaagtgtc caactctggg atagagtgcg actcctcagg tacctgcatc	1140
aaccctctta acttggtgtga tggcgtgtca cactgccccg gcggggagga cgagaatcgg	1200
tgtggtgagt cagccttgac cttgggaagg gactcctctg ctcaccttgg agacagcagc	1260
cgggtccagg ggccttttgg tgactgggcc tggcgtgcgt ccagtacgt gacacatgat	1320
gtcattgaat cctgctcca ggctgagccc tggggctcag agaggtttgtg tttccggccc	1380
aacctcacc agcagggtggg agatgacagg gccaccgagg actgtgtcat tgggaaccaca	1440
cgtgctctga actgccacag gaagtcagtt aagatgagca aactgtttat aaagtggag	1500
atgcaggcta ggaacgggtg ctcatgcctg taatcccagc actttgggag gccgaggcag	1560
atggatcacc tgaggtcagg agtttgagac cagcctgacc aatatggtga aaccttatct	1620
ccactaaaaa taaaaaatt agccaagcgc ggtggcgggt gcctgtaatt ccagctattc	1680
aggaggtga ggcaggagaa tcacttgaac ctgggaggcg gaggttcag tgagctgaga	1740
tcacgccact gcattccagc ctgggagaca gagctggctc aaaaaataaa ttaattaatt	1800
aaaaacaaaa ttggagatgc actatgttat tttcaaaca agctgccttt aaagatctat	1860
ctgttgtcac aggggtgggt catctgtttc attttatctt ctgtggttta tctatttatt	1920
cattttaatg aactaggaag cattgtcctt atttatggca taccacatga tgtttggata	1980
cgtgtatgcc tgtggcatgg ctaagtcaag ctagaacatg ggccttacct catatacgtg	2040
tcttattaag aacacataaa acctactctt gtagtgattt tcaaatatgc aacatatagt	2100
ttattaactg cagtcactat gatgtacaat agattgctcg aacttattcc tcctgtctaa	2160
ctaagatctt gtgacctctg accaaccatct cccagtggt gtcaccccc gcccccagcc	2220
tctgatagct gcctttctac tctctgcttc tgtgagtttg atgtttatac attccacatg	2280
taagtggcct catgcagtgt ttctgtctct gtgtctggct tgttcactta gcgtaatgtc	2340
ctccagcttc atctatgttg ttggaaatga caggatttcc ttctttcttg tggctgaata	2400
gtattgcctt gtgcataaac accacatttt ctttatecct tcattcactg atggactctt	2460
aggttgatgt catgtcttgg ctgttgtgaa aaatgccgca gtgagcgtgg gcgtgcaggt	2520
ccctcttcaa cacacggatt tcctttctt ttgatataaa cccagcagtg agattgctgg	2580
atcacatggc agttctgttt ctcacctttt gaggaactc catactgttt tccataatgg	2640
ctgtagcaac ttccactccc acccccacgg tgcaaagtct ccatttctct tctacaacct	2700
caccaactcc tgttattttc catctttctg atagtagcca tttgaagagg tatgagatga	2760
tacctcattg tggttttcat ttgcattttt atttgtattt ttcataatg tttgagggtg	2820

atttcaaggg tagttagtga ctggaacagg gaaacgatcc tgagtatgag ggttgtgcta 2880
 atcatccccc tctgtccagc tgcgtacgga atggggctct gcagatggca gggagctggc 2940
 tcgtttctct ttaagagctg ccttttactt ttcttctctt tcttttaaaa cttatttctt 3000
 ggccggacgc agtggctcat gcctgtaatc ccagcacttt gggaggccga ggtgggcgga 3060

 tcacgaggtc aggaattcca gaccagcctg gccaacatgg tgaaaccccg tctctactaa 3120
 aaatacaaaa attagccaga cgtggtgggtg cgggcctata gtccagcta ctcgaggagc 3180
 tgaggcagga gaatcacttg aacctgggag gaggggggttg cagtgagccg agattgcgcc 3240
 actgcactcc agcctgggag acagagccag actccatctc aaaaaacaaa aaaaagttaa 3300
 ttccaagca cagccatgta ttccaggctt gtggatcagc gttggtgggtg gtgtgtgctc 3360
 tcatacttta gtccagcta agcacactct gacatgttta cactagaacc atttgTTTT 3420
 tctagaaata gaaatttcag aattgtagag tcagaggact taccagaaat ctcttaggta 3480

 gtctctctcc cctccctcaa gtgcagtcct aacctcttgg agttttctgt agaaaccaca 3540
 agcctcagag ctggccgaga attctagcca aagatttttc catgccaaag taatccccc 3600
 tctcctaagg gccatccttg gtggggactg gtttctgtt aagccctcgc tgtcagtcct 3660
 ggctgtggaa ttctctgttg aggagcactg gcccggtggag ctcgccctc gtgccggcct 3720
 tgagcaggcc caagtgttcc gtgttcttga tacctttctt ccagcacagt ctgtcttccc 3780
 agaaaaaggt ttgcacttga aaatgatgca ttgtctgatt aaacatagtt cttttgcttt 3840
 atttggtttc taaaataaag tgggagtttt tgagattgag taacttgagg ttaagatagc 3900

 acgtggaatg gctttttctt ttctttctat tttttttttt ttttcttgg agacagggtt 3960
 tcactctgtt gccaggtctg gactgcagag gcatgacctt ggctcactgc aacttcgatg 4020
 tcttgggtt aagcgatccc ccagcctcag cccccaagt ggctgggact acaggtgctc 4080
 gccaccacac ctggctaatt ttgtattttt ttgtagaaaa tgggtttcat caatgttgtc 4140
 cagactggtc tcgaactcct gacctcaagc aattctcttg cctcagctc ccagactgct 4200
 gggattacag gcgtgaacta ccacgcctgg cctggaatgg cttttgatgt tctctatgt 4260
 gcacatgttg gtgaataaac accaacaag tccttatgtt acctgaagag ttgctctctt 4320

 cttaatatTT aagtcgtatt tatttaaata cttaatatgt tgtacactat taaagtatta 4380
 ttaggtcaaa atcaaggaag taaaaaggg tatgtgtga aaaatctctt ctctcttgtt 4440
 ctgcttactt acctaccccg catccccca tacaccagc acacacacac acacacacac 4500
 acacacacac acacacgcat cactccata catgccacc tgtttaccag ccaatcacat 4560
 ttcttggggc aactcatctg agttgcttct ctttcagag agtttttgca taaagaagca 4620
 caggtatttc tgcgttacca tgacctatt tccagtggt tccagccag ttgactctcc 4680

tgcaactggat accatcctgg acagcattcc ttagggaaat gagccccctg ttttttccca	4740
ccatggcaca gttggctcctt tgcatggacg caccattatt gccctgtct cttcttggtg	4800
gaccttaagg ttttctccat ctttttctg taacacacac tgctccaagt gtgtgagcat	4860
atcagtagga aacgcttcca ggagtagaac tgctaggtca gagggcgtgt ggatctgtaa	4920
cctgacagac ctagaccggc ttcagtttgg ttttatccag tttccatatt gattattcat	4980
ataaaaggaa acagacaaac ataacgctgt gcatgtattc tctcttagac cagaacaggc	5040
atagggtgca cttttaattt gtccatttcg tagagtagaa attgtttttg ctgaaatgaa	5100
caccttagga tgctgaagaa tatgacccgt cccatggaaa acattcaaaa atgtgtgtag	5160
cgctttcttc ccaaggggtgt gtgtgcgcat attttaacac taattcactt tctacttccg	5220
ttgctatcct ttctgtgagt ctttctcaga atctcagaaa agaaactaaa ttgttcactc	5280
tagttatcaa tgctgtactc tatacctgga atttgctaaa agggcagatt ttaagtattc	5340
tcaccacaga aaagagaaaa gaaaatggta attatgtgac gtgggtggaca tgtaactag	5400
ctttattatg gtgagcattt cacagcggat atccagtcac cacgctgtac acattaaaca	5460
tgtacaattg ggtttttttg agacaaggtc tccttctgtc acccagtcgt gagtgcagtg	5520
gctcagtcac ggctcattgc agcctcgacc tcctgggctc aatccatcct tccccctcag	5580
cctcctgaaa agctggggcc acaggcatgt accatcatgc caggctaatt catatatatt	5640
tatatTTTTT gggtggagatg gggttgggtct cgaactctgg gctcaagtga tcctcccgcc	5700
ttgcccttcc aaagtgtcga gattacaggc atgaaccaca gcaccaggcc tacatgtaaa	5760
atTTTTtatt gtcaactata ctttgacaaa gctgagaaaa aaaatcctaa tatttaaaaa	5820
aaaaaaaaa aggactagct tgagaccttt tccagctctc tggcttatca gctgccgtct	5880
cttcggggtg catagatctg gaagggaag aaaatcccta aaattacca caagccaaga	5940
atgaagtgtc tccttttgag ccacagtggc agttttgttt ttaatcatag aagtgtattt	6000
tgagccgggt gtgctggctc acgcctgtaa tccccgact ttgggaggcc gaggtggggg	6060
gcggaggggg tggggatcgc ctgaggtcag gatttcgaga ccagcctgac caacatggag	6120
aaaccccgct tctactaaaa atacaaaatt agccggcgtg gtggtgcatg cctgtaatcc	6180
cagctactca tgaggctgag tcaggagaat ctcttgaacc caggaggtgg aggttgcggt	6240
gagctgagat catgccattg cactccagcc tgggaacaag aaaaaaaaaa agaagaaga	6300
agaagtgtat tcatttcagt tacttttaaa aaagtgaaca gactttatat ttagagcgg	6360
ttttaggttt acagaaaatg aaacagacag ggcagcgagc tccttgtact cctccccagc	6420

acacagttgc cctgttatga acatcccaca tcagtgcctg gcgttcatta acaccgatga 6480
acctgatgca tacattatga tgaactgaag tcctggactt caccctttct cttgtacagt 6540
tctgtgggat ttgacaaatg cataatgctg tacagccaca atgatagtat cgtccagagt 6600
agttctcctg ccttaaaacc tcttttgctg cacctgtttc tctctcccca ctcaccccag 6660
ctatctgacg ttcttagtgc ctccgaagt ttggtctttt caggatgttg tagcgttgga 6720
atcatggagt atgtagcctt caccacatac accttccttc actttgttgg ctctctttac 6780
ttagtaatat gcattcaagt ttcttccatg ccttttcatg gcttgatagc tcatttcttt 6840

ttagcaccaa ataatatcc gttgtccaga ttagcaccaa tgtttatcca ttcattgtaac 6900
ctgtgaccga ctcacagata ggatgtggaa tcactacca cagaggcatt agacaataat 6960
cagacccaag tcatttcatg ggggaacaag cccacaggta ccagactgtc cagttagtca 7020
gggccactcg taggaagtaa gaagagaggc tagagcatag ccaggctctc actttatact 7080
ttaagcccat gtgtatttct cccaaaccac acagcattgt ttccatgctt tcagctttgc 7140
atgaataacg tgatacttga acgcatcatt taccacttgc tctctttccc acagcgtgt 7200
tttcaagctt ctctctgttc atgatgctct gcttaaccct taagctgcat gggattctgt 7260

tctgtgaata ccccccccc atgtattatc ctgccagca aaaagtcctc aaaactctgg 7320
atggtggtta cctctaggga gggagagaag agattgggaa tagggagcga cttcaacggt 7380
gtttgtaatg ttttgtttct ttaaataaaa gagctgagat catttcagca gaatgttgat 7440
ttagagtctc ctggacaatt tgttgcctca agtgcctctc taaagagcac tttaaaaaa 7500
aaaacctttt atcttattat ttatttattt atttattgag acggagtttt gctctgtcac 7560
ccaggctgga gtggagtggg gtgatctcag ctactgcaa cctttacctc ctgggttcaa 7620
gcaattcccc tgcctcagcc tccaagtag gtgggattac agatgcgtgc caccacactt 7680

ggctaatttt tgcatttttag tagagatcgg ttcttccatg ttggccaggc tgatctcaaa 7740
cgcctgacct caggtgatct gcccgcttg gcttccaaa gtgctggtat tacaggcgtg 7800
agctaccatg cctggcttat cttatatatt tttaaaaaca gcttattgag atctaattta 7860
tgtaccataa aattcaagta tataattcag tgcctttata tataaacat atatatgaaa 7920
tagcttattg agatataatt ttttatataa aacagcttat tgatatgtaa tgtatgtacc 7980
ataaaattta aatatataat tcaactggctt ttatatattc acgaatatgt gcaactatca 8040
ccacagtcaa ttttagcata ttttcatcag ctcataaaga aacccaagc ccttgaacta 8100

tcaccccata tccctctctc cagcccgctc ctctactca taagcaacca ctaatctact 8160
tagtgtctat agatttctca ctctaggcat tccatgtgag cgggatcatg caatacgtgg 8220
gctcacacaa tataagtggc attccatgtg agtcggctca tgcagtatgt ccggctcctt 8280

tcactgagca taaggtcttc agcactcatc caggttgcag cctgtgtctg aatttcattc	8340
cctctttctgg ctgaatcgta ttccattgtg tatcttggac atacctatt ctgctcaccc	8400
agccgttggg gggcgtttgg agtgttttcg cctttcagct gttttaagag ggttgcagtg	8460
aacatttgta caagttttgg acccaatgcc tgttttcaat tctcttgtgt agagagcact	8520
ttttagcaga aaaagaatag atttgtggcc tccctttgtg tgcggtcagt gccttgagaa	8580
gagtgaactg tgctgccacc tccggagccg tggagagcgc ggggcttggg tagcagctag	8640
gacgatacaa gttagggacaa ggccagggtgc aatggctcac gcctgtaatt ccaacacttt	8700
gggagaccga ggcaggggga tcacctgagg tcaggagtgc aagaccagcc tggccaacat	8760
ggtgaaaccc catctctaataaaaacagaaa aattaactgg acggggtggt ggacgcctgt	8820
aatcccagct actcgggagg ctgaggcagg agaatactt gaacctggga ggcggaggct	8880
gcagtgagtg gagatcagac cactgcactt cagcctaggt gacagagcga gactccgtct	8940
caaaaaaag aaaaaaaag aaagaaactc atggataatc ctccctctcg tgcagttcgc	9000
ctctacggac caaacttcat ctttcagggtg tactcatctc agaggaagtc ctggcacctt	9060
gtgtgccaag acgactggaa cgagaactac gggcgggcgg cctgcaggga catgggctat	9120
aagtgagtat ggggcagcac ccgccgagt acagtaacag acagcagaaa cacgagaaga	9180
ccctctctct gcctccctgt gaaagcaccg gcacatgagt gctggggaca attgtcacct	9240
tccaaaagct gagccctata accagcaggt ggaatttgtc ctgctagggc tgtgccagc	9300
acacagacct tggctcactg ccaccttgcc ctgcctctc cttggcctct atagactcct	9360
ggttgcctcg gagtgccag tgctgtggtc atctggtcag aggggtaggc tgaggcggtt	9420
aggtgcctct tttccaagg tgcctctcag ccagggtcca ttacactccc tgggtagagg	9480
ttggaccaga acagctggcg aggagggttg ggctggggag agcagcagag acaaactctg	9540
tgccagtttc acttcattcg ggagccatgg aagccttttg agctggggag agaataatc	9600
aatcagactg atacttaaaa aatgtcattc ctgctcgtag ctctgaggga aggtgggaag	9660
gcttaacagg gtgtgtgtcg cctgacagtg attcctaacg ggggtggggc ggtggttacc	9720
atttaccagc actgcctggg gagatgcggc agccctcagg catcggggga gaggggtgta	9780
ggatgctact gccactttgt tttccatggg aggggtccca ggtgatttct atgcaacttt	9840
aggttattca atatgccagt ttccagaatg aattaccact cggtagagaaa gttagcatct	9900
tagctagtca ctgtgacatc cctaacagc aggggtgaat tacacagcaa agcccccca	9960
tcacagtcca ggaacctggg ggaattgata actggggcca tgttaacatc tgtacctttt	10020
attagattaa atgtgtgtat gattatacaa tctatgtcc ttctcatagt ttcttgatcc	10080
taacctggat aagaaacacg accaatgaag gaattttgtc tgacacttta gggttattga	10140

atcgaaaaat cgttacaata ttctagcact tggttagaac gtgtgatttt ttttcctaaa	10200
tgctaagggtt tttccctctt attctgaatg tcgtatgagc ggtattatga catagtatag	10260
gatttgtgtt tgcttatgcc ttaaccatta tcacaaataa ggttttcttt tttaggaata	10320
atctttactc tagccaagga atagtggatg acagcggatc caccagcttt atgaaactga	10380
acacaagtgc cggcaatgic gatattctata aaaaactgta ccacaggtat gcagcaattt	10440
cttcttgaaa aattttggaa tgaaatcaac taggagacac catggggaat cgttgtcctg	10500
agtctgattt ctctgagctg caatactcgg tctggatggg ttttgattg ggaggagatt	10560
agagtctgac caggcctggt tactctaagc agcccttggg ttattcatag gaagtggctg	10620
aggtttctct gctatttcat tttcagcctc taccgtctgc ccttgttggg agcggctcac	10680
acttgaaca tcgacattca actctattta gttttctttc ctcttcagac atttagaggt	10740
gtacctattt tgtcaggggc tggttctagg aatccaagat aatgtctcag tgtcccagcc	10800
agggtgaccg gctcattcca gtttgccagg gacttcactg gcttgagcaa gggaaagtctt	10860
gctccattcc aggcagctgg gctggctggt cccgttagcc ccaaccccg gacagcagtg	10920
ccagagggtg ctctgtgagg gatgggcagc attctggcgg cctgggaatg agttgtggtg	10980
tttcaggggg gtagaagtgg gtacaagcca caggtcacat gatgagtggc tgacctggct	11040
gggagggcag aagaggggat ggacttaggc tcttctttt gctttgcaca tatttaggat	11100
gtttgcagac ttgctatgat tgttgctgtt atgtgttttc tgatgtgaaa gatacacagt	11160
gtcctttgcc catgagctct ccttgccctc cagggtccca gggcttatgc ctggtgtcta	11220
ggcatcacct cctgcctgc cagggtccag gtgctgcatt tcgggggagg atgaactaat	11280
caccccgcc cacttttct ctgagtggga gcctggggca ggtttgcatt cctggaggcc	11340
gctggtggag gggctctggg gcctgacttc cactgcagcc tgctgtcctg gggaatgtgg	11400
cagggcaagc ccagtgggga gggctgtgca cggccagggt cacccatcaa aacagcaggg	11460
ctgcggtttg tcctgttga gaagctaaac acagctgcct gggcactttg taaatgctga	11520
gtggttcttt gtctttctgg gttacacacg gaatcaggga gccaagtcca gccgggcagg	11580
gacgggggga ggggaggagg tgctccctc ccttgcaag agccttgga actcacaagg	11640
aggctggagg gcttgaaga aagaagagaa ggccattgtc tggtaggctc tattctatct	11700
cgggtggtgtt ggtgggggga ggcgcacttc ttttctctt tctgtgcagc agttgccctt	11760
tgatgcctga gtcttggct tgttttctgt cgggcttctg tgaataacca catgtgccct	11820
ggcgtgtga ccacacaggg ctatccctac cgaccttagg attcttagga aatgtcttct	11880

cttaaagggg acatgtcttc acttggccgt gtcagtcccc cagagccaga gtccacctgg	11940
aatgcacctg tagtcactga gaacccgggg ggtgtgcctt agtaagaagg tgcaggaag	12000
gacctattat ttagggcct gggctcctgc aaggtggttt ggggggtggtt ggaggaagca	12060
gagatttgct ctggattgga tgctgtcagg aagcaggggt aattctgtga ggctgcttta	12120
ttattttttt tctaggagga ggttggaatg aggctaggct aaagctgtga ttggtaaaga	12180
aacgtccgtc gctcaagtta gccaggacag gaggagacat cagatcgtga ttttgtggtt	12240
gtgagcaciaa gggtccgtgt ctgtctgttc agacatcatt tcggaggagg ctccttgtgt	12300
cttgccccat ctcaggcatg gaggggccta gtccgatatt gacgctcagt gaaataattc	12360
aggttccgca gagcacacgg ccagctatc agggcgggcc agctctgcat gccaggggcc	12420
gcgtcttccc ttctcagcat agcctgggaa attcactgca ggacaaaatg catcagttac	12480
ttctcttcca tccataacct gggatgtttg actcccaat gagtaactct tacgtttctt	12540
ctaactctag ggaaactatt ggttatattg ctttcaacac taaaaattta aagcagttat	12600
aggagcccag aggtttccaa atggcttcct taaaaattag aagatgattt taaattccaa	12660
gaggaaaaac aaaactagca ttattgtata cttacctca caaccgtcct aggagctggt	12720
acaattttaa gagaggttaa gtaacttgcc caaggtcaca ctgtggggat gtgagccgcg	12780
taccttggtc cagtgtctgg tctttgccac tgtccctata tggatttact taccttattg	12840
gagttgtaac tagcagacc ttctatgtct cagaagacag gagagggaac atcggaagaa	12900
atgactgatt tctaagcatg tgagaggcag gtgactccgc actatcgtga ccagaatttc	12960
ccctgttctt ttgacgtga tgctgttct tcaaaagcag tggtttcttt acgtgtata	13020
ggtaagtcca tctggagtc cccttttgat acttctaact aggaaaagct ctctactttc	13080
agaacagtac tccctgtgtc tctgggggcg tgggaggga gaaggtgggg tcacgggttg	13140
gaatgtgccc agcggcgtct cgctctttcc aaggagctcc tggtttagat ttccatggcc	13200
tgtagacacc ttacgccttg ggtccaaggg acacccctg agatcaggca cgctcaagaa	13260
gctgacaaag ccctacactt tatgccacc atgagctgga ggcccgcag gtctctttct	13320
ccagaaagca aaggggggtg gcgttagtga gccctggcag ccacctaacg tggacttgga	13380
gcatctgcgg ggctgtggtc cagcaccacc gtgtggccac caggltgtca tcagccagtg	13440
ggaccgggga ggagggacaa gaccagagaa caacagtgt cttgcctctt ctctcctgaa	13500
ttttggacgg tggcctagac ttgggtgtcc ccatctctgt gtttagagt cttacagttt	13560
ccaaactgtt tgcaaatgtg gaagccaccg tcctctcct ctgggatggc ccagtgtgt	13620
cgtggggccg tggctctgag ctcagctttt catttgaaga ggtggaagga gctgacaccg	13680
tccatcccg gcagggtggt ctcaggtctt ctttaggtcc tgagtggggg tccagcacag	13740

ccccaagggt gcgtggcacc cgccctgccc tctgcccattg cactcatctc ctggtggaga	13800
agacactcac acacaggaag caggaaggc agcagacctc actcaccct caccctca	13860
ctcaccctc actcaccctc tcaacctctc attcaccacc caccctctg cccctcact	13920
caccctca ctcctcaac cctcactcac ctcctcactc cctcaacct cactcactc	13980
ctcactcct cactctccc ctcactctc ctcacccca cccgtcacc tctcactca	14040
cctcctcacc cctcactca ccttcacc ctcactcac cactcact ctcactcac	14100
ccctactca accctcatt caccctcac cccctcactc accctgcac cccctcactc	14160
acccttcat cactcacc cctgtcacc ctcctcactc aaccctcac cccctacta	14220
atccctcact cctcacc ctcagccct cactcacacc ttacctcct cactcacc	14280
ctcaccct caaccctta cttaccct cactcatcc ttaccctc actcacc	14340
tctctcacc attcacc cactcatgc cttaccctc tactcact ctcactcac	14400
accctcacc ctcagtcacc cctcactca cccctcacc cctcaatca tgcctcact	14460
cctcactca cccctcacc cctgaatta ctcctcacc cctcactca cccctcact	14520
caccctca cccctcacc caccctca cccaccctc accaccctc taccctctt	14580
accctcacc cccctcactc accctcacc cctcactca ccactcacc caccctcac	14640
ccaccctc actcactcc tctcctc actcaccctc taccctc actcacc	14700
taccacc ctcaccacc cctcacc ctcactcacc cttaccctc ctcactcacc	14760
cctcactca cccctcacc cctcactca ccactcacc caccctcac ccaccctc	14820
actcactcc taccctc actcaccctc taccctc actcaccctc tctctctc	14880
actcaccct taccctc actcaccctc taccctc actcaccctc tgcctcctc	14940
actcaccct caccctca cccctcact caccctcac cccctgccc ctcactcac	15000
ccctgccc ctcactcac cctcacc ctcaccct cactcatcc ctcactcct	15060
cactcacc ctcactcct cactcacc ctcactcct cactcacc ctcactcct	15120
caccacc ctcactcact cctcacc ctcaccct cactcacc ctcactcct	15180
cactcacc ctcactcct caccacc ctcactcact cctcacc ctcaccct	15240
cactcacc ctcactcct cactcacc ctcactcct cactcacc ctcactcct	15300
cactcatgc ctcaccct cactcact ttactcct tgcctatcc ctcacttacc	15360
cctcacttc gtcaatcacc cccaccctc gtcaatcacc cctcactt ttactcacc	15420
cctcactca ccccttact tctcactta cctcctcacc cccactcac cccctcacc	15480
ccactcacc cctcacc acactcacc ctcaccct cactcacc ctcaccctc	15540
tactcctc actcaccctc taccctc acttatccc taccctc aattaccctc	15600

tcacccctc aattactccc tcactctttc aattaccac tcacccctc acctctcac	15660
tcctcactca ctccctcact cacccttca ctttctcact cactctctcg tctctcacc	15720
ccctcactca ctccagccc tgccctccc atcttcttt tctttgtgtg agaactctggg	15780
gtccctgagt gggtcagtc cctccaagac tcaaggagtc cccagggcct tgttatccag	15840
aacaccccca cctgggtccc gggagacccc atgggatcac aggagtgttc aggggaagtgg	15900
tgcttcctgg gtctgggtgg gctggagggg catctccct tcccaagag gagacccca	15960
ggagccccc aagtcaccc ccagcagtgg tgccctgcc ctgtccttgc agcctgggag	16020
acccttggga gggcgggcg ctgggtggct gggcggttc tctgtgtctc acccactgg	16080
ctctctgttt gtcatcctca gcctgcgggg tcaacttgaa ctcaagccgc cagagcagga	16140
ttgtgggcgg cgagagcgcg ctccggggg cctggccctg gcaggtcagc ctgcacgtcc	16200
agaacgtcca cgtgtcgga ggctccatca tcaccccca gtggatcgtg acagccgcc	16260
actgctgga aaagtatgcc agggcgggcg cggcggggt gggggtcag ggctggccta	16320
cagccacct gtgacctga gcaggtctca acccttgcag ccccgcatc cttgtgttta	16380
aatggggaga gtattgcacc tgcttcttag ggctgtgaga catcaagtgc gctcatgcca	16440
ggcagtgcac ggctgtatgc actgagtgtc cctgcacgc agggcacagg gtgcaggtgg	16500
aacattctcc acgatgtgc cgtgaccagc gtctctcca gccactgtcc tctgagctct	16560
gtctgccct tgagcaaagc cctgcccc tgaggtatcc tgtctcggg acgctagtcc	16620
caggagagg cactcaga caggctcag gctgccctgc tggaaagtcc ctggggttaa	16680
gcgttcttgg ccacagcatt gctcatgcag agggtaggt aggggtgagg ctagccgtga	16740
cagtattagc attatggac gctaccacc cctcccttt tccttaaca catagtctt	16800
ttggtcacat gctgcttgg aggagcctc acttgcgga tgtattttt tgccttagag	16860
agaggtgaa ctgggttga ctgttgccc agcctctct tctgcgtgc ccttagacga	16920
ttcactcaac gtctctgac catggcatgt acaactataa gatggcatg ccttctct	16980
ctcggctgt tatgaagtc aaggaagca gggtgttac ccaagggtgc tccctctct	17040
cccccttc acaccccag gtgctctggg cctctagga actgggttc tctcaagggc	17100
tgttaccaa ggtgtctcc ttctctccc ctctcacac cactgggtgc tctgggcca	17160
ctaggagctg gattctctt aagaggaaa ctcttgata aaggaaatgg ttgtattgat	17220
atcgacaag tctgttcatt agtatccatt tattaagcac ctacatgtg ccaggaaatg	17280
ctttgctga caaagaaaa taaggccag tctgctaga aatggcctt aaacccagg	17340

gagggatgtc ggccattgt ggggtgctgca gattccttga aggtgatgca agagccagaa	17400
agaaggatga tgtggggggc tgaggcaggg agtcgggggtt gggggagtgt gggggagaag	17460
gggagaccga gcacctcttc cactatctcc ctgtgtgggtt tttggtgaac catcctgcct	17520
ctgggtgtct tgccctccagc ttctgacgtt ggaagtcat ccactgagag ctctgtgttt	17580
atggctctga gatactgagt cttctttctc tcccagacct cttacaatc catggcattg	17640
gacggcattt gcggggattt tgagacaatc tttcatgttc tatggagccg gataccaagt	17700
agaaaaagtg atttctcatc caaattatga ctccaagacc aagaacaatg acattgcgct	17760
gatgaagctg cagaagcctc tgactttcaa cggtagctgt ggctcaggct tggcaagcag	17820
gttggcagaa tcttaaagag atgttgattg gaaatgacac ttgtgctatg ccaaattgaa	17880
gggaggcatt tgcgttgagc gaggttagcg tgcagcgggt ggccaatggg agaggctcac	17940
agaggctaag agcacctgcc gcattttggg ggaggcagca gccaccacat ctgttctgta	18000
ctgtactgag tgggtggtgat tcaagccagg catggaaaag gctagaacag ggctttccca	18060
ctgcagcacc cttagacatc ggggtggttct ctgtttagg gctctcttgt gcctttagg	18120
atgtttaaca gcgtccccag cctctacca ctggaggcca gtagctacca agctgtgaca	18180
accagtgttg cctgctgaca ttgccaaaca tccgctttga ggcaaagtca cttccagttg	18240
agaactactg gcctaaaatg tgtaaagatc cttagatttt aaagatacat tctaaaacca	18300
agttgcttaa ttcaggacaa acatgctttc tcttagcctc ttattcggtc ccactctggt	18360
ccatccaagg gtctggaatg ttctagcccc atgtggatac agaagaagca aaacctcagc	18420
cctccctaca gcatgtctgt attcacattg ggaaatggtt cacatataga agagcgaatg	18480
cctgagcaat ggctgtgtgc ctctggggcg aaagctgact ccattgactc catcggtttt	18540
ttggctgttg cctcctgtgt gtctttcccg tcttgatcac ctggagatat gtaattttgg	18600
aagcagagct agcaaataat tcctcttata agcagagcta gcaaataatt ctacttataa	18660
gtagcataac gtcttgcttg ccagaaggag aggtctggca gggggagaaa gtgagaatgt	18720
gggacttggt gggatgcagg gtcctctggg cagggtggcc agggtgccag gcccagcagc	18780
ctgcatgttg gaaggccagg tggagacata ggtgataccc gcctggctca ctgtgttttc	18840
tcttcttgaa acagacctag tgaaaccagt gtgtctgccc aaccaggca tgatgctgca	18900
gccagaacag ctctgctgga tttccgggtg gggggccacc gaggagaaag gtgaggctgc	18960
tcctgggcac acaggactgc agggccca gatggagcat tgggttcgga agtgggaggt	19020
ccaggtttta atcccagttc tactactcaa tgactggatg actttggttg attccccag	19080
tccttgtgcc tcagtttctc catctgctaa gtgggagaaa tctgcccag cctacctaat	19140
acactgtgtt cttatcgtga tcacacagag cagcatgttg aatggctttt gaagtatctg	19200

ggccatacga gtttagaggt gcaggatctc ctgtgttgca ctcatgtga gtttagagct	19260
gccctggaga tcccaccaag gcctgcgtgg ctgagtaca gggggcttgg tgaggacggg	19320
catcctggac ccatggtggc cacatctaag cctgtcctct gccctgataa ccacagagag	19380
aggctctctc caccacttc ctttgaatc tgcatttctc tctgacagtc tttcaaatga	19440
agggagcctg gctgcttcat ttttatggag ggttggagt gcttagtggc aggcacaaag	19500
gttcatttta catattgttt atatccttct caaaagcgtc taggccatac agacaacaaa	19560
tcctttcaaa caaggggaaa agtacaaagg ttgggtgatt tctggggagc gtcaggaag	19620
gtagtggggg gcatcctggc tctcatcag cagaaactta ctacagtaga gccacaggct	19680
gggcaaaaga cctcatggaa tccaagatga agggaatctc gacaaatatt tgtgcgcacc	19740
tgcacctagt acaggctggg tgctactcag gtgctgggaa tgcagaagtg aacagagtaa	19800
gacaaatgtc tctgctgtca ggagctttac ctctcttctg gatgtcgggt gtggggacgg	19860
ggcaggtgtg gtcagacaga tgggagacaa acaactgagc gagtlacttc caaacatctg	19920
agggtgggga tcacaaggtc ccggtatatt tgaaggggtg gtcaggaaag gcttctcgga	19980
agaggtggca tttagctga gactcaaatg gcaaaaatgt gtacacatca aaaaggctag	20040
tgcattgtatc ttcaggtgtg gtcaaggggc caaggagggt ggctggggcc agattgcata	20100
ggtccttgtg gattatgggt aagacaccag ctctcatct gcttaggtg gggagatcgt	20160
gagccgggga gtgccatgat ctggcagctg cgtggggagt ggggatgaat ggatggagac	20220
gaggatgatg gtgacaagtc cattgtgtg gttccttgag acaggaagcc agctcatagc	20280
agagtgcggg cgtggatgtg aagagatgag ggtacactag ggctagagcc accagactta	20340
ctgatgggtt gcatgtctgt gggagagaga gtgagaagtc agggacgatg gctttccact	20400
ctgtgctga agccccagg tggcgggtgg tgccattttt caagccagga aatattggtt	20460
ggtgagaatt tggggtggga gaaggtgtga cggagggttc tggttttgca cactaagccc	20520
acggtgcccc gaagatgccc gaggggaggc agcaaagcga gagtgggaaa tgcagagggtg	20580
gcaagtgcag gccgtgtctt gagaagctct aatgtgcagg ggagccgaga agcaggcggc	20640
ctagggaggg tcacgtgtgc tccagaagag tgtgtgcatg ccagagggga aacaggcggc	20700
tgtgtgtcct ggggtggggt cagtgaggag tgggaaattg gttcagcaga accaagccgt	20760
tgggtgaata agagggggat tccatggcac tgatagagcc ctatagtctc agagctggga	20820
atttctttcc ctgaagctga actccagagc tgcattcagc acaggcacgc ccagtgttaa	20880
ggagaatcca ggtttccag gagaggggtt ggtgctggga tgagctgacc ggggcagggc	20940
tggaaaaatg ggctgtgacc atctgtgtag tgcgtgtgga ggtctcaggg aggggaagtgt	21000
gctctccctg cgagagctgc aggcaacact gggagctcaa caagtctccc tgtccttagg	21060

gaagacctca gaagtgtctga acgctgccaa ggtgtctctc attgagacac agagatgcaa 21120

cagcagatat gtctatgaca acctgatcac accagccatg atctgtgccg gcttcctgca 21180

ggggaacgtc gattcttgcc aggttaattca acatTTTTat tctacctttg gtccttacca 21240

gatcctactg aaccccccat gagagagagg gcattcttgg ggtcagcaga gcctcctcag 21300

tgacacggag ccagctcggg gcagtcattg gaagtgcagg ccacaaacag tgcgaacgt 21360

tctgggtggca gaaggaagta cagtcaacaa atcacacaca cctctgaaa aaccggtatt 21420

tggtaaaagt gccagtggaa cagaacaag tatTTtagact attTTaaatt atgaacggca 21480

atttatttag taacttttag cttgaacaga tTaaaattca ggatgggggc tatctctttg 21540

ggggttacat ctctgttacc atcacccctt gatggtggag attcgaagcc cacacagtca 21600

ctcgtaaact acactgcgac ccccgcccc caactcctct aggcctggtc agtgggtgtgc 21660

ggcagattgt gacttgattt tctgtctctt gtaccttgct gtgtcccaca ggggtgacagt 21720

ggagggcctc tggtcacttc gaagaacaat atctgggtggc tgatagggga tacaagctgg 21780

ggttctggct gtgcccaggc ttacagacca ggagtgtacg ggaatgtgat ggtattcacg 21840

gactggattt atcgacaaat gagggttaact atcctgtcct cttctgact gtgttctccg 21900

attcctcgag ccaaagccag acatctgtta ggcgtgggtc tgctgctgga agctgactgg 21960

tgaccactgg tcagcatgaa gcaaactctg ctctctccag ccacagcccc atccccccag 22020

tgccaccca ttgcccattg cctctcactg gcttcacttg catatttccc ctgggtgttg 22080

gatgaaaagc gctggggctc agcttgtgtg aaattccttg gtgtctgtcc aaccacactt 22140

cgttctggct cagctgactc agctgttcca cccaggccac ctacatcaa actTTTTTT 22200

TTTTTTTT agatggagtc tcactgtgtc gccaggctg gagtgcagtg gcacaatctc 22260

gactcactgc aacctttgcc tcttgggttc aagtattct cctgcctcag cctcccaagt 22320

agctgggact acaggcatgc gccaccagc ccagctactt tttgtatttt tagtagagat 22380

ggggtttctc catgttgcc aggtctgtct cgaagccctg acctcaggtg attcacccac 22440

ctcagcctcc cacagtgtg ggattacaag tgtgaaccac ggtgcccggc ctacatgaa 22500

acttttgatt tatagagagc agagggaaga gccggctgtg cccatccttt tctggggcca 22560

tcgagtggct cctgggcagc cccaagggt aggaaggga ggagcagcca gggttctctg 22620

atgccccaga ctcaagcac aggggaaggc tcagggttc catgtgagcc tcatggatgt 22680

ctctgcttag cagagccctg gctttgggca ttgtccagat agggggtgag aaccagatct 22740

tctcatctcc aggacctcag acgtatagtt ttctcagatt tctgtgttt ctggggctgg 22800

gctactagtg gaagaaagca gtctattctg tcttctccca aatctcccag atgcccagtc 22860

tgttgaagga ggagcagaac cagggggcct tccccgtga ggcccacac gtgtctcctt 22920

caaatgacac gcgggactca gggccttccc atgaccatgg ggcccagggg gcgtcacctg 22980

gcccagggcc cagtgtctaga aacagatgac cccaggagga ggaggcaggg caggaggga 23040

gctggcaggg ctgggatggt cagccaggct gaggggcgga ctgcaccag gatggagcta 23100

ggaaatgac caggtgtgtt tggcggctgc aggtgggtcc gcatggctgt gcaggaggg 23160

aagggtctgc tggcaggaga gcagccgggg gagggccaga ctctgctgaa gagatgcctg 23220

tgtgtccggc ctccacatcc gctgcccgt ccttccggag ctctgcccc gccatgctca 23280

gcctgactct gaccaacacg ttggagagaa gaatgatccc tttgtctat taagcttgct 23340

tatttggttt ctaagtgtt catgcgaacc tagaggaaaa aattattttc cacctttgtt 23400

tgtcttaaga aaataacaca cttttttttt tcctatttga acaggcagac ggctaacca 23460

catggcttcc gtccctgacg tcgttttaca agaaaacaat ggggctggtt ttgcttcccc 23520

gtgcatgatt tactcttaga gatgattcag aggtcacttc attttatta aacagtgaac 23580

ttgtctggct ttggcactct ctgccattct gtgcaggctg cagtggctcc cctgcccagc 23640

ctgctctccc taacccttg tccgaaggg gtgatggccg gctggttgtg ggcactggcg 23700

gtcaagtgtg gaggagaggg gtggaggctg ccccatagag atcttctgc tgagtccttt 23760

ccaggggcca attttggatg agcatggagc tgtcacctct cagctgctgg atgacttag 23820

atgaaaaagg agagacatgg aaaggagac agccaggtgg cacctgcagc ggctgcctc 23880

tggggccact tggtagtgc cccagcctac ctctccaca ggggattttg ctgatgggtt 23940

cttagagcct tagcagccct ggatggtggc cagaaataaa gggaccagcc ctcatgggt 24000

ggtgacgtgg tagtcacttg taagggaac agaaacattt ttgttcttat ggggtgagaa 24060

tatagacagt gcccttgggtg cgagggaagc aattgaaaag gaacttgccc ttagcactec 24120

tgggtgcaggt ctccacctgc acattgggtg gggctcctgg gagggagact cagccttctt 24180

cctcactct cctgacctg ctctagcac cctggagagt gcacatgccc ctggtcctg 24240

gcaggcgcc aagtctggca ccatgttggc ctcttcaggc ctgctagtca ctggaaattg 24300

aggtccatgg gggaaatcaa ggatgctcag ttttaaggtac actgtttcca tgttatgttt 24360

ctacacattg ctacctcagt gctctggaa acttagcttt tgatgtctcc aagtagtcca 24420

ccttcattta acitcttgaa actgtatcat ctttgccaag taagatgggt ggcctatttc 24480

agctgctttg acaaaatgac tggctcctga cttaacgttc tataaatgaa tgtgctgaag 24540

caaagtgtcc atggtggcgg cgaagaagag aaagatgtgt tttgttttgg actctctgtg 24600

gtcccttcca atgctgtggg tttccaacca ggggaagggt cctttttgca ttgccaagtg 24660

ccataacat gagcactact ctacatggt tctgcctcct ggccaagcag gctggtttgc	24720
aagaatgaaa tgaatgattc tacagctagg acttaacctt gaaatggaaa gtcattgcaat	24780
cccatttgca ggatctgtct gtgcacatgc ctctgtagag agcagcattc ccagggacct	24840
tggaaacagt tggcactgta aggtgcttgc tccccaagac acatcctaaa aggtgttgta	24900
atggtgaaaa cgtcttcctt ctttattgcc ctttcttatt tatgtgaaca actgtttgtc	24960
tttttttgta tcttttttaa actgtaaagt tcaattgtga aaatgaatat catgcaaata	25020
aattatgcaa ttttttttc aaagtaacta ctgcattctt gaagtctgc ctggtgagta	25080
ggaccagcct ccatttcctt ataagggggt gatgttgagg ctgctggtca gaggacaaaa	25140
ggtgaggcaa ggccagactt ggtgctcctg tggttctcga gataacttcg tataatgtat	25200
gctatacgaa gttatatgca tggcctccgc gccgggtttt ggccctccc gcgggcgcc	25260
ccctctcac ggcgagcgt gccacgtcag acgaaggcg cagcgagcgt cctgacctt	25320
ccgcccggac gtcaggaca gcggcccgt gtcataaga ctggcctta gaacccagt	25380
atcagcagaa ggacatttta ggacgggact tgggtgactc tagggcactg gttttcttc	25440
cagagagcgg aacaggcgag gaaaagtagt ccttctcgg cgattctgcg gagggatctc	25500
cgtggggcgg tgaacgccga tgattatata aggacgcgc gccgtgtgga cagctagttc	25560
cgtcgcagcc gggttttggg tcgcggttct tgttttgga tcgctgtgat cgtcacttgg	25620
tgagtagcgg gctgctgggc tggccggggc tttcgtggcc gccgggccgc tcggtgggac	25680
ggaagcgtgt ggagagaccg ccaagggtg tagtctgggt ccgcgagcaa ggttgcctg	25740
aactgggggt tggggggagc gcagcaaat ggcggtgtt cccgagtctt gaatggaaga	25800
cgcttgtgag gcgggctgtg aggtcgttga aacaaggtgg ggggcatggt gggcggcaag	25860
aacccaaggt cttgaggcct tcgctaatgc gggaagctc ttattcgggt gagatgggct	25920
ggggcaccat ctggggacce tgacgtgaag tttgtcactg actggagaac tcggtttgtc	25980
gtctgttgcg ggggcggcag ttatggcgt gccgttgggc agtgcaccgc tacctttggg	26040
agcgcgcgcc ctctctgtgt cgtgacgtca ccggttctgt tggcttataa tgcagggtgg	26100
ggccacctgc cggtagggtg gcggtaggct tttctccgtc gcaggacga gggttcgggc	26160
ctagggtagg ctctcctgaa tcgacaggcg ccggacctct ggtgagggga gggataagtg	26220
aggcgtcagt tcttttggtc ggttttatgt acctatcttc ttaagtagct gaagctccgg	26280
ttttgaaact tgcgctcggg gttggcgagt gtgtttgtg aagttttta ggcacctttt	26340
gaaatgtaat catttgggtc aatatgtaat ttccagtgtt agactagtaa attgtccgt	26400
aaattctggc cgtttttggc ttttttgta gacgtgttga caattaatca tcggcatagt	26460
atatcggcac agtataatc gacaaggtga ggaactaac catgggatcg gccattgaac	26520

aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggtattc ggctatgact 26580

gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggtgtca gcgcaggggc 26640

gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg caggacgagg 26700

cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg 26760

tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt 26820

catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc 26880

atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcg atcgagcgag 26940

cacgtactcg gatggaagcc ggtcttgctg atcaggatga tctggacgaa gagcatcagg 27000

ggctcgcgcc agccgaactg ttgccaggc tcaaggcgcg catgcccgac ggcatgatac 27060

tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt 27120

ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg 27180

ctaccctga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt 27240

acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct 27300

tctgagggga tccgctgtaa gtctgcagaa attgatgata tattaaacaa taaagatgtc 27360

cactaaaatg gaagtttttc ctgtcatact ttgttaagaa gggtgagaac agagtaccta 27420

cattttgaat ggaaggattg gagctacggg ggtgggggtg ggggtgggatt agataaatgc 27480

ctgctcttta ctgaaggctc ttactattg ctttatgata atgtttcata gttggatatac 27540

ataatttaaa caagcaaaa caaattaagg gccagctcat tcttcccact catgatctat 27600

agatctatag atctctcgtg ggatcattgt ttttctcttg attcccactt tgtggttcta 27660

agtactgtgg tticcaaagtg tgctagtttc atagcctgaa gaacgagatc agcagcctct 27720

gttcacata cacttcattc tcagtattgt ttgccaagt tctaattcca tcagacctcg 27780

acctgcagcc ctagataaac ttctataat gtatgctata cgaagttatg ctagtaacta 27840

taacggctct aaggtagcga gctagctcca cgtggctttg tcccagactt cctttgtctt 27900

caacaacctt ctgcaagaaa accaagggcc tgaattttaa cttcctg 27947

<210> 6

<211> 25333

<212> DNA

<213> Artificial Sequence

<220><223> Recombinant polynucleotide

<400> 6

gcagagtcta agaaatcgct gtgtttagcc ctcgccctgg gcactgtcct cacgggagct 60

gctgtggctg ctgtcttgct ttggaagttc agtaagtga gggagcctcg atcccacat	120
gtgctcctgc agtccccagt gctctgagcc agaccctgct ctctgggcta ttgagacctc	180
tggaggccct ccgtgaggtt cctctcttac ataacgagge tgtctctctt cccttctctt	240
gtttagctat gagattgaca catcatgggg aaagcattta gaatgtacc agtgctttgg	300
ggtgcttggt gccaccacag actgtgagca caggttcttc taccttgggg ccacaccag	360
ttacctgtat ctactgcac agcagtggct gttggggacc agggccaccc ctccatgtcc	420
cacctcctgc aactgcagcc tgagccttcc catcagcctg ggggtggtgca gacctatgtg	480
ccattgtgga tccttcaagt tacctgtgtg gcagagagga cgtgtgagtg ccgtccaaac	540
ccaaacactg agagggtcct tccattgcc cccacggaag taaggtgccc cagtgtaat	600
tccacttata ctgtctggtg gcaaggacac ttctctcct tattaagtg ggggattggc	660
tgggtgaggt ggctcacgcc tgttatccca gcactttaag aggccaagge aggtggacca	720
cctgaggtca ggagtttgag accacaagcc tggccaacat gttgaaactc catctctact	780
aaaaatacaa aaattagtc ggcgtggtgg cgtgcacctg taatcccagc tacttaggag	840
gctggggcag gaggatcact tgaaccagg agttggaggt tgcagtgagc caagattgtg	900
cccctgcact ccagcctggg tgacagaatg agacttcac tcaaaaacaa aacaaaacaa	960
aacacagtgg ggccaggagt tggaggctgc agcgagctac agtaatgcca cgggtgtcct	1020
cactccatga ggctcattgc gtttctcagc ctgaaggga cctctctctt gttttctctg	1080
caagtgggca gcaagtgtc caactctggg atagagtgcg actcctcagg tacctgcac	1140
aacctctcta actggtgtga tggcgtgtca cactgccccg gcggggagga cgagaatcg	1200
tgtggtgagt cagccttgac ctgggaagg gactcctctg ctacacctgg agacagcagc	1260
cgggtccagg ggcttttggg tgactgggcc tggcgtgcgt ccagtacgt gacacatgat	1320
gtcattgaat cctgtctca ggctgagccc tggggctcag agaggttgtg tttccggccc	1380
aacctcacc agcaggtggg agatgacagg gccaccgagg actgtgtcat tggaaaccaca	1440
cgtgctctga actgccacag gaagtcagtt aagatgagca aactgtttat aaagttggag	1500
atgcaggcta ggaacggtgg ctcatgcctg taatcccagc actttgggag gccgaggcag	1560
atggatcacc tgaggtcagg agtttgagac cagcctgacc aatatggtga aaccttatct	1620
ccactaaaaa tacaaaaatt agccaagcgc ggtggcgggt gcctgtaatt ccagctattc	1680
aggaggctga ggcaggagaa tcacttgaac ctgggaggcg gaggttgagc tgagctgaga	1740
tcacgccact gcattccagc ctgggagaca gagctggctc aaaaaataaa ttaattaatt	1800
aaaaacaaaa ttggagatgc actatgttat ttcaaaaaca agctgccttt aaagatctat	1860

ctgttggtcac aggggtgggct catctgtttc attttatttt ctgtgggtta tctatttatt	1920
cattttaatg aactaggaag cattgtctct atttatggca taccacatga tgtttggata	1980
cgtgtatgcc tgtggcatgg ctaagtcaag ctagaacatg ggctttacct catatacgtg	2040
tcttattaag aacacataaa acctactctt gtagtgattt tcaaatatgc aacatatagt	2100
ttattaactg cagtcactat gatgtacaat agattgctcg aacttattcc tctgtctaa	2160
ctaagatfff gtgacctctg accaacatct cccagtggt gtcaccccc gcccccagcc	2220
tctgatagct gcctttctac tctctgcttc tgtgagttg atgtttatac attccacatg	2280
taagtggcct catgcagtggt ttctgtctct gtgtctggct tgttactta gcgtaatgtc	2340
ctccagcttc atctatgttg ttggaaatga caggatttcc ttctttcttg tggctgaata	2400
gtattgcctt gtgcatafac accacatttt ctttatccct tcattcactg atggactctt	2460
aggttgatgt catgtcttgg ctgttgtaaa aaatgccgca gtgagcgtgg gcgtgcaggt	2520
ccctcttcaa cacacggatt tcttttctt tggatataaa cccagcagtg agattgctgg	2580
atcacatggc agttctgttt ctcacctttt gaggaactc catactgttt tccataatgg	2640
ctgtagcaac ttccactccc acccccacgg tgcaaagtct ccatttctct tctacaacct	2700
caccaactcc tgttatfff catctttctg atagtagcca tttgaagagg tatgagatga	2760
tacctcattg tggttttcat ttgcattttt atttgtattt ttcatgaatt ttgagggtg	2820
atttcaaggg tagttatga ctgcaacagg gaaacgatcc tgagtatgag ggttgtgcta	2880
atcatcccc tcttgccagc tgcgtacgga atggggctct gcagatggca gggagctggc	2940
tcgtttctct ttaagagctg ccttttactt ttcttctct tcttttaaaa cttatttctt	3000
ggccggagcg agtggctcat gcctgtaatc ccagcacttt gggaggccga ggtgggcgga	3060
tcacgaggtc aggaattcca gaccagctg gccacatgg tgaaacccg tctctactaa	3120
aaatacaaaa attagccaga cgtggtggtg cgggcctata gtccagcta ctcgggaggc	3180
tgaggcagga gaatcactg aacctgggag gagggggtg cagtgagccg agattgcgcc	3240
actgcactcc agcctgggag acagagccag actccatctc aaaaaacaaa aaaaagtatt	3300
ttccaagca cagccatgta ttccaggctt gtggatcagc gttggtggtg gtgtgtgctc	3360
tcatactta gttccagcta agcacactct gacatgttta cactagaacc atttgttttt	3420
tctagaaata gaaatttcag aattgtagag tcagaggact taccagaaat ctcttaggta	3480
gttctctcc cctccctcaa gtgcagtcct aacctctgg agttttctgt agaaaccaca	3540
agcctcagag ctggccgaga attctagcca aagattttt catgccaaag taatcccccc	3600

tctcctaagg gccatccttg gtggggactg gtttcctgtt aagccctcgc tgtcagtcct 3660
ggctgtggaa tttcctgggtg aggagcactg gcccgtaggag ctgggccctc gtgccggcct 3720
tgagcaggcc caagtgttcc gtgttcttga tacctttcct ccagcacagt ctgtcttccc 3780
agaaaaaggt ttgcacttga aaatgatgca tttgctgatt aaacatagtt cttttgcttt 3840
atttggtttc taaaataaag tgggagtttt tgagattgag taactgagg ttaagatagc 3900
acgtggaatg gctttttctt tcttttctat tttttttttt tttttcctgg agacagggtt 3960
tcactctgtt gcccaggctg gagtgcagag gcatgacat ggctcactgc aacttcgatg 4020

tcttgggggtt aagcgatccc ccagcctcag cccccaagt ggctgggact acaggtgctc 4080
gccaccacac ctggctaatt tttgtatttt ttgtagaaaa tgggtttcat caatgttgc 4140
cagactggtc tcgaactcct gacctcaagc aattctcctg cctcagcctc ccagactgct 4200
gggattacag gcgtgaacta ccacgcctgg cctggaatgg cttttgatgt tctctatgt 4260
gcacatgtgg gtgaataaac accaacaag tccttatgtt acctgaagag ttgctctctt 4320
cttaatatTT aagtcgtatt tatttaaata cttaatatgt tgtacactat taaagtatta 4380
ttaggtcaaa atcaaggaag taaaaaggg tatgctgtga aaaatctctt ctctcttgc 4440

ctgcttactt acctaccccg catccccca tacacccag acacacacac acacacacac 4500
acacacacac acacacgcat cactcccata catgccacc tgtttaccag ccaatcacat 4560
ttcttggggc aactcatctg agttgcttct ctttcagag agtttttgca taaagaagca 4620
caggatattc tgcgttacca tgacctatt tccagtggt tccagccag ttgactctcc 4680
tgactggat accatcctgg acagattcc ttagggaaat gagccccctg ttttttccca 4740
ccatggcaca gttagtctt tgcatggacg caccattatt gccctgtct ctcttgggtg 4800
gaccttaagg ttttctcat ctttttctg taacacacac tgctccaagt gtgtgagcat 4860

atcagtagga aacgcttcca ggagtagaac tgetaggtca gagggcgtgt ggatctgtaa 4920
cctgacagac ctagaccggtc ttcagtttgg ttttatccag tttcatatt gattattcat 4980
ataaaaggaa acagacaaac ataacgctgt gcatgtattc tctcttagac cagaacaggc 5040
atagggtgca cttttaattt gtccatttcg tagagtagaa attgtttttg ctgaaatgaa 5100
caccttagga tgctgaagaa tatgaccgt cccatggaaa acattcaaaa atgtgtgtag 5160
cgctttcttc ccaagggtgt gtgtgcgcat attttaacac taattcactt tctacttccg 5220
ttgctatcct ttctgtgagt ctttctcaga atctcagaaa agaaactaaa ttgttactc 5280

tagttatcaa tgctgtactc tatacctgga atttgctaaa agggcagatt ttaagtattc 5340
tcaccacaga aaagagaaaa gaaaatggta attatgtgac gtggtggaca tgttaactag 5400
ctttattatg tgtagcattt cacagcggt atccagtcac cagctgtac acattaaaca 5460

tgtacaattg ggTTTTTTtg agacaaggTc tccttctgtc acccagtctg gagtgcagtg 5520
 gctcagtcac ggctcattgc agcctcgacc tcctgggctc aatccatcct tccccctcag 5580
 cctctgaaa agctggggcc acaggcatgt accatcatgc caggctaag catatatatt 5640
 tatatttttt ggTggagatg gggTtggTct cgaactctgg gctcaagtga tcctcccgcc 5700

 ttgcccttcc aaagtgtcga gattacaggc atgaaccaca gcaccaggcc tacatgtaaa 5760
 atttttatTT gtcaactata ctttgacaaa gctgagaaaa aaaatcctaa tttttaaaaa 5820
 aaaaaaaaa aggactagct tgagaccttt tccagctctc tggttatca gctgccgtct 5880
 ctccgggtg cagatagctg gaagggaag aaaatcccta aaattacca caagccaaga 5940
 atgaagtgtc tccctttgag ccacagtggc agttttgttt ttaatcatag aagtgtatTT 6000
 tgagccgggt gtgctggctc acgcctgtaa tccccgact ttgggaggcc gaggtggggg 6060
 gcggaggggg tggggatcgc ctgaggtcag gagttcgaga ccagcctgac caacatggag 6120

 aaaccccgTc tctactaaaa atacaaaatt agccggcgTg gtggTgcatg cctgtaatcc 6180
 cagctactca tgaggctgag tcaggagaat ctcttgaacc caggaggtgg aggttgcggt 6240
 gagctgagat catgccattg cactccagcc tgggaacaag aaaaaaaaaa agaagaaga 6300
 agaagtgtat tcatttcagt tacttttaaa aaagtgaaca gactttatat tttagagcgg 6360
 ttttaggttt acagaaaatg aaacagacag ggcagcgagc tccttgtact cctccccagc 6420
 acacagtTgc cctgttatga acatcccaca tcagtgtgtg gcgttcatta acaccgatga 6480
 acctgatgca tacattatga tgaactgaag tcctggactt caccctttct ctgttacagt 6540

 tctgtgggat ttgacaaatg cataatgtg tacagccaca atgatatgat cgtccagagt 6600
 agttctcctg ccttaaaacc tcttttctg cacctgtttc tctctccca ctcaccccag 6660
 ctatctgac tctttagtgc ctccgaagtT ttggtctttt caggatgttg tagcgttgg 6720
 atcatggagt atgtagcctt caccacatac accttccttc actttgttgg ctccctttac 6780
 ttagtaatat gcattcaagt ttcctccatg ccttttcatg gcttgatagc tcatttcttt 6840
 ttagcaccaa ataatatTcc gttgtccaga ttagcacaa tgtttatcca ttcattgTaa 6900
 ctgtgaccga ctcacagata ggatgtggaa tcactacca cagaggcatt agacaataat 6960

 cagacccaag tcatttcatg ggggaacaag cccacaggta ccagactgtc cagtgagtca 7020
 gggccactcg taggaagtga gaagagaggc tagagcatag ccaggTctc actttatact 7080
 ttaagcccat gtgtatttct cccaaccac acagcattgt ttccatgctt tcagctttgc 7140
 atgaataacg tgatactTga acgcatcatt tatcactTgc tctctttccc acagcgtgt 7200
 ttcaagctt ctccctgttc atgatgtct gcttaaccct taagctgcat gggattctgt 7260
 tctgtgaata cgcccccc atgtattatc ctgccagca aaaagtcccc aaactctgg 7320

atggtggtta ccctaggga gggagagaag agattgggaa tagggagcga cttcaacggt 7380

gtttgtaatg ttttgtttct ttaaataaaa gagctgagat catttcagca gaatgtgat 7440

ttagagtctc ctggacaatt tgttgctcaa agtgctctct taaagagcac tttaaaaaaa 7500

aaaacctttt atcttattat ttatttattt atttattgag acggagtitt gctctgtcac 7560

ccaggctgga gtggagtggg gtgatctcag ctactgcaa cctttacctc ctgggttcaa 7620

gcaattcccc tgcctcagcc tccaagtag gtgggattac agatgcgtgc caccacactt 7680

ggctaatttt tgcattttag tagagatcgg tttctccatg ttggccaggc tgatctcaaa 7740

cgctgacct caggtgatct gcccgcttg gcctcccaaa gtgctggtat tacaggcgtg 7800

agctaccatg cciggttat cttatatatt tttaaaaaca gcttattgag atctaattta 7860

tgtaccataa aattcaagta tataattcag tgcttttata tataaaacat atatatgaaa 7920

tagcttattg agatataatt ttttatataa aacagcttat tgatatgtaa tgtatgtacc 7980

ataaaattta aatataaat tcactggctt ttatatattc acgaatatgt gcaactatca 8040

ccacagtcaa ttttagcata ttttcacag ctcataaaga aacccaagc ccttgaacta 8100

tcacccata tcctctctc cagcccgctc ctctactca taagcaacca ctaatctact 8160

tagtgtctat agatttctta ctctaggcat tccatgtgag cgggatcatg caatactgg 8220

gctcacacaa tataagtggc attccatgtg agtcggctca tgcagtatgt ccggctcctt 8280

tcactgagca taaggtcttc agcactcatc caggttgag cctgtgtctg aatttcattc 8340

cctctcttgg ctgaatcgta ttccattgtg tatcttggac atacctatt ctgctcaccc 8400

agccgttggg gggcggttgg agtgttttcg cctttcagct gttttaagag ggttgagtg 8460

aacatttgta caagtttgg acccaatgcc tgttttcaat tctcttgtgt agagagcact 8520

ttttagcaga aaaaagaatg atttgggcc tcctttgtg tgcggtcagt gccttgagaa 8580

gagtgaactg tgctgccacc tccggagccg tggagagcgc ggggcttggg tagcagctag 8640

gacgatacaa gtgggacaa ggccaggtgc aatggctcac gcctgtaatt ccaacacttt 8700

gggagaccga ggcaggggga tcacctgagg tcaggagtgc aagaccagcc tggccaacat 8760

ggtgaaacct catctctaataaaaacagaaa aattaactgg acggggtggg ggacgcctgt 8820

aatccagct actcgggagg ctgaggcagg agaactactt gaacctggga ggcggaggct 8880

gcagtgagt gagatcagac cactgcactt cagcctaggt gacagagcga gactccgtct 8940

caaaaaaag aaaaaaaag aaagaaactc atggataatc ctccctctcg tgcagttcgc 9000

ctctacggac caaacttcat ctttcagggt tactcatctc agaggaagtc ctggcaccct 9060

gtgtgccaag acgactggaa cgagaactac gggcgggcgg cctgcaggga catgggctat	9120
aagttagtat ggggcagcac ccgccgagt acagtaacag acagcagaaa cagcagaaga	9180
ccctctctct gcctccctgt gaaagcaccg gcacatgagt gctggggaca attgtcacct	9240
tccaaaagct gagccctata accagcaggt ggaatttgtc ctgctagggc tgtgcccagc	9300
acacagacct tggtcactg ccaccttgcc ctgcctctc cttggcctct atagactcct	9360
ggttgctcgg gagtgccag tgctgtgttc atctggtcag aggggtaggc tgaggcgctt	9420
agggtgcctct tttccaagg tgcctctcag ccagggtcca ttcacctccc tgggtagagg	9480
ttggaccaga acagctggcg aggagggttg ggctggggag agcagcagag acaaatcctg	9540
tgccagtttc acitcattcg ggagccatgg aagccttttg agctggggag agaataatc	9600
aatcagactg atacttaaaa aatgtcattc ctgctcgtag ctctgaggga aggtgggaag	9660
gcttaacagg gtgtgtgtcg cctgacagtg attcctaacg ggggtggggc ggtggttacc	9720
atttaccagc actgcctggg gagatgcggc agccctcagg catcggggga gaggggtgta	9780
ggatgctact gccactttgt tttccatggg aggggtccca ggtgatttct atgcaacttt	9840
agggtattca atatgccagt tttcagaatg aattaccact cgggtagaaa gttagcatct	9900
tagctagtca ctgtgacatc cctaacagc aggggtgaat tacacagca agcccccca	9960
tcacagtcca ggaacctggt ggaattgata actggggcca tgtaaacatc tgtacctttt	10020
attagattaa atgtgtgtat gattatacaa tctatgtcc ttctcatagt ttcttgatcc	10080
taacctggat aagaacacg accaatgaag gaattttgtc tgacacttta gggttattga	10140
atcgaaaaat cgttacaata ttctagcact tggttagaac gtgtgatttt ttttctaaa	10200
tgctaagggt tttccctctt attctgaatg tcgtatgagc ggtattatga catagtatag	10260
gatttgtgtt tgcttatgcc ttaaccatta tcacaaataa ggttttcttt tttaggaata	10320
atttttactc tagccaagga atagtggatg acagcggatc caccagcttt atgaaactga	10380
acacaagtgc cggcaatgtc gatatctata aaaaactgta ccacaggtat gcagcaattt	10440
cttcttgaaa aattttggaa tgaaatcaac taggagacac catggggaat cgttgtcctg	10500
agtctgattt ctctgagctg caatactcgg tctggatggg ttttgattg ggaggagatt	10560
agagtctgac caggcctggt tactctaagc agcccttggg ttattcatag gaagtggctg	10620
aggtttctct gctatttcat tttcagctc taccgtctgc ccttgttggt agcggctcac	10680
acttgcaaca tcgacattca actctattta gttttctttc ctcttcagac atttagaggt	10740
gtacctattt tgtcaggcg tggttctagg aatccaagat aatgtctcag tgtcccagcc	10800
agggtgaccg gctcattcca gtttgccagg gacttcactg gcttgagcaa gggaaagtct	10860
gctccattcc aggcagctgg gctggctggt cccgttagcc ccaaccccg gacagcagtg	10920

ccagagggtg ctctgtgagg gatgggcagc attctggcgg cctgggaatg agttgtggtg	10980
tttccagggg gtagaagtgg gtacaagcca caggtcacat gatgagtggc tgacctggct	11040
gggagggcag aagaggggat ggacttaggc tcttcccttt gctttgcaca tatttaggat	11100
gtttgcagac ttgctatgat tgttgctgtt atgtgttttc tgatgtgaaa gatacacagt	11160
gtcctttgcc catgagctct ccttgccctc caggtcccca gggcttatgc ctggtgtcta	11220
ggcatcacct cctgcctgc caggtgccag gtgctgcatt tcgggggagg atgaactaat	11280
caccccgcg cacccttct ctgagtggga gcctggggca ggtttgcatt cctggaggcc	11340
gctgtgggag gggctctggg gcctgacttc cactgcagcc tgctgtcctg gggaatgtgg	11400
cagggcaagc ccagtgggga gggctgtgca cgccaggtg caccatcaa aacagcaggg	11460
ctgcggtttg tcctgttga gaagctaac acagctgcct gggcactttg taaatgctga	11520
gtggttcttt gtctttctgg gttacacacg gaatcaggga gccaagtcca gccgggcagg	11580
gacgggggga ggggaggagg tgctccctc ccttgcaag agccttgga actcacaagg	11640
aggctggagg gcttgaaga aagaagagaa ggccattgtc tggtaggctc tattctatct	11700
cgggtggtgt ggtgggggga ggcgcacttc ttttctctt tctgtgcagc agttgccctt	11760
tgatgcctga gtcttggct tgttttctgt cgggcttctg tgaataacca catgtgccct	11820
ggcgtgtga ccacacagg ctatccctac cgacctagg attcttagga aatgtcttct	11880
cttaaagggg acatgtcttc acttgccctg gtcagtcccc cagagccaga gtccacctgg	11940
aatgcacctg tagtcactga gaaccgggg ggtgtgcctt agtaagaagg tgcaggaag	12000
gacctattat tgtagggcct gggctcctgc aaggtggttt gggggtggtt ggaggaagca	12060
gagatttgct ctggattgga tgctgtcagg aagcaggggt aattctgtga ggctgcttta	12120
ttattttttt tctaggagga ggttggaatg aggctaggct aaagctgtga ttggtaaaga	12180
aacgtccgtc gctcaagtta gccaggacag gaggagacat cagatcgtga ttttgtggtt	12240
gtgagcaciaa ggttctgtt ctgtctgttc agacatcatt tcggaggagg ctcttgtgt	12300
cttgcccat ctcaggcatg gaggggccta gtccgatatt gacgtcagt gaaataattc	12360
aggttccgca gagcacacgg cccagctatc agggcgggcc agctctgcat gccaggggcc	12420
gcgtcttccc ttctcagcat agcctgggaa attcactgca ggacaaaatg catcagttac	12480
ttctcttca tcataacct gggatgtttg actcccaat gagtaactct tacgtttctt	12540
ctaactctag ggaactatt ggttatattg cttcaacac taaaattta aagcagttat	12600
aggagcccag aggtttccaa atggcttcct taaaaattag aagatgattt taaattccaa	12660
gaggaaaaac aaaactagca ttattgtata ctaccctca caaccgtcct aggagctggt	12720
acaattttaa gagaggttaa gtaacttgcc caaggtcaca ctgtgggat gtgagccgcg	12780

taccttggct cagtgtctgg tctttgccac tgcacctata tggatttact taccttattg	12840
gagttgtaac tagcagaccc ttctatgtct cagaagacag gagagggaac atcggaagaa	12900
atgactgatt tctaagcatg tgagaggcag gtgactccgc actatcgtga ccagaatttc	12960
ccctgttctt ttigcagtga tgcctgttct tcaaaagcag tggtttcttt acgctgtata	13020
ggtaagtcca tctggagtc cccttttgat acttctaact aggaaaagct ctctactttc	13080
agaacagtac tcctgtgtc tctgggggcg tgggaggga gaaggtgggg tcacgggttg	13140
gaatgtgcc agcggcgtct cgctctttcc aaggagctcc tggtttagat ttccatggcc	13200
tgtagacacc ttcagccttg ggtccaaggg acacccctg agatcaggca cgctcaagaa	13260
gctgacaaag ccctacactt tatgccaccc atgagctgga ggcccggcag gtctctttct	13320
ccagaaagca aaggggggtg gcgttagtga gccctggcag ccacctaacg tggacttgga	13380
gcatctgcgg ggctgtggtc cagcaccacc gtgtggccac caggtgtca tcagccagtg	13440
ggacccgga ggagggacaa gaccagagaa caacagtgt cttgcctctt ctctctgaa	13500
ttttggacgg tggttagac ttgggtgtcc ccctctctgt gtttagagt cttacagttt	13560
ccaaactgtt tgcaaatgtg gaagccaccg tcctctctct ctgggatggc ccagtgtgt	13620
cgtggggccg tggtcctgag ctacagcttt catttgaaga ggtggaagga gctgacaccg	13680
tcccatccc gcagggtcgg ctacagctct ctttaggtcc tgagtggggg tccagcacag	13740
ccccaagggt gcgtggcacc cgccctgccc tctgcccag cactcatctc ctggtggaga	13800
agacactcac acacaggaag cagggaaggc agcagacctc actcaccct caccctca	13860
ctcaccctc actcaccct tcaacctctc attcaccacc caccctctg cccctcact	13920
caccctca ctccctcaac cctcactcac ctctcactc cctcaacct cactcactc	13980
ctcactct cactctccc ctacacctc cctcaccct cccgtcacc tctcactca	14040
ctctcacc cctcactca ccttcacc cctcactcac cactcacct cctcactcac	14100
ccctactca accctcatt caccctcac ccctcactc accctgcac ccctcactc	14160
acccttcac cactcacc acctgctcac ctctcactc aaccctcac cccctacta	14220
atccctcact cctcaccct ctacgccct cactcacacc ttacctct cactcaccct	14280
ctcaccct caaccctta ctacccct cactcatccc ttacccctc actcaccct	14340
tctctaccc attcaccct tcaactatgc ctacccctc tcaactac cctcactcac	14400
accttcacc ctacgtcacc cctcactca ccccttacc cctcaatca tgccttact	14460
ccctactca ccccttacc ctctgaatta ctccctcacc cctcactca cccctcact	14520

caccccttca cccctcacc caccacctca cccacccctc acccaccccc tcacctcctt	14580
acccctcacc cccctcactc acccctcacc cctcactca ccacctcacc caccctcac	14640
ccacccctc actcactccc tcatccctc actcaccccc tcacccctc actcaccccc	14700
tcacccacc ctcaccacc cctcaccccc ctcaactcacc cttcaccccc ctcaactcacc	14760
ccctcactca ccccttcacc cctcactca ccacctcacc caccctcac ccacccctc	14820
actcactccc tcacccctc actcaccccc tcacccctc actcaccccc tcatctctc	14880
actcaccccc tcacctctc actcaccgc tcacctctc actcaccccc tcgccccctc	14940
actcaccct caccctca cccctcact caccctcac cccctgccc cctcactcac	15000
cccctgccc cctcactcac cctcaccccc ctcacccct cactatccc ctcacctct	15060
cactcaccccc ctcacctct cactcaccccc ctcacctct cactcaccccc ctcacctct	15120
caccacccc ctcactcact cctcaccccc ctcacccct cactcaccccc ctcacctct	15180
cactcaccccc ctcacctct caccacccc ctcactcact cctcaccccc ctcacccct	15240
cactcaccccc ctcacctct cactcaccccc ctcacctct cactcaccccc ctcacctct	15300
cactcatgcc ctcacccct cactcaccct ttcacctct tgcctatccc ctcacttacc	15360
ccctcacttc gtcaatcacc ccccacctc gtcaatcacc cctcacctt ttcactcacc	15420
ccctcactca ccccttact tctcactta cctctcacc cccactcac cccctcacc	15480
cccactcacc cctcaccccc acactcacc cctcaccccc cactcaccccc ctcacccctc	15540
tcacctctc actcaccccc tcacctctc acttatccc tcacccctc aattacccc	15600
tcacccctc aattactccc tcatctttc aattaccac tcacccctc acctctcac	15660
tctcactca cctctcact cacccttca ccttctact cacctctcg tctctcacc	15720
ccctcactca ctccagccc tgcctctcc atcttcttt tcttgtgtg agaactctggg	15780
gtccctgagt ggtgtcagtc cctcaagac tcaaggagtc cccagggcct tgttatccag	15840
aacaccccca cctgggtccc gggagacccc atgggatcac aggagtgtc agggaagtgg	15900
tgttctctgg gtctgggtgg gctggagggg cactctcct tcccaagag gagacccca	15960
ggagccccc aagtccatcc ccagcagtgg tgccctgcc ctgtccttg agcctgggag	16020
acccttggga gggcgggcg ctgggtggct gggcggttc tgcgtgttc acccactgg	16080
cctctgttt gtactctca gcctcgggg tcaactgaa ctaagccgc cagagcagga	16140
ttgtggcg cgagagcgc ctccggggg cctggcctg gcaggtcagc ctgcacgtcc	16200
agaacgtcca cgtgtcgga ggctccatca tcaccccca gtggatcgt acagccgcc	16260
actgcgtgga aaagtatgcc agggcgggcg cggccgggt gggggctcag ggctggccta	16320
cagccacct gtgacctga gcaggtctca acccttgag ccccgcatc cttgtgtta	16380

aatggggaga gtattgcacc tgcttcctag ggctgtgaga catcaagtgc gctcatgcca	16440
ggcagtgcac ggctgtatgc actgagtgtc cctgcacgc agggcacagg gtgcaggtgg	16500
aacattctcc acgatgtcgc cgtgaccagc gttccttcca gccactgtcc tctgagctct	16560
gtcctgccct tgagcaaage ccctgcccc tgaggtatcc tgtctccggg acgctagtcc	16620
caggagaggg cacactcaga caggcttcag gctgccctgc tggaaggtcc ctggggttaa	16680
gcgttcttgg ccacagcatt gctcatgcag aggggttaggt aggggtgagg ctagecgtga	16740
cagtattagc atttatggac gctaccaccc cctccccttt tccttaaaca catagtctt	16800
ttggtcacat gctgcttgg aggaggctc acttggcgga tgtatttttc tgccttagag	16860
agaggctgaa ctgggtttga ctgttggccc agccctctct tgcctcgtgc ccttagacga	16920
ttcactcaac gtctctgac catggcatgt acaactataa gatgggcatg ccttctcct	16980
ctcgggctgt tatgaaggc aaggaagcaa gggtgttac ccaaggtgc tcccttctct	17040
ccccctcttc acaccccag gtgctctggg ccctctagga actgggttct tctcaagggc	17100
tgttacccaa ggggtctccc ttctctccc ctcttcacac cactgggtgc tctgggcca	17160
ctaggagctg ggattctctt aagagggaaa ctcttgata aaggaaatgg ttgattgat	17220
atcggacaag tctgttcatt agtatccatt tattaagcac ctaccatgtg ccaggaaatg	17280
ctttggcgta caaagaaaa taagggccag tctgctaga aatggccttg aaaccccagg	17340
gagggatgtc ggccattgt ggggtctgca gattcctga aggtgatgca agagccagaa	17400
agaaggatga tttggggggc tgaggcaggg agtcggggtt gggggagtgt gggggagaag	17460
gggagaccga gcacctcttc cactatctcc ctgtgtggtt tttgtgaac catcctgcct	17520
ctgggtgtct tgcctccagc ttctgacgtt ggaagtcat ccactgagag ctctgtgttt	17580
atggctctga gatactagt ccttctcttc tccagacct cttaacaatc catggcattg	17640
gacggcattt gcggggattt tgagacaatc tttcatgttc tatggagcgg gataccaagt	17700
agaaaaagt atttctcatc caaattatga ctccaagacc aagaacaatg acattgcgt	17760
gatgaagctg cagaagcctc tgactttcaa cggtagtgt ggctcaggct tggcaagcag	17820
gttggcagaa tcttaaagag atgttgattg gaaatgacac ttgtgctatg ccaaattgaa	17880
gggaggcatt tgcgttgagc gaggttagcg tgcagcgggt ggccaatggg agaggctcac	17940
agaggctaag agcacctgcc gcatTTTggg ggaggcagca gccaccacat ctgttctgta	18000
ctgtactgag tgggtgtgat tcaagccagg catggaaaag gctagaacag ggctttccca	18060
ctgcagcacc cttgacatct ggggtggttct ctgtttagg gctctcttgt gcctttagg	18120
atgtttaaca gcgtccccag cctctacceca ctggaggcca gtagctacca agctgtgaca	18180
accagtgttg cctgctgaca ttgccaaca tccgctttga ggcaaagtca cttccagttg	18240

agaactactg gcctaaaatg tgtaaagatc cttgattttt aaagatacat tctaaaacca 18300

agttgcttaa ttcaggacaa acatgctttc tcttagcctc ttattcggtc ccactctggt 18360

ccatccaagg gtctggaatg ttctagcccc atgtggatac agaagaagca aaacctcagc 18420

cctccctaca gcatgtctgt attcacattg ggaaatggtt cacatataga agagcgaatg 18480

cctgagcaat ggcgtgggtc ctctggggcg aaagctgact ccattgactc catcggtttt 18540

ttggctgttg cctcctgtgt gtctttcccg tcttgatcac ctggagatat gtaattttgg 18600

aagcagagct agcaataat tcctcttata agcagagcta gcaataatt ctacttataa 18660

gtagcataac gtcttgctg ccagaaggag aggtctggca gggggagaaa gtgagaatgt 18720

gggacttggt gggatgcagg gtccctctggg cagggtggcc aggggtgccag gcccagcagc 18780

ctgcatgtgg gaaggccagg tggagacata ggtgataccc gcctggctca ctgtgttttc 18840

tcttcttgaa acagacctag tgaaaccagt gtgtctgccc aaccaggca tgatgtctga 18900

gccagaacag ctctgctgga tttccgggtg gggggccacc gaggagaaag gtgaggctgc 18960

tcctgggcac acaggactgc agggcccaca gatggagcat tgggttcgga agtgggaggt 19020

ccaggtttta atcccagttc tactactcaa tgactggatg actttggttg attccccag 19080

tccttgtgcc tcagtttctc catctgctaa gtgggagaaa tcctgccag cctacctaat 19140

acactgtgtt cttatcgtga tcacacagag cagcatgtgg aatggctttt gaagtatctg 19200

ggccatacga gtttagaggt gcaggatctc ctgtgttgca ctcatgtga gtttagagct 19260

gccctggaga tcccacaag gcctgcgtgg ctgagtaca gggggcttgg tgaggacggg 19320

catcctggac ccattggtgc cacatctaag cctgtcctct gccctgataa ccacagagag 19380

aggctctctc caccacttc ctttgcaatc tgcatctctc tctgacagtc tttcaaatga 19440

agggagcctg gtctcttcat ttttatggag ggttggaaat gcttagtggc aggcacaaag 19500

gttcatttta catattgttt atatccttct caaaagcgtc taggccatac agacaacaaa 19560

tcctttcaaa caaggggaaa agtacaaagg ttgggtgatt tctggggagc gtcaggaag 19620

gtagtggggg gcatcctggc tcctcatcag cagaaactta ctacagtaga gccacaggct 19680

gggcaaaaga cctcatggaa tccaagatga agggaatata gacaaatatt tgtgcgcacc 19740

tgacctagt acaggctggg tgctactcag gtgctgggaa tgcagaagt aacagagtaa 19800

gacaaatgct tctgctgta ggagctttac ctctctctg gatgtcgggt gtggggacgg 19860

ggcaggtgtg gtcagacaga tgggagacaa acaactgagc gaggtacttc caaacatctg 19920

aggggtggga tcacaagtc ccggtatatt tgaaggggtg gtcaggaaag gcttctcgga 19980

agaggtggca tttagctga gactcaaatg gcaaaaatgt gtacacatca aaaaggctag	20040
tgcattgtatc tttaggtgtg gtcaaggggc caaggagggtg ggctggggcc agattgcata	20100
ggctccttgtg gattatgggtg aagacaccag cttctcatct gcttgagggtg gggagatcgt	20160
gagccgggga gtgccatgat ctggcagctg cgtggggagt ggggatgaat ggatggagac	20220
gaggatgatg gtgacaagtc cattgtgtg gttccttgag acaggaagcc agctcatagc	20280
agagtgcggg cgtggatgtg aagagatgag ggtacactag ggctagagcc accagactta	20340
ctgatgggtt gcatgtctgt gggagagaga gtgagaagtc agggacgatg gctttccact	20400
ctgtggctga agccccaggg tggcgggtgg tgccattttt caagccagga aatattggtt	20460
gggtgagaatt tggggtggga gaaggtgtga cggagggttc tggttttgca cactaagccc	20520
acggtgcccc gaagatgccc gaggggaggc agcaaagcga gagtgggaaa tgcagagggtg	20580
gcaagtgcag gccgtgtctt gagaagctct aatgtgcagg ggagccgaga agcaggcggc	20640
ctagggaggg tcacgtgtgc tccagaagag tgtgtgcatg ccagagggga aacaggcgcc	20700
tgtgtgtcct ggggtgggtt cagtgaggag tgggaaattg gttcagcaga accaagccgt	20760
tgggtgaata agagggggat tccatggcac tgatagagcc ctatagtttc agagctggga	20820
atttctttcc ctgaagctga actccagagc tgcattcagc acaggcaccg ccagtgttaa	20880
ggagaatcca ggtttccag gagaggggtt ggtgtgtgga tgagctgacc ggggcagggc	20940
tggaaaaatg ggctgtgacc atctgtgtag tgcgtgtgga ggtctcaggg agggaaagtgt	21000
gctctccctg cgagagctgc aggcacact gggagctcaa caagtctccc tgtccttagg	21060
gaagacctca gaagtgtga acgtgccaa ggtgttctc attgagacac agagatgcaa	21120
cagcagatat gtctatgaca acctgatcac accagccatg atctgtgccg gcttcctgca	21180
ggggaacgtc gattcttgcc aggttaattca acatttttat tctacctttg gtccttacca	21240
gactctactg aaccccccat gagagagagg gcattcttgg ggtcagcaga gcctcctcag	21300
tgacacggag ccagctcggg gcagtcattg gaagtgcagg ccacaaacag tgcgaacgtt	21360
tctgtgtgga gaaggaagta cagtcaacaa atcacacaca ccctctgaaa aaccgttatt	21420
tggtaaaagt gccagtggaa cagaaacaag tatttagact attttaaat atgaacggca	21480
atttatttag taacttttag cttgaacaga ttaaaattca ggatgggggc tatctctttg	21540
ggggttacct ctctgttacc atcaccctt gatggtggag attcgaagcc cacacagtca	21600
ctcgtaactc acactgcgac ccccgcccc caactcctct aggcctggtc agtgggtgtc	21660
ggcagattgt gacttgattt tctgtctctt gtaccttgct gtgtcccaca ggggtgacagt	21720
ggagggcctc tggtaacttc gaagaacaat atctggtggc tgatagggga tacaagctgg	21780
ggttctggct gtgccaagc ttacagacca ggagtgtacg ggaatgtgat ggtattcacg	21840

gactggattt atcgacaaat gagggtaact atcctgtcct cttctgact gtgttctccg	21900
attcctcgag ccaaagccag acatctgtta ggcgtgggtc tgctgctgga agctgactgg	21960
tgaccactgg tcagcatgaa gcaaactctg ctctctccag ccacagcccc atccccccag	22020
tgtccacca ttgccattg cctctcactg gcttcacttg cataattccc ctgggtgttg	22080
gatgaaaagc gctggggctc agcttgtgtg aaattccttg gtgctctgcc aaccacactt	22140
cgttctggct cagctgactc agctgttcca cccaggccac ctacatcaa actttttttt	22200
tttttttttg agatggagtc tcactgtgtc gccaggctg gagtgcagtg gcacaatctc	22260
gactcactgc aacctttgcc tcttgggttc aagtgttct cctgcctcag cctcccaagt	22320
agctgggact acaggcatgc gccaccacgc ccagctactt tttgtatttt tagtagagat	22380
ggggtttctc catgttggcc aggcgtgtct cgaagccctg acctcaggtg attcaccac	22440
ctcagcctcc cacagtgtg ggattacaag tgtgaaccac ggtgcccggc ctacatgaa	22500
acttttgatt tatagagagc agagggaaga gccggctgtg cccatccttt tctggggcca	22560
tcgagtggct cctgggcagc cccaagggtt aggaaggga ggagcagcca gggttctctg	22620
atgccccaga ctcaagcacg aggggaaggtc tcagggttc catgtgagcc tcatggatgt	22680
ctctgcttag cagagccctg gctttgggca ttgtccagat aggggggtgag aaccagatct	22740
tctcatctcc aggacctcag acgtatagtt ttctcagatt tctgtgttt ctggggctgg	22800
gtactagtg gaagaaagca gtctattctg tcttctccca aatctccag atgccagtc	22860
tgttgaagga ggagcagaac cagggggcct ttcccgtga ggcccacct gtgtctcctt	22920
caaatgacac gcgggactca gggccttccc atgacatgg ggcccaggg gcgtcacctg	22980
gccaggggcc cagtgtaga aacagatgac cccaggagga ggaggcaggg caggaggga	23040
gctggcaggg ctgggatgt cagccaggct gaggggcgga ctgcaccag gatggagcta	23100
ggaaatgac caggtgtgtt tggcggctgc aggtgggtcc gcatggctgt gcaggagg	23160
aagggtctgc tggcaggaga gcagccgggg gagggccaga ctctgtgaa gagatgcctg	23220
ttgtgccgc ctccacatcc gctgccgct ccttccggag ctctgcccc gccatgctca	23280
gcctgactct gaccaacacg ttggagagaa gaatgatccc tttgtgtat taagcttgc	23340
tatttggttt ctaagtgtt catgcgaacc tagaggaaaa aattatttt cacctttgtt	23400
tgtcttaaga aaataacaca ctttttttt tctatttga acaggcagac ggctaateca	23460
catggtcttc gtccttgacg tcgttttaca agaaaacaat ggggctggtt ttgcttcccc	23520
gtgcatgatt tactcttaga gatgattcag aggtcacttc attttatta aacagtgaac	23580
ttgtctggct ttggcactct ctgccattct gtgcaggctg cagtggctcc cctgcccagc	23640
ctgctctccc taacccttg tccgaagggt gtgatggccg gctggttgtg ggcactggcg	23700

gtcaagtgtg gaggagaggg gtggaggctg cccattgag atcttcctgc tgagtccttt 23760

ccaggggcca attttggatg agcatggagc tgtcacctct cagctgctgg atgacttgag 23820

atgaaaaagg agagacatgg aaaggagac agccaggtgg cacctgcagc ggctgccctc 23880

tggggccact tggtagtgc cccagcctac ctctccaca ggggattttg ctgatgggtt 23940

cttagagcct tagcagccct ggatgggtgc cagaaataaa gggaccagcc cttcatgggt 24000

ggtgacgtgg tagtcacttg taagggaac agaacattt ttgttcttat ggggtgagaa 24060

tatagacagt gcccttgggt cgagggaagc aattgaaaag gaacttgccc tgagcactcc 24120

tgggtgcaggt ctccacctgc acattgggtg gggctcctgg gagggagact cagccttcct 24180

ctcatcctc ccigacctg ctctagcac cctggagagt gcacatgccc ctgggtcctg 24240

gcagggcgcc aagtctggca ccatgttggc ctcttcaggc ctgctagtca ctggaaattg 24300

aggtccatgg gggaaatcaa ggatgctcag ttaaggtac actgtttcca tgttatgttt 24360

ctacacattg ctacctcagt gctcctggaa acttagcttt tgatgtctcc aagtagtcca 24420

ccttcattta actctttgaa actgtatcat ctttgccaag taagagtggg ggcctatttc 24480

agctgctttg acaaaatgac tggctcctga cttaacgttc tataaatgaa tgtgctgaag 24540

caaagtgccc atggtggcgg cgaagaagag aaagatgtgt tttgttttgg actctctgtg 24600

gtcccttcca atgctgtggg tttccaacca ggggaagggt ccccttttga ttgccaagtg 24660

ccataacat gagcactact ctacctggt tctgcctcct ggccaagcag gctggtttgc 24720

aagaatgaaa tgaatgattc tacagctagg acttaacctt gaaatggaaa gtcacgaat 24780

cccatttga ggatctgtct gtgcacatgc ctctgtagag agcagcattc ccagggacct 24840

tggaaacagt tggcacgtga aggtgcttgc tcccaagac acatcctaaa aggtgttgta 24900

atggtgaaaa cgtcttcctt ctttattgcc ctttcttatt tatgtgaaca actgtttgtc 24960

tttttttgta tcttttttaa actgtaaagt tcaattgtga aaatgaatat catgcaaata 25020

aattatgcaa ttttttttc aaagtaacta ctgcatcttt gaagtctgc ctggtgagta 25080

ggaccagcct ccatttcctt ataagggggt gatgttgagg ctgctggtca gaggacaaa 25140

ggtgaggcaa ggccagactt ggtgctcctg tggttctcga gataacttcg tataatgtat 25200

gctatcgaa gttatgctag taactataac ggtcctaagg tagcgagcta gctccacgtg 25260

gttttgtecc agacttcctt tgtcttcaac aaccttctgc aagaaaacca agggcctgaa 25320

ttttaacttc ctg 25333

<210> 7

<211> 491

<212> PRT

<213> Artificial Sequence

<220><223> Recombinant protein

<400> 7

Met Ala Leu Asn Ser Gly Ser Pro Pro Gly Ile Gly Pro Cys Tyr Glu

1 5 10 15

Asn His Gly Tyr Gln Ser Glu His Ile Cys Pro Pro Arg Pro Pro Val

20 25 30

Ala Pro Asn Gly Tyr Asn Leu Tyr Pro Ala Gln Tyr Tyr Pro Ser Pro

35 40 45

Val Pro Gln Tyr Ala Pro Arg Ile Thr Thr Gln Ala Ser Thr Ser Val

50 55 60

Ile His Thr His Pro Lys Ser Ser Gly Ala Leu Cys Thr Ser Lys Ser

65 70 75 80

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

85 90 95

Ala Ala Val Ala Ala Val Leu Leu Trp Lys Phe Met Gly Ser Lys Cys

100 105 110

Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn Pro

115 120 125

Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp Glu

130 135 140

Asn Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Val Tyr

145 150 155 160

Ser Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp Asn

165 170 175

Glu Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn Asn

180 185 190

Phe Tyr Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser Phe

195 200 205

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

210 215 220

Tyr His Ser Asp Ala Cys Ser Ser Lys Ala Val Val Ser Leu Arg Cys
 225 230 235 240
 Ile Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile Val
 245 250 255

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

 Tyr Gly Ala Gly Tyr Gln Val Glu Lys Val Ile Ser His Pro Asn Tyr
 325 330 335
 Asp Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln Lys
 340 345 350
 Pro Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn Pro
 355 360 365
 Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp Gly
 370 375 380

 Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala Lys
 385 390 395 400
 Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr Asp
 405 410 415
 Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly Asn
 420 425 430
 Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser Lys
 435 440 445

 Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly Cys
 450 455 460
 Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe Thr

465	470	475	480
Asp Trp Ile Tyr Arg Gln Met Arg Ala Asp Gly			
	485	490	
<210> 8			
<211> 2267			
<212> DNA			
<213> Mus musculus			
<400> 8			
ccggttgtgt tataggactt gaccagcccc aatagtcctc aagtcactcc tagatacagt	60		
ggcaggtggg agctggcttg cggaaggaag aggaagaaga gaatgtgggc catcaaggag	120		
caaggccagc cttgcacttg ggccccctct gctcagtgtc gaccagggtt ttctgagccg	180		
cttcctaatt aggtcatttt gaagaccccc cccacacccc ctctgtgtgt cttgggtggc	240		
agagctagct ccaggctgta agaaaattag gaggattacc aaagcagtat ggagtcagac	300		
agtggccaac cctcaacaa ccgtgatatt gtcccttttc gcaaaccctg aaggccccag	360		
gagaccttca aaaaggtggg gatccccatc attgcagtgc tgcctgagcct gatagccctc	420		
gtgattgtgg cctttctcat caagtgatt ctggataaat actacttcat ctgcggcagt	480		
cccctgacct tcattcagag gggccagttg tctgacggcc accttgactg cgcctcaggg	540		
gaggatgagg aacactgtgt caaggacttc cctgaaaagc ccggagtggc agtccggctc	600		
tccaaggaca gatccaccct gcaggtgtgt gatgcagcca caggacactg ggcctcagtc	660		
tgtttcgaca acttcacaga agcactggcc aagacagcct gcagacagat gggctatgac	720		
agccagcccc ctttcagagc agtggagatc cgtccagatc agaacctccc tgttgctcaa	780		
gtcacaggaa acagccagga acttcagggt cagaatggaa gcagatcctg cctctcaggc	840		
tccctggttt ccttgcgttg ccttgactgt ggaaagagcc tgaagactcc tcgtgtgggtg	900		
ggtgggtggg agggccctgt ggattcttgg ccgtggcagg tcagcatcca gtacaacaag	960		
cagcatgtct gtggtgggag catctggat cccactgga tcttcacagc agcccactgc	1020		
ttcaggaagt atcttgatgt gtcaagctgg aaggtcaggg caggctcaaa catactgggt	1080		
aacttccat ccttgccctg ggccaagatc ttcactgctg aacctaatcc tctgtacccc	1140		
aaagagaagg acattgccct tgtaagctg cagatgccac tcacattctc aggtcagtc	1200		
aggcccatct gcctgccctt ctctgatgag gtgcttgtcc cagccacacc agtctgggtc	1260		
attggatggg gctttacaga agaaaacgga ggaaagatgt ctgacatgct actgcaggca	1320		
tcagtccagg tcattgacag cacacggtgc aatgcagagg atgcctacga aggggaagtg	1380		

accgctgaga tgctgtgtgc aggtacccca cagggtggca aggacacctg ccagggtgac 1440
 agtgggtgggc cttt gatgta ccattctgac aagtggcagg tagtaggcat cgtgagctgg 1500
 ggccatggat gcggcgggccc aagtactcct ggagtgtata ccaaggtcac tgcctatctc 1560
 aactggatcti acaatgttcg gaagctctgag atgtaacgct gccgtccccc acatccagaa 1620
 gctgcttccc ttcagaccta cctacggcat gacccctcaa agtcagatat gggacaagag 1680
 cctccttgaa caaactctgg taccctgca gcaagcaagg atacattgca gaggtgcccc 1740
 gagtggagtc agatgggcta gctcagccac cctgcatct cccaaaccct gggagacatg 1800

tggcccatgg gagtaaatcc aggacattga ctcaactctc agaagtgtta ttcagtcaag 1860
 gaggtctctcc ctccactga aggaaggaaa gtcagctctc tctgaaagg ccagatcact 1920
 ggctgagtag atgagacaag ggtatgaaag gcctttgcca tcttctttgc ccagtctga 1980
 aagcactgac gtaagagacc agtcagttct aatgtaaggt gtatatttta gtgtcagggt 2040
 attgcaattg tcacctctgt ggtcaatata attaaacagg tatgagaatt cgctggcata 2100
 gacttcttgg tctgcttaat aagaatcaa ctaaggatgt cacatgacag tttccagaa 2160
 aatgtgaaca agtgtccatc tgacacacgg caccaatgac aaaccaaaga agttattctg 2220

cctgagtctc agttgctgaa ctaataaatt agctgcggtt tcttgca 2267

<210> 9

<211> 435

<212> PRT

<213> Mus musculus

<400> 9

Met Glu Ser Asp Ser Gly Gln Pro Leu Asn Asn Arg Asp Ile Val Pro

1 5 10 15

Phe Arg Lys Pro Arg Arg Pro Gln Glu Thr Phe Lys Lys Val Gly Ile

20 25 30

Pro Ile Ile Ala Val Leu Leu Ser Leu Ile Ala Leu Val Ile Val Ala

35 40 45

Leu Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Ile Cys Gly Ser

50 55 60

Pro Leu Thr Phe Ile Gln Arg Gly Gln Leu Cys Asp Gly His Leu Asp

65 70 75 80

Cys Ala Ser Gly Glu Asp Glu Glu His Cys Val Lys Asp Phe Pro Glu

85 90 95

Lys Pro Gly Val Ala Val Arg Leu Ser Lys Asp Arg Ser Thr Leu Gln
 100 105 110

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

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

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

 Arg Pro Ile Cys Leu Pro Phe Ser Asp Glu Val Leu Val Pro Ala Thr
 305 310 315 320
 Pro Val Trp Val Ile Gly Trp Gly Phe Thr Glu Glu Asn Gly Gly Lys
 325 330 335
 Met Ser Asp Met Leu Leu Gln Ala Ser Val Gln Val Ile Asp Ser Thr

340 345 350
 Arg Cys Asn Ala Glu Asp Ala Tyr Glu Gly Glu Val Thr Ala Glu Met
 355 360 365

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

Ser Glu Met
 435

<210> 10

<211> 3543

<212> DNA

<213> Homo sapiens

<400> 10

atcattccag tttggcaact tcaattgtag ggctgtttta atcaagctgc ccaaagtccc 60
 ccaatcactc ctggaataca cagagagagg cagcagcttg ctcagcggac aaggatgctg 120
 ggcgtgaggg accaaggcct gccctgcact cgggcctcct ccagccagtg ctgaccaggg 180
 acttctgacc tgcctggccag ccaggacctg tgtggggagg cctcctgct gccttggggt 240
 gacaatctca gctccaggct acaggagagac cgggaggatc acagagccag catggatcct 300
 gacagtgate aacctctgaa cagcctcgat gtcaaaccce tgcgcaaacc ccgtatcccc 360

 atggagacct tcagaaaggt ggggatcccc atcatcatag cactactgag cctggcgagt 420
 atcatcattg tggttgtcct catcaagggt attctggata aatactactt cctctgcggg 480
 cagcctctcc acttcatccc gaggaagcag ctgtgtgacg gagagctgga ctgtcccttg 540
 ggggaggacg aggagcactg tgtcaagagc ttccccgaag ggctgcagt ggcagtccgc 600
 ctctccaagg accgatccac actgcagggt ctggactcgg ccacaggga ctggttctct 660
 gcctgttttc acaacttcac agaagctctc gctgagacag cctgtaggca gatgggctac 720
 agcagcaaac ccactttcag agctgtggag attggcccag accaggatct ggatgttgtt 780

gaaatcacag aaaacagcca ggagcttcgc atgcggaact caagtgggcc ctgtctctca 840
 ggctccctgg tctccctgca ctgtcttgcc tgtgggaaga gcctgaagac cccccgtgtg 900
 gtgggtgggg aggaggcctc tgtggattct tggccttggc aggtcagcat ccagtacgac 960
 aaacagcacg tctgtggagg gagcatcctg gacccccact gggtcctcac ggcagccac 1020
 tgcttcagga aacataccga tgtgttcaac tggaagggtc gggcaggctc agacaaactg 1080
 ggcagcttcc catccctggc tgtggccaag atcatcatca ttgaattcaa ccccatgtac 1140
 cccaaagaca atgacatcgc cctcatgaag ctgcagttcc cactcacttt ctcaggcaca 1200

 gtcaggccca tctgtctgcc cttctttgat gaggagctca ctccagccac cccactctgg 1260
 atcattggat ggggctttac gaagcagaat ggagggaaga tgtctgacat actgctgcag 1320
 gcgtcagtc aggtcattga cagcacacgg tgcaatgcag acgatgcgta ccagggggaa 1380
 gtcaccgaga agatgatgtg tgcaggcatc ccggaagggg gtgtggacac ctgccagggt 1440
 gacagtgggt gggccctgat gtaccaatct gaccagtggc atgtgggtgg catcgtagt 1500
 tggggctatg gctgcggggg cccgagcacc ccaggagtat acaccaaggt ctcagcctat 1560
 ctcaactgga tctacaatgt ctggaaggct gagctgtaat gctgtgccc ctttgtagt 1620

 ctgggagccg ctctcttct gccctgcca cctggggatc ccccaaagtc agacacagag 1680
 caagagtccc cttgggtaca cccctctgcc cacagcctca gcatttcttg gagcagcaaa 1740
 gggcctcaat tctataaga gacctcgca gccagaggc gccagagga agtcagcagc 1800
 cctagctcgg ccacacttgg tgtcccagc atcccaggga gagacacagc cactgaaca 1860
 aggtctcagg ggtattgcta agccaagaag gaactttccc aactactga atggaagcag 1920
 gctgtcttgt aaaagcccag atcactgtgg gctggagagg agaaggaaag ggtctgcgcc 1980
 agccctgtcc gtcttcacc atcccaagc ctactagagc aagaaccag ttgtaataa 2040

 aaatgcactg cctactgtt ggtatgacta ccgttaccta ctgtgtcat tgttattaca 2100
 gctatggcca ctattattaa agagctgtgt aacatctctg gcataggcta gctggaatgc 2160
 ttgataagaa ctgagctggg atgattgaac ttctattctt tggcttgggg agaaaagaag 2220
 tcttggggaa gcaattgagt ctcaaagtag aggcagggga aaaaagagt agggagacca 2280
 gatctgctga gtggcagcaa gaggtagctg cagattacag aaaccagggt gagcaagttt 2340
 gagtcccaca cagggccttc tcccttgcc tctttccctc cctccctgcc tgtgataatc 2400
 agccaggagc cagggataac ctatgacttg ggaaagagat gagttaggca gtcaagggtg 2460

 acattcaatc agggatccac aagtggctgg aaagaaatgc tggctctgtg tcttaacttt 2520
 ttccgcctgg agagccctca gtgtggcttc ttacatttaa aaaacaaaaa ggatcagctg 2580
 ccagggtgta ggcagtcctc aagctgagt gtgaggatgt aagcatgaat aagtcctgc 2640

actcaaatg gtcaaagaat taaaccccat ggactttttt ggcatctgta tgaaagcttg 2700
 ggttttctga ggactgtctt gctatagtta agtcagatcc tagatgaaat atacttggtc 2760
 atactgtact aggttcttag gaaacaacag aattcctcaa atgccaaaaa caaagaaaat 2820
 agaaaccag aaaacaaaac aaaataaaac aaaaccatca gaactgtgag tggaaactaa 2880

ggtgatgac tgggagcaat aactaaaaat ctgggtcga gacctatatg aaggctggca 2940
 gtggagctaa acctggacac actgaagaca agggagctga accagggctc ctacatgaag 3000
 cagggataac tgatggcagt aaatgtggc tcaaattgca gatggctctg aggaaaattt 3060
 cccaaattta gagcctcagg attcccaaag atcctccaaa tatgagctca caatcaaaga 3120
 tcagagacgt tgaaaaataa aaaacacctt aagtgggcag cataaaaaac agctaattta 3180
 gaaccccaaa ggcttcagat gtcagaatat tagagactta tgataataag caatatttgc 3240
 agagtatttg tatgtgccag acactattgt aagtgttca tcatgtactg attcatttaa 3300

tactcacaga aatctgtgag atgggtatta ttcttacct cactctatgg attaaaaaa 3360
 ctaaggcaca aagtggttaa gtccttgcc tgagattata gactgtaagt tgaacgtgag 3420
 cacttggat acagagtta tgctgtaaac taccacacta tagggcctcc aatatgataa 3480
 ttataaaat attgaataa aaaatgaata ctagttccac atttaaaaaa aaaaaaaaaa 3540
 aaa 3543

<210> 11

<211> 437

<212> PRT

<213> Homo sapiens

<400> 11

Met Leu Gln Asp Pro Asp Ser Asp Gln Pro Leu Asn Ser Leu Asp Val

1 5 10 15
 Lys Pro Leu Arg Lys Pro Arg Ile Pro Met Glu Thr Phe Arg Lys Val

20 25 30
 Gly Ile Pro Ile Ile Ile Ala Leu Leu Ser Leu Ala Ser Ile Ile Ile

35 40 45
 Val Val Val Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Leu Cys

50 55 60
 Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln Leu Cys Asp Gly Glu

65 70 75 80

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

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

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

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

325 330 335
 Gly Lys Met Ser Asp Ile Leu Leu Gln Ala Ser Val Gln Val Ile Asp
 340 345 350
 Ser Thr Arg Cys Asn Ala Asp Asp Ala Tyr Gln Gly Glu Val Thr Glu
 355 360 365
 Lys Met Met Cys Ala Gly Ile Pro Glu Gly Gly Val Asp Thr Cys Gln
 370 375 380
 Gly Asp Ser Gly Gly Pro Leu Met Tyr Gln Ser Asp Gln Trp His Val

385 390 395 400
 Val Gly Ile Val Ser Trp Gly Tyr Gly Cys Gly Gly Pro Ser Thr Pro
 405 410 415
 Gly Val Tyr Thr Lys Val Ser Ala Tyr Leu Asn Trp Ile Tyr Asn Val
 420 425 430
 Trp Lys Ala Glu Leu

435

<210> 12

<211> 20078

<212> DNA

<213> Artificial Sequence

<220><223> Recombinant polynucleotide

<400> 12

ccacccgcac aactacagt cgagataact tcgtataatg tatgctatac gaagttatat 60

gcattggcctc cgcgcgggtt ttggcgcct cccgcgggcg ccccccctct cacggcgagc 120

gctgccacgt cagacgaagg gcgcagcgag cgtcctgac cttccgcccg gacgctcagg 180

acagcggccc gctgctcata agactcggcc ttagaacccc agtatcagca gaaggacatt 240

ttaggacggg acttgggtga ctctagggca ctggttttct ttccagagag cggaacaggc 300

gaggaaaagt agtcccttct cggcgattct gcggagggat ctccgtgggg cgggtgaacgc 360

cgatgattat ataaggacgc gccgggtgtg gcacagctag ttccgtcgca gccgggattt 420

gggtcgcggt tcttgtttgt ggatcgctgt gatcgctact tggtagtag cgggctgctg 480

ggctggccgg ggctttcgtg gccgccgggc cgtcgggtgg gacggaagcg tgtggagaga 540

ccgccaaggg ctgtagtctg ggtccgcgag caaggttgcc ctgaactggg ggttgggggg 600

agcgcagcaa aatggcggtt gttcccagat cttgaatgga agacgcttgt gaggcgggct 660

gtgaggtcgt tgaacaagg tggggggcat ggtgggcggc aagaacccaa ggtcttgagg 720
ccttcgctaa tgcgggaaag ctcttattcg ggtgagatgg gctggggcac catctgggga 780
ccctgacgtg aagtttgta ctgactggag aactcggttt gtcgtctgtt gcgggggcgg 840
cagttatggc ggtgccgttg ggcagtgcac ccgtaccttt gggagcgcgc gccctcgtcg 900

tgctcgtacg tcacccgttc tgttggtta taatgcaggg tggggccacc tgccgttagg 960
tgtgcggtag gcttttctcc gtcgcaggac gcagggttcg ggccctagggt aggctctcct 1020
gaatcgacag gcgccggacc tctggtgagg ggagggataa gtgaggcgtc agtttctttg 1080
gtcggtttta tgtacctatc ttcttaagta gctgaagctc cggttttgaa ctatgcgtc 1140
ggggttggcg agtgtgtttt gtgaagtttt ttaggcacct ttgaaatgt aatcatttgg 1200
gtcaatatgt aattttcagt gttagactag taaattgtcc gctaaattct ggccgttttt 1260
ggcttttttg ttagacgtgt tgacaattaa tcatcgcat agtatatcgg catagtataa 1320

tacgacaagg tgaggaacta aaccatggga tcggccattg aacaagatgg attgcacgca 1380
ggttctccgg ccgcttgggt ggagaggcta ttcggctatg actgggcaca acagacaatc 1440
ggctgctctg atgcccgctt gttccggctg tcagcgcagg ggcgcccggt tctttttgtc 1500
aagaccgacc tgtccgtgtc cctgaatgaa ctgcaggacg aggcagcgcg gctatcgtgg 1560
ctggccacga cgggcgttcc ttgcgcagct gtgctcagc ttgtcactga agcgggaagg 1620
gactggctgc tattggcgga agtgccgggg caggatctcc tgtcatctca ccttgcctct 1680
gccgagaaag tatccatcat ggctgatgca atgcggcggc tgcatacgt tgatccggct 1740

acctgcccatt tcgaccacca agcgaacat cgcacgcagc gagcacgtac tcggatggaa 1800
gccggtcttg tcgatcagga tgatctggac gaagagcatc aggggctcgc gccagccgaa 1860
ctgttcgcca ggctcaaggc gcgcatgcc gacggcgatg atctcgtcgt gacctatggc 1920
gatgcctgct tgccgaatat catggtggaa aatggccgct tttctggatt catcgactgt 1980
ggccggcttg gtgtggcgga ccgctatcag gacatagcgt tggctaccg tgatattgct 2040
gaagagcttg gcggcgaatg ggctgaccgc ttctcgtgc ttacggtat cgcgcctccc 2100
gattcgacgc gcacgcctt ctatgcctt cttgacgagt tcttctgagg ggatecgcgtg 2160

taagtctgca gaaattgatg atctattaaa caataaagat gtccactaaa atggaagttt 2220
ttctgtcat actttgttaa gaagggtgag aacagagtac ctacattttg aatggaagga 2280
ttggagctac ggggggtggg gtgggtggg attagataaa tgcctgctct ttactgaagg 2340
ctctttacta ttgctttatg ataattttc atagttggat atcataattt aaacaagcaa 2400
aaccaaatta agggccagct cattctctcc actcatgac tatagatcta tagatctctc 2460
gtgggatcat tgtttttctc ttgattccca ctttgtggtt ctaagtactg tggtttccaa 2520

atgtgtcagt ttcatagcct gaagaacgag atcagcagcc tctgttccac atacacttca 2580

ttctcagtat tgttttgcca agtttctaatt ccatcagacc tcgacctgca gcccctagcc 2640

cgggcgccag tagcagcacc cagctccacc tttgtgtctag taatgtccaa cacctccctc 2700

agtccaaaca ctgctctgca tccatgtggc tcccatttat acctgaagca cttgatgggg 2760

cctcaatggt ttactagagc ccacccccct gcaactctga gacctctgg atttgtctgt 2820

cagtgcctca ctggggcggt ggataatttc ttaaaaggtc aagttccctc agcagcattc 2880

tctgagcagt ctgaagatgt gtgcttttca cagttcaaat ccatgtggct gtttcaccca 2940

cctgcctggc cttgggttat ctatcaggac ctgacctaga agcaggtgtg tggcacttaa 3000

cacctaaagt gagtgactaa ctgaacacitc aagtggatgc catctttgtc acttcttgac 3060

tgtgacacaa gcaactcctg atgccaaagc cctgccacc cctctcatgc ccataattgg 3120

acatggtaca ggtcctcact ggccatggtc tgtgaggtcc tggctcctctt tgacttcata 3180

attcctaggg gccactagta tctataagag gaagagggtg ctggctccca ggccacagcc 3240

cacaaaattc cacctgctca caggttggct ggctcgacc aggtgggtgc ccctgctctg 3300

agccagctcc cgccaagcc agcaccatgg gtacccccaa gaagaagagg aaggtgcgta 3360

ccgatttaaa ttccaattta ctgaccgtac accaaaattt gcctgcatta cgggtcgatg 3420

caacgagtga tgaggttcgc aagaacctga tggacatgtt cagggatcgc caggcgtttt 3480

ctgagcatac ctggaaaatg cttctgtccg tttgccggtc gtgggcggca tggtgcaagt 3540

tgaataaccg gaaatggttt cccgcagaac ctgaagatgt tcgcgattat cttctatata 3600

ttcaggcgcg cggctctggca gtaaaaacta tccagcaaca tttgggccag ctaaacaatgc 3660

ttcatcgctg gtccgggctg ccacgaccaa gtgacagcaa tgctgtttca ctggttatgc 3720

ggcggatccg aaaaagaaac gttgatgccg gtgaacgtgc aaaacaggct ctgagttcgc 3780

aacgcactga tttcgaccag gttcgttcac tcatggaaaa tagtgatcgc tgccaggata 3840

tacgtaatct ggcatttctg gggattgttt ataacacct gttacgtata gccgaaattg 3900

ccaggatcag ggttaaagat atctcacgta ctgacgtgg gagaatgtta atccatattg 3960

gcagaacgaa aacgctggtt agcaccgcag gtgtagagaa ggcacttagc ctgggggtaa 4020

ctaaactggt cgagcgaagg atttccgtct ctggtgtagc tgatgatccg aataactacc 4080

tgttttgccg ggtcagaaaa aatggtgttg ccgcgccatc tgccaccagc cagctatcaa 4140

ctcgccctt ggaagggatt tttgaagcaa ctcatcgatt gatttacggc gctaaggtaa 4200

atataaaatt ttaagtgtta taatgtgtta aactactgat tctaattgtt tgtgtatttt 4260

aggatgactc tggtcagaga tacctggcct ggctcggaca cagtgcccg gtcggagccg 4320
cgcgagatat ggccgcgcgt ggagtttcaa taccggagat catgcaagct ggtggctgga 4380
ccaatgtaaa tattgtcatg aactatatcc gtaacctgga tagtgaaaca ggggcaatgg 4440
tgcgcctgct ggaagatggc gattgatcia gataagtaat gatcataatc agccatatca 4500
catctgtaga ggttttactt gctttaaaaa acctcccaca cctccccctg aacctgaaac 4560
ataaaatgaa tgcaattgtt gttgttaaac ctgccctagt tgcggccaat tccagctgag 4620
cgtgcctccg caccattacc agttggctctg gtgtcaaaaa taataataac cgggcagggg 4680

ggatctaagc tctagataag taatgatcat aatcagccat atcacatctg tagaggtttt 4740
acttgcttta aaaaacctcc cacacctccc cctgaacctg aaacataaaa tgaatgcaat 4800
tgttgttgtt aacttgctta ttgcagctta taatggttac aaataaagca atagcatcac 4860
aaatttcaca aataaagcat ttttttactt gcattctagt tgtggtttgt ccaaactcat 4920
caatgtatct tatcatgtct ggaataactt cgtataatgt atgctatacg aagttatgct 4980
agtaactata acggtcctaa ggtagcgagc tagccaagtc tgtgtgtctac caagtagcaa 5040
aactgagcct ggaactcaca catgcgtgtc tgagagccca gcactatcgc caggaaaacc 5100

cagcgtctcc ctgctcaagc ctgacctca gccctctctg cctctccctg cacttgcctt 5160
ccagtcaagg tgattctgga taaatactac ttctctgcg ggcagcctct ccacttcac 5220
ccgaggaagc agctgtgtga cggagagctg gactgtccct tgggggagga cgaggagcac 5280
tgtgtcaaga gcttccccga agggcctgca gtggcagggt agtgcagggt ctgaggcaca 5340
agagaagtgg gccagcagg aggtctgtc agggcccccac ggccactgc atagtatctg 5400
ccccctactt gtcacttttc atccttggtt tataaggttc tttgtttgtt tgtttgtgt 5460
tgttttgagg cagagtgtc tgtggcccaa gatggagtgc agtgtcttgg tctcggtca 5520

ctgcaacctc tgcctcccag ttccaagtga ttcttctgcc tcagcctcat gagtagctgg 5580
gattacaggt gccagccacc acgcctggct aatttttata ttttagtag agacggggtt 5640
ttgccacatt ggtcaggctg atcttgaact cctgacctca ggtgatctgc ccgcctcagc 5700
ctcccaaagt gctgggatta caggcgtgag ccaccgtgcc cagctgtgta agtttcttga 5760
gagcaggacc ctgtcttctc tacctttaa tctagtact taacacacag caaacagtaa 5820
ctatttgatg accaaatgtg agccagaaag gacaggaaat tgtaactgag gctgccccat 5880
gcgtgtctcg cctgggtgat ttcaggcaga gggctagact gggtagcctt ggggcattcc 5940

tcctttctat gaaatttgtt atttcaagga gactagaaaa gagacttctc agccacttcg 6000
ccagctattg gtctttctat tcattagtgt ttgctgagac atgctatgtg acaggactga 6060
gccaggctct ttcaatggat aggagatgtt ttgagcataa aatccacgtt ctctcttggg 6120

ctgggctctt ctaccttctt cccctgggtg cttgggctct gaagaaaaaa agataggtag 6180
gagatgagt atggggcttc tgagggcagg gctgagtac tttctgtga ttgtctttt 6240
ctttatcaga agtcaaagc ccacaggcac ctgtcatcct actgccagta ggactttctca 6300
ctcaaccttc cctctgacc ttacttggag aaggacttag gtcctctct cagacatttc 6360

cccaggtcgg gcaagttgtg tggaccatgg atgggtatgt ggtccataca atttaacaa 6420
gctgtatatg gtcgctgggt agagtacca cataattgat catcaaaact gatacctga 6480
agagcaaaag ggggcactat taaccattgg gtcagggcaa caggtaaaaa tggagaccta 6540
ccctgggact tctggtcaca ctagctactg tcaaaatggg gcccaaatag acaaagccaa 6600
atggaagaaa ttcccttgac attgaaagtg ttgggctct gtggcacccc cagttctagg 6660
ttgggggagc ttgggctggt ctcatgatga gttctgaggg ggatgggcca gttgggcccc 6720
ccgttccatc taactcaggt tcttttctc ccagtccgcc tctccaagga ccgatccaca 6780

ctgcagggtc tggactcggc cacagggaac tggttctctg cctgtttcga caacttcaca 6840
gaagctctcg ctgagacagc ctgtaggcag atgggctaca gcaggtaacc aacctgggcc 6900
tctctctttt ttccctctt cctcttctt cctcttctc ctttcttcc tcccttctt 6960
tctcttctt aaaaattacg ggcataggag ccaggcagaa tggcttttga atcccagcat 7020
ttcacttata agcaacatga agttaaat tctaagcctc aggttctca ggagttaatt 7080
gggggaacta atgccaacct cataggatag ttttgcaatg ccagtggag aatgtgtgct 7140
gccctccaac acacacacac acattctag cgtctatgca gtcctctctt tcttttact 7200

cctcaacctt cactccttg tgctggcttt gcaagaaact gttcctgccc agtaatacaa 7260
aagctaagtt aacttattca aagtctggt agttaagatt tagcttaagt gaggctagtt 7320
tcagtggggc ccatcttca gcaatcccag ctctctctgc aaatttcaa agcagttcca 7380
aatctggagt ggatgaaaag gtgtaagatg atagtaagag taatttgcat tctatatatt 7440
tatattcact tgattttggc agaaaaccaa aaagatagtt attatatctt atatatagat 7500
atatattata tctatttcat aaataggctc aaacaaagta agtaacttgc tagggacta 7560
gctgggaggt agagggtag aatttgagcc caagaccct aattcttgcg cattaggagt 7620

tcccacattg tttctgttc tagactgagt aattctttat tctcatgtag gacatcatct 7680
ctaagggaag gggctaata gatggttgat cactcagaga gtttagctgg agaggatgga 7740
aaagaacca tacattcagt tgcagattga gatagcctat ctctggcagg cctcagattt 7800
cttcaggatt ctaacagact ggaccagag actaggccaa acaacaaac aaacaaaaac 7860
tctactaggc agacatcacc aaccaatcac agaactctt cccatggatc cctaatacag 7920
cctcaaagtc cttttcagta aatgctccag gcagccatta caaatcaatc agaattattt 7980

gcctttctct tctctgctca acgggcttct gctgctctct actttccata gggggcaact 8040

tccattaccc tctagaaagc acaccccacc accttcattt caaggagagt gaggaactca 8100

tgcccagcac ctgctattct cccctcttcc tgcagccacg gagcccagcc tcgctgcagc 8160

cagccctgcc tccccactgt agtccagtca actgctgcat cagccgttcc tggcacagca 8220

ggctgagcct tgattatgaa acctgggtgt ctccaggggt tcttaagatg ataggctcct 8280

ggaatttctg tctttttgga gctcagtaag gcaccaaacc acctgagtct tgtgcttcac 8340

aaaatcaaag ttcatcagaa tcattcattg ggatggaatt ggtgaacaga agttaacttt 8400

cctgggaatg tccatttcca ccatattccg tcttctagg tctcagactt ctctactttc 8460

tttctctct ctagatcgga ggcccttctt gtcttagaac cataggcatt tcaagatgtg 8520

ggagacccta gggatcatct agtccacgca tctttttttt ttttttttga cagagtctca 8580

ctctgtcacc caggctggag tgcaatggca ccatctctgc ttactgcaac ctccacctcc 8640

caggttcaag tgattctttc gcctcagcct cccaagtagc tgggattaca ggcacgcacc 8700

atcatgccc gctaattttt atatttttgt agagaccgag tttcaccatg ttggccaggc 8760

tggctttgaa ctctgacct caggatgatc acccacctcg gcctccaaa gtgctgggat 8820

tacaggcgtg agccactgca ccagcccccg tgcattcttt tatagagggg gaaactgagg 8880

cttggagaga ccagaaaaa gaatatgacc tgcccaaggc cacacatcaa actagtcca 8940

gagccaggga cagaacctag atcatgagga ctcttaaaat gcactctagt cctcccaggt 9000

ctgagacttg ggtccttcca ggaagtcca gcattcctgc ctgagaatgt gccaatccac 9060

cagtattgcc aatgactcag ccctccatgg agagcttcta ctaacattac tagcatagtt 9120

agggatggaa ggaaaagatt tagaagaggc agattcagta aaggaacaat cagagagatg 9180

gaattaatca aggaaggctt cctggaggag gaaaaacttc aaccaaggt ttgaaagtag 9240

caagcatgga ttagcaggga gaaagaggga gattggtcca gttgagagaa acgtttgtct 9300

ggattcatat gaagacagat ctagtctgt tctattaaat atctctaagg gggccaaaaa 9360

cataccccg ctatcaaagt cagaccagat gctttgtttg gagaacgaaa tatccacatt 9420

ccaactccct cccaggtag aaggagctc acctgagccc ctatgcctct ttgtttccct 9480

gctgtgaacc agaagacatt gctgggatat ttgaaatagg gacagagctg ggaatatgga 9540

aaggagacc ctaacatttc tccagggtc tgggttctgg atttgattc cccaccaag 9600

aaagcaagtt acatcagcaa tgcactgagg gttgagtcct gggatgcaa gggtcggttc 9660

tttattgtat agcaaagcag gcccactct cactgactaa gaccatctcc actccctggc 9720

cactccccac caagcattct ctgccactct ttctcctgaa agtgggggcc aactctacca	9780
tcttgttcta accccctgcc ccagctcaca actctctctc cctcttgatg tgagcagcaa	9840
accactttc agagctgtgg agattggccc agaccaggat ctggatgttg ttgaaatcac	9900
agaaaacagc caggagcttc gcatgctgaa ctcaagtggg taagtgaggg gacaccttct	9960
ggcctacaga aggccccac atggacgtg ctcttcaggt tgcaaccagc tcacctgga	10020
ccccaaagcag ccaggggaat gtaagcagac atcaggaaga actcctagcc agatggatca	10080
ttcaatgcca agagctatag actcacattt tggagagggt ttctgtgttg acttgttttt	10140
aatacaatgg acagctggac aaagtgtgtt gtcctactca gagccagagg gatggataat	10200
gtgacctttc catcaatctg gatagtaa atgttttgc actgctgtag gttttcta	10260
aaattgcccc ataggcaaga ttccaaagtc actttgtcct tccctaccac ttaccagcc	10320
agagctcccc accttcttga tgctccaggg aagaggtcc atggcccttg tgggtggcct	10380
gttcttgagc ctgcccacc tgtgttagag cagagcatcc agatgaaatc tgtcacactg	10440
tggcaaagtg gctcagagag gaggtggct tcctagcatt caggagcgtt gctgagggcc	10500
gcttattcac cgaaaataaa tcttgaaaag gacagggtg gtagcagaat gatcctttac	10560
ctaaaattct atcaaaatcc cattcttcca ttggaaagc ccacagtgc acagactctg	10620
ttccgggtc tgcctcttc cctcttgggt ccagagagcc caggctgggc ttgaaagcag	10680
gcagggccca gcacacagta ggtactcagc agtgggggtg ttgaatcaa tcaaacggaa	10740
gtgtcaatgc aggaaatgca atggatgtca atgcagtctc caaatgttcc cactgtgca	10800
gcttccacat tcccaggta ttgggagggg acttgaatta acagcttcgg gaggcctgag	10860
tccctgcctc ccagctgagg aagaagctta aatcacaggc cgctgtgtct gtcttccagg	10920
ccctgtctct caggctccct ggtctccctg cactgtcttg gtgagtacc ccaatctctg	10980
agggtttggg gcctgggcca gcaatgagca gggaggaaga ccttcattct cactcctaaa	11040
tttctgggac tccaagtctc attctgcctt ggtctacagc ccttgggctt gtcggtcaat	11100
gccccctcga gttgttggtg gccttgggca ggtcacattc ttttctggg tctttccaag	11160
ccccagtctc ccccttctac catctgtgca tggctccatg acctaaagtg agacctggga	11220
gagagtgtta ggaagaccga aaagggcagg acggggcctc cactgcctcc catccctggt	11280
ccggggccac atagccttct ttgtcacaat cagctcaggt atccaagatc agattacca	11340
cattcattat ttgagcaact attcattgaa cagttagaat atgtctact ctgtcagttg	11400
ctggctagaa gtagaaagta ccagatgagt gaaataattg gccactatcc ttggtagctg	11460
atgactaagt aagagagaga tgcaagacaa catgtggaaa atgccaaact gagtagcagt	11520
cacagttgac atgtgcaga gagagctggc cgggggtcag aagacctggg caccagtcct	11580

gttcatttcc agtgtggcct cgagtcattc acctgacctc cctgaagttc attttcccaa	11640
gaagttgttt agtccaactg cccatcaagg atcttttaggg acccttctag ctctaacaga	11700
ggagatcaga aaagaaaaca agcaatgtgg ctgagctcat cctacaagct tcatagagaa	11760
ctgagactgg cciggaagca tagccagaaa ttagaacgcc taagggaaga aggtcacaaac	11820
gctgcctctg caatttagga gtgtatatgc tttcctgcag gatgttgaga gtttcattca	11880
ttatcgtatg cccctaccc cggtcccccaca atacctagtg cgtgggatct gacacgtggt	11940
ggctgggtcaa tgaatgaatg aatgaatggt cacaccatct gaggttctgc actgagtagc	12000
cctgaaggct tgaagcagca taagtacag gtctctccctt gaggggcctc tgttttacc	12060
ataagccaag acctaaagctc aacaacactg aaagggtggc caataccag gacagcctgt	12120
gggaattcca gagaaaggga gattcccagg gactgggggc ccaggctaaa cactgaaaaa	12180
tgcactctgta ggctcaagga ggaaaagccc atgtctgtct gtcttgccca ccactctctc	12240
ccagcaccca gcactgcccc aggacagaga gcacttgaca caagttggtt agattaatga	12300
atgatttaga gttcagtggt cccaacctt ttggcaca gagactggtt gcatggaaga	12360
caatttttcc gcaaaccaag agggggatag agagcattag attctctctt ttttttttt	12420
ttgagaccaa gtctggctct gtctactcag cctggagtaa agtgttgca tctcggtca	12480
ctgcaacctc cgcctctgg attcaagcga ttctctgcc tcagccccct aaatagctgg	12540
gattacaggc acccgtcacc agcccagctg ggactatagg catgtgccac catgccggc	12600
taatttttgt atttttagta gagacggcgt ttcacatgt tggccaggct agtctcgaac	12660
tcctgacctc aggtgatctg cccgcctgag cctcccaaag tgctgggatt acaggcatga	12720
gtcgcctcac ccagcciaaa gtctcataag gaacgtacag catagatccc tcacatgtgc	12780
agttcacaat aaggtttgic tcctacaaga atctaagcc acctctgac tgacaggagg	12840
tgaagctcag gtggteatgc tcgctgtcc ctgccactca cttcctaag tacagccagg	12900
ttcctaacag gccacgaacc agtgggaagg gcatctttt ggatcaaaaa cagaattact	12960
ttttagagaa ctacaagcag atcaatttg ctagacagag actttatatg aaacagcagg	13020
aggctgctag gaggagtga aactctactt tgccctcaag ggagatccc aagggtttg	13080
caggagcggg caaggtggca tgaagaaagc agtgtttgaa atcaggtggt atttgaag	13140
cccagccctt ccccttagaa tggcccttct accatctgtg catggctcca caaccgtggt	13200
ggtggctgcc agaagaattg gaaaggcaga gcatgggtgg agagggggga cctgagggt	13260
ttacaggagt tccgggggtg gtgagggtgt gaaagccagg tcagtcagta ggaagacagg	13320
atgtcagatt gagagactcc cctggccggg gaaacagact tggagaaggg ggagttttg	13380
atgagacagt ccacttccga gtcacaaaat agcttgtggg tgtctgttta ctgttactca	13440

gtgggagtgg ctggggacac gccacctggg cagggtttc gtaattctgc atcacttgtg	13500
aaggtcacag attcccagca caacggacac acccatgttc atagtctgaa ctccaaaca	13560
catcttaaac caaaataaaa aaaaaagaaa gaaagaaaga aaaaggagag ggaggtttga	13620
ggaaagccta tggctctggga cactcaatac ctcccatgaa tatctcatat tgggctggtc	13680
ctctctccac tctggcccca gccataaggg ccttgcttag agcagatfff gggtgctgag	13740
tggaggcagc ctcatcccca acagcctgac ttctgcctc ctcctgcct ctgccttgt	13800
ccagcctgtg ggaagagcct gaagaccccc cgtgtggtgg gtgtggagga ggcctctgtg	13860
gattcttggc cttggcaggt cagcatccag tacgacaaac agcacgtctg tggagggagc	13920
atcctggacc cccactgggt cctcacggca gccactgct tcaggtaga cccagctgt	13980
aaggaggtct ctggggacca aggccagtca gggaccagag agcttgggt cctgtctct	14040
ggcacgtcc ttctcttca tctccacta gagacgttt ccaggttgt gtggcccaa	14100
tgagacaatg gccatgatgc cttttgttag gcttttgggt gtctgagcag aggtgctgg	14160
tcaccaagca tggcctcttc ctggtgggac accagcagat acccagagtc ctcacccac	14220
ccccatatcg ttcaagctac aaaagctctt cccacctgcc tcaacttca agaactcact	14280
ctctttttgc ttgtttccag gaagttgttc cagggtctag agtcatagcc acgtcctcat	14340
tatgtctgga aactttaaaa aaattaaaga gcataggttc ctttcagtcc acagagaagc	14400
ctggccttac ctgagggaag ggctactccc agacccctt cactttttt tttttttt	14460
ttttttttt ttttgagaca gagtcttct ctgttgctta ggctggagcg cagcagcatg	14520
atcttggctc actgcaacct ccgctctctg agttcaagca attctctgc ctcagcttc	14580
caagtagctg ggactatagg catgggccac catgccggc taattttgt atttttgta	14640
gagacagggt ttaccatgt tggccaggct gatctctaac tctgacctc aagtgatctg	14700
cccacctcag cctcccaaac tgctgggatt acaggcatga gccaggcat ccggtttta	14760
tttattcatt cattcaatat ctaatgagca cctaccaggt accaaacacc agatgatgcg	14820
cccaagtcca ttagaccca ccgtgtctt caaggcactc atgatctagg ccagcgttt	14880
ttaaccactt tttttttt ttttttgag attctggtga gagctataaa ttctttctg	14940
gaaaaacatc tctgcacact aagctgtgcc tggcattggg aaaaagaaag cacgtaatgt	15000
aactgacagc atgagtaaca cagtgagaaa ggttggagga gagagcgcca ggacctcaga	15060
actcaggcat tagaggagcc ctttccccag cctccttga ggtttcgtt ggaggtttc	15120
actgaggaaa aagggtcaaa tcccttttc gaatttgact tcttgaagt gccagaagac	15180

tgcccccttct ccaccatccc tgcctcacca tcatctttcc tcccaaggca gtgacatcca	15240
gcaccccgat ccctagggcc ctggggaccc agcctttggc aaagtctcct caggcttgga	15300
tcaggcctga acccagctgt ctctaccccc aggaacata ccgatgtgtt caactggaag	15360
gtgcgggcag gctcagacaa actgggcagc ttcccatccc tggctgtggc caagatcatc	15420
atcattgaat tcaaccccat gtaccccaaa gacaatgaca tcgccctcat gaagctgcag	15480
ttcccactca ctttctcagg tgagaagcag ggcccaaggc cactcaagcc tcttacatca	15540
gttttcacgc ccactctgct attagctcac tgaccgccct tggcacataa tgtctcctct	15600
caagtctca gcttgcccat ttgtctctaa tacgtcagcc taacatcact gatgccatga	15660
ggcctcctca agctgtcagc taacacctcc actccattcc ctgccagaga ttcttccaag	15720
gcctgtcttc cctatgtgga gccctcgag tgagaactgg agtttcatcc aatcttgag	15780
ttttaggaga ctttttaaaa agattatcga gctaattccc caccactgac caacacgcaa	15840
gagcctgctc agtatccctg ccaaggagtc attgtgcccc tgtttgtctt cctccagggg	15900
caggaaccc attacctgtg aggcagccca cagagtcttt gaacagctct gttagatgcc	15960
ttgtgcttat actgaaatgt atttagatca ggattcccaa ctgtggggtc cacaagacac	16020
tgcccccttg gagaagagag gattccattg tcaaataagt ttggggaaca tttcatact	16080
acagctccct tcttggaaca cattagtta ttaaaggtag gagaagtttt taaaataatc	16140
tgttttattg cgtttaacct acatttttta aattttattg accacagaat cttttttca	16200
tgctacttct attagcatcc catagaacaa gtgttctaga gacctggtg tgacctttt	16260
cagagagctt aactgccagg ctctcctgag ccctggtgtg tgtttcaaga tttgtgcctg	16320
ggaattgttt taatcaggta tggcaagggt acagatacag acacagctat ctttgaaaga	16380
agagtttatt attataatt cctgagagaa agggacatac cccaccccc aacacaggga	16440
cacccgggga agcagctggg tccaccagga ggcaggagtg aggggaaggc atggcccaga	16500
gccacctgtg gcttccatgg gcaggtctgg ccaaggtagg gtaggcaaga ttgagcatgc	16560
tcaggattgg atagtgtgga caattctcta ggctatagat gtcagcctct ggttgtctag	16620
tatctgtccc tggggtgatt tagggcaggg aaaatatgg cttggtgtct gagagtcaga	16680
taaaggaagt ggttggggat atgggctttg ggttggctgg tttgcctatt aaaggcgtgc	16740
ccaaagccaa gttgtttact atctgcagga attagctaac ccagtctctc ccagaccagc	16800
aagatcccca taatcataaa gcatcataat ttacagaaaa ttaacactta tgatgaataa	16860
aagatctcct tcttctctg tgctcctggc aggcacagtc aggccatct gtctgccctt	16920
ctttgatgag gagctcactc cagccacccc actctggatc attggatggg gctttacgaa	16980
gcagaatgga ggtaagtctt ggggtgcagga ccacagggca ggagatgccc ttgtatagg	17040

gagcagcttc cagaagtaat gggaaggagg accacccttc agagaaaccc atcctggagg	17100
accaagcacc aaggcgccag gcagaaagca aagtggtttg gcaatccagg gctgggggat	17160
agaaggcaag gatgggaatg tgagtgtttt tacctccca gggaagatgt ctgacatact	17220
gctgcaggcg tcagtccagg tcatigacag cacacggtgc aatgcagacg atgcgtacca	17280
gggggaagtc accgagaaga tgatgtgtgc aggcattccc gaagggggtg tggacacctg	17340
ccagggtgggg cctccaagaa tcatggggag ttctaagaat agggtttagg tcctagagag	17400
atgagaaaac ccagaggctg catgcctac aggaagcctt gcatatcatg ggcactcaat	17460
gtgtgatgat gggaggaaga gagggaggga aggaagagat agtcagataa aagtgtacca	17520
atagatgagt ggggtgatgg atggatgcag acaagcagag agatttcaa tgtctcttc	17580
acattcgaag atgatgttac tggcctggca tgggtggctc cgcttgtaat cccagcactt	17640
tgggaggctg aggcgggcag gtgatttgag gtcaggaatt caagaccagc ctggccaaca	17700
tggtgaaatc ccatctctac taaaaagaat acaaaaatta gctgggcgtg gtggcacgtg	17760
cctgtaatcc cagctacttg ggaggctgag gcaggagaat tgcttgaacc caggaggcag	17820
aggttgacgt aagctgagat tgcgccactg cactccagcc tgggtgaccc agcaagactc	17880
catctgaaaa caacaacaac aacaagatg acattactca tccacccac ccaccctct	17940
cactagctac agaattgatta gccccttgag gtcaggaatc ccaggtctat tttctctgtg	18000
actctccca agctgctgaa ctacactagg aaagaattac cgctgcaga atgctggaag	18060
cacatctgtg tgtgccctca ccccgccctc attggccatc aggactgctt agcaatccct	18120
gtagaccttc ttctccccc atacttccag aggatcttct gaactatctt cttttttat	18180
ttttctttt atgttttta acagagacag ggtcactatg ttgccagtc tggctcaca	18240
ctcctgggtt caagggttc tccacctca gctttccaaa atgctgggat tacaggcatg	18300
agccatcgtg cttggcctga accattttca ttaaaacccc taccctactc tcacctccat	18360
ttccagtcat taaattcctt catttaagag gcatctctta gtcacgcat gtgtgccatg	18420
aacatggtag tctttggaga cccctcaggg agctcacagt ggttggggga aaggggggca	18480
ttaaacagac atttaagcta tagttttggg ttcagaggga ggaagccca ggggctaaaa	18540
cagctgataa ggactccag ataagtgcac ttttactat ctggcatttt cttgtttgt	18600
tatttgcttg ttactgtct ctcacccat ttgatctaa gctttctgag ggcagggatc	18660
tttgttttt ttcatcagtt ggatcccaat tgcttagaac actacctggc acaaaatagg	18720
cactctataa gtgattacac aaattttgga acgactaggt taaacaatga taaccaggct	18780
ttttttttt tttttgagac tgagtctcac tctgttgcac aggetagagt gaagtggttt	18840
gatctcggct cactgcagcc tccgctctg ggttcgaatg attctccacc tcagctcct	18900

gagtagctgg gattacaggt gcctgccact atgcccagct aatttttgta ttgtagtag 18960

agacgggttt caccatgttg gccaggtg tcttgaactc ctgacctcaa gtgattcacc 19020

cgcctcagcc tcccaaggtg ctgggattac aggtgtgagc caccgtcct ggccaacaac 19080

caggcttttt taagacatca ctgagagcct ttaatttgct aatgtgagtt gtgaatctct 19140

gagagaaggc taacggcatg cttgcaactt acttgtccac agacaagcct ttctgcccc 19200

gaagagaaga ccattctagg gtgctaata gcaaagaggg tgagggtgga atacggaga 19260

gcagcaggga gtgcaggga acagataggc cagttcagg agcagagaag gagaagcccc 19320

cccacctcac ctgccctccc cagcagtctc tgttctggtc tctcacagg tgacagtgtt 19380

gggcccctga tgiaccaatc tgaccagtgg catgtgtgg gcatcgtag ttggggctat 19440

ggctgcgggg gcccgagcac cccaggagta tacaccaagg tctcagcta tctcaactgg 19500

atctacaatg tctggaaggt aaggtacctt tgcctaccc actgtgctt ccctccagtc 19560

ctctacctgg ggggtgcca tccatcctca gggttgattt aatggttct gacaactctt 19620

tacatcccaa ataactttcc ctccaagcaa gggacagcct gagattgcac tattaaggct 19680

gaaattcctt aggtcagaga ttcttgataa atgcaaatac cttagggaat agaacacacc 19740

aagcctttct ttctctttc tgacagaatg agactatcag atcctttcta gagagaagat 19800

tctgataagg aagagagtgg aaaggtcat gagacctct ggccctctgc aggttaggga 19860

gagaagcaaa gtgtttcaga aaaggaagac tcacgttaca catgtcacca ctttgtccag 19920

tttcagataa tctgacttct tcttcacgg tctctcttat tctaggctga gctgtaacgc 19980

tgccgtcccc cacatccaga agctgcttcc cttcagacct acctacggca tgaccctca 20040

aagtcagata tgggacaaga gcctccttga acaaacctc 20078

<210> 13

<211> 15159

<212> DNA

<213> Artificial Sequence

<220><223> Recombinant polynucleotide

<400> 13

ccaccgcac aactacagt cgagataact tcgtataatg tatgtatag gaagttatgc 60

tagtaactat aacggctcta aggtagcgag ctgccaagt ctgtgtgcta ccaagtagca 120

aaactgagcc tggaactcac acatgcgtgt ctgagagccc agcactatcg ccaggaaaac 180

ccagcgtctc cctgctcaag cctgacctc agccctctct gcctctccct gcaattgcct 240

tccagtcaag gtgattctgg ataaatacta ctctctctgc gggcagcctc tccacttcat 300

cccgaggaag cagctgtgtg acggagagct ggactgtccc ttgggggagg acgaggagca 360
 ctgtgtcaag agcttccccg aagggcctgc agtggcaggt gagtgcaggg tctgaggcac 420
 aagagaagtg ggcccagcag gaggtctgct caggcccccga cgcccactg catagtatct 480

 gccccctact tgtcacitit catccttgtt gtataaggtt ctttgtttgt ttgtttgttg 540
 ttgttttgag gcagagtgt ctgtggccca agatggagtg cagtgtcttg gtctcggtc 600
 actgcaacct ctgcctccca gtttcaagtg attcttctgc ctcagcctca tgagtagctg 660
 ggattacagg tgccagccac cagcctggc taatttttat atttttagta gagacggggt 720
 ttgtccacat tggtcaggct gatcttgaac tcctgacctc aggtgatctg cccgcctcag 780
 cctcccaaag tgctgggatt acaggcgtga gccaccgtgc ccagctgtgt aagtttcttg 840
 agagcaggac cctgtcttgt ctaccttta atcctagtac ttaacacaca gcaaacagta 900

 actatttgat gaccaaagt gagccagaaa ggacaggaaa ttgtaactga ggctgcccc 960
 tgcgtgctgc gcctgggtga ttccaggcag agggctagac tgggtgacct tggggcattc 1020
 ctcttttcta tgaattttgt tatttcaagg agactagaaa agagacttct cagccacttc 1080
 gccagctatt ggtccttcta ttcatatgt ttgtctgaga catgtatgt gacaggactg 1140
 agccaggctc ttcaatgga taggatgt tttagcata aaatccacgt tctctcttgg 1200
 gctgggctct tctacctct tccccctgt gcttgggctc tgaagaaaaa aagataggta 1260
 ggagatgagt gatggggctt ctgagggcag ggctgagtga ctttctgtgt atttgcctt 1320

 tctttatcag aagtcaaatg cccacaggca cctgtcatcc tactgccagt aggacttctc 1380
 actcaacctt cccctctgac ctactttgga gaaggactta ggtccctctc tcagacattt 1440
 cccaggtctg ggcaagtgt gtggaccatg gatgggtatg tggccatc aatttaaaca 1500
 agctgtatat ggtcgtctgg tagagtacc acataattga tcatcaaac tgatactgt 1560
 aagagcaaaa gggggcacta ttaaccattg ggtcaggga acaggtcaaa atggagacct 1620
 accctgggac ttctgggtac actagctact gtcaaatgg ggcccaata gacaaagcca 1680
 aatggaagaa attcccttga cattgaaagt gttggggctc tgtggcacc ccagtctag 1740

 gtgggggag ctgggctgg tctcatgatg agttctgagg gggatgggc agttgggcc 1800
 cccgttccat ctaactcagg ttcttttct cccagtcgc ctctccaagg accgatccac 1860
 actgcagggt ctggactcg ccacaggga ctggttctct gcctgtttcg acaacttcac 1920
 agaagctctc gctgagacag cctgtaggca gatgggtac agcaggtaac caacctgggc 1980
 ctctctctt tttccctct tctctcttcc tctcttctt ctttcttct cttcttctt 2040
 ctctcttct taaaaattac gggcatigga gccaggcaga atggttttg aatcccagca 2100

tttacttat aagcaacatg aagttaaatt tcctaagcct caggttcctc aggagttaat	2160
tgggggaact aatgccaaacc tcataggata gttttgcaat gccagtgaga gaatgtgtgc	2220
tgccctccaa cacacacaca cacacttcta gcgtctatgc agtcctctcc tttcctttac	2280
tctcaacct tactccttt gtgctggcct tgcaagaaac tgttcctgcc cagtaataca	2340
aaagctaagt taacttattc aaagtttcgt tagttaagat ttagtctaag tgagcctagt	2400
ttcagtgggg ccccatcttc agcaatccca gctctctctg caaatttcaa aagcagttcc	2460
aaatctggag tggatgaaaa ggtgtaagat gatagtaaga gtaatttgca ttctatatat	2520
ttatatccac ttgattttgg cagaaaacca aaaagatagt tattatatct tatatataga	2580
tatatattat atctatttca taaataggct caaacaagt aagtaacttg ctagggtact	2640
agctgggagg tagagggcta gaatttgagc ccaagacccc taattcttgc gcattaggag	2700
ttccacatt gtttctgttt ctgactgag taattcttta ttctcatgta ggacatcatc	2760
tctaagggaa ggggctaata gatggttga tctctcagag agtttagctg gagaggatgg	2820
aaaagaacc atacattcag ttgcagattg agatagccta tctctggcag gcctcagatt	2880
tcttcaggat tctaacagac tggaccaga gactaggcca aacaacaaa caaacaaaaa	2940
ctctactagg cagacatcac caaccaatca cagaactctc tcccatggat ccctaataca	3000
gcctcaaagt ctttttcagt aaatgctcca ggcagccatt acaaatcaat cagaattatt	3060
tgcccttctc ttctctgctc aacgggcttc tgcgtctctc tactttccat agggggcaac	3120
ttccattacc ctctagaaag cacacccac cacttcatt tcaaggagag tgaggaaactc	3180
atgccacga cctgctattc tcccctcttc ctgcagccac ggagcccagc ctgctgcag	3240
ccagccctgc ctccccactg tagtcagtc aactgctgca tcagccgttc ctggcacagc	3300
aggctgagcc ttgattatga aacctgggtg tctccagggg ttcttaagat gataggtcc	3360
tggaatttct gtccttttgg agctcagtaa ggcaccaaac cactgagtc ttgtgcttca	3420
caaatcaaa gtcatcaga atcattcatt gggatggaat tggtaacag aagttaactt	3480
tcctgggaat gtccatttcc accatattcc gtcttcttag gtctcagact tctctacttt	3540
ctttctctc tctagatcgg aggcccttct tgccttagaa ccataggcat ttcaagatgt	3600
gggagaccct agggatcatc tagtcacgc atctttttt ttttttttg acagagtctc	3660
actctgtcac ccaggctgga gtgcaatggc accatctctg ctactgcaa cctccacctc	3720
ccaggttcaa gtgattcttt cgcctcagcc tccaagtag ctgggattac aggcacgcac	3780
catcatgcc agctaatttt tatattttt tagagaccga gtttcacat gttggccagg	3840

ctggtcttga actcctgacc tcaggtgac caccacctc ggctcccaa agtgctggga 3900

ttacaggcgt gagccactgc acccagcccc gtgcattctt ttatagagg ggaaactgag 3960

gcttggagag acccagaaaa agaatatgac ctgcccaagg ccacacatca aactagtgcc 4020

agagccaggg acagaacctg gatcatgagg actcttaaaa tgcactctag tcctcccagg 4080

tctgagactt gggtccttcc aggaagtgcc agcattcctg cctgagaatg tgccaatcca 4140

ccagtattgc caatgactca gccctccatg gagagcttct actaacatta ctagcatagt 4200

tagggatgga aggaaaagat ttagaagagg cagattcagt aaaggaacaa tcagagagat 4260

ggaattaatc aaggaaggct tcctggagga ggaaaaactt caaccaagg ttgaaagta 4320

gcaagcatgg attagcaggg agaaagagg agagtgggcc agttgagaga aacgtttgtc 4380

tggattcata tgaagacaga tctagtcctg ttctattaaa tatctctaag ggggccaaaa 4440

acatacccc gctatcaaa tcagaccaga tgccttgggt ggagaacgaa atatccacat 4500

tccaactccc tcccaggatga gaaggagct aacctgagcc cctatgcctc ttgtttccc 4560

tgctgtgaac cagaagacat tgctgggata ttgaaatag ggacagagct gggaatatgg 4620

aaaggagacc cctaacattt ctccagggtc ctgggttctg gatttggatt cccacccaa 4680

gaaagcaagt tacatcagca atgcactgag ggttgagtc tgggatgcca agggtcggtt 4740

ctttattgta tagcaaagca ggccccatct tctactgacta agaccatctc cactccctgg 4800

ccactccca ccaagcattc tctgccactc ttctctctga aagtgggggc caactctacc 4860

atcttgttct aacccccctgc ccagctcac aactctctct cctcttgat gtgagcagca 4920

aaccacttt cagagctgtg gagattggcc cagaccagga tctggatgtt gttgaaatca 4980

cagaaaacag ccaggagctt cgcatgcgga actcaagtgg gtaagtgagg ggacaccttc 5040

tggcctacag aaggccccc catggacgtc gctcttcagg ttgcaaccag ctacactgga 5100

acccaagca gccaggggaa tgtaagcaga catcaggaag aactcctagc cagatggatc 5160

attcaatgcc aagagctata gactcacatt ttggagagg tttctgtgtt gacttgtttt 5220

taatacaatg gacagctgga caaagtgtgt tgcctactc agagccagag gtagtgataa 5280

tgtgaccttt ccatcaatct ggatagtaaa tagtttttgc tactgctgta ggttttctaa 5340

taaattgccc aataggcaag attccaaagt cactttgtcc ttcctacca ctaccagc 5400

cagagctccc caccttcttg atgctccagg gaagaggctc catggccctt gtgggtggcc 5460

tgttctgag cctcgccacc ctgtgttaga gcagagcatc cagatgaaat ctgtcacact 5520

giggcaaagt ggctcagaga ggagctggc ttcttagcat tcagggacgt tgctgagggc 5580

cgcttattca ccgaaaataa atcttgaaaa ggacagggct ggtagcagaa tgatccttta 5640

cctaaaattc tatcaaaatc ccattcttcc atttgaaag cccacagtgt cacagactct 5700

gttccgggct ctgtcctctt ccctcttggg tcccaggagc ccaggctggg ctttgaagca 5760
ggcagggccc agcacacagt aggtactcag cagtgggggt gttgaatcca atcaaacgga 5820
agtgtcaatg caggaaatgc aatggatgtc aatgcagtct ccaaatgttc cccactgtgc 5880
agcttccaca ttcccagggt attgggaggg gacttgaatt aacagcttcg ggaggcctga 5940

gtccctgcct ccagctgag gaagaagctt aaatcacagg gcgctgtgtc tgtcttcag 6000
gccctgtctc tcaggtccc tggctccct gcactgtctt ggtgagtacc cccaatctct 6060
gagggtttgg ggcttggcc agcaatgagc agggaggaag acctcatct tcaactcctaa 6120
atttctggga ctccaagttt cattctgcct tggctacag cccttgggct tgcgggtcaa 6180
tgccccctcg agttgttggg ggcttgggc aggtcacatt cttttctgg gtctttccaa 6240
gccccagttt ccccttcta ccatctgtgc atggctccat gacctaatg gagacctggg 6300
agagagtgtt aggaagaccg aaaagggcag gacggggcct ccactgcctc ccatccctgg 6360

tccgggcccc catagccttc ttgtcacaa tcagctcagg tatccaagat cagattaccc 6420
acattcatta tttagcaac tattcattga acagttagaa tatgtctcac tctgtcagtt 6480
gttggttaga agtagaaagt accagatgag tgaataaatt ggccactatc cttggtagct 6540
gatgactaag taagagagag atgcaagaca acatgtggaa aatgccaaac tgagtagcag 6600
tcacagttga catgtgcag agagagctgg ccgggggtca gaagacctgg gcaccagtc 6660
tgttcatttc cagtgtggcc tcgagtcatt cacctgacct cctgaagtt cattttccca 6720
agaagttgtt tagtccaact gcccatcaag gatctttagg gaccttcta gctctaacag 6780

aggagatcag aaaagaaaac aagcaatgtg gctcagctca tctacaagc ttcatagaga 6840
actgagactg gcctggaagc atagccagaa attagaacgc ctaagggaag aaggtcacia 6900
cgctgcctct gcaatttagg agtgtatatg ctttctgca ggatgttag agtttcattc 6960
attatcgtat gccccctacc ccggccccac aatacctagt gcgtgggac tgacacgtgg 7020
tggctggta atgaatgaat gaatgaatgg tcacaccatc tgaggttctg cactgagtag 7080
ccctgaaggc ttgaagcagc ataagtgaca ggtcctcct tgaggggcct ctgttttacc 7140
aataagccaa gacctaatg caacaacact gaaaggtgg ccaataacca ggacagcctg 7200

tgggaattcc agagaaagg agattcccag ggactggggg cccaggctaa aactgaaaa 7260
atgcatctgt aggtcaagg aggaaaagcc catgtctgtc tgtcttgcac accactctct 7320
cccagcacc agcactgccc caggacagag agcacttgac acaagttgg tagattaatg 7380
aatgatttag agttcagtg tcccaacct ttttggcaca agagactgg tgcattggaag 7440
acaatttttc cgcaaaccaa gagggggata gagagcatta gattctctct ttttttttt 7500
tttgagacca agtctggctc ttgtactca gcctggagta aagtgttgcg atctcggtc 7560

actgcaacct ccgcctcctg gattcaagcg attctcctgc ctcagccccc taaatagctg 7620

ggattacagg caccggtcac cagcccagct gggactatag gcatgtgcca ccatgcccgg 7680

ctaatttttg tatttttagt agagacggcg ttccaccatg ttggccaggc tagtctcgaa 7740

ctcctgacct caggtgatct gcccgcctga gcctcccaaa gtgctgggat tacaggcatg 7800

agctgcctca cccagcctaa agtctcataa ggaacgtaca gcatagatcc ctcacatgtg 7860

cagttcacia taaggttgtg ctctacaag aatctaacgc cacctctgat ctgacaggag 7920

gtgaagctca ggtgggtcatg ctgcttgtc cctgccactc acttcctaat gtacagccag 7980

gttcctaaca ggccacgaac cagtgggaag ggcatctttt tggatcaaaa acagaattac 8040

tttttagaga actacaagca gatcaatttg gctagacaga gactttatat gaaacagcag 8100

gaggctgcta ggaggagtgg aaactctact ttgccctcaa gggagatccc gaagggtttt 8160

gcaggagcgg gcaaggtggc atgaagaaag cagtgtttga aatcaggtgg tatttgaaaa 8220

gcccagccct tccccttaga atggcccttc taccatctgt gcatggctcc acaaccgtgg 8280

tgggtgctgc cagaagaatt ggaaaggcag agcatgggtg gagagggggg acctgagggc 8340

tttacaggag ttccgggggt ggtgagggtg tgaagccag gtcagtcagt aggaagacag 8400

gatgtcagat tgagagactc ccctggccgg ggaaacagac ttggagaagg gggagttttg 8460

gatgagacag tccacttccg agtcacaaaa tagcttgtgg gtgtctgttt actgttactc 8520

agtgggagtg gctggggaca cgccacctgg gcagggtttt cgtaattctg catcacttgt 8580

gaaggtcaca gattcccagc acaacggaca caccatgtt catagtctga actcctaaac 8640

acatcttaaa ccaaaataaa aaaaaaagaa agaaagaaag aaaaaggaga gggaggtttg 8700

aggaaagcct atggtctggg acactcaata cctcccatga atatctcata ttgggtggtt 8760

cctctctcca ctctggcccc agccataagg gcctgtctta gagcagattt tgggtgctga 8820

gtggaggcag cctcatcccc aacagcctga ctctctgcct cctccctgcc tetgcctgtg 8880

tccagcctgt gggaagagcc tgaagacccc ccgtgtggtg ggtgtggagg aggcctctgt 8940

ggattcttgg ctttggcagg tcagcatcca gtacgacaaa cagcacgtct gtggaggagg 9000

catcctggac cccactggg tctcacggc agccactgc ttcaggtaag accccagctg 9060

taaggaggtc tctggggacc aaggccagtc agggaccaga gagcttgggg tctgtctcc 9120

tggcacctgc ctctcttca ctctccact agagacgttt tccaggttgt ggtggcccca 9180

atgagacaat ggccatgatg ccttttgtta ggcttttggg tgtctgagca gaggggtctg 9240

gtaccaagc atggcctctt cctggtggga caccagcaga taccagagt ctcaccca 9300

ccccatatac gttaagcta caaaagctct tcccacctgc ctcaacttcc aagaactcac	9360
tctctttttg ctgttttcca ggaagttgtt ccagggtcta gagtcatagc cacgtcctca	9420
ttatgtctgg aaactttaaa aaaattaaag agcatagggt cctttcagtc cacagagaag	9480
cctggcctta cctcaggga gggctactcc cagacccctc tcaacttttt ttttttttt	9540
tttttttttt tttttgagac agagtcctgc tctgttgctt aggcctggagc gcagcagcat	9600
gatcttggt cactgcaacc tccgcctcct gagttcaagc aattctcctg cctcagcttc	9660
ccaagtagct gggactatag gcatgggcca ccatgcccg ctaatttttg ttttttggt	9720
agagacaggg tttcaccatg ttggccaggc tgatctctaa ctctgacct caagtgatct	9780
gcccacctca gcctccaaa ctgctgggat tacaggcatg agccaggga tccggctttt	9840
atttattcat tcattcaata tctaagagc acctaccagg taccaaacc cagatgatgc	9900
gccaagtgc attagacccc accgctgtct tcaaggcact catgatctag gccagcgttt	9960
tttaaccact ttttttttt tttttttga gattctggtg agagctataa attctttcct	10020
ggaaaaacat ctctgcacac taagctgtgc ctggcattgg gaaaaagaaa gcacgtaatg	10080
taactgacag catgagtaac acagtgagaa aggttgagg agagagcgcc aggacctcag	10140
aactcaggca ttagaggagc ccttcccca gccctccttg aggtttcgtt gggcaggttt	10200
cactgaggaa aaaggtcaa atcccttttt cgaatttgac ttcttgtaag tgccagaaga	10260
ctgccccctc tcaccatcc ctgctcacc atcatcttc ctccaaggc agtgacatcc	10320
agcaccctga tccctagggc cctggggacc cagcctttgg caaagtctcc tcaggcttgg	10380
atcaggcctg aaccagctg tctctacccc caggaaacat accgatgtgt tcaactggaa	10440
ggtgctggca ggctcagaca aactgggcag ctcccatcc ctggctgtgg ccaagatcat	10500
catcattgaa ttcaaccca tgtacccaa agacaatgac atgcctca tgaagctgca	10560
gttccactc actttctcag gtgagaagca gggcccaagg ccaactcaagc ctcttacatc	10620
agttttcag cccactctgc tattagctca ctgaccgcc ttggcacata atgtctctc	10680
tcaagctctc agcttgccca ttgtctctc atactcagc ctaacatcac tgatgccatg	10740
aggcctctc aagctgtcag ctaacacctc cactccatc cctgccagag attcttccaa	10800
ggcctgtctt ccctatgtgg agccctcga gtgagaactg gagtttcac caatcttga	10860
gttttaggag accttttaa aagattatcg agctaattcc ccaccactga ccaacacgca	10920
agagcctgct cagtatccct gccaaaggat cattgtgccc ctgtttgtc tctccaggg	10980
gcagggaacc cattacctgt gaggcagccc acagagtctt tgaacagctc tgttgatgc	11040
cttgtctta tactgaaatg ttttagatc aggtatccca actgtgggt ccacaagaca	11100
ctggccctt ggagaagaga ggattccatt gtcaataaag ttggggaac attttcatac	11160

tacagctccc ttcttggaac acattagttt attaaaggta ggagaagttt ttaaaataat	11220
ctgttttatt gcgtttaacc tacatttttt aaattttatt gaccacagaa tccttttttc	11280
atgctacttc tattagcatc ccatagaaca agtgttctag agaccctggt gtgacccctt	11340
tcagagagct taactgccag gctctcctga gccctggtgt gtgtttcaag atttgtgcct	11400
gggaattggt ttaatcaggt atggcaaggt gacagatata gacacagcta tctttgaaag	11460
aagagtttat tatttataat tcctgagaga aaggacata cccaccccc caacacaggg	11520
acaccggggg aagcagctgg gtccaccagg aggcaggagt gaggggaagg catggcccag	11580
agccacctgt ggcttccatg ggcaggtctg gccaaagtag ggtaggcaag attgagcatg	11640
ctcaggattg gatagtgtgg acaattctct aggctataga tgtcagcctc tggttgtcta	11700
glatctgtcc ctggggtgat ttagggcagg gaaaatattg gcttggtgtc tgagagtcag	11760
ataaaggaag tggttgggga tatgggcttt gggttggtctg gtttgcctat taaaggcgtg	11820
cccaaagcca agttgtttac tatctgcagg aattagctaa ccagtcctc cccagaccag	11880
caagatcccc ataatacata agcatcataa ttacagaaa attaacactt atgatgaata	11940
aaagatctcc ttcttctct gtgctcctgg caggcacagt caggcccatc tgtctgccct	12000
tctttgatga ggagctcact ccagccacc cactctggat cattggatgg ggctttacga	12060
agcagaatgg aggtaaagtc tgggtgcagg accacagggc aggagatgcc ctgtgatgag	12120
ggagcagctt ccagaagtaa tgggaaggag gaccaccctt cagagaaacc catcctggag	12180
gaccaagcac caaggcgcca ggcagaaagc aaagtggttt ggcaatccag ggctggggga	12240
tagaaggcaa ggatgggaat gtgagtgttt ttaccctccc agggaagatg tctgacatac	12300
tgctgcaggc gtcagtccag gtcattgaca gcacacggtg caatgcagac gatgcgtacc	12360
agggggaagt caccgagaag atgatgtgtg caggcatccc ggaagggggt gtggacacct	12420
gccaggtggg gcctccaaga atcatgggga gttctaagaa tagggtttag gtcctagaga	12480
gatgagaaaa ccagaggct gcatgcccta caggaagcct tgcatatcat gggcactcaa	12540
tgtgtgatga tgggaggaag agaggagagg aaggaaagga tagtcagata aaagtgtacc	12600
aatagatgag tgggtggatg gatggatgca gacaagcaga gagatttcaa atgtctcttt	12660
cacattcgaa gatgatgta ctggcctggc atggtggctc acgctttaa tcccagcact	12720
ttgggaggct gaggcgggca ggtgatttga ggtcaggaat tcaagaccag cctggccaac	12780
atggtgaaat cccatctcta ctaaaaagaa taaaaaatt agctgggcgt ggtggcacgt	12840
gcctgtaatc ccagctactt gggaggctga ggcaggagaa ttgcttgaac ccaggaggca	12900
gaggttcgag taagctgaga ttgcgccact gcactccagc ctgggtgacc cagcaagact	12960
ccatctgaaa acaacaaca caacaaagat gacattactc atccaccca cccacccttc	13020

tcactagcta cagaatgatt agcccccttga ggtcaggaat cccaggtcta ttttctctgt	13080
gactctcccc aagctgctga actacactag gaaagaatta ccgcctgcag aatgctggaa	13140
gcacatctgt gtgtgccctc accccggcct catggccat caggactgct tagcaatccc	13200
tgtagacctt ctctctcccc catacttcca gaggatcttc tgaactatit tcttttttta	13260
ttttttcttt tatgtttttt aacagagaca gggtcactat gttgccagct ctggtctcaa	13320
actcctgggt tcaagggatt ctcccacctc agctttccaa aatgctggga ttacaggcat	13380
gagccatcgt gcttggcctg aaccattttc attaaaacc ctaacctact ctcacctcca	13440
tttccagtca ttaaattcct tcatttaaga ggcatctctt agtcacgca tgtgtgccat	13500
gaacatggta gtctttggag acccctcagg gagctcacag tggttggggg aaaggggggc	13560
attaaacaga catttaagct atagtitttg gttcagaggg aggaagcccc aggggctaaa	13620
acagctgata aggactccca gataagtgca cttttcacta tctggcattt tcttgttttg	13680
ttatttgctt gttcactgtc tctcacccca ttgatccta agctttctga gggcagggat	13740
ctttgttttt tttcatcagt tggatcccaa ttgcttagaa cactacctgg cacaaaatag	13800
gcactctata agtgattaca caaatttttg aacgactagg ttaaacaatg ataaccaggc	13860
tttttttttt ttttttgaga ctgagtctca ctctgttgcc caggctagag tgaagtgggt	13920
tgatcteggc tcaactgcagc ctccgcctct gggttcgaat gattctccac ctacgcctcc	13980
tgagtagctg ggattacagg tgccctgccac tatgccagc taatttttgt atttgiagta	14040
gagacgggtt tcaccatgtt ggccaggctg gtcttgaact cctgacctca agtgattcac	14100
ccgcctcagc ctcccaaggt gctgggatta cagggtgtgag ccaccgctcc tggccaacaa	14160
ccaggctttt ttaagacatc actcagagcc ttttaattgc taatgtgagt tgtgaatctc	14220
tgagagaagg ctaacggcat gcttgcaact tacttgtcca cagacaagcc tttctgcccc	14280
agaagagaag accattctag ggtgctaatt agcaaagagg gtgagggtgg aatatcgag	14340
agcagcaggg agtgcagggg aacagatagg ccagttcagg gagcagagaa ggagaagccc	14400
ccccacctca cctgccctcc ccagcagctc ctgttctggt ctctcacagg gtgacagtgg	14460
tgggcccctg atgtaccaat ctgaccagtg gcatgtggtg ggcatcgta gttggggcta	14520
tggctgcggg ggcccagca cccaggagt atacaccaag gtctcagcct atctcaactg	14580
gatctacaat gtctggaagg taaggtacct ttgccctacc cactgtgcct tccctccagt	14640
cctctacctg ggggggtcca atccatctc aggtttgatt taaatggttc tgacaactct	14700
ttacatccca aataactttc cctccaagca agggacagcc tgagattgca ctattaaggc	14760

tgaatttcct taggtcagag atttctgata aatgcaaata ccttagggaa tagaacacac 14820
 caagcctttc ttctctttt ctgacagaat gagactatca gacctttct agagagaaga 14880
 ttctgataag gaagagagtg gaaaggtca tgagacctcc tggccctctg cagggtaggg 14940
 agagaagcaa agtgtttcag aaaaggaaga ctcacgttac acatgtcacc actttgtcca 15000
 gtttcagata atctgacttt ctcttcacg gtctctctta ttctaggctg agctgtaacg 15060
 ctgccgtccc ccacatccag aagctgcttc ccttcagacc tacctacggc atgacccctc 15120
 aaagtcagat atgggacaag agcctccttg aacaaactc 15159

<210> 14

<211> 435

<212> PRT

<213> Artificial Sequence

<220><223> Recombinant protein

<400> 14

Met Glu Ser Asp Ser Gly Gln Pro Leu Asn Asn Arg Asp Ile Val Pro

1 5 10 15

Phe Arg Lys Pro Arg Arg Pro Gln Glu Thr Phe Lys Lys Val Gly Ile

20 25 30

Pro Ile Ile Ala Val Leu Leu Ser Leu Ile Ala Leu Val Ile Val Ala

35 40 45

Leu Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Leu Cys Gly Gln

50 55 60

Pro Leu His Phe Ile Pro Arg Lys Gln Leu Cys Asp Gly Glu Leu Asp

65 70 75 80

Cys Pro Leu Gly Glu Asp Glu Glu His Cys Val Lys Ser Phe Pro Glu

85 90 95

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

100 105 110

Val Leu Asp Ser Ala Thr Gly Asn Trp Phe Ser Ala Cys Phe Asp Asn

115 120 125

Phe Thr Glu Ala Leu Ala Glu Thr Ala Cys Arg Gln Met Gly Tyr Ser

130 135 140

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

145 150 155 160
 Asp Val Val Glu Ile Thr Glu Asn Ser Gln Glu Leu Arg Met Arg Asn
 165 170 175
 Ser Ser Gly Pro Cys Leu Ser Gly Ser Leu Val Ser Leu His Cys Leu

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

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

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

 370 375 380
 Ser Gly Gly Pro Leu Met Tyr Gln Ser Asp Gln Trp His Val Val Gly
 385 390 395 400

Ile Val Ser Trp Gly Tyr Gly Cys Gly Gly Pro Ser Thr Pro Gly Val

405 410 415

Tyr Thr Lys Val Ser Ala Tyr Leu Asn Trp Ile Tyr Asn Val Trp Lys

420 425 430

Ala Glu Leu

435

<210> 15

<211> 2046

<212> DNA

<213> Mus musculus

<400> 15

cagaaacaag gacctcttca ttattcaaga gtaaaatgta taggccaaga ccaatgctat	60
caccgtcaag attcttcaact ccctttgcag tagctttcgt tgtcataata acggtagggc	120
tcctggccat gatggcaggc ctacttattc actttttagc ttttgacaag aaagcttact	180
tttatcatag cagctttcaa atcctaaacg ttgaatacac tgaggcttta aactcaccag	240
ctacacacga atacagaacc ttgagtgaag gaattgaggc tatgattact gatgaatttc	300
gaggatcaag tctaaaaagt gagtttatca ggacacatgt tgtcaaaact agaaaagaag	360
ggactggtgt gggtgcggat gttgtcatga aatttcgac tagtaaactg aacaacagaa	420
aggtaatgaa aaccagaatt caatctgtgc tacgaagact cagcagctct ggaaacttgg	480
aaatagcccc ttcgatgag ataacatcac tcaactgacca ggatacagaa aatgttttga	540
ctcaagaatg tggagcagct ccagacctta taacactgtc agaagagaga atcattggag	600
gcatgcaagc tgagcccggt gactggccct ggcaagtcag tctacagctc aataatgtcc	660
accactgtgg aggtgccctg atcagtaaca tgtgggtcct gacagcagct cattgcttca	720
aaagctatcc taatctcaa tattggacag ccacctttgg ggtttctaca atgagcccta	780
ggctgagagt gagagtaagg gctattttag ccacgacgg gtacagctcc gtaactcgtg	840
acaatgacat cgagtttga caacttgaca gatctgtcgc cttttccaga aatatccata	900
gggtatgtct ccagcagca acccaaaata tcatccctgg ttctgtcgca tatgttacag	960
gatggggatc tctcacatat ggaggcaacg cagtcacaaa tctacggcaa ggagaggcca	1020
gaataataag ttcagaggaa tgcaatacgc cagctggtta cagtgggaagt gtcttgccag	1080
gaatgtctgt tgctggaatg cgttcagggg ccgtggatgc atgccagggt gattcaggtg	1140
gcccgtagt acaagaagac tcaaggcggc ttggtttgt tgtgggcatt gtgagctggg	1200

gatatcagtg tggcctccca aataagccag gcgtgtatac tcgagtgaca gcctaccgca 1260

actggatcag acagcagacg ggaatctagt gcaaccgagg aaaaaacgtg ccatgaggtc 1320

tctgtatcca agtgtgactg actcggatgc catggcttca catttcaact gcaaaggaga 1380

ctggaaatgc cccttctgaa cgtccattta cataaatatg gtttaactgt ttagtatitc 1440

tttgtcggtg cagatittta ctttcttgag gaaaaaaaaa acatgaacat ggctaagtaa 1500

gaattatgtt aggctagtaa caggaagaca ttattacat ggggtgtcag gtgtagtagt 1560

gagaagtcag gtaagttaag tcaataattt acagaaaata atgtcaggtg gtcctaactg 1620

taaatatgtg aggccacaga acaaatagtg ttagaactga agccatccca agtatttaac 1680

atttgttttc aagtgaact aagaacaga cttacatata gttttaatgg tgaattttca 1740

ttttaaatat tttatctaca tagaaaagac atatctcctt catgaagaag ctgaggtgat 1800

gaatcaacac agcctcttca gctatgtttg caaccacaag atttgtggga aagaaatccc 1860

tactaccaac ttctactgt tggcattatt ttttagagta acacgacgca caatagcaaa 1920

atttaagtaa caaattaaaa gttaatgatg aagaagaagt aaagagtttg ttgcaaaga 1980

caaaaattaa acagattaat atcaataaat ctggagacag aagggtctca gattcatatt 2040

ctctct 2046

<210> 16

<211> 417

<212> PRT

<213> Mus musculus

<400> 16

Met Tyr Arg Pro Arg Pro Met Leu Ser Pro Ser Arg Phe Phe Thr Pro

1 5 10 15

Phe Ala Val Ala Phe Val Val Ile Ile Thr Val Gly Leu Leu Ala Met

20 25 30

Met Ala Gly Leu Leu Ile His Phe Leu Ala Phe Asp Lys Lys Ala Tyr

35 40 45

Phe Tyr His Ser Ser Phe Gln Ile Leu Asn Val Glu Tyr Thr Glu Ala

50 55 60

Leu Asn Ser Pro Ala Thr His Glu Tyr Arg Thr Leu Ser Glu Arg Ile

65 70 75 80

Glu Ala Met Ile Thr Asp Glu Phe Arg Gly Ser Ser Leu Lys Ser Glu

85 90 95
 Phe Ile Arg Thr His Val Val Lys Leu Arg Lys Glu Gly Thr Gly Val
 100 105 110
 Val Ala Asp Val Val Met Lys Phe Arg Ser Ser Lys Arg Asn Asn Arg
 115 120 125
 Lys Val Met Lys Thr Arg Ile Gln Ser Val Leu Arg Arg Leu Ser Ser
 130 135 140
 Ser Gly Asn Leu Glu Ile Ala Pro Ser Asn Glu Ile Thr Ser Leu Thr
 145 150 155 160
 Asp Gln Asp Thr Glu Asn Val Leu Thr Gln Glu Cys Gly Ala Arg Pro
 165 170 175
 Asp Leu Ile Thr Leu Ser Glu Glu Arg Ile Ile Gly Gly Met Gln Ala
 180 185 190
 Glu Pro Gly Asp Trp Pro Trp Gln Val Ser Leu Gln Leu Asn Asn Val
 195 200 205
 His His Cys Gly Gly Ala Leu Ile Ser Asn Met Trp Val Leu Thr Ala
 210 215 220
 Ala His Cys Phe Lys Ser Tyr Pro Asn Pro Gln Tyr Trp Thr Ala Thr
 225 230 235 240
 Phe Gly Val Ser Thr Met Ser Pro Arg Leu Arg Val Arg Val Arg Ala
 245 250 255
 Ile Leu Ala His Asp Gly Tyr Ser Ser Val Thr Arg Asp Asn Asp Ile
 260 265 270
 Ala Val Val Gln Leu Asp Arg Ser Val Ala Phe Ser Arg Asn Ile His
 275 280 285
 Arg Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Ile Pro Gly Ser Val
 290 295 300
 Ala Tyr Val Thr Gly Trp Gly Ser Leu Thr Tyr Gly Gly Asn Ala Val
 305 310 315 320
 Thr Asn Leu Arg Gln Gly Glu Val Arg Ile Ile Ser Ser Glu Glu Cys
 325 330 335

Asn Thr Pro Ala Gly Tyr Ser Gly Ser Val Leu Pro Gly Met Leu Cys
 340 345 350
 Ala Gly Met Arg Ser Gly Ala Val Asp Ala Cys Gln Gly Asp Ser Gly
 355 360 365
 Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Val Val Gly

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

<210> 17

<211> 2800

<212> DNA

<213> Homo sapiens

<400> 17

atttgagtgg gaatctcaaa gcagttgagt aggcagaaaa aagaacctct tcattaagga 60
 ttaaaatgta taggccagca cgtgtaactt cgacttcaag atttctgaat ccatatgtag 120

tatgtttcat tgcgtcgca ggggtagtga tcttggcagt caccatagct ctactgttt 180
 acttttttagc ttttgatcaa aaatcttact tttataggag cagtitttcaa ctctaaatg 240
 ttgaatataa tagtcagtta aattcaccag ctacacagga atacaggact ttgagtggaa 300
 gaattgaatc tcigattact aaaacattca aagaatcaaa ttttaagaaat cagttcatca 360
 gagctcatgt tgccaaactg aggcagatg gtagtggtgt gagagcggat gttgtcatga 420
 aatttcaatt cactagaaat aacaatggag catcaatgaa aagcagaatt gagtctgttt 480
 tacgacaaat gcigaataac tctggaaacc tggaaataaa cccttcaact gagataacat 540

cacttactga ccaggctgca gcaaattggc ttattaatga atgtggggcc ggtccagacc 600
 taataacatt gtctgagcag agaatccttg gaggcactga ggctgaggag ggaagctggc 660
 cgtggcaagt cagtctgcgg ctcaataatg cccaccactg tggaggcagc ctgatcaata 720
 acatgtggat cctgacagca gctcactgct tcagaagcaa ctctaactct cgtgactgga 780
 ttgccacgtc tggatatttc acaacatttc ctaaactaag aatgagagta agaaatattt 840
 taattcataa caattataaa tctgcaactc atgaaaatga cattgcactt gtgagacttg 900

agaacagtgt cacctttacc aaagatatcc atagtgtgtg tctcccagct gctaccaga 960

atattccacc tggtcttact gcttatgtaa caggatgggg cgctcaagaa tatgctggcc 1020

acacagttcc agagctaagg caaggacagg tcagaataat aagtaatgat gtatgtaatg 1080

caccacatag ttataatgga gccatcttgi ctggaatgct gtgtgctgga gtacctcaag 1140

gtggagtgga cgcatgtcag ggtgactctg gtggcccact agtacaagaa gactcacggc 1200

ggctttgggt tattgtgggg atagtaagct ggggagatca gtgtggcctg cgggataagc 1260

caggagtgtg tactcgagtg acagcctacc ttgactggat taggcaacaa actgggatct 1320

agtgaacaaa gtgcatccct gttgcaaagt ctgtatgcag gtgtgcctgt cttaaattcc 1380

aaagctttac atttcaactg aaaaagaaac tagaaatgtc ctaatttaac atcttgttac 1440

ataaatatgg ttaacaaac actgtttaac ctttctttat tattaaggt tttctatfff 1500

ctccagagaa ctatatgaat gttgcatagt actgtggctg tgtaacagaa gaaacacact 1560

aaactaatta caaagttaac aatttcatta cagtgtgtgt aaatgcccgt agtgagaaga 1620

acaggaacct tgagcatgta tagtagagga acctgcacag gtctgatggg tcagaggggt 1680

cttctctggg tttcactgag gatgagaagt aagcaaactg tggaaacatg caaaggaaaa 1740

agtgatagaa taatattcaa gacaaaaaga acagtatgag gcaagagaaa taatatgtat 1800

ttaaaatfff tggttactca atatcttata cttagtatga gtcctaaaat taaaaatgtg 1860

aaactgttgt actatacgta taacctaac ttaattattc tgtaagaaca tgcttccata 1920

ggaaatagtg gataatfff agctatftaa ggcaaaagct aaaatagttc actcctcaac 1980

tgagacccaa agaattatag atatfffca tgatgacca tgaaaaatat cactcatcta 2040

cataaaggag agactatatt tattttatag agaagctaag aaatatacct acacaaactt 2100

gtcaggtgct ttacaactac atagtacttt ttaacaacaa aataataatt ttaagaatga 2160

aaaatftaat catcgggaag aacgtccac tacagacttc ctatcactgg cagttatatt 2220

tttgagcgta aaagggtcgt caaacgctaa atctaagtaa cgaattgaaa gtttaaagag 2280

ggggaagagt tggtttgcaa aggaaaagt taaatagctt aatatcaata gaatgacct 2340

gaagacagaa aaaactttgt cactcttctt ctctcatfff ctttctctct ctctccctt 2400

ctcatacaca tgcttcccc accaaagaat ataatgtaaa ttaaatccac taaaatgtaa 2460

tggcatgaaa atctctgtag tctgaatcac taatatctct gagttttat gagtcctag 2520

tacagctaaa gtttgcctat gcatgatcat ctatgcgtca gagcttctct cttctacaag 2580

ctaactccct gcatctgggc atcaggactg ctccatacat ttgctgaaaa cttcttgtat 2640

ttcctgatgt aaaattgtgc aaacacctac aataaagcca tctactttta gggaaaggga 2700
gttgaaaatg caaccaactc ttggcgaact gtacaaacaa atctttgcta tactttattt 2760
caaataaatt ctttttaaaa taaaaaaaaa aaaaaaaaaa 2800

<210> 18

<211> 418

<212> PRT

<213> Homo sapiens

<400> 18

Met Tyr Arg Pro Ala Arg Val Thr Ser Thr Ser Arg Phe Leu Asn Pro

1 5 10 15

Tyr Val Val Cys Phe Ile Val Val Ala Gly Val Val Ile Leu Ala Val

20 25 30

Thr Ile Ala Leu Leu Val Tyr Phe Leu Ala Phe Asp Gln Lys Ser Tyr

35 40 45

Phe Tyr Arg Ser Ser Phe Gln Leu Leu Asn Val Glu Tyr Asn Ser Gln

50 55 60

Leu Asn Ser Pro Ala Thr Gln Glu Tyr Arg Thr Leu Ser Gly Arg Ile

65 70 75 80

Glu Ser Leu Ile Thr Lys Thr Phe Lys Glu Ser Asn Leu Arg Asn Gln

85 90 95

Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val

100 105 110

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

115 120 125

Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn

130 135 140

Asn Ser Gly Asn Leu Glu Ile Asn Pro Ser Thr Glu Ile Thr Ser Leu

145 150 155 160

Thr Asp Gln Ala Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly

165 170 175

Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu

180 185 190

Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn
195 200 205

Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr
210 215 220

Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala
225 230 235 240

Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg
245 250 255

Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp
260 265 270

Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile
275 280 285

His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser
290 295 300

Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr
305 310 315 320

Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val
325 330 335

Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu
340 345 350

Cys Ala Gly Val Pro Gln Gly Gly Val Asp Ala Cys Gln Gly Asp Ser
355 360 365

Gly Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Ile Val
370 375 380

Gly Ile Val Ser Trp Gly Asp Gln Cys Gly Leu Pro Asp Lys Pro Gly
385 390 395 400

Val Tyr Thr Arg Val Thr Ala Tyr Leu Asp Trp Ile Arg Gln Gln Thr
405 410 415

Gly Ile

<210> 19

<211> 38992

<212> DNA

<213> Artificial Sequence

<220><223> Recombinant polynucleotide

<400> 19

```

gagggagggt ggtgctttgc taatggtgaa ttactaactc ctcaataaag aatattattt      60
gaaataattt ttgaaatttc ataattactt tgggttcttt cttaatgata aataaataat      120
agtatattac aaacatacat taatatttcc tgaatgaata caccacaaat ctcccttaaa      180
atatagcaag aataaaaaatt atactatttc tgacaatttt taatttctca aataataata      240
ccactctgat ttttaacat ctacaccact ctggctttgc caatcttttt aaaaattgaa      300

aagataataa ttttatcata attacactga agcatagaac ttttctttc aaggaaagca      360
aatTTTTgaa attctataat ataacctccc ataatcctga ataaattaaa ggttcaacaa      420
cttagtaaag taagactgac ctcccttttt atttcttttt cagatcaaaa atcttacttt      480
tataggagca gttttcaact cctaaatgtt gaataataata gtcagttaaa ttcaccagct      540
acacaggaat acaggacttt gagtggaaga attgaatctc tggtaagtta atatttgtct      600
ttgctcttta ttccattata aaatgaatat gataataaac ctaatgtttt gtaatatatt      660
ttcagttgct aagtgtctca cataatttcc ttcccttgaat ggtgaaacat gtgtttctct      720

ctgcttttat ccagttagtt tactcatata ctggttctta ttcacatctt tgtcatgagt      780
aaaaagtgtt agaaaggcca cgagtaaata tgcattttat ttgtttatga attcaaatac      840
taaaagtttt ttatttgttt aattaagcat tgacattgtc tttttaaatt cttttcattt      900
taccttcttc cctcttctt atccaactaa agacgcaaag caggaggtgt taaaaaacag      960
gtttaccata tcagcagtaa catagtttgg acaacattac actttggttc aatgatagac      1020
atagaagttt gaacagaaat atgcaaagca agtttgagct ctaacttgaa gagagcctct      1080
gggtgcctgc caggaaacct cagagtgga cccttaacat tcatgtgtca ccacaaacta      1140

ggggctgccc tttagttttg accagtctca gtgtcactca cttaccctta ctttttcaaa      1200
aaaaagtcct aagaataata agtaattcaa tggttctaca attttagcat gtaactgagt      1260
cacctggcag ggttgctttg gtgagctcaa gataaaattt tatcagcatt tctacatttt      1320
ctggaatatt ccttaatcca ggcttttaat cccttggtgc ttttctgaac cactgcaatg      1380
agcttctaac tgttctcact gtgtgcaggc tcttttctt ctaatctaatt ttacacactt      1440
ctgaacacaa atctctcaca gctgtttcc ttcattgtac ctccagctca agactttttg      1500
cctacaaaat aaaattcaaa cttgttagct aagcaccttc tcatgtctat gctttggctc      1560

```

atatttcagc catcgtgtgc cccacttatt cttatagcca acctgaaaag ccatctttta	1620
taagaaacta cctctgctct ccatgattgg atataattaa tctccttcc acatcacctc	1680
gccacaaaat tgtatctgtg ttgatctcat gccacatacc tgtatgtatt ttatattata	1740
aatatttgca gacttgttta atttgccatg ttagactaag ttccatgaag acagctccat	1800
atccattcca tttttatata tccacaacat ttggtcgggt tgatgcttaa taaatgttta	1860
ttgaaggaaac aggagtctcc cacttctgac ataatgaact tatttcccc agtggttaacc	1920
ctacatctgg ttctctgcca agagtctctt cccaaatcat tctgattcaa ctgttcattc	1980
tgatctcatt aaacatttaa atgatataac taacttcgct tgctttatc tatgctcatc	2040
ctgcagtctc ctcataactt gggttcaatg atgcttgctt ctagagaaaa aaatgtatta	2100
aataagctta tgattcagtc ctccagctgt gatggttctc actgaacatt agctcagtgg	2160
ttttcgaagt atggctctca gcataaccta gaaacttgtt agaaatgcaa attcttgggc	2220
tcaccaagac atactaaac aaaaattctg acattggggc ctagaaatct gtgttttaac	2280
aagcctgcca gtgcagcctg gtccctttc ttctcggagc cccactcaa gctttcagtg	2340
ctcatctccc accaatgaca gggctctcta tggaaaccgg caggacggtt tccaactcta	2400
actacgtttt agagtttgct tcctagggct atccaggcac caaglatcac aggttagttt	2460
cccagggag cagactctga gacttgcatg caggagtggt ctctgggtg ctctcaacca	2520
acaccttcag gaagagaagg aagcagcatt gggcagaggc atagtcaaac tacagtctg	2580
ttggcacaga agactgaagg gagtcagagc caggggtag aggtgggcc ttagcatcca	2640
tccttcacca ttaggtgtga gttgccccac ctcttgatg gtgtaacctc agtccaagg	2700
tgggtgggag tgcagcagag cagcccctac aagggccaaa ccagagatac accaggcgcc	2760
agaagtgtg ccagggata gagaggaaag gatgggctta aggtaggatc cacagaactt	2820
ggcaatggat tagaagacag gatgagaagt gacaggttaa cactaacaca gaaatgtcta	2880
acttcggtag ataattgtgc cattggctag aagaggaaac cgaaatgaaa gcaggttgtt	2940
caggagaca aaagttcact gtggacatct cagcagagtg attcagtggg gaaaggaatg	3000
gatgcccaga ccacctcaga ggaagatcta agctggagcc agcaataaag atacaagatg	3060
aacaatccct aacgaactgc tctcagcca tgctcccag acacgtgct tcagatttat	3120
agtccgggtg aggctaggag gtgcgcctcc ctcagtggag gacagcaaag caccagtggc	3180
tccaggagat taaaatcttt tgataatttt tgttctagca tctgtctgca gagctgtctc	3240
tcagccattg cctgccttta cacaggagtg cagtccgaaa ttgggagatg agtgaaattt	3300
attatgccta gagatctgga tccccagttg ttgggagta tattttctga accattgtt	3360
ggtttaagta atgcagattt attgatgcca ctctcttga atctgtgact ctggaccac	3420

catctaagtg aatgtgcaga gggaacggaa tggctgcaat agatctccat taaaaccagt 3480
gcacacctcc agacacatac agtagtaggg aggtgagtca atgtcaggac agcaccagct 3540
cccgttcgg tacatttcca aagtctcag tctgtgtaca aaggtttgct ctggggcagc 3600
agaaatagcc ctgggcagggt agtcaaaggc ctggtttgat ttcctccact tccaggcaag 3660

tcactcgaag gtcacaggc ttttctca cctgccacat ggggccagt agatctactg 3720
agctgtaaat aatgaaatga gtgtgtgtgc agtcactat aagttgtaa gtactagaaa 3780
atggtgaaac ttgggattt gggctattta aggtgaaatg ctaaaatgt caggcattgt 3840
ggagaaagga atttaaatat aagattgatt gactgggatt taaagacaaa tgaaggcaca 3900
cacgcaagtg cacaccaca ctgacactgc acagctcccg ttggaggcat atcctgacca 3960
tgcagacctg gggctctgcc tgtccaagt cactccttta ctacataaac cctccttctc 4020
ttttggggct gtcacccac cagagctggc accgagccct tgcgtctgcg ctccctggg 4080

gtgtcagctt ttgacagggt gtttctccc tctgcaggag ccttaacatc cttggactt 4140
ccttcccccc accaccccc agcagtttta tctcttcta actcgggacc cttttttcc 4200
cacacaaagt ttattgtcag ttgctgggtt catctgtttg agcggctgca acaaaatacc 4260
atagactggg tggcatatgc acgacaaaaa ttatttctc acaggagaag tcaaagatta 4320
atgcaccagc agatctggtg tctgaggggc caccttctgg ttgtagatg atgctttcta 4380
gttaaaacac ctatttaaca cactattaaa cactaagtgt gttaaatagt gcagttgatg 4440
tatttgcatt gtcaccttta tcatacata aatccttctt tgtcttttt tctgtactct 4500

aatctctttc tgtaagtaat ctttgcttgc agcagtagga tatttagagt actgtggctt 4560
gacaatatat ttagtatttc aagatttcca tgaattctt ctgatgtatg agttccctag 4620
ttaatcttac atatgtatcc ctttgtaaaa acactttgaa catttaaaat gatacatgaa 4680
tagtactcta atacaatgcc ataaaaatta taaatcattt gtatagactg gtaagtaaag 4740
attgtgagat taagaaacgc atcaaaggcc attgagctgg aaagtggat aatgagaatt 4800
caaaccaggg tctcttgact caaaatctaa ggatcatacc atttctcatg ataatatgag 4860
tattattgtt atctctatcc catagacaaa gtgttaacac tgaatgagca gtgaaatagt 4920

ctcagaattt ttattttat ttagcaattc acttgtcatt tctggctctc agtttattca 4980
cgagtaaaat aaaaatgtg gactagataa tttctatagt acattcttac acaaaaaatc 5040
tatgattttg ttatttttaa tgtgatatac tcatggcact cattcacctc attttcccag 5100
cctgcctcac tggtcattac ttctctgtgt tctttacagg ctccccctcc tctacactgc 5160
cattaaatat tgaaacacct caaagcttta ctatgtcca cctctctct gacactatca 5220
ttctgtctag atgatcccat acatacatgc ccattacttc aacctgtatt tatacgccaa 5280

tgattcacta tatttccagc ctagacattc tttgtactc tagttaccag ctgatatacc	5340
ttacatggct gtttcaaac aactcaaata tattatctct caaaatcaaa ctcatgatgt	5400
ccccacacca tctagcttt ccaccaacaa tacctatccc tattaatagc aataccattt	5460
attcagttat ccaaatacaa aacctagaat tcaccttaa aatttacta tcattccaaa	5520
tatctatcc atcagcagcc actgtattct taatccccctg tatttcttc aaatccattc	5580
acctctctcc atatccattg ctgcatgact atccaagcca tcgcctctac cctagggtac	5640
caaaatagca acaaacctaa tctgttcatt tgcattatit tttctccaaa actgattatc	5700
tatatgtagc aagacagatt gttctcaaat tgcaaatccc actatattat cctcttgctt	5760
caaacacttc catggtttcc cattgtttat gataaaacca aatgcttcaa gttcgaagac	5820
cggcatgatt gggaatttcc tgtcaccta gcctacttgc tctccatggt acagttgcac	5880
tggttttctt tcatttctta agtacaacct gtttctctcc acctcaggac tgtgcatgtg	5940
ccattcattc tgctgaggag cttttttcct tccacttcaa tcagctaagt ctgattcttc	6000
ctgacaatct cagctcaata agcatttctt ctaagaaatg tctctaatat cattaattgg	6060
ctcaggtccc tctactgtat tgctgcactt ttcacagtta taattttact taattatgaa	6120
tgattatttg attaggtcta tttccatcca ttagacataa gcttcatgat ggccagatta	6180
ctgttttcta tccatcgttg tattccaata cctgacagaa ggagggcggg aggtggtggc	6240
acacaagaga tgctcaaaaa caattgttga ataagtaa atgaatgaggcc atttagaaat	6300
aacgaaagta cctgtttaca aagtacatgt atcaaaacta tgaatgcatt ctacttacat	6360
ggttttctcc aaataaaaca aaagacttca atcaggatta atacctggga taaactgagt	6420
cattaatatct ctcttttccc atcaggagt acattgaaac aaatgtctgc aaacaacaaa	6480
tacttttttc caaaatata ttgaatggca tttccataaa caaactagaa catgggagga	6540
gaaagaaagc aatattaatt taaaattaat cttatcacat aacttatacc atcagggtatt	6600
tcgggtaaaa ttcctttcag gcacatccat ttaacaagaa ttgattgtta ctgaaagcct	6660
agaagagaat ttggcacata cttggtgttc aaatatttgt tgactgagtg aataaatgat	6720
gcaagtgtct aagaaacaca aaataaggac atgattacag tcacggtgga gttcacagtc	6780
atctccaaaa tgaggatatg catcccaggg aggaccaaca attcattgga gtgctgaaat	6840
aaaatactca aaggtcattt tacatgtatt ttttctctaa attacttttc ttaagacaca	6900
gaaaacaaaa aaagaaactt agctttgtta ctttctaaca aatagttaaa tcattaacaa	6960
ggattgacac tagcatcctt gtttggctct atgccttagg ggaacatgaa atgtgtgaag	7020

acattctgag atctgagga agggtagaca gtaatacagt gggactgacc aggcttcagc	7080
acacctttac ctctctcag cagatttcag tgatgagcag tttaacta gattgaaaga	7140
ttatattatc tagttctaaa agaaaactaa gcctcccaa agcaacaagg gaactgagag	7200
gaatcctgca aaacaaaaac aaattttaaa acttgacatt tgaataacc ctaatatgta	7260
atcacagtaa tgaacagtaa gataatgaca gaactgacat atttccttat ctattaaagc	7320
catattaaca ggtaaagcaa tgccagtcag tggtagactt cttagaagat atttaataca	7380
tactagacac atacacacac acaacatttt cttcaaggt gtatgtatca gaaaatcact	7440
ttttaaggcc ggatgcagtg gctcaggcct gtaatcccag cactttggga ggccgacgtg	7500
ggcggatcat ctgaggtcag gagttcaaga ccagcctgcc caacatggcg aaaccccatc	7560
tctacaaaaa taaaaaatt agccagggat gatggtggat gctttagtc ccagctactc	7620
aagaggcaga ggcaggagaa tcacttgaac ctgggaggca gaggttcag tgagccaaga	7680
tcacccattg cactccagcc tgggcaacag agtgagactc tgtctcaaaa aaaaaaaat	7740
cactttttag ataaaattca tgctatagag agaagactat gaaaatatgt ttagcaatgt	7800
gtccatcatt aggtgattga gtttcctttt gttttgtttt actgaaaatc atataaagta	7860
tgttatctgt aaaagtcttc tgacatgcac acataaaaat ttgggagaaa agattaacta	7920
taatgtttta tagattttgt acacatttct ttaaaaatat ataaacaca acacctttca	7980
attggtttgc aagaataacc aattgacatc atggaaaaatg gaaattcact tgctgaattt	8040
taacaaaaat ttgcatgatg agtgagactg acaacttagt gtcatgattt aatgaattat	8100
gccaatggta aacttcatgc acatggggcc aggtaatat gtggaaactt tttcaatgct	8160
taaagccaag tattgaaatt aaacttagaa tcagaccttt gaaccatttt atgacaatgt	8220
tcaaaaatta taaattctat ccattatat tataatatta aaaatatcat tacaaaaaaa	8280
acctgtgttt atttataac tcagcctttt taatttctaa tttcataaat atattataat	8340
ggatatgtgt agtaatgtag tattattaca tgtatataat ttataagtaa atatacatgt	8400
tttggtact catgcataaa atgtttcacc cataggagca cataatcaga aatgtctgga	8460
gaccattata gtaatagata gatcatattg ccacatatat tatctctcc ttgacaactg	8520
agctttccag atcttctggt gaaacgaaag agaaagtgt aacagaagag tgattaaaat	8580
gacaaaagca ttacttctat tacttctatt ctaataatat gagcaaagct ataactatca	8640
agtaataatg cactaaagaa ggtgattaat ctgatatatt cacaggcaac taataagacc	8700
tttctattgc agccatgaaa aatatgtgac aattatagat atcctgtgtg cagtgtttca	8760
acctttatgt gacctgttct actaacagat ttagtgatgt tcactttgtt agaattttct	8820
tacacatgcc ataacttgct tcagtccttt gattatgaat attatggata ttaaggattc	8880

tagactattc tagatttaaa aaataatatt gtcacctcaa tcagaaggga aatattaaat	8940
agttctcatt ttttcaatgt ttactcagtt ttgtccaat gtaatgaaag tgtcagcagt	9000
acaggttaca aaataaaatg tgtattaaag taaactcatt tgaacaggtt aataattgta	9060
gaggaggga aaaggctaaa agattgaatg taaaacttat gaaaagtaga tacatcgtct	9120
ctatgatttg cagtagtcaa ctgcatacag atgaatcatt ttaatacacg ttaactactt	9180
tccttttaca gatggagaaa ctgagaggaa gaaagtttat atggttcatt aaactttgtg	9240
atgcaagcta aactaacctg tctctgtatt ttccatctac tgcccttata actatctcat	9300
tagaatactc ttcaagcatc tccttactga ttttcttacc aagcatttgt taagtcttaa	9360
tgagagttgg tagtaacatt ttacccact ctgtgaaata tgaaatctta ttcataggcc	9420
tcttctttta ttcttgtatt tgcatatcaa ccaattaatc aacttgcttt ctttatgttg	9480
cttattatct tagtcttac taaattgcct cttaatgttg tccacataac agaaatgtta	9540
agggtgatac ttaacatfff agtccagtct agccgggtgcc agtgcaatgc caaatcatga	9600
attaaaatat aattacaaga accacttatc aaattttaac aattccttca gctttgtgac	9660
agttttttct acttcgatta aagtcaagta aaattaaagt taaatatttt tattaaaata	9720
tctcctttaa cattccatat taataaacat attaaagctc atgcttctaa gtagattact	9780
agaagtact ttatcgaatt acagcaatgg ttaattctag atcatagaat ttagaatgac	9840
tttttgctt cttctttttt ttctttttt ttaaacagag tcttgctctg ttgtccaggc	9900
tggagtgtac tggcgcgac ttgactcact gcgacctctg ccttcaggt tcaagtgatt	9960
ctctgcccc agcctcttaa gtagttggga ttacaggtgc ctgccaccac acctggctaa	10020
ttttttttt gtatttttag gagagacagg gtttcacat gttggccaga ctggctcga	10080
actctgacc tcaagtgate cacttgctc agcctcccaa agtgcctggga ttacaggtgt	10140
gagccactgt gcttgccctg actttttgt tttctcttaa tacttactag tatttcttga	10200
atttttaaaa aagaacata aagtactttg ataaaacca cagtctcatt gttcttaaaa	10260
ttgttcaaag gttctctgga aaaaaaaaag aaaattatca ttgtgttaag aatcatgttg	10320
gtctgacac aatcatccta taggagtga tattgaaaa gtaagatata ttgtgtata	10380
atcgagattg cataaatttt accatttttg agaagaatct gtcctaaatc ctggcttaat	10440
gtaatatcca gcatctact taattttctt gtcttcacct ttcatatcc acatccacct	10500
agggtgccacc tcacagtata agccagcata atccattctt ctcaatgaaa ccacaatata	10560
tctgaccctg catctcagga gaactgtatc agccacagca cttccagttg actatgaatc	10620
tgaatgttat gctcaggag aaacatcctt gctgggactg agtagtgatt caaggagata	10680
gttatgattc agtcaagaaa ttaataatta gtgttatttt tattattgag acagagtctc	10740

gttctgtagc ccaggctgga gtacagtggc atgatctcgg ctactgcaa cctctacctc	10800
cccggttcaa gtgattctcc tgcctcagcc tcccaaataa ctgggacagc aggcacttgc	10860
caccacgcct agctaatttt ttgtattttt agtagagacg gagtttcacc gtgttagcca	10920
ggatggcttc gatctcctga cctcaaggtc cacctgcctc agcctcccaa agtgctggga	10980
ttacaggcgt gagccactgc gcccggccat aaattattaa ctgagccagg cacagtggta	11040
cacacttata gtcccagata ctgaggagac tgaggttgga gtatcctttt ttatgttatt	11100
ttatttttaa ttattatggg tacataatag gtgtacatac ccatggagta caagtcatgt	11160
tctgatacag acacataatg ttttaataatc acatcagggt aattgggata tccatcacct	11220
caagcattta tctttctttg tgttaggaac attccacctc cactcttgga ataggcaccc	11280
tgttgtgcta ttaaatacga ggtcttattc atttcacta actatatttt tctaccatt	11340
aaccatcacc tcttttcccc tcttccccac tacctttcct gtgaggctgc aggattctta	11400
agcacaacag ttagaggcca gcctggacaa catagtgaga ctcaatttct aaaaaataaa	11460
aaagaaatta ccaactaatg ctaaaaaat agtctctgat gcttaggtat gaattagaaa	11520
tgaccaaaaa aaaaaaaaaa aaaaagactg ccttttgctt ccttctcccc ttctcttcaa	11580
gttttccatt gctactcatt ttagtctggi ttaatcagggt ttcattccatt aaaagcaatt	11640
gttgggatca cacattttga gtttgtcag tggacttccc tcatgctggc atgattcctg	11700
ccccaagccc ttagtaaaag ccaccaagcc atataacata atctctcatt gagtaaaaca	11760
tctgatgtgt ttagaatgac ttctagcaaa aaaccagcct gtccagcatc atctctgtat	11820
aacagataaa ggaataggta ctgcatcaaa aggttataga acctgcccac atcaatccca	11880
tgtgttttgc aatggaatta ggttgaacta aagtgaatc tcagttttct actcctcatt	11940
aacatgtctc atgttgcaag gttgagagga aggagaagaa gaactgtatt tacagagaga	12000
ttccccctct ctttctttct acagattact aaaacattca aagaatcaaa ttttaagaaat	12060
cagttcatca gagctcatgt tgccaaactg aggtgagtgg aactgtagaa aaaatattta	12120
agtatagata caatgtggca tacttgactt ttgtcacag aatgaatagt aatgacatg	12180
ttcagataag ttgttgtaat attatgaaaa tagtatttta gtcagcttaa aaaccaatgc	12240
caaaaaagcc aaacatafga tctatttagc tactaatgta aataaccata ttatatctat	12300
tcttatitggg aagaggaaga aggggtggag agagagtgg ggtgaaggta cagtaacaag	12360
gccatcctat tgiaaaactc cagtggatat cattcacagt gcagcctatg taaacagtcc	12420
ctcctggagt tgtacaatgc tgtggttttg gtgtatccat ccaagatcaa gacactatga	12480

ccaacatcaa aagtggcttt ttggttttat ctgcctgatg tgctataata aaagggtatt	12540
atggccaaat ccaaggcatg tctatcatga attaataata ggaggagtag cagcatgcat	12600
gctagtattt tgccattcct gccttagtta aatatgatgt gataaaacca gcctttccaa	12660
ctgaaatagi cacccttact gactctcccg caaatgtctc aaatgaccac attgctctag	12720
tctttaaata atatgcaata gttcttttgg agaagaggaa ttatactaata tctttctcaa	12780
atactagcat cacaagaaaa ttaattcttg ttctctggag agtcacctag taagtatctg	12840
gagcacagat gtctggtcag gtaagttttg atgaggagtt aaagggataa gaagagtcca	12900
tgagaagggg attttccaaa acacctttcg gtcaattcag tgcacattca cttagtactt	12960
tcttgtcagt atctgtatca gccactaatg ttcaaaagtg agtaagccct gaaaacctgt	13020
aggactacat gagccttctg ccttttctct ccttttgttc acttcccact tatcactcaa	13080
tcctctgcaa cctggcttca ataccacat aaaatatcaa ctgctcttgc cgattcaaca	13140
atgacatcca gataacaaaa tccaaagaaa ccacatcagt cctattcttg gacctttcaa	13200
cagtatttgg tctgtttggc ctgtcactcc ttgaaatagg actatccctt ggtttgcattg	13260
gccttgtata cctgattttt ccccttacct ccctagctat tcttcttag ttctctttac	13320
taggtcttac ttctttglat attccttaaa tgttgcgtga catcaggctg tgctctaggc	13380
ctctcatctt ctccaggtcac actctctect ttcccttgcc ttcactgcca cccatatgct	13440
gagtgtcttc aaagtgtgat ctctaggcca gtccctcttt gcctccaaac atgaatatat	13500
gcagccatct acttgggtacc atcacatgga taattctcat gatctcttcc agtatgactg	13560
cttctttatt ttttctggg ctctttttta gcattgcttt acatggaact ttatcatgtc	13620
tctcaacctc tattttatct ttatctatg tatgtagagt ctgtgtaatt tcttcatctc	13680
ttttagataa ctaatatctc ttcagctttg acttgtattc tgtgtaacct atttattgcg	13740
ttttcaattt caatgagtat gttttcctat ctgcaagttc tatttgtttc ttttgagaat	13800
cttcttggtc ttttaaacac atttcttatt ttaatttttg ggggtacctt gtagttgtat	13860
gtatttttgg agtacatgag atgttttgat acaagcaaac aatgcataat aatcacattg	13920
tgtaaaatgg ggtatccatc ccctcaagca ttatctctt gtgttacaaa caatccaatt	13980
atattctttt agttattttt aaatgtacaa ttaaattatt attgaccata gtgactctgt	14040
tgtgctatca gatactaggt gatcttttaa aaataatgtt ttctacttaa tctcattttt	14100
atgattccct cttttacgtc atttgtcatt tcaaatacag tcaattgtct gttgattcta	14160
ttatgtgaag tttttgagga taatcttttt gttactttga ttccaccttg gtatggtttg	14220
gctgtgcccc cactaaaatc tcactttgaa ctctggttcc cataataccc acatgttgtg	14280
ggagggacct tgtgggaggt gattagatta tagggacgtt tcccccttt gctctgttct	14340

ttttctgcc accatgtaag aaagatgtgt ttgcttcccc ttctgccatg attgtaaatt	14400
tcctgaggcc tccgcagcca tgcaggacct cttttctttg taaattaccc agtctccggc	14460
ggttctttat agctccgtga gaaaaaacta atacacacct catgatgtat tgtttaccac	14520
tgaaattgta tgcttaaatt taatctcact tgggaccttg tacaacctag acttaacata	14580
tctacctcca gagcagttac atctgtcaga cattctagag gaatcagcag cacatggact	14640
ttgtttgtgt taattttgtg tcgggggagg ggggagggat agcattagga gatacaccta	14700
atgctaaatg acgagttaat ggggtgcagca caccaacatg gcacatgtat acatatgtaa	14760
caaacctgca cgttgtgcac atatacccta aaacttaaag tataataata ataaaattaa	14820
aaaaaaaaag gtcttgggag tattcaggta gtattaatga agattcagac atcgtgcagc	14880
caggcccatg cttatgaatt ttcaggatgat acttcttttt cttttttctt aatttaaagc	14940
tggatctcgg aaacagataa atttatTTTT ttatgacatg acgagcattt ttttcattct	15000
agttcatgct gttattgggt gtttagttct ttgagactcc tggccttttt ctaaaacctc	15060
aagtccaact tcctatTTTg cactggccca aggtcccatc tccagtctct atgtaaatgc	15120
taaacataag cctgtggaat attctagtct caccacatac tattcacatt cttctttgtt	15180
tttggcttc caggattttc cttacttttc tatgaacca gtcttgcatg tgaatggaa	15240
tttattatat attatctatc ctttctatTT gttttatgca gaaagtgttt tctaaaatta	15300
tttaggcttc catattgcta gacatggaag ttgtaattat ttgttcagtg cctgtttcta	15360
catctaaact gcaagacca tatggcaact gtgaatctta gtcccagcta atttctgaag	15420
cttagaatag tgcctagcac aagaagttgt ttatctaaca tttttaaaaa taaatattaa	15480
attcatatct ggaatgaata ttaagttaga gctggtcatt gaggtgagag gaggaagcca	15540
agagagaata tgagagcctc aaagccaaat atctttaatg tactttttca gaaaagaaga	15600
cagccaatgt cagggtggagg aactggttta tgaggtaact ttcttggag aaaatagaaa	15660
ttactgaggt ttagataaat ccaaataTTT aatcaagtca ccaaggTTta ttgtggggaa	15720
tctttattat taattaaaaat gagtgatgaa atcttaatat acgacaaaag ttaaaatttg	15780
cttttgcagg cagatgaatg gtctaggtat caaaaaatta agttgagtct ctaactcaca	15840
caaatttaca accctatcac tttatgaatt tgtttaggag attattttta ataacttg	15900
tgaagtctaa gaatagctaa aatttatagt acacttattg tgtgctattg actcttcttt	15960
gaagttttgc atatagtgat tcatctaac ttcataaccc attttacatg tgaagaaact	16020
tagatataga aagattaaga aacttacata acttatccaa agttacacag taaaactctg	16080
gcattataac ttcaaaatca gctatcctac agtgagtaca gtgttctgtg cattgaaatc	16140
aaataagtga gatagcatcg tgatatagta ttacgtatgc aaactgtt acagagatct	16200

gtctaaagtt aaattccaca aatgaattct ttaaaagggt ttaatcaaga agaatatata	16260
aacaggatgg tgaaaaattg tcatattatt tgttttttaa aatatcttta tgatttacag	16320
gcaagatggg agtgggtgta gagcggatgt tgtcatgaaa tttcaattca ctagaaataa	16380
caatggagca tcaatgaaaa gcagaattga gtctgtttta cgacaaatgc tgaataactc	16440
tggaaacctg gaaataaacc cttcaactga gataacatgt aagtataatt tttcataaac	16500
aattttattt caatataacc ctcaagttta ccaattcaaa ttcataattt aattgagagg	16560
ctgacttttc tttctttgaa actaaactgt gaaaacaatc cattaaaaag ctaaataatac	16620
catatagctc cctaacgtaa atcattctaa gacttaaaga atcatttggc atttatatag	16680
taaattttat ttgctaaaaa ttctcattaa ttatccctgc aacattcctt atgagtgatg	16740
ttactgtcag atgtcattag tggataggcc ataggagggg tacatagatg ctcaaggcca	16800
gagaactatt taattaatga tccacctcag aggcttcttc atttttcttt gtaacattta	16860
tcacaattga aattacaaag ttatctgtgt aaattttgta ttgtttggct tcatectaca	16920
ctgtaatcat cctaaaagaa agaaccagtc aaccttcttc atcctactac cctcctacca	16980
cccagctctc atcatataac acatatitcaa taaataattc ttgcatgact gaaagaaaag	17040
aaataatata tgcatagaat ttaaggacat tcctccaagt tggttacatt ctgctagttt	17100
aataagccat tatttcttct cgatgagctc aagattaaaa ggattttgat gattcccata	17160
ctagactggg aggtaccagt tacagatgia ctaactgtta aatattgaaa tgctttccta	17220
tttgtttgta aacaattact gcatcaggcc cacaaagttg tcttccgaga tgtttcaaat	17280
ccactgcccc tgetgctaaa gagttatgct tagcaaagca aagcactcta agacactgct	17340
ccaactccat ggctgattg catcttttat gactggccaa tgctcagca ctgcagtttg	17400
ttaggtagtt gaatattacc tctgtttcca cacattaagg aatgctccg aacgcacttc	17460
ccaagtgttt atttatattt cattatacta gacaatatgg tgatagatg gtcacagaat	17520
agcggtttcc acctccagag cccataatct agttgaaggg aaagatatc caacacaaga	17580
gtgttgacaa tcaagataga atatgatcaa gggcccagtg tgaggcccag gcaatgatca	17640
ctgcaggaat ctggggaaga aagagaccag cgtgcttggg atatctagca aaagtttcat	17700
gaaggagaat ggactttgac tttgaaatat gggtaggatt tacatatttt gagatgagaa	17760
aaagaaagtt cccagagaag gaaagcatga aaaggcaaac agtctgtact gaacgcgatg	17820
ctttgacaga ataataga aagggacctg ctggaatgat tgatcagtg tcatcattca	17880
caccatcatc atcaaaacac ttatttaatg agaacttact gtttttagg catggcttta	17940

atgcctata tgaatTTTT tcttgattaa tccttacaac aaacatatcc catagatagt	18000
tttattgtcc cccttagaaa agataaattg cctaggctga cacagtcagt atatgaggca	18060
gtcaggattc aaactaagtc tgtttgttca aaaaattaag aatggccagc tttttaaaat	18120
tttctgtctc cagaagtatg atttggtccc actgaagttt gcaaaacaaa tgtgataccc	18180
aaacctgtgt aaacttttag tgggaaataa ctttgcataa gtcggtttga gagagcgtgg	18240
aaacctgtct tgaagagttt taatttaact tgcaggaaat aaaaatgatg ggtttctcaa	18300
ttaaaaattt caatcaagga aggatatgag ctaacataac atttttttaa aaagatcagt	18360
ctggtaaggt agaggtgcat aaactgaaaa ggagcaaaag tgggtgaatt cagttagaaa	18420
attattgtaa ctgtactgat gtcaaatgat gaaacatga actaaagtag taccaaaagg	18480
agtgaggagg atggaataat tcaaaagata gaggacagat gtgcagaacc tggagattat	18540
aagatgtgaa aggaggagtt tgagaaaatt tcagattttg gaagtgggtg cattttacta	18600
aaaggatata ataagtagca aattttggat aaagttgggt cccactgagt ttgagatggc	18660
tgttggacat gcagagaaaa ctgtcttgta tgctgttctt aaattgaaat agacagacct	18720
ttacctctg atactgacat attttcttt ccaggctcac cctccatttc cctaaacaca	18780
acacatgcac tagctctcct tactttattg ctccacaaac atcttacacc tccaagcatt	18840
tgtgcccact gtaccttcta tctggaatct cttttgtcct cttgtgtgcc tgaaaaattc	18900
ctttcagatc ttcaaaatac agtgcagatg ctattttctt tagctcaaat attatctcct	18960
ccatataatt taattactct cttttttctt ttctctactt tgcacttaca tttatttgaa	19020
tgattgcttg attaatttct acctgtaaat tatgtgaggg caggctctct atattttgct	19080
cgcagttaaa tctgcagcac ttattataga tgggtatcat tagagtaata tacatatatt	19140
tgaggacatg ataaattaac ttcccctata gtatttatca cattgcatct caatgacttg	19200
cttatgtttc tgttttccca tataaattga gtaacttgaa aaaagagata tctattaagt	19260
atttaatgag aaattaaagt acaaaactta gtatgcataa caacaaattg ggaaaagggt	19320
gtaaacaaag agatttttag ggcccatgag ttagagatcg tttcagcagg tctgaaagga	19380
agcctaggaa tctgcatttt agaggaccac ctcccaacce caacaagtaa ttctgcttct	19440
tgttgtctgg gtactgtact ttaagaaatt atggtgaaat gatatcagcc tttattgtat	19500
ttatcttatt ctcatTTTT aatactagca ctactgacc aggtctgcagc aaattggctt	19560
attaatggta agttttaata ttattttgta actgtaattt gccaaatcat aaagagtaaa	19620
agtgaagtc ttttgtgtac ttttggccaa ggcagtatct atcaagttga tgtctttgtt	19680
cttagttcgc tcaggtggtg ttgaaacaag acagtgtctga tccaagtggt cccatggagt	19740
ggactttagg ttccccctt ctttttagaa aaaggaagaa gttgtagtgg aggactaccc	19800

actctgcact caaaattgcc ctcatgaaaa tttctttggc agctttgaga accttttact	19860
gccctggttc taagggtggca tttctgtaga cttacaaatt atgtttgatg acaccgttta	19920
tgtagcttct cctaaccacc agagtagctt gctttgttgt gaattcaggt taatcacaaa	19980
gtataataaa aaagaattgt cagaagtctt cccagctttg ggtctataac ctgaaggaaa	20040
agtcactact cttcaacatc atcctatgta ctctcaggct aggatagcag aaatgcaatc	20100
cctagaaaa acgaacttac ttctctgacc aaaaaaatgc agttaaaaat tagttcaatg	20160
tacctggtag ctggcctatc ttaggtactt cagtgat ttt acaaagtgat ggtagtccta	20220
tgggtgtttt tcagcttcac tacgtattta attcatgctt attgttaatg aaactgtgat	20280
aagcaattta ctagggtatt tgtttgggag atgccacaaa ggaacacatg tatctcttaa	20340
tggaagcctg gtctctcttt atccaggaaa tttgctagga aaaaaagcc tttaggtggt	20400
tgtgctatta aaccagggca ctacttaaaa gccagcccag caatagtgtg gtgatttacc	20460
attaatttct tagtaataga ccacacaaaa gaagaaaatt atgggaatgc gagttgagag	20520
gaattgggtg atcagcctac cccagcccg ttcagctctg gccagtagac tattcacgag	20580
ctctttgaaa acatttaaat aaaccttatt tagatactag aaacctctg tcacctcaa	20640
gaatatcttg tgglatagcg actcctttat gagggcatgt ttggtlaatac agcatcagtc	20700
ttggaggtgg actggattct acaagggtgaa ctgcagtcac taaggagtct tttggatgag	20760
accagttttc ctccaacttc aatgtgtgca tgaacctcac atcaaatgt agcttttagat	20820
ttgtcccatg atgtggttcc aagaatcagc acttctaata agtttccagg ggatgccc	20880
gctgcaggcc cacaaccac actgagcata gcaagactat tgagaaaaag gaaatttccc	20940
aggagtctgt ggctgagct ggcacatcca ataatacct atcttaacct caactcatga	21000
ggaattccag ggaactctga agctgctcaa aatttgaagc ctatatgcca actaaattca	21060
gaaatgttct ccaaatgct atctataagc aacagtagtc acaaatgcat ttagaaaata	21120
tatcgatcat gctttttgga aaatccagca tgcctgagg aagaatgtat aagacataaa	21180
agtcataaat tatggaaaga ctcttcagct tcttccaaat gtaaaggaat catgatcttc	21240
ccagcacatt aatgcccttt ctcatagaa tgtggggccg gtcagacct aataacattg	21300
tctgagcaga gaatccttgg aggactgag gctgaggagg gaagctggcc gtggcaagtc	21360
agtctcgccg tcaataatgc ccaccactgt ggaggcagcc tgatcaataa catgtggatc	21420
ctgacagcag ctactgctt cagaagggtga ggccaccact acctacccat ctgggaacaa	21480
ttagaataga caggatcatga agactgcacc ctctacccta ggattgaatt gagccagaaa	21540
taattcaatg caaaaaatc agtaagaatt ttttctctat tcatgaaagg aaaaggattt	21600
ttccccttta gcatgctaatt ttagtgctat ttctctgttt caggtaataa tatattagca	21660

cagtaaagaa caaagattta tatgtcagaa tgttttttaa atcctagcta taaaagctta	21720
agaaattttac taaatctcca taagctttat tttttttcca aattaaggga caacactgtt	21780
atctgtgact tagtgttact ggtagcattg agtacactaa tgtaaacata cgttaaatgt	21840
tagcgaaacg aattgctgtg gaagatttgc acattatatc atgggagctg atggctaacc	21900
tagagactgc cccatgccat taattttatc attcataaag attattgagt atctagtatg	21960
agcacagtgt tatatatgtt agaagctact agtataaaca aagtattgcc tctgccttca	22020
aagagcttac actcgaatgt tggaatcaga atgcacaaaa ataatgatca attacaatga	22080
gtagcataaa taaaattaat gtaggcaact tacaagaatt ctttaattgag gtgactaaac	22140
tattgccaac actagggtga tatgtacca gtggcgagta ggttgcataa acttacctta	22200
ttggtaaaaa gaaaagtcca cattgtcat aaaagaagga ttttagattt cagcataact	22260
aaaatctgtt tcaaacctgc ctgtttactg gggcatcgca gaccacaaca gttgttgga	22320
acttaactca aaaagttcac ccagaaaaat aatggagatt tgaactcgtg tgcccctgac	22380
catatcaatt ttcttctcag actcttactc taaactggac ctccttatca cacacacaaa	22440
gccttcata ggcagatcaa tccagtctta tttctcaaag catgtacctt gagcttcaga	22500
taaacagcat tgttctcttc ccctggactc ttctacatt tccctaccta tgagtatctg	22560
atcaatctgc ttatccttga aatgttaata tatttaccac atctctatct gaattttatg	22620
aaatttttga taattttctaa gtagtttttt cagatttata ggcactactt catggtacag	22680
tgactgttac aaacgtatct gttaaattta gaaggaataa agatttaaaa gactagggtta	22740
gttactgaac taaagtttta ggaaatccca aattatttca aatttttctt atggtaattt	22800
tatgacttaa tatttttata tgcagtgaac aaatttgaaa cttttaaaga tactcccaga	22860
attatcagtt ttctgatgta gattggcaaa ttattacta tatcccaaat aaccaagag	22920
acaaaattca caaaaacatt tcaattttca ttgccacttg aaaggccaaa aagcagaaat	22980
ggcacgcatt gatttcaatc gtactcttga gtgtgggaac caggaattaa aatacctgga	23040
cttatcagcg acttagcata accaagaacg gaatagaaac ctccttgat tctaagccct	23100
attcagtccc aatcaccaaa aaccaagtaa acgatatcac tataatgaaa gccacagtta	23160
taaatatcga caacgattac caaaggaatc catggaactt tgaattttgc caccacat	23220
ccttctatc attacatga ttgatccact aaagctaaca gactctgtga acctgtatt	23280
ggacccctcc ctaaagacct gattgtcact gagaaccatc agtgaggatt tgtttggggc	23340
atgaccagcc ttacatcaaa gtacatagaa gtgatgaggt cttatcaaag aggattattg	23400

aattatcacc tcttctatgt agctttccct gatactctct ttcctctcca ttgagttcca	23460
cagaaatfff tttatctgcc tttaacagtt gtccctcatga tttgtgatat ttgacttacc	23520
tcttgtcagt ttccttccact agtgtagagt tccctcaaaga aagagacat aattacttat	23580
atftttatct ctggagactc atactattcc ttatacaaag tagacactta acaatggctt	23640
gttgaactat aattaatgaa aataatagct accttcatga aagttcactt tgtgccaaac	23700
actatagttg acataatata tttgtctcat taatacttaa caattgtgtg agaaggtatc	23760
accaatcaca ttttatatgt aaataaaccc cagagctatt aattaaacttg tcataaataa	23820
cacttttcat atgtggcata gccaaagattt aaatataaat gttactggtt ccaaaatgat	23880
gtcttaattc acttgcctga aagaaggaaa ggaagaaaat aaacgagtgg aaggaagaga	23940
gggagggag agagaaaagg aaggaaagaa aaaagagtct cttcagaacc ttcactgtaa	24000
agactccgag caaaagaagt tgaatataaa aacaacatag gtttgtttgt tttctaata	24060
tttttcttca aaatftttta ctcaggttca ctcttacaca aactactgtg tcttataaaa	24120
gtatttccgg tcatagaatt tttattttct gtattaaact cactatctaa tctccataaa	24180
actcctaaat tggatattatc ggtaacattt tgtttttact caacccttag gaacaatgtt	24240
aagttaatca gccctccaca tcacagatcc ttattttcat cagtctgtac aaggcatttc	24300
tctcatftta atfttttttc ctctgtcat ccttggttt cactttcact gccctccttc	24360
cacccatag ctcatacta atatatctga aatatacatg tcttaaaggt acatgcacgc	24420
acctacaaaa cctatagtgt tttttgtat gtatatgtct ttaatttaaa taagtagcat	24480
tgtgtaaaag tctaataatg tttcttactg ttttactca attcttgaa ttttcatctg	24540
atgcactgct gcatagcacc ccatggatg cagccacat atttcttca tccaattagg	24600
ttgatgacc taccttccca ttgccacaaa gactacacac aaaatattg tacttatctt	24660
tctgtaaacc ttcaggaatt tcagaagcac acatgcaggc tgctaaatat accagaatac	24720
tttccagcca cttaaatctt taccagtatt gcaaaagagg ccccatfttc ctccacatca	24780
acatttagta ttattctttt gtttaagttt tatcaatctt ttaaatgtac acaagatgct	24840
catttttata atfttaattt ctcagattac tagtttgagt atcttttcat atatctaaga	24900
gctgttttga tctcccctac catgaactgc cactaatatt ctttgctat tttaaatgg	24960
tttttctgct tatttattac tggtttacag actfttaaaa tatattctac aaaaatttta	25020
gacattaaac attaccaata ttttccatg gttctctatc catctggtaa acttgtctat	25080
ggtatatcta atfttgattt aatagaattc attctattt taccttttag tttgtgtttt	25140
tgttgtttag ccaaaaagtc cccatttcta ggtcataaag gtaatgtcct ttttttttt	25200
ttaacgtac tgttctctct ctgtctcccc ctatgtatat aggtgcacat atactgtac	25260

acacatacat atacctatat atgaggggag ttcgataagt ttatggaaaa taaaattaaa	25320
agataaaata aaaaattata aactttatit ctcaacataa gtcctttcaa gttcaagaca	25380
cttttgtaag caataatacc agccatatcg tccatcccta aagaactgag ggtcctgaga	25440
atttaactat gtcaatgcag tcttttttac attacttttt tacagtactt attgatgaaa	25500
aatgggtgcc ttttaaagat tgttttaaga ttagggaaca aaaataagtc agaggaagtc	25560
aaatcaggac tgaaaggagg atgcctagt atttattgct gaaactttca taaaactaac	25620
cttatttgat gagaggaatg agcatgagca tggttgtgat ggagaagaac tctggaggag	25680
ctttcctgga cactttttct actaaagctt tggctaactt tcttactctc ataagaagaa	25740
gatgttattt ttcactgacc ctttagaagg tcaacaagca aaatgccttc agcatccaa	25800
atgtctgttg tcatgacttt tgttcttgac tagtctgggt ttgctttgac tggaccactt	25860
ctacctcttt atagccattg ctttgatggt gctttgtctt caagattgta ttagtaaaagc	25920
catatttcat ctctgtttac aattcttcaa agaaatactt cagaatcttg atctgacatg	25980
tttaaaattt ctattggaag ctctgacctt ggggtgcagct gatctgggag aaacagtttt	26040
ggcatccatc aagtagaaaag tttgtctaac tttagttttt cagtcagaat tgtataagct	26100
gaaccagtgg agatgtctat ggtgttgtct attgtttctc acagtttaatt gttggctctc	26160
tttgagacat gaacaagatg aaatttttcc tagcaaaactg atgtggatga tctgttgctg	26220
cgggcttcac cctcaacaac atctctttct tctttgaaac aaattatcca ttagtaaaact	26280
gatgattggg ggagatgctg tccccataaa ctttttgtaa ggcataaata atttcacat	26340
tcttccagtt tcaccataaa tttgacgttt ttttgcttca attttagcag cattcatgtt	26400
gctttgataa gagctctttt caaatctatg tcttattcct cttagtgcct caaactagat	26460
cttgttcagt atgacaagtt agtatgagtt tatctgcatg caaaaatctt tgaaatccat	26520
gcatagtttg tttataatat acattttcaa tgaacttttg aagaccccat acatacatat	26580
gtatatatat gcacacacac acacacacac acacaaaaat cttcaacat tatcagactt	26640
agtcagaaaa aattattcat ccattaacaa gataagaatg ccccttatca tcaactactat	26700
ttaaatggag ctcttggtta aaggaaaaga cagggttgta aaaaaattag ttaaatctaa	26760
aatgtttatt atttcaggtt tcttagttgc ttaaatggga agggaggtat ggacaaaaga	26820
gaaatcaaag atatttgtgt tatgtactt atcattaaag tatcagaata acttcattgg	26880
aatagaaaaa caccaagatc accccacgat atgttttcta aaatcttctc catttcttta	26940
gacaagtgac catgtattcg gccagtgaag aattaaactc acttgccagc ttataatgca	27000
ggaaaatata gcaaagagat gtggatccaa tagtttctag atagtgttac aggatggcta	27060
agatgaattt atatatctga aatgttcaca aattccctac tcatatagca tgttttcata	27120

atgttttagc aactctaata ctcgtgactg gattgccacg tctggatatt ccacaacatt	27180
tcctaaacta agaatgagag taagaataat ttttaattcat aacaattata aatctgcaac	27240
tcatgaaaat gacattgcac ttgtgagact tgagaacagt gtcaccttta ccaaagatat	27300
ccatagtgtg tgcctcccag ctgctaccca gaataatcca cctggctcta ctgcttatgt	27360
aacaggatgg ggcgctcaag aatatgctgg taagtgtctc ggaaaaaaaa attaacaata	27420
gaaatgtctt atatttgcta ttaggtaatt ttttaatta ggaaacatct ggaataggtg	27480
tttctattct tctacagaca gaaccattct atattctgct cagcccaagc tctggctacc	27540
cctgagtctc cttagcaaag caaagcaatg ctccagaaac tatgggaatt ctcaaatata	27600
gtaataggaa aatgtaaaag aaagtatga agacacgagt tctttaataa tccagagatt	27660
ctataagatt caaatagctt ccctataaac aataaaaaag attttgtttg ttgtttgtt	27720
tgcttgtttt ttagagacaa agactttctc agactggagt gcagtgggtc aatcatggct	27780
tactgcagcc tcaaactctg gtcttaagaa atcctcttgc ttcagcctcc caagtagcta	27840
gaattataaa taagtgtgta ccaccatacc cagctttttt ttttttttc tacagacagg	27900
ttcttgctct gttgcccagg ctggcttgga attcctgccc tcaagccatc ctctgcctt	27960
gttggcctcc caagcaatg ggaggattta gattagacat tgtatgaggg ctttaataatc	28020
cttaaggtat taactgcctt ttaaagtatt ctgggatatg gcaaaaactc gatgtgtata	28080
taaacattgg tcatatttgt ttattgaatg aataaaatgg aaactaaat gaggacaatg	28140
cacaagact actagaacca gtaagagtat cagcgaagga gtggaagggt agcattgaca	28200
atttccctgg gcttttacc atgttgtaga ttgtctctcc aaggaataat acaaagcctt	28260
aatagtccta gaacacattc tatttgttgc ttatggccca aagtaaatg gtgtagtaga	28320
taacatttgc accagtcagt aaaaactatt ggtgtcattc tgagagtaca tcaatataaa	28380
atagactagt tctttagcct tgaaactaga ctggtttctc ttttgctgct aggttaaagg	28440
ttattcaata tgtaatcttc caatccaaaa tctgtcagtg gataatttaa aagcttttag	28500
tcaattttta gatatttgtt ttcttaaaat ttttaagggc actgtgtcac aaagctaaag	28560
aaaaaaaaaga aaaaaaaact gatctgtgaa aggggttata ctcactact tggggaattt	28620
tggctgcgaa gaaactcaa agtaaatctt tagaagcctt cattgttaaa tatgaaataa	28680
tgtttgaggt acatttatit ctctcaaat ttattatagg gtcaataatg tacacatctt	28740
gaagtccatt ttttctgc ttttataaca aacaggccac acagttccag agctaaggca	28800
aggacaggtc agaataataa gtaatgatgt atgtaatgca ccacatagtt ataatggagc	28860

catcttgtct ggaatgctgt gtgctggagt acctcaaggt ggagtggacg catgtcaggt	28920
aagctcaaga caatctcatc catgtcatca tccaagaagt gtataagcac ttcctagtat	28980
gtgataatgt gatagacata agtftaacag ttacaataca cagccctgtt cctctaaaat	29040
ttataatcta gattttagaa ataaattttt ttatgaatga agtttatcta tcatgaaagc	29100
attaactctg agaggccaaa ttacagagta gttaacctac caaagctcaa gaatcagaaa	29160
gacctcgatt tgaattcctt aacctctatt accaagtctc taactaaaag ctggggataa	29220
tcataatagc acctaacctt ttgggtacta agaaaagtta aatgaagact aaatatatca	29280
ggcacatggt aaacaacaaa gaaatctcat ctatttcact attattaatg tagaccatgg	29340
tcactcgtgt taataacttt aacctcaacc ttttaactgc tgtgaaggat taaataaaaa	29400
attaatcact atattataaa aattaattga tatataataa atgaatttta agagatagct	29460
aataattcat ggactccttg aagatagaaa atttatacaa aatcctagta atttgagtca	29520
caaaagctcc tacaataatg aaacagtatg aatgaaaaag aaaagaaata actattatat	29580
ttggatctag ccataatth ttaaccaaat gcacaaaaac aaacaacaaa tatgaaattc	29640
tcactgtaaa gtgattaaaa tcaaatthga attctaaaat tttaaattaa attatctaaa	29700
cataattgat gcagttatat gttttaatag gttttgttca catactgaa atccaactcc	29760
acacagtagc aggaacagct ggtgtcagaa attaaatatt cttttagtct ggagttttaa	29820
aaaatcaatc tgtttacttg agtaatttgt tgcgtttttc atgggtgaat tgtatacaga	29880
aggataagaa ttattcttcg catcaaaagg tcactgactt tcatatttag tgctcatggt	29940
ctttaaaaag tggataaaaa gtagttctca catthcatgg aaagcccca atccatgagc	30000
acatttccca aaatgaaaca tttttatcaa ctgcaagtth tgtgtaggth gagatttgtt	30060
thtcaattgt caagatactg ttaattacc agtcttht atctctthtgg tggagatgth	30120
tcgtgctag gaaaccttc ttgctctct tctgtttct cttttactac tggccctgaa	30180
acaacaaatt ctcaagthc atgacagctt tcaaagaat ccatcaatca aataagcaac	30240
acaactgac actgacaatt ccagacctac taagagcatt aattaagact taaaaataaa	30300
catgagthtt aaaagggtgt tattcattat thtccattt ataacgtccc thaccttctg	30360
tccttcagth catacaaatt attatcttc ttgaagcca gttcaagccg tacctacca	30420
tgatacttc catgtatatt ccactctagg cctcactgat ttttaactga aatactataa	30480
tgcatagth acacttaaaa aaaaaaaaa aacacagcac thtacataag agcttacagg	30540
atctattth ttttatccat tctthtthc atthttacaa tcattaatc aaaggaatta	30600
tattaattac thtctatgca cccgagtht tthtaacaca acaatactat cctgcatc	30660
agcaagthta tggthtcaa gagaggacac aaattcaaat gthtthtagth aagcagthga	30720

gctggctaga tatggaaaaa ttacaagtcc ctcttgcttt aacatttgct tgcccacatt	30780
tggtcagaca tcatgcaaaa taattttctca ctatagaaaa aaaaacacta caaaaacaat	30840
aatataaaga actgagaact ggttaactga agcatgcata tgtcatctaa aagaagcagg	30900
tgacgaccag cticatgaag tacttgccat gcatattggc acttcacaca ctgacccttc	30960
tccccaccta gaccagtaat taaacaggta tggatgagct agctactaag agcagccaac	31020
tgaatagctg actaacttag aagcacactt ggtaataata gctgactttt attagtactg	31080
actatactat atgctaagct gtactcaaag tgctttgagt tttaaaactga taaaacatt	31140
atatgaggaa acagaggtag agagagctat tcaccagctt accaaaggtc acatagctgg	31200
taagtggagg acttaaaccc agactatctt gtttcagaac ccacagactt aatccatcgt	31260
gcagaacata agacatactc catctgtctc cccaactagg ttattatgtg cacaatat	31320
tattggttgg ttggttcatt attatgactg ggtggtaagt atgtcattag gagtgttttg	31380
cttatgacta tataaatttc ttcacaaaa gaagactttc tgatgatata ctatgcatca	31440
gacaccacgc aggggtgctaa ggtaggaag ataagtgaga cttctagaaa ctcattcatt	31500
caacaaatat ctctaaggg ctagaagctt aggtttcagc agtgaacaga ataggtatgt	31560
tctctttcgt gtiggacctt atagtatctc tgggaaaaca gacattgaat aaatatcaca	31620
aatgcaagtg agtgtttcag agacatgcag ctgctacatc aaaacaaaac agaacaaaac	31680
aaacaaaaca aaactgacca gtgggattaa gtgtaaatag gcacacaaat gcacaaatat	31740
gcttttataa aatagtgaag cagtgcaga gacacacaca agatataaag acacaatgaa	31800
gaacaattga gcccaaagct ggaaagggtg agagtgtgaa ggaaaaaggt tgatcagaga	31860
agttttcccg aaggagagaa agcctggatg attaggaggc aaccactcgg tgactgaggg	31920
aaatctgaaa aatgtatttg tcatcttctc agacttgctg aaggaatgac ttgggtactt	31980
tgaggatttc agtaattttt ccatgacttg gtataatatt tcaaaaggaa ataggtgac	32040
tttatttgta taatgaatgt gactccttc tcgactgcca tagaaataa ctccttaata	32100
ttttgggttt gtctttgcac ttaagtaatc agtcattctg ttttttaca gggtgactct	32160
ggtggccac tagtacaaga agactcacgg cggcttttgt ttattgtggg gatagtaagc	32220
tggggagatc agtgtggcct gccggataag ccaggagtgt atactcgagt gacagcctac	32280
cttgactgga ttaggcaaca aactgggatc tagtgcaaca agtgcacccc tgttgcaaag	32340
tctgtatgca gggtgcctg tcttaaatc caaagcttta catttcaact gaaaaagaaa	32400
ctagaaatgt cctaatttaa catcttgta cataaatatg gtttaacaaa cactgtttaa	32460
cctttcttta ttattaaagg ttttctattt tctccagaga actatatgaa tgttgcatag	32520
tactgtggct gtgtaacaga agaaacacac taaactaatt acaaagttaa caatttcatt	32580

acagttgtgc taaatgcccg tagtgagaag aacaggaacc ttgagcatgt atagtagagg 32640

aacctgcaca ggtctgatgg gtcagagggg tcttctctgg gtttactga ggatgagaag 32700

taagcaaact gtggaaacat gcaaaggaaa aagtgataga ataatttca agacaaaaag 32760

aacagtatga ggcaagagaa ataatatgta tttaaaattt ttggttactc aatatcttat 32820

acttagtatg agtcctaaaa ttaaaaatgt gaaactgttg tactatacgt ataacctaac 32880

cttaattatt ctgtaagaac atgcttccat aggaaatagt ggataatttt cagctattta 32940

aggcaaaagc taaaatagtt cactcctcaa ctgagaccca aagaattata gatatttttc 33000

atgatgaccc atgaaaaata tcactcatct acataaagga gagactatat ctattttata 33060

gagaagctaa gaaatatacc tacacaaact tgtcaggtgc tttacaacta catagtactt 33120

tttaacaaca aaataataat tttaagaatg aaaaatttaa tcatcgggaa gaacgtccca 33180

ctacagactt cctatcactg gcagttatat ttttgagcgt aaaagggtcg tcaaacgcta 33240

aatctaagta acgaattgaa agtttaaaga gggggaagag ttggtttgca aaggaaaagt 33300

ttaaatagct taatatcaat agaattgatcc tgaagacaga aaaaactttg tcactcttcc 33360

tctctcattt tctttctctc tctctccctt tctcatcac atgcctcccc caccaaagaa 33420

tataatgtaa attaaatcca ctaaaatgta atggcatgaa aatctctgta gtctgaatca 33480

ctaattattc tgagttttta tgagctccta gtacagctaa agtttgcta tgcattgatca 33540

tctatgcgtc agagcttccct ccttctacaa gctaactccc tgcatctggg catcaggact 33600

gctccataca tttgctgaaa acttcttgta tttcctgatg taaaattgtg caaacaccta 33660

caataaagcc atctactttt agggaaaggg agttgaaaat gcaaccaact ctggcggaac 33720

tgtacaaaca aatctttgct atactttatt tcaataaat tctttttaaa ataatttccc 33780

tgccataatta tttatggaag ttatgacttt tgaaggacaa ttcaaaacca tttatttaat 33840

tggttctgca atgaaagaac tgccccatat actctactaa aggettgga ctttctgctg 33900

ccttttaac cagcgtata attgaggcaa gcgtccagct tgacacctcg agataacttc 33960

gtataatgta tgctatacga agttatatgc atggcctccg cgccgggttt tggcgctcc 34020

cgcgggcgcc cccctcctca cggcgagcgc tgccacgtca gacgaagggc gcagcgagcg 34080

tcctgatcct tccgcccga cgctcaggac agcggcccg tgctcataag actcggcctt 34140

agaaccccag tatcagcaga aggacatttt aggacgggac ttgggtgact ctagggcact 34200

ggttttcttt ccagagagcg gaacaggcga ggaaaagtag tcccttctcg gcgattctgc 34260

ggagggatct ccgtggggcg gtgaacgccg atgattatat aaggacgcgc cgggtgtggc 34320

acagctagtt ccgtcgagc cggtatgttg gtcgcggttc ttgtttgtgg atcgctgtga	34380
tcgtcacttg gtgagtagcg ggctgctggg ctggccgggg ctttcgtggc cgccgggccc	34440
ctcggtggga cggaagcgtg tggagagacc gccaagggtg gtagtctggg tccgcgagca	34500
aggttgccct gaactggggg ttggggggag cgcagcaaaa tggcggctgt tcccagttct	34560
tgaatggaag acgcttgta ggcgggctgt gaggtcgttg aaacaagggt gggggcatgg	34620
tgggcggcaa gaacccaagg tcttgaggcc ttctgtaatg cgggaaagct cttattcggg	34680
tgagatgggc tggggcacca tctggggacc ctgacgtgaa gtttgtact gactggagaa	34740
ctcggtttgt cgtctgttc gggggcggca gttatggcgg tgccgttggg cagtgcaccc	34800
gtacctttgg gagcgcgcgc cctcgtcgtg tcgtgacgtc acccgttctg ttggcttata	34860
atgcagggtg gggccacctg ccggtagggt tgcggtaggc ttttctccgt cgcaggacgc	34920
agggttcggg cctaggtag gctctcctga atcgacaggc gccggacctc tggtaggggg	34980
agggataagt gaggcgtcag tttcttttgt cggttttatg tacctatctt cttaaagtagc	35040
tgaagctccg gttttgaact atgcgctcgg ggttggcgag tgtgtttgt gaagttttt	35100
aggcaccttt tgaatgtaa tcatttgggt caatatgtaa ttttcagtgt tagactagta	35160
aattgtccgc taaattctgg ccgttttttg cttttttgtt agacgtgttg acaattaatc	35220
atcgcatag tatatcgga tagtataata cgacaagggt aggaactaaa ccatgggac	35280
ggccattgaa caagatggat tgcacgcagg ttctccggcc gcttgggtgg agaggctatt	35340
cggctatgac tgggcacaa agacaatcgg ctgctctgat gccgccgtgt tccggctgtc	35400
agcgcagggg cgcccggttc tttttgtcaa gaccgacctg tccggtgccc tgaatgaact	35460
gcaggacgag gcagcgcggc tatcgtggct ggccacgacg ggcgttcctt gcgcagctgt	35520
gctcgacgtt gtactgaag cgggaaggga ctggctgcta ttgggcgaag tgccggggca	35580
ggatctcctg tcatctcacc ttgctcctgc cgagaaagta tccatcatgg ctgatcaat	35640
gcggcggtg catagcttg atccggctac ctgcccattc gaccaccaag cgaaacatcg	35700
catcgagcga gcacgtactc ggatggaagc cggtcttgtc gatcaggatg atctggacga	35760
agagcatcag gggtcgcgc cagccgaact gttcgccagg ctcaaggcgc gcatgcccga	35820
cggcgatgat ctgctcgtga cccatggcga tgctgtcttg ccgaatatca tggtagaaaa	35880
tggccgcttt tctgattca tcgactgtgg ccggtgggt gtggcgacc gctatcagga	35940
catagcgttg gctaccctg atattgctga agagcttggc ggcaatggg ctgaccgctt	36000
cctcgtgctt tacggtatcg ccgtcccgga ttgcgcgc atgccttct atgccttct	36060
tgacgagttc ttctgagggg atccgctgta agtctgcaga aattgatgat ctattaaaca	36120
ataaagatgt ccactaaaat ggaagttttt cctgtcatc tttgttaaga agggtgagaa	36180

cagagtacct acatittgaa tggaaggatt ggagctacgg ggggtgggggt ggggtgggat	36240
tagataaatg cctgctcttt actgaaggct ctttactatt gctttatgat aatgtttcat	36300
agttggatat cataatttaa acaagcaaaa ccaaattaag ggccagctca ttcctccac	36360
tcatgatcta tagatctata gatctctcgt gggatcattg tttttctctt gattccact	36420
ttgtggttct aagtactgtg gtttccaaat gtgtcagttt catagcctga agaacgagat	36480
cagcagcctc tgttccacat acatttcatt ctgagtattg ttttgccaag ttctaattcc	36540
atcagacctc gacctgcagc ccctagcccg ggcgccagta gcagcaccca cgtccacctt	36600
ctgtctagta atgtccaaca cctccctcag tccaaacact gctctgcac catgtggctc	36660
ccatttatac ctgaagcact tgatggggcc tcaatgtttt actagagccc accccctgc	36720
aactctgaga ccctctggat ttgtctgtca gtgcctcact ggggcgttgg ataatttctt	36780
aaaaggtcaa gttccctcag cagcattctc tgagcagtct gaagatgtgt gcttttcaca	36840
gttcaaatcc atgtggctgt ttcaccacc tgctggcct tgggttatct atcaggacct	36900
agcctagaag caggtgtgtg gcacttaaca cctaagctga gtgactaact gaacactcaa	36960
gtggatgcca tctttgtcac tctttgactg tgacacaagc aactcctgat gccaaagccc	37020
tgtccacccc tctcatgccc atatttggac atggtacagg tctcactgg ccatggtctg	37080
tgaggtcctg gtccctcttg acttcataat tcttaggggc cactagtatc tataagagga	37140
agaggtgctt ggctcccagg ccacagccca caaaattcca cctgtctaca ggttggctgg	37200
ctcgaccag gtggtgtccc ctgctctgag ccagctccc gccaaagccag cccatgggt	37260
acccccaaga agaagaggaa ggtgcgtacc gatttaaatt ccaatttact gaccgtacac	37320
caaaatttgc ctgcattacc ggtcgtatga acgagtgatg aggttcgcaa gaacctgatg	37380
gacatgttca gggatcgcca ggcgttttct gagcatacct ggaaaatgct tctgtccgtt	37440
tgccgtctgt gggcggcatg gtgcaagttg aataaccgga aatggtttcc cgcagaacct	37500
gaagatgttc gcgattatct tctatatctt caggcgcgcg gtctggcagt aaaaactatc	37560
cagcaacatt tgggccagct aaacatgctt catcgtcggc ccgggctgcc acgaccaagt	37620
gacagcaatg ctgtttcact ggttatgcgg cggatccgaa aagaaaacgt tgatgccggt	37680
gaacgtgcaa aacaggctct agcgttcgaa cgcactgatt tcgaccaggt tcgttcactc	37740
atggaaaata gtgatcgctg ccaggatata cgtaatctgg ctttctggg gattgcttat	37800
aacaccctgt tacgtatagc cgaaattgcc aggatcaggg ttaaagatat ctcacgtact	37860
gacggtggga gaatgttaat ccatattggc agaacgaaaa cgctggttag caccgcaggt	37920
gtagagaagg cacttagcct gggggttaact aaactggtcg agcgtatggat ttccgtctct	37980
ggtgtagctg atgatccgaa taactacctg ttttgcggg tcagaaaaaa tgggtgtgcc	38040

gcgccatctg ccaccagcca gctatcaact cgcgccctgg aagggatttt tgaagcaact 38100

catcgattga tttacggcgc taaggtaaata ataaaatttt taagtgtata atgtgttaaa 38160

ctactgattc taattgtttg tgtatttttag gatgactctg gtcagagata cctggcctgg 38220

tctggacaca gtgcccgtgt cggagccgcg cgagatatgg cccgcgctgg agtttcaata 38280

ccggagatca tgcaagctgg tggctggacc aatgtaaata ttgtcatgaa ctatatccgt 38340

aacctggata gtgaaacagg ggcaatggtg cgcctgctgg aagatggcga ttgatctaga 38400

taagtaatga tcataatcag ccatatcaca tctgtagagg ttttacttgc tttaaaaaac 38460

ctccacacc tccccctgaa cctgaaacat aaaatgaatg caattgttgt tgttaaacct 38520

gccctagtgt cggccaattc cagctgagcg tgccctcgca ccattaccag ttggtctggt 38580

gtcaaaaata ataataaccg ggcagggggg atctaagctc tagataagta atgatacataa 38640

tcagccatat cacatctgta gaggttttac ttgcttttaa aaacctccca cacctcccc 38700

tgaacctgaa acataaaatg aatgcaattg ttgttgttaa cttgtttatt gcagcttata 38760

atggttaciaa ataaagcaat agcatcaciaa atttcacaaa taaagcattt ttttactgc 38820

attctagtgt tggtttgtcc aaactcatca atgtatctta tcatgtctgg aataacttcg 38880

tataatgtat gctatacgaa gttatgctag taactataac ggtcctaagg tagcgagcta 38940

gctgcaaccg aggaaaaaac gtgccatgag gtctctgtat ccaagtgtga ct 38992

<210> 20

<211> 34073

<212> DNA

<213> Artificial Sequence

<220><223> Recombinant polynucleotide

<400> 20

gaggagggtt ggtgctttgc taatggtgaa ttactaactc ctcaataaag aatattattt 60

gaaataattt ttgaaatttc ataattactt tgggttcttt cttaatgata aataaataat 120

agtatattac aaacatacat taatatcttc tgaatgaata caccacaaat ctcccttaaa 180

atatagcaag aataaaaaat atactatttc tgacaatttt taatttctca aataataata 240

ccactctgat ttttaaacat ctacaccact ctggctttgc caatcttttt aaaaattgaa 300

aagataataa ttttatcata attacactga agcatagaac ttttctttc aaggaaagca 360

aatttttgaa attctataat ataacctccc ataactctga ataaattaaa ggttcaacaa 420

cttagtaaag taagactgac ctcccttttt atttcttttt cagatcaaaa atcttacttt 480

tataggagca gttttcaact cctaaatgtt gaataataata gtcagttaaa ttcaccagct 540

acacaggaat acaggacttt gagggaaga attgaatctc tggtaagtta atatttgtct 600
 ttgctcttta ttccattata aaatgaatat gataataaac ctaatgtttt gtaatatatt 660
 ttcagtgtct aagtgtctta catattttcc ttccttgaat ggtgaaacat gtgtttctct 720

ctgcttttat ccagttagtt tactcataa ctggttctta ttcacatctt tgtcatgagt 780
 aaaaagtgtt agaaaggcca cgagtaaata tgcattttat ttgtttatga attcaaatac 840
 taaaagtttt ttatttgttt aattaagcat tgacattgtc tttttaaatt cttttcattt 900
 taccttcttc cctcttctct atccaactaa agacgcaaag caggaggtgt taaaaacag 960
 gtttaccata tcagcagtaa catagtttgg acaacattac actttggttc aatgatagac 1020
 atagaagttt gaacagaaat atgcaaagca agtttgagct ctaacttgaa gagagcctct 1080
 ggggtgcctgc caggaaacct cagcagtgga cccttaacat tcatgtgtca ccacaaacta 1140

ggggctgccc tttagttttg accagtctca gtgtcactca cttaccctta ctttttcaaa 1200
 aaaaagtcct aagaatataa agtaattcaa tggttctaca attttagcat gtaactgagt 1260
 cacctggcag ggttgctttg gtgagctcaa gataaaattt tatcagcatt tctacatttt 1320
 ctggaatatt ccttaatcca ggcttttaat cctttggtgc ttttctgaac cactgcaatg 1380
 agcttctaac tgttctcact gtgtgcaggc tcttttctct ctaactaat ttacacactt 1440
 ctgaacacaa atctctcaca gctgtttcc ttcattgtac ctccagctca agactttttg 1500
 cctacaaat aaaattcaaa cttgttagct aagcaccttc tcatgtctat gctttggctc 1560

atatttcagc catcgtgtgc cccacttatt cttatagcca acctgaaaag ccacttttta 1620
 taagaaacta cctctgtctt ccatgattgg atataattaa tctccttcc acatcacctc 1680
 gccacaaaat tgiatctgtg ttgatctcat gccacatacc tgiatgtatt ttatattata 1740
 aatatttgca gacttgttta atttgcatg ttagactaag ttccatgaag acagctccat 1800
 atccattcca tttttatata tccacaacat ttggctgggt tgatgcttaa taaatgttta 1860
 ttgaaggaac aggagtctcc cacttctgac ataatgaact tatttcccc agtggttaacc 1920
 ctacatctgg ttctgtcca agagtctctt cccaaatcat tctgattcaa ctgttcattc 1980

tgatctcatt aaacatttaa atgatatac taacttcgct tgctttattc tatgtcacc 2040
 ctgcagtctc ctcataactt ggtttcaatg atgcttgctt ctagagaaaa aaatgtatta 2100
 aataagctta tgattcagtc ctccagctgt gatgggtctc actgaacatt agctcagtgg 2160
 ttttcgaagt atggtctcta gcataacctt gaaacttggt agaaatgcaa attcttgggc 2220
 tcaccaagac atactaaac aaaaattctg acattggggc ctagaaatct gtgttttaac 2280
 aagctgcca gtgcagcctg gtcccttttc ttctcgagc cccactcaaa gctttcagtg 2340

ctcatctccc accaatgaca gggtcctctia tggaaaccgg caggacgggtt tccaactcta 2400

actacgtttt agagtttict tccatagggt atccaggcac caagtatcac aggttagttt 2460

cccaggggaag cagactctga gacttgcatt cagggagtggt ctctgggtg ctctcaacca 2520

acaccttcag gaagagaagg aagcagcatt gggcagaggc atagtcaaac tacagtgtctg 2580

ttggcacaga agactgaagg gactcagagc caggggtag aggtggggcc ttagcatcca 2640

tccttcacca ttaggtgtga gttgccccac ctcttgatg gtgtaacctc agtccaagg 2700

tgggtgggag tgcagcagag cagccctac aagggccaaa ccagagatac accaggcgcc 2760

agaagtgtg ccaggaata gagaggaaag gatgggtta aggtaggatc cacagaactt 2820

ggcaatggat tagaagacag gatgagaagt gacagggtta cactaacaca gaaatgtcta 2880

acttcggtag ataattgtgc cattggctag aagaggaaac cgaaatgaaa gcaggttgtt 2940

caggagaca aaagtctact gtggacatct cagcagagtg attcagtggg gaaaggaatg 3000

gatgccaga ccacctcaga ggaagatcta agctggagcc agcaataaag atacaagatg 3060

aacaatccct aacgaactgc tctcagcca tgctccccag acacgtgtct tcagatttat 3120

agtccgggtg aggttaggag gtgcgcctcc ctctgtggag gacagcaaag caccagtggc 3180

tccagggagt taaaatcttt tgataatttt tgttctagca tctgtctgca gagctgtctc 3240

tcagccattg cctgccttta cacaggagtg cagtcgaaa ttgggagatg agtgaaattt 3300

attatgccta gagatctgga tccccagttg ttgggagta tattttctga accacttgtt 3360

ggtttaagta atgcagattt attgatcca ctctcttga atctgtgact ctggaccac 3420

catctaagt aatgtgcaga gggaacggaa tgctgcaat agatctccat taaaaccagt 3480

gcatcctccc agacacatac agtagtaggg aggtgagtca atgtcaggac agcaccagct 3540

cccgttcgg tacatttcca aagttctcag tctgtgtaca aaggtttgct ctggggcagc 3600

agaaatagcc ctgggcaggt agtcaaaggc ctggtttgat ttctccact tccaggcaag 3660

tcactcgaag gctcacaggc ttttctca cctgccat gggccagtg agatctactg 3720

agctgtaaat aatgaaatga gtgtgtgtgc agtcatctat aagttgtaaa gtactagaaa 3780

atggtgaaac ttgggattt gggctattta aggtggaatg ctaaaaatgt caggcattgt 3840

ggagaaagga atttaaatat aagattgatt gactgggatt taaagacaaa tgaaggcaca 3900

cacgcaagt caccaccaca ctgacactgc acagctcccg ttggaggcat atctgacca 3960

tgcagacctg gggctctgcc tgtccaagt cactccttta ctacataaac cctccttctc 4020

ttttggggt gtcacccac cagagctggc accgagccct tgctgtgctg cttccctggg 4080

gtgtcagctt ttgacagggt gtttcctccc tctgcaggag ccttaacatc ccttggactt 4140
 cttcccccc acccaccccc agcagtttta tctcttcta actcgggacc ctttttttcc 4200
 cacacaaagt ttattgtcag ttgctggttt catctgtttg agcggctgca acaaaatacc 4260
 atagactggg tggcataatgc acgacaaaaa ttattttctc acaggagaag tcaaagatta 4320
 atgcaccagc agatctgggtg tctgaggggc caccttctgg tttgtagatg atgctttcta 4380
 gttaaacac ctatttaaca cactattaaa cactaagtgt gttaaatagt gcagttgatg 4440
 tatttgtcat gtcaccttta tcatacacia aatccttctt tgtctttttt tctgtactct 4500

 aatctctttc tgtaagtaat ctttgccttg agcagtagga tatttagagt actgtggctt 4560
 gacaatatat ttagtatttc aagatttcca tgaatttctt ctgatgtatg agttccctag 4620
 ttaatcttac atatgtatcc ctttgtaaaa acactttgaa catttaaaat gatacatgaa 4680
 tagtactcta atacaatgcc ataaaaatta taaatcattt gtatagactg gtaagtaaag 4740
 attgtgagat taagaaacgc atcaaaggcc attgagctgg aaagtggat aatgagaatt 4800
 caaaccaggg tctcttgact caaaatctaa ggatcatacc atttctcatg ataatatgag 4860
 tattatgttt atctctatcc catagacaaa gtgttaacac tgaatgagca gtgaaatagt 4920

 ctgagaattt tttattttat ttagcaattc acttgtcatt tctggtcctc agttttattca 4980
 cgagtaaaat aaaatagttg gactagataa tttctatagt acattcttac acaaaaaatc 5040
 tatgattttg ttatttttaa tgtgatatac tcatggcact cattcacctc attttcccag 5100
 cctgcctcac tggtcattac ttctctgtgt tctttacagg ctcctccctc tctacactgc 5160
 cattaaatat tgaacacct caaagcttta ctatgtcca cctctctct gacactatca 5220
 ttctgtctag atgateccat acatacatgc ccattacttc aacctgtatt tatacgccaa 5280
 tgattcacta tatttccagc ctagacattc tttgtactc tagttaccag ctgtatatcc 5340

 ttacatggct gtttcaaac aactcaaata tattatctct caaaatcaaa ctcatgatgt 5400
 cccacacca tctagcttt ccaccaaaa tacctatccc tattaatagc aataccattt 5460
 attcagttat ccaaataaaa aacctagaat tcattcttaa aattctacta tcattccaaa 5520
 tatectatcc atcagcagcc actgtattct taatcccctg tatttcttc aaatecattc 5580
 acctctctcc atatccattg ctgcatgact atccaagcca tcgcctctac ctaggggtac 5640
 caaaatagca acaaacctaa tctgttcatt tgcattattt tttctccaaa actgattatc 5700
 tatatgtagc aagacagatt gtttcaaat tgcaaatccc actatattat cctcttgctt 5760

 caaacacttc catggtttcc cattgtttat gataaaacca aatgcttcaa gttcgaagac 5820
 cggcatgatt gggaatttcc tgtcacccta gctacttgc tctccatggt acagttgcac 5880
 tggctttctt tcatttctta agtacaacct gtttctccc acctcaggac tgtgcatgtg 5940

ccattcattc tgctgaggag cctttttcct tccacttcaa tcagctaagt ctgattcttc 6000
ctgacaatct cagctcaata agcatttcct ctaagaaatg tctctaatat cattaattgg 6060
ctcaggtecc tctactgtat tgctgcactt ttcacagtta taattttact taattatgaa 6120
tgattatttg attaggtcta tttccatcca ttagacataa gcttcatgat ggccagatta 6180

ctgttttcta tccatcgttg tattccaata cctgacagaa ggagggcggg aggtggtggc 6240
acacaagaga tgctcaaaaa caattgttga ataagtaa atgaatgaggcc atttagaaat 6300
aacgaaagta ccgttttaca aagtacatgt atcaaaacta tgaatgcatt ctacttacat 6360
ggttttctcc aaataaaaaca aaagacttca atcaggatta atacctggga taaactgagt 6420
cattaatatct ctcttttgc atcaggagt acattgaaac aaatgtctgc aaacaacaaa 6480
tacttttttc ccaaaatata ttgaatggca tttccataaa caaactagaa catgggagga 6540
gaaagaaagc aatattaatt taaaattaat ctatcacat aacttatacc atcagggtatt 6600

tcgggtaaaa ttcctttcag gcacatccat ttaacaagaa ttgattgtta ctgaaagcct 6660
agaagagaat ttggcacata cttggtgttc aaatatttgt tgactgagtg aataaatgat 6720
gcaagtgtct aagaacaca aaataaggac atgattacag tcacgggtgga gttcacagtc 6780
atctccaaaa tgaggatatg catcccaggg aggaccaaca attcattgga gtgctgaaat 6840
aaaatactca aaggtcattt tacatgtatt ttttctctaa attacttttc ttaagacaca 6900
gaaaacaaaa aaagaaactt agctttgtta ctttctaaca aatagttaaa tcattaacaa 6960
ggattgacac tagcatcctt gtttggtctt atgccttagg ggaacatgaa atgtgtgaag 7020

acattctgag atctgaggga agggtagaca gtaatacagt gggactgacc aggccttcagc 7080
acacctttac ctctctcag cagatttcag tgatgagcag tttacaacta gattgaaaga 7140
ttatattatc tagttctaaa agaaaactaa gcctcccaaa agcaacaagg gaactgagag 7200
gaatcctgca aaacaaaaac aaattttaaa acttgacatt tgtaataacc ctaatatgta 7260
atcacagtaa tgaacagtaa gataatgaca gaactgacat atttccttat ctattaaagc 7320
catattaaca ggtaaagcaa tgccagtcag tggtagactt cttagaagat atttaataca 7380
tactagacac atacacacac acaacatttt cttcaagggt gtatgtatca gaaaatcact 7440

ttttaaggcc ggatgcagtg gctcaggcct gtaatcccag cactttggga ggccgacgtg 7500
ggcggatcat ctgaggtcag gagttaaga ccagcctgcc caacatggcg aaaccccatc 7560
tctacaaaaa tacaaaaatt agccagggat gatggtggat gctttagtgc ccagctactc 7620
aagaggcaga ggcaggagaa tcacttgaac ctgggaggca gaggttgagc tgagccaaga 7680
tcacccattg cactccagcc tgggcaacag agtgagactc tgtctcaaaa aaaaaaaaaat 7740
cactttttag ataaaaatca tgctatagag agaagactat gaaaatatgt ttagcaatgt 7800

gtccatcatt aggtgattga gtttcctttt gttttgtttt actgaaaatc atataaagta	7860
tgttatctgt aaaagtcttc tgacatgcac acataaaaat ttgggagaaa agattaacta	7920
taatgtttta tagattttgt acacatttct ttaaaaatat ataaaacaca acacctttca	7980
attggtttgc aagaataacc aattgacatc atggaaaatg gaaattcact tgctgaattt	8040
taacaaaaat ttgcatgatg agtgagactg acaacttagt gtcatgattt aatgaattat	8100
gccaatggta aacttcatgc acatggggcc aggttaattat gtggaaactt tttcaatgct	8160
taaagccaag tattgaaatt aaacttagaa tcagaccttt gaaccatttt atgacaatgt	8220
tcaaaaatta taaattctat ccacttata tataatatta aaaatatcat tacaaaaaaa	8280
acctgtgttt attttataac tcagcctttt taatttctaa tttcataaat atattataat	8340
ggatattgtt agtaatgtag tattattaca tgtatataat ttataagtaa atatacatgt	8400
tttggtctact catgcataaa atgtttcacc cataggagca cataatcaga aatgtctgga	8460
gaccattata gtaatagata gatcatattg ccacatattt tatctcctcc ttgacaactg	8520
agctttccag atcttctggt gaaacgaaag agaaagtgtt aacagaagag tgattaaaat	8580
gacaaaagca ttacttctat tacttctatt ctaataatat gagcaaagct ataactatca	8640
agtaataatg cactaaagaa ggtgattaat ctgatatatt cacaggcaac taataagacc	8700
tttctattgc agccatgaaa aatatgtgac aattatagat atcctgtgtg cagtgtttca	8760
acctttatgt gacctgttct actaacagat ttagtgatgt tcactttgtt agaattttct	8820
tacacatgcc ataacttget tcagtccttt gattatgaat attatggata ttaaggattc	8880
tagactattc tagattttaa aaataatatt gtcacctcaa tcagaaggga aatattaaat	8940
agttctcatt tttcaatgt ttactcagtt ttgtccaat gtaatgaaag tgtcagcagt	9000
acaggttaca aaataaaatg tgtattaaag taaactcatt tgaacaggtt aataattgta	9060
gagggaggga aaaggctaaa agattgaatg taaaacttat gaaaagtaga tacatcgtct	9120
ctatgatttg cagtagtcaa ctgcatacag atgaatcatt ttaatacacg ttaactactt	9180
tccttttaca gatggagaaa ctgagaggaa gaaagtttat atggttcatt aaactttgtg	9240
atgcaagcta aactaacctg tctctgtatt ttccatctac tgcccttacc actatctcat	9300
tagaatactc ttcaagcatc tccttactga ttttcttacc aagcatttgt taagttctaa	9360
tgagagtgg tagtaacatt ttaccacct ctgtgaaata tgaaatctta ttcataggcc	9420
tcttctttta ttcttgtatt tgcatatcaa ccaattaatc aacttgcttt ctttatgttg	9480
cttattatct tagtccttac taaattgcct cttaatgttg tccacataac agaaatgtta	9540

agggtggatac ttaacatttt agtccagtct agccgggtgcc agtgcaatgc caaatcatga	9600
attaaaaatat aattacaaga accacttatc aaattttaac aattccttca gctttgtgac	9660
agttttttct acttcgatta aagtcaagta aaattaaagt taaatatttt tattaaaata	9720
tctcctttaa cattccatat taataaacat attaaagctc atgcttctaa gtagattact	9780
agaagttact ttatcgaatt acagcaatgg ttaattctag atcatagaat ttagaatgac	9840
tttttgccctt cttctttttt ttcccttttt ttaaacagag tcttgctctg ttgtccaggc	9900
tggagtgtac tggcgcgac ttgactcact gcgacctctg ccttgcagg tcaagtgatt	9960
ctcctgcccc agcctcttaa gtagttggga ttacagggtc ctgccaccac acctggctaa	10020
tttttttttt gtatttttag gagagacagg gtttcacat gttggccaga ctggctcga	10080
actcctgacc tcaagtgatc cacttgcctc agcctcccaa agtgctggga ttacagggtg	10140
gagccactgt gcctggcctg actttttgct ttcttcttaa tacttactag tatttcttga	10200
atttttaaaa aagaaacata aagtactttg ataaaaccaa cagtctcatt gtctttaaaa	10260
ttgttcaaag gtctcttga aaaaaaaag aaaattatca tttggttaag aatcatgttg	10320
gtctgacatc aatcatccta taggagtga tattgaaaaa gtaagatata ttgtggtata	10380
atcgagattg cataaatttt accatttttg agaagaatct gctccaaatc ctggcttaat	10440
gtaatatcca gcatgctact taattttctt gtcttcacct ttccatatcc acatccacct	10500
agggtgccacc tcacagtata agccagcata atccattctt ctcaatgaaa ccacaataca	10560
tctgaccctg catctcagga gaactgtatc agccacagca cttccagttg actatgaatc	10620
tgaatgttat gcctcaggag aaacatcctt gctgggactg agtagtgatt caaggagata	10680
gttatgattc agtcaagaaa ttaataatta gtgttatttt tattattgag acagagctc	10740
gttctgtagc ccaggctgga gtacagtggc atgatctcgg ctactgcaa cctctacctc	10800
cccggttcaa gtgattctcc tgcctcagcc tcccaaataa ctgggacagc aggcaattgc	10860
caccacgcct agctaatttt ttgtattttt agtagagacg gagtttcacc gtgttagcca	10920
ggatgggtctc gatctctga cctcaaggtc cactgcctc agcctcccaa agtgctggga	10980
ttacaggcgt gagccactgc gcccggccat aaattattaa ctgagccagg cacagtggta	11040
cacacttata gtcccagata ctgaggagac tgaggttgga gtatcctttt ttatgttatt	11100
ttatttttaa ttattatggg tacataatag gtgtacatac ccatggagta caagtcatgt	11160
tctgatacag acacataatg ttttaataatc acatcagggt aattgggata tccatcacct	11220
caagcattta tctttctttg tgtttagaac attccacctc cactcttgga ataggcaccc	11280
tgttgtgcta ttaaatacga ggtcttattc atttcatcta actatatttt tctacccatt	11340
aaccatcacc tcttttcccc tcttccccac tacctttcct gtgaggctgc aggattctta	11400

agcacaacag ttagaggcca gcctggacaa catagtgaga ctcaatttct aaaaaataaa 11460
 aaagaaatta ccaactaatg ctaaaaaat agtctctgat gcttaggtat gaattagaaa 11520
 tgaccaaaaa aaaaaaaaaa aaaaagactg ccttttgctt ccttctcccc ttctcttcaa 11580
 gttttccatt gctactcatt ttagtctggt ttaatcaggt ttcattccatt aaaagcaatt 11640

 gttgggatca cacattttga gttgtgtcag tggacttccc tcatgctggc atgattcctg 11700
 ccccaagccc ttagtaaaag ccaccaagcc atataacata atctctcatt gagtaaaaca 11760
 tctgatgtgt ttagaatgac ttctagcaaa aaaccagcct gtccagcatc atctctgtat 11820
 aacagataaa ggaataggtg ctgcatcaaa aggttataga acctgcccac atcaatccca 11880
 tgtgttttgc aatggaatta ggttgaacta aagtgaatat tcagttttct actcctcatt 11940
 aacatgtctc atgttgcaag gttgagagga aggagaagaa gaactgtatt tacagagaga 12000
 ttccccctct ctttctttct acagattact aaaacattca aagaatcaaa ttttaagaaat 12060

 cagttcatca gagctcatgt tgccaaactg aggtgagtg aactgtagaa aaaatattta 12120
 agtatagata caatgtggca tacttgactt ttgtcacag aatgaatagt aaatgacatg 12180
 ttcagataag ttgttgtaat attatgaaaa tagtatttta gtcagcttaa aaaccaatgc 12240
 caaaaaagcc aaacatatga tctatttagc tactaatgta aataaccata ttatatctat 12300
 tcttattggg aagaggaaga aggggtggag agagagttgg ggtgaaggta cagtaacaag 12360
 gccatcctat tgtaaaactc cagtgagatg cattcacagt gcagcctatg taaacagtcc 12420
 ctcttgaggt tgtacaatgc tgtggtttgg gtgtatccat ccaagatcaa gacactatga 12480

 ccaacatcaa aagtggcttt ttggttttat ctgcctgatg tgctataata aaagggtatt 12540
 atggccaaat ccaaggcatg tctatcatga attaataata ggaggagtag cagcatgcat 12600
 gctagttatt tgccattcct gccttagtta aatatgatgt gataaaacca gcctttccaa 12660
 ctgaaatagt cacctttact gactctcccg caaatgtctc aaatgaccac attgctctag 12720
 tctttaata atagcaata gttctttggt agaagaggaa ttataactat tctttctcaa 12780
 atactagcat cacaagaaaa ttaattcttg ttctctggag agtcacctag taagtatctg 12840
 gagcacagat gtctggtcag gtaagttttg atgaggagtt aaagggataa gaagagtcca 12900

 tgagaagggg attttccaaa acacctttcg gtcaattcag tgcacattca cttagtactt 12960
 tcttgtcagt atctgtatca gccactaatg ttcaaaagtg agtaagccct gaaaacctgt 13020
 aggactacat gagccttctg ccttttctct ccttttgctt acttcccact tatcactcaa 13080
 tcctctgcaa cctggcttca ataccacat aaaatatcaa ctgctcttgc cgattcaaca 13140
 atgacatcca gataacaaaa tccaaagaaa ccacatcagt cctattcttg gacctttcaa 13200
 cagtatttgg tcctgttggc ctgtcactcc ttgaaatagg actatccctt ggtttgcatt 13260

gccttgata cctgatttt ccccttacct ccctagctat tccttcttag ttccctttac 13320

taggtcttac ttctttgtat attcctttaa tgttgctgaa catcaggctg tgctctaggc 13380

ctctcatctt ctcaggcac actctctect ttccctggcc ttcactgcca cccatatgct 13440

gagtgcctc aaagtgtgat ctctaggcca gtctctttt gccccaac atgaatatat 13500

gcagccatct acttggtacc atcacatgga taattctcat gatctctcc agtatgactg 13560

cttctttatt ttttctggg ctctttttta gcattgcttt acatggaact ttatcatgtc 13620

tctcaacctc tattttatct tttatctatg tatgtagagt ctgtgtaatt tcttcatctc 13680

tttagataa ctaatatctc ttcagctttg acttgatttc tgtgtaaccc atttattgcg 13740

tttcaattt caatgagtat gtttccctat ctgcaagttc tatttgtttc ttttgagaat 13800

cttctggtc ttttaaacac atttcttatt ttaatttttg ggggtacctg gtagttgtat 13860

gtatttttg agtacctgag atgttttgat acaagcaaac aatgcataat aatcacattg 13920

tgtaaatgg ggtatccatc cctcaagca tttatccttt gtgttataaa caatccaatt 13980

atattctttt agttattttt aaatgtacaa ttaaattatt attgaccata gtgactctgt 14040

tgtgctatca gatactaggc gatcttttaa aaataatgtt tttacttaa tctattttt 14100

atgattccct cttttacgtc atttgcatt tcaaatacag tcaattgtct gttgattcta 14160

ttatgtgaag ttttgagga taatctttt gttactttga ttccaccttg gtatggtttg 14220

gctgtgccc cactaaaac tcatcttgaa ctctgggtcc cataataccc acatgttgtg 14280

ggaggagctt tgtgggaggt gattagatta tagggacgtt tcccccttt gctctgttct 14340

tttctctgcc accatgtaag aaagatgtgt ttgcttcccc ttctgcatg attgtaaatt 14400

tcctgaggcc tccgcagcca tgcaggacct cttttctttg taaattacc agtctccggc 14460

ggttctttat agtctcgta gaaaaaacta atcacacct catgatgtat tgtttaccac 14520

tgaaattgta tgcttaaat taatctcact tgggacctg tacaacctag acttaacata 14580

tctacctca gagcagttac atctgtcaga cattctagag gaatcagcag cacatggact 14640

ttgttgttgt taatttgttg tggggggagg ggggaggat agcattagga gatacaccta 14700

atgctaaatg acgagttaat ggggtgcagca caccaacatg gcacatgtat acatatgtaa 14760

caaacctgca cgttgtgcac atatacccta aaacttaaag tataataata ataaaattaa 14820

aaaaaaaaag gttctgggag tattcaggta gtattaaatga agattcagac atcgtgcagc 14880

caggcccatg ctatgaatt ttcaggatgat acttcttttt ctttttctt aatttaaagc 14940

tggatctcgg aaacagataa atttattttt ttatgacatg acgagcattt ttttcattct 15000

agttcatgct gttattgggt gtttagttct ttgagactcc tggccttttt ctaaaacctc	15060
aagttcaact tcctattttg cactggccca aggtcccatc tccagtctct atgtaaatgc	15120
taaacataag cctgtggaat attctagtct caccacatac tattcacatt cttctttgtt	15180
tttggcttc caggattttc cttacttttc tatgaacca gtcttgcat tgaatggaa	15240
tttattatat attatctatc ctttctatit gttttatgca gaaagtgttt tctaaaatta	15300
tttagcttc catattgcta gacatggaag ttgtaattat ttgttcagt cctgtttcta	15360
catctaaact gcaagacca tatggcaact gtgaatctta gtcccagcta atttctgaag	15420
cttagaatag tgcctagcac aagaagttgt ttatctaaca tttttaaaaa taaatattaa	15480
attcatatct ggaatgaata ttaagttaga gctggtcatt gaggtgagag gaggaagcca	15540
agagagaata tgagagcctc aaagccaaat atctttaatg tactttttca gaaaagaaga	15600
cagccaatgt cagggtggagg aactggttta tgaggtaact ttcctggaag aaaatagaaa	15660
ttactgaggt ttiagataat ccaaataatt aatcaagtca ccaaggttta ttgtggggaa	15720
tctttattat taattaaaat gagtgatgaa atcttaatat acgacaaaag ttaaaatttg	15780
cttttgcagg cagatgaatg gtctaggtat caaaaaatta agttgagtct ctaactcaca	15840
caaatttaca accctatcac tttatgaatt tgtttaggag attattttta ataacttg	15900
tgaagtctaa gaatagctaa aatttatagt acacttattg tgtgctattg actcttcttt	15960
gaagttttgc atatagtat tcatctaac ttcataacce attttacatg tgaagaaact	16020
tagatataga aagattaaga aacttacata acttatccaa agttacacag taaaactctg	16080
gcattataac ttcaaatca gctatcctac agtgagtaca gtgttctgtg cattgaaatc	16140
aaataagtga gatagcatcg tgatatagta ttacgtatgc aaacactgtt acagagatct	16200
gtctaaagtt aaattccaca aatgaattct ttaaaagggt ttaatcaaga agaatatata	16260
aacaggatgg tgaaaaattg tcatattatt tgttttttaa aatatcttta tgatttacag	16320
gcaagatggt agtgggtgta gagcgatgt tgcctatgaaa tttcaattca ctagaataa	16380
caatggagca tcaatgaaaa gcagaattga gtctgtttta cgacaaatgc tgaataactc	16440
tggaaacctg gaaataaacc cttcaactga gataacatgt aagtataatt tttcataaac	16500
aattttattt caatataatc ctcaagttta ccaattcaaa ttcataattt aattgagagg	16560
ctgacttttc tttctttgaa actaaactgt gaaaacaatc cattaaaaag ctaaatatac	16620
catatagctc cctaactgaa atcatcttaa gacttaaaga atcatttggc atttatatag	16680
taaattttat ttgctaaaaa ttctcattaa ttatccctgc aacattcctt atgagtgatg	16740
ttactgtcag atgtcattag tggataggcc ataggagggg tacatagatg ctcaaggtea	16800
gagaactatt taattaatga tccactcag aggccttctc atttttcttt gtaacattta	16860

tcacaattga aattacaaag ttatctgtgt aaattttgta ttgtttggct tcatcctaca	16920
ctgtaatcat cctaaaagaa agaaccagtc aaccttcttc atcctactac cctcctacca	16980
cccagtcctc atcatataac acatattcaa taaataattc ttgcatgact gaaagaaaag	17040
aaataatata tgcatagaat ttaaggacat tcctccaagt tggttacatt ctgctagttt	17100
aataagccat tatttcttct cgatgagctc aagattaaaa ggattttgat gattcccata	17160
ctagactggg aggtaccagt tacagatgta ctaactgtta aatatgaaa tgctttccta	17220
tttgttggta aacaattact gcacaggcc cacaaagttg tcttccgaga tgtttcaaat	17280
ccactgcccc tgcigctaaa gagttatgct tagcaaagca aagcactcta agacactgct	17340
ccaactccat ggctgattg catcttttat gactggccaa tgctcacgca ctgcagtttg	17400
ttaggtagtt gaatattacc tctgcttcca cacattaagg aatgctcccg aacgcacttc	17460
ccaagtgttt atttatttat cattatacta gacaatatgg tgatacgatg gtcacagaat	17520
agcggtttcc acctccagag ccataaatct agttgaaggg aaagatatc caacacaaga	17580
gtgttgacaa tcaagataga atatgatcaa gggcccagtg tgaggcccag gcaatgatca	17640
ctgcaggaat ctggggaaga aagagaccag ctgcttggg atatctagca aaagtttcat	17700
gaaggagaat ggactttgac ttgaaatat gggtaggatt tacatatatt gagatgagaa	17760
aaagaaagtt ccagagaag gaaagcatga aaaggcaaac agtctgtact gaacgcgatg	17820
ctttgacaga ataatgaaga aagggacctg ctggaatgat tgatcagtg tcatcattca	17880
caccatcatc atcaaaacac ttatttaatg agaacttact gttttttagg catggcttta	17940
atgcctata tgaatTTTT tcttgattaa tccttacaac aaacatatcc catagatagt	18000
tttattgtcc cccttagaaa agataaatig ctaggctga cacagtcagt atatgaggca	18060
gtcaggattc aaactaagtc tgtttgttca aaaaattaag aatggccagc tttttaaat	18120
tttctgtctc cagaagtatg atttggctcc actgaagttt gcaaaacaaa tgtgataccc	18180
aaacctgtgt aaacttttag tgggaaataa ctttgcataa gtcggtttga gagagcgtgg	18240
aaacctgtct tgaagagttt taatttaact tgcaggaaat aaaaatgatg ggttttcaa	18300
ttaaaaaatt caatcaagga aggatatgag ctaacataac atttttttaa aaagatcagt	18360
ctggttaaggt agaggtgcat aaactgaaaa ggagcaaaag tgggtgaatt cagttagaaa	18420
attattgtaa ctgtactgat gtcaaatgat gaaacatga actaaagtag taccaaaagg	18480
agtgaggagg atggaataat tcaaaagata gaggacagat gtgcagaacc tggagattat	18540
aagatgtgaa aggaggagtt tgagaaaatt tcagattttg gaagtgggtg cattttacta	18600
aaaggatata ataagtagca aattttggat aaagttgggt cccactgagt ttgagatggc	18660
tgttggacat gcagagaaaa ctgtcttgta tgctgttctt aaattgaaat agacagacct	18720

ttaccctctg atactgacat attttccttt ccaggctcac cctccatttc cctaaacaca	18780
acacatgcac tagctctcct tactttattg ctccacaaac atcttacacc tccaagcatt	18840
tgtgcccact gtaccttcta tctggaatct cttttgtcct cttgtgtgcc tgaaaaattc	18900
ctttcagatc ttcaaaatac agtgcagatg ctatttcttc tagctcaaat attatctcct	18960
ccatataatt taattactct cttttttctt ttctctactt tgcacttaca tttatttgaa	19020
tgattgcttg attaatttct acctgtaaat tatgtgaggg caggctctct atattttgct	19080
cgcagttaaa tctgcagcac ttattataga gtggtatcat tagagtaata tacatatatt	19140
tgaggacatg ataaattaac ttccctata gtatttatca cattgcatct caatgacttg	19200
cttatgtttc tgttttccca tataaattga gtaacttgaa aaaagagata tctattaaagt	19260
atttaatgag aaattaaagt acaaacttta gtatgcataa caacaattg ggaaaagggt	19320
gtaaacaaag agatttttag ggcccatgag tttagagatcg tttcagcagg tctgaaagga	19380
agcctaggaa tctgcatttt agaggaccac ctcccaaccc caacaagtaa ttctgcttct	19440
tgttgtctgg gtactgtact ttaagaaatt atggtgaaat gatacagcc tttattgtat	19500
ttatcttatt ctcatTTTTT aatactagca ctactgacc aggtctgcagc aaattggctt	19560
attaatggta agttttaata ttattttgta actgtaattt gccaatcat aaagagtaaa	19620
agtgaagtc ttttgtgtac ttttggccaa ggagctatct atcaagttga tgtctttgtt	19680
cttagttcgc tcagggtggtg ttgaaacaag acagtgtga tccaagtggt cccatggagt	19740
ggactttagg tttccctttt ctttttagaa aaaggaagaa gttgtagtgg aggactacc	19800
actctgact caaaattgcc ctcatgaaaa tttctttggc agctttgaga accttttact	19860
gccctggttc taagggtgca tttctgtaga cttacaaatt atgtttgatg acaccgttta	19920
tgtagcttct cctaaccacc agagtactt gctttgttgt gaattcaggt taatcacaaa	19980
gtataataaa aaagaattgt cagaagtctt cccagctttg ggtctataac ctgaaggaaa	20040
agtcactact cttaacatc atcctatgta ctctcaggct aggatagcag aaatgcaatc	20100
cctagaaaac agcaacttac ttctctgacc aaaaaaatgc agttaaaaat tagttcaatg	20160
tacctggtag ctggcctatc ttaggtactt cagtgatTTT acaaagtgat ggtagtccta	20220
tgggtgtttt tcagcttcac tacgtattta attcatgctt attgttaatg aaactgtgat	20280
aagcaattta ctagggtatt tgtttgggag atgccacaaa ggaacacatg tatctcttaa	20340
tggaagcctg gtctctcttt atccaggaaa tttgctagga aaaaaagcc tttaggtggt	20400
tgtgctatta aaccagggca ctacttaaaa gccagcccag caatagtTgt gtgatttacc	20460

attaatttct tagtaataga ccacacaaaa gaagaaaatt atgggaatgc gagttgagag	20520
gaattgggtg atcagcctac cccagcccg ttcagctctg gccagtagac tattcacgag	20580
ctctttgaaa acatttaaat aaaccttatt tagatactag aaacctctg tcacctcaa	20640
gaatatcttg tggatagcg actcctttat gagggcatgt ttggtaatag agcatcagtc	20700
ttggaggtgg actgattctt acaaggtgaa ctgcagtcac taaggagtct tttggatgag	20760
accagttttc ctccaacttc aatgtgtgca tgaacctcac atcaaatgt agcttttagat	20820
ttgtcccatg atgtggttcc aagaatcagc acttctaata agtttccagg ggatgcccat	20880
gctgcaggcc cacaaccac actgagcata gcaagactat tgagaaaaag gaaatttccc	20940
aggagtcctgt ggctgagct ggcacatcca ataatgacct atcttaacct caactcatga	21000
ggaattccag ggaactctga agctgctcaa aatttgaagc ctatatgcca actaaattca	21060
gaaatgttct ccaaatgct atctataagc aacagtagtc acaaatgcat ttagaaaata	21120
tatcgatcat gctttttgga aaatccagca tgcctgagg aagaatgtat aagacataaa	21180
agtcataaat tatggaaaga ctcttcagct tcttccaaat gtaaaggaat catgatcttc	21240
ccagcacatt aatgcccttt ctcatagaa tgtggggccg gtccagacct aataacattg	21300
tctgagcaga gaatccttgg aggcactgag gctgaggagg gaagctggcc gtggcaagtc	21360
agtctcgccg tcaataatgc ccaccactgt ggaggcagcc tgatcaataa catgtggatc	21420
ctgacagcag ctactgctt cagaagggtga ggccaccact acctacccat ctgggaacaa	21480
ttagaataga caggatcatga agactgcacc ctctacccta ggattgaatt gagccagaaa	21540
taattcaatg caaaaaaatc agtaagaatt ttcttctat tcatgaaagg aaaaggattt	21600
ttcccttta gcatgcta ttagtgctat ttctctgttt caggtaataa tatattagca	21660
cagtaagaa caaagattta tatgtcagaa tgttttttaa atcctagcta taaaagctta	21720
agaaatttac taaatctcca taagctttat tttttttcca aattaaggga caacactgtt	21780
atctgtgact tagtgttact ggtagcattg agtacctaa tgtaaacata cgttaatgt	21840
tagcgaacg aattgctgtg gaagatttgc acattatct atgggagctg atggctaacc	21900
tagagactgc cccatgccat taatttattc attcataaag attattgagt atctagtatg	21960
agcacagtgt tatatatgtt agaagctact agtataaaca aagtattgcc tctgccttca	22020
aagagcttac actcgaatgt tggaaatcaga atgcacaaaa ataagatca attacaatga	22080
gtagcataaa taaaattaat gtaggcaact tacaagaatt ctttaattgag gtgactaaac	22140
tattgccaac actagggtga tatgtacca gtggcgagta ggttgcataa acttacctta	22200
ttggtaaaaa gaaaagtcca cattgctcat aaaagaagga ttttagattt cagcataact	22260
aaaatctgtt tcaaacctgc cttgttactg gggcatcgca gaccacaaca gttgttggga	22320

acttaactca aaaagttcac ccagaaaaat aatggagatt tgaactcgtg tgcccctgac 22380
catatcaatt ttcttctcag actcttactc taaactggac ctccttatca cacacacaaa 22440
gccttcata ggcagatcaa tccagtctta ttctcacaag catgtacctt gagcttcaga 22500
taaacagcat tgttctcttc ccctggactc ttcttacatt tccctaccta tgagtatctg 22560

atcaatctgc ttatccttga aatgttaata tatttaccac atctctatct gaattttatg 22620
aaatttttga taattttctaa gtagtttttt cagatttata ggcactactt catggtacag 22680
tgactgttac aaacgtatct gttaaattta gaaggaataa agatttataa gactagggtta 22740
gttactgaac taaagtttta ggaaatccca aattatttca aatttttctt atggtaattt 22800
tatgacttaa tatttttata tgcagtgaac aaatttgaaa ctttaaaaga tactcccaga 22860
attatcagtt ttctgatgta gattggcaaa ttattacta tatcccaaat aaccaagag 22920
acaaaattca caaaaacatt tcaattttca ttgccacttg aaaggccaaa aagcagaaat 22980

ggcacgcatt gatttcaatc gtactcttga gtgtgggaac caggaattaa aatacctgga 23040
cttatcaggc acttagcata accaagaacg gaatagaaac ctccttgat tctaagccct 23100
attcagtcac aatcacaaa aaccaagtaa acgatatcac tataatgaaa gccacagtta 23160
taaataatga caacgattac caaaggaatc catggaactt tgaattttgc caccacacat 23220
ccttctattc attaccatga ttgatccact aaagctaaca gactctgtga accttgtatt 23280
ggacccctcc ctaaagacct gattgtcact gagaaccatc agtgaggatt tgtttggggc 23340
atgaccagcc ttacatcaaa gtacatagaa gtgatgaggt cttatcaaag aggattattg 23400

aattatcacc tcttctatgt agctttccct gatctctct ttctctcca ttgagttcca 23460
cagaaatctt ttatctgcc tttaacagtt gtctcatga ttgtgatat ttgacttacc 23520
tcttctcagt ttcttctact agtgtagagt tctcaaaga aagagaccat aattacttat 23580
atctttatct ctggagactc atactattcc ttatacaaag tagacactta acaatggctt 23640
gttgaactat aattaatgaa aataatagct accttcatga aagttcactt tgtgccaac 23700
actatagttg acataataga ttgtctcat taatacttaa caattgtgtg agaaggtatc 23760
accaatcaca ttctatctgt aaataaacc cagagctatt aattaacttg tcataaataa 23820

cacttttcat atgtggcata gccaaagattt aaatataaat gttactggtt ccaaatgat 23880
gtctaatctc acttgctgga aagaaggaaa ggaagaaaat aaacgagtgg aaggaagaga 23940
gggagggaag agagaaaagg aaggaaagaa aaaagagtct cttcagaacc ttcactgtaa 24000
agactccgag caaaagaagt tgaatataaa aacaacatag gttgtttgt tttctaatat 24060
tttttcttca aaatttttaa ctgaggttca ctcttacaca aactactgtg tcttataaaa 24120
gtatttccgg tcatagaatt ttatttttct gtattaaactc cactatctaa tctccataaa 24180

actcctaaat tggattatc ggtaacattt tgtttttact caacccttag gaacaatgtt	24240
aagttaatca gccctccaca tcacagatcc ttattttcat cagtctgtac aaggcatttc	24300
tctcatttta attttttttc ctectgtcat ccttggtttt cactttcact gccctccttc	24360
cacccatatg cctcatacta atatattcga aatatacatg tcttaaaggt acatgcacgc	24420
acctacaaaa cctatagtgt ttttttgtat gtatatgtct ttaatttaaa taagtagcat	24480
tgtgtaaaag tctaataatg tttcttactg ttttactca attcttggaa ttttcatctg	24540
atgcactgct gcatagcacc ccatggtagt cagccacat atttccttca tccaattagg	24600
ttgcatgacc taccttccca ttgccacaaa gactacacac aaaatatttg tacttatctt	24660
tctgtaaacc ttcaggaatt tcagaagcac acatgcaggc tgctaaatat accagaatac	24720
tttcagcca cttaaatctt taccagtatt gcaaaagagg ccccatctcc ctccacatca	24780
acatttagta ttattctttt gtttaagttt tatcaatctt ttaaatgtac acaagatgct	24840
catttttata attttaattt ctacagattac tagtttgagt atcttttcat atatctaaga	24900
gctgttttga tctcccctac catgaactgc cactaatatt ctttgccat tttacaatgg	24960
ttttctgct tattttattac tggtttacag acttttaaaa tatattctac aaaaatttta	25020
gacattaaac attaccaata ttttcccatg gttcctcatc catctggtaa acttgtctat	25080
ggtatatcta attttgattt aatagaattc attctatttt taccttttag tttgtgtttt	25140
tgttgttttag ccaaaaagtc cccattccca ggctataaag gtaatgtcct ttttttttt	25200
ttaacgtac tgttctctct ctgtctcccc ctatgtatat aggtgcacat atactgtac	25260
acacatacat atacctatat atgaggggag ttcgataagt ttatggaaaa taaaattaaa	25320
agataaaata aaaaattata aactttattt ctcaacataa gctccttcaa gttcaagaca	25380
cttttgtaag caataatacc agccatatcg tccatcccta aagaactgag ggtcctgaga	25440
atttaactat gtcaatgcag tcttttttac attacttttt tacagtactt attgatgaaa	25500
aatgggtgcc ttttaaagat tgttttaaga ttagggaaca aaaataagtc agaggaagtc	25560
aaatcaggac tgaaggtgg atgcctagtg atttattgct gaaactttca taaaactaac	25620
cttatttgat gagaggaatg agcatgagca tggtttgat ggagaagaac tctggaggag	25680
ctttcctgga cactttttct actaaagctt tggttaactt tcttactctc ataagaagaa	25740
gatgttattt ttactgacc ctttagaagg tcaacaagca aaatgccttc agcatcccaa	25800
atgtctgttg tcatgacttt tgttcttgac tagtctgggt ttgctttgac tggaccactt	25860
ctacctcttt atagccattg ctttgatggt gctttgtctt caagattgta ttagtaaagc	25920

catatttcat ctctgttac aattcttcaa agaaatactt cagaatcttg atctgacatg	25980
tttaaaattt ctattggaag ctctgacctt ggggtgcagct gatctgggcg aaacagtttt	26040
ggcatccatc aagtagaaaag ttgtctcaac tttagttttt cagtcagaat tgtataagct	26100
gaaccagttg agatgtctat ggtgttgtct attgtttctc acagttaatt gttggctctc	26160
tttgagacat gaacaagatg aaatttttcc tagcaaactg atgtggatga tctgttgctg	26220
cgggcttcac cctcaacaac atctctttct ttcttgaaac aaattatcca ttagtaaaact	26280
gatgattggg ggagatgctg tccccataaa ctttttgtaa ggcataaata atttcacat	26340
tcttccagtt tcaccataaa ttgacgttt ttttgcttca attttagcag cattcatgtt	26400
gctttgataa gagctctttt caaattcatg tcttattcct cttagtgcct caaactagat	26460
cttgttcagt atgacaagtt agtatgagtt tatctgcatg caaaaatctt tgaaatccat	26520
gcatagtttg ttataatat acattttcaa tgaacttttg aagaccccat acatacatat	26580
gtatatatat gcacacacac acacacacac acacaaaaat cttcaacat tatcagactt	26640
agtcagaaa aattattcat ccattaacaa gataagaatg ccccttatca tcactactat	26700
ttaaatggag ctcttgcta aaggaaaaga cagggatga aaaaaattag ttaaatctaa	26760
aatgtttatt atttcaggtt tcttagttgc ttaaatggga agggaggtat ggacaaaaga	26820
gaaatcaaag atatttgtgt tatgtctactt atcattaaag tadcagaata acttcattgg	26880
aatagaaaaa caccaagatc accccacgat atgttttcta aaatcttctc catttcttta	26940
gacaagtgc catgtattcg gccagtgaag aattaaactc acttgccagc ttataatgca	27000
ggaaaatata gcaaagagat gtggatccaa tagtttctag atagtgttac aggatggcta	27060
agatgaattt atatatctga aatgttcaca aattccctac tcatatagca tgttttcata	27120
atgttttagc aactctaate ctctgactg gattgccacg tctggtatth ccacaacatt	27180
tcctaaacta agaatgagag taagaatat tttaattcat aacaattata aatctgcaac	27240
tcatgaaaat gacattgcac ttgtgagact tgagaacagt gtcaccttta ccaaagatat	27300
ccatagtgtg tgtctcccg ctgctacca gaatatcca cctggctcta ctgcttatgt	27360
aacaggatgg ggcgctcaag aatatgctgg taagtgtctc ggaaaaaaaa attaacaata	27420
gaaatgtctt atatttgcta ttaggtaatt tttaaatga ggaaacatct ggaataggtg	27480
tttctattct tctacagaca gaaccattct atattctgct cagcccaagc tctggctacc	27540
cctgagtctc cttagcaaaag caaagcaatg ctccagaaac tatgggaatt ctcaaatata	27600
gtaataggaa aatgtaaaag aaagtatga agacacgagt tctttaataa tccagagatt	27660
ctataagatt caaatagctt ccctataaac aataaaaaag attttgtttg ttgtttgtt	27720
tgcttgtttt ttagagacaa agactttctc agactggagt gcagtgggtc aatcatggct	27780

tactgcagcc tcaaactctg gtcttaagaa atcctcttgc ttcagcctcc caagtagcta	27840
gaattataaa taagtgtgta ccaccatacc cagctttttt ttttttttc tacagacagg	27900
ttcttgcctt gttgccagg ctggctgga attcctgccc tcaagccatc ctctgcctt	27960
gttggcctcc caaagcaatg ggaggattta gattagacat tgtatgaggg cttataatc	28020
cttaaggtat taactgccct ttaaagtatt ctgggatatg gcaaaaactc gatgtgtata	28080
taaacattgg tcatatttgt ttattgaatg aataaaatgg aaactaaat gaggacaatg	28140
cacaagagct actagaacca gtaagagtat cagcgaagga gtggaagggt agcattgaca	28200
atttccctgg gcttttacc atgtttaga ttgtctctcc aaggaataat acaaagcctt	28260
aatagtccta gaacacattc tatttgtttc ttatggcca aagtaaattg gtgtagtaga	28320
taacatttgc accagtcatg aaaaactatt ggtgtcattc tgagagtaca tcaatataaa	28380
atagactagt tctttagcct tgaaactaga ctggtttctc ttttgctgct aggttaaagg	28440
ttattcaata tgtaatcttc caatccaaaa tctgtcagtg gataatttaa aagcttttag	28500
tcaattttaa gatatttgtt ttcttaaaat ttaaggggc actgtgtcac aaagctaaag	28560
aaaaaaaaa aaaaaaaact gatctgtgaa aggggttctc ctcatctact tggggaattt	28620
tggctgcgaa gaaactccaa agtaaatctt tagaagcctt cattgttaaa tatgaaataa	28680
tgtttggagt acattttatt ctctcaaat ttattatagg gtcaataatg tacacatctt	28740
gaagtccatt ttttctcgc ttttataaca aacaggccac acagttccag agctaaggca	28800
aggacaggtc agaataataa gtaatgatgt atgtaatgca ccacatagtt ataatggagc	28860
catcttgtct ggaatgctgt gtgctggagt acctcaaggt ggagtggagc catgtcaggt	28920
aagctcaaga caatctcatc catgtcatca tccaagaagt gtataagcac ttcctagtat	28980
gtgataatgt gatagacata agtgtaacag ttacaataca cagccctgtt cctctaaaat	29040
ttataatcta gattttagaa ataaattttt ttatgaatga agtttatcta tcatgaaagc	29100
attaactctg agaggccaaa ttacagagta gttaccatc caaagctcaa gaatcagaaa	29160
gacctcgatt tgaattcctt aacctctatt accaagtcct taactaaaag ctggggataa	29220
tcataatagc acctaaactt ttgggtacta agaaaagtta aatgaagact aaatatatca	29280
ggcacatggt aaacaacaaa gaaatctcat ctatttcact attattaatg tagaccatgg	29340
tcactcgtgt taataacttt aacctcaacc ttttaactgc tgtgaaggat taaataaaaa	29400
attaatcact atattataaa aattaattga tatataataa atgaatttta agagatcgt	29460
aataattcat ggactccttg aagatagaaa atttatcaa aatcctagta atttgagtca	29520
caaaagctcc tacaataatg aaacagtatg aatgaaaaag aaaagaaata actattatat	29580
ttggatctag ccataattt ttaaccaaat gcacaaaaac aaacaacaaa tatgaaattc	29640

tcactgtaaa gtgattaaaa tcaaatttga attctaaaaat tttaaattaa attatctaaa	29700
cataattgat gcagttatat gttttaatag gttttgttca cataatctgaa atccaactcc	29760
acacagtagc aggaacagct ggtgtcagaa attaaatatt cttttagtct ggagttttaa	29820
aaaatcaatc tgtttacttg agtaatttgi tgctgttttc atgggtgaat tgtatacaga	29880
aggataagaa ttattcttcg catcaaaagg tcaactgactt tcatatttag tgctcatggt	29940
ctttaaaaag tggataaaaa gtagttctca catttcatgg aaagcccca atccatgagc	30000
acatttccca aaatgaaaca tttttatcaa ctgcaagttg tgtglagggtg gagatttggt	30060
tttcaattgt caagatactg ttaattaccc agtcctttat ctcttttgg tggagatgtc	30120
tctgtgctag gaaacccttc ttgctctcct tectgtttct cttttactac tggccctgaa	30180
acaacaaatt ctcaagtttc atgacagctt tccaaagaat ccatcaatca aataagcaac	30240
acaactgcac actgacaatt ccagacctac taagagcatt aattaagact taaaaataaa	30300
catgagtttt aaaagggtgt tattcattat ttcccattht ataacgtccc ttaccttctg	30360
tccttcagtg catacaaat attatcttcc ttgaagccca gttcaagccg tacctcacca	30420
tgatacttc catgtatatt ccactctagg cctcactgat ttttaactga aatactataa	30480
tgcatagttc acacttaaaa aaaaaaaaaa aacacagcac ttacataag agcttacagg	30540
atcctatttg ttttatccat tctttgttc atttttaca tcattaattc aaaggaatta	30600
tattaattac ttictatgca cccgacgttg tgtaacaca acaatactat ccctgcattc	30660
agcaagtcta tggctctaca gagaggacac aaattcaa atgtctgtagtc aagcagtga	30720
gctggctaga tatggaaaa ttacaagtcc ctcttgcttt aacatttget tgcccacatt	30780
tggtcagaca tcatgcaaaa taatttctca ctatagaaaa aaaaacacta caaaaacaat	30840
aataataaga actgagaact ggttaactga agcatgcata tgtcatctaa aagaagcagg	30900
tgacgaccag ctcatgaag tacttgccat gcataattggc acttcacaca ctgacccttc	30960
tccccaccta gaccagtaat taaacaggta tggatgagct agctactaag agcagccaac	31020
tgaatagctg actaacttag aagcacactt ggtaataata gctgactttt attagtactg	31080
actatactat atgctaagct gtactcaaag tgctttgagt tttaaactga taaaacatt	31140
atataggaa acagaggtag agagagctat tcaccagctt accaaaggtc acatagctgg	31200
taagtggagg acttaaaccc agactatcta gtttcagaac ccacagactt aatccatcgt	31260
gcagaacata agacatactc catctgtctc cccaactagg ttattatgtg cacaatatt	31320
tattggttgg ttggttcatt attatgactg ggtggtaagt atgtcattag gagtgttttg	31380

cttatgacta tataaatttc ttcaccaaaa gaagactttc tgatgatata ctatgcatca	31440
gacaccacgc aggggtgctaa ggtaggaag ataagtgaga cttctagaaa ctcattcatt	31500
caacaaatat ctctaaggg ctagaagctt aggtttcagc agtgaacaga ataggtatgt	31560
tctctttcgt gttagacctt atagtatac tgggaaaaca gacattgaat aaatatcaca	31620
aatgcaagtg agtgtttcag agacatgcag ctgctacac aaaacaaaac agaacaaaac	31680
aaacaaacaa aaactgacca gtgggattaa gtgtaaatag gcacacaaat gcacaaatat	31740
gcttttataa aatagtgaag cagtgcacaga gacacacaca agatataaag acacaatgaa	31800
gaacaattga gcccaaagct ggaaaggggtg agagtgtgaa ggaaaaaggt tgatcagaga	31860
agttttcccg aaggagagaa agcctggatg attaggaggc aaccactcgg tgactgaggg	31920
aaatctgaaa aatgtatttg tcatcttctc agacttgctg aaggaaatgac ttgggtactt	31980
tgaggatttc agtaattttt ccatgacttg gtataatatt tcaaaaggaa ataggctgac	32040
tttatttgta taatgaatgt gactccttcc tcgactgcca tagaaataaa ctccttaata	32100
ttttgggttt gtctttgcac ttaagtaac agtcattctg tttttttaca ggggtgactct	32160
ggtagggccac tagtacaaga agactcacgg cggctttgggt ttattgtggg gatagtaagc	32220
tgaggagatc agtgtggcct gccggataag ccaggagtgt ataactcgagt gacagcctac	32280
cttgactgga ttaggcaaca aactgggatc tagtgcaaca agtgcacccc tgttgcaaag	32340
tctgtatgca ggtgtgcctg tcttaaattc caaagcttta catttcaact gaaaaagaaa	32400
ctagaaatgt cctaatttaa catcttgta cataaatatg gtttaacaaa cactgtttaa	32460
cctttcttta ttattaaagg ttttctattt tctccagaga actatatgaa tgttgcatag	32520
tactgtggct gtgtaacaga agaaacacac taaactaatt acaaagttaa caatttcatt	32580
acagttgtgc taaatgcccc tagtgagaag aacaggaacc ttgagcatgt atagtagagg	32640
aacctgcaca ggtctgatgg gtcagagggg tcttctctgg gtttactga ggatgagaag	32700
taagcaaact gtggaaacat gcaaaggaaa aagtgataga ataatttca agacaaaaag	32760
aacagtatga ggcaagagaa ataatatgta tttaaaattt ttgggttactc aatatcttat	32820
acttagtatg agtcctaaaa ttaaaaatgt gaaactgttg tactatacgt ataacctaac	32880
cttaattatt ctglaagaac atgcttccat aggaaatagt ggataatttt cagctattta	32940
aggcaaaagc taaaatagtt cactcctcaa ctgagaccca aagaattata gatatttttc	33000
atgatgaccc atgaaaaata tcactcatct acataaagga gagactatat ctattttata	33060
gagaagctaa gaaatatacc tacacaaact tgtcaggtgc tttacaacta catagtactt	33120
tttaacaaca aaataataat tttaagaatg aaaaatttaa tcatcgggaa gaacgtccca	33180
ctacagactt cctatcactg gcagttatat ttttgagcgt aaaaggtcg tcaaacgcta	33240

aatctaagta acgaattgaa agtttaaaga gggggaagag ttggtttgca aaggaaaagt 33300
 ttaaatagct taatatcaat agaatgatcc tgaagacaga aaaaactttg tcactcttcc 33360
 tctctcatTT tctttctctc tctctcccct tctcatacac atgcctcccc caccaaagaa 33420
 tataatgtaa attaaatcca ctaaaatgta atggcatgaa aatctctgta gtctgaatca 33480

ctaataattcc tgagttttta tgagctccta gtacagctaa agtttgccta tgcattgatca 33540
 tctatgcgtc agagcttccct cttctacaa gctaactccc tgcattctggg catcaggact 33600
 gctccataca ttigtgaaa acttcttgta tttctgatg taaaattgtg caaacaccta 33660
 caataaagcc atctactttt agggaaaggg agttgaaaat gcaaccaact ctggcggaac 33720
 tgtacaaaca aatctttgct atacittatt tcaaataaat tctttttaaa ataatttccc 33780
 tgcctaatta tttatggaag ttatgacttt tgaaggacaa ttcaaaacca tttatttaat 33840
 tggttctgca atgaaagaac tgccccatat actctactaa aggcttggca ctttctgctg 33900

ccttttaatc cagcgtata attgaggcaa gcgtccagct tgacacctcg agataacttc 33960
 gtataatgta tgctatacga agttatgcta gtaactataa cggtcctaag gtagcgagct 34020
 agctgcaacc gaggaaaaaa cgtgccatga ggtctctgta tccaagtgtg act 34073

<210> 21

<211> 418

<212> PRT

<213> Artificial Sequence

<220><223> Recombinant protein

<400> 21

Met Tyr Arg Pro Arg Pro Met Leu Ser Pro Ser Arg Phe Phe Thr Pro

1 5 10 15

Phe Ala Val Ala Phe Val Val Ile Ile Thr Val Gly Leu Leu Ala Met

20 25 30

Met Ala Gly Leu Leu Ile His Phe Leu Ala Phe Asp Gln Lys Ser Tyr

35 40 45

Phe Tyr Arg Ser Ser Phe Gln Leu Leu Asn Val Glu Tyr Asn Ser Gln

50 55 60

Leu Asn Ser Pro Ala Thr Gln Glu Tyr Arg Thr Leu Ser Gly Arg Ile

65 70 75 80

Glu Ser Leu Ile Thr Lys Thr Phe Lys Glu Ser Asn Leu Arg Asn Gln

85	90	95	
Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val			
100	105	110	
Arg Ala Asp Val Val Met Lys Phe Gln Phe Thr Arg Asn Asn Asn Gly			
115	120	125	
Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn			
130	135	140	
Asn Ser Gly Asn Leu Glu Ile Asn Pro Ser Thr Glu Ile Thr Ser Leu			
145	150	155	160
Thr Asp Gln Ala Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly			
165	170	175	
Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu			
180	185	190	
Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn			
195	200	205	
Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr			
210	215	220	
Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala			
225	230	235	240
Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg			
245	250	255	
Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp			
260	265	270	
Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile			
275	280	285	
His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser			
290	295	300	
Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr			
305	310	315	320
Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val			
325	330	335	

Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu

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

405 410 415
Gly Ile

<210> 22

<211> 257

<212> DNA

<213> Artificial Sequence

<220><223> synthetic oligonucleotide

<400> 22

agcaccctc tcttcgcag agtctaagaa atcgtgtgt ttagccctcg cctgggcac 60
tgtctcagc ggagctgctg tggctgctgt ctgtcttgg aagttcagta agtcaggga 120
gcctgatcc caccatgtgc tctgcagtc ccagtgctc tgagccagac cctgtctct 180
gggctattga gacctctgga ggccctccgt gaggttctc tttacataa cgagctgtc 240

tctcttcct tctcttg 257

<210> 23

<211> 190

<212> DNA

<213> Artificial Sequence

<220><223> synthetic oligonucleotide

<400> 23

ggtcagagga ccaaaggtga ggcaaggcca gacttggtgc tctgtggtt ctcagataa 60
cttcgtataa tgtatgctat acgaagttat atgcatggcc tccgcgccg gttttggcgc 120
ctcccgccgg cgccccctc ctcacggcga gcgtgccac gtcagacgaa gggcgcagcg 180

agcgtcctga 190

<210> 24

<211> 171

<212> DNA

<213> Artificial Sequence

<220><223> synthetic oligonucleotide

<400> 24

attgttttgc caagttctaa ttccatcaga cctcgacctg cagcccctag ataacttcgt 60

ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt agcgagctag 120

ctccacgtgg ctttgtccca gacttccttt gtcttcaaca accttctgca a 171

<210> 25

<211> 177

<212> DNA

<213> Artificial Sequence

<220><223> synthetic oligonucleotide

<400> 25

ggtcagagga ccaaaggatga ggcaaggcca gacttggtgc tctgtgggtt ctcgagataa 60

cttcgtataa tgtatgctat acgaagttat gctagtaact ataacgggcc taaggtagcg 120

agctagctcc acgtggcttt gtcccagact tcctttgtct tcaacaacct tctgcaa 177

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> synthetic oligonucleotide

<400> 26

gccgtgactg tgaccttctc 20

<210> 27

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> synthetic oligonucleotide

<400> 27

tggaggagcc acctgatgcc tc	22
<210> 28	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 28	
gccttgccct caatggaaac	20
<210> 29	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 29	
ggttgcacag caaggaagaa g	21
<210> 30	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 30	
ccaggagttc ctgtgagcct accc	24
<210> 31	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 31	
tggaatggaa ggagctggag	20
<210> 32	
<211> 19	
<212> DNA	
<213> Artificial Sequence	

<220><223> synthetic oligonucleotide
 <400> 32
 gtccacctc ctgcaactg 19
 <210> 33
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 33
 tgagccttc catcagcctg gg 22
 <210> 34
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 34
 ccacaatggc acatgggtct g 21
 <210> 35
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 35
 ggtgcttgct cccaaga 18
 <210> 36
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 36
 cctaaaaggt gttgtaatgg 20
 <210> 37
 <211> 24

<212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 37
 ggcaataaag aaggaagacg tttt 24
 <210> 38
 <211> 120
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 38
 ccagtcaggg acacacatgc tcacacgccc gccacccgc acacactaca gtcgagataa 60
 cttcgtataa tgtatgctat acgaagttat atgcatggcc tccgcgccgg gttttggcgc 120
 <210> 39
 <211> 198
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 39
 attctagttg tggtttgtec aaactcatca atgtatctta tcatgtctgg aataacttcg 60
 tataatgtat gctatacgaa gttatgctag taactataac ggtcctaagg tagcgagcta 120
 gccaaagtctg tgtgctacca agtagcaaaa ctgagcctgg aactcacaca tgcgtgtctg 180
 agagcccagc actatcgc 198
 <210> 40
 <211> 100
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 40
 taatctgact ttctcttcat cggctctctt tattctaggc tgagctgtaa cgctgccgtc 60
 cccacatcc agaagctgct tcccttcaga cctacctacg 100
 <210> 41
 <211> 177

<212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 41
 ccagtcaggg acacacatgc tcacagccc gccaccgc acacactaca gtcgagataa 60
 ctctgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggtagcg 120
 agctagccaa gtctgtgtgc taccaagtag caaaactgag cctggaactc acacatg 177
 <210> 42
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 42
 gagcagggcc atgacacat 19
 <210> 43
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 43
 accattagat cccagcactg gaca 24
 <210> 44
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 44
 aaacccttcc cgagagagaa 20
 <210> 45
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide

<400> 45	
gaggaacact gtgtcaagga ctt	23
<210> 46	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 46	
cctgaaaagc ccggagtggc ag	22
<210> 47	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 47	
gggcagagac cacatctga	19
<210> 48	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 48	
ggaagccctc tctcgatact tg	22
<210> 49	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> synthetic oligonucleotide	
<400> 49	
ttctaccctg agggcatgca gc	22
<210> 50	
<211> 22	
<212> DNA	
<213> Artificial Sequence	

<220><223> synthetic oligonucleotide
 <400> 50
 tgggatgtag aaggttgtca ga 22

<210> 51
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 51
 ctgagcctgg aactcacaca tg 22

<210> 52
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 52
 tctgagagcc cagcactatc gcc 23

<210> 53
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 53
 gctgagggtc aggcttgag 19

<210> 54
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 54
 tctgcagggt agggagagaa g 21

<210> 55
 <211> 29

<212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 55
 tgtttcagaa aaggaagact cacgttaca 29
 <210> 56
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 56
 gagaccgatg aagagaaagt caga 24
 <210> 57
 <211> 100
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 57
 gaccatttta aggttttgct tggttgtttt ggaggaggagg tgggtgctttg ctaatggtga 60
 attactaact cctcaataaa gaatattatt tgaataaatt 100
 <210> 58
 <211> 190
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 58
 gctgcctttt aatccagcgc tataattgag gcaagcgtcc agcttgacac ctcgagataa 60
 cttegtataa tgtatgctat acgaagtat atgcatggcc tccgcgccgg gttttggcgc 120
 ctcccgcggg cgccccctc ctcacggcga gcgctgccac gtcagacgaa gggcgcagcg 180
 agcgtcctga 190
 <210> 59
 <211> 171
 <212> DNA

<213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 59
 attgttttgc caagtttctaa ttccatcaga cctcgacctg cagcccctag ataacttcgt 60
 ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt agcgagctag 120
 ctgcaaccga ggaaaaaacg tgccatgagg tctctgtatc caagtgtgac t 171
 <210> 60

<211> 177
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 60
 ccagtcaggg acacacatgc tcacacgccc gccaccgcg acacactaca ctcgagataa 60
 ctctgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggtagcg 120
 agctagctgc aaccgaggaa aaaacgtgcc atgaggcttc tgtatccaag tgtgact 177
 <210> 61

<211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 61
 tcctctccag acaagaaagc t 21

<210> 62
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 62
 tcatagcagc tticaaatcc taaacgttga 30
 <210> 63
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> synthetic oligonucleotide
 <400> 63
 tcgtgtgttag ctggtgagtt 20
 <210> 64
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 64
 catgcgatca caggaggaga tc 22
 <210> 65
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 65
 aattggggccc gaagccagat gc 22
 <210> 66
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 66
 cggaaggctt ctgtgacttc 20
 <210> 67
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 67
 gtctcccact tctgacataa tgaac 25
 <210> 68
 <211> 27

<212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 68
 cccagtgtta accctacatc tggttcc 27
 <210> 69
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 69
 tgggaagaga ctcttgaca 20
 <210> 70
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 70
 atgagctcct agtacagcta aagtt 25

 <210> 71
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 71
 atgcatgata atctatgcgt cagagc 26
 <210> 72
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> synthetic oligonucleotide
 <400> 72
 tgcccagatg caggagatta g 21