(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.^{7}
C12N 15/12

(11) 공개번호 특2003-0074643
(43) 공개일자 2003년09월19일

---

(21) 출원번호 10-2003-7007511
(22) 출원일자 2003년06월05일
번역문 제출일자 2003년06월05일
(86) 국제출원번호 PCT/IB2001/02798
(86) 국제출원출원일자 2001년12월04일
(87) 국제공개번호 WO 2002/46404
(87) 국제공개일자 2002년06월13일

---

(81) 지정국 국내특허 : 아랍에미리트, 안티구아바부다, 알바니아, 아르메니아, 오스트리아, 오스트레일리아, 아제르바이잔, 보스니아-헤르체고비나, 바베이도스, 불가리아, 브라질, 벨라루스, 벨리즈, 캐나다, 스위스, 중국, 코스타리카, 쿠바, 체코, 독일, 덴마크, 도미니카연방, 알제리, 에스토니아, 스페인, 핀랜드, 영국, 그루지야, 헝가리, 이스라엘, 아이슬란드, 일본, 케냐, 키르기즈, 북한, 대한민국, 카자흐스탄, 세인트루시아, 스리랑카, 라이베리아, 레소토, 리투아니아, 룩셈부르크, 라트비아, 모로코, 몰도바, 마다가스카르, 마케도니아, 몽고, 말라위, 멕시코, 모잠비크, 노르웨이, 뉴질랜드, 폴란드, 포르투칼, 루마니아, 러시아, 수단, 스웨덴, 싱가포르, 슬로베니아, 슬로바키아, 타지키스탄, 투르크메니스탄, 터어키, 트리니다드토바고, 탄자니아, 우크라이나, 우간다, 미국, 우즈베키스탄, 베트남, 남아프리카,

AP ARIPO특허: 가나, 감비아, 케냐, 레소토, 말라위, 수단, 시에라리온, 스와질랜드, 우간다, 짐바브웨,

EA 유라시아특허: 아르메니아, 아제르바이잔, 벨라루스, 키르기즈, 카자흐스탄, 몰도바, 러시아, 타지키스탄, 투르크메니스탄,

EP 유럽특허: 오스트리아, 벨기에, 스위스, 사이프러스, 독일, 덴마크, 스페인, 핀랜드, 프랑스, 영국, 그리스, 아일랜드, 이탈리아, 룩셈부르크, 모나코, 네덜란드, 포르투칼, 스웨덴, 터어키,

OA OAPI특허: 부르키나파소, 베넹, 중앙아프리카, 콩고, 코트디브와르, 카메룬, 가봉, 기네, 기네비쏘, 말리, 모리타니, 니제르, 세네갈, 차드, 토고,

(30) 우선권주장 60/251,317 2000년12월05일 미국(US)

(71) 출원인 장세
프랑스 파리 75008 뤼 로얄 24

(72) 발명자 코헨,다니엘
프랑스,누에일리-수르-세인에프-92200,보울레바드리차드왈라스1

추마코브,일리아
프랑스,바욱스-레-페닐에프-77000,루에데스체브레페우일레스196

시몬,안네-마리아
프랑스,파리에프-75012,루에데레우일리,41아파트먼트86

아브데라히,하우디
프랑스,차렌톤레폰트에프-94220,루에데파리스145

(74) 대리인 강명구

심사청구 : 없음

(54) 정신분열증-관련된 전압-개폐 이온 채널 유전자와 단백질

요약

본원 발명은 전압-개폐 이온 채널 단백질인 CanIon의 게놈 DNA, cDNA, 폴리펩티드 서열에 관한다. 본원 발명은 또한, CanIon 유전자의 이중대립인자 마커에 관한다. CanIon 유전자는 정신분열증, 양극성 장애, 다른 질환과 이상의 치료 및 진단을 위한 생물학적 표적으로 이용될 수 있다.

대표도

도 1

명세서

기술분야

본원 발명은 전압-개폐 이온 채널 유전자와 단백질 및 질환에서 이의 역할에 관한다. 본원 발명은 CanIon 폴리펩티드를 인코드하는 폴리뉴클레오티드 및 상기 코딩 영역의 5'-와 3'-말단에 위치하는 조절 영역에 관한다. 본원 발명은 또한, CanIon 채널 조절물질(예, 길항물질)의 스크리닝 방법 및 다양한 질환이나 이상의 치료 또는 예방에 이런 조절물질을 사용하는 방법에 관한다. 본원 발명은 또한, 진단 시약으로 유용한이런 폴리펩티드를 특이적으로 지향하는 항체에 관한다. 본원 발명은 또한, 유전자 분석에 유용한 CanIon 유전자의 이중대립인자 마커에 관한다.

배경기술

임상 연자에게 가용한 기술 설비의 발전으로 건강 상태와 질환 상태에서 뇌와 신경계 기능의 정교한 연구가 가능하게 되었다. 정신 질환과 신경퇴행성 질환을 비롯한 중추신경계(CNS) 질환에 관여하는 신경화학적 · 유전학적 기전의 측면에서 많은 신경학적 · 약리학적 가설이 제시되고 있다. 하지만, CNS 질환은 복합적이고 이해하기 어려운 병인 및 중복되고 거의 특성화되지 않으며 측정하기 어려운 증상을 보인다. 결과적으로, 앞으로의 치료 섭생과 약물 개발 노력은 좀더 정교하고 다중유전적 원인에 좀더 집중되어야 하고, 질환 개체군을 구분하는 신규한 분석법을 필요로 하며, CNS 질환 환자에 대한 좀더 정확한 진단과 예후 정보를 제공해야 한다.

CNS 질환은 광범위한 질환 및 상응하는 광범위한 유전적 요인을 동반할 수 있다. CNS 질환의 예에는 신경퇴행성 질환, 정신 질환, 기분 장애, 자폐증, 약물 의존과 알코올 중독, 정신 지체 및 인식 장애, 정신 장애, 불안 장애, 섭식 장애, 충동 조절 장애, 인격 장애를 비롯한 다른 정신 질환이 포함된다. 질환은 Diagnosis Statistical Manual of Mental Disorder fourth edition(DSM-IV) 분류에 따라 정의할 수 있다.

CNS 질환의 소규모 부분집합을 고려하는 경우에도 정신 분열증 및 양극성 장애에 대한 적절한 치료 및 분자적 기초의 이해 부족으로 인해 치료 발명의 새로운 표적 및 개선된 치료 방법이 필요하다는 것은 명백하다. 정신분열증과 양극성 장애의 치료에 사용되는 현재 공지된 모든 분자는 부작용을 보이고 효과가 이들의 증상에만 한정된다. 관련된 부작용이 없으면서 정신분열증과 양극성 장애의 원인 기전에 관여하는 표적을 저해하는 신규한 분자가 절실히 필요하다. 따라서, 이들 표적의 발견과 특성화를 용이하게 하는 도구는 유용하다.

전압-개폐 이온 채널

전압-개폐 이온 채널은 세포막에서 전위의 조절과 유지, 분비, 신호 전달을 비롯한 여러 기능을 수행하는 거대분자의 대과이다. 이들 채널 단백질은 뉴우런으로부터 신경전달물질의 조절에 관여하고 막 흥분성, 근육 수축, 시냅스 전달을 비 롯한 다양한 세포 기능의 조절에 중요한 역할을 수행한다. Na+ 채널의 주요 알파-아단위와 Ca+ 채널의 알파-1 아단위는 대략 2000개의 아미노산으로 구성되고 이온 전도 경로를 보유한다. 생화학적 분석에서 생리활성 이온 채널이 여러 개의 상이한 아단위로 구성되는 것으로 밝혀졌다. Na+ 채널의 알파 아단위를 보조하는 하는 2개의 보조 아단위, 베타-1과 베타-2 아단위가 존재한다. Ca+ 채널에서 추가의 아단위(알파-2, 베타, 감마, 시그마)는 조절 작용을 갖는 것으로 확인되었다. 알파-2와 베타-아단위는 Ca+ 알파-1 아단위의 기능적 활성을 강화시키는 것으로 보인다. K+ 채널의 알파-아단위는 1:1 방식으로 베타 아단위와 결합하여 $(알파)_4(베타)_4$ 입체화학적 성상을 보이는 K+

채널 복합체를 유도한다(Terlau et al., Naturwissenschaften 85: 437-444 (1998)). 칼슘과 나트륨 이온 채널의 기본 구조와 예는 예로써 Williams, et al. Science 257: 389-395 (1992); Mori, et al., Nature 350: 398-402 (1991) ; Koch, et al., J.Biol.Chem. 265 (29): 17786-17791 (1990)에서 기술한다. CanIon 채널과 높은 수준의 서열 상동성을 공유하고 쥐로부터 유래된 Ca+와 Na+ 이온 채널 핵산 서열은 Lee et al., FEBS Lett. 445 231:236 (1999)에서 기술한다.

알파 아단위는 모든 전압-의존 양이온 채널과 서열 특징을 공유하고 전위 막 스패닝(spanning) 도메인의 6개-나선 번들을 포함하는 동일한 구조 모티프를 이용한다. 나트륨과 칼슘 양 채널에서, 상기 모티프는 서열내에서 4회 반복되고 24개-나선 번들을 제공한다. 아미노산 서열은 종(예, 사람과 초파리)간에 그리고 상이 한 이온 채널간에 고도로 보존된다.

다중-유전자 군에서 개별적인 유전자에 의해 코드되는 칼슘 채널의 조직-특이적인 약리학적 · 전기생리적으로 구별되는 여러 동소체가 존재한다. 골격근에서, 알파, 베타, 감마 아단위의 단단하게 결합된 각 조립체는 4개의 다른 조립체와 결합하여 오합체의 거대분자를 형성한다. 가령, 신경 칼슘 채널 알파-1 아단위는 적어도 7개의 상이한 유전자(알파 1 A에서 H까지)의 산물이다. 면역세포화학적 연구에서는 알파-1 칼슘 채널 아단위의 차별적인 분포를 보였다. 알파-1A와 알파-1B는 주로 수상돌기와 전시냅스 말단에서 발현되는데, 알파-1A가 알파-1B보다 좀더 많은 수의 신경 말단에 농축된다. 쥐와 사람 신경근 연접(neuromuscular junction)에서, 전시냅스에는 알파-1A가 집중적으로 위치하는 반면, 축삭-결합된 슈반세포(Schwann cell)에는 알파-1B와 알파-1A가 모두 존재한다. 알파-1E는 세포체에 집중적으로 위치하고, 일부 경우에 근접 수지상 돌기 및 푸르키니에세포(Purkinje cell)의 말단 가지에 위치한다. 알파-1C와 알파-1D는 세포체 및 중심성 뉴우런의 근접 수지상 돌기에 집중적으로 위치한다.

고유 칼슘 채널은 약리학적 및/또는 전기생리학적 특성에 기초하여 분류되었다. 전압 의존한 칼슘 채널은 다음과 같이 구분된다: L-, N-, P-, Q-형을 비롯한 고전압-활성화 채널(HVA); 중간 채널(IVA, R-형); 저전압-활성화 채널(LVA, T-형)(Morena et al., Annals NY Acad. Sci. 102-117(1999)).

주요 아단위(알파-1)는 이질성 발현계에서 발현되는 경우에 막이 자체적으로 기능적 채널을 형성할 수 있는 유전자에 속한다(Zhang et al., Neuropharmacology 32(11): 1075-1088(1993). 고유 세포에서, 알파-1 아단위는 알파-1 단위의 기능적 특성을 변화시키는 보조 아단위를 보유하는 다중아단위 복합체로 발현된다. 많은 경우에, 보조 아단위의 공동발현은 채널의 생물물리학적 특성에 영향을 준다. 특히, 베타 아단위는 알파-1 아단위에 상당한 영향을 준다; 베타 아단위는 활성화 특성, 항정상태 비활성화, 비활성화 동력학, 피크 전류를 변화시키는 것으로 밝혀졌다.

채널의 분자 다양성중 대부분은 복합적 형태의 알파-1 아단위의 존재에 의해 발생한다. 가령, 상이하게 절단접합된 형태의 칼슘 채널은 차별적으로 발현되고 세린-트레오닌 키나아제에 의한 인산화에 대하여 독특한 감수성을 보유한다(Hell et al., Annals N.Y. Acad. Sci. 747: 282-293(1994)). 이온 채널 유전자에서 돌연변이는 여러 중추 신경계 질환을 비롯한 광범위한 질환에 관여하는 것으로 밝혀졌다. 주기성 마비, 발작성 운동실조증, 편두통, 연장된 QT 증후군, 발작성 운동장애를 비롯한 다양한 발작성 질환을 유발하는 이온 채널 돌연변이의 예는 Bulman et al., Hum. Mol. Gen. 6(10) 1679-1685(1997)에서 자세하게 기술한다. 몇몇 Ca+ 채널 돌연변이 질환의 예는 표 A에 도시한다(Moreno, supra).

표 A

질환	칼슘 채널 아단위	모델
가족성 반측마비성 편두통	알파-1A	사람
발작성 운동실조증 2형	알파-1A	사람
척수소뇌성 운동실조증 6형	알파-1A	사람
토터링(Tottering) 및 러너(Learner) 표현형	알파-1A	생쥐
람버트-이튼 증후군	알파-1A, α-1B	사람
저칼륨성 주기성 마비	알파-1S	사람
근육 이형성	알파-1S	생쥐
Zucker 당뇨성 지방과다 표현형	α-1C; α-1D	쥐
혼수상태	베타-4	생쥐
악성 저체온증	알파-2/δ	사람
스타가제르(Stargazer)	감마	생쥐

칼슘과 나트륨 채널의 조절물질은 다양한 질환과 이상의 치료에 일반적으로 사용된다. 가령, 칼슘 및/또는 나트륨 채널 차단제는 다양한 질환이나 이상, 예를 들면 다양한 심장 질환과 이상(예, 협심증, 부정맥), 고혈압, 편두통, 발작의 신경학적 효과, 조병, 신경이완제-유도된 지발성 운동장애, 양극성 장애, 통증, 간질 등과 관련된 하나 또는 복수의 증상을 치료 또는 예방하는데 유용한 것으로 밝혀졌다.

신경계에서 현저한 기능적 차이가 상이한 이온 채널간에 존재하는 것으로 보인다. 이에 더하여, 동일한 이온 채널 유전자에서 상이한 돌연변이간에 그리고 동일한 이온 채널의 절단접합 변이체간에 기능적 차이가 존재한다. 따라서, CNS 질환에서 이온 채널의 연관에도 불구하고 어떤 이온 채널이 특정 질환에 대한 치료 개입에 효과적인 표적인 가를 예측하는 것은 어려운 일이다. 한가지 난제는 정신분열증, 양극성 장애 또는 이들과 관련된 질환에 연관된 이온 채널 유전자를 제공하는 것이다.

본원 발명의 요약

본원 발명은 CanIon이라고 하는 전압-개폐 이온 채널 단백질을 인코드하는 신규한 사람 유전자의 게놈 서열을 포함하는 핵산 분자에 관한다. 상기 CanIon 게놈 서열은 상기 유전자의 전사 영역의 상류(5' 말단)와 하류(3' 말단)에 위치하는 조절 서열을 포함하는데, 이들 조절 서열 역시 본원 발명에 속한다.

본원 발명은 또한, CanIon 단백질을 인코드하는 완전 cDNA 서열 및 상응하는 번역 산물을 제공한다.

CanIon 게놈이나 cDNA 서열과 특이적으로 혼성화되는 올리고뉴클레오티드 프로브와 프라이머 및 이런 프라이머와 프로브를 이용한 DNA 증폭과 검출 방법 역시 본원 발명에 속한다.

본원 발명의 다른 목적은 전술한 핵산 서열중 한가지를 포함하는 재조합 벡터, 특히 CanIon 조절 서열 또는 CanIon 단백질을 인코드하는 서열을 포함하는 재조합 벡터 및 상기 핵산 서열이나 재조합 벡터를 포함하는 세포 숙주와 유전자도입 사람제외 동물을 제공하는 것이다.

또한, 본원 발명은 CanIon 유전자의 이중대립인자 마커 및 이의 용도에 관한다.

최종적으로, 본원 발명은 CanIon의 발현이나 활성을 조절하는 물질이나 분자의 스크리닝 방법 및 CanIon 폴리펩티드와 상호작용하는 물질이나 분자의 스크리닝 방법에 관한다.

이들 방법으로 동정된 물질을 이용하는 방법을 또한 제공한다. 가령, CanIon 채널 길항물질을 이용하여 정신분열증이나 양극성 장애를 비롯한 질환이나 이상을 치료 또는 예방하는 방법을 제공한다.

따라서, 한 측면에서 본원 발명은 SEQ ID NO 1 내지 4 또는 6으로 도시된 뉴클레오티드 서열중 하나, 또는 이들 서열에 상보적인 서열을 포함하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드를 제공한다.

다른 측면에서, 본원 발명은 SEQ ID NO 4의 적어도 50개 뉴클레오티드의 연속 스팬을 포함하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드를 제공하는데, 상기 폴리뉴클레오티드는 생물학적 활성 CanIon 폴리펩티드를 인코드한다.

다른 측면에서, 본원 발명은 SEQ ID NO 5의 아미노산 서열 또는 이의 생물학적 활성 단편을 포함하는 사람 CanIon 폴리펩티드를 인코드하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드를 제공한다.

한 구체예에서, 전술한 폴리뉴클레오티드는 고체 서포트에 부착된다. 다른 구체예에서, 폴리뉴클레오티드는 라벨을 포함한다.

다른 측면에서, 본원 발명은 전술한 폴리뉴클레오티드중 적어도 한가지를 포함하는 폴리뉴클레오티드의 어레이를 제공한다. 한 구체예에서, 어레이는 어드레스(address)가능하다.

다른 측면에서, 본원 발명은 프로모터에 동작가능하게 연결되고 전술한 폴리뉴클레오티드중 한가지를 포함하는 재조합 벡터를 제공한다.

다른 측면에서, 본원 발명은 세포에서 CanIon 유전자의 발현 수준을 변화시키는 폴리뉴클레오티드를 제공한다. 한 구체예에서, 상기 폴리뉴클레오티드는 CanIon 유전자 또는 CanIon 게놈 영역에 삽입된다. 한 구체예에서, 상기 폴리

뉴클레오티드는 CanIon 유전자 프로모터에 삽입된다. 한 구체예에서, 상기 폴리뉴클레오티드는 상동성 재조합, 예를 들면 내인성 CanIon 프로모터 또는 인헨서 영역의 한가지이상 요소를 치환하여 삽입한다.

다른 측면에서, 본원발명은 전술한 재조합 벡터 또는 폴리뉴클레오티드중 한가지를 포함하는 사람제외 숙주 세포를 제공한다.

다른 측면에서, 본원 발명은 적중(knock out) 벡터와의 상동성 재조합에 의해 파괴된 *CanIon* 유전자를 포함하는 포유동물 숙주 세포 또는 사람제외 숙주 포유동물을 제공한다.

다른 측면에서, 본원 발명은 SEQ ID NO 5에 도시된 아미노산 서열 또는 이의 생물학적 활성 단편을 포함하는 분리된, 정제된 또는 재조합된 폴리펩티드를 제공한다.

다른 측면에서, 본원 발명은 폴리펩티드를 만드는 방법을 제공하는데, 상기 방법은 a) 프로모터에 동작가능하게 연결된 본원 발명에 따른 폴리펩티드를 인코드하는 폴리뉴클레오티드를 포함하는 세포 개체군을 제공하고; b) 세포내에서 상기 폴리펩티드의 생산에 도움이 되는 조건하에 세포 개체군을 배양하며; c) 상기 세포 개체군으로부터 폴리펩티드를 정제하는 단계로 구성된다.

다른 측면에서, 본원 발명은 CanIon 폴리펩티드에 항-CanIon 항체를 결합시키는 방법을 제공하는데, 상기 방법은 항체와 폴리펩티드가 특이적으로 결합할 수 있는 조건하에 상기 항체와 전술한 CanIon 폴리펩티드를 접촉시키는 단계로 구성된 다. 다른 측면에서, CanIon 단백질이나 에피토프를 특이적으로 인식하는 항체 또는 이의 면역학적 활성 단편을 제공한다.

다른 측면에서, 본원 발명은 세포내에서 CanIon 유전자의 발현을 검출하는 방법을 제공하는데, 상기 방법은 a) i) 엄격한 조건하에 전술한 CanIon 폴리뉴클레오티드와 혼성화되는 폴리뉴클레오티드; 또는 ii) 전술한 CanIon 폴리펩티드중 한가지에 특이적으로 결합하는 폴리펩티드를 세포 혹은 세포 추출물과 접촉시키고; b) 상기 세포 혹은 세포 추출물에서 폴리뉴클레오티드와 RNA 종간 혼성화의 존부, 또는 상기 세포 혹은 세포 추출물에서 폴리펩티드와 단백질 결합의 존부를 확인하는 단계로 구성되고, 여기서 혼성화 또는 결합의 존재는 상기 CanIon 유전자가 세포내에서 발현된다는 것을 암시한다. 한 구체예에서, 폴리뉴클레오티드는 프라이머이고, 혼성화는 상기 프라이머의 서열을 포함하는 증폭 산물의 존재에 의해 확인된다. 다른 구체예에서, 폴리펩티드는 항체, 예를 들면 항-CanIon 항체이다.

다른 측면에서, 본원 발명은 CanIon 폴리펩티드의 후보 조절물질을 동정하는 방법을 제공하는데, 상기 방법은 a) 전술한 CanIon 폴리펩티드중 한가지와 검사 화합물을 접촉시키고; b) 상기 화합물이 상기 폴리펩티드에 특이적으로 결합하는 지를 확인하는 단계로 구성되고, 여기서 상기 화합물이 상기 폴리펩티드에 특이적으로 결합한다는 확인은 상기 화합물이 CanIon 폴리펩티드의 후보 조절물질이라는 것을 암시한다.

한 구체예에서, 상기 방법에는 후보 조절물질의 존재하의 CanIon 폴리펩티드의 생물활성을 측정하는 단계가 추가로 포함되는데, 후보 조절물질 부재하의 활성 과 비교하여 후보 조절물질 존재하의 CanIon 폴리펩티드의 생물활성 변화는 후보 조절물질이 CanIon 폴리펩티드의 조절물질이라는 것을 암시한다.

다른 측면에서, 본원 발명은 CanIon 폴리펩티드의 조절물질을 동정하는 방법을 제공하는데, 상기 방법은 a) 전술한 CanIon 폴리펩티드중 한가지와 검사 화합물을 접촉시키고; b) 상기 화합물의 존부하에 상기 폴리펩티드의 활성을 확인하는 단계로 구성되고, 여기서 상기 화합물 부재하의 활성과 비교하여 상기 화합물 존재하의 활성에서 차이의 확인은 상기 화합물이 CanIon 폴리펩티드의 조절물질이라는 것을 암시한다.

본원 발명의 한 구체예에서, 폴리펩티드는 세포 혹은 세포막에 존재하고, 생물활성은 전압-개폐 이온 채널 활성이다.

다른 측면에서, 본원 발명은 제약학적 조성물을 제조하는 방법을 제공하는데, 상기 방법은 a) 전술한 방법중 한가지를 이용하여 CanIon 폴리펩티드의 조절물질을 동정하고; b) 상기 조절물질과 생리학적으로 수용가능한 담체를 혼합하는 단계로 구성된다. 제약학적 조성물을 이용하는 방법 역시 제공한다.

또한, 예로써 인체의 치료 또는 전술한 질환이나 이상의 치료를 위한 약물의 제조에서 전술한 CanIon 조절물질, 폴리펩티드, 폴리뉴클레오티드 또는 항체중 한가지의 용도를 제공한다.

또한, 시험관내 또는 생체내에서 본원 발명의 CanIon 폴리뉴클레오티드와 폴리펩티드를 이용 및 검출하는 키트를 제공한다.

도면의 간단한 설명

도 1 은 CanIon 유전자를 포함하는 크로모좀 13q 영역의 BAC 지도를 도시하는 다이어그램이다.

도 2 는 전형적 컴퓨터 시스템의 블록 다이어그램이다.

도 3 은 신규한 뉴클레오티드 또는 단백질 서열을 데이터베이스의 서열과 비교하여 신규한 서열과 데이터베이스의 서열간 상동성 수준을 측정하기 위한 프로세스(200)의 한 구체예를 예시하는 흐름도이다.

도 4 는 2개의 서열이 상동한 지를 측정하기 위한 컴퓨터에서 프로세스(250)의 한 구체예를 예시하는 흐름도이다.

도 5 는 서열에서 지형(feature)의 존재를 검출하기 위한 확인 프로세스(300)의 한 구체예를 예시하는 흐름도이다.

서열 리스트에 제시된 서열의 간단한 설명

SEQ ID No 1은 5' 조절영역(상류의 전사되지 않은 영역)과 엑손 1 내지 7을 포함하는 CanIon의 게놈 서열을 보유한다.

SEQ ID No 2는 엑손 8 내지 27을 포함하는 CanIon의 게놈 서열을 보유한다.

SEQ ID No 3은 엑손 28 내지 44 및 3' 조절영역(하류의 전사되지 않은 영역)을 포함하는 CanIon의 게놈 DNA 서열을 보유한다.

SEQ ID No 4는 CanIon의 cDNA 서열을 보유한다.

SEQ ID No 5는 SEQ ID No 4에 의해 인코드되는 아미노산 서열을 보유한다.

SEQ ID No 6은 이중대립인자 마커 A18을 포함하는 앰플리콘의 뉴클레오티드 서열을 보유한다.

SEQ ID No 7은 실시예 2에 자세하게 기술된 PU 5' 서열을 추가로 포함하는 프라이머를 보유한다.

SEQ ID No 8은 실시예 2에 자세하게 기술된 RP 5' 서열을 추가로 포함하는 프라이머를 보유한다.

서열 리스트와 관련된 규정에 따라, 서열내에서 이중대립인자 마커의 위치를 표시하고 다형적염기에 존재하는 대립유전자 각각을 동정하기 위하여 서열 리스트에 다음과 같은 코드가 이용되었다. 서열에서 코드 'r'은 다형적염기의 한 대립유전자가 구아닌이고 다른 대립유전자가 아데닌이라는 것을 나타낸다. 서열에서 코드 'y'는 다형적염기의 한 대립유전자가 티민이고 다른 대립유전자가 시토신이라는 것을 나타낸다. 서열에서 코드 'm'은 다형적염기의 한 대립유전자가 아데닌이고 다른 대립유전자가 시토신이라는 것을 나타낸다. 서열에서 코드 'k'는 다형적염기의 한 대립유전자가 구아닌이고 다른 대립유전자가 티민이라는 것을 나타낸다. 서열에서 코드 's'는 다형적염기의 한 대립유전자가 구아닌이고 다른 대립유전자가 시토신이라는 것을 나타낸다. 서열에서 코드 'w'는 다형적염기의 한 대립유전자가 아데닌이고 다른 대립유전자가 티민이라는 것을 나타낸다. 각 이중대립인자 마커에 대한 원형 대립유전자의 뉴클레오티드 코드는 다음과 같다:

이중대립인자 마커 원형 대립유전자

5-124-273 -----------A(예를 들면)

일부 경우에, 이중대립인자 마커의 다형적염기는 인코드된 폴리펩티드에서 아미노산의 본성을 변화시킨다. 이는 변이(VARIANT) 특징, 다형적 아미노산의 위 치에 Xaa의 배치 및 2개의 양자택일적 아미노산으로 Xaa의 정의를 이용하여 첨부된 서열 리스트에 나타낸다. 가령, 이중대립인자 마커의 한 대립유전자가 히스티딘을 인코드하는 코돈 CAC이고 이중대립인자 마커의 다른 대립유전자가 글루타민을 인코드하는 CAA이면, 인코드된 폴리펩티드에 대한 서열 리스트는 다형적 아미노산의 위치에 Xaa를 보유하게 된다. 이런 경우에, Xaa는 히스티딘 또는 글루타민으로 정의된다.

다른 경우에, Xaa는 뉴클레오티드 서열 다의성으로 인해 실체가 불명확한 아미노산을 표시할 수 있다. 이런 경우에, 불명확(UNSURE) 특징, 불명확 아미노산의 위치에 Xaa의 배치 및 유전자 코드에 의한 20개 아미노산이나 한정된 수의 아미노산으로 Xaa의 정의가 이용된다.

발명의 상세한 설명

가족에서 정신분열증과 양극성 장애의 집중, 쌍둥이와 입양 연구의 증거, 전세계적 발병률에서 변화 결여는 환경적 위험 요인이 일정 수준으로 발병에 관여한다 하더라도 정신분열증과 양극성 장애가 일차적으로 유전적 질환이라는 것을 암시한다. 가령, 정신분열증은 전체 인구의 1%에서 발병한다. 하지만, 조부모중 1명에게 정신분열증이 있으면, 병에 걸릴 위험이 3% 정도; 부모중 1명에게 정신분열증이 있으면 10% 정도로 증가한다. 부모 모두에게 정신분열증이 있으면 위험은 40% 정도까지 상승한다.

특정 형질, 예를 들면 정신분열증과 같은 특이적인 중추신경계 질환에 관여하는 유전자의 동정은 현재 유전자 지도작성에 활용되고 있는 2가지 주요 전략: 연 쇄 분석 및 연관 연구를 통하여 실시할 수 있다. 연쇄 분석은 여러 환자의 가족 연구를 필요로 하고, 현재 단일- 또는 저인자(oligogenic) 유전된 형질의 검출에 이용된다. 반대로, 연관 연구는 형질 음성 (T-) 대조군과 비교하여 무관한 형질 (T+) 개체에서 마커 대립유전자의 빈도를 검사하고, 일반적으로 다인자(polygenic) 유전의 검출에 이용된다.

유전적 링크 또는 '연쇄'는 크로모좀에서 2개의 인접한 서열이 감수분열동안 유전자 교환에 의한 재조합을 최소한으로 보유하는 지에 대한 분석에 기초한다. 이를 위하여, 마이크로위성(microsatellite) 마커와 같은 크로모좀 마커는 게놈에 정확하게 위치시킨다. 유전자 링크 분석은 계통도(genealogical tree)에 따라 사용된 크로모좀 마커와 표적 유전자에서 재조합, 상기 질환의 전달, 마커의 전달 가능성을 측정한다. 따라서, 상기 질환에서 임의 마커의 특정 대립유전자가 통상적 수준(0 내지 0.5의 재조합 수준)보다 더 높은 빈도로 전달되면, 표적 유전자가 마커의 인근에서 발견된다고 추론하는 것이 가능하다. 이런 기술을 활용하여, 가족성 암의 유전적 소인을 보이는 몇몇 유전자의 위치를 확인하였다. 유전자 연쇄 연구에 포함되기 위해서는 유전 질환의 영향을 받은 가족은 '정보적 요소' 기준을 충족시켜야 한다: 세대마다 여러 명의 영향을 받은 개체(구성 DNA가 가용함) 및 많은 후손.

기존 연쇄 연구의 결과는 크로모좀 13이 13q32에서 정신분열증 감수성 좌위를 보유할 가능성이 높다는 가설을 뒷받침하였다(Blouin JL et al., 1998, Nature Genetics, 20:70-73; Lin MW et al., 1997, Hum. Genet., 99(3):417-420; Brzustowicz et al., Am. J. Hum. Genet. 65:1096-1103(1999)). 연쇄 분석은 형질(trait)에 관여하는 유전자를 검출하는데 효과적인 방법이긴 하지만, 분석능력이 메가염기 수준을 넘지 못하고 상기 방법으로 처음에 동정된 영역의 분석결과를 정련하는데 상보성 연구가 종종 요구된다.

크로모좀 13q-31-q33 좌위의 후보 게놈 영역을 포함하는 BAC contig(콘틱)은 7개의 게놈 등가의 독점적 BAC 라이브러리를 스크린하기 위하여 크로모좀 13q31-33q 영역에 위치하는 공개된 STS로 작제하였다. 이들 물질로부터 신규한 STS를 만들고, 상기 영역의 밀집된 물리적 지도를 작성하였다. BAC는 검증을 위하여 in situ 크로모좀 혼성화로 크기를 측정하고 지도를 작성하였다. 최소한의 BAC가 동정되고 완전하게 서열분석되었는데, 여기에서 4Mb이상 서열로 구성되는 콘틱이 작성되었다. 이런 지도의 작성은 정신분열증과 현저한 연쇄를 보이는 게놈 영역내에 위치하는 CanIon 유전자의 동정을 결과하였다.

CanIon 아미노산 서열은 칼슘 채널의 알파-1 아단위에 대한 기호(signature)를 제공하는 7개-구성요소 핑거프린트인 Ca 채널의 특징이다. 이런 핑거프린트는 6개 서열로 구성되는 초기 정렬로부터 유래되었다: 모티프는 이들과 다른 양이온 채널을 구별할 수 있는 보존된 루프 영역으로부터 도출되었다; 모티프 1과 2는 막통 분절 4와 5 및 5와 6(첫 번째 내부 반복서열)간 보존된 루프 영역을 인코드한다; 모티프 3은 반복서열 1의 분절 6과 반복서열 2의 분절 1간 보존된 루프 영역에 상응한다; 모티프 4는 반복서열 2의 분절 5와 6간 보존된 루프 영역을 인코드한다; 모티프 5는 반복서열 3의 분절 6과 반복서열 4의 분절 1간 보존된 루프 영역에 상 응한다; 모티프 6과 7은 반복서열 4의 분절 4와 5 및 5와 6간 보존된 루프 영역을 인코드한다

도 1 은 CanIon 유전자를 포함하는 관심있는 크로모좀 13 영역을 포괄하는 BAC 콘틱을 도시하고, 연쇄 연구에서 가장 높은 중요성을 보이는 유전자 마커와 관련하여 CanIon 유전자의 게놈 위치를 보여준다. 특히, Blouin 등(1998)은 54개 복합 혈통에서 452개의 마이크로위성 마커를 이용하여 정신분열증 감수성 좌위에 대한 전체 게놈 검색(genome wide scan)을 실시하였다. 가족에서 정신분열증간 가장 현저한 연쇄는 마커 D13S174 주변의 크로모좀 13q32에서 확인되었다. Brzustowics 등(1999)은 21개의 확대된 캐나다 가계에서 크로모좀 8과 13에 걸쳐 마이크로위성 마커를 평가하였다. 크로모좀 13q 영역에서 마커는 이용된 각 분석 모델에서 파지티브 LOD 스코어를 보였다: 정신분열증의 좁은 또는 넓은 정의에 따라 상염색체 우성과 열성. 최대 3점 LOD 스코어는 열성의 넓은 모델하의 마커 D13S798로 얻었다: 상동성하의 재조합 분율(fraction)(θ).1에서 3.92; 이질성하의 α=.65와 θ=0에서 4.42. 도 1을 참고하면, CanIon 유전자는 콘틱 'E 영역' 및 콘틱 C0001A10에 부분적으로 위치한다. CanIon 유전자는 연쇄 연구에서 가장 높은 중요성을 보인 2개의 마커가 측면에 접한다. 마커 D13S174는 콘틱 C0001A10에 위치하고, 마커 D13S793은 CanIon 유전자의 대략 3.5 Mb 중심에 위치한다.

정신분열증, 양극성 장애 및 다른 CNS와 심혈관의 질환과 이상에 관여하는 유전자를 동정하는 하는 것이 필요하다. 또한, 병에 관여하는 신규한 이온 채널을 동정하는 것이 필요하다. 이들 유전자와 단백질은 정신분열증, 양극성 장애 또는 다른 CNS 질환 및 다른 질환, 예를 들면 심장 질환과 고혈압의 치료에 새로운 개입지점을 제공하고, CanIon 유전자 및 관련된 생물학적 경로의 추가적인 연구와 특성화를 가능하게 한다. 이들 질환과 이상에 관여하는 이들 유전자 및 관련된 생물학적 경로에 관한 지식은 연자들이 예로써 정신분열증과 양극성 장애의 병인을 이해하도록 하고, 상기

질환의 원인에 대한 약물과 약물 치료를 결과할 것이다. 가령, CanIon 채널을 차단하는 화합물은 다양한 질환이나 이상, 바람직하게는 정신분열증이나 양극성 장애, 통증 질환, 간질, 다양한 심혈관 질환(예, 심장 부정맥), 조병, 고혈압을 치료하는데 사용될 수 있다. 또한, 정신분열증, 양극성 장애 및 다른 이상에 대한 감수성을 검출하고 이들 질환의 발병을 예방 또는 추적하는 신규한 방법이 필요하다. 진단 도구 역시 매우 유용하다. 실제로, 정신분열증이 발병할 위험이 높은 개체의 조기 확인은 진단 및/또는 예방 치료를 가능하게 한다. 게다가, 약물의 효능과 이에 대한 환자의 내약성에 대한 정확한 평가는 의사들이 정신분열증과 양극성 장애 치료의 이익/위험 비율을 향상시킬 수 있도록 할 것이다.

본원 발명은 CanIon 유전자와 관련된 폴리뉴클레오티드와 폴리펩티드에 관한다. CanIon의 게놈이나 cDNA 서열과 특이적으로 혼성화되는 올리고뉴클레오티드 프로브와 프라이머 역시 본원 발명에 속한다. 본원 발명의 다른 목적은 본원에서 밝힌 핵산 서열중 한가지를 포함하는 재조합 벡터, 특히 CanIon의 조절영역 또는 CanIon 단백질을 인코드하는 서열을 포함하는 재조합 벡터 및 상기 핵산 서열 또는 재조합 벡터를 포함하는 숙주 세포를 제공하는 것이다. 본원 발명은 또한, CanIon 유전자나 단백질의 발현 또는 활성을 조절하는 능력에 대한 분자의 스크리닝 방법 및 정신분열증, 양극성 장애 또는 다양한 다른 질환이나 이상을 치료 또는 예방하기 위하여 이런 분자를 이용하는 방법에 관한다. 본원 발명은 또한, 진단 시약으로 유용한이런 폴리펩티드를 특이적으로 지향하는 항체에 관한다.

본원 발명은 또한, 가족에서 연쇄 연구, 개체군에서 연쇄 불균형 연구, 사례 대조 개체군에서 연관 연구를 비롯한 유전자 분석 방법에서 CanIon-관련된 이중대립인자 마커 및 이들의 용도에 관한다. 본원 발명의 중요한 측면은 이중대립인자 마커를 통하여 복잡한 형질에 관여하는 유전자의 역할을 확인하는 연관 연구를 실시할 수 있다는 점이다.

정의

본원 발명을 좀더 자세하게 설명하기에 앞서, 본원 발명을 설명하기 위하여 이용된 용어의 의미와 범위를 한정하기 위하여 다음과 같은 정의를 제시한다.

본원에서 'CanIon 유전자'에는 CanIon 단백질을 인코드하고 게놈 DNA의 번역되지 않은 조절영역을 포함하는 게놈, mRNA, cDNA 서열이 포함된다.

본원에서 '이질성 단백질'은 CanIon 단백질을 제외한 임의의 단백질 또는 폴리펩티드를 의미한다. 가령, 이질성 단백질은 CanIon 조절 영역과의 추가적인 검사에 마커로 사용할 수 있는 화합물일 수 있다.

'분리된'은 물질이 원래의 환경(예, 자연 발생하는 경우에 고유 환경)으로부터 이동되는 것을 의미한다. 가령, 살아있는 동물에 존재하는 자연-발생 폴리뉴클레오티드 또는 폴리펩티드는 분리된 것이 아니지만, 고유 시스템에서 공존하는 물질의 일부 또는 전부로부터 격리된 상기와 동일한 폴리뉴클레오티드, DNA 또는 폴 리펩티드는 분리된 것이다. 이런 폴리뉴클레오티드는 벡터의 일부가 되거나 또는 이런 폴리뉴클레오티드 혹은 폴리펩티드는 조성물의 일부가 될 수 있는데, 벡터나 조성물이 고유 환경의 일부가 아니라는 점에서 이 역시 분리된 것이다.

가령, 살아있는 동물에 존재하는 자연-발생 폴리펩티드는 분리된 것이 아니지만, 고유 시스템에서 공존하는 물질의 일부 또는 전부로부터 격리된 상기와 동일한 폴리뉴클레오티드는 분리된 것이다. '분리된'의 정의에서 자연-발생 크로모좀(예, 크로모좀 스프레드); 인공 크로모좀 라이브러리; 게놈 라이브러리; 시험관내 핵산 제형 또는 트랜스펙션/형질전환된 숙주 세포 제형(여기서, 숙주 세포는 시험관내 이질성 제형이거나 또는 단일 콜로니의 이질성 개체군으로 도말된다)으로 존재하는 cDNA 라이브러리는 제외된다. 특정된 폴리뉴클레오티드가 벡터 분자에서 핵산 삽입체 수의 5% 미만을 구성하는 라이브러리 역시 제외된다. 전체 세포 게놈 DNA 또는 전체 세포 RNA 제형(기계적으로 전단된 또는 효소적으로 절단된 전체 세포 제형 포함) 역시 제외된다. 전기영동(블랏 전달)에 의해 분리되는 시험관내 제형 또는 이질성 혼합물로서 상기 전체 세포 제형 역시 배제되는데, 여기서 본원 발명의 폴리뉴클레오티드는 전기영동 매체에서 이질성 폴리뉴클레오티드로부터 더 이상 분리되지 않는다(예, 아가로즈 겔 또는 나일론 블랏에서 이질성 밴드 개체군으로부터 단일 밴드를 절단함으로써 추가적인 분리).

본원에서 '정제된'은 절대 순도를 요구하지는 않는다; 오히려, 이는 상대적 정의로서 이해한다. 출발 물질 또는 천연 물질의 정제는 최소한 1 크기 자릿수, 바람직하게는 2 또는 3 크기 자릿수, 좀더 바람직하게는 4 또는 5 크기 자릿수로 생각한다. 예로써, 0.1% 농도에서 10% 농도로의 정제는 2 크기 자릿수이다. 예시하면, cDNA 라이브러리로부터 개별 cDNA 클론은 통상적으로 전기영동적 동질성까지 정제된다. 이들 클론으로부터 수득된 서열은 라이브러리로부터 또는 전체 사람 DNA로부터 구할 수 없었다. cDNA 클론은 자연 발생하지 않고, 부분적으로 정제된 자연 발생 물질(메신저 RNA)의 조작을 통하여 얻는다. mRNA의 cDNA 라이브러리로의 전환에는 합성 물질(cDNA)의 생성이 포함되고, 순수한 개별 cDNA 클론은 클론 선별로 합성 라이브러리로부터 분리할 수 있다. 따라서, 메신저 RNA로부터 cDNA 라이브러리를 만들고, 후속으로 상기 라이브러리로부터 개별 클론을 분리하면 고유 메시지가 10^4-10^6 배 정제된다.

이에 더하여, '정제된'은 폴리펩티드, 폴리뉴클레오티드, 탄수화물, 지질 등을 포함하나 이에 국한되지 않는 다른 화합물로부터 격리되는 본원 발명의 폴리펩티드 또는 폴리뉴클레오티드를 설명하는데 이용될 수 있다. '정제된'은 호모- 또는 헤테로- 이합체, 삼합체 등과 같은 올리고형으로부터 본원 발명의 단량체형 폴리펩티드의 분리를 명기하는데 이용될 수도 있다. '정제된'은 선형 폴리뉴클레오티드로부터 공유 결합된 폴리뉴클레오티드의 분리를 명기하는데 이용될 수도 있다. 폴리펩티드는 적어도 50%, 바람직하게는 60 내지 75%의 시료가 단일 폴리뉴클레오티드 서열과 입체형태를 보이는 경우에 실질적으로 순수하다. 실질적으로 순수한 폴리펩티드 또는 폴리뉴클레오티드는 각각 대략 50%, 바람직하게는 60 내지 90%, 좀더 바람직하게는 대략 95%, 가장 바람직하게는 99% w/w이상의 폴리펩티드 또는 폴리뉴클레오티드 시료로 구성된다. 폴리펩티드와 폴리뉴클레오티드 순도 또는 동질성은 당분야에 공지된 방법, 예를 들면 시료의 아가로즈 또는 폴리아크릴아마이드 겔 전기영동 및 후속으로 겔을 염색한 직후에 단일 밴드 가시화로 표시한다. 특정 목적을 위한 더 높은 분석은 HPLC 또는 다른 공지된 방법으로 제공할 수 있다. 대안적 구체예로 본원 발명에 따른 폴리펩티드와 폴리뉴클레오티드의 정제는 이질성 폴리펩티드와 폴리뉴클레오티드(DNA, RNA 또는 둘 모두)와 관련하여 '적어도' 순도%로 표시할 수도 있다. 적절한 구체예에서, 본원 발명의 폴리펩티드와 폴리뉴클레오티드는 각각 이질성 폴리펩티드와 폴리뉴클레오티드와 관련하여 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 96%, 98%, 99%, 또는 100% 순도이다. 다른 적절한 구체예에서, 폴리펩티드와 폴리뉴클레오티드는 각각 이질성 폴리펩티드 또는 폴리뉴클레오티드와 관련하여 90% 내지 100%사이에서 1/1000 자릿수까지의 순도(예, 적어도 99.995% 순도의 폴리펩티드 또는 폴리뉴클레오티드), 또는 담체에 존재하는 화합물 및 분자를 제외한 모든 화합물 및 분자와 관련하여 w/w 비율로 순도를 갖는다. 1/1000 자릿수까지 순도 %를 표시하는 각 숫자는 순도의 개별 형식으로 주장할 수 있다.

본원에서 '폴리펩티드'는 중합체의 길이와 상관없이 아미노산 중합체를 의미한다; 따라서, 펩티드, 올리고펩티드, 단백질은 폴리펩티드의 정의에 포함된다. 이 용어는 폴리펩티드의 발현-후 수식을 명기 또는 배제하지 않는다, 예를 들면 글리코실기, 아세틸기, 포스페이트기, 지질기 등을 보유하는 폴리펩티드는 폴리펩티드에 포함된다. 하나이상의 아미노산 유사체를 보유하는 폴리펩티드(예, 비-자연 발생 아미노산, 무관한 생물체에서만 자연적으로 발생하는 아미노산, 포유동물체로부터 유래된 수식된 아미노산 등), 치환된 연쇄를 보유하는 폴리펩티드, 자연 발생하고 자연-발생하지 않는 당분야에 공지된 다른 변형체 역시 폴리펩티드에 포함된다.

본원에서 '재조합 폴리펩티드'는 인위적으로 설계되고 최초 고유 환경의 연속 폴리펩티드 서열에서 발견되지 않는 적어도 2개의 폴리펩티드 서열을 포함하는 폴리펩티드, 또는 재조합 폴리뉴클레오티드로부터 발현되는 폴리펩티드를 의미한다.

본원에서 '사람제외 동물'은 사람제외 척추동물, 조류, 포유동물, 바람직하게는 영장류, 가금(예, 돼지, 염소, 양, 당나귀, 말, 토끼) 또는 설취류, 좀더 바람직하게는 쥐 또는 생쥐를 의미한다. 본원에서 '동물'은 임의의 척추동물, 바람직하게는 포유동물을 의미한다. '동물'과 '포유동물'에는 '사람제외'를 명시하지 않는 경우 사람이 포함된다.

본원에서 '항체'는 한 개이상의 결합 도메인으로 구성되는 폴리펩티드 또는 폴리펩티드 군을 의미하는데, 여기서 항체 결합 도메인은 항체 분자에서 다양한 도메인의 접힙(folding)으로 생성되고 항원의 항원 결정기 외형에 상보적인 내부 표면 형태 및 전하 분포를 갖는 3차원 결합 공간을 형성하는데, 이는 항원과의 면역학적 반응을 가능하게 한다. 항체에는 Fab, Fab', $F(ab)_2$, $F(ab')_2$ 를 비롯하여 결합 도메인을 보유하는 재조합 단백질이 포함된다.

본원에서 '항원 결정기'는 항원-항체 반응의 특이성을 결정하는 항원 분자의 일부분, 본원에서는 CanIon 폴리펩티드를 의미한다. '에피토프'는 폴리펩티드의 항원 결정기를 의미한다. 에피토프는 상기 에피토프에 독특한 공간적 입체형태를 제공하는 적게는 3개의 아미노산으로 구성된다. 일반적으로, 에피토프는 적어도 6개의 이런 아미노산, 바람직하게는 적어도 8-10개의 이런 아미노산으로 구성된다. 에피토프를 구성하는 아미노산을 측정하는 방법에는 x-레이 결정법; 2-차원 핵 자기 공명법; 에피토프 지도작성, 예를 들면 Geysen et al. 1984; PCT 출원 WO 84/03564; PCT 출원 WO 84/03506에서 밝힌 Pepscan 방법이 포함된다.

본원에서, '뉴클레오티드 서열'은 폴리뉴클레오티드 또는 핵산을 의미한다. 좀더 구체적으로, '뉴클레오티드 서열'에는 핵산 물질 자체가 포함되고, 따라서 특정 DNA 또는 RNA 분자를 생화학적으로 특성화하는 서열 정보(즉, 4개의 염기 문자에서 선택되는 문자의 연속)에 제한되지 않는다.

본원에서, '핵산', '올리고뉴클레오티드', '폴리뉴클레오티드'는 서로 동일한 의미를 가지고, 여기에는 단일 사슬 또는 이중 나선 형태로 한 개이상의 뉴클레오티드를 포함하는 RNA, DNA, 또는 RNA/DNA 하이브리드 서열이 포함된다. 본원에서 '뉴클레오티드'에는 단일 사슬 또는 이중 나선 형태로 임의 길이의 RNA, DNA, 또는 RNA/DNA 하이브리드 서열을 포함하는 서열 역시 포함된다. 또한, 본원에서 '뉴클레오티드'는 개별 뉴클레오티드 또는 다수의 뉴클레오티드, 다시 말하면 올리고뉴클레오티드 또는 폴리뉴클레오티드의 뉴클레오티드의 경우에 퓨린 또는 피리미딘, 리보즈 또는 디옥시리보즈 당 성분, 인산염 기 또는 포스포디에스터 결합으로 구성되는 분자 또는 좀더 큰 핵산 분자에서 개별 단위를 의미한다. '뉴클레오티드'에는 적어도 한 개의 변형체 (a) 다른 결합기, (b) 퓨린의 유사체, (c) 피리미딘의 유사체 또는 (d) 당 유사체를 포함하는 '변형된 뉴클레오티드'가 포함된다(참조: WO95/04064). 본원 발명의 폴리뉴클레오티드 서열은 합성, 재조합, *ex vivo* 발생 또는 이들의 조합을 비롯한 임의의 공지된 방법 및 당분야에 공지된 정

제 방법을 활용하여 만들 수 있다.

조절 서열, 예를 들면 프로모터에 '동작가능하게 연결된' 서열은 조절 요소가 핵산에서 정확한 위치와 방향으로 존재하여 RNA 중합효소 개시 및 관심있는 핵산의 발현을 조절한다는 것을 의미한다. 본원에서 '동작가능하게 연결된'은 기능적 상관관계에서 폴리뉴클레오티드 요소의 연쇄를 의미한다. 가령, 프로모터 또는 인헨서는 코딩 서열에 동작가능하게 연결되고 코딩 서열의 전사에 영향을 준다.

'형질'과 '표현형'은 본원에서 동일한 의미를 갖고 미생물의 가시적인, 검출가능한 또는 달리 측정가능한 특성, 예를 들면 질환의 증상 또는 질환에 대한 감수성을 의미한다. 일반적으로, '형질' 또는 '표현형'은 질환의 증상 또는 이에 대한 감수성, 치료와 관련된 유익한 반응 또는 치료와 관련된 부작용을 의미한다. 적절하게는, 이런 형질은 정신분열증이나 양극성 장애와 같은 정신 질환; 간질이나 통증 질환과 같은 다른 CNS 또는 신경 질환; 협심증, 고혈압, 부정맥과 같은 심혈관 질환; 이들 질환이나 이상중 한가지의 임의 측면, 특징 또는 특질일 수 있지만, 이들에 국한되지는 않는다.

'대립유전자'는 뉴클레오티드 서열의 변이를 의미한다. 이중대립인자 다형 성은 2가지 형태를 갖는다. 이배체 미생물은 하나의 대립유전자형에 대하여 동형접합성 또는 이형접합성일 수 있다.

본원에서 '이형접합률(heterozygosity rate)'은 특정 대립유전자에서 이형접합성인 개체군에서 개체 빈도를 의미한다. 이중대립인자 시스템에서, 이형접합률은 평균 $2P_a(1-P_a)$인데, 여기서 P_a는 최소 공통 대립유전자의 빈도이다. 유전자 연구에 유용한 유전자 마커는 무작위로 선택되는 개체가 이형접합일 가능성을 온당하게 허용하는 적합한 수준의 이형접합성을 보유한다.

본원에서 '유전형'은 개체 또는 시료에 존재하는 대립유전자의 본성을 의미한다. 적절하게는, 유전형은 개체 또는 시료에 존재하는 이중대립인자 마커 대립유전자, 예를 들면 CanIon 유전자 또는 게놈 영역에서 이중대립인자 마커의 대립유전자의 의미한다. '이중대립인자 마커'에 대한 시료 또는 개체의 '유전자형판별(genotyping)'에는 이중대립인자 마커에서 개체가 보유하는 특정 대립유전자 또는 특정 뉴클레오티드를 결정하는 것이 포함된다.

본원에서 '돌연변이'는 1% 미만의 빈도로 상이한 게놈 또는 개체간 DNA 서열에서 차이를 의미한다.

본원에서 '일배체형'은 개체 또는 시료에 존재하는 대립유전자의 조합을 의미한다. 적절하게는, 일배체형은 임의의 개체에서 발견되고 표현형과 연관된 이중대립인자 마커 대립유전자의 조합을 의미한다.

본원에서 '다형성'은 상이한 게놈 또는 개체간 2개이상의 게놈 서열 또는 대 립유전자의 발생을 의미한다. '다형적(polymorphic)'은 특정 게놈 서열에서 2개 이상 변이가 개체군에서 발견되는 상태를 의미한다. '다형적 부위'는 변이가 발생하는 좌위이다. 단일 뉴클레오티드 다형성은 다형적 부위에서 1개 뉴클레오티드의 다른 뉴클레오티드로의 치환이다. 단일 뉴클레오티드의 결실 또는 단일 뉴클레오티드의 부가 역시 단일 뉴클레오티드 다형성을 유발한다. 본원에서, '단일 뉴클레오티드 다형성'은 가급적 단일 뉴클레오티드 치환을 의미한다. 전형적으로, 상이한 개체간 다형적 부위가 2개의 상이한 뉴클레오티드에 의해 점유될 수도 있다. 본원에서 '이중대립인자 다형성'과 '이중대립인자 마커'는 동일한 의미를 갖고 개체군에서 상당히 높은 빈도로 2개의 대립유전자를 갖는 단일 뉴클레오티드 다형성을 의미한다. '이중대립인자 마커 대립유전자'는 이중대립인자 마커 부위에 존재하는 뉴클레오티드 변이를 의미한다. 전형적으로, 본원 발명에 따른 이중대립인자 마커의 조금 덜 공통된 대립유전자의 빈도는 1%이상, 바람직하게는 10%이상, 좀더 바람직하게는 20%이상(즉, 적어도 0.32의 이형접합률), 가장 바람직하게는 30%이상(즉, 적어도 0.42의 이형접합률)이다. 조금 덜 공통된 대립유전자의 빈도가 30%이상인 이중대립인자 마커는 '고 품질 이중대립인자 마커'라고 한다.

폴리뉴클레오티드의 중심과 관련하여 폴리뉴클레오티드에서 뉴클레오티드의 위치는 다음의 방식으로 설명한다. 폴리뉴클레오티드가 홀수의 뉴클레오티드를 보유하면, 폴리뉴클레오티드의 3'와 5' 말단으로부터 동등한 거리에 있는 뉴클레오티드는 폴리뉴클레오티드의 '중심'으로 간주하고, 중심에 위치한 뉴클레오티드에 바로 인접하는 뉴클레오티드 또는 중심 자체에 위치하는 뉴클레오티드는 중심으로부 터 1개 뉴클레오티드 이내'에 위치하는 것으로 간주한다. 홀수개의 뉴클레오티드를 갖는 폴리뉴클레오티드 중간에서 5개의 뉴클레오티드 위치중 하나는 중심으로부터 2개 뉴클레오티드 이내에 위치하는 것으로 간주한다. 폴리뉴클레오티드가 짝수개의 뉴클레오티드를 보유하면, 폴리뉴클레오티드의 중심에는 뉴클레오티드가 아닌 결합이 존재한다. 따라서, 2개의 중심 뉴클레오티드중 하나는 '중심으로부터 1개 뉴클레오티드 이내'에 위치하는 것으로 간주하고, 폴리뉴클레오티드 중간에서 4개의 뉴클레오티드중 하나는 '중심으로부터 2개 뉴클레오티드 이내'에 위치하는 것으로 간주한다. 1개이상 뉴클레오티드의 치환, 결실 또는 부가를 비롯한 다형성의 경우에, 다형성의 치환된, 삽입된 또는 결실된 폴리뉴클레오티드로부터 거리와 상기 폴리뉴클레오티드의 3' 말단간 차이 및 다형성의 치환된, 삽입된 또는 결실된 폴리뉴클레오티드로부터 거리와 상기 폴리뉴클레오티드의 5' 말단간 차이가 0이나 1개 뉴클레오티드이면 다형성, 대립유전자 또는 이중대립인자 마커는 폴리뉴클레오티드의 '중심'에 위치한다. 이런 차이가 0 내지 3이면, 다형성은 '중심으로부터 1개 뉴클레오티드 이내'에 위치하는 것으로 간주한다. 이런 차이가 0 내지 5이면, 다형성은 '중심으로부터 2개 뉴클레오티드 이내'에 위치하는 것으로 간주

한다. 이런 차이가 0 내지 7이면, 다형성은 '중심으로부터 3개 뉴클레오티드 이내'에 위치하는 것으로 간주한다.

본원에서 '상류'는 특정 기준지점으로부터 폴리뉴클레오티드의 5'말단 방향, 또는 유전자의 경우에 코딩 서열로부터 프로모터 방향으로의 위치를 의미한다.

본원에서 '염기쌍'과 'Watson amp; Crick 염기쌍'은 동일한 의미를 갖고 이중나 선 DNA에서 발견되는 것과 유사한 방식으로 서열 상동성으로 인해 서로 수소 결합된 뉴클레오티드를 의미하는데, 여기서 티민 또는 우라실 잔기는 2개의 수소 결합에 의해 아데닌 잔기에 결합하고, 사이토신 잔기는 3개의 수소 결합에 의해 구아닌 잔기에 결합한다(Stryer, L., Biochemistry, 4 th edtion, 1995)

'상보성' 또는 '이의 보체'는 상보성 전체 영역에서 다른 특정된 폴리뉴클레오티드와 Watson amp; Crick 염기쌍을 형성할 수 있는 폴리뉴클레오티드 서열을 의미한다. 본원 발명에서는 첫 번째 폴리뉴클레오티드에서 각 염기가 상보성 염기와 쌍을 이루는 경우에 첫 번째 폴리뉴클레오티드는 두 번째 폴리뉴클레오티드와 상보성이라고 한다. 상보성 염기는 A와 T(또는 A와 U), 또는 C와 G이다. 본원에서 '보체'는 '상보성 뉴클레오티드', '상보성 핵산', '상보성 뉴클레오티드 서열과 동의어로 사용된다. 이들 용어는 2개의 폴리뉴클레오티드가 실제로 반응하는 특정 조건이 아닌 순전히 서열에 기초한 폴리뉴클레오티드의 쌍에 적용된다.

변이와 단편

본원 발명은 전술한 폴리뉴클레오티드, 특히 본원 발명에 따른 하나이상의 이중대립인자 마커를 보유하는 CanIon 유전자의 변이와 단편에 관한다.

본원에서 폴리뉴클레오티드의 변이는 기준 폴리뉴클레오티드와 상이한 폴리뉴클레오티드이다. 폴리뉴클레오티드의 변이는 자연 발생 변이, 예를 들면 자연 발생 대립유전자 변이이거나 또는 자연 발생하는 것으로 알려져 있지 않은 변이이다. 폴리뉴클레오티드의 이런 비-자연 발생 변이는 폴리뉴클레오티드, 세포 또는 미생물에 적용되는 것을 비롯하여 돌연변이유발 기술로 만들 수 있다. 일반적으로, 차이는 제한적이고, 기준 뉴클레오티드 서열과 변이는 전반적으로 유사하고 다수의 영역에서 상동하다.

본원 발명에 따른 폴리뉴클레오티드의 변이에는 SEQ ID No 1 내지 4 뉴클레오티드 서열과 적어도 95% 상동한 뉴클레오티드 서열 혹은 SEQ ID No 1 내지 4 뉴클레오티드 서열의 적어도 12개 연속 뉴클레오티드의 폴리뉴클레오티드 단편 및 SEQ ID No 1 내지 4 뉴클레오티드 서열 또는 SEQ ID No 1 내지 4 뉴클레오티드 서열의 적어도 12개 연속 뉴클레오티드 서열의 폴리뉴클레오티드 단편과 바람직하게는 적어도 99% 상동한, 좀더 바람직하게는 99.5% 상동한, 가장 바람직하게는 적어도 99.8% 상동한 뉴클레오티드 서열이 포함되지만, 이들에 국한되지 않는다.

변이 폴리뉴클레오티드에 존재하는 뉴클레오티드 변화는 침묵할 수도 있는데, 이는 이들이 폴리뉴클레오티드에 의해 인코드되는 아미노산을 변화시키지 않는다는 것을 의미한다. 하지만, 뉴클레오티드 변화는 기준 서열에 의해 인코드되는 폴리펩티드에서 아미노산 치환, 부가, 결실, 융합, 절두를 유발할 수도 있다. 치환, 결실 또는 부가에는 하나이상의 뉴클레오티드가 참여할 수 있다. 변이는 코딩 영역, 비-코딩 영역 또는 둘 모두에서 발생할 수 있다. 코딩 영역에서 변화는 보존성 또는 비-보존성 아미노산 치환, 결실 또는 부가를 유발할 수 있다.

본원 발명의 특히 적절한 구체예에서, 폴리뉴클레오티드는 성숙 CanIon 단백질과 동일한 생물학적 기능이나 활성을 실질적으로 보유하는 폴리펩티드, 또는 특정 생물 활성을 유지 혹은 증가시키면서 이차적인 생물 활성을 감소시키는 폴리펩 티드를 인코드한다.

폴리뉴클레오티드 단편은 임의의 뉴클레오티드 서열, 바람직하게는 CanIon 유전자의 뉴클레오티드 서열 및 이의 변이의 일부분과 완전히 상동한 서열을 보유하는 폴리뉴클레오티드이다. 단편은 CanIon 유전자의 인트론이나 엑손의 일부일 수 있다. 이는 CanIon의 조절 영역의 일부일 수도 있다. 적절하게는, 이런 단편은 이중대립인자 마커 A1 내지 A17중 적어도 한가지 혹은 이들의 보체, 또는 이중대립인자 마커 A1 내지 A17중 적어도 한가지와 연쇄 불균형한 이중대립인자 마커를 포함한다.

이런 단편은 '프리-스탠딩(free-standing)', 다시 말하면 다른 폴리뉴클레오티드의 일부 또는 다른 폴리뉴클레오티드에 융합되지 않거나, 또는 좀더 큰 단일 폴리뉴클레오티드에서 일부분 또는 영역을 형성할 수도 있다. 실제로, 이들 단편중 몇몇 단편은 좀더 큰 폴리뉴클레오티드상에 존재한다.

선택적으로, 이런 단편은 적어도 8, 10, 12, 15, 18, 20, 25, 35, 40, 50, 70, 80, 100, 250, 500 또는 1000개의 연속 뉴클레오티드로 구성될 수 있다. 일단의 적합한 단편은 본원에서 밝힌 CanIon 유전자의 이중대립인자 A1 내지 A17 중 적어도 한가지 또는 이의 보체를 보유한다.

2-폴리펩티드

또한, 본원 발명은 돌연변이된 CanIon 단백질을 비롯하여, 전술한 폴리펩티드의 변이체, 단편, 유사체, 유도체에 관한다.

변이는 하나이상의 아미노산 잔기가 보존된 또는 비-보존된 아미노산 잔기로 치환되고 이런 치환된 아미노산 잔기가 유전자 코드에 의해 인코드되는 또는 인코드되지 않는 변이; 2) 하나이상의 아미노산 잔기가 치환기를 보유하는 변이; 3) 돌연변이된 CanIon이 다른 화합물, 예를 들면 폴리펩티드의 반감기를 증가시키는 화합물(예, 폴리에틸렌 글리콜)과 융합되는 변이; 또는 4) 부가된 아미노산이 돌연변이된 CanIon, 예를 들면 리더나 분비 서열 또는 돌연변이된 CanIon 혹은 전단백질(preprotein) 서열의 정제에 사용되는 서열에 융합되는 변이일 수 있다.

폴리펩티드 단편은 임의의 폴리펩티드 서열, 바람직하게는 CanIon 유전자에 의해 인코드되는 폴리펩티드 및 이의 변이체의 일부분과 완전히 상동한 서열을 보유하는 폴리펩티드이다.

본원 발명에 따른 폴리펩티드의 아미노산 서열에서 아미노산 치환의 경우에, 하나 또는 수 개의 아미노산은 '등가의' 아미노산으로 치환될 수 있다. 본원에서 '등가의' 아미노산은 펩티드 화학분야의 당업자가 폴리펩티드의 이차 구조와 수치(hydropathic) 성질이 실질적으로 변화지 않을 것으로 예상할 수 있을 정도로, 유사한 특성을 보유하는 아미노산을 대체할 수 있는 임의의 아미노산을 의미한다. 일반적으로, 다음의 아미노산 군은 등가의 변화를 나타낸다: (1) Ala, Pro, Gly, Glu, Asp, Gln, Asn, Ser, Thr; (2) Cys, Ser, Tyr, Thr; (3) Val, Il3, Leu, Met, Ala, Phe; (4) Lys, Arg, His; (5) Phe, Tyr, Trp, His.

본원 발명에 따른 변형된 CanIon 펩티드 분자의 특정 구체예에는 단백분해에 저항하는 펩티드 분자, 다시 말하면 -CONH-펩티드 결합이 변화되고 (CH2NH) 환원된 결합, (NHCO) 레트로 인벌소 결합, (CH2-O) 메틸렌-옥시 결합, (CH2-S) 티오메틸렌 결합, (CH2CH2) 카바 결합, (CO-CH2) 케토메틸렌 결합, (CHOH-CH2) 하이드록시에틸렌 결합, (N-N) 결합, E-알센 결합 또는 -CH=CH 결합으로 치환되는 펩티드가 포함되지만, 이들에 국한되지 않는다. 또한, 본원 발명에는 사람 CanIon 폴리펩티드 또는 적어도 한 개의 펩티드 결합이 전술한 바와 같이 변형된 이의 단편이나 변이가 포함된다.

이런 단편은 '프리-스탠딩(free-standing)', 다시 말하면 다른 폴리펩티드의 일부 또는 다른 폴리펩티드에 융합되지 않거나, 또는 좀더 큰 단일 폴리펩티드에서 일부분 또는 영역을 형성할 수도 있다. 실제로, 몇몇 단편은 좀더 큰 단일 폴리펩티드에 포함된다.

본원 발명에 다른 폴리펩티드 단편의 대표적 예는 대략 5, 6, 7, 8, 9, 10 내지 15, 10 내지 20, 15 내지 40, 또는 30 내지 55개의 아미노산을 갖는 폴리펩티드이다. 적절하게는, CanIon 단백질에서 적어도 한 개의 아미노산 돌연변이를 갖는 폴리펩티드 단편이다.

핵산 또는 폴리펩티드간 상동성

본원에서 '서열 상동성 %'와 '동질성 %'는 동일한 의미를 갖고 폴리뉴클레오티드와 폴리펩티드간 비교를 의미하며 비교 윈도우에서 2개의 최적으로 정렬된 서열을 비교하여 결정하는데, 여기서 비교 윈도우상의 폴리뉴클레오티드 또는 폴리펩티드 서열의 일부분은 두 서열의 최적 정렬을 위하여 기준 서열(부가 또는 결실을 보유하지 않음)과 비교하여 부가 또는 결실(즉, 갭)을 보유할 수도 있다. 서열 상동성 %는 양 서열에서 동일한 핵산 염기 또는 아미노산 잔기가 발생하는 위치의 총 수를 측정하여 대합된 위치의 총수를 산출하고, 대합된 위치의 총수를 비교 윈도우에서 위치의 총수로 나누고, 결과에 100을 곱하여 계산한다. 동질성은 당분야에 공지된 다양한 서열 비교 알고리즘과 프로그램을 활용하여 평가한다. 이런 알고리즘과 프로그램에는 TBLASTN, BLASTIP, FASTA, TFASTA, CLUSTALW (Pearson and Lipman, 1988; Altschul et al., 1990; Thompson et al., 1994; Higgins et al., 1996; Altschul et al., 1990; Altschul et al., 1993)이 포함되지만, 이들에 국한되지 않는다. 특히 적절한 구체예에서, 단백질과 핵산 서열 동질성은 당분야에 공지된 Basic Local Alignment Tool('BLAST')를 활용하여 평가한다(Karlin and Altschul, 1990; Altschul et al., 1990 1993, 1997).

특히, 5종류의 특정 BLAST 프로그램이 다음의 작업을 실행하는데 활용되고 있다:

(1) BLASTP와 BLAST3은 단백질 서열 데이터베이스와 아미노산 쿼리 서열을 비교한다;

(2) BLASTN은 뉴클레오티드 서열 데이터베이스와 뉴클레오티드 쿼리 서열을 비교한다;

(3) BLASTX는 단백질 서열 데이터베이스와 쿼리 뉴클레오티드 서열(양 가닥)의 6개-프레임 추정 변역 산물을 비교한다;

(4) TBLASTN은 6개의 해독틀(양 가닥)에서 번역된 뉴클레오티드 서열 데이터베이스와 쿼리 단백질 서열을 비교한다;

(5) TBLASTX는 뉴클레오티드 서열 데이터베이스의 6개-프레임 번역과 뉴클레 오티드 쿼리 서열의 6개-프레임 번역을 비교한다.

BLAST 프로그램은 유사한 분절을 동정하여 상동성 서열을 동정하는데, 이는 이후 쿼리 아미노 서열이나 핵산 서열과 단백질이나 핵산 서열 데이터베이스로부터 얻은 시험 서열간 '하이-스코어링 분절 쌍'이라고 한다. 하이-스코어링 분절 쌍은 가급적 당분야에 공지된 스코어링 매트릭스로 동정(즉, 정렬)한다. 이용되는 스코어링 매트릭스는 바람직하게는 BLOSUM62 매트릭스(Gonnet et al., 1992; Henikoff and Henikoff, 1993)이다. 조금 덜 선호되긴 하지만, PAM 또는 PAM250 매트릭스 역시 이용할 수 있다(Schwartz and Dayhoff, eds., 1978). BLAST 프로그램은 확인된 모든 하이-스코어링 분절 쌍의 통계학적 유의성을 평가하고, 가급적 사용자-특정된 유의성 한계, 예를 들면 사용자-특정된 동질성 %를 만족시키는 분절을 선택한다. 적절하게는, 하이-스코어링 분절 쌍의 통계학적 유의성은 Karlin의 통계 유의성 공식(Karlin and Altschul, 1990)을 활용하여 평가한다.

BLAST 프로그램은 디폴트 변수 또는 사용자가 제공한 파라미터로 이용할 수 있다.

엄격한 혼성화 조건

본원 발명에 따른 혼성화 핵산을 정의하기 위한 엄격한 혼성화 조건은 다음과 같다:

혼성화 단계는 6 x SSC 완충액, 5 x 덴하르트 용액, 0.5% SDS, 100㎍/㎖ 연어 정자 DNA의 존재하에 65℃에서 달성한다.

혼성화 단계에 이어 4가지 세척 단계를 실시한다:

- 2 x SSC와 0.1% SDS 완충액의 존재하에 65℃에서 5분동안 2회 세척;

- 2 x SSC와 0.1% SDS 완충액의 존재하에 65℃에서 30분동안 1회 세척;

- 2 x SSC와 0.1% SDS 완충액의 존재하에 65℃에서 10분동안 1회 세척.

이들 혼성화 조건은 대략 20개 뉴클레오티드로 구성되는 핵산 분자에 적합하다. 전술한 혼성화 조건은 당업자에게 공지된 기술로 원하는 핵산의 길이에 따라 조정할 수 있다. 가령, 적합한 혼성화 조건은 Hames와 Higgins(1985)의 책자에서 공개한 내용에 따라 조정할 수 있다.

CanIon 유전자의 게놈 서열

본원 발명은 CanIon의 게놈 서열에 관한다. 본원 발명에는 CanIon 유전자, SEQ ID No 1 내지 3 서열을 포함하는 CanIon 게놈 서열, 이들의 상보적 서열, 이들의 단편과 변이체가 포함된다. 이들 폴리뉴클레오티드는 정제, 분리 또는 재조합될 수 있다.

또한, 본원 발명에는 SEQ ID No 1 내지 3의 뉴클레오티드 서열과 적어도 70, 75, 80, 85, 90 또는 95% 뉴클레오티드 상동성을 갖는 뉴클레오티드 서열, 이의 상보적 서열 또는 이의 단편을 포함하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드가 포함된다. 일반적으로, SEQ ID No 1 내지 3 뉴클레오티드 서열과 관련된 뉴클레오티드 차이는 핵산 전체에 무작위로 분포한다. 그럼에도 불구하고, 바람직한 핵산은 SEQ ID No 1 내지 3 뉴클레오티드 서열과 관련된 뉴클레오티드 차이가 엑손에 포함되는 코딩 서열 외부에 위치하는 핵산이다. 이들 핵산, 이들의 단편과 변이체는 테스트 시료에서 CanIon 유전자 사본의 존재를 검출하는, 또는 대안으로 CanIon 서열에서 표적 뉴클레오티드 서열을 증폭하기 위한 올리고뉴클레오티드 프라이머 또는 프로브로 사용될 수 있다.

본원 발명의 다른 목적은 전술한 엄격한 혼성화 조건하에 SEQ ID No 1 내지 3 뉴클레오티드 서열, 이의 상보적 서열 또는 변이체와 혼성화되는 분리된, 정제된, 재조합된 핵산에 관한다. 적절한 구체예에서, 상기 정제된, 분리된 또는 재조합된 핵산은 사람 CanIon 유전자의 폴리뉴클레오티드와 특이적으로 혼성화되고, 좀더 적절하게는 상기 핵산은 사람 CanIon 유전자의 뉴클레오티드와 혼성화할 수 있지만 쥐 CanIon 유전자의 핵선 서열과는 혼성화될 수 있다.

본원 발명의 특히 적합한 핵산에는 SEQ ID No 1 내지 3의 적어도 12, 15, 18, 20, 25, 35, 40, 50, 70, 80, 100, 250, 500 또는 1000개의 연속 뉴클레오티드를 보유하는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함된다. 임의 크기와 서열의 핵산 단편 역시 본 섹션에서 밝힌 폴리뉴클레오티드를 포함할 수 있다.

CanIon 게놈 핵산은 44개의 엑손으로 구성된다. SEQ ID No 1 내지 3에서 엑손 위치는 하기 표 B에 상세하게 기술한다.

표 B

엑손	SEQ ID No 1에서 위치		인트론	SEQ ID No 1에서 위치	
	시작	종결		시작	종결
1	2001	2026	1	2027	19187
2	19188	19334	2	19335	22995
3	22996	23178	3	23179	39730
4	39731	39814	4	39815	41336
5	41337	41476	5	41477	41564
6	41565	41693	6	41694	73012
7	73013	73167			

엑손	SEQ ID No 2에서 위치		인트론	SEQ ID No 2에서 위치	
	시작	종결		시작	종결
8	43726	43868	7	43869	43997
9	43998	44102	8	44103	52092
10	52093	52179	9	52180	77567
11	77568	77699	10	77700	98225
12	98226	98393	11	98394	106566
13	106567	106758	12	106759	144108
14	144109	144246	13	144247	159793
15	159794	159868	14	159869	191291
16	191292	191428	15	191429	192966
17	192967	193108	16	193109	211539
18	211540	211613	17	211614	225005
19	225006	225107	18	225108	225543
20	225544	225613	19	225614	228449
21	228450	228541	20	228542	228629
22	228630	228752	21	228753	231288
23	231289	231345	22	231346	231588
24	231589	231709	23	231710	231812
25	231813	231944	24	231945	232899
26	232900	233067	25	233068	235354
27	235355	235459			

엑손	SEQ ID No 3에서 위치		인트론	SEQ ID No 3에서 위치	
	시작	종결		시작	종결
28	3895	4001	26	4002	9610
29	9611	9731	27	9732	9815
30	9816	9914	28	9915	15775
31	15776	15869	29	15870	16381
32	16382	16488	30	16489	16696
33	16697	16771	31	16772	17933
34	17934	18053	32	18054	23643
35	23644	23712	33	23713	24927
36	24928	25076	34	25077	25912
37	25913	26006	35	26007	30766
38	30767	30899	36	30900	31560
39	31561	31676	37	31677	34043
40	34044	34201	38	34202	37492
41	37493	37643	39	37644	39651
42	39652	39801	40	39802	41562
43	41563	41680	41	41681	44130
44	44131	45841			

따라서, 본원 발명에는 CanIon 유전자의 44개 엑손에서 선택되는 뉴클레오티드 서열 또는 이의 보체 서열을 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함된다. 또한, 본원 발명은 CanIon 유전자의 적어도 2개 엑손의 조합을 포함하는 정제된, 분리된, 재조합된 핵산을 제공하는데, 여기서 폴리뉴클레오티드는 SEQ ID No 1 내지 3에서와 동일한 순서로 상기 핵산에서 5' 말단에서 3' 말단으로 정렬된다.

인트론 1은 엑손 1과 엑손 2 사이에 위치하는 뉴클레오티드 서열을 나타낸다. 인트론의 위치는 표 A에 상세하게 기술한다. 따라서, 본원 발명에는 CanIon 유전자의 43개 인트론에서 선택되는 뉴클레오티드 서열 또는 이의 보체 서열을 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함된다.

본 섹션의 타이틀이 'CanIon 유전자의 게놈 서열'이긴 하지만, 양 측면에서 CanIon의 게놈 서열에 접하는 또는 2개이상의 이런 게놈 서열간 임의의 크기와 서열의 핵산 단편 역시 본 섹션에서 밝힌 폴리뉴클레오티드로 구성될 수 있다.

CanIon cDNA 서열

CanIon 유전자의 발현은 적어도 한가지 종류의 mRNA를 생산하는데, 이의 핵산 서열은 SEQ ID No 4에 제시한다.

본원 발명의 다른 목적은 SEQ ID No 4 뉴클레오티드 서열, 이의 상보적 서열, 대립유전자 변이체, 단편을 포함하는 분리된, 정제된, 재조합된 핵산이다. 적절하게는, 본원 발명의 폴리뉴클레오티드에는 SEQ ID No 4의 서열을 포함하는 분리된, 정제된, 재조합된 CanIon cDNA가 포함된다. 특히 적절하게는, 본원 발명의 핵산에는 SEQ ID No 4의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000개 뉴클레오티드의 연속 스팬(span) 또는 이의 보체를 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함된다. 적절한 구체예에서, 상기 스팬은 A12 내지 A16에서 선택되는 CanIon-관련된 이중대립인자 마커를 포함한다.

또한, 본원 발명은 SEQ ID No 4 폴리뉴클레오티드와 적어도 95% 뉴클레오티드 상동성, 바람직하게는 99% 뉴클레오티드 상동성, 좀더 바람직하게는 99.5% 뉴클레오티드 상동성, 가장 바람직하게는 99.8% 뉴클레오티드 상동성을 갖는 폴리뉴클레오티드, 이의 상보적 서열 또는 생물학적 활성 단편을 포함하는 분리되거나 정제된 핵산에 관한다.

본원 발명의 다른 목적은 전술한 엄격한 조건하에 SEQ ID No 4 폴리뉴클레오티드와 혼성화되는 폴리뉴클레오티드, 이의 상보적 서열, 변이체 또는 생물학적 활성 단편을 포함하는 분리된, 정제된, 재조합된 핵산에 관한다.

본원 발명의 다른 목적은 SEQ ID No 5의 적어도 6개 아미노산의 연속 스팬을 포함하는 CanIon 폴리펩티드를 인코드하는 분리된, 정제된, 재조합된 폴리뉴클레오티드에 관하는데, 여기서 상기 연속 스팬은 SEQ ID No 5의 아미노산 위치 277, 338, 574, 678, 680, 683, 691, 692, 695, 696, 894, 1480, 1481, 1483, 1484, 1485, 1630, 1631, 1632, 1636, 1660, 1667, 1707, 1709중 적어도 1, 2, 3, 5 또는 10개를 보유한다. 또한, SEQ ID No 5의 아미노산 서열, 이의 유도체 또는 생물학적 활성 단편을 인코드하는 분리된, 정제된, 재조합된 폴리뉴클레오티드 및 SEQ ID No 5의 아미노산 서열과 적어도 80, 85, 90, 95, 98, 99, 99.5 또는 99.8% 상동한 CanIon 폴리펩티드를 인코드하는 분리된,

정제된, 재조합된 폴리뉴클레오티드가 본원 발명에 포함된다.

SEQ ID No 4의 cDNA는 SEQ ID No 4의 1번 위치 뉴클레오티드에서 시작하여 65번 위치 뉴클레오티드에서 종결되는 5'-UTR 영역을 포함한다. SEQ ID No 4의 cDNA는 5283번 위치 뉴클레오티드에서 시작하여 6799 위치 뉴클레오티드에서 종결되는 3'-UTR 영역을 포함한다.

결과적으로, 본원 발명은 CanIon cDNA 5'UTR의 뉴클레오티드 서열, 이의 상보적 서열 또는 대립유전자 변이체를 포함하는 분리된, 정제된, 재조합된 핵산에 관한다. 또한, 본원 발명은 CanIon cDNA 3'UTR의 뉴클레오티드 서열, 이의 상보적 서열 또는 대립유전자 변이체를 포함하는 분리된, 정제된, 재조합된 핵산에 관한다.

본 섹션의 타이틀이 'CanIon cDNA 서열'이긴 하지만, 양 측면에서 CanIon의 게놈 서열에 접하는 또는 2개이상의 이런 게놈 서열간 임의의 크기와 서열의 핵산 단편 역시 본 섹션에서 밝힌 폴리뉴클레오티드로 구성될 수 있다.

코딩 영역

CanIon 개방해독틀은 SEQ ID No 4 cDNA의 상응하는 mRNA에 포함된다. 좀더 구체적으로, 효과적인 CanIon 코딩 서열(CDS)은 SEQ ID No 4의 66번 위치 뉴클레오티드(ATG 코돈의 첫 번째 뉴클레오티드)와 5282번 위치 뉴클레오티드(TGA 코돈의 마지막 뉴클레오티드)간 영역을 포함한다. 본원 발명은 SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 12, 15, 20, 25, 30, 40, 50, 100개의 연속 스팬을 포함하는 폴리펩티드를 인코드하는 분리된, 정제된, 재조합된 폴리뉴클레오티드에 관한다.

CanIon 유전자의 코딩 서열을 포함하는 전술한 폴리뉴클레오티드는 적합한 발현 신호의 조절하에 원하는 숙주 세포 또는 숙주 미생물에서 발현될 수 있다. 발현 신호는 본원 발명의 CanIon 유전자상의 조절 영역에 포함된 발현 신호이거나 또는 외인성 조절 핵산 서열이다. 이런 폴리뉴클레오티드는 적합한 발현 신호하에 발현 및/또는 증폭을 위한 벡터에 삽입될 수도 있다.

CanIon의 조절 서열

전술한 바와 같이, CanIon 유전자의 게놈 서열은 상기 유전자의 44개 엑손을 포함하는 CanIon 코딩 영역에 접하는 비-코딩 5-측부영역과 비-코딩 3-측부영역에 조절 서열을 보유한다.

5'와 3' 조절 영역으로부터 유래된 폴리뉴클레오티드는 테스트 시료에서 CanIon 뉴클레오티드 서열 또는 이의 단편의 1개이상 사본의 존재를 검출하는데 유용하다.

CanIon에 포함된 5'조절 영역의 프로모터 활성은 후술한 바와 같이 평가할 수 있다.

SEQ ID No 1의 유관한 생물학적 활성 폴리뉴클레오티드 단편 또는 변이체를 동정하기 위하여, 당업자는 마커 유전자(즉, 베타 갈락토시다제, 클로람페니콜 아세틸 트랜스퍼라제 등)를 보유하는 재조합 벡터의 용도를 설명하는 Sambrook et al.(Sambrook, 1989)을 참조하는데, 상기 마커 유전자의 발현은 SEQ ID No 1의 생물학적 활성 폴리뉴클레오티드 단편 또는 변이체의 조절하에 검출될 수 있다. CanIon 유전자의 첫 번째 엑손의 상류에 위치한 게놈 서열은 적절한 프로모터 리포터 벡터, 예를 들면 pSEAP-Basic, pSEAP-Enhancer, pβgal-Basic, pβgal-Enhancer, pEGEP-1 프로모터 리포터 벡터(Clontech); 또는 pGL2-Basic, pGL3-Basic 무-프로모터 루시페라제 리포터 유전자 벡터(Promega)로 클론시킨다. 간단히 말하면, 이들 프로모터 리포터 벡터 각각은 쉽게 분석가능한 단백질, 예를 들면 배출된 알칼리성 포스파타제, 루시페라제, β-갈락토시다제 또는 녹색 형광 단백질을 인코드하는 리포터 유전자의 상류에 위치하는 다중 클로닝 부위를 보유한다. CanIon 코딩 영역의 상류 서열은 양 방향으로 리포터 유전자 상류의 클로닝 부위로 삽입되 고 적합한 숙주 세포에 도입된다. 리포터 단백질의 수준은 클로닝 부위에 삽입체가 없는 벡터로부터 얻어진 수준과 비교 분석한다. 대조군 벡터에서 수준과 비교하여 삽입체를 보유하는 벡터에서 상승된 발현 수준의 존재는 삽입체에서 프로모터의 존재를 암시한다. 필요한 경우, 상류 서열은 약한 프로모터 서열로부터 전사 수준을 증가시키기 위한 인헨서를 보유하는 벡터에 클론할 수 있다. 삽입체가 없는 벡터에서 관찰되는 수준보다 현저하게 높은 발현 수준은 프로모터 서열이 삽입된 상류 서열에 존재한다는 것을 암시한다.

상류 게놈 DNA에서 프로모터 서열은 통상적인 기술, 예를 들면 엑소뉴클레아제 III 또는 적합한 제한 엔도뉴클레아제 절단으로 상류 DNA에서 중첩(nested) 5' 및/또는 3' 결실을 작제함으로써 추가로 정의한다. 생성된 결실 단편은 프로모터 리포터 벡터에 삽입하여 결실이 프로모터 활성을 감소 또는 소멸시키는 지를 확인할 수 있다(Coles et al.(1998)). 이런 방식으로, 경계를 정의할 수 있다. 원하는 경우, 프로모터에서 잠재적인 개별 조절 부위는 개별적으로 또는 복합적으로 프로모터내에서 잠재적인 전사 인자를 소멸시키는 특정 부위 돌연변이유발 또는 링커 스캐닝으로 확인할 수 있다. 전사 수준에 대한 이들 돌연변이의 효과는 돌연변이를 프로모터 리포터 벡터상의 클로닝 부위로 삽입하여 확인할 수 있다. 이런 유형의 분석법은 당분야에 공지된 것이다(WO 97/17359, US Patent No. 5, 374,544; EP 582

796; US Patent No. 5,698,389; US 5,643,746; US Patent No. 5,502,176; US Patent 5,266,488;).

CanIon 유전자의 프로모터의 강도와 특이성은 상이한 유형의 세포와 조직에 서 CanIon 프로모터에 동작가능하게 연결된 검출가능한 폴리뉴클레오티드의 발현 수준을 통하여 평가할 수 있다. 검출가능한 폴리뉴클레오티드는 사전결정된 올리고뉴클레오티드 프로브와 특이적으로 혼성화되는 폴리뉴클레오티드, 또는 CanIon 폴리펩티드, 이의 단편 혹은 변이체를 비롯한 검출가능한 단백질을 인코드하는 폴리뉴클레오티드일 수 있다. 이런 유형의 분석법은 당분야에 공지된 것이다(US Patent No. 5,502,176; US Patent No. 5,266,488). 이런 방법중 일부는 하기에 상술한다.

CanIon 코딩 영역의 5' 말단과 3' 말단에 위치하는 조절 요소를 보유하는 폴리뉴클레오티드는 관심있는 이질성 폴리뉴클레오티드의 전사와 번역 활성을 조절하는데 효과적으로 사용할 수 있다.

따라서, 본원 발명은 5'와 3' 조절 영역에서 선택되는 폴리뉴클레오티드를 포함하는 분리되거나 정제된 핵산, 이의 상보적 서열, 생물학적 활성 단편 또는 변이체에 관한다. 한 측면에서, '5' 조절 영역'은 SEQ ID No 1의 1 내지 2000번 위치에 존재하는 뉴클레오티드에 존재한다. 한 측면에서, '3' 조절 영역은 SEQ ID No 3의 45842 내지 47841번 위치에 존재하는 뉴클레오티드 서열에 존재한다.

또한, 본원 발명은 5'와 3' 조절 영역에서 선택되는 폴리뉴클레오티드와 적어도 95% 뉴클레오티드 상동성, 바람직하게는 99% 뉴클레오티드 상동성, 좀더 바람직하게는 99.5% 뉴클레오티드 상동성, 가장 바람직하게는 99.8% 뉴클레오티드 상동성을 갖는 폴리뉴클레오티드를 포함하는 분리되거나 정제된 핵산, 이의 상보적 서열, 변이체 또는 생물학적 활성 단편에 관한다.

본원 발명의 다른 목적은 전술한 엄격한 조건하에 5'와 3' 조절 영역의 뉴클레오티드 서열, 이의 상보적 서열, 변이체 또는 생물학적 활성 단편에서 선택되는 폴리뉴클레오티드와 혼성화되는 폴리뉴클레오티드를 포함하는 분리된, 정제된, 재조합된 핵산에 관한다.

적절하게는, 5' 조절 영역의 단편은 대략 1500 내지 1000개, 바람직하게는 500개, 좀더 바람직하게는 400개, 이보다 좀더 바람직하게는 300개, 가장 바람직하게는 200개의 뉴클레오티드를 보유한다.

적절하게는, 3' 조절 영역의 단편은 적어도 50, 100, 150, 200, 300, 400개의 염기이다.

SEQ ID No 1과 3의 '생물학적 활성' 폴리뉴클레오티드 유도체는 재조합 숙주 세포에서 재조합 폴리펩티드 또는 재조합 폴리뉴클레오티드를 발현하기 위한 조절 영역으로 기능하는 상기 폴리뉴클레오티드의 단편을 포함하는 또는 상기 단편으로 구성되는 폴리뉴클레오티드이다. 이는 인헨서 또는 억제체로 작용할 수 있다.

본원 발명에서, 조절 폴리뉴클레오티드가 전사와 번역 조절 정보를 담고 있는 뉴클레오티드 서열을 보유하고 이런 서열이 원하는 폴리펩티드를 인코드하는 뉴클레오티드 서열 또는 원하는 폴리뉴클레오티드에 '동작가능하게 연결되는' 경우에 이런 핵산 또는 폴리뉴클레오티드는 재조합 폴리펩티드 또는 재조합 폴리뉴클레오티드를 발현하기 위한 조절 영역으로 '기능한다'.

본원 발명의 조절 뉴클레오티드는 Sambrook et al.(1989)에서 밝힌 바와 같이 적절한 제한 효소를 이용한 절단으로 SEQ ID No 1 내지 3의 뉴클레오티드 서열 로부터 만들 수 있다. 조절 폴리뉴클레오티드는 엑소뉴클레아제 효소, 예를 들면 Bal31(Wabiko et al., 1986)로 SEQ ID No 1 내지 3을 절단하여 만들 수도 있다. 또한, 이들 조절 폴리뉴클레오티드는 본원에서 밝힌 핵산 화학적 합성으로 만들 수 있다.

본원 발명에 따른 조절 폴리뉴클레오티드는 원하는 숙주 세포 또는 숙주 미생물에서 코딩 서열을 발현하는데 사용할 수 있는 재조합 발현 벡터의 일부분이 될 수 있다. 본원 발명에 따른 재조합 발현 벡터는 본원에서 제시한다.

적절하게는, 본원 발명의 5'-조절 폴리뉴클레오티드는 CanIon cDNA의 5-비번역 영역(5'-UTR), 이의 생물학적 활성 단편 또는 변이체를 포함한다.

적절하게는, 본원 발명의 3'-조절 폴리뉴클레오티드는 CanIon cDNA의 3-비번역 영역(3'-UTR), 이의 생물학적 활성 단편 또는 변이체를 포함한다.

본원 발명의 다른 목적은 다음과 같이 구성되는 분리된 또는 정제된 핵산을 제공하는 것이다:

a) 다음에서 선택되는 조절 뉴클레오티드 서열을 포함하는 핵산:

(i) 5' 조절 영역의 폴리뉴클레오티드 또는 이의 상보적 서열을

포함하는 뉴클레오티드 서열 ;

(ii) 5' 조절 영역의 뉴클레오티드 서열과 적어도 95%의

뉴클레오티드 상동성을 갖는 폴리뉴클레오티드 또는 이의 상보성

서열을 포함하는 뉴클레오티드 서열;

(iii) 엄격한 혼성화 조건하에 5' 조절 영역의 뉴클레오티드 서열과

혼성화되는 폴리뉴클레오티드 또는 이의 상보적 서열을 포함하는

뉴클레오티드 서열 ;

(iv) (i),(ii), (iii) 뉴클레오티드의 생물학적 활성 단편 또는 변이체

b) 상기 (a) 정의된 핵산에 동작가능하게 연결되고 원하는 폴리펩티드 또는 핵산을 인코드하는 폴리뉴클레오티드;

c) 선택적으로, 3'-조절 폴리뉴클레오티드, 바람직하게는 CanIon 유전자의 3'-조절 폴리뉴클레오티드를 포함하는 핵산.

상기 정의된 핵산의 특정 구체예에서, 핵산은 CanIon cDNA의 5'-비번역 영역(5'-UTR), 이의 생물학적 활성 단편 또는 변이체를 포함한다.

상기 정의된 핵산의 다른 특정 구체예에서, 핵산은 CanIon cDNA의 3'-비번역 영역(3'-UTR), 이의 생물학적 활성 단편 또는 변이체를 포함한다.

5'-비번역 영역(5'-UTR)의 조절 폴리뉴클레오티드, 이의 생물학적 활성 단편 또는 변이체는 원하는 폴리펩티드 또는 폴리뉴클레오티드를 인코드하는 폴리뉴클레오티드의 5'-말단에 동작가능하게 연결된다.

3'-비번역 영역(3'-UTR)의 조절 폴리뉴클레오티드, 이의 생물학적 활성 단편 또는 변이체는 원하는 폴리펩티드 또는 폴리뉴클레오티드를 인코드하는 폴리뉴클레오티드의 3'-말단에 동작가능하게 연결된다.

전술한 핵산에 의해 인코드되는 원하는 폴리펩티드는 다양한 특성 또는 기원을 보일 수 있는데, 여기에는 진핵과 원핵 기원의 단백질이 포함된다. CanIon 조 절 영역의 조절하에 발현되는 폴리펩티드에는 박테리아, 진균 또는 바이러스 항원이 포함된다. 또한, 진핵 단백질, 예를 들면 '하우스 키핑' 단백질과 같은 세포내 단백질; 수용체와 같은 막-결합된 단백질; 내인성 매개물질(예, 사이토킨)과 같은 배출된 단백질이 포함된다. 원하는 폴리펩티드는 CanIon 단백질, 특히 SEQ ID No 5의 아미노산 서열로 구성되는 단백질, 이의 단편 또는 변이체일 수 있다.

전술한 폴리뉴클레오티드에 의해 인코드되는 원하는 핵산, 일반적으로 RNA 분자는 원하는 코딩 폴리뉴클레오티드, 예를 들면 CanIon 코딩 서열에 상보적이고, 따라서 안티센스 폴리뉴클레오티드로 유용하다.

이런 폴리뉴클레오티드는 숙주 세포 또는 숙주 미생물에서 원하는 폴리펩티드 또는 원하는 핵산을 발현하기 위한 재조합 발현 벡터에 포함될 수 있다. 이런 폴리뉴클레오티드를 포함하는데 적합한 재조합 벡터는 전술한 바 있다.

폴리뉴클레오티드 구조체

본원에서 ''폴리뉴클레오티드 구조체'와 '재조합 폴리뉴클레오티드'는 동일한 의미를 갖고 인위적으로 설계되고 최초 고유 환경에서 연속 뉴클레오티드 서열로 발견되지 않는 적어도 2개의 뉴클레오티드 서열을 포함하는 선형이나 환형의 분리되거나 정제된 폴리뉴클레오티드를 의미한다.

재조합 세포 숙주와 유전자도입 동물에서 시간적, 공간적 CanIon 유전자 발현을 조절하는 DNA 구조체

세포 수준과 다세포 미생물 수준에서 CanIon 단백질 합성의 결핍으로 인한 생리적, 표현형 결과를 조사하기 위하여, 본원 발명에는 CanIon 게놈 서열이나 cDNA의 특정 대립유전자 및 SEQ ID No 1 내지 4의 CanIon 뉴클레오티드 서열이나 이의 단편과 관련하여 하나이상 염기의 치환, 결실 혹은 부가를 보유하는 상기 게놈 서열이나 cDNA 사본의 조건 발현을 가능하게 하는 DNA 구조체와 재조합 벡터가 포함되는데, 여기서 염기 치환, 결실, 부가는 엑손, 인트론 또는 조절 서열, 바람직하게는 5'-조절 서열, CanIon 게놈 서열의 엑손 또는 SEQ ID No 4의 CanIon cDNA에 위치한

다. 적절한 구체예에서, CanIon 서열은 본원 발명에 따른 이중대립인자 마커를 포함한다. 적절한 구체예에서, CanIon 서열은 본원 발명에 따른 이중대립인자 마커, 바람직하게는 이중대립중 마커 A1 내지 A17중 한가지를 포함한다.

본원 발명에는 본원 발명에서 밝힌 폴리뉴클레오티드중 임의 하나로 구성되는 재조합 벡터가 포함된다. 좀더 구체적으로, 본원 발명에 다른 폴리뉴클레오티드 구조체는 'CanIon 유전자의 게놈 서열' 섹션, 'CanIon cDNA 서열' 섹션, '코딩 영역' 섹션, '올리고뉴클레오티드 프로브와 프라이머' 섹션에서 밝힌 폴리뉴클레오티드중 임의 하나로 구성될 수 있다.

바람직한 첫 번째 DNA 구조체는 Gossen et al.(1992, 1995)과 Furth et al.(1994)에서 밝힌 바와 같이 CanIon 유전자 발현을 조절하는 대장균(*E. Coli*) 트랜스포손 Tn10으로부터 유래된 테트라사이클린 저항성 오페론 *tet* 에 기초한다. 이런 DNA는 Tn10으로부터 유래된 7개의 *tet* 오퍼레이터 서열(tetop)을 보유하는데, 이들 서열은 CanIon 유전자의 최소 프로모터 또는 5'-조절 서열에 융합되고, 상기 최소 프로모터 또는 CanIon 조절 서열은 센스 혹은 안티센스 올리고뉴클레오티드 또는 CanIon 폴리펩티드 혹은 이의 펩티드 단편을 비롯한 폴리펩티드를 코딩하는 폴리뉴클레오티드에 동작가능하게 연결된다. 이런 DNA 구조체는 동일 세포가 프로모터, 예를 들면 HCMVIE1 인헨서/프로모터 또는 MMTV-LTR의 조절하에 단순 포진 바이러스의 바이러스 단백질 VP16의 활성화 도메인에 융합된 야생형(tTA) 또는 돌연변이(rTA) 억제체를 코딩하는 뉴클레오티드 서열을 포함하는 경우에, 관심있는 뉴클레오티드 서열에 대한 조건 발현 시스템으로 기능한다. 실제로, 본원 발명의 바람직한 DNA 구조체는 *tet* 오퍼레이터 서열을 포함하는 폴리뉴클레오티드 및 tTA 또는 rTA 억제체를 코딩하는 서열을 포함하는 폴리뉴클레오티드로 구성된다.

특정 구체예에서, 조건 발현 DNA 구조체는 돌연변이 테트라사이클린 억제체 rTA를 인코드하는 서열을 보유하는데, 관심있는 폴리뉴클레오티드의 발현은 테트라사이클린의 부재하에서는 침묵하고 이의 존재하에서는 유도된다.

상동성 재조합을 가능하게 하는 DNA 구조체: 치환 벡터

바람직한 두 번째 DNA 구조체는 5'-말단에서 3'-말단까지 다음과 같이 구성된다: (a) CanIon 게놈 서열에 포함되는 제 1 뉴클레오티드 서열; (b) 양성(positive) 선별 마커, 예를 들면 네오마이신 저항성(neo) 마커를 포함하는 뉴클레오티드 서열; (c) CanIon 게놈 서열에 포함되고 게놈에서 (a) CanIon 뉴클레오티드 서열의 하류 게놈에 위치하는 제 2 뉴클레오티드 서열.

적절한 구체예에서, 상기 DNA 구조체는 (a) 뉴클레오티드 서열 상류 또는 (c) 뉴클레오티드 서열 하류에 위치하는 음성(negative) 선별 마커로 구성된다. 적절하게는, 음성 선별 마커는 티미딘 키나아제(tk) 유전자(Thomas et al., 1986), 히그로마이신 베타 유전자(Te Riele et al., 1990), hprt 유전자(Van der Lugt et al., 1991; Reid et al., 1990) 또는 디프테리아 독소 A 단편(Dt-A) 유전자(Nada et al., 1993; Yagi et al. 1990)로 구성된다. 적절하게는, 양성 선별 마커는 CanIon 단백질을 인코드하는 서열을 깨뜨리기 위하여 CanIon 엑손 서열내에 위치시킨다. 이들 치환 벡터는 Thomas et al.(1986; 1987); Mansour et al.,(988); Koller et al.(1992)에서 기술한다.

제 1과 제 2 뉴클레오티드 서열 (a)와 (c)는 CanIon 조절 서열, 인트론 서열, 엑손 서열, 또는 조절 서열, 인트론 서열 및/또는 엑손 서열을 포함하는 서열내에 독립적으로 위치시킬 수 있다. 뉴클레오티드 서열 (a)와 (c)의 크기는 1 내지 50kb, 바람직하게는 1 내지 10kb, 좀더 바람직하게는 2 내지 6kb, 가장 바람직하게는 2 내지 4kb이다.

상동성 재조합을 가능하게 하는 DNA 구조체: Cre-LoxP 시스템

이들 신규한 DNA 구조체는 P1 파지의 부위 특이적 재조합 시스템을 이용한다. P1 파지는 34 염기쌍 loxP와 특이적으로 상호작용하는 Cre라고 하는 재조합효소(recombinase)를 보유한다. loxP 부위는 8bp 보존된 서열에 의해 분리되는 13bp의 팰린드롬 서열 2개로 구성된다(Hoess et al., 1986). 동일 배향된 2개의 loxP 부위사이에서 Cre 효소에 의한 재조합은 DNA 단편의 결실을 초래한다.

상동성 재조합 기술과 짝으로 사용되는 Cre-loxP 시스템은 Gu et al.(1993, 1994)에서 처음으로 기술되었다. 간단히 말하면, 게놈의 표적 위치로 삽입되는 뉴클레오티드 서열은 동일 배향되고 재조합 게놈으로부터 절단되는 뉴클레오티드 서열의 각 말단에 위치하는 적어도 2개의 loxP 부위를 보유한다. 절단은 재조합 세 포 숙주의 핵내에 재조합효소(*Cre*) 효소의 존재를 필요로 한다. 재조합효소 효소는 (a) Araki et al.(1995)에서 밝힌 바와 같이 Cre 효소를 원하는 세포에 직접 주사하여 또는 Baubonis et al.(1993)에서 밝힌 바와 같이 상기 효소의 세포로의 리포펙션으로 이런 효소를 함유하는 배양 배지에서 재조합 세포 숙주를 배양하고; (b) 재조합 세포 숙주에서 기능하는 프로모터에 동작가능하게 연결된 *Cre* 코딩 서열을 포함하는 벡터로 숙주 세포를 트랜스펙션시키고, 상기 프로모터는 상기 세포 숙주에서 선택적으로 유도되고, 상기 벡터는 Gu et al.(1993)와 Sauer et al.(1988)에서 밝힌 바와 같이 재조합 세포 숙주에 도입되고; (c) 재조합 숙주 세포에서 기능하는 프로모터에 동작가능하게 연결된 *Cre* 코딩 서열을 포함하는 폴리뉴클레오티드를 세포 숙주의 게놈에 도입하고, 상기 프로모터는 상기 숙주 세포에서 선택적으로 유도되고, 상기 폴리뉴클레오티드는 Gu et al.(1994)에서 밝힌 바와 같이 무작위 삽입 또는 상동성 재조합으로 세포 숙주의 게놈에 삽

입된다.

특정 구체예에서, 상동성 재조합으로 CanIon 유전자에 삽입되는 서열을 포함하는 벡터는 선별 마커가 동일 배향의 loxP가 측면에 접하도록 작제하는데, Cre 효소 처리로 선별 마커는 제거하면서 상동성 재조합에 의해 삽입된 CanIon 서열은 존속하도록 하는 것이 가능하다. 한번 더, 2개의 선택 마커가 필요하다: 재조합 현상을 선별하는 양성 선별 마커 및 상동성 재조합 현상을 선별하는 음성 선별 마커. Cre-loxP 시스템을 이용한 벡터와 방법은 Zou et al.(1994)에서 기술한다.

따라서, 바람직한 본원 발명의 세 번째 DNA 구조체는 5'-말단에서 3'-말단까지 다음과 같이 구성된다: (a) CanIon 게놈 서열을 포함하는 제 1 뉴클레오티드 서 열; (b) 양성 선별 마커를 인코드하는 폴리뉴클레오티드를 포함하는 뉴클레오티드 서열, 상기 뉴클레오티드 서열은 재조합효소에 의해 인식되는 부위, 예를 들면 loxP 부위를 정의하는 2개의 서열을 추가로 포함하고, 상기 두 부위는 동일 배향으로 위치하고; (c) CanIon 게놈 서열로 구성되고 게놈에서 (a) CanIon 뉴클레오티드 서열의 하류에 위치하는 제 2 뉴클레오티드 서열.

적절하게는, 재조합효소에 의해 인식되는 부위, 예를 들면 loxP 부위를 정의하는 서열은 (b) 뉴클레오티드 서열에서, 조건 절단이 요구되는 뉴클레오티드 서열에 인접하는 위치에서 존재한다. 특정 구체예에서, 2개의 loxP 부위는 상동성 재조합이후 원하는 시점에 절단을 가능하게 하기 위하여 각각의 양성 선별 마크 서열에 위치시킨다.

전술한 세 번째 DNA 구조체를 이용한 방법의 적절한 구체예에서, 재조합효소에 의해 인식되는 두 부위, 바람직하게는 두 개의 loxP 부위가 측면에서 접하는 폴리뉴클레오티드 단편의 절단은 Gu et al.(1994)에서 밝힌 바와 같이 재조합 숙주 세포 게놈에서 프로모터 서열, 바람직하게는 유도성 프로모터, 좀더 바람직하게는 조직-특이적 프로모터 서열, 가장 바람직하게는 유도성이고 조직-특이적인 프로모터 서열에 동작가능하게 연결된 Cre 효소를 인코드하는 서열의 존재로 인하여 원하는 시점에 실행된다.

재조합 숙주 세포의 게놈에서 Cre 효소의 존재는 Gu et al.(1994)에서 밝힌 바와 같이 2마리의 유전자도입 동물, 전술한 바와 같이 loxP 부위를 보유하는 CanIon-유래된 서열을 포함하는 제 1 유전자도입 동물 및 적합한 프로모터 서열에 동작가능하게 연결된 *Cre* 코딩 서열을 포함하는 제 2 유전자도입 동물의 번식으로부터 유래할 수 있다.

또한, Cre 효소 발현의 공간적-시간적 조절은 Graham(1995)과 Kanegae et al.(1995)에서 밝힌 바와 같이 Cre 유전자를 보유하여 세포의 감염 또는 장기의 *in vivo* 감염, Cre 효소의 전달을 가능하게 하는 아데노바이러스 기초한 벡터로 달성할 수 있다.

전술한 DNA 구조체는 본원 발명의 뉴클레오티드 서열, 바람직하게는 CanIon 게놈 서열이나 CanIon cDNA 서열, 가장 바람직하게는 CanIon 게놈이나 cDNA 서열의 변형된 사본을 표적 게놈의 사전결정된 위치에 도입하여 표적 유전자의 변형된 사본 발생 또는 상동성 재조합이 일어나도록 할 만큼 상동한 다른 사본에 의한 표적 유전자 사본의 치환(적중 상동성 재조합)을 유도하는데 사용할 수 있다. 특정 구체예에서, 전술한 DNA 구조체는 적어도 1개의 이중대립인자 마커를 포함하는 CanIon 게놈 서열 또는 CanIon cDNA 서열을 도입하는데 사용할 수 있다.

핵 안티센스 DNA 구조체

본원 발명의 벡터를 함유하는 다른 조성물은 핵산 서열 SEQ ID No 4의 올리고뉴클레오티드 단편, 바람직하게는 CanIon 유전자의 개시 코돈을 포함하는 단편을 상응하는 CanIon 유전자의 발현을 저해하는 안티센스 도구로 포함한다. 본원 발명에 다른 안티센스 폴리뉴클레오티드를 이용하는 바람직하는 방법은 Sczakiel et al.(1995)에서 밝힌 과정 또는 WO 95/24223에서 밝힌 과정이다.

적절하게는, 안티센스 도구는 CanIon mRNA의 5'말단에 상보적인 폴리뉴클레 오티드(15-200bp)에서 선택된다. 한 구체예에서, 원하는 표적 유전자의 상이한 부분에 상보적인 개별 안티센스 폴리뉴클레오티드의 조합을 사용한다.

본원 발명에 따른 바람직한 안티센스 폴리뉴클레오티드는 번역 개시 코돈 ATG 또는 접합 부위를 보유하는 CanIon mRNA의 서열에 상보적이다. 본원 발명에 따른 다른 바람직한 안티센스 폴리뉴클레오티드는 CanIon mRNA의 접합 부위에 상보적이다.

적절하게는, 본원 발명의 안티센스 폴리뉴클레오티드는 Liu et al.(1994)에서 밝힌 바와 같이 자가-절단 리보자임 서열로 치환되는 3' 폴리아데닐화 신호를 갖고, 따라서 RNA 중합효소 II 전사체는 3' 말단에 poly(A)가 없고, 이들 안티센스 폴리뉴클레오티드는 핵으로 배출될 수 없다. 적절한 구체예에서, 이들 CanIon 안티센스 폴리뉴클레오티드는 Eckner et al.(1991)에서 밝힌 구조와 같이 절단된 전사체를 3'-5' 엑소뉴클레아제 분해로부터 안정화시키는 히스톤 스템-루프 구조를 리보자임 카세트에 포함한다.

올리고뉴클레오티드 프로브와 프라이머

CanIon 유전자로부터 유래된 폴리뉴클레오티드는 테스트 시료에서 SEQ ID No 1내지 4와 6 뉴클레오티드 서열, 이의 단편, 보체 또는 변이체의 1개이상 사본의 존재를 검출하는데 유용하다.

본원 발명의 특히 바람직한 프로브와 프라이머는 SEQ ID No 1 내지 3의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000또는 2000개 뉴클레오티드의 연속 스팬 또는 이의 보체를 포함하는 분리된, 정 제된, 재조합된 폴리뉴클레오티드를 보유한다. 본원 발명의 좀더 바람직한 프로브와 프라이머는 상기 연속 스팬이 A1 내지 A17에서 선택되는 이중대립인자 마커를 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드를 보유한다.

본원 발명의 다른 목적은 SEQ ID No 4의 뉴클레오티드 서열, 이의 보체 서열, 대립유전자 변이체, 단편을 포함하는 분리된, 정제된, 재조합된 핵산이다. 이에 더하여, 본원 발명의 바람직한 프로브와 프라이머는 SEQ ID 4를 포함하는 정제된, 분리된, 재조합된 CanIon cDNA를 보유한다. 본원 발명의 특히 바람직한 프로브와 프라이머는 SEQ ID 4의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000개, 2000, 3000, 4000, 5000 또는 6000개의 연속 뉴클레오티드 또는 이의 보체를 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드를 보유한다. 본원 발명의 좀더 바람직한 프로브와 프라이머는 SEQ ID 4의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000개, 2000, 3000, 4000, 5000 또는 6000개 뉴클레오티드의 연속 스팬 또는 이의 보체를 포함하고 상기 연속 스팬이 A1 내지 A17에서 선택되는 이중대립인자 마커를 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드를 보유한다.

다른 구체예에서, 본원 발명의 프로브와 프라이머는 SEQ ID 6의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300 또는 400개 뉴클레오티드의 연속 스팬 또는 이의 보체를 포함하는 분리된, 정제된, 재조합된 폴리뉴클레오티드를 보유한다. 적절한 구체예에서, SEQ ID No 6의 연속 스팬은 이중대립인자 마커 A18을 포함한다.

따라서, 본원 발명은 엄격한 혼성화 조건하에 SEQ ID No 1과 3의 사람 CanIon 뉴클레오티드 서열, 이의 상보적 서열 또는 변이체와 특이적으로 혼성화되는 핵산 프로브에 관한다.

한 구체예에서, 본원 발명에는 SEQ ID No 1 내지 4와 6중 한가지의 8 내지 50개 뉴클레오티드의 연속 스팬으로 구성되는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함되는데, 여기서 상기 스팬은 서열에 CanIon-관련된 이중대립인자 마커를 보유하고, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형의 이중대립인자 마커에서 선택된다. 선택적으로, 연속 스팬은 18 내지 35개 뉴클레오티드이고 이중대립인자 마커는 상기 폴리뉴클레오티드 중심으로부터 4개 뉴클레오티드 이내에 존재한다; 선택적으로, 연속 스팬은 25개 뉴클레오티드이고 이중대립인자 마커는 상기 폴리뉴클레오티드 중심에 존재한다; 선택적으로, 연속 뉴클레오티드의 3'말단은 상기 뉴클레오티드의 3'말단에 존재한다; 선택적으로, 연속 뉴클레오티드의 3'말단은 상기 뉴클레오티드의 3'말단에 존재하고 이중대립인자 마커는 상기 폴리뉴클레오티드의 3'말단에 존재한다. 적절한 구체예에서, 프로브는 다음의 서열: P1 내지 P18과 이의 상보적 서열에서 선택되는 서열로 구성된다.

다른 구체예에서, 본원 발명에는 SEQ ID No 1 내지 4의 8 내지 50개 뉴클레오티드의 연속 스팬 또는 이의 보체로 구성되는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함되는데, 여기서 상기 연속 스팬의 3'말단은 상기 폴리뉴클레오티드의 3'말단에 위치하고, 상기 폴리뉴클레오티드의 3'말단은 서열에서 CanIon-관련 된 이중대립인자 마커 상류의 20개 뉴클레오티드이내에 위치한다; 선택적으로, CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형의 이중대립인자 마커에서 선택된다; 선택적으로, 상기 폴리뉴클레오티드의 3'말단은 서열에서 CanIon-관련된 이중대립인자 마커 상류의 1 뉴클레오티드에 위치한다; 선택적으로, 상기 폴리뉴클레오티드는 다음의 서열: D1 내지 D18과 E1 내지 E18에서 선택되는 서열로 구성된다.

다른 구체예에서, 본원 발명에는 다음의 서열: B1 내지 B17과 C1 내지 C17에서 선택되는 서열로 구성되는 분리된, 정제된, 재조합된 폴리뉴클레오티드가 포함된다.

다른 구체예에서, 본원 발명에는 혼성화 분석, 염기서열분석, SEQ ID No 1 내지 4와 6 또는 이의 보체에서 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드의 정체를 확인하기 위한 효소-기초한 미스매치 검출 방법에 사용되는 폴리뉴클레오티드 및 SEQ ID No 1 내지 4와 6 또는 이의 보체에서 CanIon-관련된 이중대립인자 마커를 포함하는 뉴클레오티드의 분절을 증폭하는데 사용되는 폴리뉴클레오티드가 포함된다.

본원 발명은 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드의 실체를 확인에서 분석, 바람직하게는 혼성화 분석, 염기서열분석, 마이크로염기서열분석 또는 효소-기초한 미스매치 검출 분석 및 CanIon-관련된 이중대립인자 마커를 포함하는 뉴클레오티드 분절의 증폭에서 본원 발명에 따른 폴리뉴클레오티드의 용도에 관한다.

본원 발명에 따른 프로브 또는 프라이머는 8 내지 1000개의 뉴클레오티드, 좀더 구체적으로, 적어도 12, 15, 18, 20, 25, 35, 40, 50, 60, 70, 80, 90, 100, 250, 500, 1000의 뉴클레오티드를 갖는다. 좀더 구체적으로, 이들 프로브와 프라이머의 길이는 8, 10, 15, 20 또는 30 내지 100, 바람직하게는 10 내지 50, 좀더 바람직하게는 15 내지 30개의 뉴클레오티드이다. 좀더 짧은 프로브와 프라이머는 표적 핵산 서열에 대한 특이성이 결핍되는 경향이 있고, 일반적으로 주형과 안정적인 하이브리드 복합체를 형성하는데 좀더 낮은 온도를 필요로 한다. 좀더 긴 프로브와 프라이머는 만드는데 비용이 많이 들고, 때때로 자가-혼성화되어 헤어핀 구조를 형성할 수 있다. 특정한 분석 조건하에 프라이머와 프로브의 적합한 길이는 당업자가 경험적으로 판단할 수 있다. 적합한 프로브 또는 프라이머는 서열 리스트에서 개별 위치가 각각 표 1, 2, 3에 제시된 P1 내지 P18과 이의 상보적 서열, B1 내지 B17, C1 내지 C17, D1 내지 D18, E1 내지 E18의 뉴클레오티드 서열에서 선택되는 폴리뉴클레오티드를 포함하는 핵산으로 구성된다.

안정한 하이브리드의 형성은 DNA의 융점(Tm)에 의존한다. Tm은 프라이머 또는 프로브의 길이, 용액의 이온력, G+C 함량에 따라 달라진다. 프라이머 또는 프로브의 G+C 함량이 높을수록 융점이 높은데, 그 이유는 G:C 쌍은 3개의 수소 결합을 갖는 반면 A:T는 2개의 수소 결합을 갖기 때문이다. 본원 발명의 프로브에서 GC 함량은 10 내지 75%, 바람직하게는 35 내지 60%, 좀더 바람직하게는 40 내지 55%이다.

프라이머와 프로브는 임의의 적절한 방법, 예를 들면 적합한 서열의 클로닝 과 제한 및 Narang et al.(1979)의 포스포디에스터 방법, Brown et al.(1979)의 포스포디에스터 방법, Beaucage et al.(1981)의 디에틸포스포라미디트 방법, EP 0 707 592에서 밝힌 고체 서포트 방법과 같은 방법에 의한 직접적인 화학적 합성으로 만들 수 있다.

검출 프로브는 일반적으로 핵산 서열 또는 하전되지 않은 핵산 유사체, 예를 들면 국제 특허 출원 WO 92/20702에 개시된 펩티드 핵산, 미국 특허 5, 185,44; 5,034,506, 5,142,047에서 기술된 모폴리노 유사체이다. 프로브는 dNTP가 추가로 부가될 수 없다는 점에서 '확장-불가능'하다. 본질적으로 유사체는 확장-불가능하고, 핵산 프로브는 3'말단을 변형시켜 하이드록실기가 더 이상 신장에 참여할 수 없도록 함으로써 확장-불가능하게 만들 수 있다. 가령, 프로브의 3'말단은 포획이나 검출 라벨로 기능화시켜 하이드록실기를 소모 또는 차단할 수 있다. 대안으로, 3'하이드록실기는 절단, 치환 또는 변형할 수 있다(U.S. Patent No. 07/049,061, filed April 19, 1993).

본원 발명의 폴리뉴클레오티드는 현미경적, 광화학적, 생화학적, 면역화학적 또는 화학적 수단으로 탐지할 수 있는 당분야에 공지된 라벨을 통합하여 표지할 수 있다. 가령, 유용한 라벨에는 방사성 물질(예, ^{32}P, ^{35}S, ^{3}H, ^{125}I), 형광 염료(예, 5-브로모디속시우리딘, 플루레신, 아세틸아미노플루오렌, 디옥시제닌) 또는 비오틴이 포함된다. 적절하게는, 폴리뉴클레오티드는 3'와 5'말단에 표지한다. 핵산 단편의 비-방사성 표지는 프랑스 특허 No. FR-7810975, Urdea et al(1988); Sanchez- Pescador et al.(1988)에서 기술한다. 이에 더하여, 본원 발명에 따른 프로브는 신호증폭을 가능하게 하는 구조적 특징을 가질 수 있는데, 이런 구조적 특징은 예로써 Urdea et al.(1991) 또는 EP 0 225 807 (Chiron)에서 밝힌 것과 같은 분지쇄 DNA 프로브이다.

라벨은 프라이머를 포획하여 고형 서포트에 프라이머 또는 프라이머 신장 산물, 예를 들면 증폭된 DNA의 고정을 용이하게 하는데 사용할 수 있다. 포획 라벨은 프라이머 또는 프로브에 부착되는데, 고체상 시약의 특이적 결합 구성원(예, 비오틴과 스트렙타비딘)과 결합쌍을 형성하는 특이적인 결합 구성원일 수 있다. 따라서, 폴리뉴클레오티드 또는 프로브에 의해 포함되는 라벨의 종류에 따라, 표적 DNA를 포획 또는 검출하는 것이 가능하다. 게다가, 본원에 제시된 폴리뉴클레오티드, 프라이머 또는 프로브는 포획 라벨로서 기능한다. 가령, 고체상 시약의 결합 구성원이 핵산 서열인 경우에, 프라이머 또는 프로브의 상보성 영역에 결합하여 상기 프라이머 또는 프로브를 고체상에 고정시키는 포획 라벨을 선택한다. 폴리뉴클레오티드 프로브 자체가 결합 구성원으로 기능하는 경우에, 이런 프로브가 표적과 상보적이지 않은 서열 또는 '꼬리'를 보유한다. 폴리뉴클레오티드 프라이머 자체가 포획 라벨로 기능하는 경우에, 프라이머의 적어도 일부는 고체상에서 핵산과 자유롭게 혼성화된다. DNA 표지 기술은 당업자에게 공지된 것이다.

본원 발명의 프로브는 다수의 목적에 유용하다. 이들은 특히, 게놈 DNA에 대한 서던 혼성화에 사용할 수 있다. 프로브는 PCR 증폭 산물을 검출하는데도 사용할 수 있다. 또한, 이들은 다른 기술을 활용하여 CanIon 유전자 또는 mRNA에서 미스매치를 검출하는데 사용할 수 있다. 이들은 예로써 노던 블랏에서 CanIon 유전자의 발현을 검출하는데 사용될 수도 있다.

본원 발명의 폴리뉴클레오티드, 프라이머, 프로브는 고체상에 쉽게 고정시킬 수 있다. 고체 서포트는 당업자에게 공지된 것으로, 여기에는 반응 트레이의 웰 벽, 테스트 튜브, 폴리스티렌 비드, 자성 비드, 니트로셀룰로오스 스트립, 막, 마이크로입자(예, 라텍스 입자), 양(또는 다른 동물) 적혈구, 두라사이트(duracyte) 등이 포함된다. 고형 서포트는 중요하지 않고 당업자가 선택할 수 있다. 따라서, 라텍스 입자, 마이크로입자, 자성이나 비-자성 비드, 막, 플라스틱 튜브, 마이크로역가 웰의 벽, 유리나 실리콘 칩, 양(또는 다른 동물) 적혈구, 두라사이트 모두 적합하다. 핵산을 고체상에 고정시키는데 적합한 방법에는 이온 결합, 소수성 결합, 공유 결합 등이 포함된다. 본원에서 고체 서포트는 불용성이거나 또는 후속 반응으로 불용화시킬 수 있는 물질을 의미한다. 고체 서포트는 포획 시약을 유인하고 고정시키는 내적 능력으로 선택할 수 있다. 대안으로, 고체상은 포획 시약을 유인하고 고정시키는 능력을 갖는 추가적인 수용체를 보유할 수 있다. 추가적인 수용체에는 포획 시약 자체 또는 포획 시약에 공액되는 하전된 물질과 관련하여 반대로 하전된 물질이 포함될 수 있다. 또 다른 대안으로, 수용체 분자는 고체 서포트에 고정되고(부착되고) 특이적인 결합 반응

을 통하여 포획 시약을 고정시킬 수 있는 능력을 갖는 임의의 특이적인 결합 구성원일 수 있다. 수용체 분자는 분석의 실시에 앞서 또는 실시동안 고체 서포트 물질에 대한 포획 시약의 간접적인 결합을 가능하게 한다. 따라서, 고체상은 당업자에게 공지된 플라스틱, 파생된 플라스틱, 자성이나 비-자성 물질, 테스트 튜브의 유리나 실리콘 표면, 마이크로역가 웰, 시트, 비드, 마이크로입자, 칩, 양(또는 다른 동물) 적혈구, 두라사이트 및 다른 형태일 수 있다. 본원 발명의 폴리뉴클레오티드는 고체 서포트에 개별적으로 또는 본원 발명에 따른 2, 5, 8, 10, 12, 15, 20 또는 25개의 개별 폴리뉴클레오티드의 집합으로 단일 고체 서포트에 고정 또는 부착시킬 수 있다. 이에 더하여, 본원 발명에 따른 폴리뉴클레오티드를 제외한 다른 폴리뉴클레오티드를 본원 발명에 따른 하나 또는 복수의 폴리뉴클레오티드와 동일한 고체 서포트에 부착시킬 수 있다.

결과적으로, 본원 발명은 시료에서 SEQ ID No 1 내지 4와 6, 이의 단편 또는 변이체에서 선택되는 뉴클레오티드 서열 및 이의 상보적 서열을 포함하는 핵산의 존재를 검출하는 방법에 관하는데, 상기 방법은 다음의 단계로 구성된다.

a) SEQ ID No 1 내지 4와 6의 핵산, 이의 단편 또는 변이체에서 선택되는 핵산에 포함되는 뉴클레오티드 서열 및 이의 상보적 서열과 혼성화될 수 있는 핵산 프로브 또는 복수의 핵산 프로브를 분석 시료와 접촉시키고;

b) 시료에서 프로브와 핵산간에 형성된 하이브리드 복합체를 검출한다.

또한, 본원 발명은 시료에서 SEQ ID No 1 내지 4와 6, 이의 단편 또는 변이체에서 선택되는 뉴클레오티드 서열 및 이의 상보적 서열을 포함하는 핵산의 존재를 검출하는 키트에 관하는데, 상기 키트는 다음과 같이 구성된다:

a) SEQ ID No 1 내지 4와 6, 이의 단편 또는 변이체에서 선택되는 핵산에 포함되는 뉴클레오티드 서열 및 이의 상보적 서열과 혼성화될 수 있는 핵산 프로브 또는 복수의 핵산 프로브;

b) 선택적으로, 혼성화 반응을 실행하는데 필요한 시약.

이런 검출 방법과 키트의 적절한 첫 번째 구체예에서, 핵산 프로브 또는 복수의 핵산 프로브는 검출가능한 분자로 표지한다. 상기 방법과 키트의 적절한 두 번째 구체예에서, 핵산 프로브 또는 복수의 핵산 프로브는 기질에 고정된다. 세 번째 구체예에서, 핵산 프로브 또는 복수의 핵산 프로브는 P1 내지 P18과 이의 상보적 서열, B1 내지 B17, C1 내지 C17, D1 내지 D18, E1 내지 E18의 뉴클레오티드 서열, 또는 A1 내지 A18과 이의 보체에서 선택되는 이중대립인자 마커에서 선택되는 서열을 포함한다.

올리고뉴클레오티드 분석

본원 발명에 따른 복수의 올리고뉴클레오티드 프라이머 또는 프로브를 포함하는 기질은 CanIon 유전자에서 표적 서열을 검출 또는 증폭하는데 사용할 수 있고, CanIon 유전자의 코딩이나 비-코딩 서열에서 돌연변이를 검출하는데 사용할 수 있다.

본원에 제시된 임의의 폴리뉴클레오티드는 고체 서포트에서 중복 지역이나 또는 무작위 위치에 부착시킬 수 있다. 대안으로, 본원 발명의 폴리뉴클레오티드는 정렬된 어레이에 부착시킬 수 있는데, 여기서 각 폴리뉴클레오티드는 다른 폴리뉴클레오티드의 부착 위치와 중복되지 않는 고체상의 개별 위치에 부착된다. 적절하게는, 폴리뉴클레오티드의 이런 정렬된 어레이는 '어드레스가능(addressable)'한데, 여기서 개별 위치는 기록되고 분석 과정의 일부로서 액세스할 수 있다. 전형적으로, 어드레스가능 폴리뉴클레오티드 어레이는 상이한 공지된 위치에서 기질의 표면에 결합하는 복수의 올리고뉴클레오티드프로브로 구성된다. 각 폴리뉴클레오티드의 정확한 위치에 관한 지식은 이들 '어드레스가능' 어레이가 혼성화 분석에 특히 유용하도록 한다. 당분야에 공지된 어드레스가능 어레이 기술은 본원 발명의 폴리뉴클레오티드에 활용할 수 있다. 이런 폴리뉴클레오티드 어레이의 특정 구체예는 Genechips ™ 으로, 이는 US Patent 5,143,854; PCT 국제 출원 WO 90/15070과 92/10092에 전반적으로 기술되어 있다. 이들 어레이는 기계적 합성 방법 또는 광석판술과 고체상 올리고뉴클레오티드 합성의 조합을 통합하는 광 유도된 합성 방법으로 만들 수 있다(Fodor et al., 1991). 고체 서포트에 올리고뉴클레오티드 어레이의 고정은 'Very Large Scale Immobilized Polymer Synthesis'(VLSIPS ™) 기술의 개발로 가능하게 되었는데, 여기서 프로브는 칩의 고체 표면에 고밀도 어레이로 고정된다. VLSIPS ™ 기술의 예는 US Patent 5,143,854; 5,412,087; PCT 출원 WO 90/15070; WO 92/10092; WO 95/11995에서 제공하는데, 이들 문헌은 광 유도된 합성 기술과 같은 기술을 통하여 올리고뉴클레오티드 어레이를 형성하는 방법을 기술한다. 고체 서포트에 고정된 뉴클레오티드 어레이를 제공하기 위한 전략의 마련하는데 있어, 혼성화 패턴과 서열 정보를 극대화시키기 위하여 칩에 올리고뉴클레오티드 어레이를 정렬하고 전시하기 위한 추가적인 프레젠테이션 전략이 개발되었다. 이런 프리젠테이션 전략의 예는 PCT 출원 WO 94/12305; WO 94/11530; WO 97/29212; WO 97/31256에서 개시한다.

본원 발명에 따른 올리고뉴클레오티드 어레이의 다른 구체예에서, 올리고뉴 클레오티드 프로브 매트릭스는 CanIon 유전자, 바람직하게는 이의 조절 영역에서 발생하는 돌연변이를 검출하는데 유용하게 사용할 수 있다. 이를 위해, 프

로브는 공지된 돌연변이(예, 하나 또는 복수 뉴클레오티드의 결실, 삽입 또는 부가)를 보유하는 유전자에 혼성화되는 뉴클레오티드 서열을 갖는 특이적인 프로브를 설계한다. 공지된 돌연변이는 예로써 Huang et al.(1996) 또는 Samson et al.(1996)이 활용한 기술에 따라 동정된 CanIon 유전자에서 돌연변이를 의미한다.

CanIon 유전자에서 돌연변이를 검출하는데 사용되는 다른 기술은 고밀도 DNA 어레이이다. 고밀도 DNA 어레이의 단위 요소를 구성하는 각 올리고뉴클레오티드 프로브는 CanIon 게놈 DNA 또는 cDNA의 특정 서열에 대합하도록 설계한다. 따라서, 표적 유전자서열에 상보적인 올리고뉴클레오티드를 포함하는 어레이는 야생 유전자 서열과 표적 서열의 상동성을 확인하고, 이의 함량을 측정하며 표적 서열과 CanIon 유전자의 기준 야생형 유전자 서열간 차이를 검출하는데 사용한다. 이런 설계에서 4L 타일드(tiled) 어레이는 한 세트의 4개 프로브(A, C, G, T), 바람직하게는 15개-뉴클레오티드 올리고머로 실행한다. 각 세트의 4개 프로브에서, 완전 보체는 미스매치된 프로브보다 좀더 강하게 혼성화된다. 결과적으로, 길이 L의 핵산 표적은 4L 프로브를 보유하는 타일드 어레이로 돌연변이를 스캔하고, 전체 프로브 세트는 공지된 야생형 참고 서열에서 모든 가능한 돌연변이를 포함한다. 15-mer 프로브 세트 타일드 어레이의 혼성화 신호는 표적 서열에서 단일 염기 변화로 섭동한다. 결과로써, 돌연변이 위치에 인접하는 프로브에서 신호의 특징적인 소멸 또는 '푸트프린트(footprint)'가 존재한다. 이런 기술은 Chee et al.(1996)에 기 술되었다.

결과적으로, 본원 발명은 적어도 하나의 전술한 폴리뉴클레오티드를 프로브와 프라이머로 포함하는 핵산 분자의 어레이에 관한다. 적절하게는, 본원 발명은 2개이상의 전술한 폴리뉴클레오티드를 프로브와 프라이머로 포함하는 핵산 분자의 어레이에 관한다.

본원 발명의 다른 목적은 P1 내지 P18, B1 내지 B17, C1 내지 C17, D1 내지 D18, E1 내지 E18, 이의 상보적 서열, 이의 적어도 8, 10, 12, 15, 18, 20, 25, 30 또는 40개의 연속 뉴클레오티드 단편 및 A1 내지 A18과 이의 보체에서 선택되는 이중대립인자 마커를 보유하는 적어도 한가지 서열에서 선택되는 서열을 1가지이상 포함하는 핵산 서열의 어레이에 관한다.

또한, 본원 발명은 P1 내지 P18, B1 내지 B17, C1 내지 C17, D1 내지 D18, E1 내지 E18, 이의 상보적 서열, 이의 적어도 8, 10, 12, 15, 18, 20, 25, 30 또는 40개의 연속 뉴클레오티드 단편 및 A1 내지 A18과 이의 보체에서 선택되는 이중대립인자 마커를 보유하는 적어도 2개의 서열에서 선택되는 서열을 2가지이상 포함하는 핵산 서열의 어레이에 관한다.

CanIon 단백질과 폴리펩티드 단편

본원에서 'CanIon 폴리펩티드'에는 본원 발명의 모든 단백질과 폴리펩티드가 포함된다. 본원 발명에 따른 폴리뉴클레오티드로 인코드되는 폴리펩티드 및 이런 폴리펩티드를 포함하는 융합 펩티드 역시 본원 발명에 포함된다. 본원 발명에는 SEQ ID No 5 서열로 구성되는 분리되거나 정제된 CanIon 단백질을 비롯하여 사람으 로부터 유래된 CanIon 단백질이 포함된다.

본원 발명은 SEQ ID No 1 내지 4와 6의 뉴클레오티드 서열, 이의 상보적 서열 또는 단편에 의해 인코드되는 폴리펩티드에 관한다.

본원 발명에는 SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개 아미노산, 좀더 바람직하게는 12, 15, 20, 25, 30, 40, 50, 100, 150, 200, 300, 400, 500, 700, 1000, 1200, 1400, 1600 또는 1700개 아미노산의 연속 스팬을 포함하는 분리된, 정제된, 재조합된 폴리펩티드가 포함된다. 다른 적절한 구체예에서, 아미노산의 연속 스팬은 CanIon 단백질 서열에서 아미노산의 결실, 부가, 치환 또는 절두를 비롯한 돌연변이 또는 기능적 돌연변이의 부위를 포함한다.

적절한 구체예에서, 본원 발명에는 SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개 아미노산, 좀더 바람직하게는 12, 15, 20, 25, 30, 40, 50, 100, 150, 200, 300, 400, 500, 700, 1000, 1200, 1400, 1600 또는 1700개 아미노산의 연속 스팬을 포함하는 분리된, 정제된, 재조합된 폴리펩티드가 포함되는데, 여기서 아미노산의 연속 스팬은 아미노산 위치 277, 338, 574, 678, 680, 683, 691, 692, 695, 696, 894, 1480, 1481, 1483, 1484, 1485, 1630, 1631, 1632, 1636, 1660, 1667, 1707, 1709중 적어도 1, 2, 3, 5 또는 10개를 보유한다. 적절하게는, SEQ ID No 5의 연속 스팬은 277 위치에서 알라닌 잔기; 338 위치에서 세린 잔기; 574 위치에서 발린 잔기; 678 위치에서 루이신 잔기; 680 위치에서 세린 잔기; 683 위치에서 트레오닌 잔기; 691 위치에서 히스티딘 잔기; 692 위치에서 세린 잔기; 695 위치에서 알세린 잔기; 696 위치에서 알라닌 잔기; 894 위치에서 이소루이신 잔기; 1480 위치에서 리신 잔기; 1481 위치에서 아르기닌 잔기; 1483 위치에서 글리신 잔기; 1484 위치에서 발린 잔기; 1485 위치에서 이소루이신 잔기; 1630 위치에서 아스파라긴 잔기; 1631 위치에서 세린 잔기; 1632 위치에서 메티오닌 잔기; 1636 위치에서 트레오닌 잔기; 1660 위치에서 알라닌 잔기; 1667 위치에서 페닐알라닌 잔기; 1707 위치에서 트레오닌 잔기; 1709 위치에서 알라닌 잔기를 포함한다. 이들 폴리펩티드중 한가지를 인코드하는 폴리뉴클레오티드를 또한 제시한다.

또한, 본원 발명에는 SEQ ID No 5의 아미노산 서열과 적어도 70, 75, 80, 85, 90, 95, 98 또는 99%의 아미노산 상동성을 갖는 아미노산 서열 또는 이의 단편을 포함하는 분리된, 정제된, 재조합된 폴리펩티드가 포함된다.

CanIon 단백질은 가급적 사람이나 포유동물 조직 시료로부터 분리되거나 또는 사람이나 포유동물 유전자로부터 발현된다. 본원 발명의 CanIon 폴리펩티드는 당분야에 공지된 통상적인 방법으로 만들 수 있다. 원하는 폴리펩티드를 인코드하는 폴리뉴클레오티드는 임의의 사용하기 편한 숙주에 적합한 발현 벡터에 결찰시킨다. 진핵과 원핵 숙주계가 재조합 폴리펩티드를 생성하는데 사용된다. 이후, 폴리펩티드는 용해된 세포 또는 배양 배지로부터 분리하고 원하는 목적에 필요한 수준으로 정제한다. 정제는 당분야에 공지된 임의의 기술, 예를 들면 분별 추출, 염 분획, 크로마토그래피, 원심분리 등으로 실시한다.

이에 더하여, 좀더 짧은 단백질 단편은 화학적 합성으로 만든다. 대안으로, 본원 발명의 단백질은 사람 또는 사람제외 동물의 세포 혹은 조직으로부터 추출한다. 단백질을 정제하는 방법을 당분야에 공지된 것으로, 여기에는 세정제나 카오트로픽(chaotropic) 약물을 이용한 입자 파괴, 분별 추출 및 이온 교환 크로마토그래피, 친화성 크로마토그래피, 밀도에 따른 침강, 겔 전기영동에 의한 폴리펩티드의 분리가 포함된다.

SEQ ID No 4를 비롯한 CanIon cDNA는 CanIon 단백질과 폴리펩티드를 발현하는데 사용될 수 있다. 발현되는 CanIon 단백질이나 폴리펩티드를 인코드하는 핵산은 통상적인 클로닝 기술을 활용하여 발현 벡터상의 프로모터에 동작가능하게 연결시킨다. 발현 벡터에서 CanIon 삽입체는 CanIon 단백질에 대한 전체 코딩 서열 또는 이의 일부분으로 구성될 수 있다.

발현 벡터는 당분야에 공지된 포유동물, 이스트, 곤충 또는 박테리아 발현 시스템이다. 상업적으로 가용한 벡터와 발현 시스템은 Genetics Institute(Cambridge, MA), Stratagenen(La Jolla, California), Promega (Madison, Wisconsin), Invitrogen (San Diego, California)을 비롯한 다양한 공급업체로부터 구할 수 있다. 원하는 경우, 발현을 강화하고 적절한 단백질 접힘을 촉진하기 위하여 서열의 코돈 환경과 코돈 쌍은 발현 벡터가 도입되는 특정 발현 미생물에 최적화시킨다(U.S. Patent No. 5,082,767).

한 구체예에서, cDNA의 폴리 A 신호를 통하여 CanIon cDNA의 전체 코딩 서열은 발현 벡터상의 프로모터와 동작가능하게 연결시킨다. 대안으로, CanIon 단백질의 일부분을 인코드하는 핵산에서 개시 부위로 작용하는 메티오닌이 부재하는 경우 에 통상적인 기술을 활용하여 핵산의 제 1 코돈에 측부에 개시 메티오닌을 도입할 수 있다. 유사하게, CanIon cDNA로부터 유래된 삽입체에 폴리 A 신호가 부재하는 경우에, 상기 서열은 BglI과 SalI 제한 엔도뉴클레아제 효소로 pSG5(Stratagene)로부터 폴리 A 신호를 절단하고 이를 포유동물 벡터 pXT1(Stratagene)에 통합시켜 구조체에 부가할 수 있다. pXT1은 LRT 및 Moloney 뮤린 백혈병 바이러스로부터 유래된 개그 유전자 일부를 포함한다. 구조체에서 LTR의 존재는 효과적이고 안정적인 트랜스펙션을 가능하게 한다. 상기 벡터는 단순 포진 티미딘 키나아제 프로모터와 선별가능 네오마이신 유전자를 보유한다. CanIon 단백질 또는 이의 일부분을 인코드하는 핵산은 SEQ ID No 5의 CanIon cDNA를 보유하고 5'프라이머에 통합된 Pst I 및 상응하는 cDNA 3' 프라이머의 5'말단에서 BglⅡ에 대한 제한 엔도뉴클레아제 서열을 보유하는 박테리아 벡터로부터, CanIon cDNA 또는 이의 일부분에 상보적인 올리고뉴클레오티드 프라이머를 이용한 PCR로 수득하는데, 여기서 CanIon 단백질 또는 이의 일부분을 인코드하는 서열은 폴리 A 신호와 관련하여 정확하게 위치하도록 담보한다. PCR 반응으로 수득된 정제된 단편은 Pst I로 절단하고 엑소뉴클레아제로 평활 말단이 되게 하며 BglⅡ로 절단하고 정제하며 폴리 A 신호를 보유하는 pXT1에 결찰시키고 BglⅡ로 절단한다.

결찰된 산물은 제품 명세서에 명시된 조건하에 리포펙틴(Life Technologies, Inc., Grand Island, New York)을 이용하여 NIH 3T3 세포로 트랜스펙션될 수 있다. 양성 형질감염체(transfectanct)는 트랜스펙션 세포를 600ug/㎖ G418 (Sigma, St. Louis, Missouri)에서 배양하여 선별한다.

상기 과정은 검출가능한 표현형을 담당하는 돌연변이 CanIon 단백질 또는 이의 일부분을 발현하는데 활용할 수도 있다.

발현된 단백질은 통상적인 정제 기술, 예를 들면 황산암모늄 침전 또는 크기나 전하에 기초한 크로마토그래피 분리로 정제될 수 있다. 핵산 삽입체에 의해 인코드되는 단백질은 표준 면역크로마토그래피 기술을 활용하여 정제할 수도 있다. 이런 과정에서, 발현된 CanIon 단백질 또는 이의 일부분을 함유하는 용액, 예를 들면 세포 추출액은 CanIon 단백질 또는 이의 일부분에 대한 항체를 갖는 칼럼에 가한다. 발현된 단백질은 면역크로마토그래피 칼럼에 결합하게 된다. 이후, 칼럼은 세척하여 비-특이적으로 결합된 단백질을 제거한다. 그 다음, 특이적으로 결합된 발현된 단백질은 칼럼으로부터 방출시키고 표준 기술로 회수한다.

CanIon 단백질 또는 이의 일부분을 발현을 확인하기 위하여, CanIon 단백질 또는 이의 일부분을 인코드하는 삽입체를 포함하는 발현 벡터를 보유하는 숙주 세포로부터 발현된 단백질은 삽입체가 없는 발현 벡터를 보유하는 숙주 세포에서 발현된 단백질과 비교할 수 있다. 삽입체가 없는 발현 벡터를 보유하는 세포의 시료에서는 나타나지 않는 삽입

체를 보유하는 발현 벡터를 보유하는 세포의 시료에서 밴드의 존재는 CanIon 단백질 또는 이의 일부분이 발현된다는 것을 암시한다. 일반적으로, 이런 밴드는 CanIon 단백질 또는 이의 일부분에서 예상되는 이동성을 갖는다. 하지만, 밴드는 글리코실레이션, 유비퀴티네이션 또는 효소적 절단과 같은 수식의 결과로 예상한 것과는 다른 이동성을 가질 수도 있다.

발현된 CanIon 단백질 또는 이의 일부분을 특이적으로 인식할 수 있는 항체 는 후술한다.

항체 생산이 불가능하면, CanIon 단백질 또는 이의 일부분을 인코드하는 핵산은 키메라 폴리펩티드를 활용한 정제 계획에 사용하기 위하여 설계된 발현 벡터에 통합시킨다. 이런 전략에서 CanIon 단백질 또는 이의 일부분을 인코드하는 핵산은 키메라의 다른 절반을 인코드하는 유전자와 맞추어 삽입한다. 키메라의 다른 절반은 β-글로빈 또는 니켈 결합 폴리펩티드 인코딩 서열이다. 이후, β-글로빈 또는 여기에 부착된 니켈에 대한 항체를 갖는 크로마토그래피 매트릭스는 키메라 단백질을 정제하는데 사용한다. β-글로빈 유전자 또는 니켈 결합 폴리펩티드와 CanIon 단백질 또는 이의 일부분사이에 프로테아제 절단 부위를 조작한다. 따라서, 키메라의 두 폴리펩티드는 프로테아제 절단에 의해 서로 분리된다.

β-글로빈 키메라 단백질을 만드는데 유용한 발현 벡터는 pSG5(Stratagene)로, 이는 토끼 β-글로빈을 인코드한다. 토끼 β-글로빈 유전자의 인트론 II는 발현된 전사체의 절단접합을 촉진하고, 구조체에 통합된 폴리아데닐화 신호는 발현의 수준을 증가시킨다. 이들 기술은 분자 생물학 분야의 당업자에게 공지된 것이다. 표준 방법은 Davis et al.,(1986)과 같은 방법 텍스트에 공개되어 있고, 많은 방법은 Stratagene, Life Technologies, Inc., 또는 Promega로부터 가용하다. 또한, 폴리펩티드는 In vitro Express ™ Translation Kit(Stratagene)과 같은 시험관내 번역 시스템을 활용하여 구조체로부터 만들 수 있다.

본원 발명의 CanIon 폴리펩티드에 결합하는 항체

CanIon 폴리펩티드 또는 전체 단백질은 발현된 CanIon 단백질 또는 이의 단편에 특이적으로 결합할 수 있는 항체를 만드는데 사용할 수 있다.

본원 발명의 항체 조성물은 SEQ ID No 5의 CanIon 단백질에 특이적으로 또는 선택적으로 결합할 수 있다. CanIon의 제 1 변이체에 특이적으로 결합하는 항체 조성물은 ELISA, RIA 또는 다른 항체-기초된 결합 분석에서 CanIon 단백질의 전장 제 2 변이체보다 CanIon 단백질의 전장 제 1 변이체에 대하여 적어도 5%, 10%, 15%, 20%, 25%, 50% 또는 100% 더 높은 결합 친화성을 보여야 한다. 적절한 구체예에서, 항체 조성물은 사람 CanIon 단백질에 특이적으로 결합할 수 있다.

적절한 구체예에서, 본원 발명은 SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 12, 15, 20, 25, 30, 40, 50, 100, 200, 300, 400, 500, 700 또는 1000개 아미노산의 연속 스팬을 포함하는 에피토프-포함 폴리펩티드에 특이적으로 결합하는 다클론이나 단클론 항체 조성물에 관한다. 적절한 구체예에서, 상기 연속 스팬은 SEQ ID No 5의 아미노산 위치 277, 338, 574, 678, 680, 683, 691, 692, 695, 696, 894, 1480, 1481, 1483, 1484, 1485, 1630, 1631, 1632, 1636, 1660, 1667, 1707, 1709중 적어도 1, 2, 3, 5 또는 10개를 보유한다.

임의의 CanIon 폴리펩티드 또는 전체 단백질은 본원에서 밝힌 발현된 CanIon 단백질 또는 이의 단편에 특이적으로 결합할 수 있는 항체를 만드는데 사용될 수 있다.

에피토프는 상기 에피토프에 독특한 공간적 입체형태를 제공하는 적게는 3개 의 아미노산으로 구성된다. 일반적으로, 에피토프는 적어도 6개의 이런 아미노산, 바람직하게는 적어도 8-10개의 이런 아미노산으로 구성된다. 적절한 구체예에서, 항원성 에피토프는 3 내지 50개의 아미노산으로 구성된다. 에피토프로 기능하는 단편은 임의의 통상적 수단으로 만들 수 있다. 에피토프는 디폴트 변수를 이용한 PROTEAN 컴퓨터 프로그램 (Version 4.0 Windows, DNASTAR, Inc., 1228 South Park Street Madison, WI)을 활용하여 Jameson-Wolf 항원 분석으로 확인할 수 있다.

본원 발명은 또한, 돌연변이된 CanIon 단백질의 에피토프로 구성되고 돌연변이된 CanIon 단백질에 특이적으로 결합하는 정제되거나 분리된 항체, 이의 단편이나 변이체에 관한다. 다른 적절한 구체예에서, 본원 발명은 CanIon 단백질의 10개이상 연속 아미노산으로 구성되고 형질 유발 돌연변이에 의해 인코드될 수 있는 적어도 1개의 아미노산을 보유하는 폴리펩티드에 결합할 수 있는 항체에 관한다.

항체 결합이 요구되는 것과 상이한 종류의 CanIon을 발현하는 사람제외 동물 또는 포유동물 및 CanIon을 발현하지 않는 동물(즉, CanIon 적중 동물)은 항체를 만드는데 특히 유용하다. CanIon 적중 동물은 CanIon의 단배질의 노출된 영역 대부분을 외래 항원으로 인식하고, 따라서 좀더 광범위한 CanIon 에피토프를 갖는 항체를 만들 수 있다. 게다가, 10 내지 30 아미노산을 갖는 좀더 작은 폴리펩티드는 CanIon 단백질중 임의 한가지에 대한 특이적인 결합을 달성하

는데 유용할 수 있다. 이에 더하여, 이런 항원 서열과 유사한 종류의 CanIon을 생산하는 동물의 체액 면역계는 고유 CanIon와 항원 서열간의 차이를 인식하고, 항원 서열에서 독특한 부위에 대한 항체를 생산하게 된다. 이런 기술은 CanIon 단백질중 임의 한가지에 특이 적으로 결합하는 항체를 수득하는데 특히 유용하다.

이런 프로토콜에 의해 만들어진 항체 제형은 생물 시료에서 항원-보유 물질의 농도를 측정하는 정량적 면역분석에 유용하다; 이들은 또한, 생물 시료에서 항원의 존재를 반-정량적으로 또는 정성적으로 확인하는 사용된다. 항체는 체내에서 단백질을 발현하는 세포를 죽이거나 또는 단백질의 수준을 감소시키는 치료요법적 조성물에 사용될 수도 있다.

본원 발명의 항체는 당분야에 공지된 방사성, 형광 또는 효소적 라벨로 표지할 수 있다.

따라서, 본원 발명은 생물 시료에서 본원 발명에 따른 CanIon 폴리펩티드의 존재를 특이적으로 검출하는 방법에 관하는데, 상기 방법은 다음의 단계로 구성된다:

a) SEQ ID No 5의 아미노산 서열로 구성되고 CanIon 폴리펩티드, 펩티드 단편 또는 이의 변이체에 특이적으로 결합하는 다클론이나 단클론 항체를 생물 시료와 접촉시키고;

b) 형성된 항원-항체 복합체를 검출한다.

또한, 본원 발명은 시료에서 본원 발명에 따른 CanIon 폴리펩티드의 시험관내 존재를 검출하는 진단 키트에 관하는데, 상기 키트는 다음과 같이 구성된다:

a) SEQ ID No 5의 아미노산 서열로 구성되고 CanIon 폴리펩티드에 특이적으로 결합하는 다클론이나 단클론 항체, 펩티드 단편 또는 이의 변이체;

b) 형성된 항원-항체 복합체의 검출을 가능하게 하는 시약, 상기 시약은 선 택적으로 라벨을 보유하거나, 또는 특히 전술한 단클론이나 다클론 항체가 자체적으로 표지되지 않은 경우에 임의의 표지된 시약으로 자신을 인식할 수 있다.

따라서, 본원 발명은 항체 및 T-세포 항원 수용체(TCR)에 관하는데, 이는 폴리펩티드, 좀더 구체적으로 IgG(IgG1, IgG2, IgG3, IgG4 포함), IgA(IgA1, IgA2 포함), IgD, IgE, IgM, IgY를 포함하는 이들에 국한되지 않는 본원 발명에 따른 폴리펩티드의 에피토프에 특이적으로 결합한다. 적절한 구체예에서, 항체는 본원 발명의 사람 항원 결합 항체 단편으로, 여기에는 Fab, Fab', F(ab)2, F(ab')2, Fd, 단일-사슬 Fvs(scFv), 단일-사슬 항체, 다이설파이드-결합된 Fvs(sdFv), V_L 이나 V_H 을 포함하는 단편이 포함되지만 이들에 국한되지 않는다. 항체는 조류와 포유동물을 비롯한 임의의 동물로부터 유래될 수 있다. 적절하게는, 항체는 사람, 뮤린, 토끼, 염소, 기니피그, 낙타, 말 또는 닭으로부터 유래된다.

단일-사슬 항체를 비롯한 항원-결합 항체 단편은 가변 영역 단독으로 또는 다음 영역: 힌지 영역, CH1, CH2, CH3 도메인의 전체 혹은 일부와 함께 구성될 수 있다. 본원 발명에는 가변 영역 및 힌지 영역, CH1, CH2, CH3 도메인의 임의 조합 역시 포함된다. 또한, 본원 발명에는 본원 발명의 폴리펩티드에 특이적으로 결합하는 키메라, 인화, 사람 단클론과 다클론 항체가 포함된다. 본원 발명에는 본원 발명의 항체에 대한 항-이디오타입 항체도 포함된다.

본원 발명의 항체는 단일특이적, 이중특이적, 삼중특이적 또는 다중특이적일 수 있다. 다중특이적 항체는 본원 발명에 따른 폴리펩티드의 상이한 에피토프에 특이적이거나 또는 본원 발명에 따른 폴리펩티드와 이질성 조성물, 예를 들면 이질성 폴리펩티드 또는 고체 서포트 물질 모두에 특이적일 수 있다(참조: WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, A. et al.(1991) J. Immunol. 147:60-69; US 특허 5,573,920, 4,474,893, 5,601,819, 4,714,681, 4,925,648; Kostelny, S.A. et al.(1992) J. Immunol. 148:1547-1553).

본원 발명의 항체는 항체에 인식되거나 또는 여기에 특이적으로 결합하는 본원 발명에 따른 폴리펩티드의 에피토프 또는 에피토프-보유 영역의 측면에서 설명한다. 본원 발명에 따른 단백질의 경우에, 항체는 본원 발명의 핵산에 의해 인코드되는 전장 단백질, 본원 발명의 핵산에 의해 인코드되는 성숙 단백질(즉, 신호 펩티드의 절단으로 생성된 단백질), 본원 발명의 핵산에 의해 인코드되는 신호 펩티드 또는 본원 발명의 임의 다른 폴리펩티드에 특이적으로 결합할 수 있다. 따라서, 이런 에피토프 또는 폴리펩티드 일부분을 보유하는 에피토프는 예로써 N-말단과 C-말단 위치, 연속 아미노산 잔기의 크기 등으로 기술할 수 있다. 본원 발명의 에피토프 또는 폴리펩티드에 특이적으로 결합하는 항체 역시 개별적으로 구별할 수 있다. 따라서, 본원 발명은 본원 발명에 따른 특정 폴리펩티드에 특이적으로 결합하고 이의 구별을 가능하게 하는 항체에 관한다.

본원 발명의 항체는 교차-반응성의 측면에서 기술할 수도 있다. 본원 발명에 따른 폴리펩티드의 임의 다른 유사체, 오트로그(ortholog) 또는 상동체와 특이적으로 결합하지 않는 항체는 본원 발명에 포함된다. 본원 발명의 폴리펩티드

와 95% 미만, 90% 미만, 85% 미만, 80% 미만, 75% 미만, 70% 미만, 65% 미만, 60% 미 만, 55% 미만, 50% 미만의 상동성(당분야에 공지된 방법으로 측정)을 갖는 폴리펩티드와 결합하지 않는 항체 역시 본원 발명에 포함된다. 또한, 본원 발명에는 엄격한 혼성화 조건하에 본원 발명의 폴리뉴클레오티드와 혼성화되는 폴리뉴클레오티드에 의해 인코드되는 폴리펩티드에만 결합 항체가 포함된다. 본원 발명의 항체는 결합 친화성의 측면에서 기술할 수도 있다. 바람직한 결합 친화성 항체는 분리 상수(Kd)가 $5X10^{-6}$ M, 10^{-6} M, $5X10^{-7}$ M, 10^{-7} M, $5X10^{-8}$ M, 10^{-8} M, $5X10^{-9}$ M, 10^{-9} M, $5X10^{-10}$, 10^{-10} M, $5X10^{-11}$ M, 10^{-11} M, $5X10^{-12}$ M, 10^{-12} M, $5X10^{-13}$ M, 10^{-13} M, $5X10^{-14}$ M, 10^{-14} M, $5X10^{-15}$ M, 10^{-15} M 미만인 항체이다.

본원 발명의 항체는 시험관내와 생체내 진단 방법과 치료 방법을 비롯하여 본원 발명에 따른 폴리펩티드를 정제, 검출, 표적하는 당분야에 공지된 방법에 사용된다. 가령, 항체는 생물 시료에서 본원 발명에 따른 폴리펩티드의 수준을 정량적으로, 정성적으로 측정하는 면역분석에 사용된다(Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988).

본원 발명의 항체는 단독으로 또는 다른 조성물과 함께 사용할 수 있다. 항체는 N- 또는 C-말단에서 이질성 폴리펩티드에 재조합으로 융합되거나 또는 폴리펩티드 혹은 다른 조성물에 화학적으로 공액(공유와 비-공유 공액 포함)될 수 있다. 가령, 본원 발명의 항체는 검출 분석에서 라벨로 유용한 분자 및 효과기 분자, 예를 들면 이질성 폴리펩티드, 약물 또는 독소에 재조합으로 융합되거나 또는 공액될 수 있다(WO 92/08495; WO 91/14438; WO 89/12624; US Patent 5,314,995; EP 0 396 387).

본원 발명의 항체는 당분야에 공지된 임의의 적절한 방법으로 만들 수 있다. 가령, 본원 발명의 폴리펩티드 또는 이의 항원성 단편은 동물에 투여하여 다클론 항체를 함유하는 혈청의 생산을 유도하는데 사용할 수 있다. '단클론 항체'는 하이브리도마 기술을 통하여 만들어진 항체에 국한되지 않는다. '항체'는 한 개이상의 결합 도메인을 포함하는 폴리펩티드 또는 폴리펩티드 군을 의미하는데, 여기서 항체 결합 도메인은 항체 분자에서 다양한 도메인의 접힙(folding)으로 생성되고 항원의 항원 결정기 외형에 상보적인 내부 표면 형태 및 전하 분포를 갖는 3차원 결합 공간을 형성하는데, 이는 항원과의 면역학적 반응을 가능하게 한다. '단클론 항체'는 진핵, 원핵 또는 파지 클론을 비롯한 단일 클론으로부터 유래된 항체를 의미하고, 이를 생산하는 방법을 의미하지 않는다. 단클론 항체는 하이브리도마 사용, 재조합, 파지 디스플레이 기술을 비롯한 당분야에 공지된 다양한 기술을 활용하여 만들 수 있다

하이브리도마 기술은 당분야에 공지된 것이다(Harlow et al.(1988); Hammerling, et. al.(1981)). Fab와 F(ab')2 단편은 예로써 파파인(Fab 단편 생성) 또는 펩신(F(ab')2 단편 생성)과 같은 효소를 이용한 단백질분해 절단으로 하이브리도마-생성된 항체로부터 만들 수 있다.

대안으로, 본원 발명의 항체는 재조합 DNA 기술 또는 당분야에 공지된 합성 화학을 적용하여 만들 수 있다. 가령, 본원 발명의 항체는 당분야에 공지된 다양한 파지 디스플레이 방법으로 만들 수 있다. 파지 디스플레이 방법에서, 기능적 항체 도메인은 파지 입자의 표면에 전시되고, 상기 파지 입자는 이들을 인코드하는 폴리뉴클레오티드 서열을 보유한다. 원하는 결합 특성을 갖는 파지는 항원, 특히 고체 표면이나 비드에 결합되거나 또는 포획된 항원으로 직접 선별함으로써 레파토리 또는 조합 항체 라이브러리(예, 사람 또는 뮤린)로부터 선택된다. 전형적으로, 이들 방법에 사용되는 파지는 파지 유전자 III 또는 유전자 VII 단백질에 재조합으로 융합된 Fab, Fv 또는 다이설파이드 안정화된 Fv 항체 도메인을 갖는 fd와 M13을 비롯한 섬유성 파지이다. 본원 발명의 항체를 만드는데 이용할 수 있는 파지 디스플레이 방법은 Brinkman U. et al.(1995); Ames, R.S. et al.(1995); Kettleborough, C.A.. et al. (1994); Persic, L. et al.(1997); Burton, D.R. et al.(1994); PCT/GB91/01134; WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; US Patents 5,698,426, 5,223,409, 5,403,484, 5,580,717, 5,427,908, 5,750,753, 5,821,047, 5,571,698, 5,427,908, 5,516,637, 5,780,225, 5,658,727, 5,733,743에서 개시한다.

상기 참고문헌에서 밝힌 바와 같이, 파지 선별후 파지로부터 항체 코딩 영역을 분리하는데, 이는 사람 항체 또는 임의의 다른 항원 결합 단편을 비롯하여 포유동물 세포, 곤충 세포, 식물 세포, 이스트, 박테리아에서 발현되는 전체 항체를 만드는데 사용할 수 있다. 가령, Fab, Fab'F(ab)2, F(ab')2 단편을 재조합으로 생산하는 기술은 당분야에 공지된 방법으로 활용할 수 있다(WO 92/22324; Mullinax, R.L. et al.(1992); Sawai, H. et al.(1995); Better, M. et al.(1988)).

단일-사슬 Fvs와 항체를 생산하는데 사용할 수 있는 기술은 U.S. Patents 4,946,778과 5,258,498; Huston et al.(1991); Shu, L. et al.(1993); Skerra, A. et al.(1988)에서 제시한다. 사람에서 항체의 생체내 용도 및 시험관내 검출 분석을 비롯한 일부 용도에서는 키메라, 인화 또는 사람 항체를 사용하는 것이 바람직할 수도 있다(Morrison(1985); Oi et al.,(1986); Gillies, S.D. et al.(1989); US Patent 5,807,715). 항체는 다양한 기술, 예를 들면 CDR-조직 이식(EP 0 239 400; WO 91/09967; US Patent 5,530,101; 5,585,089), 베니어링(veneering)이나 박피(resurfacing)(EP 0 592 106; EP 0 519 596; Padlan E.A.,(1991); Studnicka G.M. et al.(1994); Roguska M.A. et al.(1994)), 사슬 셔플링(shuffling)(US Patent 5,565,332)으로 인화시킬 수 있다. 사람 항체는 전술한 파지 디스플레이 방법을 비롯하여 당분야에 공지된 다양한 방법으로 만들 수 있다(US Patents 4,444,887, 4,716,111, 5,545,806, 5,814,318;

WO 98/46645; WO 98/50433; WO 98/24893; WO 96/34096; WO 96/33735; WO 91/10741).

또한, 본원 발명에는 본원 발명의 폴리펩티드에 재조합으로 융합된 또는 화학적으로 공액된(공유와 비-공유 공액 포함) 항체가 포함된다. 항체는 본원 발명의 폴리펩티드를 제외한 다른 항원에 특이적일 수도 있다. 가령, 항체는 본원 발명의 폴리펩티드를 특정 세포 표면 수용체에 특이적인 항체에 융합 또는 공액시킴으로써 시험관내 또는 생체내에서 본원 발명의 폴리펩티드가 특정 세포형으로 향하도록 하는데 사용할 수 있다. 본원 발명의 폴리펩티드에 융합된 또는 공액된 항체 역시 당분야에 공지된 방법으로 시험관내 면역분석 및 정제에 사용할 수 있다(Harbor et al. supra and WO 93/21232; EP 0 430 095; Naramura M. et al.(1994); US Patent 5,474,981; Gillies, S.O. et al.(1992); Fell, H.P. et al.(1991)).

또한, 본원 발명은 변이 영역을 제외한 다른 도메인에 융합된 또는 공액된 본원 발명의 폴리펩티드을 포함하는 조성물에 관한다. 가령, 본원 발명의 폴리펩티드는 항체 Fc 영역 또는 이의 일부분에 융합 또는 공액될 수 있다. 본원 발명의 폴리펩티드에 융합된 항체 일부분은 힌지 영역, CH1 도메인, CH2 도메인, CH3 도메인, 전체 도메인의 임의 조합 또는 이들의 일부분으로 구성될 수 있다. 본원 발명의 폴리펩티드는 상기 폴리펩티드의 반감기를 증가시키기 위하여 또는 공지된 방법을 이용한 면역 분석에 사용하기 위하여 항체 일부분과 융합 또는 공액될 수 있다. 폴리펩티드는 또한, 다합체를 형성하기 위하여 항체 일부분과 융합 또는 공액될 수 있다. 가령, 본원 발명의 폴리펩티드에 융합된 Fc 영역은 Fc 영역간 다이설파이드 결합을 통하여 이합체를 형성할 수 있다. 좀더 상등의 다합체 형태는 본원 발명의 폴리펩티드를 IgA와 IgM의 일부분에 융합시켜 만들 수 있다. 본원 발명의 폴리펩티드를 항체 일부분에 융합 또는 공액하는 방법은 당분야에 공지된 것이다(US Patents 5,336,603, 5,622,929, 5,359,046, 5,349,053, 5,447,851, 5,112,946; EP 0 307 434, EP 0 367 166; WO 96/04388, WO 91/06570; Ashkenazi A. et al.(1991); Zheng, X.X. et al.(1995); Vil, H. et al.(1992)).

또한, 본원 발명은 본원 발명에 따른 폴리펩티드의 작용제 또는 길항제로 작용하는 항체에 관한다. 가령, 본원 발명에는 본원 발명에 따른 폴리펩티드와 수용체/리간드 상호작용을 완전히 또는 부분적으로 파괴하는 항체가 포함된다. 수용체-특이적 항체와 리간드-특이적 항체 역시 본원 발명에 포함된다. 또한, 본원 발명에는 리간드 결합을 방해하지 않으면서 수용체 활성화를 예방하는 수용체-특이적 항체가 포함된다. 수용체 활성화(즉, 신호 전달)는 당분야에 공지된 기술로 확인할 수 있다. 본원 발명에는 리간드 결합과 수용체 활성화를 모두 차단하는 수용체-특이적 항체도 포함된다. 유사하게, 본원 발명에는 리간드와 결합하여 수용체와 리간드의 결합을 차단하는 중화 항체 및 리간드와 결합하여 수용체 활성화를 차단하지만 리간드가 수용체와 결합하는 것을 차단하지는 않는 항체가 포함된다. 또한, 본원 발명에는 수용체를 활성화시키는 항체가 포함된다. 이들 항체는 리간드-매개된 수용체 활성화의 영향을 받는 생물 활성에 대한 작용제 역할을 할 수도 있다. 이런 항체는 본원에 개시된 특이적 활성을 포함하는 생물학적 활성에 대한 작용제 또는 길항제로 기술할 수 있다. 상기 항체 작용제는 당분야에 공지된 방법으로 만들 수 있다(WO 96/40281; US Patent 5,811,097; Deng B. et al.(1998); Chen, Z. et al.(1998); Harrop J.A. et al.(1998); Zhu, Z. et al.(1998); Yoon, D.Y. et al.(1998); Prat, M. et al.(1998); Pitard, V. et al.(1997); Liautard, J. et al.(1997); Carlson, N.G. et al.(1997); Taryman, R.E. et al.(1995); Muller, Y.A. et al(1998); Bartunek, P. Et al.(1996)).

전술한 바와 같이, 본원 발명에 따른 폴리펩티드의 항체는 당분야에 공지된 기술로 본원 발명의 폴리펩티드를 '모방'하는 항-이디오타입 항체를 만드는데 사용할 수 있다(Greenspan and Bona, (1989); Nissinoff, (1991)). 가령, 결합을 통하여 폴리펩티드 다합체화(multimerization) 또는 본원 발명에 따른 폴리펩티드와 리 간드의 결합을 경쟁적으로 저해하는 항체는 폴리펩티드 다합체화 또는 결합 도메인을 '모방'하고 결론적으로 폴리펩티드 또는 이의 리간드와 결합하여 이들을 중화시키는 항-이디오타입 항체를 만드는데 사용할 수 있다. 이런 중화 항-이디오타입 항체는 본원 발명의 폴리펩티드에 결합시키거나 또는 이의 리간드/수용체 결합시켜 이의 생물 활성을 차단하는데 사용할 수 있다.

CanIon-관련된 이중대립인자 마커

본원 발명에 따른 이중대립인자 마커의 이점

본원 발명의 CanIon-관련된 이중대립인자 마커는 RFLP(제한 단편 길이 다형성) 및 VNTR(직렬 반복서열의 변이 총수) 마커와 같은 다른 유전자 마커에 비하여 다수의 중요한 이점을 제공한다.

1세대 마커는 RFLP로, 이는 제한 단편의 길이를 변화시키는 변수이다. 하지만, RFLP를 확인하고 분류하는데 이용되는 방법은 상대적으로 많은 재료, 노력, 시간을 소모시킨다. 2세대 유전자 마커는 VNTR로, 이는 미니위성(minisatellite) 또는 마이크로위성(microsatellite)으로 분류할 수 있다. 미니위성은 0.1 내지 20 Kilobase 길이의 사람 크로모좀의 영역을 따라 분포되는 5-50개 반복서열 단위에 존재하는 직렬 반복된 DNA 서열이다. 이들은 많은 가능 대립유전자를 제공하기 때문에, 이들의 정보 내용은 가치가 높다. 미니위성은 검사 개체로부터 얻은 핵산 시료에 존재하는 직렬 반복서열의 총수를 확인하는 서던 블랏을 실시하여 기록한다. 하지만, 10^4 잠재적 VNTR만 서던 블랏팅으로 분류될 수 있다. 게다가, RELP와 VNTR은 개발하고 대량으로 분석하는데 많은 비용과 시간이 소모된다.

단일 뉴클레오티드 다형성 또는 이중대립인자 마커는 RELP 및 VNTR과 동일한 방식으로 사용되면서도 다양한 이점을 제공할 수 있다. SNP는 사람 게놈에 집중적으로 위치하고 가능한 빈번한 형태의 변이이다. 대략 10^{7} 개 이상의 부위가 사람 게놈의 3x10^{9} 개 염기쌍을 따라 분산되어 있다. 따라서, SNP는 RFLP 또는 VNTR 마커보다 좀더 높은 빈도와 통일성으로 발생하는데, 이는 이런 마커가 관심있는 유전자 좌위에 근접하여 발견될 가능성이 좀더 높다는 것을 의미한다.

또한, 본원 발명의 이중대립인자 마커와 같은 특성화된 단일 뉴클레오티드 다형성의 다른 형태는 구별하는 것이 좀더 용이하고, 따라서 일상적인 기초하에 용이하게 분류될 수 있다. 이중대립인자 마커는 단일 뉴클레오티드 기초한 대립유전자 및 단지 2개의 공통 대립유전자를 보유하는데, 이는 매우 병렬적인 검출 및 자동화된 기록을 가능하게 한다. 본원 발명의 이중대립인자 마커는 다수 개체의 신속한 고성능 유전자형판별(genotyping)를 가능하게 한다.

이중대립인자 마커는 게놈에 집중적으로 위치하고 충분히 유익하며 대량으로 분석할 수 있다. 이런 이점의 종합적 효과로 인해 이중대립인자 마커는 유전자 연구에 극히 중요하다. 이중대립인자 마커는 가족에서 연쇄 연구, 대립유전자 공유 방법, 개체군에서 연쇄 불균형 연구 및 사례-대조 개체군 또는 형질 양성과 형질 음성 개체군의 연관 연구에 사용될 수 있다. 본원 발명의 중요한 측면은 복잡한 형질에 관여하는 유전자를 동정하는 연관 연구가 이중대립인자 마커를 통하여 실시 될 수 있다는 점이다. 연관 연구는 무관한 사례-와 대조-개체군에서 마커 대립유전자의 빈도를 검사하는데, 일반적으로 다인자(polygenic) 또는 산발적 형질의 검출에 이용된다. 연관 연구는 전체 개체군에서 실시할 수 있고, 영향을 받은 가족에서 관련된 개체에 실시된 연구(연쇄 연구)로 한정되지 않는다. 상이한 유전자에서 이중대립인자 마커는 질병과의 직접적인 연관 또는 치료에 대한 반응을 병행으로 스크린할 수 있다. 이런 다중 유전자 방식은 특정 표현형, 약물 반응, 산발적 형질, 또는 복잡한 유전자 병인을 갖는 질병상태에 대한 다중적 유전자 인자의 상승적 효과를 검사하는데 필요한 통계학적 능력을 제공한다는 점에서, 사람 유전자 연구에 유력한 도구이다.

본원 발명의 후보 유전자

연관 방법을 실시하는데 상이한 방식이 이용될 수 있다: 게놈-전체 연관 연구, 후보 영역 연관 연구, 후보 유전자 연관 연구. 게놈-전체 연관 연구는 균일하게 분포하고 전체 게놈을 커버하는 유전자 마커의 스크리닝에 의존한다. 후보 유전자 방식은 관심있는 형질과 관련된 생물학적 경로에 잠재적으로 관여하는 유전자에 특이적으로 위치하는 유전자 마커의 연구에 기초한다. 본원 발명에서, CanIon이 후보 유전자이다. 이런 후보 유전자 분석은 형질의 생태에 관한 정보가 일부 가용한 경우에, 특정 형질과 관련된 유전자 및 유전자 다형성의 동정에 지름길을 제공한다. 하지만, 본원에 개시된 모든 이중대립인자 마커는 게놈-전체 연관 연구 또는 후보 영역 연관 연구의 일부로 활용될 수 있는데, 이런 활용은 본원 발명과 특허청구범위에서 특이적으로 정관한다.

CanIon-관련된 이중대립인자 마커 및 이와 관련된 폴리뉴클레오티드

본원 발명은 또한, CanIon-관련된 이중대립인자 마커에 관한다. 본원에서 'CanIon-관련된 이중대립인자 마커'는 CanIon 유전자와 연쇄 불균형인 일단의 이중대립인자 마커를 의미한다. CanIon-관련된 이중대립인자 마커에는 A1 내지 A17로 명명된 이중대립인자 마커가 포함된다.

본원 발명에 따른 이중대립인자 마커의 일부분은 표 2에 개시한다. 이들은 관련된 1 내지 4와 6에서 단일 염기 다형성으로 기술된다. 한가지 CanIon 이중대립인자 마커의 다형적 염기를 보유하는 핵산의 증폭을 가능하게 하는 프라이머쌍은 실시예 2의 표 1에 기재한다.

17개의 CanIon-관련된 이중대립인자 마커, A1 내지 A17은 CanIon의 게놈 서열에 위치한다. 이중대립인자 마커 A12와 A16은 CanIon의 엑손에 위치한다. 이중대립인자 마커 A18은 CanIon 유전자와 측면에서 접한다.

본원 발명은 또한, CanIon-관련된 이중대립인자 마커의 다형적 염기를 포함하는 정제된 및/또는 분리된 뉴클레오티드 서열에 관한다. 적절한 구체예에서, 이중대립인자 마커는 A1 내지 A18, 이의 보체에서 선택된다. 상기 서열은 SEQ ID 1 내지 4와 6, 이의 변이체 또는 상보적 서열에서 선택되는 뉴클레오티드 서열의 8 내지 1000개 길이의 뉴클레오티드, 바람직하게는 적어도 8, 10, 12, 15, 18, 20, 25, 35, 40, 50, 60, 70, 80, 100, 250, 500, 1000개 뉴클레오티드의 연속 스팬을 보유한다. 이들 뉴클레오티드 서열은 이중대립인자 마커의 대립유전자 1 또는 2중 한가지의 다형적 염기를 포함한다. 선택적으로, 상기 대립유전인자 마커는 상기 폴리뉴클레오티드의 중심으로부터 6, 5, 4, 3, 2, 1개 뉴클레오티드 이내, 또는 상기 폴리뉴클레오티드의 중심에 위치할 수 있다. 선택적으로, 상기 스팬의 3' 말단은 상기 폴리뉴클레오티드의 3' 말단에 존재할 수 있다. 선택적으로, 이중대립인자 마커는 상기 폴리뉴클레오티드의 3' 말단에 존재할 수 있다. 선택적으로, 상기 폴리뉴클레오티드는 라벨을 추가로 포함할 수 있다. 선택적으로, 상기 폴리뉴클레오티드는 고체 서포트에 부착될 수 있다. 다른 구체예에서, 전술한 폴리뉴클레오티드는 단독으로 또는 임의의 조합으로 사용될 수 있다.

본원 발명은 또한, SEQ ID 1 내지 4 또는 이의 상보적 서열에서 선택되는 뉴클레오티드 서열의 8 내지 1000개 연속

뉴클레오티드, 바람직하게는 적어도 8, 10, 12, 15, 18, 20, 25, 35, 40, 50, 60, 70, 80, 100, 250, 500, 1000개 뉴클레오티드의 연속 스팬을 포함하는 정제된 및/또는 분리된 뉴클레오티드 서열에 관한다. 상기 뉴클레오티드의 3' 말단은 서열에서 CanIon-관련된 이중대립인자 마커 이내에 또는 상기 마커의 적어도 2, 4, 6, 8, 10, 12, 15, 18, 20, 25, 50, 100, 250, 500 또는 1000개 뉴클레오티드 상류에 위치할 수 있다. 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17에서 선택된다; 선택적으로, 상기 폴리뉴클레오티드의 3' 말단은 서열에서 CanIon-관련된 이중대립인자 마커의 1개 뉴클레오티드 상류에 위치할 수 있다. 선택적으로, 상기 폴리뉴클레오티드는 라벨을 추가로 포함할 수 있다. 선택적으로, 상기 폴리뉴클레오티드는 고체 서포트에 부착될 수 있다. 다른 구체예에서, 전술한 폴리뉴클레오티드는 단독으로 또는 임의의 조합으로 사용될 수 있다.

적절한 구체예에서, 표 2에 기재된 이중대립인자 마커중 한가지의 다형적 염기를 포함하는 서열은 표 1에 기재된 앰플리콘 서열의 핵산, 이의 변이체 또는 상보적 서열에서 선택되는 폴리뉴클레오티드의 연속 스팬을 보유하는, 이런 폴리뉴클레오티드로 구성되는, 이런 폴리뉴클레오티드에 포함되는, 또는 이런 폴리뉴클레오티드를 포함하는 뉴클레오티드 서열에서 선택된다.

본원 발명은 또한, CanIon 단백질을 인코드하는 핵산에 관하는데, 상기 핵산은 A12와 A16 및 이의 보체에서 선택되는 이중대립인자 마커의 다형적 염기를 포함한다.

본원 발명에는 CanIon-관련된 이중대립인자 마커에서 한가지이상 뉴클레오티드의 실체를 확인하는데 사용되는 임의의 폴리뉴클레오티드, 또는 이런 폴리뉴클레오티드의 용도가 포함된다. 이에 더하여, CanIon-관련된 이중대립인자 마커에서 한가지이상 뉴클레오티드의 실체를 확인하는데 사용되는 본원 발명의 폴리뉴클레오티드에는 추가적인 제한없이 본원에 개시된 단독의 또는 조합된 폴리뉴클레오티드가 포함된다. 선택적으로 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 폴리뉴클레오티드는 본원에 개시된 단독의 또는 조합된 서열을 포함한다; 선택적으로, 상기 폴리뉴클레오티드는 본원에 개시된 임의의 폴리뉴클레오티드로 구성된다; 선택적으로, 확인은 혼성화 분석, 염기서열분석, 마이크로염기서열 분석, 또는 효소-기초된 미스매치 검출 분석으로 실시할 수 있다; 선택적으로, 상기 폴리뉴클레오티드는 고체 서포트, 어레이, 또는 어드레스가능 어레이에 부착될 수 있다; 선택적으로, 상기 폴리뉴클레오티드는 표지될 수 있다. 적절한 폴리뉴클레오티드는 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드의 실체를 확인하기 위한 혼성화 분석에 사용될 수 있다. 다른 적합한 폴리뉴클레오티드는 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드의 실체를 확인하기 위한 염기서열 분석 또는 마이크로염기서열 분석에 사용될 수 있다. 다른 적절한 폴리뉴클레오티드는 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드의 실체를 확인하기 위한 효소-기초한 미스매치 검출 분석에 사용될 수 있다. 다른 적절한 폴리뉴클레오티드는 CanIon-관련된 이중대립인자 마커를 포함하는 폴리뉴클레오티드의 분절을 증폭하는데 사용될 수 있다. 선택적으로, 상기 폴리뉴클레오티드는 고체 서포트, 어레이, 또는 어드레스가능 어레이에 부착될 수 있다; 선택적으로, 상기 폴리뉴클레오티드는 표지될 수 있다.

또한, 본원 발명에는 CanIon-관련된 이중대립인자 마커를 포함하는 뉴클레오티드를 증폭하는데 사용되는 임의의 폴리뉴클레오티드, 또는 이런 폴리뉴클레오티드의 용도가 포함된다. 이에 더하여, CanIon-관련된 이중대립인자 마커를 포함하는 뉴클레오티드의 분절을 증폭하는데 사용되는 본원 발명의 폴리뉴클레오티드에는 추가적인 제한없이 본원에 개시된 단독의 또는 조합된 폴리뉴클레오티드가 포함된 다. 선택적으로 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 폴리뉴클레오티드는 본원에 개시된 단독의 또는 조합된 서열을 포함한다; 선택적으로, 상기 폴리뉴클레오티드는 본원에 개시된 임의의 폴리뉴클레오티드로 구성된다; 선택적으로, 증폭은 PCR 또는 LCR로 실시할 수 있다; 선택적으로, 상기 폴리뉴클레오티드는 고체 서포트, 어레이, 또는 어드레스가능 어레이에 부착될 수 있다; 선택적으로, 상기 폴리뉴클레오티드는 표지될 수 있다.

본원 발명의 이중대립인자 마커를 포함하는 폴리뉴클레오티드의 증폭이나 염기서열 반응을 위한 프라이머는 당분야에 공지된 서열로부터 설계할 수 있다. 적절한 프라이머 세트에서 SEQ ID 1 내지 4와 6, 이의 상보적 서열 또는 변이체에서 선택되는 서열과 상동한 연속 스팬의 3' 말단은 프라이머의 3' 말단에 존재한다. 이런 형태는 프라이머의 3' 말단이 선별된 핵산 서열과 혼성화되도록 하고 증폭이나 염기서열 반응을 위한 프라이머의 효능을 극적으로 증가시킨다. 대립유전자 특이적 프라이머는 이중대립인자 마커의 다형적 염기가 연속 스팬의 3' 말단에 위치하고 연속 스팬이 프라이머의 3' 말단에 존재하도록 설계할 수 있다. 이런 대립유전자 특이적 프라이머는 이중대립인자 마커에 존재하는 대립유전자 2개중 하나를 보유하는 핵산 시료와 함께 사용되면 증폭이나 염기서열 반응을 선별적으로 기폭하는 경향이 있다. 본원 발명에 따른 프라이머의 3' 말단은 상기 서열에서 또는 염기서열, 증폭, 신규한 서열이나 마커의 위치 선정에 사용하기 적합한 임의의 다른 위치에서, CanIon-관련된 이중대립인자 마커 이내에 또는 상기 마커의 적어도 2, 4, 6, 8, 10, 12, 15, 18, 20, 25, 50, 100, 250, 500 또는 1000개 뉴클레오티드 상류에 위치할 수 있다. 따라서,

다른 일단의 적절한 증폭 프라이머는 SEQ ID 1 내지 4와 6, 이의 상보적 서열 또는 변이체에서 선택되는 서열에서 8 내지 50개 뉴클레오티드의 연속 스팬으로 구성되는 분리된 폴리뉴클레오티드를 포함하는데, 여기서 상기 연속 스팬의 3' 말단은 상기 폴리뉴클레오티드의 3' 말단에 위치하고, 상기 폴리뉴클레오티드의 3' 말단은 서열에서 CanIon-관련된 이중대립인자 마커의 상류에 위치한다. 적절하게는, 이들 증폭 프라이머는 서열 B1 내지 B17 및 C1 내지 C17에서 선택되는 서열을 포함한다. CanIon의 이중대립인자 마커의 1개 뉴클레오티드 상류에 3'말단이 위치하는 프라이머는 마이크로염기서열 분석에 특히 유용하다. 적절한 마이크로염기서열 프라이머는 표 4에 기술한다. 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 마이크로염기서열 프라이머는 뉴클레오 티드 서열 D1 내지 D18과 E1 내지 E18에서 선택된다.

본원 발명의 프로브는 당분야에 공지된 방법, 특히 본원에 개시된 마커가 존재하는 지를 검사할 수 있는 방법으로부터 설계할 수 있다. 적절한 프로브 세트는 임의 특정한 분석 조건하에 이중대립인자 마커의 한 대립유전자에는 선별적으로 결합하지만 다른 대립유전자에는 결합하지 않도록 하는 당분야에 공지된 방식으로 본원 발명의 혼성화 분석에 사용되도록 설계할 수 있다. 적절한 혼성화 프로브는 이중대립인자 마커의 대립유전자 1 또는 2중 한가지의 다형적 염기를 포함한다. 선택적으로, 상기 대립유전인자 마커는 상기 폴리뉴클레오티드의 중심으로부터 6, 5, 4, 3, 2, 1개 뉴클레오티드 이내, 또는 상기 폴리뉴클레오티드의 중심에 위치할 수 있다. 적절한 구체예에서, 프로브는 서열 P1 내지 P18 각각 및 이의 상보적 서열 각각에서 선택된다.

본원 발명의 폴리뉴클레오티드는 서열 리스트에 명시된 다형적 염기를 둘러싸는 정확한 측부 서열을 갖는 것으로 한정되지 않는다. 오히려, 이중대립인자 마커를 둘러싸는 측부 서열은 의도한 용도에 따라 얼마간 길어지거나 또는 짧아질 수 있는데, 본원 발명은 이런 서열을 구체적으로 정관한다. 연속 스팬 외부의 측부 영역은 사람 개체에서 실제로 발생하는 고유 측부 영역과는 상동하지 않아야 한다. 의도된 뉴클레오티드와 양립하는 임의 뉴클레오티드 서열의 첨가를 구체적으로 정관한다.

프라이머와 프로브는 '올리고뉴클레오티드 프로브와 프라이머'에서 밝힌 바와 같이 표지하거나 또는 고체 서포트에 고정시킬 수 있다.

고체 서포트에 부착되는 본원 발명의 폴리뉴클레오티드에는 제한없이 본원에 개시된 단독의 또는 조합된 폴리뉴클레오티드가 포함된다: 선택적으로, 상기 폴리뉴클레오티드는 단일 고체 서포트에 개별적으로, 또는 본원 발명에 따른 적어도 2, 5, 8, 10, 12, 15, 20 또는 25개의 상이한 폴리뉴클레오티드의 군으로 부착될 수 있다. 선택적으로, 본원 발명에 따른 폴리뉴클레오티드가 아닌 다른 폴리뉴클레오티드가 본원 발명에 따른 폴리뉴클레오티드에서와 동일한 고체 서포트에 부착될 수 있다. 선택적으로, 다중 폴리뉴클레오티드가 고체 서포트에 부착되는 경우에, 이들은 무작위 위치에, 또는 정렬된 어레이에 부착될 수 있다. 선택적으로, 상기 정렬된 어레이는 어드레스가능할 수 있다.

또한, 본원 발명에는 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드의 실체를 확인함으로써 검사 개체의 유전형분류에 필요한 모든 또는 일부 반응물과 함께 본원 발명에 따른 하나 또는 복수의 폴리뉴클레오티드를 포함하는 키트가 포함된다. 상기 키트의 폴리뉴클레오티드는 고체 서포트에 부착되거나, 또는 폴리뉴클레오티드의 어레이나 어드레스가능 어레이의 일부일 수 있다. 상기 키트는 염기서열분석 방법, 마이크로염기서열분석 방법, 혼성화 분석 또는 효소-기초한 미스매치 검출 분석 방법을 비롯한 당분야에 공지된 방법으로 마커 위치에서 뉴클레오티드의 실체를 확인할 수 있다.

이중대립인자 마커의 De Novo 동정을 위한 방법

다양한 방법, 예를 들면 올리고뉴클레오티드 프로브와 혼성화, 겔 전기영동으로 측정된 이동도에서 변화 검출 또는 증폭된 핵산의 직접 서열분석을 활용하여 단일 뉴클레오티드 다형성의 게놈 단편을 스크린할 수 있다. 이중대립인자 마커를 동정하는데 적합한 방법에는 다수의 무관한 개체로부터 게놈 DNA 단편의 비교 서열분석이 포함된다.

첫 번째 구체예에서, 무관한 개체로부터 DNA 시료는 함께 모으고, 이후 관심있는 게놈 DNA을 증폭하고 서열분석한다. 이렇게 수득된 뉴클레오티드 서열은 분석하여 현저한 다형성을 동정한다. 이런 방법의 중요한 이점중의 한가지는 DNA 시료의 집합(pooling)은 DNA 증폭 반응과 염기서열 반응의 회수를 실질적으로 감소시킨다는 점이다. 게다가. 이런 방법은 이렇게 수득된 이중대립인자 마커가 연관 연구를 실시하는데 사용되는 조금 덜 일반적인 대립유전자의 충분한 빈도를 보일 만큼 민감하다.

두 번째 구체예에서, DNA 시료는 모으지 않고, 따라서 개별적으로 증폭과 서열분석한다. 일반적으로, 이런 방법은 후보 유전자내에서 연관 연구를 실시하기 위하여 이중대립인자 마커를 동정할 필요가 있는 경우에 적합하다. 적절하게는, 매우 유관한 유전자 영역, 예를 들면 프로모터 영역이나 엑손 영역은 이중대립인자 마커에 대하여 스크린할 수 있

다. 이런 방법으로 얻은 이중대립인자 마커는 예로써 조금 빈도가 낮은 대립유전자의 빈도가 10% 이하이면, 연관 연구를 실시하는데 좀더 낮은 수준의 정보성(informativeness)을 보일 수 있다. 하지만, 이런 이중대립인자 마커는 연관연구를 실시할 만큼 충분한 정보를 제공하는데, 본원 발명의 유전자 분석 연구에 조금 덜 유익한 마커를 포함시키는 것은 일부 경우에 발현빈도(penetrance)에서 드물게 나타나는 원인 돌연변이의 직접적인 동정을 가능 하게 한다.

다음의 본원 발명에 따른 이중대립인자 마커의 동정에 사용되는데 적합한 방법의 다양한 파라미터를 설명한다.

게놈 DNA 서열

본원 발명의 이중대립인자 마커가 만들어지는 게놈 DNA 시료는 가급적, 공지된 인종 배경의 이질성 개체군에 상응하는 무관한 개체로부터 얻는다. DNA 시료를 얻는 개체수는 실질적으로 변화될 수 있는데, 바람직하게는 약 10 내지 1000명, 좀더 바람직하게는 50 내지 200명이다. 임의의 개체군에서 가능한 많은 마커를 동정하고 통계학적으로 유의한 결과를 발생시키는 충분한 다형적 다양성을 얻기 위하여 적어도 100명의 개체로부터 DNA 시료를 수거하는 것이 바람직하다.

분석할 게놈 DNA의 공급원과 관련하여, 제한없이 임의의 검사 시료를 사용할 수 있다. 이들 검사 시료는 본원 발명의 방법으로 검사할 수 있는 생물학적 시료를 함유하는데, 여기에는 사람과 동물 체액, 예를 들면 전체 혈액, 혈청, 혈장, 뇌척수, 뇨, 림프액 및 호흡기, 장, 비뇨생식기의 다양한 외부 분비물, 눈물, 타액, 우유, 백혈구, 골수종 등; 생물학적 유체, 예를 들면 세포 배양 상층액; 고정 조직 견본, 예를 들면 종양과 비-종양 조직과 림프절 조직; 골수 천자종(bone marrow aspirate)과 고정된 세포 견본이 포함된다. 본원 발명에 사용되는 게놈 DNA의 적합한 공급원은 각 공여자의 말초 정맥혈이다. 생물학적 시료로부터 게놈 DNA를 준비하는 기술은 당업자에게 공지된 것이다. 적절한 구체예의 상세한 내용은 실시예 1에 제시한다. 당업자는 집합된 또는 집합되지 않은 시료를 증폭할 지를 선택할 수 있다.

DNA 증폭

게놈 DNA의 시료에서 이중대립인자 마커의 동정은 DNA 증폭 방법으로 이용으로 조장할 수 있다. DNA 시료는 증폭 단계를 위하여 모으거나 또는 모으지 않을 수 있다. DNA 증폭 기술은 당업자에게 공지된 것이다.

본원 발명에 사용될 수 있는 증폭 기술에는 리가아제 연쇄 반응(LCR(EP-A-320 308, WO 9320227, EP-A 439 182), 중합효소 연쇄 반응(PCR, RT-PCR), 핵산 서열 기초한 증폭(NASBA)과 같은 기술(Guatelli J.C., et al.(1990); Compton J.(1991), Q-베타 증폭(유럽 특허 출원 No 4544610), 가닥 치환 증폭(Walker et al.(1996); EP A 684 315), 표적 매개된 증폭(PCT 출원 WO 9322461)이 포함되지만, 이들에 국한되지 않는다.

LCR과 Gap LCR은 기하급수적 증폭 기술로, DNA 분자에 어닐링된 인접 프라이머를 결합시키는데 DNA 리가아제에 의존한다. 리가아제 연쇄 반응(LCR)에서 프로브 쌍이 사용되고, 여기에는 2개의 일차(제1과 제2) 프로브 및 2개의 이차(제3과 제4) 프로브가 포함되는데, 이들 모두는 표적에 대하여 과잉 몰로 사용된다. 제1 프로브는 표적 가닥의 제1 분절과 혼성화되고 제2 프로브는 표적 가닥의 제2 분절과 혼성화되는데, 제1과 제2 분절은 인접하고, 따라서 일차 프로브는 5'인산염-3'하이드록실 관계로 접하고 리가아제는 2개의 프로브를 융합된 산물을 공유 융합 또는 결찰시킬 수 있다. 이에 더하여, 제3(이차) 프로브는 제1 프로브의 일부분과 혼성화되고 제4(이차) 프로브는 유사한 접경 방식으로 제2 분절의 일부분과 혼성화 된다. 물론, 표적이 처음에 이중 가닥이면, 이차 프로브는 일차적으로 표적 보체와 혼성화된다. 일차 프로브의 결찰된 가닥이 표적 가닥으로부터 분리되면, 제3과 제4 프로브와 혼성화되는데, 이들은 결찰시켜 상보적인 이차 결찰 산물을 만들 수 있다. 결찰된 산물이 표적이나 이의 보체와 기능적으로 등가라는 것을 인식하는 것이 중요하다. 혼성화와 결찰 사이클의 반복으로, 표적 서열의 증폭을 달성한다. 다중 LCR 방법 역시 보고되었다(WO 9320227). Gap LCR(GLCR)은 프로브가 인접하지 않고 2 내지 3개의 염기에 의해 분리되는 변형 LCR이다.

본원 발명에서 mRNA의 증폭을 위하여, mRNA를 cDNA를 역전사시키고 이후 이를 중합효소 연쇄 반응시키는 방법(RT-PCR), 단계에 단일 효소를 이용한 방법(미국 특허 5,322,770), 또는 비대칭 Gap LCR(RT-AGLCR)을 이용한 방법(Marsshall et al.(1994))이 활용된다. AGLCR은 RNA를 증폭할 수 있는 변형 GLCR이다.

PCR 기술은 본원 발명에 적합한 증폭 기술이다. 다양한 PCR 기술은 당분야에 공지된 것이다(White(1997); 'PCR Methods and Application', 1991, Cold Spring Harbor laboratory Press). 이들 각 PCR 과정에서, 증폭되는 핵산 서열의 양면에 PCR 프라이머는 dNTP 및 열안정성 중합효소, 예를 들면 Taq 중합효소, Pfu 중합효소 또는 Vent 중합효소와 함께 적절히 준비된 핵산 시료에 첨가한다. 혼성화된 프라이머는 신장시킨다. 이후, 변성, 혼성화, 신장의 사이클을 다시 시작한다. 사이클은 여러 번 반복하여 프라이머 부위사이의 핵산 서열을 보유하는 증폭된 단편을 만든다. PCR은 미국 특허 4,683,195; 4,683,202; 4,965,188을 비롯한 여러 특허에서 좀더 상세하게 기술되어 있다.

PCR 기술은 신규한 이중대립인자 마커를 동정하는데 적합한 증폭 기술이다. 본원 발명의 목적에 적합한 PCR 반응의 전형적인 예는 실시예 2에 제시한다.

본원 발명의 한 측면은 검사 시료에서 가급적 PCR을 이용하여 사람 CanIon 유전자, 특히 SEQ ID No 1 내지 3의 게놈 서열이나 SEQ ID 4의 cDNA 서열, 이의 단편 또는 변이체의 증폭 방법이다. 상기 방법은 다음의 단계로 구성된다:

a) 증폭되는 폴리뉴클레오티드 영역의 양면에 위치하는 전술한 증폭 프라이머 쌍을 함유하는 증폭 반응물과 검사 시료를 접촉시키고;

b) 선택적으로, 상기 증폭 산물을 검출한다.

본원 발명은 또한, 검사 시료에서 사람 CanIon 유전자, 특히 SEQ ID No 1 내지 3의 게놈 서열이나 SEQ ID 4의 cDNA 서열, 이의 단편 또는 변이체의 증폭 키트에 관한는데, 상기 키트는 다음과 같이 구성된다:

a) 증폭되는 CanIon 영역의 양면에 위치하는 한쌍의 올리고뉴클레오티드 프라이머;

b) 선택적으로, 증폭 반응을 실시하는데 필요한 반응물.

상기 증폭 방법과 키트의 한 구체예에서, 증폭 산물은 증폭된 영역과 상보적인 서열을 갖는 표지된 프로브와의 혼성화로 검출된다. 상기 증폭 방법과 키트의 다른 구체예에서, 프라이머는 B1 내지 B17, C1 내지 C17, D1 내지 D18, E1 내지 E18의 뉴클레오티드 서열에서 선택되는 서열을 포함한다.

본원 발명의 첫 번째 구체예에서, 이중대립인자 마커는 본 발명자들에 의해 얻어진 게놈 서열 정보를 이용하여 동정한다. 서열분석된 게놈 DNA 단편은 500 bp 단편의 증폭을 위한 프라이머를 설계하는데 사용된다. 이들 500 bp 단편은 게놈 DNA로부터 증폭하고 이중대립인자 마커를 스크린한다. 모든 프라이머는 특정 표적 염기의 상류에, 염기서열 프라이머로 기능하는 공통 올리고뉴클레오티드 꼬리를 보유한다.

CanIon 유전자를 인코드하는 게놈 서열의 증폭에 적합한 프라이머는 유전자의 프로모터, 엑손, 절단접합 부위에 집중한다. 이중대립인자 마커는 유전자의 이들 기능 영역에 위치하면 돌연변이를 초래할 가능성이 높아지게 된다. 본원 발명의 적절한 증폭 프라이머는 실시예 2, 표 1에 상세하게 기술된 뉴클레오티드 서열 B1 내지 B17과 C1 내지 C17을 보유한다.

증폭된 게놈 DNA의 서열분석 및 단일 뉴클레오티드 다형성의 동정

이후, 전술된 바와 같이 만들어진 증폭 산물은 당업자에게 공지된 임의의 방법으로 서열분석한다. 디데옥시-매개된 방법(Sanger 방법) 또는 Maxam-Gilbert 방법을 이용하여 DNA를 서열 분석하는 방법은 당업자에게 공지된 것이다. 이런 방법은 예로써 Sambrook et al.(1989)에 개시되어 있다. 대안적 방법에는 Chee et al.(1996)에서 밝힌 바와 같은 고밀도 DNA 어레이에 혼성화가 포함된다.

적절하게는, 증폭된 DNA는 염료 종결기 사이클 서열분석 프로토콜(dye terminator cycle sequencing protocol)을 이용한 자동화 디데옥시 종결기 염기서열 반응으로 측정한다. 다형성 검색은 동일한 위치에서 발생하는 상이한 염기로부터 기인하는 전기영동 패턴에서 중첩된 피크의 존재에 기초한다. 각 디데옥시 종결기는 상이한 형광 분자로 표지되기 때문에, 이중대립인자 부위에 상응하는 2개의 피크는 서열상의 동일한 위치에서 2개의 상이한 뉴클레오티드에 상응하는 차별적인 색깔을 나타낸다. 하지만, 2개 피크의 존재는 배경 소음(background noise)으로 인한 인공물(artifact)일 수 있다. 이런 인공물을 배제하기 위하여, 2개의 DNA 가닥은 서열분석하고 피크간 비교를 실시한다. 다형적 서열로 등록되기 위해서는 다형성이 양 가닥에서 모두 검출되어야 한다.

상기 과정으로 이중대립인자 마커를 보유하는 이들 증폭 산물을 확인할 수 있다. 100명 개체의 염기서열 풀에 의해 검출되는 이중대립인자 다형성 빈도의 검출 한계는 공지된 대립유전자 빈도의 염기서열 풀로 검증하는 경우에, 마이너 대립유전자에서 대략 0.1이다. 하지만, 상기 집합(pooling) 방법에 의해 검출되는 이중대립인자 다형성의 90%이상은 마이너 대립유전자에 대하여 0.25이상의 빈도를 갖는다. 따라서, 상기 방법에 의해 선별된 이중대립인자 마커는 마이너 대립유전자에서 적어도 0.1 및 메이저 대립유전자에서 0.9 이하의 빈도, 바람직하게는 마이너 대립유전자에서 적어도 0.2 및 메이저 대립유전자에서 0.8 이하의 빈도, 좀더 바람직하게는 마이너 대립유전자에서 적어도 0.3 및 메이저 대립유전자에서 0.7 이하의 빈도를 갖고, 따라서 0.18 이상, 바람직하게는 0.32 이상, 좀더 바람직하게는 0.42 이상의 이질접합성을 갖는다.

다른 구체예에서, 이중대립인자 마커는 개별 DNA 시료를 서열분석하여 검출하는데, 이런 이중대립인자 마커의 마이너 대립유전자의 빈도는 0.1 이하이다.

본원 발명에 따른 이중대립인자 마커의 검증

다형성은 양 대립유전자가 개체군에 존재하는 지를 검증함으로써 유전자 마 커로서의 활용도를 평가한다. 이중대립인자 마커의 검증은 본원 발명의 방법으로 일군의 개체를 유전자형판별하여 달성한다. 마이크로염기서열분석은 대립유전자를 유전자형판별하는데 적합한 방법은 유전자형판별 단계에 의한 검증은 상기 군에서 각 개체로부터 유래된 개별 시료에서 또는 1명이상의 개체로부터 유래된 집합된 시료에서 실시할 수 있다. 상기 군은 1명의 개체가 연구중인 대립유전자에 이형접합성이면 1명의 개체일 수 있다. 적절하게는, 상기 군에는 적어도 3명의 개체, 좀더 바람직하게는 적어도 5-6명의 개체가 포함되고, 따라서 단일 검증 검사로 복수의 이중대립인자 마커를 검증할 수 있다. 하지만, 검증 검사가 소규모 군에서 실시되면 샘플링 오류로 인해 검사 개체가 2개의 대립유전자중 한가지를 보유하지 않는 경우에 가음성(false negative) 결과가 초래될 수 있다. 따라서, 검증 과정은 특정 초기 결과가 인공물이고 서열상의 특정 위치에서 이중대립인자 마커가 실제로(*Bona fide*) 존재한다는 것을 입증하는 데에는 그다지 효과적이지 못하다. 선택적으로, 본원 발명의 유전자형판별, 일배체형분류(haplotyping), 연관, 상호작용 연구 방법은 검증된 이중대립인자 마커로만 실시할 수도 있다.

본원 발명에 따른 이중대립인자 마커의 빈도 평가

검증된 이중대립인자 마커는 이중대립인자 마커 위치에서 최소한의 공통 대립유전자의 빈도를 측정함으로써 유전자 마커로서 이들의 활용도를 추가로 평가한다. 조금 덜 공통된 대립유전자가 빈도가 높을수록, 연관 연구와 상호작용 연구에서 이중대립인자 마커의 활용도는 더 커진다. 최소한의 공통 대립유전자의 측정은 본원 발명의 방법으로 일군의 개체를 유전자형판별하고 양 대립유전자가 존재한다 는 것을 입증함으로써 달성한다. 유전자형판별 단계에 의한 빈도의 측정은 상기 군에서 각 개체로부터 유래된 개별 시료에서 또는 1명이상의 개체로부터 유래된 집합된 시료에서 실시할 수 있다. 상기 군은 전체 개체군을 대표할 수 있을 만큼 충분히 커야 한다. 적절하게는, 상기 군에는 적어도 20명의 개체, 좀더 바람직하게는 적어도 50명의 개체, 가장 바람직하게는 적어도 100명의 개체가 포함된다. 물론, 군의 규모가 더 클수록, 감소된 샘플링 오류로 인해 빈도 측정의 정확도가 커진다. 조금 덜 공통된 대립유전자의 빈도가 30 % 이상인 이중대립인자 마커는 '고 품질 이중대립인자 마커'라고 한다. 선택적으로, 본원 발명의 유전자형판별, 일배체형판별(haplotyping), 연관, 상호작용 연구 방법은 고 품질 이중대립인자 마커로만 실시할 수도 있다.

이중대립인자 마커에 대하여 개체를 유전자형판별하는 방법

본원 발명의 한가지이상 이중대립인자 마커에 대하여 생물학적 시료를 유전자형판별하는 방법을 제시하는데, 이들 방법은 시험관내에서 실시한다. 이런 유전자형판별 방법은 당분야에 공지된 방법으로 CanIon 이중대립인자 마커 위치에서 뉴클레오티드의 실체를 확인하는 단계로 구성된다. 이들 방법은 연관 연구에서 사례-대조 개체군 및 임의의 형질과 연관된 것으로 알려진 이중대립인자 마커의 대립유전자의 검출에서 개체를 유전자형판별하는데 이용되는데, 여기서 개체의 게놈에 존재하는 이중대립인자 마커의 양 사본이 측정되고, 따라서 개체는 특정 대립유전자에 대하여 동형접합성 또는 이형접합성으로 분류될 수 있다.

이들 유전자형판별 방법은 단일 개체 또는 집합된 DNA 시료로부터 유래된 핵 산 시료에서 실시할 수 있다.

유전자형판별은 이중대립인자 마커의 동정을 위한 전술된 방법과 유사한 방법, 또는 후술된 방법과 같은 유전자형판별 방법을 이용하여 실시할 수 있다. 적절한 구체예에서, 상이한 개체로부터 증폭된 게놈 단편 단편의 서열 비교는 신규한 이중대립인자 마커를 동정하는데 이용되는 반면, 마이크로염기서열분석은 진단과 연관 연구에서 공지된 이중대립인자 마커를 유전자형판별하는데 이용된다.

한 구체예에서, 본원 발명에는 생물학적 시료에서 CanIon-관련된 이중대립인자 마커에서 뉴클레오티드 또는 이의 보체의 실체를 확인하는 유전자형판별 방법이 포함된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 생물학적 시료는 단일 개체로부터 유래된다; 선택적으로, 상기 생물학적 시료는 다중 개체로부터 유래된다; 선택적으로, 상기 이중대립인자 마커에서 뉴클레오티드의 실체는 개체의 게놈에 존재하는 이중대립인자 마커의 양 사본에서 확인된다; 선택적으로, 본원 발명의 유전자형판별 방법에는 임의의 제한없이 본원에 개시된 단독의 또는 조합된 방법이 포함된다; 선택적으로, 상기 방법은 시험관내에서 실시된다; 선택적으로, 확인 단계에 앞서 이중대립인자 마커를 보유하는 서열의 일부분을 증 폭하는 단계가 추가로 포함된다. 선택적으로, 상기 증폭은 PCR, LCR, 또는 숙주 세포에서 복제 기점과 상기 단편을 포함하는 재조합 벡터의 복제로 실시한다; 선택적으로, 확인은 혼성화 분석, 염기서열분석, 마이크로염기서열분석 또는 효소-기초한 미스매치 검출 분석으로 실시한다.

유전자형판별을 위한 핵산의 공급원

정제된 또는 비-정제된 핵산의 공급원은 원하는 특이적인 핵산 서열을 보유하는 또는 보유한 것으로 생각되면 출발

핵산으로 사용될 수 있다. DNA 또는 RNA는 전술한 바와 같이 세포, 조직, 체액 등으로부터 추출할 수 있다. 본원 발명의 유전자형판별 방법에 사용되는 핵산이 임의의 포유동물 공급원으로부터 유래될 수 있다면, 핵산 시료를 채취하는 검사 대상과 개체는 사람으로 간주한다.

이중대립인자 마커를 포함하는 DNA 단편의 증폭

본원 발명의 한가지이상 대립유전자 마커를 포함하는 뉴클레오티드 분절을 증폭하기 위한 방법과 폴리뉴클레오티드를 제공한다. 이중대립인자 마커를 포함하는 DNA 단편의 증폭은 다양한 방법과 목적에 활용될 수 있으며, 유전자형판별에 한정되지 않는다. 하지만, 대부분의 유전자형판별 방법은 관심있는 이중대립인자 마커를 보유하는 DNA 영역의 사전 증폭을 필요로 한다. 이런 방법은 이중대립인자 마커에 걸쳐 있거나 또는 상기 부위를 보유하는 서열 및 여기에 멀리 또는 가까이 위치하는 서열의 농도나 전체 개체수를 특이적으로 증가시킨다. 진단 방법 역시 본원 발명의 이중대립인자 마커를 보유하는 DNA 분절의 증폭에 의존할 수 있다. DNA의 증폭은 당분야에 공지된 임의의 방법으로 달성할 수 있다. 증폭 기술은 'DNA 증폭' 섹션에서 전술한다.

이들 증폭 방법중 일부는 단일 뉴클레오티드 다형성의 검출에 특히 적합하고, 표적 서열의 동시 증폭 및 다형적 뉴클레오티드의 동정을 가능하게 한다.

전술한 이중대립인자 마커의 동정은 적합한 올리고뉴클레오티드의 설계를 가능하게 하는데, 상기 올리고뉴클레오티드는 본원 발명의 이중대립인자 마커를 포함하는 DNA 단편을 증폭하는 프라이머로 사용될 수 있다. 증폭은 본원에서 밝힌 신규한 이중대립인자 마커를 발견되는 처음 사용된 프라이머, 또는 본원 발명의 이중대립인자 마커를 포함하는 DNA 단편의 증폭을 가능하게 하는 일단의 프라이머를 이용하여 실시할 수 있다.

일부 구체예에서, 본원 발명은 본원 발명의 한가지이상 이중대립인자 마커를 보유하는 DNA 단편을 증폭하기 위한 프라이머를 제공한다. 적절한 증폭 프라이머는 실시예 2에 기재한다. 기재된 프라이머는 단지 예를 든 것으로, 본원 발명의 한가지이상 이중대립인자 마커를 보유하는 증폭 산물을 생산하는 다른 일단의 프라이머 역시 사용 가능하다.

프라이머의 간격은 증폭되는 분절의 길이를 결정한다. 본원 발명에서, 이중대립인자 마커를 보유하는 증폭된 분절은 적어도 25 bp 내지 35 bp 크기를 갖는다. 증폭 단편은 25-3000 bp, 바람직하게는 50-1000 bp, 좀더 바람직하게는 100-600 bp이다. 이중대립인자 마커에 대한 증폭 프라이머는 마커를 보유하는 DNA 단편의 특이적인 증폭을 가능하게 하는 임의의 서열일 수 있다. 증폭 프라이머는 '올리고뉴클레오티드 프로브와 프라이머'에서 밝힌 바와 같이 표지되거나 또는 고체 서포트 에 고정될 수 있다.

이중대립인자 마커에 대하여 DNA 시료를 유전자형판별하는 방법

이중대립인자 마커 부위에 존재하는 뉴클레오티드를 동정하기 위하여 당분야에 공지된 방법을 이용할 수 있다. 검출되는 이중대립인자 마커 대립유전자가 본원 발명에서 동정되고 구체화되었기 때문에, 당업자는 다양한 기술을 활용하여 검출을 용이하게 실시할 수 있다. 많은 유전자형판별 방법은 관심있는 이중대립인자 마커를 보유하는 DNA 영역의 사전 증폭을 요구한다. 표적이나 신호의 증폭이 현재 선호되긴 하지만, 증폭을 필요로 하지 않는 초감수성 검출 방법 역시 본원 발명의 유전자형판별 방법에 포함된다. 이중대립인자 다형성을 검출하는데 사용될 수 있는 당업자에게 공지된 방법에는 통상적인 점 블랏 분석, 단일 가닥 형태 다형성 분석(SSCP)(Orita et al.(1989)), 변성 농도구배 겔 전기영동(DGGE), 헤테로이중나선 분석, 미스매치 제거 검출 및 다른 통상적인 기술(Sheffield et al.(1991), White et al.(1992), Grompe et al.(1989과 1993))과 같은 방법이 포함된다. 특정 다형적 부위에 존재하는 뉴클레오티드의 실체를 확인하는 다른 방법에서는 미국 특허 4,656,127에서 밝힌 바와 같은 전문화된 엑소뉴클레아제-저항성 뉴클레오티드 유도체를 이용한다.

적절한 방법은 염기서열 분석, 효소-기초한 미스매치 검출 분석 또는 혼성화 분석으로 이중대립인자 마커 부위에 존재하는 뉴클레오티드의 실체를 직접적으로 측정하는 단계를 포함한다. 다음은 일부 적절한 방법의 설명이다. 특히 적절한 방법은 마이크로염기서열분석 기술이다. 본원에서 '염기서열분석'은 이중나선 프 라이머/주형 복합체의 중합효소 신장을 의미하는데, 여기에는 통상적인 염기서열분석과 마이크로염기서열분석이 포함된다.

1) 염기서열분석

다형적 부위에 존재하는 뉴클레오티드는 염기서열분석으로 확인할 수 있다. 적절한 구체예에서, DNA 시료는 전술한 염기서열분석에 앞서 PCR 증폭한다. DNA 염기서열분석 방법은 '증폭된 게놈 DNA의 증폭 및 단일 뉴클레오티드 다형성의 동정'에서 기술한다.

적절하게는, 증폭된 DNA는 염료-프라이머 사이클 염기서열분석 프로토콜을 이용한 자동화 디데옥시 종료기 서열분석 반응을 실시한다. 서열 분석은 이중대립인자 마커 부위에 존재하는 염기의 동정을 가능하게 한다.

2) 마이크로염기서열분석

마이크로염기서열분석 방법에서, 표적 DNA의 다형성 부위에서 뉴클레오티드는 단일 뉴클레오티드 프라이머 신장 반응으로 검출한다. 상기 방법에는 표적 핵산에서 관심있는 다형적 염기의 바로 상류와 혼성화되는 적합한 마이크로염기서열분석 프라이머가 관여한다. 중합효소는 다형적 부위에서 뉴클레오티드에 상보적인 1개의 단일 ddNTP(사슬종료기)로 프라이머의 3'말단을 특이적으로 신장시키는데 사용된다. 이후, 통합된 뉴클레오티드의 실체는 임의의 적합한 방법으로 확인한다.

전형적으로, 마이크로염기서열 반응은 형광 ddNTP로 실시하고 신장된 마이크로염기서열분석 프라이머는 EP 412 883에서 밝힌 바와 같이, ABI377 염기서열분석 기계에서 전기영동으로 분석하여 통합된 뉴클레오티드의 실체를 확인한다. 대안으로, 모세관 전기영동은 좀더 많은 수의 분석을 동시에 처리하는데 사용될 수 있다. 본원 발명에서 이용될 수 있는 전형적인 마이크로염기서열분석 과정의 예는 실시예 4에 제시한다.

ddNTP의 표지와 검출에 상이한 접근방식을 활용할 수 있다. 형광 공명 에너지 전달에 기초한 상동성 상 검출 방법은 Chen and Kwok(1997) 및 Chen et al.(1997)에서 기술한다. 상기 방법에서, 다형적 부위를 보유하는 증폭된 게놈 DNA 단편은 대립유전자 염료-표지된 디데옥시리보뉴클레오시드 삼인산염과 변형된 Taq 중합효소의 존재하에 5'-플루레신-표지된 프라이머와 함께 배양한다. 염료-표지된 프라이머는 주형에 존재하는 대립유전자에 특이적인 염료-종료기에 의해 1염기 신장된다. 유전자형판별 반응의 종결시점에서, 반응 혼합물에서 2가지 염료의 형광 강도는 분리 또는 정제없이 직접 분석한다. 이들 모든 단계는 동일한 튜브에서 실시할 수 있고, 형광 변화는 실시간으로 모니터할 수 있다. 대안으로, 신장된 프라이머는 MALDI-TOF 질량 분광법으로 분석할 수 있다. 다형적 부위에서 염기는 마이크로염기서열분석 프라이머에 더해진 질량에 의해 동정된다(Haff and Smirnov, 1997).

마이크로염기서열분석은 확립된 염기서열분석 방법 또는 이로부터 파생된 방법의 개발로 달성할 수 있다. 대안적 방법에는 여러 고체상 마이크로염기서열분석 기술이 포함된다. 기본 마이크로염기서열분석 프로토콜은 프라이머 또는 표적 분자가 고체 서포트에 고정이나 포획되는 이형이질성 상 분석으로 상기 방법이 실시 되는 점을 제외하고, 전술한 바와 동일하다. 프라이머 분리와 종결 뉴클레오티드 첨가 분석을 단순화시키기 위하여, 올리고뉴클레오티드는 고체 서포트에 부착시키거나, 또는 친화성 분리와 중합효소 신장을 가능하게 하는 형태로 변형시킨다. 합성 올리고뉴클레오티드의 5' 말단과 내부 뉴클레오티드는 상이한 친화성 분리 방식을 가능하게 하는 다양한 방식, 예를 들면 비오틴화로 변형시킬 수 있다. 단일 친화성 기가 올리고뉴클레오티드에 사용되면, 올리고뉴클레오티드는 통합된 종료기 반응물로부터 분리할 수 있다. 이는 물리적 또는 크기 배제의 필요성을 제거한다. 1개이상의 올리고뉴클레오티드는 종료기 반응물로부터 분리할 수 있고 1개이상의 친화성 기를 사용하면 동시에 분석할 수 있다. 이는 신장 반응당 여러 핵산 종류 또는 다른 핵산 서열 정보의 분석을 가능하게 한다. 친화성 기는 기폭 올리고뉴클레오티드에는 필요하지 않지만 주형에는 존재할 수 있다. 가령, 고정화는 비오틴화된 DNA와 스트렙타비딘-피복된 마이크로적정(microtitration) 웰 또는 아비딘-피복된 폴리스티렌 입자간 상호작용을 통하여 실시할 수 있다. 동일한 방식으로, 올리고뉴클레오티드 또는 주형은 고밀도 형식으로 고체 서포트에 부착될 수 있다. 이런 고체 상 마이크로염기서열 반응에서, 통합된 ddNTP는 방사성표지하거나(Syvanen, 1994) 또는 플루레신에 결합시킬 수 있다(Livak and Hainer, 1994). 방사성 표지된 ddNTP의 검출은 신털레이션-기초한 기술을 통하여 달성할 수 있다. 플루레신-결합된 ddNTP의 검출은 알칼린 포스파타제와 공액된 항플루레신 항체의 결합 및 이후 색소(chromogenic) 기질(예, p-니트로페닐 인산염)과의 배양에 기초할 수 있다. 다른 가능한 리포터-검출쌍은 다음과 같다: 디니트 로페닐(DNP)에 결합된 ddNTP 및 항-DNP 알칼린 포스파타제 공액체(Harju et al., 1993), 또는 비오틴화된 ddNTP 및 기질로 o-페닐렌디아민을 함유하는 양고추냉이 과산화효소-공액된 스트렙타비딘(쨰 92/15712). 다른 대안적 고체-상 마이크로염기서열분석 과정으로, Nyren et al.(1993)은 효소 발광계 무기 피로포스페이트 검출 분석(ELIDA)에 의한 DNA 중합효소 활성의 검출에 기초한 방법을 기술하였다.

Pastinen et al.(1997)에서는 단일 뉴클레오티드 다형성의 다중 검출 방법을 기술하는데, 여기서 고체 상 미니염기서열분석 원리가 올리고뉴클레오티드 어레이 포맷에 적용된다. 고체 서포트(DNA 칩)에 부착된 DNA 프로브의 고밀도 어레이는 하기에 추가로 설명한다.

한 측면에서, 본원 발명은 마이크로염기서열분석을 실시함으로써 본원 발명의 한가지이상 이중대립인자 마커를 유전자형판별하기 위한 폴리뉴클레오티드와 방법을 제공한다. 적절한 마이크로염기서열분석 프라이머에는 뉴클레오티드 서열 D1 내지 D18과 E1 내지 E18이 포함된다. 실시예 4에 기재된 마이크로염기서열분석 프라이머는 단지 예를 든 것으로, 다형적 폴리뉴클레오티드에 인접하여 3' 말단을 보유하는 임의의 프라이머를 사용할 수 있다. 유사하게, 마이크로염기서열분석은 본원 발명에 따른 임의의 이중대립인자 마커 또는 이런 마커의 조합에 대하여 실시할 수 있다. 본원 발명의 한 측면은 이중대립인자 마커 부위에서 뉴클레오티드의 실체를 확인하기 위한 실시에 4에 기재된 한가지이상 마이크로염기서열분석 프라이머, 또는 적어도 8, 12, 15, 20, 25, 30, 40 또는 50개의 연속 뉴클레오티드를 포함하는 이의 단편을 보유하고(여기서 단편 길이는 전술한 프라이머와 일치한다) 상 응하는 이중대립인자 마커의 바로 상류에 3' 말단을 보유하는 고체 서포트이다.

3) 중합효소와 리가아제에 기초한 미스매치 검출 분석

한 측면에서, 본원 발명은 중합효소 및/또는 리가아제 기초한 미스매치 검출 분석으로 생물학적 시료에서 본원 발명의 한가지이상 이중대립인자 마커의 대립유전자를 확인하기 위한 폴리뉴클레오티드와 방법을 제공한다. 이들 분석은 중합효소와 리가아제의 특이성에 기초한다. 중합화 반응에서는 증폭 프라이머 3' 말단의 정확한 염기 대합(paring)을 특히 엄격하게 요구하고, 표적 DNA 서열에 혼성화된 2가지 올리고뉴클레오티드의 결합은 결찰 부위 인근, 특히 3' 말단에서 미스매체에 매우 민감하다. 본원 발명의 이중대립인자 마커를 포함하는 DNA 단편을 증폭하기 위한 방법, 프라이머, 다양한 파리미터는 '이중대립인자 마커를 포함하는 DNA 단편의 증폭'에서 전술한다.

대립유전자 특이적 증폭 프라이머

이중대립인자 마커의 2가지 대립유전자간 구별 역시 대립유전자 특이적 증폭, 선별 전략으로 달성할 수 있는데, 여기서 대립유전자중 한가지는 다른 대립유전자의 증폭없이 증폭된다. 대립유전자 특이적 증폭을 위하여, 프라이머쌍중 적어도 한가지는 혼성화되고 증폭을 유도할 만큼 본원 발명에 따른 이중대립인자 마커의 다형적 염기를 포함하는 CanIon 유전자 영역과 충분히 상보적이다. 이런 프라이머는 이중대립인자 마커의 2가지 대립유전자를 구별할 수 있다.

이는 증폭 프라이머중 한가지의 3' 말단에 다형적 염기를 위치시켜 달성한다. 프라이머의 3' 말단으로부터 신장이 이루어지기 때문에, 이 위치 인근에서 미 스매치는 증폭에 저해 효과를 보인다. 따라서, 적절한 증폭 조건하에 이들 프라이머는 이들 상보적 대립유전자에 증폭을 유도한다. 미스매체의 정확한 위치 및 상응하는 분석 조건을 결정하는 것은 당분의 통상적 지식에 속한다.

결찰/증폭 기초한 방법

'올리고뉴클레오티드 결찰 분석'(OLA)에서는 표적 분자의 단일 가닥의 인접 서열에 혼성화되도록 설계된 2가지 올리고뉴클레오티드를 사용한다. 올리고뉴클레오티드중 한가지는 비오틴화되고, 다른 한가지는 검출가능하게 표지된다. 정확한 상보적 서열이 표적 분자에서 발견되면, 올리고뉴클레오티드는 혼성화되어 말단이 접하게 되고 포획 및 검출가능한 결찰 기질이 만들어진다. OLA는 단일 뉴클레오티드 다형성을 검출할 수 있는데, Nickerson et al.(1990)에서 밝힌 바와 같이 PCR를 병행하면 유리하다. 이런 방법에서, PCR은 표적 DNA의 기하급수적 증폭을 달성하는데 사용되고, 상기 표적 DNA는 이후 OLA에 의해 검출된다.

단일 뉴클레오티드 다형성의 검출에 특히 적합한 다른 증폭 방법에는 LCR(리가아제 연쇄 반응), 'DNA 증폭'에 전술된 Gap LCR(GLCR)이 포함된다. LCR에서는 2쌍의 프로브를 이용하여 특정 표적을 기하급수적으로 증폭시킨다. 올리고뉴클레오티드 각 쌍의 서열은 표적의 동일 가닥의 인접 서열과의 혼성화가 가능하도록 선별된다. 이런 혼성화는 주형-의존한 리가아제에 대한 기질을 형성한다. 본원 발명에 따라, LCR은 이중대립인자 마커 부위의 동일 가닥의 인접 서열과 원심 서열을 보유하는 올리고뉴클레오티드로 실시할 수 있다. 한 구체예에서, 올리고뉴클레오티드는 이중대립인자 마커 부위를 보유하도록 설계된다. 이런 구체예에서, 반응 조건은 표적 분자가 올리고뉴클레오티드에서 이중대립인자 마커에 상보적인 특정 뉴클레오티드를 보유하거나 또는 결핍하는 경우에만 올리고뉴클레오티드가 서로 결찰될 수 있도록 선별된다. 대안적 구체예에서, 올리고뉴클레오티드는 이중대립인자 마커를 보유하지 않고, 따라서 이들이 표적 분자와 혼성화될 때 '공백'이 만들어진다(WO 90/01069). 이후, 이런 공백은 상보적 dNTP(DNA 중합효소로 매개됨), 또는 다른 올리고뉴클레오티드 쌍으로 '채워진다'. 따라서, 각 사이클의 종결시점에서 각 단일 가닥은 다음 사이클동안 표적으로 기능할 수 있는 보체를 보유하고 원하는 서열의 기하급수적 대립유전자-특이적 증폭이 달성된다.

리가아제/중합효소-매개된 Genetic Bit Analysis ™ 은 핵산 분자의 사전선택된 부위에서 뉴클레오티드의 실체를 확인하는 다른 방법이다. 이런 방법에는 사전선택된 부위에 존재하는 뉴클레오티드에 상보적인 뉴클레오시드 삼인산염의 프라이머 분자 말단으로의 통합 단계 및 두 번째 올리고뉴클레오티드로의 후속적 결찰 단계가 포함된다. 반응은 고체상에 부착된 특정 라벨을 검출하여, 또는 용액에서 검출로 모니터한다.

4) 혼성화 분석 방법

이중대립인자 마커 부위에 존재하는 뉴클레오티드의 실체를 확인하는데 적합한 방법에는 핵산 혼성화가 포함된다. 이런 반응에 간편하게 사용될 수 있는 혼성화 프로브에는 가급적, 본원에 정의된 프로브가 포함된다. 임의의 혼성화 분석, 예를 들면 서던 혼성화, 노던 혼성화, 점 블랏 혼성화, 고체상 혼성화를 이용할 수 있다(Sambrook et al., 1989).

혼성화는 상보적 염기 대합에 기인한 2가지 단일 가닥 핵산의 이중나선구조 형성을 의미한다. 혼성화는 정확하게 상보적 핵산 가닥 사이에 또는 약간의 미스매치 영역을 보유하는 핵산 가닥 사이에서 발생할 수 있다. 이중대립인자 마커의 한가지 형태에는 혼성화되지만 다른 형태에는 혼성화되지 않고, 따라서 상이한 대립유전자 형태를 구별할 수 있는 특이적인 프로브를 설계할 수 있다. 대립유전자-특이적 프로브는 종종 쌍으로 사용되는데, 이런 쌍의 한 구성원은 원형 대립유전자를 보유하는 표적 서열에 완전한 매치를 보이고, 다른 구성원은 대안적 대립유전자를 보유하는 표적

서열과 완전한 매치를 보인다. 혼성화 조건은 대립유전자간 혼성화 강도에서 현저한 차이, 바람직하게는 프로브가 대립유전자중 한가지에만 혼성화되는 이항 반응이 존재할 만큼 충분히 엄격해야 한다. 프로브가 정확하게 상보적 표적 서열에만 혼성화되는 엄격한 서열 특이적 혼성화 조건은 당분야에 공지된 것이다(Sambrook et al., 1989). 엄격한 조건은 서열 의존성으로, 개별 환경에 따라 달라진다. 일반적으로, 엄격한 조건은 정의된 이온 강도와 pH에서 특정 서열에 대한 열 융점(Tm)보다 5℃ 정도 낮게 선택된다. 이런 혼성화가 용액에서 실시될 수 있긴 하지만, 고체상 혼성화 분석을 이용하는 것이 바람직하다. 본원 발명의 이중대립인자 마커를 포함하는 표적 DNA는 혼성화 반응에 앞서 증폭할 수 있다. 시료에서 특정 대립유전자의 존재는 프로브와 표적 DNA간에 형성된 안정한 하이브리드 이중나선의 존부를 검출하여 확인한다. 하이브리드 이중나선의 검출은 다양한 방법으로 실시할 수 있다. 하이브리드 이중나선을 검출을 가능하게 하는 표적 이나 프로브에 결합된 검출가능한 라벨을 사용한 다양한 검출 분석 포맷은 당분야에 공지된 것이다. 전형적으로, 혼성화 이중나선은 혼성화되지 않은 핵산으로부터 분리하고, 이후 이중나선에 결합된 라벨은 검출한다. 당업자는 인지하는 바와 같이, 세척 단계를 활용하여 과잉 DNA 또는 프로브 및 결합되지 않은 공액체를 씻어낼 수 있다. 또한, 표준 이형이질성 분석 포맷은 프라이머와 프로브에 존재하는 라벨을 이용하여 하이브리드를 검출하는데 적합하다.

최근에 개발된 2가지 분석법은 분리 또는 세척없이 혼성화-기초된 대립유전자 구별이 가능하다(Landegren U. et al., 1998). TaqMan 분석은 축적 증폭 산물에 특이적으로 어닐링되는 DNA 프로브를 절단하는 Taq DNA 중합효소의 5' 뉴클레아제 활성을 활용한다. TaqMan 프로브는 형광 에너지 전달을 통하여 상호작용하는 공여자-수용자 염료 쌍으로 표지한다. 증폭동안 진행(advancing) 중합효소에 의한 TaqMan 프로브의 절단은 켄칭(quenching) 수용자 염료로부터 공여자 염료를 분리시키고 공여자 형광을 현저하게 증가시킨다. 2가지 대립유전자 변이체를 검출하는데 필요한 모든 반응물은 반응의 시작시점에서 집합시킬 수 있고, 결과는 실시간으로 모니터한다(Livak et al., 1995). 대안적인 상동성 혼성화 기초한 과정에서, 분자 표지(molecular beacon)가 대립유전자 구별에 사용된다. 분자 표지는 상동성 용액에서 특정 핵산의 존재를 보고하는 헤어핀-형태의 올리고뉴클레오티드 프로브이다. 이들이 표적에 결합하면, 내부적으로 켄칭된 형광단의 형광을 회복하는 입체형태적 재조직이 진행된다(Tyagi et al., 1998).

본원에 제시된 폴리뉴클레오티드는 생물학적 시료에서 이중대립인자 마커의 검출을 위한 혼성화 분석에 사용될 수 있는 프로브를 생산하는데 이용될 수 있다. 이들 프로브는 8 내지 50개의 뉴클레오티드를 포함하고 본원 발명의 이중대립인자 마커를 포함하는 서열과 혼성화될 만큼 충분히 상보적이며 표적 서열에서 단지 1개의 뉴클레오티드 변이를 구별할 수 있을 만큼 충분히 특이적이라는 점으로 특징지어진다. 특히 적절한 프로브는 길이가 25개 뉴클레오티드이다. 적절하게는, 이중대립인자 마커는 폴리뉴클레오티드 프로브의 중심으로부터 4개 뉴클레오티드 이내에 위치한다. 특히 적절한 프로브에서, 이중대립인자 마커는 상기 폴리뉴클레오티드의 중심에 위치한다. 적절한 프로브는 표 1에 기재된 앰플리콘, 이의 상보적 서열 또는 이의 단편에서 선택되는 뉴클레오티드를 포함하는데, 상기 단편은 적어도 8개의 연속 뉴클레오티드, 바람직하게는 10, 15, 20개 연속 뉴클레오티드, 좀더 바람직하게는 25, 30, 40, 47 또는 50개 연속 뉴클레오티드를 포함하고 다형적 염기를 보유한다. 적절한 프로브는 P1 내지 P8 및 이의 상보적 서열에서 선택되는 뉴클레오티드 서열을 포함한다. 적절한 구체예에서, 다형적 염기는 상기 폴리뉴클레오티드의 중심으로부터 5, 4, 3, 2, 1개 뉴클레오티드 이내, 좀더 바람직하게는 상기 폴리뉴클레오티드의 중심에 위치한다.

적절하게는, 본원 발명의 프로브는 표지되거나 또는 고체 서포트에 고정된다. 라벨과 고체 서포트는 '올리고뉴클레오티드 프로브와 프라이머'에서 추가로 설명한다. 프로브는 '올리고뉴클레오티드 프로브와 프라이머'에서 밝힌 바와 같이 신장-불가능할 수 있다.

대립유전자 특이적 프로브에 대한 혼성화를 분석함으로써, 일정한 시료에서 이중대립인자 마커 대립유전자의 존부를 검출할 수 있다. 어레이 포맷에서 고성능 병렬 혼성화는 '혼성화 분석법'에 포함되고 하기에 상술한다.

5) 올리고뉴클레오티드의 어드레스가능 어레이의 혼성화

올리고뉴클레오티드 어레이에 기초한 혼성화 분석은 완전하게 대합되고 미스매치된 표적 서열 변이체에 대한 짧은 올리고뉴클레오티드의 혼성화 안정성에서 차이에 기초한다. 다형성 정보에 대한 효과적 접근은 선택된 위치에서 고체 서포트에 부착된 올리고뉴클레오티드 프로브의 고밀도 어레이(예, 칩)를 포함하는 기본 구조를 통하여 달성한다. 각 DNA 칩은 격자-유사 패턴으로 정렬되고 동전 크기로 축소된 수천 내지 수백만개의 개별 합성 DNA 프로브를 보유할 수 있다.

칩 기술은 다양한 사례에 성공적으로 적용되었다. 가령, 돌연변이의 스크리닝은 맥주 효모균(*S. cerevisiae*) 변이 균주에서 BRCA1 유전자 및 HIV-1 바이러스의 프로테아제 유전자에서 실시되었다(Hacia et al., 1996; Shoemaker et al., 1996; Kozal et al., 1996). 이중대립인자 다형성을 검출하는데 사용되는 다양한 포맷의 칩은 Affymetrix(Gene Chip ™), Hyseq(HyChip and HyGnostics), Protogene Laboratories에 의해 맞춤 생산될 수 있다.

일반적으로, 이들 방법에서는 표적 서열이 다형적 마커를 보유하는 개체로부터 유래된 표적 핵산 서열 분절에 상보적인 올리고뉴클레오티드 프로브의 어레이를 이용한다. EP 785280에서는 단일 뉴클레오티드 다형성의 검출을 위한 기

와(tiling) 전략을 기술한다. 간단히 말하면, 어레이는 다수의 특이적 다형성이 '기와처럼 배열된다'. 일반적으로, '기와'는 관심있는 표적 서열에 상보적인 서열 및 이런 서열의 사전선별된 변이, 예를 들면 기초 뉴클레오티드의 한가지이상 구성원으로 한가지이상 정해진 위치의 치환으로 구성되는 정의된 일단의 올리고뉴클레오티드 프로브의 합성을 의미한다. 기와 전략은 PCR 출원 95/11955에서 좀더 상세하게 기술한다. 특정 측면에서, 어레이는 다수의 특이적인 동정된 이중대립인자 마커 서열이 기와처럼 배열된다. 특히, 어레이는 특정 다형성을 보유하는 서열 분절을 스팬하는 다수의 프로브가 기와처럼 배열된다. 각 대립유전자에 상보적인 프로브를 담보하기 위하여, 상기 프로브는 이중대립인자 마커에서 차별되는 쌍으로 합성된다. 다형적 염기에서 차별되는 프로브 이외에, 단일치환된 프로브 역시 검출 블록에 기와처럼 배열된다. 이들 단일치환된 프로브는 다형성으로부터 양방향으로, 다른 뉴클레오티드(A, T, G, C(또는 U)에서 선택)로 치환된 일정한 수의 염기를 보유한다. 전형적으로, 기와처럼 배열되는 검출 블록에서 프로브는 이중대립인자 마커로부터 최대 5개 염기까지 서열의 치환을 보유한다. 단일치환된 프로브는 인공 교차-혼성화(cross-hybridization)로부터 실제 혼성화를 구별하는 기와처럼 배열된 어레이에 대한 내부 대조군을 제공한다. 표적 서열과의 혼성화의 완결 및 어레이의 세척 직후, 어레이는 표적 서열이 혼성화되는 위치를 스캔한다. 이후, 스캔된 어레이로부터 얻은 혼성화 데이터는 분석하여 이중대립인자 마커의 대립유전자가 시료에 존재하는 지를 확인한다. 혼성화와 스캐닝(scanning)은 PCR 출원 WO 92/10092와 WO 95/11955; 미국 특허 5,424,186에서 밝힌 바와 같이 실시할 수 있다.

따라서, 일부 구체예에서 칩은 대략 15개 뉴클레오티드 길이의 핵산 단편 서열의 어레이를 포함할 수 있다. 다른 구체예에서, 칩은 표 1에 기재된 앰플리콘, 이의 상보적 서열 또는 이의 단편에서 선택되는 서열중 적어도 한가지를 보유하는 어레이를 포함할 수 있는데, 상기 단편은 적어도 8개의 연속 뉴클레오티드, 바람직하게는 10, 15, 20개 연속 뉴클레오티드, 좀더 바람직하게는 25, 30, 40, 47 또는 50개 연속 뉴클레오티드를 포함하고 다형적 염기를 보유한다. 적절한 프로브는 P1 내지 P8 및 이의 상보적 서열에서 선택되는 뉴클레오티드 서열을 포함한다. 적절한 구체예에서, 다형적 염기는 상기 폴리뉴클레오티드의 중심으로부터 5, 4, 3, 2, 1개 뉴클레오티드 이내, 좀더 바람직하게는 상기 폴리뉴클레오티드의 중심에 위치한다. 일부 구체예에서, 칩은 본원 발명에 따른 폴리뉴클레오티드중 적어도 2, 3, 4, 5, 6, 7, 8 또는 이보다 많은 폴리뉴클레오티드로 구성되는 어레이를 포함한다.

고체 서포트 및 고체 서포트에 부착되는 본원 발명의 폴리뉴클레오티드는 '올리고뉴클레오티드 프로브와 프라이머'에서 추가로 설명한다.

6) 통합 시스템

다형성을 분석하는데 사용할 수 있는 다른 기술에는 단일 기능 장치에서 PCR과 모세관 전기영동 반응과 같은 과정을 축소되고 구획한 다성분 통합 시스템이 포함된다. 이런 기술의 예는 미국 특허 5,589,136에서 개시하는데, 이는 칩에서 PCR 증폭. 및 모세관 전기영동의 통합을 기술한다.

통합 시스템은 마이크로유체 시스템을 사용하는 경우에 주로 고려할 수 있 다. 이들 시스템은 마이크로칩에 포함되는 유리, 실리콘, 석영 또는 플라스틱 와이퍼에 설계된 마이크로채널의 패턴을 포함한다. 시료의 움직임은 마이크로칩의 전체 영역에 가해지는 전기력, 전기삼투력 또는 부력으로 조절하여, 회전체(moving part)가 없는 기능성 현미경적 밸브와 펌프를 만든다.

유전자형판별 이중대립인자 마커에서, 마이크로유체 시스템은 핵산 증폭, 마이크로염기서열분석, 모세관 전기영동 및 레이저-유도된 형광 검출과 같은 검출 방법을 통합할 수 있다.

본원 발명에 따른 이중대립인자 마커를 이용한 유전자 분석 방법

복합 형질의 유전자 분석에 여러 방법이 가용하다(Lander and Schork, 1994). 질병-취약성 유전자의 검색은 2가지 주요 방법으로 실시한다: 가족 연구를 활용하여 좌위와 추정 형질 좌위간 동시분리(cosegregation)의 증거를 추적하는 연쇄 방식 및 대립유전자와 대립유전자를 유발하는 형질간 통계학적으로 유의한 연관의 증거를 추적하는 연관 방식(Khoury et al., 1993). 일반적으로, 본원 발명의 이중대립인자 마커는 유전형과 표현형간 통계학적으로 유의한 상관관계를 예시하는 당분야에 공지된 방법에 사용될 수 있다. 이중대립인자 마커는 파라메트릭(parametric)과 비-파라메트릭 연쇄 분석 방법에 사용될 수 있다. 적절하게는, 본원 발명의 이중대립인자 마커는 연쇄 연구를 이용하여 검출가능한 형질과 연관된 유전자를 동정하는데 사용되는데, 이런 방식은 영향을 받은 가족의 참여를 필요로 하지 않고 복합적 형질과 산발적 형질과 연관된 유전자의 동정이 가능하다.

본원 발명의 이중대립인자 마커를 이용한 유전자 분석은 임의의 스케일로 실시할 수 있다. 본원 발명에 따른 이중대립인자의 전체 세트, 또는 후보 유전자에 상응하는 본원 발명에 따른 이중대립인자 마커의 임의 서브세트를 사용할 수 있다. 또한, 본원 발명의 이중대립인자 마커를 보유하는 유전자 마커의 임의 세트를 사용할 수도 있다. 본원 발명의 이중대립인자 마커와 함께 유전자 마커로 사용될 수 있는 일단의 이중대립인자 다형성은 WO 98/20165에서 기술한다. 전술한 바와 같이, 본원 발명의 이중대립인자 마커는 사람 게놈의 전체 또는 부분 유전자 지도에 포함될 수 있다.

이들 별개의 용도는 본원 발명 및 특허청구범위에서 특이적으로 정관한다.

연쇄 연구

연쇄 연구는 가족내 여러 세대에서 유전자 마커의 전달과 특정 형질의 전달간 상관관계를 확립하는데 기초한다. 따라서, 연쇄 연구의 목적은 가계에서 관심있는 형질로 동시분리(codegregation)를 보이는 마커 좌위를 검출하는 것이다.

파라메트릭 방법

데이터가 연속 세대로부터 가능하면, 여러 쌍의 좌위간 연쇄 정도를 연구할 수 있다. 재조합 분류(recombinant fraction)의 평가로 좌위를 정리하고 유전자 지도에 배치할 수 있다. 유전자 마커인 좌위로 유전자 지도를 확립할 수 있고, 이후 마커와 형질간 연쇄 강도를 계산할 수 있는데, 이는 이들 형질에 영향을 주는 마커와 유전자의 상대적 위치를 표시하는데 사용된다(Weir, 1996). 연쇄 분석의 고전적 방법은 홀수(lod) 스코어 방법의 대수이다(Morton, 1955; Ott, 1991). lod 스코어의 계산은 질환에 대한 유전 모드의 명시를 필요로 한다(파라메트릭 방법). 일반적으로, 연쇄 연구로 동정된 후보 영역의 길이는 2 내지 20 Mb이다. 후보 영역이 전술한 바와 같이 동정되면, 추가의 마커를 이용한 재조합 개체의 분석으로 후보 영역의 상세한 설명이 가능하다. 일반적으로, 연쇄 분석 연구는 최대 5,000개 마이크로위성 마커의 이용에 기초하고, 따라서 연쇄 분석의 이론적으로 가능한 최대 분석능(resolution)은 평균 600 kb이다.

연쇄 연구는 분명한 멘델식 유전 패턴을 보이고 높은 발현빈도(penetrance)(즉, 대립유전자 *a* 의 형질 파지티브 보인자의 개체수와 개체군에서 *a* 보인자의 총 개체수간 비율)를 갖는 간단한 유전 형질의 지도를 작성하는데 성공적으로 이용되고 있다. 하지만, 파라메트릭 연쇄 분석은 많은 단점이 있다. 먼저, 이는 각 연구 형질에 적합한 유전자 모델의 선택에 대한 의존으로 제한된다. 게다가, 전술한 바와 같이 연쇄 분석을 이용하여 달성가능한 분석능(resolution)이 제한되고, 상보성 연구는 연쇄 분석을 통하여 처음에 동정된 2Mb 내지 20Mb의 분석을 정련해야 한다. 이에 더하여, 파라메트릭 연쇄 분석 방식은 복합적 유전 형질, 예를 들면 다중 유전자 및/또는 환경 인자의 복합 작용에 기인한 형질에는 적용하기 어렵다. 로드(lod) 스코어 분석에서 이들 인자를 적절히 설계하는 것은 매우 어렵다. 이런 경우에, 이런 상황에 연쇄 분석을 적용하는데 필요한 적절한 수의 가족을 확보하는데 너무 많은 노력과 비용이 소요된다(Risch and Merikangas. (1996)).

비-파라메트릭 방법

연쇄 분석을 위한 소위 비-파라메트릭 방법의 이점은 이들이 질환에 대한 유 전 모드의 명시를 필요로 하지 않고 복합적 형질의 분석에 좀더 적합하다는 점이다. 비-파라메트릭 방법에서는 영향을 받은 근친이 우연히 기대되는 것보다 좀더 빈번하게 크로모좀 영역의 동일한 사본을 유전한다는 것을 보임으로써 크로모좀 영역의 유전 패턴이 무작위 만델식 분리와 일치하지 않는다는 것을 입증한다. 영향을 받은 근친은 불완전한 발현빈도 및 다인자 유전(polygenic inheritance)의 존재에서도 과도한 '대립유전자 공유(allele sharing)'를 보인다. 비-파라메트릭 연쇄 연구에서 2명의 개체에서 마커 좌위에서 일치 정도는 상태에서 동일한(IBS) 대립유전자 총수 또는 세대에서 동일한(IBD) 대립유전자 총수로 측정할 수 있다. 영향을 받은 근친 쌍 분석(affected sib pair analysis)은 공지된 특별한 사례로서, 이들 방법 중 가장 간단한 형태이다.

본원 발명의 이중대립인자 마커는 파라메트릭과 비-파라메트릭 연쇄 분석에 모두 사용될 수 있다. 적절하게는, 이중대립인자 마커는 복합적 형질에 관여하는 유전자의 지도작성을 가능하게 하는 비-파라메트릭 방법에 사용될 수 있다. 본원 발명의 이중대립인자 마커는 복합적 형질에 영향을 주는 유전자의 지도를 작성하는 IBD-와 IBS-방법에 사용될 수 있다. 이런 연구에서, 고밀도의 이중대립인자 마커로 여러 인접 이중대립인자 마커 좌위를 집합시켜, 다중-대립유전자 마커에 의해 달성되는 효율을 성취할 수 있다(Zhao et al., 1998).

개체군 연관 연구

본원 발명에는 본원 발명의 이중대립인자 마커를 이용하여 CanIon 유전자가 검출가능한 형질과 연관하는 지를 확인하는 방법이 포함된다. 한 구체예에서, 본 원 발명에는 이중대립인자 마커 대립유전자 또는 이중대립인자 마커 일배체형 및 형질간 연관을 검출하는 방법이 포함된다. 게다가, 본원 발명에는 본원 발명의 이중대립인자 마커와 연쇄 불균형한 형질 유발 대립유전자를 동정하는 방법이 포함된다.

전술한 바와 같이, 연관 연구를 실시하는데 대안적 방법을 이용할 수 있다: 게놈-전체 연관 연구, 후보 영역 연관 연구, 후보 유전자 연관 연구. 적절한 구체예에서, 본원 발명의 이중대립인자 마커는 후보 유전자 연관 연구를 실시하는데 사용된다. 후보 유전자 분석은 형질의 생태에 관한 정보가 일부 가용한 경우에, 특정 형질과 관련된 유전자 및 유전자 다형성의 동정에 지름길을 제공한다. 게다가, 본원 발명의 이중대립인자 마커는 게놈전체 연관 연구를 실시하기 위하여 사람 게놈의 유전자 마커의 임의 지도에 통합될 수 있다. 이중대립인자 마커의 고밀도 지도를 만드는 방법은 미국 특허 분할 출원 60/082,614에서 기술한다. 본원 발명의 이중대립인자 마커는 게놈의 특정 후보 영역(예, 특정

크로모좀 또는 특정 크로모좀 분절)의 임의 지도에 통합될 수도 있다.

전술한 바와 같이, 연관 연구는 전체 개체군에서 실시할 수 있고 영향을 받은 가족에서 관련된 개체에 실시된 연구로 한정되지 않는다. 연관 연구는 산발적 또는 다중적 형질의 분석이 가능하다는 점에서 매우 귀중하다. 게다가, 연관 연구는 연쇄 연구보다 형질 유발 대립유전자의 좀더 세밀한 유전지도작성을 가능하게 하는 소규모 유전지도작성(fine-scale mapping)에 유력한 방법이다. 가계에 기초한 연구는 종종 형질 유발 대립유전자의 위치를 정밀하게 좁힌다. 따라서, 본원 발명의 이중대립인자 마커를 이용한 연관 연구는 연쇄 분석 방법으로 동정된 후보 영역에서 형질 유발 대립유전자의 위치를 정련하는데 사용될 수 있다. 관심있는 크로모좀 분절이 동정되면, 관심있는 영역에서 후보 유전자, 예를 들면 본원 발명의 후보 유전자의 존재는 형질 유발 대립유전자의 동정에 지름길을 제공할 수 있다. 본원 발명의 이중대립인자 마커는 후보 유전자가 형질과 연관한다는 것을 입증하는데 사용될 수 있다. 이런 용도는 본원에서 구체적으로 정관한다.

개체군에서 이중대립인자 마커 대립유전자 또는 이중대립인자 마커 일배체형의 빈도 측정

연관 연구는 좌위간 여러 대립유전자의 빈도간 상관관계를 탐구한다.

<u>개체군에서 대립유전자의 빈도 측정</u>

개체군에서 이중대립인자 마커의 대립유전자 빈도는 '이중대립인자 마커에 대하여 개체를 유전자형판별하는 방법'에서 전술한 방법, 또는 이런 의도한 목적에 적합한 유전자형판별 과정으로 측정할 수 있다. 집합된 시료 또는 개별 시료의 유전자형판별은 개체군에서 이중대립인자 마커의 빈도를 측정할 수 있다. 집합된 시료를 이용할 때 주요 장애물은 풀을 정립하는데 있어 정확한 DNA 농도를 측정하기 위한 정확도와 반복재현성이다. 개별 시료의 유전자형판별은 좀더 높은 감수성, 반복 재현성, 정확도를 제공하고, 본원 발명에 사용하기 적합한 방법이다. 적절하게는, 각 개체는 개별적으로 유전자형판별하고, 간단한 유전자 계수(conuting)를 적용하여 임의의 개체군에서 이중대립인자 마커의 대립유전자 또는 유전형의 빈도를 측정한다.

또한, 본원 발명은 개체군에서 대립유전자의 빈도를 측정하는 방법에 관하는데, 상기 방법은 다음과 같이 구성된다: a) 본원 발명에 따른 이중대립인자 마커의 개체군으로부터 개체를 유전자형판별하고; b) 상기 개체군에서 이중대립인자 마커의 비례 표현(proportional representation)을 측정한다. 이에 더하여, 본원 발명의 개체군에서 대립유전자의 빈도를 측정하는 방법에는 제한없이 본원 발명에 개시된 단독의 또는 조합된 방법이 포함되는데, 여기서 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 개체군에서 이중대립인자 마커 대립유전자의 빈도 측정은 개체군에서 각 개체의 게놈에 존재하는 이중대립인자 마커의 양 사본에 대한 뉴클레오티드의 실체를 확인하고 상기 개체군에 대하여 CanIon-관련된 이중대립인자 마커에서 상기 뉴클레오티드의 비례 표현을 계산하여 달성할 수 있다; 선택적으로, 비례 표현 확인은 상기 개체군에서 대표적 인원수의 개체 또는 각 개체로부터 유래된 집합된 생물학적 시료에 본원의 유전자형판별 방법을 실시하고 전체와 비교된 상기 뉴클레오티드의 비례 함량을 계산하여 달성할 수 있다.

<u>개체군에서 일배체형의 빈도 측정</u>

이배체형 개체가 하나이상의 좌위에서 이형접합성이면 일배체형의 배우자 위 상은 알지 못한다. 가족에서 가계 정보를 이용하면 배우자 위상을 추론할 수 있다(Perlin et al., 1994). 가계 정보가 가용하지 않으면 상이한 전략을 이용할 수 있다. 한가지 가능성은 분석으로부터 다중-부위 이형접합성 이배체형을 배제하고 동형접합체와 단일-부위 이형접합 개체만을 유지하는 것이지만, 이런 방식은 시료 조성에서 선입견 및 저밀도 일배체형의 과소평가를 초래할 가능성이 있다. 다른 가능성은 예로써 비대칭 PCR 증폭(Newton et al, 1989; Wu et al., 1989) 또는 제한 희석과 이후 PCR 증폭에 의한 단일 크로모좀의 분리(Ruano et al., 1990)로 독립적으로 단일 크로모좀을 조사하는 것이다. 게다가, 시료는 특정 대립유전자의 이중 PCR 증폭으로 충분히 밀접한 이중대립인자 마커에 대하여 일배체형화시킬 수 있다(Sarkar, G. and Sommer S. S., 1991). 이들 방식은 기술적 복잡성, 추가적인 비용, 대규모로 보편화의 부족 또는 가능한 선입견으로 인해 완전히 만족스럽지 못하다. 이런 어려움을 극복하기 위하여, Clark, A.G.(1990)에 의해 도입된 PCR-증폭된 DNA 유전형의 위상을 추론하는 대수를 이용할 수 있다. 간단히 말하면, 이런 원리는 분명한 개체, 다시 말하면 완전 동형접합체와 단일-부위 이형접합체를 검사함으로써 시료에 존재하는 일배체형의 예비 리스트를 기입하는 것으로 시작한다. 이후, 동일 시료에서 다른 개체는 이전에 인식된 일배체형의 가능 발생을 스크린한다. 각 파지티브 동정을 위하여, 상보적 일배체형은 모든 개체에 대한 위상 정보가 해결되거나, 또는 해결되지 않는 것으로 인식될 때까지 인식된 일배체형의 리스트에 추가한다. 상기 방법에서는 각 다중 이형접합성 개체에 단일 일배체형을 할당하는데, 여기서 1개이상의 이형접합성 부위가 존재하면 여러 일배체형이 가능하 다. 대안으로, 각 개체에 일배체형을 할당하지 않고 개체군에서 일배체형 빈도를 측정하는 방법을 이용할 수 있다. 적절하게는, Hardy-Weinberg 비율(임

의 교배(random mating))의 가정하에 일배체형 빈도의 최대 공산치(maximum likelihood estimate)를 결과하는 기대치-최대화(expectation-maximization)(EM) 대수에 기초한 방법이 이용된다(Exocoffier L. and Slatkin M., 1995). EM 대수는 데이터가 불명확/불완전한 경우에 유용한 일반화된 반복 최대 공산 측정 방법이다. EM 대수는 이형접합체를 일배체형으로 분석하는데 사용된다. 일배체형 측정은 '통계학적 방법'에서 상세하게 후술한다. 개체군에서 일배체형의 빈도를 측정하거나 또는 평가하는 다른 공지된 방법을 이용할 수 있다.

본원 발명에는 또한, 개체군에서 일단의 이중대립인자 마커에 대한 일배체형의 빈도를 측정하는 방법이 포함되는데, 상기 방법은 다음의 단계로 구성된다: a) 개체군에서 각 개체에 대하여 본원 발명의 방법에 따라 적어도 1개의 CanIon-관련된 이중대립인자 마커를 유전자형판별하고; b) 상기 개체군에서 각 개체의 게놈에 존재하는 두 번째 이중대립인자 마커의 양 사본에 대하여 두 번째 이중대립인자 마커에서 뉴클레오티드의 실체를 확인함으로써 두 번째 이중대립인자 마커를 유전자형판별하며; c) a)와 b) 단계에서 확인된 뉴클레오티드의 실체에 일배체형 확인 방법을 적용하고; d) 상기 빈도의 추정값을 얻는다. 이에 더하여, 본원 발명의 일배체형의 빈도를 측정하는 방법에는 제한없이 본원 발명에 개시된 단독의 또는 조합된 방법이 포함되는데, 여기서 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선 택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 일배체형 확인 방법은 비대칭 PCR 증폭, 특정 대립유전자의 이중 PCR 증폭, Clark 대수, 또는 기대치-최대화 대수로 실시한다.

연쇄 불균형 분석

연쇄 불균형은 2개이상 좌위에서 대립유전자의 비-무작위 연관으로, 질환 형질에 관여하는 유전자의 지도 작성에 유력한 도구이다(Ajioka R.S. et al., 1997). 이중대립인자 마커는 사람 게놈에서 밀집하여 위치하고 다른 유형의 유전자 마커보다 좀더 많은 수량으로 유전자형판별될 수 있기 때문에, 연쇄 불균형에 기초한 유전자 분석에 특히 유용하다.

질환 돌연변이가 개체군에 먼저 도입되면(새로운 돌연변이 또는 돌연변이 담체의 이동으로 인해), 이는 단일 크로모좀 및 연결된 마커의 '배경'이나 '조상' 일배체형에 위치하게 된다. 결과적으로, 이들 마커와 질환 돌연변이간에는 완전 불균형이 존재한다: 질환 돌연변이는 일단의 특이적인 마커 대립유전자가 존재하는 경우에만 확인된다. 이후 세대에서 질환 돌연변이와 이들 마커다형성간에 재조합 현상이 발생하고, 불균형은 점진적으로 소멸된다. 이런 소멸의 속도는 재조합 빈도의 함수인데, 질환 유전자에 가장 가까운 마커가 좀더 멀리 떨어진 마커보다 좀더 높은 수준의 불균형을 나타낸다. 재조합으로 파괴되지 않으면 '조상' 일배체형 및 상이한 좌위에서 마커 대립유전자간 연쇄 불균형은 가계뿐만 아니라 개체군을 통하여 추적할 수 있다. 일반적으로, 연쇄 불균형은 한 좌위에서 한가지 특이적 대립유전자와 두 번째 좌위에서 다른 특이적 대립유전자간 연관으로 간주된다.

질환과 마커 좌위간 불균형의 패턴이나 곡선은 질환 좌위에서 최대값을 보일 것으로 예상된다. 결과적으로, 질환 대립유전자와 밀접하게 연관된 유전자 마커간 연쇄 불균형의 총계는 질환 유전자의 위치와 관련된 귀중한 정보를 제공할 수 있다. 질환 좌위의 소규모 유전지도작성을 위하여, 연구 영역에서 마커간에 존재하는 연쇄 불균형의 패턴과 관련된 얼마간의 지식을 확보하는 것이 유리하다. 전술한 바와 같이, 연쇄 불균형 분석을 통하여 달성된 유전지도작성 분석능은 연쇄 연구에서보다 훨씬 높다. 연쇄 불균형 분석과 함께 고밀도의 이중대립인자 마커는 소규모 유전지도작성에 유력한 도구를 제공한다. 연쇄 불균형을 측정하는 다른 방법은 '통계학적 방법'에서 상세하게 후술한다.

형질-마커 연관의 개체군-기초한 사례-대조 연구

전술한 바와 같이, 동일한 크로모좀의 상이한 좌위에서 여러 쌍의 특정 대립유전자의 빈도는 무작위가 아니고, 무작위로부터 편차는 연쇄 불균형(linkage disequilibrium)이라 한다. 연관 연구는 개체군 빈도에 집중하고 연쇄 불균형 현상에 기초한다. 일정한 유전자에서 특정 대립유전자가 특정 형질을 초래하는데 직접 관여하면, 이의 빈도는 형질 네거티브 개체군 또는 무작위 대조 개체군에서 빈도와 비교하여 영향을 받은(형질 파지티브) 개체군에서 통계학적으로 증가하게 된다. 연쇄 불균형의 존재로 인하여, 형질-유발 대립유전자를 보유하는 일배체형에 존재하는 모든 다른 대립유전자의 빈도 역시 형질 네거티브 개체 또는 무작위 대조군과 비교하여 형질 파지티브 개체에서 증가한다. 따라서, 형질 및 형질-유발 대립유전자와 연쇄 불균형인 임의의 대립유전자(특히, 이중대립인자 마커 대립유전자)간 연관은 특정 영역에서 형질-관련된 유전자의 존재를 암시한다. 사례-대조 개체군은 이중대립인자 마커에 대하여 유전자형판별하여 형질 유발 대립유전자에 근접하여 위치하는 연관을 확인할 수 있다. 형질과 연관된 일정한 한가지 마커와 연쇄 불균형인 임의의 마커는 형질과 연관한다. 연쇄 불균형은 형질-유발 대립유전자를 발견하기 위한 모든 가능한 기능적 다형성의 스크리닝에 대한 대안으로 한정된 유전자 다형성(특히, 이중대립인자 마커)의 사례-대조 개체군에서 상대적 빈도를 분석하는 것을 가능하게 한다. 연관 연구는 무관한 사례-대조 개체군에서 마커 대립유전자의 빈도를 비교하고, 복합적 형질의 정밀한 분석에 유력한 도구를 제공한다.

사례-대조 개체군(산입 기준)

개체군-기초한 연관 연구는 가족성 유전과는 무관하지만 사례-대조 개체군에서 특정 유전자 마커 또는 일단의 마커의 우세를 비교한다. 이들은 무관한 사례(영향을 받은 또는 형질 파지티브) 개체와 무관한 대조군(영향을 받지 않은, 형질 네거티브 또는 무작위) 개체의 비교에 기초한 사례-대조 연구이다. 적절하게는, 대조군은 영향을 받지 않은 또는 형질 네거티브 개체로 구성된다. 또한, 대조군은 사례 개체군과 인종적으로 대합된다. 이에 더하여, 대조군은 연구중인 형질에 대한 공지된 주요 혼란 인자(confusion factor)에 대한 사례-개체군에 대합된다(예, 연령-의존성 형질에 대하여 연령-대합된다). 이상적으로, 2가지 시료에서 개체는 질환 상태에서만 구별되는 방식으로 쌍을 이룬다. '형질 파지티브 개체군', '사례 개체군', '영향을 받은 개체군'은 동일한 의미이다.

연관 연구를 이용한 복합적 형질의 정밀한 분석에서 중요한 단계는 사례-대조 개체군의 선택이다(Lander and Schork, 1994). 사례-대조 개체군의 선택에서 주요 단계는 일정한 형질이나 표현형의 임상적 정의이다. 임의의 유전 형질은 형질 파지티브와 형질 네거티브 군에 포함되는 개체를 신중하게 선별하고 본원에 제시된 연관 방법을 이용하여 분석할 수 있다. 일반적으로, 4가지 기준이 사용된다: 임상적 표현형, 발병 연령, 가족력, 심각도. 연속 또는 정량적 형질(예, 혈압)에 대한 선별 과정은 연구중인 형질의 표현형 분포에서 양 극단의 개체를 선별하여 중복되지 않은 표현형을 갖는 개체가 형질 파지티브와 형질 네거티브 개체군에 포함되도록 하는 단계를 포함한다. 적절하게는, 사례-대조 개체군은 표현형에서 상동한 개체군으로 구성된다. 형질 파지티브와 형질 네거티브 개체군은 각각, 연구중인 전체 개체군중 1 내지 98%, 바람직하게는 1 내지 80%, 좀더 바람직하게는 1 내지 50%, 이보다 좀더 바람직하게는 1 내지 30%, 가장 바람직하게는 1 내지 20%을 대표하는 개체의 표현형에서 상동한 개체군으로 구성되고, 선별된 개체는 중복되지 않은 표현형을 보인다. 2가지 형질 표현형간 차이가 분명할수록 이중대립인자 마커로 연관을 검출할 가능성이 높아진다. 현저하게 상이하지만 상대적으로 상동한 표현형의 선별은 연구중인 개체군의 시료 크기가 충분히 크다는 조건하에, 연과 연구에서 효율적인 비교 및 유전자 수준에서 현저한 차이의 검출을 가능하게 한다.

적절한 구체예에서, 50 내지 300명의 형질 파지티브 개체, 바람직하게는 대략 100명의 개체로 구성되는 첫 번째 군을 표현형에 따라 참여시킨다. 유사한 인원수의 대조 개체를 이런 연구에 참여시킨다.

<u>연관 분석</u>

본원 발명에는 또한, 유전형과 표현형간 연관을 검출하는 방법이 포함되는데, 상기 방법은 다음의 단계로 구성된다: a) 본원 발명의 유전자형판별 방법에 따라 형질 파지티브 개체군에서 적어도 1개의 CanIon-관련된 이중대립인자 마커의 빈도를 측정하고; b) 본원 발명의 유전자형판별 방법에 따라 대조 개체군에서 상기 CanIon-관련된 이중대립인자 마커의 빈도를 측정하고; c) 유전형과 표현형간 통계학적으로 유의한 상관관계가 존재하는 지를 측정한다. 이에 더하여, 본원 발명의 유전형과 표현형간 연관을 검출하는 방법에는 제한없이 본원 발명에 개시된 단독의 또는 조합된 방법이 포함되는데, 여기서 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 대조 개체군은 형질 네거티브 대조군, 또는 무작위 개체군일 수 있다; 선택적으로, a)와 b)의 유전자형판별 단계는 상기 개체군 각각으로부터 유래된 집합된 생물학적 시료에서 실시할 수 있다; 선택적으로, a)와 b)의 유전자형판별 단계 는 상기 개체군에서 각 개체, 또는 이의 부차시료(subsample)로부터 유래된 생물학적 시료에서 개별적으로 실시한다.

후보 유전자를 보유하는 영역으로부터 유래된 이중대립인자 마커를 이용한 연관 연구를 실시하는데 일반적인 전략은 양 군에서 본원 발명의 이중대립인자 마커의 대립유전자 빈도를 측정하고 통계학적으로 비교하기 위하여 두 개체군(사례-대조 개체군)을 스캔한다.

형질과의 통계학적으로 유의한 연관이 분석된 한가지이상의 이중대립인자 마커에서 확인되면, 연관된 대립유전자가 형질을 직접적으로 유발하거나(즉, 연관된 대립유전자가 형질 유발 대립유전자이다), 또는 연관된 대립유전자가 형질 유발 대립유전자와 연쇄 불균형하다라고 추정할 수 있다. 후보 유전자 기능과 관련하여 연관된 대립유전자의 특징적인 특성은 연관된 대립유전자와 형질간 상관관계(인과 또는 연쇄 불균형)에 대한 더욱 심화된 통찰을 제공한다. 이런 증거가 후보 유전자내에서 연관된 대립유전자가 형질 유발 대립유전자가 아니고 실제 형질 유발 대립유전자와 연쇄 불균형 상태에 있음을 시사하면, 형질 유발 대립유전자는 연관된 마커의 주변을 서열분석하고 반복적인 방식으로 확인된 다형성으로 추가적인 연관 연구를 실시하여 발견할 수 있다.

일반적으로, 연관 연구는 2가지 연속 단계로 실시한다. 첫 번째 단계에서, 후보 유전자로부터 감소된 수량의 이중대립인자 마커의 빈도는 형질 파지티브와 대조 개체군에서 측정한다. 분석의 두 번째 단계에서, 일정한 형질을 주도하는 유전자 좌위의 위치는 유관한 영역으로부터 좀더 높은 밀도의 마커를 이용하여 추가로 정련한다. 하지만, CanIon에서처럼 연구중인 후보 유전자의 길이가 상대적으로 짧으면, 유의한 연관을 확립하는데 단일 단계로도 충분하다.

일배체형 분석

전술한 바와 같이, 질환 대립유전자를 보유하는 크로모좀이 돌연변이 또는 이동의 결과로 개체군에서 먼저 나타나면, 변이 대립유전자는 일단의 연결된 마커: 조상 일배체형을 보유하는 크로모좀에 존재한다. 이런 일배체형은 개체군을 통하여 추적할 수 있고, 정해진 형질과의 통계학적 연관을 분석할 수 있다. 일배체형 연구라고 하는 다중-포인트 연관 연구로 단일 포인트(대립유전자) 연관 연구를 보충하면 연관 연구의 통계학적 능력이 증가한다. 따라서, 일배체형 연관 연구로 조상 보인자 일배체형의 빈도와 유형을 정의할 수 있다. 일배체형 분석은 개별 마커가 포함된 분석의 통계학적 능력을 증가시킨다는 점에서 중요하다.

일배체형 빈도 분석의 첫 번째 단계에서, 본원 발명의 동정된 이중대립인자 마커의 다양한 조합에 기초한 가능 일배체형의 빈도를 측정한다. 이후, 일배체형 빈도는 형질 파지티브와 대조 개체의 별도의 개체군에 대하여 비교한다. 통계학적으로 유의한 결과를 얻기 위하여 이런 분석이 적용되는 형질 파지티브 개체의 인원수는 일반적으로, 30 내지 300, 바람직하게는 50 내지 150이다. 동일한 고려는 본 연구에 참여한 영향을 받지 않은 개체(또는 무작위 대조)의 인원수에도 적용된다. 상기 첫 번째 분석의 결과는 사례-대조 개체군에서 일배체형 빈도를 제공하는데, 각 평가된 일배체형 빈도에서 p-값과 교차비(odd ratio)를 계산한다. 통계학적으로 유의한 연관이 확인되면, 연구중인 형질로 영향을 받는 일정한 일배체형을 보유 하는 개체에 대한 비교위험도(relative risk)를 근사할 수 있다.

본원 발명의 다른 구체예에는 일배체형과 표현형간 연관을 검출하는 방법이 포함되는데, 상기 방법은 다음의 단계로 구성된다: a) 일배체형의 빈도를 측정하기 위한 본원 발명의 방법에 따라 형질 파지티브 개체군에서 적어도 1개의 일배체형의 빈도를 측정하고; b) 일배체형의 빈도를 측정하기 위한 본원 발명의 방법에 따라 대조 개체군에서 상기 일배체형의 빈도를 측정하고; c) 일배체형과 표현형간 통계학적으로 유의한 연관이 존재하는 지를 측정한다. 이에 더하여, 본원 발명의 일배체형과 표현형간 연관을 검출하는 방법에는 제한없이 본원 발명에 개시된 단독의 또는 조합된 방법이 포함되는데, 여기서 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A1 내지 A17과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 CanIon-관련된 이중대립인자 마커는 A12와 A16과 이의 보체, 또는 선택적으로 이와 연쇄 불균형인 이중대립인자 마커에서 선택된다; 선택적으로, 상기 대조 개체군은 형질 네거티브 대조군, 또는 무작위 개체군일 수 있다; 선택적으로, 상기 대조 개체군은 형질 네거티브 개체군, 또는 무작위 개체군이다. 선택적으로, 상기 방법에는 c) 단계에 앞서 형질 파지티브와 대조 개체군에서 표현형을 확인하는 단계가 추가로 포함된다.

상호작용 분석

본원 발명의 이중대립인자 마커는 다인자 상호작용에 기인하는 검출가능한 형질과 연관된 이중대립인자 마커의 패턴을 확인하는 데에도 사용될 수 있다. 동일 연쇄군에 속하지 않는 좌위에서 대립유전자간 유전자 상호작용의 분석은 본원에서 밝힌 기술을 이용한 개별적인 유전자형판별을 필요로 한다. 적절한 수준의 통계학적 유의성을 갖는 선별된 일단의 이중대립인자 마커간 대립유전자 상호작용의 분석은 일배체형 분석으로 간주할 수 있다. 상호작용 분석은 첫 번째 좌위에 대한 일정한 일배체형과 관하여 사례-대조 개체군을 계층화시키고 각 부차개체군(subpopulation)에서 두 번째 좌위로 일배체형 분석을 실시하는 단계로 구성된다.

연관 연구에 이용된 통계학적 방법은 하기에 상세하게 기술한다.

연관의 존재에서 연쇄의 검사

본원 발명의 이중대립인자 마커는 TDT(전달/불균형 검사)에도 사용될 수 있다. 연쇄와 연관에 대한 TDT 검사는 개체군 계층화(stratification)의 영향을 받지 않는다. TDT는 영향을 받은 개체와 이들의 부모에 대한 데이터 또는 부모 대신에 영향을 받지 않은 근친으로부터 얻은 데이터를 필요로 한다(Spielmann S. et al., 1993; Schaid D.J. et al., 1996, Spielmann S. and Ewens W.J., 1998). 이런 통합된 검사는 개별적인 분석에 의한 가 양성 오류를 전반적으로 감소시킨다.

통계학적 방법

일반적으로, 형질과 유전형이 통계학적으로 유의한 상관관계를 보이는 지를 검사하는 당분야에 공지된 임의의 방법을 이용할 수 있다.

1) 연쇄 분석에서 방법

연쇄 분석에 유용한 통계학적 방법과 컴퓨터 프로그램은 당업자에게 공지되어 있다(Terwilliger J.D. and Ott J., 1994; Ott J., 1991).

2) 개체군에서 일배체형 빈도를 측정하는 방법

전술한 바와 같이, 유전형을 스코어하는 경우에 이형접합체가 구별되지 않아 일배체형 빈도를 쉽게 추론할 수 없는 경우가 있다. 배우자 위상이 알려져 있지 않는 경우에, 일배체형 빈도는 다중좌위 유전형 데이터로부터 평가할 수 있다. 당업자에게 공지된 임의의 방법을 이용하여 일배체형 빈도를 평가할 수 있다(Lange K., 1997; Weir, B.S., 1996). 적절하게는, 최대-공산 일배체형 빈도는 기대치-최대화(expectation-maximization)(EM) 대수(Dempster et al., 1977; Exocoffier L. and Slatkin M., 1995). 이런 과정은 배우자 위상이 알려져 있지 않은 경우에 다중-좌위 유전형 데이터로부터 일배체형 빈도의 최대-공산치를 얻기 위한 반복 과정이다. 일반적으로, 일배체형 평가는 예로써 EM-HAPLO 프로그램(Hawley M. E. et al., 19994) 또는 Arlequin 프로그램(Schneider et al., 1997)을 이용한 EM 대수를 적용하여 실시한다. EM 대수는 일반화된 반복 최대 공산 측정 방법으로, 하기에 간단하게 기술한다.

본 섹션, '개체군에서 일배체형 빈도를 측정하는 방법'에서 표현형은 미지의 일배체형 위상을 갖는 다중-좌위 유전형을 의미한다. 유전형은 공지된 일배체형 위상을 갖는 다중-좌위 유전형을 의미한다.

K 마커 타입의 N 명의 무관한 개체의 시료를 가정한다. 관찰된 데이터는 F 상이한 표현형으로 분류할 수 있는 미지의 위상 K-좌위 표현형이다. 또한, H 가능 한 일배체형을 보유한다고 가정한다(K 이중대립인자 마커의 경우에, 가능 일배체형 $H=2^K$의 최대값을 보유한다).

C_j 가능 유전형을 갖는 표현형 j 에서

$$P_j = \sum_{i=1}^{c_j} P(genotype(i)) = \sum_{i=1}^{c_j} P(h_k, h_l). \qquad \text{방정식 1}$$

여기서, P_j 는 j th 표현형의 확률이고, $P(h_k, h_l)$는 일배체형 h_k 와 h_l 로 구성되는 i th 유전형의 확률이다. 무작위 대합[즉, 하디-와인버그 평형(Hardy-Weinberg Equilibrium)]하에, $P(h_k, h_l)$는 다음과 같이 표시된다:

$$\begin{aligned} &P(h_k, h_l) = P(h_k)^2 \text{ for } h_k = h_l, \\ &P(h_k, h_l) = 2P(h_k)P(h_l) \text{ for } h_k \neq h_l. \qquad \text{방정식 2} \end{aligned}$$

E-M 대수는 다음과 같은 단계로 구성된다: 먼저, 유전형 빈도는 일배체형 빈도의 일단의 초기값으로부터 추정한다. 이들 일배체형 빈도는 $P_1^{(0)}, P_2^{(0)}, P_3^{(0)}, \ldots, P_H^{(0)}$ 라고 한다. 일배체형 빈도에 대한 초기값은 무작위 번호 발생기 또는 당분야에 공지된 다른 방식으로 얻을 수 있다. 최대화 단계라고 하는 두 번째 단계는 유전형 빈도에 대한 추정값을 이용하여 일배체형 빈도를 재-계산한다. 첫 번째 반복 일배체형 빈도 추정값은 $P_1^{(1)}, P_2^{(1)}, P_3^{(1)}, \ldots, P_H^{(1)}$ 라고 한다. 전반적으 로, s th 반복에서 기대 단계는 이전 반복의 일배체형 빈도에 기초하여 각 표현형을 별개의 가능 유전형으로 정하는 확률을 계산한다:

$$P(h_k, h_l)^{(s)} = \frac{n_j}{N}\left[\frac{P_j(h_k, h_l)^{(s)}}{P_j}\right], \qquad \text{방정식 3}$$

여기서, n_j 는 j th 표현형을 갖는 개체의 인원수이고, $P(h_k, h_l)^{(s)}$ 는 표현형 j 에서 유전형 h_k, h_i 의 확률이다. 유전자-계수(counting) 방법(Smith, Ann. Hum. Genet., 21:254-276, 1957)에 대응하는 최대화 단계에서, 일배체형 빈도는 유전형 추정값에 기초하여 재-평가한다:

$$P_t^{(s+1)} = \frac{1}{2}\sum_{j=1}^{F}\sum_{i=1}^{c_j} \delta_{it} P_j(h_k, h_l)^{(s)}. \qquad \text{방정식 4}$$

여기서, δ_{it} 는 일배체형 t 가 i th 유전형에 존재하는 발생 회수를 계산하는 지시 요소이다; it은 0, 1, 2이다.

E-M 반복은 다음의 기준에 도달하면 중단한다. 최대 공산 측정(MLE) 이론을 이용하여, 표현형 j 는 다항 분포(multinomial distribution)한다고 가정한다. 각 반복 s 에서, 우도 함수(likelihood function) L을 계산할 수 있다. 2개의 연속 반복간 로그-우도의 차이가 작은 수치, 바람직하게는 10^{-7} 이면 수렴이 달성된다.

3) 마커간 연쇄 불균형을 계산하는 방법

다양한 방법을 이용하여 임의 2개의 유전자 위치간 연쇄 불균형을 계산할 수 있다. 실제로, 연쇄 불균형은 개체군으로부터 얻은 일배체형 데이터에 통계학적 연관 검사를 적용하여 측정한다.

마커 M_i에서 대립유전자(a_i/b_i) 및 마커 M_j에서 대립유전자(a_j/b_j)를 보유하는 본원 발명의 이중대립인자 마커(M_i, M_j)중 적어도 한가지를 포함하는 임의의 이중대립인자 마커쌍 사이의 연쇄 불균형은 Piazza 공식에 따라 모든 대립유전자 조합(a_i,a_j; a_i,b_j; b_i,a_j; b_i,b_j)에서 계산할 수 있다:

$\Delta a_i a_j = \sqrt{\theta 4} - \sqrt{(\theta 4+\theta 3)(\theta 4+\theta 2)}$, 여기서

$\theta 4 = - - =$ M_i에서 대립유전자 a_i를 보유하지 않고 M_j에서 대립유전자 a_j를 보유하지 않는 유전형의 빈도;

$\theta 3 = - + =$ M_i에서 대립유전자 a_i를 보유하지 않고 M_j에서 대립유전자 a_j를 보유하는 유전형의 빈도;

$\theta 2 = + - =$ M_i에서 대립유전자 a_i를 보유하고 M_j에서 대립유전자 a_j를 보유하지 않는 유전형의 빈도.

Weir B. S., 1996에서 밝힌 바와 같이, 델타(혼성 유전형 불균형 계수)에 대한 최대 공산치(MLE)에 따라 모든 대립유전자 조합(a_i,a_j; a_i,b_j; b_i,a_j; b_i,b_j)에서 여러 쌍의 이중대립인자 마커(M_i, M_j)간 연쇄 불균형(LD)을 계산할 수도 있다. 혼성 연쇄 불균형에 대한 MLE은 다음과 같다:

$Da_i a_j = (2n_1 + n_2 + n_3 + n_4/2)/N - 2(pr(a_i).\ pr(a_j))$

여기서, n_1은 Σ표현형(a_i/a_i, a_j/a_j), n_2는 Σ표현형(a_i/a_i, a_j/b_j), n_3은 Σ표현형(a_i/b_i, a_j/a_j), n_4는 Σ표현형(a_i/b_i, a_j/b_j), N은 시료에서 개체의 인원수이다.

이 공식으로, 단지 유전형 데이터만 가용하고 일배체형 데이터가 가용하지 않은 경우에 대립유전자간 연쇄 불균형을 평가할 수 있다.

마커간 연쇄 불균형을 계산하는 다른 수단은 다음과 같다. 한 쌍의 이중대립인자 마커, $Mi(a_i/b_i)$ 및 $M_j(a_j/b_j)$에 대하여 Hardy-Weinberg 평형을 적합시켜, 전술한 방식에 따라 일정한 개체군에서 4가지 가능 일배체형 빈도를 평가할 수 있다.

배우자 불균형의 추정치는 다음과 같다:

$Da_i a_j = pr(일배체형(a_i, a_j)) - pr(a_i).pr(a_j)$.

여기서, $pr(a_i)$은 대립유전자 a_i의 확률이고 $pr(a_j)$은 대립유전자 a_j의 확률이며, $pr(a_i).pr(a_j)$은 상기 방정식 3에서와 같이 추정한다.

한 쌍의 이중대립인자 마커에서는 Mi와 M_j간 연관을 설명하는 한가지 불균형 측도(measure of disequilibrium)만 필요하다.

이후, 이의 표준화된 값은 다음과 같이 계산하다:

$$D'_{aiaj} = D_{aiaj} / \max(-pr(a_i).\ pr(a_j),\ -pr(b_i).\ pr(b_j)) \quad \text{with } D_{aiaj}<0$$
$$D'_{aiaj} = D_{aiaj} / \max(pr(b_i).\ pr(a_j),\ pr(a_i).\ pr(b_j)) \quad \text{with } D_{aiaj}>0$$

다른 연쇄 불균형 계산 방법을 이용할 수도 있다.

적절한 예상 이형접합률(heterozygosity rate)을 갖는 일단의 이중대립인자 마커간 연쇄 불균형은 50 내지 1000명, 바람직하게는 75 내지 200명, 좀더 바람직하게는 대략 100명의 무관한 개체를 유전자형판별하여 측정할 수 있다.

4) 연관 검사

이중대립인자 마커에서 대립유전자 또는 이런 대립유전자로 구성되는 일배체형에서 표현형과 유전자간 상관관계의 통계학적 유의성을 측정하는 방법은 당분야에 공지되고 통계학적 유의성의 용인된 경계값이 요구되는 임의의 통계학적 검증으로 달성할 수 있다. 특정 방법 및 유의 경계값의 이용은 당업자에게 공지된 것이다.

연관 검사는 사례와 대조 개체군에서 이중대립인자 마커 대립유전자의 빈도를 측정하고, 통계학적 검증으로 이들 빈도를 비교하여 이들이 형질 및 연구중인 이중대립인자 마커 대립유전자간 상관관계를 암시하는 빈도의 통계학적으로 유의한 차이인 지를 확인하는 단계로 구성된다. 유사하게, 일배체형 분석은 사례와 대조 개체군에서 일정한 일단의

이중대립인자 마커에 대한 모든 가능 일배체형의 빈도를 측정하고, 통계학적 검증으로 이들 빈도를 비교하여 이들이 일배체형 및 연구중인 표현형(형질)간 통계학적으로 유의한 상관관계가 존재하는 지를 확인하는 단계로 구성된다. 유전형과 표현형간 통계학적으로 유의한 상관간계를 검증하는데 유용한 임의의 통계학적 도구를 이용할 수 있다. 적절하게는, 이용된 통계학적 검증은 1 자유도(degree of freedom)를 갖는 카이제곱검증(chi-square test)이다. P-값이 계산된다(P-값은 관찰값 이상의 통계치가 우연히 발생하는 확률이다).

<u>통계학적 유의성</u>

추가적인 진단 검사를 위한 파지티브 기초 또는 조기 예방 요법을 위한 예비 출발점으로서 진단 목적을 위한 유의성의 적절한 구체예에서, 이중대립인자 마커 연관과 관련된 p 값은 단일 이중대립인자 마커 분석에서 바람직하게는 1 x 10^{-2} 이하, 좀더 바람직하게는 1 x10^{-4} 이하이고, 2개이상의 마커를 포함하는 일배체형 분석에서 1 x10^{-3} 이하, 바람직하게는 1 x10^{-6} 이하이다. 이들 값은 단일 또는 다중 마커 조합을 포함하는 임의의 연관 연구에 적용가능하다.

당업자는 본원 발명의 이중대립인자 마커로 연관 연구를 실시하기 위한 출발점으로 전술한 범위의 값을 이용할 수 있다. 이렇게 함으로써 본원 발명의 이중대립인자 마커 및 형질간 유의한 상관관계가 확인될 수 있는데, 이는 진단 및 약물 스크리닝 목적에 이용될 수 있다.

<u>표현형 교환(permutation)</u>

전술한 1단계 일배체형 분석의 통계학적 유의성을 검정하기 위하여, 추가적인 분석을 실시하는 것이 바람직한데, 여기서 사례-대조 개체로부터 얻은 유전자형판별 데이터는 수집되고 형질 표현형에 대하여 무작위화된다. 각 개체 유전자형판별 데이터는 사례-대조 개체군과 동일한 인원수의 두 군에 무작위로 할당하여, 1단계에서 얻은 데이터를 편집한다. 2단계 일배체형 분석은 가급적, 최대 비교위험도 계수를 보인 1단계 분석의 일배체형에 포함된 마커에 대하여 이들 인위적 군에서 실시한다. 적절하게는, 이 실험은 적어도 100 내지 10000회 반복한다. 이런 반복 은 검사된 일배체형을 우연히 얻을 확률의 측정을 가능하게 한다.

<u>통계학적 연관의 평가</u>

가 양성 문제를 해결하기 위하여, 무작위 게놈 영역에서 동일한 사례-대조 개체군으로 유사한 분석을 실시할 수 있다. 무작위 영역과 후보 영역에서 결과는 미국 분할 출원 'Methods, software and apparati for identifying genomic regions harboring a gene associated with a dectable trait'(60/107,986; 1998.11.10일 출원)에서 기술한 바와 같이 비교한다.

<u>5) 위험 인자의 평가</u>

위험 인자(유전독성학에서 위험 인자는 마커 좌위에서 특정 대립유전자 또는 일배체형의 존부이다)와 질환간 연관은 교차비(odds ratio) 및 비교위험도(RR)로 측정한다. P(R+)가 R과 질환이 발병할 확률이고 P(R-)가 위험 인자가 없는 개체에서 확률이면, 비교위험도는 이들 2가지 확률의 비율이다:

RR= P(R+)/P(R-)

사례-대조 연구에서, 비교위험도의 직접적인 측도(measure)는 샘플링 디자인으로 인하여 얻을 수 없다. 하지만, 교차비는 낮은-발병률 질환에 대한 비교위험도의 효과적인 근사를 가능하게 하는데, 다음과 같이 계산할 수 있다:

$$OR=\left[\frac{F^{+}}{1-F^{+}}\right]\Big/\left[\frac{F^{-}}{(1-F^{-})}\right]$$

$$OR= (F^{+}/(1-F^{+}))/(F^{-}/(1-F^{-}))$$

F^+ 는 사례에서 위험 인자에 대한 노출 빈도이고, F^- 는 대조에서 위험 인자에 대한 노출 빈도이다. F^+ 와 F^- 는 본 연구의 대립유전자 또는 일배체형 빈도를 이용하여 계산하고, 기초적인 유전 모델(우성, 열성, 부가...)에 추가적으로 의존한다.

개체군에서 일정한 위험 인자로 인한 형질을 보이는 개체의 비율을 의미하는 기여위험도(AR)를 평가할 수 있다. 이런 측정은 질환 병리 및 위험 인자의 공중 보건 영향의 측면에서 특정 인자의 역할을 정량하는데 중요하다. 이런 측정의 공중 보건 상관(relevance)은 노출이 없는 경우에 예방될 수 있는 개체군의 질환 사례의 비율을 평가하는데 있다. AR

은 다음과 같이 측정한다:

$AR= P_E (RR-1)/(P_E (PR-1)+1)$

AR은 이중대립인자 마커 대립유전자 또는 이중대립인자 마커 일배체형에 대한 기여위험도이다. P_E 는 전체 개체군에서 대립유전자 또는 일배체형에 대한 노출 빈도이다; RR은 연구중인 형질이 전체 개체군에서 상대적으로 낮은 발병률을 갖는 경우에 교차비에 근사하는 비교위험도이다.

본원 발명의 이중대립인자 마커로 연쇄 불균형인 이중대립인자 마커의 동정

첫 번째 이중대립인자 마커가 관심있는 게놈 영역에서 동정되면, 당업자는 본원 발명의 기술적 사상을 이용하여 첫 번째 마커와 연쇄 불균형인 다른 이중대립인자 마커를 용이하게 동정할 수 있다. 전술한 바와 같이, 첫 번째 마커와 연쇄 불균형인 임의의 마커는 형질과 연관하게 된다. 따라서, 연관이 일정한 이중대립인자 마커와 형질간에 입증되면, 상기 특정 영역에서 이중대립인자 마커의 밀도를 증가시키기 위하여 이런 형질과 연관된 추가적인 이중대립인자 마커의 발견이 많은 관심을 받는다. 인과 유전자 또는 돌연변이가 형질과 최대 상관관계를 보이는 마커 일단의 마커의 주변에서 발견된다.

일정한 마커와 연쇄 불균형인 추가적인 마커의 동정은 다음의 단계로 구성된다: (a) 복수의 개체로부터 첫 번째 이중대립인자 마커를 포함하는 게놈 단편을 증폭하고; (b) 첫 번째 이중대립인자 마커를 보유하는 게놈 영역에서 두 번째 이중대립인자 마커를 동정하고; (c) 첫 번째 이중대립인자 마커와 두 번째 이중대립인자 마커간 연쇄 불균형 분석을 실시하고; (d) 첫 번째 마커와 연쇄 불균형인 두 번째 이중대립인자 마커를 선별한다. (b)와 (c) 단계로 구성되는 부차 조합 역시 정관한다.

이중대립인자 마커를 동정하고 연쇄 불균형 분석을 실시하는 방법은 본원에 기술되어 있어 당업자가 과도한 실험없이 실시할 수 있다. 본원 발명은 또한, 이중대립인자 마커 A1 내지 A18과 연쇄 불균형이고 일정한 형질과의 개별 연관의 측면에서 유사한 특성을 나타낼 것으로 예상되는 이중대립인자 마커에 관한다.

기능적 돌연변이의 동정

검출가능한 표현형 또는 형질을 주도하는 CanIon 유전자에서 돌연변이는 형질 파지티브와 대조 개체로부터 CanIon 유전자의 서열을 비교하여 동정할 수 있다. 파지티브 연관이 본원 발명의 이중대립인자 마커로 확인되면, 동정된 좌위는 돌연변이를 스캔할 수 있다. 적절한 구체예에서, CanIon 유전자의 엑손과 절단접합 부위와 같은 기능 영역, 프로모터 및 다른 조절 영역은 돌연변이를 스캔한다. 적절한 구체예에서, CanIon 유전자의 서열은 형질 파지티브와 대조 개체에서 비교한다. 적절하게는, 형질 파지티브 개체는 형질과 연관된 것으로 보이는 일배체형을 보유하고, 형질 네거티브 개체는 형질과 연딴된 일배체형 또는 대립유전자를 보유하지 않는다. 검출가능한 형질 또는 표현형에 변형된 CanIon 기능의 다양한 발현(manifestation)이 포함된다.

돌연변이 검출 과정은 이중대립인자 마커 동정에 사용되는 것과 유사하다. 이런 돌연변이를 검출하는데 이용되는 방법은 일반적으로, 다음의 단계로 구성된다:

- 형질 파지티브 환자와 형질 네거티브 대조의 DNA 시료로부터 형질과 연관된 이중대립인자 마커 또는 이중대립인자 마커군을 포함하는 CanIon 유전자 영역의 증폭;

- 증폭된 영역의 염기서열분석;

- 형질 파지티브와 대조 개체로부터 DNA 서열의 비교;

- 형질-파지티브 환자에 특이적인 돌연변이의 확인

한 구체예에서, 이중대립인자 마커는 A1 내지 A18 및 이의 보체에서 선택된다. 이후, 후보 다형성은 본원에서 밝힌 유전자형판별 과정, 바람직하게는 개별 검사 포맷에서 마이크로염기서열분석 기술로 사례와 대조의 대규모 개체군을 스크린하여 검증한다. 다형성은 사례와 대조에서 예상된 연관 결과와 일치하는 빈도로 존재하면 후보 돌연변이로 간주된다. 다형성은 검출가능한 표현형과 통계학적으로 유의한 상관관계를 보이면 후보 '형질-유발' 돌연변이로 간주된다.

유전자 진단의 방법에서 본원 발명의 이중대립인자 마커

본원 발명의 CanIon 핵산 서열과 이중대립인자는 특정 유전형의 결과로 검출가능한 형질을 발현하는 개체, 또는 유전형으로 인해 이후의 삶에서 검출가능한 형질이 발생할 위험이 높은 개체를 동정할 수 있는 진단 검사를 개발하는 데에도 사용될 수 있다. 이런 진단은 정신분열증, 양극성 장애, 다른 CNS 질환(예, 간질), 통증 질환, 심혈관 질환(예, 심장 질환, 고혈압, 부정맥) 및 다른 질환과 이상을 비롯한 다양한 질환이나 이상의 시기결정(staging), 모니터링, 예측 및/또는 예방이나 치료 요법에 사용될 수 있다.

본원 발명의 진단 기술에는 일배체형판별을 위한 개별 크로모좀의 분석이 가능한 방법, 예를 들면 가족 연구, 단순 정자 DNA 분석 또는 체세포 하이브리드를 비롯하여, 검체가 검출가능한 형질이 발현할 위험의 증가와 연관된 이중대립인자 마커 패턴을 보유하는지 또는 개체가 특정 돌연변이의 결과로 검출가능한 형질로 고생하는지를 확인하기 위한 다양한 방법을 이용할 수 있다.

본원 발명은 개체가 발병할 위험이 있는지 또는 CanIon 유전자에서 돌연변이나 다형성으로 인해 질환에 걸릴 지를 확인하는 진단 방법을 제공한다. 본원 발명은 또한, 개체가 정신분열증과 양극성 장애, 또는 당분야에 공지되거나 본원에서 밝힌 칼슘-채널 관련된 이상에 감수성을 보유하는지를 확인하는 방법을 제공한다.

이들 방법에는 개체로부터 핵산 시료를 수득하고, 상기 핵산 시료가 한 개이상의 대립유전자 또는 한 개이상의 이중대립인자 마커 일배체형을 보유하는 지를 확인 단계가 포함되는데, 이들의 존재는 형질의 발현 위험을 시사하거나 또는 개체가 특정 CanIon 다형성이나 돌연변이를 보유하는 결과로서 형질을 발현한다는 것을 시사한다(형질-유발 대립유전자).

적절하게는, 이런 진단 방법에서 핵산 시료는 개체로부터 수득되고, 상기 시료는 '이중대립인자 마커에 대하여 DNA 시료를 유전자형판별하는 방법'에서 전술한 방법으로 유전자형판별한다. 진단은 단일 이중대립인자 마커 또는 이중대립인자 마커군에 기초한다.

이들 방법에서, 핵산 시료는 검체로부터 수득하고, 이중대립인자 마커 A1 내지 A18중 적어도 한가지의 이중대립인자 마커 패턴을 확인한다.

한 구체예에서, PCR 증폭은 검출가능한 표현형과 연관된 다형성이 동정된 영역을 증폭하기 위하여 핵산 시료에서 실시한다. 증폭 산물은 염기서열분석하여 개체가 검출가능한 표현형과 연관된 하나 또는 복수의 CanIon 다형성을 보유하는지를 확인한다. 증폭 산물을 만들기 위하여 사용된 프라이머에는 표 1에 기재된 프라이머가 포함된다. 대안으로, 핵산 시료는 전술한 마이크로염기서열분석 반응을 실시하여 개체가 CanIon 유전자에서 돌연변이나 다형성으로 기인한 검출가능한 표현형과 연관된 한 개이상의 CanIon 다형성을 보유하는지를 확인한다. 마이크로염기서열분석 반응에 사용된 프라이머는 표 4에 기재된 프라이머가 포함된다. 다른 구체예에서, 핵산 시료는 검출가능한 표현형과 연관된 한 개이상의 CanIon 대립유전자에 특이적으로 혼성화되는 하나 또는 복수의 대립유전자 특이적 올리고뉴클레오티드 프로브와 접촉시킨다. 혼성화 분석에 사용되는 프로브에는 표 3에 기재된 프라이머가 포함된다. 다른 구체예에서, 핵산 시료는 증폭 반응에서 대립유전자 특이적 올리고뉴클레오티드와 함께 사용되는 경우에 증폭 산물을 생산할 수 있는 두 번 째 CanIon 올리고뉴클레오티드와 접촉시킨다. 증폭 반응에서 증폭 산물의 존재는 개체가 검출가능한 표현형과 연관된 한 개이상의 CanIon 대립유전자를 보유한다는 것을 암시한다.

적절한 구체예에서, A1 내지 A18 및 이의 보체에서 선택되는 한 개이상의 이중대립인자 마커에 존재하는 뉴클레오티드의 실체를 확인하는데, 검출가능한 형질은 정신분열증과 양극성 장애이다. 진단 키트는 본원 발명의 폴리뉴클레오티드를 포함한다.

이들 진단 방법은 특정 환경에서 예방 치료를 시작하거나 또는 유의한 일배체형을 보유하는 개체가 경미한 증상과 같은 경고 신호를 인지할 수 있도록 하는데 이용될 수 있다는 점에서 매우 유용하다.

약물에 대한 반응 또는 약물에 대한 부작용을 분석과 예측하는 진단은 특정 약물로 개체를 치료할 것인가를 결정하는데 이용될 수 있다. 가령, 진단에서 개체가 특정 약물 치료에 긍정적으로 반응할 가능성이 있는 것으로 나타나면, 이 약물은 개체에 투여될 수 있다. 반대로, 진단에서 개체가 특정 약물 치료에 부정적으로 반응할 가능성이 있는 것으로 나타나면, 다른 치료 과정을 처방할 수 있다. 부정적인 반응은 유익한 반응의 부재 또는 독성 부작용의 존재로 정의할 수 있다.

본원 발명의 마커는 임상적 약물 실험에도 이용될 수 있다. 정신분열증, 양극성 장애 또는 다른 칼슘-채널 관련된 이상을 억제하는 작용제에 반응, 또는 정신분열증, 양극성 장애 또는 다른 칼슘-채널 관련된 이상을 억제하는 작용제에 부작용을 암시하는 한 개이상의 마커는 전술한 방법으로 동정할 수 있다. 이후, 이런 작용제의 임상적 실험에 잠재적 참여자를 스크린하여, 약물에 긍정적으로 반응할 가능성이 가장 높은 개체를 확인하고 부작용을 보일 가능성이 높은 개체는 배제시킨다. 이런 방식으로, 연구중에 긍정적으로 반응할 가능성이 낮은 개체의 참여로 인한 측정의 저하없이 또는 원치않는 안전 문제의 위험없이, 약물에 긍정적으로 반응하는 개체에서 치료의 유효성을 측정할 수 있다.

특히 적절한 구체예에서, 이런 진단으로 분석되는 형질은 정신분열증 또는 양극성 장애이다. 하지만, 본원 발명에는 관련된 질환, 특히 관련된 CNS 질환을 예방, 진단, 관리, 치료하는 방법에서 본원 발명의 이중대립인자 마커를 이용한 예방, 진단, 예측, 치료 방법이 포함된다. 예로써, 관련된 질환에는 정신 질환, 기분 장애, 자폐증, 약물 의존과 알코올 중독, 통증 질환, 간질, 정신 지체 및 인식 장애, 정신 장애, 불안 장애, 섭식 장애, 충동 조절 장애, 인격 장애를 비롯한 다른 정신 질환이 포함된다. 질환은 Diagnosis Statistical Manual of Mental Disorder fourth edition(DSM-IV) 분류에 따라 정의할 수 있다. 다른 질환에는 심혈관 질환, 예를 들면 협심증, 고혈압 또는 부정맥이 포함된다.

재조합 벡터

'벡터'는 환형 또는 선형 DNA 또는 RNA 분자를 의미하는데, 이는 이중-가닥 또는 단일 가닥이고 세포 숙주 또는 단세포나 다세포 숙주 미생물로 전달되는 적어도 한 개의 폴리뉴클레오티드를 포함한다.

본원 발명에는 CanIon 게놈 서열로부터 유래된 조절 폴리뉴클레오티드 및/또는 CanIon 게놈 서열이나 cDNA 서열로부터 유래된 코딩 폴리뉴클레오티드를 포함하 는 재조합 벡터 군이 포함된다.

일반적으로, 본원 발명의 재조합 벡터는 조절 서열, 코딩 서열, 폴리뉴클레오티드 구조체, 전술한 CanIon 프라이머나 프로브를 비롯한 본원 발명의 폴리뉴클레오티드를 포함한다. 좀더 구체적으로, 본원 발명의 재조합 벡터는 'CanIon 유전자의 게놈 서열' 섹션, 'CanIon cDNA 서열' 섹션, '코딩 영역' 섹션, '폴리뉴클레오티드 구조체' 섹션, '올리고뉴클레오티드 프로브와 프라이머' 섹션에서 밝힌 임의의 폴리뉴클레오티드를 포함할 수 있다.

적절한 첫 번째 구체예에서, 본원 발명의 재조합 벡터는 적합한 숙주 세포에서 SEQ ID No 1 내지 3 또는 6의 CanIon 게놈 서열이나 CanIon cDNA, 예를 들면 SEQ ID No 4의 cDNA로부터 유래된 삽입된 폴리뉴클레오티드를 증폭하는데 사용되고, 상기 폴리뉴클레오티드는 재조합 벡터가 복제할 때마다 증폭된다.

적절한 두 번째 구체예에서, 본원 발명에 따른 재조합 벡터는 본원 발명의 조절 폴리뉴클레오티드나 코딩 핵산, 또는 둘 모두를 포함하는 발현 벡터로 구성된다. 특정 구체예에서, 발현 벡터는 CanIon 폴리펩티드를 발현하는데 사용되는데, 이후 상기 폴리펩티드는 정제하고 리간드 스크리닝 분석에 또는 CanIon 단백질에 대한 항체를 조성하기 위한 면역원으로 사용할 수 있다. 다른 구체예에서, 발현 벡터는 유전자도입 동물의 생성 및 유전자 요법에 사용된다. 발현은 적절한 신호가 벡터에 제공되어야 가능한데, 상기 신호에는 다양한 조절 요소, 예를 들면 숙주 세포의 유전자의 발현을 유도하는 바이러스와 포유동물로부터 유래된 인헨서/프로모터가 포함된다. 산물을 발현하는 영구적이고 안정적인 세포 클론을 확립하기 위 한 가장 우세한 약물 선별 마커는 본원 발명의 발현 벡터에 포함되는데, 그 이유는 이들이 약물 선별 마커의 발현을 폴리펩티드의 발현과 연계시키는 요소이기 때문이다.

좀더 구체적으로, 본원 발명은 CanIon 단백질, 바람직하게는 SEQ ID No 5 아미노산 서열의 CanIon 단백질, 이의 변이체 또는 단편을 인코드하는 핵산을 포함하는 발현 벡터에 관한다.

또한, 본원 발명은 CanIon 코딩 서열의 발현에 유용한 재조합 발현 벡터에 관하는데, 여기서 상기 벡터는 SEQ ID No 4 핵산을 포함한다.

CanIon-관련된 이중대립인자 마커를 보유하는 핵산을 포함하는 재조합 벡터 역시 본원 발명에 속한다. 적절한 구체예에서, 상기 이중대립인자 마커는 A1 내지 A18과 이의 보체, 또는 선택적으로 이와 연쇄 불균형의 이중대립인자 마커에서 선택된다.

본원 발명의 벡터에서 발견할 수 있는 요소중 일부는 하기 섹션에 좀더 상세하게 설명한다

또한, 본원 발명은 a) 표적 유전자의 내인성 크로모좀 DNA로 외인성(이질성) 폴리뉴클레오티드 삽입, b) 내인성 크로모좀 DNA 결실, 및/또는 c) 내인성 크로모좀 DNA의 외인성 폴리뉴클레오티드로의 치환된 척추동물 기원, 바람직하게는 포유동물 기원, 특히 사람 기원의 일차, 이차, 영속화된 상동성 재조합된 숙주 세포에 관한다. 폴리뉴클레오티드 서열의 삽입, 결실 또는 치환은 표적 유전자의 코딩 서열 및/또는 조절 영역, 예를 들면 표적 유전자와 동작가능하게 연결된 프로모터와 인헨서 서열에서 발생할 수 있다.

또한, 본원 발명은 시험관내 또는 생체내에서 상동성 재조합된 숙주 세포를 만드는 방법에 관하는데, 여기서 세포에서 정상적으로 발현되지 않는 표적 유전자의 발현이 변화된다. 적절하게는, 이런 변형은 정상 성장 조건 또는 표적 유전자에 의해 인코드되는 폴리펩티드를 생산하는데 적합한 조건하에 표적 유전자의 발현을 유도한다. 상기 방법은 다음과 같은 단계로 구성된다: (a) 시험관내 또는 생체내에서 세포를 폴리뉴클레오티드 구조체로 트랜스펙션시키고, 상기 폴리뉴클레오티드 구조체는 (i) 표적화 서열; (ii) 조절 서열 및/또는 코딩 서열; (iii) 필요한 경우 트랜스펙션된 세포를 생산하기 위한 비대합된 스플라이스 공여자 부위로 구성되고; (b) 시험관내 또는 생체내에서 트랜스펙션된 세포를 상동성 재조합에 적합한 조건하에 유지시킨다.

또한, 본원 발명은 시험관내 또는 생체내 세포에서 표적 유전자의 발현을 변화시키는 방법에 관하는데, 여기서 상기 유전자는 세포에서 정상적으로 발현되지 않고, 상기 방법은 다음의 단계로 구성된다: (a) 시험관내 또는 생체내에서 세포를 폴리뉴클레오티드 구조체로 트랜스펙션시키고, 상기 폴리뉴클레오티드 구조체는 (i) 표적화 서열; (ii) 조절 서열 및/또는 코딩 서열; (iii) 필요한 경우 트랜스펙션된 세포를 생산하기 위한 비대합된 스플라이스 공여자 부위로 구성되고; (b) 시험관내 또는 생체내에서 트랜스펙션된 세포를 유전자 발현에 적합한 조건하에 유지시킨다.

또한, 본원 발명은 시험관내 또는 생체내 세포에서 표적화된 내인성 유전자 의 발현을 변화시켜 본원 발명에 따른 폴리펩티드를 만드는 방법에 관하는데, 여기서 상기 유전자는 세포에서 정상적으로 발현되지 않고, 상기 방법은 다음의 단계로 구성된다: (a) 시험관내에서 세포를 폴리뉴클레오티드 구조체로 트랜스펙션시키고, 상기 폴리뉴클레오티드 구조체는 (i) 표적화 서열; (ii) 조절 서열 및/또는 코딩 서열; (iii) 필요한 경우 트랜스펙션된 세포를 생산하기 위한 비대합된 스플라이스 공여자 부위로 구성되고; (b) 시험관내 또는 생체내에서 트랜스펙션된 세포를 상동성 재조합에 적합한 조건하에 유지시켜 상동성 재조합된 세포를 만들고; c) 시험관내 또는 생체내에서 트랜스펙션된 세포를 유전자 발현에 적합한 조건하에 유지시켜 폴리펩티드를 생산한다.

또한, 본원 발명은 세포에서 표적화된 유전자의 발현을 변화시키는 폴리뉴클레오티드 구조체에 관하는데, 여기서 상기 유전자는 정상적으로 발현되지 않는다. 이는 폴리뉴클레오티드 구조체가 표적 세포의 크로모좀 DNA에 삽입되는 경우에 발생하는데, 여기서 폴리뉴클레오티드 구조체는 다음과 같이 구성된다: (i) 표적화 서열; (ii) 조절 서열 및/또는 코딩 서열; (iii) 필요한 경우, 비대합된 스플라이스-공여자 부위. 본원 발명에는 전술한 폴리뉴클레오티드 구조체도 포함되는데, 여기서 상기 구조체는 폴리펩티드를 인코드하는 폴리뉴클레오티드를 추가로 포함하고 크로모좀 DNA와 상동성 재조합이후 표적화된 내인성 유전자와 쌍을 이룬다.

조성물은 당분야에 공지된 방법과 기술로 생산할 수 있다(미국 특허 출원 6,054,288; 6,048,729; 6,048,724; 6,048,524; 5,994,127; 5,968,502; 5,965,125; 5,869,239; 5,817,789; 5,783,385; 5,733,761; 5,641,670; 5,580,734; PCT 출원 WO 96/29411, WO 94/12650; 과학 논문 Koller et al. Proc, 1994, Natl. Acad. Sci. USA 86:8932-8935 (1989)).

1. 본원 발명에 다른 발현 벡터의 전반적인 특징

본원 발명에 따른 재조합 벡터는 YAC(이스트 인공 크로모좀), BAC(박테리아 인공 크로모좀), 파지, 파게미드, 코스미드, 플라스미드, 또는 크로모좀, 비-크로모좀, 반-합성, 합성 DNA를 포함할 수 있는 선형 DNA 분자 등으로 구성된다.

이런 재조합 벡터는 다음과 같은 어셈블리를 포함하는 전사 단위를 포함할 수 있다:

(1) 유전자 발현에서 조절 역할을 하는 유전 요소, 예를 들면 프로모터 또는 인헨서. 인헨서는 프로모터에 작용하여 전사를 증가시키는 10 내지 300bp 정도의 DNA을 포함하는 cis-작용 요소이다.

(2) mRNA로 전사되고 최종적으로 폴리펩티드로 번역되는 구조 또는 코딩 서열, 상기 구조 또는 코딩 서열은 (1)에서 밝힌 조절 요소에 동작가능하게 연결된다.

(3) 적절한 전사 개시와 종결 서열. 이스트 또는 진핵 발현 시스템에 사용되는 구조 단위에는 숙주 세포에 의해 번역된 단백질의 세포외 배출을 가능하게 하는 리더 서열이 포함된다. 대안으로, 재조합 단백질이 리더 또는 전달 서열없이 발현되는 경우에는 N-말단 잔기가 포함될 수 있다. 이런 잔기는 최종 산물을 제공하기 위하여 발현된 재조합 단백질로부터 후속으로 절단될 수도 또는 절단되지 않을 수도 있다.

일반적으로, 재조합 발현 벡터에는 복제 기점, 숙주 세포의 형질전환을 가능하게 하는 선별 마커, 하류 구조 서열의 전사를 유도하기 위한 고도로 발현된 유전자로부터 유래된 프로모터가 포함된다. 이질성 구조 서열은 적절한 단계에서 번역 개시와 종결 서열 및 번역된 단백질의 표층공간(periplasmic space)이나 세포외 매체로의 배출을 유도할 수 있는 리더 서열로 조합된다. 벡터가 포유동물 숙주 세포에서 원하는 서열을 트랜스펙션시키고 발현하는 특정 구체예에서, 바람직한 벡터는 원하는 숙주에서 복제 기점, 적합한 프로모터, 인헨서, 임의의 필요한 리보좀 결합 부위, 폴리아데닐화 신호, 스플라이스 공여자와 수용자 부위, 전사 종결 서열, 5'-측부 비-전사 서열을 포함한다. SV40 바이러스 게놈으로부터 유래된 DNA 서열, 예를 들면 SV40 기점, 초기 프로모터, 인헨서, 스플라이스, 폴리아데닐화 신호는 비-전사된 유전 요소를 제공하는데 사용될 수 있다.

SEQ ID No 5의 CanIon 폴리펩티드, 이의 단편 또는 변이체의 생체내 발현은 숙주 미생물에서 고유 유전자의 발현 또는 생물학적 불활성 CanIon 단백질의 생산과 관련된 유전적 결함을 수정하는데 유용할 수 있다.

따라서, 본원 발명은 치료할 환자에 적절한 유전 물질을 도입하여 SEQ ID No 2의 CanIon 폴리펩티드, 이의 단편 또

는 변이체의 생체내 생산을 위해 설계된 재조합 발현 벡터에 관한다. 상기 유전 물질은 환자로부터 사전에 추출된 세포에 시험관내에서 도입할 수 있는데, 이후 변형된 세포는 환자의 적합한 조직에 직접 생체내 재도입할 수 있다.

2. 조절 요소

프로모터

본원 발명에 따른 발현 벡터에 사용하는데 적합한 프로모터 영역은 이질성 유전자가 발현되는 세포 숙주를 고려하여 선택한다. 관심있는 핵산 서열의 발현을 조절하는데 사용되는 프로모터는 표적 세포에서 핵산의 발현을 유도할 수만 있다면, 큰 의미를 갖지 않는다. 따라서, 사람 세포를 표적하는 경우에는 사람 세포에서 발현될 수 있는 프로모터, 예를 들면 사람이나 바이러스 프로모터에 인접하여 또는 상기 프로모터하에 핵산 코딩 영역을 위치시키는 것이 바람직하다.

적합한 프로모터는 발현을 조절할 핵산에 대하여 이질성이거나 또는 대안으로 발현되는 코딩 서열을 포함하는 고유 폴리뉴클레오티드에 대하여 내인성일 수 있다. 이에 더하여, 프로모터는 구조 프로모터/코딩 서열이 삽입되는 재조합 벡터 서열에 대하여 이질성이다.

프로모터 영역은 예로써 CAT(클로람페니콜 전이효소) 벡터, 바람직하게는 pKK232-8과 pCM7 벡터를 이용하여 임의의 원하는 유전자로부터 선택할 수 있다.

바람직한 박테리아 프로모터는 LacI, LacZ, T3이나 T7 박테리아파지 RNA 중합효소 프로모터, gpt, 람다 PR, PL, trp 프로모터(EP 0036776), 폴리헤드린 프로모터, 바큘로바이러스로부터 유래된 p10 단백질 프로모터(Kit Novagen)(Smith et al., 1983; O'Reilly et al., 1992), 람다 PR 프로모터 또는 trc 프로모터이다.

진핵 프로모터에는 극초기 CMV, HSV 티미딘 키나아제, 초기와 후기 SV40, 레트로바이러스로부터 유래된 LTR, 생쥐 메탈로티오네인-L이 포함된다. 간편한 벡터와 프로모터의 선택은 당업자에게 자명하다.

프로모터의 선택은 유전공학 분야의 당업자에게 자명하다. 가령, Sambrook et al.(1989) 또는 Fuller et al.(1996)에서 밝힌 과정을 참조한다.

다른 조절 요소

cDNA 삽입체를 이용하는 경우에는, 유전자 전사체의 적절한 폴리아데닐화를 달성하기 위하여 폴라아데닐화 신호를 포함시키는 것이 바람직하다. 폴리아데닐화 신호의 특성은 본원 발명의 성공적인 실행에 중요한 의미를 갖지 않고, 임의의 서열, 예를 들면 사람 성장 호르몬과 SV40 폴리아데닐화 신호를 사용할 수 있다. 종결자 역시 발현 카세트의 요소로 고려된다. 이들 요소는 메시지 수준을 강화시키고 카세트로부터 다른 서열로의 통독을 최소화시키는 역할을 한다.

3. 선별성 마커

이런 마커는 세포에 확인가능한 변화를 전달하여 발현 구조체를 포함하는 세포의 조기 동정을 가능하게 한다. 형질전환된 숙주 세포의 선별을 위한 선별가능한 마커 유전자는 바람직하게는, 진핵 세포 배양액에서 다이하이드로폴레이트 환원효소 또는 네오마이신 저항성; 맥주 효모균(*S. cerevisiae*)에서 TRP1; 대장균(*E. coli*)에서 테트라사이클린, 리팜프신 또는 앰피실린 저항성 또는 마이코박테리아(*mycobacteria*)에서 대한 레반 사카라제인데, 상기 후자 마커는 음성 선별 마커이다.

4. 선호되는 벡터

박테리아 벡터

박테리아에 사용되는데 유용한 발현 벡터는 선별가능한 마커 및 pBR322(ATCC 37017) 유전 요소를 보유하는 상업적으로 가용한 플라스미드로부터 유래된 박테리아 복제 기점을 포함할 수 있다. 이런 상업적 벡터에는 예로써 pKK223-2(Pharmacia, Uppsala, Sweden)과 GEM1(Promega Biotec, Madison, WI, USA)이 포함된다.

다수의 다른 적합한 벡터가 당분야에 공지되어 있는데, 예를 들면 다음과 같은 박테리아 벡터가 상업적으로 가용하다 : pQE70, pQE60, pQE-9(Qiagen), pbs, pD10, phagescript, psiX174 pbluescript SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A(Stratagene); ptrc99a pKK223-3, pKK233-3, pDR540, pRIT5(Pharmacia); pWLNEO, pSV2CAT, pOG44, pXT1, pSG(Stratagene); pSVK3 pBPV, pMSG, pSVL(Pharmacia); pQE-30(QIAexpress).

박테리아파지 벡터

P1 박테리아파지 벡터는 80 내지 100kb의 많은 삽입체를 포함할 수 있다.

P1 박테리아파지 벡터, 예를 들면 p158 또는 p158/neo8의 작제는 Stemberg(1992, 1994)에서 기술한다. CanIon 뉴클레오티드 서열을 포함하는 재조합 P1 클론은 40kb이상의 폴리뉴클레오티드를 삽입하는데 고려할 수 있다(Linton et al., 1993). 유전자도입 실험을 위한 P1 DNA를 만드는데 적합한 프로토콜은 McCormick et al.(1994)에서 기술한다. 간단히 말하면, P1 플라스미드를 보유하는 대장균(*E. coli*)(바람직하게는, NS3529)은 25㎍/㎖ 카나마이신을 함유하는 액체 배지에서 하룻밤동안 배양한다. P1 DNA는 Qiagen Plasmid Maxi kit(Qiagen, Chatsworth, CA USA)를 이용한 알칼리성 세포용해로 대장균(*E. coli*)으로부터 준비 한다. P1 DNA는 2개의 Qiagen-tip 500 칼럼에서 키트에 포함된 세척과 용출 완충액을 이용하여 박테리아 용해질로부터 정제한다. 이후, 페놀/클로로포름 추출을 실시하고, 후속으로 70% 에탄올로 DNA를 침전시킨다. TE(10 mM Tris-HCL, pH 7.4, 1mM EDTA)에 DNA를 용해시킨 이후, DNA의 농도는 분광광도법으로 평가한다.

유전자도입 동물, 전형적으로 유전자도입 생쥐에서 CanIon 뉴클레오티드 서열을 보유하는 P1 클론을 발현시키는 경우에는 예로써 P1 폴리링커(SfiI, NotI 또는 SalI)내 빈도가 낮은 절단 부위에서 P1 DNA를 절단하여 P1 DNA 단편으로부터 벡터 서열을 제거하는 것이 바람직하다. 이후, P1 삽입체는 YAC로부터 DNA의 분리에서 적용된 방법(Schedl et al. 1993a; Peterson et al., 1993)으로 펄스-필드 아가로즈 겔에서 벡터 서열로부터 정제한다. 필요한 경우에, 정제된 삽입체 DNA는 Millipore Ultrafree-MC Filter Unit(Millipore Bedford, MA USA-30,000 분자량 한계)에서 농축하고, 후속으로 마이크로투석 막(type VS, 0.025 ㎛, Millipore)에서 100mM NaCl, 30μM 스페르민, 70μM 스프레미딘을 함유하는 마이크로주사 완충액(10mM Tris-HCl, pH 7.4, 250㎛ EDTA)에 투석한다. 정제된 P1 DNA 삽입체의 보존도(intactness)는 1% 아가로즈(Sea Kem GTG: FMC Bio-products) 펄스-필드 겔에서 전기영동 및 브롬화에티듐 염색으로 평가한다.

바쿨로바이러스 벡터

SEQ ID No 5의 CanIon 폴리펩티드, 이의 단편 또는 변이체에 적합한 벡터는 곤충 세포와 곤충 세포주에서 증식할 수 있는 바쿨로바이러스 벡터이다. 특히 적합한 숙주 벡터 시스템은 스포도프테라 프루기페르다(*Spodoptera frugiperda*)로부 터 유래된 SF9 세포주(ATCC N° CRL 1711)를 트랜스펙션시키는데 사용되는 pVL1392/1393 바쿨로바이러스 전달 벡터(Pharmingen)이다.

바쿨로바이러스 발현 시스템에서 SEQ ID No 5의 CanIon 폴리펩티드, 이의 단편 또는 변이체에 적합한 다른 벡터에는 Chai et al.(1993), Vlasak et al.(1983), Lenhard et al.(1996)에서 밝힌 벡터가 포함된다.

바이러스 벡터

특정 구체예에서, 벡터는 아데노바이러스로부터 유래된다. 본원 발명에 따른 적절한 아데노바이러스 벡터는 Feldman와 Steg(1996) 또는 Ohno et al.(1994)에서 밝힌 벡터이다. 본원 발명의 특정 구체예에 따른 다른 적절한 재조합 아데노바이러스는 2형이나 5형 사람 아데노바이러스(Ad2, Ad 5) 또는 동물 기원의 아데노바이러스(French N° FR-93.05954)이다.

일반적으로, 레트로바이러스와 아데노-관련된 바이러스 벡터는 생체내에서 외인성 폴리뉴클레오티드를 사람을 비롯한 포유동물에 전달하는데 선택되는 재조합 유전자 전달 시스템이다. 이들 벡터는 유전자의 세포로의 효과적인 전달을 제공하고, 전달된 핵산은 숙주의 크로모좀 DNA에 안정적으로 통합된다.

본원 발명에 따른 레트로바이러스 시험관내 또는 생체내 유전자 전달체의 제조나 작제에 특히 적절한 레트로바이러스는 Mink-Cell 집중 유도 바이러스, 뮤린 육종 바이러스, 레티큘로엔도텔리오시스 바이러스, 라우스 육종 바이러스에서 선택되는 레트로바이러스이다. 특히 바람직한 뮤린 백혈병 바이러스에는 4070A와 1504A 바이러스, Abelson(ATCC No VR-999), Friend(ATCC No VR-245), Gross(ATCC No VR-590), Rauscher(ATCC No VR-998), Moloney 뮤린 백혈병 바이러스(ATCC No VR-190; PCT 출원 WO 94/24298)가 포함된다. 특히 바람직한 라우스 육종 바이러스에는 Bryan high titer(ATCC Nos VR-334, VR-657, VR-726, VR-659, VR-728)가 포함된다. 다른 적절한 레트로바이러스 벡터는 Roth et al.(1996), PCT 출원 WO 93/25234, PCT 출원 WO 94/06920 Roux et al., 1989, Julan et al., 1992, Neda et al., 1991에서 기술한다.

본원 발명에서 또 다른 바이러스 벡터 시스템은 아데노-관련된 바이러스(AAV)로 구성된다. 아데노-관련된 바이러스는 효과적인 복제 및 생명 사이클을 위한 보조 바이러스로 다른 바이러스, 예를 들면 아데노바이러스 또는 포진 바이러스를 필요로 하는 자연 발생 결함 바이러스이다(Muzyczka et al., 1992). 이는 DNA를 비-분열 세포에 통합시키고 높은 빈도의 안정적 통합을 보이는 극소수 바이러스중 하나이다(Flotte et al., 1992; Samulski et al., 1989; McLau ghlin et al., 1989). AAV의 다른 장점은 형질전환된 세포와 관련된 일차 세포(primary cell)를 형질도입하는 능력의 감소에 기인한다.

BAC 벡터

박테리아 인공 크로모좀(BAC) 클로닝 시스템(Shizuya t al., 1992)은 큰 게놈 DNA 단편(100-300kb)을 대장균(*E. coli*)에 안정적으로 유지시키기 위하여 개발되었다. 적절한 BAC 벡터는 Kim et al.(1996)에서 밝힌 pBeloBAC11 벡터를 포함한다. BAC 라이브러리는 벡터에서 Bam HI 또는 HindIII 부위로의 결찰을 허용하는 효소로 부분적으로 절단된 크기-선별된 게놈 DNA를 이용하여 상기 벡터로 준비한다. 이들 클로닝 부위의 측부는 RNA 전사 또는 PCR 방법으로 최종 프로브를 만드는데 사용할 수 있는 T7과 SP6 RNA 중합효소 전사 개시 부위이다. 대장균(*E. coli*)에서 BAC 라이브러리의 작제이후, BAC DNA는 숙주 세포로부터 초나선 환형으로 정제된다. 이들 환형 분자를 선형으로 전환시키기에 앞서, 크기 결정 및 BAC의 수용체 세포로의 도입이 진행된다. 클로닝 부위는 2개의 Not I 부위가 측면에 접하는데, 이들 부위는 Not I 절단에 의해 클론된 분절이 벡터로부터 절단되도록 한다. 대안으로, pBeloBAC11 벡터에 포함된 DNA 삽입체는 단일 cosN 부위에서 절단을 유발하는 상업적으로 가용한 효소 람다 테르미나제(terminase)로 BAC 벡터를 처리하여 선형화시킬 수 있지만, 이런 절단 방법은 삽입체 DNA와 BAC 서열을 모두 포함하는 전장 BAC 클론을 결과한다.

5. 재조합 벡터의 전달

본원 발명에 따른 폴리뉴클레오티드와 폴리뉴클레오티드 구조체의 발현을 달성하기 위하여, 이들 구조체는 세포로 전달되어야 한다. 이런 전달은 세포주를 형질전환시키는 실험실 과정에서와 같이 시험관내에서, 또는 특정 질환의 치료에서와 같이 생체내외에서 달성할 수 있다.

한가지 기전은 바이러스 감염인데, 여기서 발현 구조체는 감염성 바이러스 입자에 통합된다.

본원 발명에서는 폴리뉴클레오티드를 배양된 포유동물 세포에 전달하기 위한 몇몇 비-바이러스 방법을 고려하는데, 여기에는 칼슘 포스페이트 침전(Graham et al., 1973; Chen et al., 1987), DEAE-덱스트란(Gopal, 1985), 에렉트로포레이션(Tur-Kaspa et al., 1986; Potter et al., 1984), 직접 미세주사(Harland et al., 1985), DNA-적하된 리포좀(Nicolau et al., 1982; Fraley et al., 1979), 수용체-매개된 트랜스펙션(Wu and Wu, 1987; 1988) 등이 포함된다. 이들 기술중 일부는 생체내외 용도에 성공적으로 적용될 수 있다.

발현 폴리뉴클레오티드가 세포에 전달되면, 이는 수용자 세포의 게놈에 안정적으로 통합될 수 있다. 이런 통합은 상동성 재조합을 통하여 동원(cognate)위치와 배향으로 진행되거나(유전자 치환) 또는 무작위의 비-특이적인 위치로 통합될 수 있다(유전자 보충). 다른 구체예에서, 핵산은 별도의 에피좀 DNA 분절로 세포에서 안정적으로 유지될 수도 있다. 이런 핵산 분절 또는 '에피좀'은 숙주 세포 사이클과 독립적으로 또는 이런 사이클과 동조하여 복제와 유지를 가능하게 하는 서열을 인코드한다.

특정 구체예에서, 생체내에서 단백질 또는 펩티드를 척추동물의 세포내로 전달하는 방법은 생리적으로 수용가능한 담체 및 관심있는 폴리펩티드를 동작가능하게 코딩하는 나신 폴리뉴클레오티드를 함유하는 제형을 세포를 포함하는 조직의 간질성 공간에 도입하는 단계로 구성되고, 여기서 나신 폴리뉴클레오티드는 세포의 내부로 흡수되고 생리 효과를 보인다. 이는 특히 시험관내 전달에 적용할 수 있지만, 시험관내 전달에도 적용할 수 있다.

'나신' 폴리뉴클레오티드를 포함하는 시험관내와 생체내에서 사용되는 조성물은 PCT 출원 WO 90/11092(Vical Inc.), PCT 출원 WO 95/11307(Institut Pasteur, INSERM, Universite d'Ottawa), Tacson et al.(1996), Huygen et al.(1996)에서 기 술한다.

본원 발명의 다른 구체예에서, 본원 발명의 폴리뉴클레오티드 구조체를 비롯하여 본원 발명에 따른 나신 폴리뉴클레오티드의 세포로의 전달은 입자 폭탄(바이오리스틱)으로 실시하는데, 상기 입자는 고속으로 가속되는 DNA-피복된 마이크로발사체로서, 이들 DNA가 세포막을 통과하여 세포에 진입하도록 한다(Klein et al.(1987)).

적절한 구체예에서, 본원 발명의 폴리뉴클레오티드는 리포좀에 포집된다(Ghosh and Bacchawat, 1991; Wong et al., 1980; Nicolau et al., 1987).

특정 구체예에서, 본원 발명은 본원 발명에 따른 CanIon 단백질 또는 폴리펩티드의 생체내 생산을 위한 조성물을 제시한다. 상기 조성물은 생리학적으로 수용가능한 담체에 용해되고 조직에 도입되면 조직의 세포가 상기 단백질 또는 폴리펩티드를 발현하도록 하는 이런 폴리펩티드를 동작가능하게 코딩하는 나신 폴리뉴클레오티드로 구성된다.

원하는 숙주 미생물에 주사되는 벡터의 함량은 주사 부위에 따라 달라진다. 일반적으로, 동물 신체, 바람직하게는 포

유동물 신체, 예를 들면 생쥐 신체에 0.1 내지 100μg이 주사된다.

다른 구체예에서, 본원 발명에 따른 벡터는 숙주 세포, 바람직하게는 처리된 동물로부터 수거된 숙주 세포, 좀더 바람직하게는 체세포(예, 근세포)에 도입할 수 있다. 후속 단계에서, 세포는 원하는 CanIon 폴리펩티드 또는 이의 단편을 코딩하는 벡터로 형질전환되고, 상기 세포는 동물 신체에 재도입하여 상기 재조합 단백질 을 국소적으로 또는 전신적으로 체내에 전달한다.

<u>**세포 숙주**</u>

본원 발명의 다른 목적은 본원 발명의 폴리뉴클레오티드, 특히 CanIon 조절 폴리뉴클레오티드, 또는 SEQ ID No 1 내지 4의 CanIon 폴리펩티드, 이의 단편 혹은 변이체의 코딩 서열을 포함하는 폴리뉴클레오티드로 형질전환 또는 트랜스펙션되는 숙주 세포에 관한다. 본원 발명에는 전술한 재조합 벡터로 형질전환(원핵 세포) 또는 트랜스펙션(진핵 세포)된 숙주 세포 역시 포함된다. 좀더 구체적으로, 본원 발명의 세포 숙주는 'CanIon 유전자의 게놈 서열' 섹션, 'CanIon cDNA 서열' 섹션, '코딩 영역' 섹션, '폴리뉴클레오티드 구조체' 섹션, '올리고뉴클레오티드 프로브와 프라이머' 섹션에서 밝힌 폴리뉴클레오티드를 보유할 수 있다.

본원 발명에 따른 다른 재조합 숙주 세포는 A1 내지 A18과 이의 보체에서 선택되는 상기 이중대립인자 마커를 보유하는 폴리뉴클레오티드를 포함한다.

본원 발명에 따른 다른 재조합 세포 숙주는 본원에서 밝힌 임의의 벡터, 좀더 구체적으로 '재조합 벡터' 섹션에서 밝힌 임의의 벡터를 보유할 수 있다.

본원 발명의 발현 벡터에 대한 수용자로 사용하는데 적합한 숙주 세포는 다음과 같다:

a) 원핵 숙주 세포: 대장균(*Escherichia coli*)(I.E.DH5-α균주), 바실러스 서브틸리스(*Bacillus subtilis*), 살모넬라 티피무림(*Salmonella typhimurium*) 및 슈도모나스(*Pseudomonas*), 스트렙토마이세스(*Streptomyces*), 스태필로코커스(*Staphylococus*)와 같은 종으로부터 유래된 균주;

b) 진핵 숙주 세포: HeLa 세포(ATCC N° CCL2; N° CCL2.1; N° CCL2.2), Cv 1 세포(ATCC N° CCL70), COS 세포(ATCC N° CRL1650; N° CRL1651), Sf-9 세포(ATCC ° CRL1711), C127 세포(ATCC N° CRL-1804); 3T3(ATCC N° CRL-6361), CHO(ATCC N° CCL-61), 사람 신장 293(ATCC N° 45504; N° CRL-1573), BHK(ECACC N° 84100501; N° 84111301);

c) 다른 포유동물 숙주 세포.

포유동물, 전형적으로 사람 세포에서 CanIon 유전자 발현은 불완전하거나, 또는 대안으로 본원 발명에 따른 CanIon 폴리뉴클레오티드로 동물 세포의 게놈에서 CanIon 유전자 상대를 대체하는 게놈 서열이나 cDNA 서열의 삽입으로 진행된다. 이런 유전적 변형은 전술한 특이적인 DNA 구조체를 이용한 상동성 재조합으로 달성할 수 있다.

사용할 수 있는 세포 숙주중 한가지는 포유동물 접합자, 예를 들면 뮤린 접합자이다. 가령, 뮤린 접합자는 관심있는 정제된 DNA 분자, 예를 들면 100mM NaCl 30μM 스페르민, 70μM 스페르미딘을 함유하는 10mM Tris-HCl, pH 7.4, 250μm EDTA에서 1 ng/mℓ(BAC 삽입체) 내지 3ng/μℓ(P1 박테리오파지 삽입체) 농도 범위로 조정된 정제된 DNA 분자를 미세주사할 수 있다. 미세주사되는 DNA가 대형인 경우에는 폴리아민과 높은 염 농도를 이용하여 DNA의 물리적 파괴를 피할 수 있다(Schedl et al.(1993b)).

본원에서 밝힌 DNA 구조체를 비롯한 본원 발명에 따른 폴리뉴클레오티드는 배아 줄기(ES) 세포주, 바람직하게는 생쥐 ES 세포주에 도입할 수 있다. ES 세포 주는 착상전 배반포의 내세포괴(inner cell mass)의 전능한 비수임된 세포로부터 유래된다. 적절한 ES 세포주는 다음과 같다: ES-E14TG2a(ATCC n° CRL-1821), ES-D3(ATCC n° CRL1934, n° CRL-11632), YS001(ATCC n° CRL-11776), 36.5(ATCC n° CRL-11116). ES 세포는 비수임된 상태로 유지하기 위하여, 배아 표현형을 보존하는데 적합한 신호를 제공하고 ES 세포 유착을 위한 매트릭스로 기능하는 성정 저해된 영양세포의 존재하에 배양한다. 적절한 영양 세포는 Abbondanzo et al.(1993)에서 밝힌 바와 같이 지속 배양되고 Robertson(1987)에서 밝힌 바와 같이 자외선 또는 Pease와 Williams(1990)가 밝힌 바와 같이 저해 농도의 LIF 존재에 의해 성장이 저해되는 거의 모든 생쥐 계통의 13일 내지 14일된 배아의 조직으로부터 확립된 원시 배아 섬유아세포이다.

숙주 세포에서 구조체는 재조합 서열에 의해 인코드되는 유전자 산물을 생산하는 통상적인 방법에 사용할 수 있다.

적합한 숙주의 형질전환 및 적절한 세포밀도까지 숙주의 배양이후, 적당한 수단, 예를 들면 온도 변화 또는 화학적 유도로 선별 프로모터를 유도하고, 세포는 추가로 일정 기간동안 배양한다.

전형적으로, 세포는 원심분리로 수거하고 물리적 또는 화학적 수단으로 파괴하며, 생성된 정제되지 않은 추출물은 추가 정제를 위해 보관한다.

단백질의 발현에 사용되는 미생물 세포는 동결-해동 사이클링, 초음파처리, 기계적 파괴 또는 세포용해제의 이용을 비롯한 임의의 간편한 방법으로 파괴할 수 있다. 이런 방법은 당분야에 공지된 것이다.

유전자도입 동물

본원에서 '유전자도입 동물' 또는 '숙주 동물'은 게놈이 유전적으로, 인위적으로 조작되어 본원 발명에 따른 핵산중 한 가지를 포함하는 동물을 의미한다. 바람직한 동물은 사람제외 동물인데, 여기에는 본원 발명에 따른 핵산의 삽입에 의해 게놈이 인위적으로, 유전적으로 변형되고 *Mus* (예, 생쥐), *Rattus* (예, 쥐), *Oryctogalus* (예, 토끼)에서 선택되는 속(genus)에 속하는 동물이 포함된다. 한 구체예에서, 본원 발명에는 본원 발명에 따른 재조합 벡터 또는 본원 발명의 적중 벡터와의 상동성 재조합으로 파괴된 CanIon 유전자를 보유하는 사람제외 숙주 포유동물 및 동물이 포함된다.

본원 발명의 유전자도입 동물은 복수의 세포에 클론된 재조합 또는 합성 DNA 서열, 좀더 구체적으로 CanIon 코딩 서열, CanIon 조절 폴리뉴클레오티드, 폴리뉴클레오티드 구조체 또는 본원에서 밝힌 안티센스 폴리뉴클레오티드를 인코드하는 DNA 서열을 포함하는 분리되거나 정제된 핵산을 보유한다.

일반적으로, 본원 발명에 따른 유전자도입 동물은 본원에서 밝힌 폴리뉴클레오티드, 재조합 벡터, 세포 숙주를 보유한다. 특히, 본원 발명의 유전자도입 동물은 'CanIon 유전자의 게놈 서열' 섹션, 'CanIon cDNA 서열' 섹션, '코딩 영역' 섹션, '폴리뉴클레오티드 구조체' 섹션, '올리고뉴클레오티드 프로브와 프라이머' 섹션, '재조합 벡터' 섹션, '세포 숙주' 섹션에서 밝힌 폴리뉴클레오티드를 보유할 수 있다.

본원 발명에 따른 다른 유전자도입 동물은 체세포 및/또는 생식 세포에 A1 내지 A18과 이의 보체에서 선택되는 상기 이중대립인자 마커를 보유하는 폴리뉴클레오티드를 포함한다.

적절한 첫 번째 구체예에서, 이들 유전자도입 동물 특히, 고유 CanIon 단백질 또는 대안으로 돌연변이 CanIon 단백질을 인코드하는 폴리뉴클레오티드의 한 개 또는 수 개의 사본이 삽입된 게놈을 갖는 유전자도입 동물은 세포 분화와 관련된 다양한 병리를 연구하기 위한 훌륭한 실험 모델일 될 수 있다.

적절한 두 번째 구체예에서, 이들 유전자도입 동물은 CanIon 유전자의 조절 폴리뉴클레오티드의 조절하에 원하는 폴리펩티드를 발현하여, 궁극적으로 높은 수율로 상기 단백질 합성 및 이런 단백질의 조직 특이적 발현을 달성한다.

본원 발명에 따른 유전자도입 동물의 설계는 당분야에 공지된 통상적 기술로 만들 수 있다. 유전자도입 동물, 특히 유전자도입 생쥐의 생산과 관련하여 미국 특허 4,873,191(Oct. 10, 1989); 5,464,764(Nov 7, 1995); 5,789,215(Aug 4, 1998)를 참조할 수 있다.

본원 발명의 유전자도입 동물은 외인성 유전자 물질을 통합하는 게놈을 갖는 동물을 결과하는 과정으로 만들 수 있다. 이런 과정에는 CanIon 코딩 서열을 인코드하는 유전 물질이나 이의 일부분, CanIon 조절 폴리뉴클레오티드, 또는 본원에서 밝힌 CanIon 안티센스 폴리뉴클레오티드를 인코드하는 DNA 서열을 수득하는 단계가 포함된다.

본원 발명의 재조합 폴리뉴클레오티드는 배아 또는 ES 줄기 세포주에 삽입한다. 이런 삽입은 에렉트로포레이션으로 실시한다(Thomas et al.(1987)). 에렉트 로포레이션되는 세포는 스크리닝하여(선별가능 마커를 통한 선별, PCR 또는 서던 블랏 분석), 가급적 상동성 재조합으로 자신의 게놈에 외인성 재조합 폴리뉴클레오티드를 통합하는 양성 세포를 찾는다. 본원 발명에 따라 사용할 수 있는 양성-음성 선별 과정은 Mansour et al.(1988)에서 기술한다.

이후, 양성 세포는 분리하고 클론하며 생쥐로부터 얻은 3.5일된 배반포에 주사한다(Bradley(1987)). 이후, 배반포는 암컷 숙주 동물에 삽입하고 일정 기간 성장시킨다.

대안으로, 양성 ES 세포는 2.5 일된 8-16 세포 단계 배아(상실배)와 접촉시키고(Wood et al.(1993), Nagy et al.(1993)), ES 세포는 흡수하여 생식세포로 발달하는 세포를 비롯한 배반포의 광범위한 콜로니를 형성한다.

암컷 숙주의 후손은 검사하여 야생형 동물과 유전자도입된, 다시 말하면 삽입된 외인성 DNA를 보유하는 동물을 확인한다.

따라서, 본원 발명은 본원 발명에 다른 핵산, 재조합 발현 벡터 또는 재조합 숙주 세포를 보유하는 유전자도입 동물에 관한다.

본원 발명의 유전자도입 동물로부터 유래된 재조합 세포주

본원 발명의 다른 목적은 본원에서 밝힌 유전자도입 동물로부터 수득된 재조합 숙주 세포에 관한다. 한 구체예에서, 본원 발명에는 본원 발명에 따른 재조합 벡터 또는 적중 벡터와의 상동성 재조합으로 파괴된 CanIon 유전자를 보유하는 사람제외 숙주 포유동물 및 동물로부터 유래된 세포가 포함된다.

재조합 세포주는 본원 발명에 다른 유전자도입 동물의 임의 조직으로부터 수 득되는 세포로부터 시험관내에서 예로써 SV40 거대 T 항원과 같은 *onc* -유전자를 발현하는 벡터로 원시 세포 배양액의 트랜스펙션으로 확립할 수 있다(Chou(1989), Shay et al.(1991))

CanIon 조절물질 및 상호작용 화합물의 스크리닝 방법

다양한 구체예에서, 본원 발명은 CanIon 폴리펩티드, 채널, 폴리뉴클레오티드와 상호작용하거나, 이들과 결합하거나, 활성화시키거나 또는 이들의 발현이나 활성을 저해하는 화합물을 제공한다. 이런 화합물은 유기 또는 무기 화합물일 수 있는데, 여기에는 폴리펩티드, 폴리뉴클레오티드, 지질, 탄수화물, 뉴클레오티드, 아미노산 또는 소형 분자 저해물질이나 활성물질이 포함된다. 본원에서 밝힌 바와 같이, 이런 화합물은 다양한 질환이나 이상의 치료 또는 예방에 유용하다. 적절하게는, CanIon 활성이나 발현의 저해물질이 정신분열증 또는 양극성 장애와 같은 정신 질환의 치료 또는 예방에 사용된다.

CanIon 채널 조절물질을 스크리닝하는 방법

CanIon에 결합할 수 있는 화합물 및 CanIon 기능을 조절할 수 있는 화합물은 질환의 치료에 중요하다. 전압-개폐 이온 채널은 약리학적으로 접근가능하고 다양한 유전자에 의해 인코드되며 다중결합의 단백질 조립체로 움직이고, 따라서 높은 기능적 · 해부학적 특이성을 결과한다는 점에서, 일반적으로 확립된 약물 표적이다. 게다가, 하전된 전압 감수성 아미노산의 움직임에 영향을 주는 이온 채널 개폐는 입체형태에서 변화를 유도하기 때문에, 이온 채널은 예로써 전도(활성화) 또는 비-전도(불활성화) 상태의 채널에만 결합하는 상태 의존성 분자의 설계를 가능하게 한 다.

이에 더하여, 다수의 칼슘 채널 조절물질이 다양한 질환과 이상의 치료 또는 예방에 효과적인 것으로 입증되었다. 가령, 칼슘 채널 저해물질은 다양한 심장 질환과 이상(예, 협심증, 부정맥, 고혈압) 및 CNS와 신경 질환(예, 편두통, 발작의 신경학적 효과, 조병, 신경이완제-유도된 지발성 운동장애, 양극성 장애, 통증, 간질 등)에 효과적인 것으로 밝혀졌다. 이에 더하여, 칼슘 채널 작용물질은 다양한 용도, 예를 들면 국소 마취제의 지속기간을 감소시키고 희석시키는데 효과적인 것으로 밝혀졌다. CasIon 채널의 길항물질과 작용물질은 이들 질환과 이상의 치료 또는 예방에 유사하게 효과적이다. 가령, CanIon 길항물질은 정신분열증과 양극성 장애의 치료 또는 예방에 효과적이다.

전압 개폐 이온 채널은 활성화를 위한 작용물질 결합을 요구하지 않기 때문에, 화합물은 기능적 CanIon 채널에 대하여 스크린한다. 분석에는 고유 또는 클론된 채널을 발현하는 세포(소포 또는 막), 또는 전체 세포에 적용되는 기능적 방사성리간드 결합 분석이 포함된다. 기능적 전세 세포 분석에서는 패치 클램핑(patch clamping)과 같은 전기물리적 기술을 이용한다. 분석은 전압-개폐 채널 유형, 바람직하게는 L, N, T 유형 채널에 관계한다. 채널을 통한 동적 이온 유동은 예로써 형광, 종점 방사성추적기(end-point radiotracer), 세포 생존능 기술로 측정할 수 있다.

또한, 분석에서는 채널에 결합하고 이를 개폐하는 다양한 독소, 독액 또는 화합물을 사용할 수 있다(Denyer et al., Drug Disc. Today 3(7):323-332(1998). 한 구체예에서, 본 분석에서는 예로써 파지티브 또는 네거티브 대조군으로 다수의 공지된 칼슘 채널 작용물질과 길항물질을 이용한다. 적절한 공지된 칼슘채널 길항물질의 예에는 페닐알킬아민(예, 베라파밀), 벤조티아제핀(예, 딜티아젬), 디하이드로피리딘(예, 니페디핀)이 포함되고; 칼슘 채널 작용물질에는 FPL-64176과 BAYK 8644가 포함된다; 나트륨 채널 작용물질에는 베트라초톡신이 포함되고; 나트륨 채널 길항물질에는 스피라돌린, 멕실레틴, U-54494A((+/-)-cis-3,4-디클로로-N-메틸-N-[2-(1-피롤리디닐)-사이클로헥실]-벤즈아마이드)가 포함된다. 이런 화합물은 CanIon 채널에 특이적으로 결합하거나 이를 조절하는 유도체 분자의 설계 또는 발견을 위한 출발 분자 역할을 하는 '선도' 화합물로 사용될 수도 있다.

적절한 구체예에서, 본원 발명의 분석에는 후보 물질의 스크리닝 방법이 포함되는데, 상기 방법은 다음의 단계로 구성된다:

a) (i) CanIon 단백질이나 이의 단편을 포함하는 폴리펩티드를 보유하는 시료 혹은 숙주 세포, 또는 (ii) CanIon 단백질이나 이의 단편을 포함하는 폴리펩티드를 인코드하는 폴리뉴클레오티드를 발현하는 재조합 숙주 세포를 제공하고;

b) 후보 물질을 수득하며;

c) 상기 후보 물질과 숙주 세포를 접촉시키고;

d) CanIon 활성에 대한 상기 후보 물질의 효과를 측정한다.

CanIon 활성에 대한 후보 물질 효과의 측정은 당분야에 공지된 방법으로 달성할 수 있다. 적절하게는, CanIon 활성에 대한 후보 물질의 효과는 작용물질 또는 길항물질 효과이다. 일반적으로, 이온(예, Ca2+ 또는 Na+)을 전달하는 능력이 감소하면 화합물은 CanIon을 저해한다. 이온을 전달하는 능력이 증가하면 화합물은 CanIon을 촉진한다.

CanIon 활성은 임의의 적합한 수단으로 검출할 수 있다. 적절한 예에서, CanIon 활성은 신호전달 현상을 측정하여 검출한다. 신호전달 현상에 세포의 분자 특징이나 파라미터의 적절한 변화가 포함될 수 있는데, 신호전달 현상의 무제한적 예에는 이온 유동에서 변화, 예를 들면 Ca2+ 혹은 Na+의 변화나 발생, 또는 K+ 유동이나 효소 활성화가 포함된다.

한 측면에서, 이온 유동은 예로써 단일 세포 또는 막 패치에서 전체 세포 전류를 측정하는 기술을 이용하여 CanIon 채널의 전기물리적 특성을 측정하여 모니터할 수 있다. 다른 예에서, 형광이나 방사성 라벨은 공지된 CanIon-결합 화합물의 이동, 또는 세포에서 이온(예, 표지된 Ca2+ 또는 Na+) 유동을 검출하는데 사용될 수 있다. 세포의 생리학적 파라미터에 대한 지표, 예를 들면 세포 생존에 대한 형광 지표를 사용할 수 있다. 다른 예에서, 지표의 물리적 위치에서 변화는 예로써 생리학적 지표의 배출이나 흡수를 확인하는 형광 활성화된 세포 정렬(sorting)의 이용으로 확인할 수 있다.

본원 발명의 분석에 사용되는 시료는 i) CanIon 단백질이나 이의 단편을 포함하는 폴리펩티드를 발현하는 숙주세포, 또는 ii) CanIon 단백질이나 이의 단편을 포함하는 폴리펩티드를 인코드하는 폴리뉴클레오티드를 발현하는 재조합 숙주 세포를 함유한다. 적절하게는, 본원 발명의 CanIon 분석에는 기능적 CanIon 폴리펩티드를 발현하는 재조합 숙주 세포의 이용이 포함된다. 숙주 세포는 CanIon 채널의 기능적 알파 아단위, 한가지이상 추가적인 이온 채널 아단위, 또는 CanIon을 포함하는 이온 채널 복합체를 발현 혹은 포함할 수 있다. 적절하게는, 낮은 내인성 이온 채널 발현을 보유하거나, 또는 낮은 배경 이온, 특히 Ca2+ 및/또는 Na+ 전도도를 갖는 숙주 세포가 사용된다.

방사성리간드 결합

한 측면에서, CanIon 채널은 CanIon의 관심 부위에 결합하고 바람직하게는 원하는 조절 효과를 보유하는 고 친화성 리간드를 동정하고 상기 표지된 리간드를 치환화는 검사 화합물의 능력을 검출하여 스크린할 수 있다. 전압 개폐(Ca2+, Na+, K+)에 사용되는 독성학적/약리학적 작용제의 목록은 Denyer et al.(supra)에서 제시한다. 일반적으로, 이런 방법은 표지된 리간드로 동일 부위에 결합하거나 또는 상기 부위에 알로스테릭하게 결합하지만 작용물질이나 길항물질 특성과 관련된 정보를 제공하지 않는 화합물을 검출하는데 적합하다.

세포 기초한 형광과 방사성 추적기 분석

다른 분석에서, CanIon 기능은 형광-이온 지표 또는 방사성표지된 이온으로 투과성 이온의 세포내 농도에서 변화를 측정하여 모니터할 수 있다.

전형적으로, Na+ 채널과 같은 이온 채널은 전압 자극(voltage stimulation)후 수초내에 불활성화된다. Ca2+ 채널은 거의 불활성화를 보이지 않고 높은 K+ 탈분극으로 개방할 수 있다. 세포 기초한 형광과 방사성 추적기 분석에서, CanIon 채널은 독소 혹은 임의의 검사 화합물 또는 높은 K+ 탈분극으로 활성화시켜, 연장된 기간(최대 수분)동안 개방시킬 수 있다.

형광 기초한 분석에서, 형광 Ca2+ 염료(Fluo-3, 칼슘 그린-1, Molecular Probes, OR, U.S.A)가 사용 가능하다. Ca2+ 채널은 등장성 용액으로 막을 탈분극시키거나, 또는 독소나 다른 화합물로 Na+ 채널을 탈분극시켜 활성화시킬 수 있는데, 세포에서 형광의 일시적 움직임은 20 내지 60초동안 측정된다. 형광 측정 시스템과 장치는 Denyer et al.(supra)에서 상세하게 기술한다. 방사성 추적기 22Na+와 14C-구아니딘은 Na+ 채널 분석에 통상적으로 사용되고 45Ca2+는 Ca2+ 채널 분석에 사용된다. Denyer et al.(supra)에서 밝힌 적절한 구체예에서, 고성능 CanIon 세포 기초한 분석에 *Cytostar* -T 신틸레이션 마이크로플레이트(Amersham International, U.K.)가 사용된다.

추가 분석에서, 이온 채널의 Ca2+ 기능은 막 전위 지표로 막 전위를 측정하여 모니터한다. 생물학적 막의 높은 전기 저항성은 적은 이온 전류가 혈장 막을 통과하여 막 전위에 많은 변화를 유도할 수 있도록 한다. 따라서, 전압 분석은 막을 통한 전반적인 이온 유동을 검출하는데 용이하게 이용될 수 있다. 일반적으로, 세포주는 내생적 이온 채널로부터 효과가 최소화되도록 선택된다. 다양한 염료가 막 전위 지표 염료로 가용하는데, 이들은 빠른 반응 염료와 느린 반응

염료 및 FRET-기초한 전압 센서 염료로 세분된다(Auora Biosciences, CA, USA; Gonzalez et al., Drug Disc. Today 4(9): 431-439(1999).

세포 생존능

세포 생존능 분석에서, 이온 채널 활성과 이온 유동은 세포 생존능과 직접적으로 관련된다. 이온 채널 표적의 검사에 효모와 포유동물 세포 시스템이 가용하 다. 가령, 이온-특이적 K+ 흡수 결핍성 맥주 효모균(*Saccharomyces cerevisiae*) 세포주를 이용한 효모 시스템이 사용되는데, 여기서 기능적 K+ 채널은 상기 세포주에서 발현되고, 따라서 K+ 흡수를 회복하고 세포 생존을 촉진한다(Anderson et al., Symp. Soc. Exp. Biol. 48: 85-97(1994)). Ca2+ 또는 Na+ 채널에 대한 이런 분석은 CanIon 기능을 차단할 수 있는 화합물을 동정하는데 이용될 수 있다. 포유동물 세포 시스템, 예를 들면 비색 세포 생존능 판독으로 포유동물 신경아세포종 세포를 이용한 Na+ 채널 분석 역시 가용하다. Na+ 채널로 개방물질과 Na+/K+ 펌프 저해물질로 처리된 세포는 치명적인 세포내 Na+ 과부하를 촉진한다. 채널을 차단할 수 있는 검사 화합물로 처리하면 세포 생존능이 향상되는 반면, 채널 개방을 강화시키는 화합물은 세포 사멸을 더욱 촉진한다(Manger et al., Anal. Biochem. 214: 190-194(1993)).

전기생리(electrophysiology)

전기생리학적 전압-클램핑 기술에는 한가지이상 채널을 통한 이온 전류 유동의 측정이 포함된다. 막 전압은 단일 마이크로전극으로 조절하고, 전류는 단일 세포 또는 막 패치를 통하여 측정한다(Hamil, Pfugers Arch. 391, 85-100(1981)). 따라서, 이온 전류는 관심있는 검사 화합물의 존부하에 측정할 수 있다. 대규모 화합물 스크리닝 시스템이 설계되었다(Neurosearch A/S, Glostrup, Denmark; Olesen et al., Voltage gated ion channel modulators, 7-8 December, Philadelphia PA, USA(1995); Denyer et al., supra).

CanIon 폴리펩티드와 상호작용하는 물질의 스크리닝 방법

본원 발명에서, 리간드는 CanIon 단백질, 이의 단편 또는 변이체와 결합하거나 또는 CanIon, 이의 단편 또는 이의 변이체를 코딩하는 폴리뉴클레오티드의 발현을 조절할 수 있는 단백질, 펩티드, 항체 또는 임의의 화학적 합성화합물과 같은 분자를 의미한다.

본원 발명에 따른 리간드 스크리닝 방법에서, CanIon 단백질의 추정 리간드로 검사할 생물 시료 또는 정의된 분자는 상응하는 정제된 CanIon 단백질, 예를 들면 재조합 세포 숙주에 의해 생산되는 상응하는 정제된 재조합 CanIon 단백질과 접촉시켜 상기 단백질과 검사할 추정 리간드 분자간 복합체를 형성시킨다.

가령, SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편과 약물이나 소형 분자, 예를 들면 조합 화학, HPLC와 결부된 마이크로투석(Wang et al.(1997)) 또는 친화성 모세관 전기영동 방법(Bush et al.(1997))을 통하여 만들어진 분자의 복합체를 형성시킨다.

다른 방법에서, SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편과 상호작용하는 펩티드, 약물, 지방산, 리포프로테인 또는 작은 분자는 다음과 같은 분석법으로 확인할 수 있다. 결합을 검사할 분자는 검출가능한 라벨, 예를 들면 형광, 방사성 또는 효소 태그로 표지하고, 특이적인 결합이 가능한 조건하에 고정된 CanIon 단백질이나 이의 단편과 접촉시킨다. 비-특이적으로 결합된 분자를 제거한 이후, 결합된 분자는 적절한 수단으로 검출한다.

본원 발명의 다른 목적은 CanIon 폴리펩티드와 상호작용하는 후보 물질을 스크리닝하는 방법과 키트에 관한다.

본원 발명은 CanIon 단백질, 이의 단편 또는 변이체와 상호작용하는 물질을 스크리닝하는 방법을 제시한다. CanIon 단백질, 이의 단편 또는 변이체와 공유 또는 비-공유 결합하는 능력으로 인해, 이들 물질은 시험관내와 생체내에서 효과적으로 사용할 수 있다.

시험관내에서, 이런 상호작용 분자는 시료, 바람직하게는 생물 시료에서 CanIon 단백질의 존재를 확인하는 검출 수단으로 사용할 수 있다.

후보 물질을 스크리닝하는 방법은 다음의 단계로 구성된다.

a) SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편을 포함하는 또

는 이들을 포함하는 폴리펩티드를 제공하고;

b) 후보 물질을 수득하고;

c) 상기 폴리펩티드와 후보 물질을 접촉시키고;

d) 상기 폴리펩티드와 후보 물질간에 형성된 복합체를 검출한다.

또한, 본원 발명은 CanIon 폴리펩티드와 상호작용하는 후보 물질을 스크리닝하는 키트에 관하는데, 상기 키트는 다음과 같이 구성된다.

a) SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편;

b) 선택적으로, CanIon 단백질, 이의 단편 또는 변이체와 후보 물질간에 형성된 복합체를 검출하는데 유용한 수단.

전술한 키트의 적절한 구체예에서, 검출 수단은 CanIon 단백질, 이의 단편 또는 변이체에 대한 단클론이나 다클론 항체을 포함한다.

다양한 후보 물질이나 분자는 CanIon 폴리펩티드와의 상호작용을 분석할 수 있다. 이들 물질이나 분자에는 고유 혹은 합성 유기 화합물 또는 생물 기원의 분자, 예를 들면 폴리펩티드가 포함되지만 이들에 국한되지 않는다. 후보 물질이나 분자가 폴리펩티드를 포함하는 경우에, 상기 폴리펩티드는 파지-기초된 무작위 펩티드 라이브러리에 속하는 파지 클론의 발현 산물 또는 대안으로 이중-하이브리드 스크리닝 분석을 실시하는데 적합한 벡터에 클론된 cDNA 라이브러리의 발현 산물일 수 있다.

또한, 본원 발명은 전술한 스크리닝 방법을 실행하는데 유용한 키트에 관한다. 적절하게는, 이런 키트는 CanIon 폴리펩티드, 이의 단편 또는 변이체 및 선택적으로 CanIon 단백질, 이의 단편 또는 변이체와 후보 물질간에 형성된 복합체를 검출하는데 유용한 수단으로 구성된다. 적절한 구체예에서, 검출 수단은 CanIon 단백질, 이의 단편 또는 변이체에 대한 단클론이나 다클론 항체로 구성된다.

A. 무작위 펩티드 라이브러리로부터 수득된 후보 리간드

스크리닝 방법의 특정 구체예에서, 추정 리간드는 파지 벡터에 포함된 DNA 삽입체의 발현 산물이다(Parmley and Smith, 1988). 특히, 무작위 펩티드 파지 라이브러리가 사용된다. 무작위 DNA 삽입체는 8 내지 20개 아미노산 길이의 펩티드를 인코드한다(Oldenburg K.R. et al., 1992; Valadon P., et al., 1996; Lucas A.H., 1994; Westerink M.A.J., 1995; Felici F. et al., 1991). 이런 특정 구체예에 따라, 고정된 CanIon 단백질에 결합하는 단백질을 발현하는 재조합 파지는 유지시키고, CanIon 단백질과 재조합 파지간에 형성된 복합체는 CanIon 단백질에 대한 다클론이나 단클론 항체로 면역침전시킬 수 있다.

일단 재조합 파지에서 리간드 라이브러리가 작제되면, 파지 개체군은 고정된 CanIon 단백질과 접촉시킨다. 이후, 복합체는 세척하여 비-특이적으로 결합된 재조합 파지를 제거한다. 그 다음, CanIon 단백질에 특이적으로 결합하는 파지는 완충액(산성 pH)으로 용출하거나 또는 하이브리도마 항-CanIon에 의해 만들어진 단클론 항체로 면역침전시키고, 이런 파지 개체군은 박테리아(가령, 대장균(*E. coli*))의 과잉-감염(over-infection)으로 증폭시킨다. 선별 단계는 수회, 바람직하게는 2-4회 반복하여 좀더 특이적인 재조합 파지 클론을 선별한다. 최종 단계는 선별된 재조합 파지 클론에 의해 생성된 펩티드를 감염된 박테리아에서 발현과 분리로 특성화시키는 단계, 다른 숙주-벡터 시스템에서 파지 삽입체를 발현시키는 단계 또는 선별된 재조합 파지에 포함된 삽입체를 서열분석하는 단계로 구성된다.

B. 경합 실험으로 수득된 후보 리간드

대안으로, SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내 지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편과 결합하는 펩티드, 약물 또는 소형 분자들은 경합 실험으로 확인할 수 있다. 이런 분석에서, CanIon 단백질 또는 이의 단편은 표면, 예를 들면 플라스틱 평판에 고정시킨다. 펩티드, 약물 또는 소형 분자들은 CanIon 단백질 리간드로 알려진 검출가능한 라벨의 존재하에 함량을 점진적으로 증가시키면서 고정된 CanIon 단백질 또는 이의 단편과 접촉시킨다. 가령, CanIon 리간드는 형광, 방사성 또는 효소 태그로 검출가능하게 표지할 수 있다. CanIon 단백질이나 이의 단편에 결합하는 검사 분자의 능력은 검사 분자의 존재하에 결합된 검출가능하게 표지된 리간드의 함량을 측정하여 확인한다. 검사 분자가 존재하는 경우에 CanIon 단백질이나 이의 단편에 결합하는 공지된 리간드의 함량이 줄어드는데, 이는 검사 분자가 CanIon 단백질이나 이의 단편과 결합할 수 있다는 것을 암시한다.

C. 친화성 크로마토그래피에 의해 수득된 후보 리간드

SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편과 상호작용하는 단백질 또는 다른 분자는 CanIon 단백질이나 이의 단편을 보유하는 친화성 칼럼을 이용하여 발견할 수 있다. CanIon 단백질이나 이의 단편은 통상적 기술, 예를 들면 아가로즈 Affi Gel 또는 당업자에게 공지된 다른 매트릭스와 같은 적합한 칼럼 매트릭스에 대한 화학적 결합으로 칼럼에 부착시킬 수 있다. 상기 방법의 일부 구체예에서, 친화성 칼럼은 CanIon 단백질이나 이의 단편이 글루타티온 S 트랜스퍼라제(GST)와 융합된 키메라 단백질을 함유한다. 세포 단백질의 혼합물 또는 전술한 발현된 단백질의 풀(pool)을 친화성 칼럼에 가한다. 이후, 칼럼에 부착된 CanIon 단백질이나 이의 단편과 상호작용하는 단백질 또는 다른 분자는 분리하고 Ramunsen et al.(1997)에서 밝힌 바와 같이 2-D 전기영동 겔에서 분석할 수 있다. 대안으로, 친화성 칼럼에 잔존하는 펩티드는 전기영동 기초한 방법으로 정제하고 서열 분석할 수 있다. 상기 방법은 항체를 분리하거나, 파지 디스플레이 산물을 스크리닝하거나 또는 파지 디스플레이 사람 항체를 스크리닝하는데 활용할 수 있다.

D. 광학적 바이오센서 방법으로 수득된 후보 리간드

SEQ ID No 5의 적어도 6개 아미노산, 바람직하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편과 상호작용하는 단백질은 광학적 바이오센서로 스크리닝할 수도 있다(Edwards and Leatherbarrow(1997), Szabo et al.(1995)). 이 기술은 실시간으로 분자간 상호작용의 검출이 가능하고, 표지된 라벨을 필요로 하지 않는다. 이 기술은 표면 플라스몬 공명(SPR) 현상에 기초한다. 간단히 말하면, 검사할 후보 리간드 분자는 표면(예, 카복시메틸 덱스트란 매트릭스)에 부착시킨다. 광빔은 검사 시료를 함유하지 않은 표면의 측면으로 향하고 상기 표면에 의해 반사된다. SPR 현상은 각도 및 파장과 특이적으로 연관하여 반사된 광의 밀도를 감소시킨다. 후보 리간드 물질의 결합은 표면에서 굴절율에 변화를 초래하고, 이런 변화는 SPR 신호에서 변화로 감지된다. CanIon 단백질 이나 이의 단편과 상호작용할 수 있는 후보 리간드 분자 또는 물질의 스크리닝에서, CanIon 단백질이나 이의 단편은 표면에 고정시킨다. 상기 표면은 분석할 후보 물질이 통과하게 되는 세포로 구성된다. CanIon 단백질이나 이의 단편에 후보 물질의 결합은 SPR 신호의 변화로 감지된다. 검사된 후보 물질은 조합 화학으로 만들어진 단백질, 펩티드, 탄수화물, 지질 또는 소형 분자일 수 있다. 이 기술은 내인성 또는 재조합으로 발현된 CanIon 단백질을 보이는 진핵이나 원핵 세포 또는 지질 소포를 표면에 고정시켜 실시할 수도 있다.

상기 방법의 최대 이점은 CanIon 단백질 및 CanIon 단백질과 상호작용하는 분자간 결합 속도를 측정할 수 있다는 점이다. 따라서, 강한 또는 상대적으로 약한 결합 상수를 통하여 CanIon 단백질이나 이의 단편과 상호작용하는 리간드 분자를 특이적으로 선별하는 것이 가능하다.

E. 이중-하이브리드 스크리닝 분석을 통하여 수득된 후보 리간드

이중-하이브리드 시스템은 생체내에서 단백질-단백질 상호작용을 연구하기 위한 것으로(Fields and Song 1989), 이스트 Gal4 단백질의 DNA 결합 도메인에 미끼 단백질의 융합에 의존한다. 이 기술은 미국 특허 5,667,973과 5,283,173(Fields et al.)에서도 기술한다.

이중-하이브리드 분석에 의한 라이브러리 스크리닝의 전반적 과정은 Harper et al.(1993), Cho et al.(1998) 또는 Fromont-Racine et al.(1997)에서 밝힌 과정으로 실시할 수 있다.

미끼 단백질이나 폴리펩티드는 SEQ ID No 5의 적어도 6개 아미노산, 바람직 하게는 적어도 8개 내지 10개의 아미노산, 좀더 바람직하게는 적어도 12, 15, 20, 25, 30, 40, 50 또는 100개의 연속 아미노산을 포함하는 CanIon 단백질이나 이의 단편으로 구성되거나 또는 이를 포함한다.

좀더 구체적으로, CanIon 단백질, 이의 단편 또는 변이체를 인코드하는 뉴클레오티드 서열은 GAL4 단백질의 DNA 결합 도메인을 인코드하는 폴리뉴클레오티드에 융합되는데, 상기 융합된 뉴클레오티드 서열은 적합한 발현 벡터, 예를 들면 pAS2 또는 pM3에 삽입한다.

이후, 특별하게 설계된 벡터에 사람 cDNA 라이브러리를 작제하여, 사람 cDNA 삽입체가 GAL4 단백질의 전사 도메인을 인코드하는 벡터상의 뉴클레오티드 서열과 융합되도록 한다. 적절하게는, 사용된 벡터는 pACT 벡터이다. 사람 cDNA 라이브러리의 뉴클레오티드 삽입체에 의해 인코드되는 폴리펩티드는 '먹이' 폴리펩티드라 한다.

세 번째 벡터는 검출가능한 마커 유전자, 예를 들면 전사 활성 도메인과 DNA 결합 도메인 모두를 포함하는 완전 Gal4 단백질의 결합에 반응하는 조절 서열의 조절하에 위치하는 베타 갈락토시다제 유전자 또는 CAT 유전자를 보유한다. 가령, 벡터 pG5EC를 사용할 수 있다.

또한, 2가지 상이한 이스트 균주가 사용된다. 예로써, 2가지 상이한 이스트 균주는 다음과 같다:

- Y190, 이의 표현형은 (MATa, Leu2-3, 112 ura3-12, trpl-901 his3-D200 ade2-101, gal4Dgall80D URA3 GAL-LacZ, LYS GAL-HIS3, cyh)이다;

- Y187, 이의 표현형은 (MATa gal4 gal80 his3 trpl-901 ade2-101 ura3-52 leu2-3, -112 URA3 GAL-lacZmet)으로, Y190의 반대 교배형(mating type)이다.

간단히 말하면, 20㎍의 pAS2/CanIon와 20㎍의 pACT-cDNA 라이브러리는 이스트 균주 Y190에 공동-형질전환시킨다. 형질전환체는 히스티딘, 루이신, 트립토판이 결핍되지만 히스티딘 합성 저해물질 3-AT(50mM)을 함유하는 최소 배지에서 성장에 대하여 선별한다. 양성 콜로니는 필터 리프트(lift) 분석으로 베타 갈락토시다제에 대하여 선별한다. 이후, 이중 양성 콜로니(*His* $^{+}$, *beta-gal* $^{+}$)는 히스티딘과 루이신이 결핍되지만 트립토판과 사이클로헥사미드(10mg/㎖)를 함유하는 평판에 배양하여 pAS2/CanIon 플라스미드는 없지만 pACT-cDNA 라이브러리 플라스미드는 계속 유지하는 콜로니를 선별한다. 결과의 Y190 균주는 Gal4 융합체로 CanIon 또는 무관한 대조군 단백질; 예를 들면 사이클로필린 B, 라민 또는 SNF1을 발현하는 Y187 균주와 교배시키고[Harper et al.(1993), Bram et al.(Bram RJ et al., 1993)], 필터 리프트(lift) 분석으로 베타 갈락토시다제에 대하여 선별한다. 대조군 Gal4 융합체와의 교배후 *beta-gal* $^{-}$ 인 이스트 클론은 가양성(false positive)이다.

본원 발명에 따른 이중-하이브리드 방법의 다른 구체예에서, CanIon, 이의 단편이나 변이체와 세포 단백질간 상호작용은 매치메이커 Hybrid System 2(Catalog No. K1604-1, Clontech)를 이용하여 평가할 수 있다. 매치메이커 Two Hybrid System 2(Catalog No. K1604-1, Clotech)에 동반된 매뉴얼에 밝힌 바와 같이, CanIon 단백질이나 이의 일부분을 인코드하는 핵산은 발현 벡터에 삽입시켜 이들이 이스트 전사 활성자 GAL4의 DNA 코딩 영역을 인코드하는 DNA와 쌍을 이루게 한다. 원하는 cDNA, 바람직하게는 사람 cDNA는 두 번째 발현 벡터에 삽입시켜 이들이 GAL4의 활성화 도메인을 인코드하는 DNA와 쌍을 이루게 한다. 이중 발현 플라스미드는 이스트로 형질전환시키고, 이스트는 HIS3 유전자의 GAL4 의존성 발현뿐만 아니라 각 발현 벡터에서 선별 마커의 발현을 선별하는 선별 배지에 도말한다. 히스티딘이 결핍된 배지에서 성장할 수 있는 형질전환체는 GAL4 의존성 lacZ 발현에 대하여 스크리닝한다. 히스티딘 선별과 lacZ 분석 모두에서 양성인 세포는 CanIon과 초기에 선별된 cDNA 삽입체에 의해 인코드되는 단백질이나 펩티드간 상호작용을 보유한다.

CanIon의 조절 서열과 상호작용하는 물질을 스크리닝 하는 방법

또한, 본원 발명은 CanIon 유전자의 조절 서열, 예를 들면 프로모터 또는 인헨서 서열과 상호작용 할 수 있는 물질이나 분자를 스크리닝하는 방법에 관한다.

CanIon 유전자의 조절 서열, 좀더 구체적으로 5'와 3' 조절 영역의 폴리뉴클레오티드, 이의 단편 또는 변이체에서 선택되는 뉴클레오티드 서열, 바람직하게는 본원 발명의 이중대립인자 마커중 한가지를 포함하는 변이체와 상호작용할 수 있는 단백질을 인코드하는 핵산은 Matchmaker One-Hybrid System kit(Catalog Ref. n° K1603-1, Clontech)에 동봉된 책자에 기술되어 있는 바와 같이 단일-하이브리드 시스템으로 동정할 수 있다. 간단히 말하면, 표적 뉴클레오티드 서열은 선별가능 리포터 서열의 상류에 클론시키고, 생성된 DNA 구조체는 맥주 효모균(*Saccharomyces cerevisiae*) 게놈에 통합시킨다. 이후, 게놈에 리포터 서 열을 포함하는 이스트 세포는 CanIon 유전자의 조절 서열에 결합하는 후보 단백질을 인코드하는 cDNA와 GAL4와 같은 이스트 전사 인자의 활성자 도메인을 인코드하는 서열간 융합 분자을 포함하는 라이브러리로 형질전환시킨다. 재조합 이스트 세포는 리포터서열을 발현하는 세포를 선별하는 액체 배지에 도말한다. 이렇게 선별된 재조합 이스트 세포는 CanIon 유전자의 조절 서열에 결합할 수 있는 융합 단백질을 보유한다. 이후, 상기 융합 단백질을 인코드하는 cDNA는 서열분석하고 시험관내에서 발현이나 전사 벡터에 클론시킬 수 있다. CanIon 유전자의 표적 조절 서열에 상기 인코드된 폴리펩티드의 결합은 당업자에게 공지된 기술, 예를 들면 겔 지체 분석 또는 DNAse 보호 분석으로 확인할 수 있다.

겔 지체 분석은 CanIon 유전자의 조절 서열과 상호작용할 수 있는 후보 분자를 스크리닝하기 위하여 독립적으로 실시할 수 있다(Fried and Crothers(1981); Garner and Revzin(1981); Dent and Latchman(1993)). 이들 기술은 단백질에 결합하는 DNA 단편이 결합하지 않은 동일한 DNA 단편보다 느리게 이동한다는 원칙에 기초한다. 간단히 말하면, 표적 뉴클레오티드 서열을 표지한다. 이후, 표지된 뉴클레오티드 서열은 전사 인자를 포함하는 세포의 전체 핵 추출물 또는 검사할 상이한 후보 물질과 접촉시킨다. CanIon 유전자의 표적 조절 서열과 후보 물질이나 전사 인자간 상호작용은 겔이나 모세관 전기영동이후 이동 지체로 감지된다.

CanIon 유전자의 발현을 조절하는 리간드를 스크리닝하는 방법

본원 발명의 다른 목적은 CanIon 단백질의 발현을 조절하는 분자를 스크리닝하는 방법이다. 이런 스크리닝 방법은 다음의 단계로 구성된다:

a) 자체 프로모터의 조절하에 CanIon 단백질, 이의 단편 또는 변이체를 인코드하는 뉴클레오티드 서열로 트랜스펙션된 원핵이나 진핵 세포를 배양하고;

b) 배양된 세포를 검사할 분자와 접촉시키고;

c) CanIon 단백질, 이의 단편 또는 변이체의 발현을 정량한다.

한 구체예에서, CanIon 단백질, 이의 단편 또는 변이체를 인코드하는 뉴클레오티드 서열은 이중대립인자 마커 A12 또는 A16중 적어도 한가지의 대립유전자 및 이의 보체를 포함한다.

당업자에게 공지된 DNA 재조합 기술을 활용하여, DNA 서열을 인코드하는 CanIon 단백질은 발현 벡터에서 프로모터 서열로부터 하류에 삽입한다. 가령, CanIon 유전자의 프로모터 서열은 5'조절 영역의 핵산에 포함시킨다.

CanIon 단백질의 발현 정량은 mRNA 수준 또는 단백질 수준에서 달성할 수 있다. 후자의 경우에, 예로써 ELISA 또는 RIA 분석에서 생산된 CanIon 단백질의 함량을 정량하는데 다클론이나 단클론 항체를 사용할 수 있다.

적절한 구체예에서, CanIon mRNA의 정량은 CanIon에 특이적인 프라이머를 이용한 배양된 CanIon-트랜스펙션된 숙주 세포의 전체 mRNA의 역전사로 수득되는 cDNA의 정량적 PCR 증폭으로 달성한다.

또한, 본원 발명은 CanIon 유전자의 발현 수준을 증가 또는 감소시킬 수 있는 물질이나 분자를 스크리닝하는 방법에 관한다. 이런 방법으로, 당업자는 CanIon 유전자의 발현 수준에 조절 효과를 보이고 전술한 질환을 앓고 있는 환자를 치료하는 제약학적 조성물에 함유되는 활성 성분으로 유용한 물질을 선별할 수 있 다.

따라서, 본원 발명에는 CanIon 유전자의 발현을 조절하는 후보 물질이나 분자를 스크리닝하는 방법이 포함되는데, 상기 방법은 다음의 단계로 구성된다:

- 핵산을 보유하는 재조합 세포 숙주를 제공하고, 여기서 상기 핵산은 검출가능한 단백질을 인코드하는 폴리펩티드의 상류에 위치하는 5'조절 영역의 뉴클레오티드 서열, 이의 생물학적 활성 단편 또는 변이체를 포함하고;

- 후보 물질을 수득하고;

- 검출가능한 단백질을 인코드하는 폴리뉴클레오티드의 발현 수준을 조절하는 후보 물질의 능력을 확인한다.

다른 구체예에서, 5'조절 영역의 뉴클레오티드 서열, 이의 생물학적 활성 단편 또는 변이체를 포함하는 핵산에는 SEQ ID No 4의 CanIon cDNA의 5'UTR 영역, 이의 생물학적 활성 단편 또는 변이체가 포함된다.

검출가능한 단백질을 인코드하는 적절한 폴리뉴클레오티드에는 베타 갈락토시다제, 녹색 형광 단백질(GFP), 클로람페니콜 아세틸 트랜스퍼라제(CAT)를 인코드하는 폴리뉴클레오티드가 포함된다.

또한, 본원 발명은 전술한 스크리닝 방법을 실행하는데 유용한 키트에 관한다. 적절하게는, 이런 키트는 검출가능한 단백질 또는 CanIon 단백질, 이의 단편이나 변이체를 인코드하는 폴리펩티드의 상류에 위치하고 여기에 동작가능하게 연결된 5'조절 영역의 뉴클레오티드 서열, 이의 생물학적 활성 단편 또는 변이체의 발현을 가능하게 하는 재조합 벡터로 구성된다.

CanIon 유전자의 발현을 조절하는 후보 물질이나 분자를 스크리닝하는 방법의 다른 구체예에서, 상기 방법은 다음과 같은 단계로 구성된다:

a) 핵산을 보유하는 재조합 세포 숙주를 제공하고, 여기서 상기 핵산은 SEQ ID No 4의 CanIon cDNA의 5'UTR 서열, 이의 생물학적 활성 단편 또는 변이체를 포함하고, 상기 5'UTR 서열, 이의 생물학적 활성 단편 또는 변이체는 검출가능한 단백질을 인코드하는 폴리펩티드에 동작가능하게 연결되고;

b) 후보 물질을 수득하고;

c) 검출가능한 단백질을 인코드하는 폴리뉴클레오티드의 발현 수준을 조절하는 후보 물질의 능력을 확인한다.

상기 스크리닝 방법의 특정 구체예에서, SEQ ID No 4의 CanIon cDNA의 5'UTR 서열, 이의 생물학적 활성 단편 또는 변이체에서 선택되는 뉴클레오티드 서열을 포함하는 핵산은 전술한 CanIon 5'UTR 서열과 내인성의 프로모터 서열을 보유한다.

상기 스크리닝 방법의 다른 특정 구체예에서, SEQ ID No 4의 CanIon cDNA의 5'UTR 서열, 이의 생물학적 활성 단편 또는 변이체에서 선택되는 뉴클레오티드 서열을 포함하는 핵산은 전술한 CanIon 5'UTR 서열과 외인성의 프로모터 서열을 보유한다.

다른 적절한 구체예에서, SEQ ID No 4의 CanIon cDNA의 5'UTR 서열 또는 이의 생물학적 활성 단편을 포함하는 핵산은 A12 또는 A16에서 선택되는 이중대립인자 마커 또는 이의 보체를 보유한다.

또한, 본원 발명은 CanIon 유전자의 발현을 조절하는 후보 물질을 스크리닝 하는 키트에 관하는데, 상기 키트는 상기 핵산은 SEQ ID No 4의 CanIon cDNA의 5'UTR 서열, 이의 생물학적 활성 단편 또는 변이체를 보유하는 핵산을 포함하는 재조합 벡터로 구성되고, 상기 5'UTR 서열, 이의 생물학적 활성 단편 또는 변이체는 검출가능한 단백질을 인코드하는 폴리펩티드에 동작가능하게 연결된다.

CanIon의 발현 수준과 패턴은 국제 특허 출원 WO 97/05277에서 밝힌 긴 프로브와의 용액 혼성화로 분석할 수 있다. 간단히 말하면, 전술한 CanIon cDNA이나 CanIon 게놈 DNA 또는 이들의 단편은 박테리오파지(T3, T7 or SP6) RNA 중합효소 프로모터에서 매우 가까운 하류의 클로닝 부위에 삽입시켜 안티센스 RNA를 만든다. 적절하게는, CanIon 삽입체는 게놈 DNA 서열 또는 cDNA 서열의 100개이상 연속 뉴클레오티드로 구성된다. 상기 플라스미드는 선형화시키고, 변형된 리보뉴클레오티드(즉, 비오틴-UTP와 DIG-UTP)을 포함하는 리보뉴클레오티드의 존재하에 전사시킨다. 과량의 이중 표지된 RNA는 관심있는 세포 또는 조직으로부터 분리된 mRNA와 용액에서 혼성화된다. 혼성화는 평균적인 엄격한 조건(40-50℃, 16시간, 80% 폼아마이드, 0.4M NaCl 완충액 pH 7.8)하에 실시한다. 혼성화되지 않은 프로브는 단일-사슬 RNA에 특이적인 리보뉴클레아제(즉, RNases CL3, T1, Phy M, U2 또는 A)로 절단하여 제거한다. 비오틴-UTP 수식의 존재는 스트렙타비딘으로 피복된 마이크로적정 평판에 하이브리드의 포획을 가능하게 한다. DIG 수식의 존재는 알칼리성 포스파타제에 결합된 항-DIG 항체를 이용한 ELISA로 하이브리드를 검출하고 정량하는 것을 가능하게 한다.

CanIon 유전자 발현의 정량적 분석은 어레이를 이용하여 실시할 수도 있다. 본원에서 어레이는 mRNA 발현을 특이적으로 검출할 수 있는 충분한 길이의 핵산을 포함하는 일차원, 이차원 또는 다차원 정렬을 의미한다. 가령, 어레이는 발현 수준을 평가해야 하는 유전자로부터 유래된 복수의 핵산을 포함할 수 있다. 어레이는 CanIon 게놈 DNA, CanIon cDNA 서열, 이의 상보적 서열 또는 단편, 특히 본원 발명에 따른 이중대립인자 마커중 적어도 한가지, 바람직하게는 이중대립인자 A1 내지 A17중 적어도 한가지를 포함하는 서열을 포함할 수 있다. 바람직하게는 단편은 적어도 15개의 뉴클레오티드이다. 다른 구체예에서, 단편은 적어도 25개의 뉴클레오티드이다. 일부 구체예에서, 단편은 적어도 50개의 뉴클레오티드이다. 좀더 적절하게는, 단편은 적어도 100개의 뉴클레오티드이다. 다른 적절한 구체예에서, 단편은 100개이상의 뉴클레오티드이다. 일부 구체예에서, 단편은 500개이상의 뉴클레오티드이다.

가령, CanIon 유전자 발현의 정량적 분석은 상보성 DNA 마이크로어레이로 실시할 수 있다(Schena et al.(1995 and 1996)). 전장 CanIon cDNA 또는 이의 단편은 PCR로 증폭하고, 고속 로봇공학을 이용하여 96-웰 마이크로역가 평판으로부터 실릴레이트 현미경 슬라이드로 정렬시킨다. 프린트된 어레이는 습한 체임버에서 배양하여 어레이 요소를 수화시키고 1분동안 0.2% SDS에서 1회, 1분동안 물에서 2회, 5분동안 소디움 보로하이드라이드 용액에서 1회 세척한다. 어레이는 95℃에서 2분동안 물에 담그고 1분동안 0.2% SDS로 이전하며 물로 2회 세척하고 대기 건조시키며 25℃에서 암실에 보관한다.

세포 또는 조직 mRNA는 분리하거나 또는 상업적으로 구입하고, 프로브는 1회 의 역전사로 준비한다. 프로브는 60℃에서 14 x 14㎜ 유리 커브슬립하의 1㎠ 마이크로어레이에 6-12시간동안 혼성화시킨다. 어레이는 25℃에서 낮은 엄밀도 세척 완충액(1 x SSC/0.2% SDS)에 5분동안, 이후 실온에서 높은 엄밀도 세척 완충액(0.1 x SSC/0.2% SDS)에 10분동안 세척한다. 어레이는 맞춤 필터 세트를 갖춘 형광 레이저 스캐닝으로 0.1 x SSC에서 스캔한다. 정확한 특이적 발현 측정은 독립된 2회 혼성화의 비율을 평균하여 얻는다.

CanIon 유전자 발현의 정량 분석은 상보성 DNA 어레이에서 전장 CanIon cDNA 또는 이의 단편으로 실시할 수도 있다(Pietu et al.(1996)). 전장 CanIon cDNA 또는 이의 단편은 PCR 증폭하고 막에 스팟한다. 이후, 다양한 조직이나 세포로부터 기원한 mRNA는 방사성 뉴클레오티드로 표지한다. 통제된 조건하에 혼성화 및 세척이후, 혼성화된 mRNA는 포스포이미징(phosphoimaging) 또는 자가방사법으로 검출한다. 중복 실험을 실시하고, 이후 상이하게 발현된 mRNA의 정량 분석을 실시한다.

대안으로, CanIon 게놈 DNA, CanIon cDNA 또는 이의 단편을 이용한 발현 분석은 고밀도 뉴클레오티드 분석으로 실시할 수 있다(Lockhrt et al.(1996); Sosnowsky et al.(1997)). CanIon 게놈 DNA 혹은 CanIon cDNA 서열, 특히 본원 발명에 따른 이중대립인자 마커중 적어도 한가지를 포함하는 서열, 바람직하게는 A1 내지 A17에서 선택되는 이중대립인자 마커중 적어도 한가지를 포함하는 서열, 또는 이의 상보적 서열로부터 유래된 15-50개의 뉴클레오티드를 포함하는 올리고뉴클레오티드는 칩에서 직접 합성하거나(Lockhart et al. supra), 또는 합성하고 이후 칩에 어드레스한다(Sosnowski et al., supra). 적절하게는, 올리고뉴클레오티드 대 략 20개의 뉴클레오티드이다.

적합한 화합물, 예를 들면 비오틴, 디곡시제닌 또는 형광 염료로 표지된 CanIon cDNA 프로브는 적합한 mRNA 개체군으로부터 합성하고, 이후 평균 50 내지 100개의 뉴클레오티드 크기로 무작위로 단편화시킨다. 그 다음, 상기 프로브는 칩에 혼성화시킨다. 세척(Lockhart et al., supra) 및 상이한 전기장의 적용(Sosnowsky et al., 1997)이후, 염료 또는 표지화 화합물은 검출하고 정량한다. 중복 혼성화를 실시하다. 상이한 cDNA 시료상의 동일한 표적 올리고뉴클레오티드에서 cDNA 프로브로부터 기원하는 신호의 비교 강도 분석은 CanIon mRNA의 상이한 발현을 암시한다.

CanIon 유전자의 발현을 저해하는 방법

본원 발명에 따른 다른 치료 조성물은 상응하는 CanIon 유전자의 발현을 저해하는 CanIon 뉴클레오티드 서열의 올리고뉴클레오티드 단편, 안티센스 또는 삼중 나선을 함유한다. CanIon 핵산 서열의 적합한 단편은 이중대립인자 마커 A1 내지 A17중 적어도 한가지의 대립유전자를 포함한다.

안티센스

본원 발명에 따른 안티센스 폴리뉴클레오티드를 이용하는데 적합한 방법은 Sczakiel et al.(1995)에서 밝힌 과정이다.

적절하게는, 안티센스는 CanIon mRNA의 5'말단에 상보적인 폴리뉴클레오티드(15-200bp)에서 선택된다. 다른 구체예에서, 원하는 표적 유전자의 상이한 부분에 상보적인 별개의 안티센스 폴리뉴클레오티드의 조합을 사용한다.

본원 발명에 따른 적절한 안티센스 폴리뉴클레오티드는 번역 개시 코돈 ATG 또는 절단접합 공여자나 수용자 부위를 보유하는 CanIon의 mRNA 서열에 상보적이다.

안티센스 핵산은 이중으로 CanIon mRNA의 발현을 저해하는데 충분한 안정성을 갖는 세포내 이중나선의 형성을 가능하게 하는 길이와 융점을 보유해야 한다. 유전자 요법에 사용하기 적합한 안티센스 핵산을 설계하는 전략은 Green et al., (1986); Izant and Weintraub, (1984)에서 개시한다.

일부 전략에서, 안티센스 분자는 프로모터와 관련하여 CanIon 코딩 영역의 배향을 역전시킴으로써 세포에서 정상적으로 전사되는 것과 반대의 가닥을 전사시켜 얻는다. 안티센스 분자는 시험관내 전사 시스템, 예를 들면 전사체를 얻기 위하여 T7 또는 SP6 중합효소를 이용하는 전사 시스템으로 전사시킬 수 있다. 다른 방식은 적합한 발현 벡터에서 안티센스 서열을 보유하는 DNA를 안티센스 서열에 동작가능하게 연결시켜 생체내에서 CanIon 안티센스 핵산을 전사시키는 것이다.

다른 적합한 안티센스 전략은 Rossi et al.(1991); 국제 특허 출원 WO 94/23026, WO 95/04141, WO 92/18522; 유럽 특허 출원 EP 0 572 287 A2에서 기술한다.

본원 발명에 활용되는 안티센스 기술에 대한 대안은 상보성 폴리뉴클레오티드를 통하여 표적 서열에 결합하고 표적 부위를 가수분해함으로써 상응하는 RNA를 절단하는 리보자임(즉, '해머헤드 리보자임')을 사용하는 것이다. 간단히 말하면, 해머헤드 리보자임의 사이클은 다음과 같이 구성된다: (1) 상보성 안티센스 서열을 통하여 표적 RNA에 서열 특이적 결합; (2) 표적 가닥의 절단 모티프의 부위-특이적 가수분해; (3) 절단 산물의 방출, 이는 다른 촉매 사이클을 유발한다. 실제로, 장쇄 안티센스 폴리뉴클레오티드(적어도 30개 염기) 또는 긴 안티센스 팔을 갖는 리보자임이 효과적이다. 안티센스 리보자임에 적합한 전달 시스템은 이들 안티센스 리보자임을 지질친화성기에 공유 결합시키거나 또는 리포좀을 벡터로 이용하여 달성한다. 본원 발명에 따른 적절한 안티센스 리보자임 Sczakiel et al.(1995)에서 밝힌 바와 같이 준비한다.

삼중 나선 방법

CanIon 게놈 DNA는 세포내 삼중 나선 형성에 기초하여 CanIon 유전자의 발현을 저해하는 데에도 사용할 수 있다.

삼중 나선 올리고뉴클레오티드는 게놈으로부터 전사를 저해하는데 사용된다. 이들은 특정 유전자와 관련하여 세포 활성에서 변화를 연구하는데 특히 유용하다.

유사하게, CanIon 게놈 DNA의 일부분은 세포내에서 CanIon 전사를 저해하는 효과를 연구하는데 사용할 수 있다. 전통적으로, 호모퓨린 서열이 삼중 나전 전략에 가장 적합한 것으로 생각되었다. 하지만, 호모피리미딘 서열 역시 유전자 발현을 저해할 수 있다. 이런 호모피리미딘 올리고뉴클레오티드는 호모퓨린 : 호모피리미딘 서열에서 대홈(major groove)에 결합한다. 따라서, CanIon 게놈 DNA로부터 유래된 양 유형의 서열은 본원 발명의 범주에 포함된다.

삼중 나선을 이용한 유전자 요법 전략을 실행하기 위하여, CanIon 게놈 DNA 서열은 먼저 스캔하여 CanIon 발현을 저해하는 삼중-나선 기초한 전략에 사용할 수 있는 10-mer 내지 20-mer 호모피리미딘 또는 호모퓨린 스트레치를 동정한다. 후보 호모피리미딘 또는 호모퓨린 스트레치의 동정이후, CanIon 발현을 저해하는 이들의 효능은 CanIon 유전자를 발현하는 조직 배양 세포에 후보 서열을 보유하는 다양한 함량의 올리고뉴클레오티드를 도입하여 평가한다.

상기 올리고뉴클레오티드는 당업자에게 공지된 다양한 방법, 예를 들면 칼슘 포스페이트 침전, DEAE-덱스트란, 에렉트로포레이션, 리포좀-매개된 트랜스펙션 또는 자연적인 흡수로 세포에 도입할 수 있다.

처리된 세포는 노던 블랏팅, RNase 보호 분석, 또는 올리고뉴클레오티드로 처리된 세포에서 CanIon 유전자의 전사 수준을 모니터하는 PCR 기초한 전략과 같은 기술을 활용하여 변화된 세포 기능 또는 감소된 CanIon 발현을 모니터한다.

이후, 조직 배양 세포에서 유전자 발현을 저해하는데 효과적인 올리고뉴클레오티드는 안티센스에서 밝힌 바와 같이 시험관내 결과에 기초하여 계산된 용량에서 전술한 기술을 활용하여 생체내 도입할 수 있다.

일부 구체예에서, 올리고뉴클레오티드 단위의 천연(베타) 아노머(anomer)는 알파 아노머로 치환하여 올리고뉴클레오티드가 뉴클레아제에 좀더 저항하도록 할 수 있다. 이에 더하여, 브롬화에티듐 등과 같은 삽입성 물질을 알파 올리고뉴클레오티드의 3'말단에 부착하여 삼중 나선을 안정화시킬 수 있다. 삼중 나선 형성에 적합한 올리고뉴클레오티드의 발생에 관한 정보는 Griffin et al.(1989)을 참조한다.

제약학적 조성물과 제형

CanIon-조절 화합물

본원에 개시된 방법을 활용하여, 시험관내와 생체내에서 CanIon 활성을 선별적으로 조절하는 CanIon 작용물질이나 길항물질을 동정할 수 있다. 따라서, 본원 발명에는 환자에서 정신분열증, 양극성 장애, 또는 다른 전술한 질환이나 이상을 치료하는 방법이 포함되는데, 상기 방법은 CanIon-조절 화합물의 효과량을 투여하는 단계로 구성된다. 적절하게는, 상기 화합물은 선택성 CanIon 조절 화합물이다. 본원 발명의 방법으로 동정된 화합물에는 사람 CanIon 폴리펩티드에 결합 특이성을 갖는 항체가 포함된다. 또한, CanIon의 유사체가 정신분열증 또는 양극성 장애와 연관된 CanIon-매개된 활성 및 관련된 생리 이상을 조절하는데 효과적일 것으로 기대된다. 일반적으로, 중추신경계 질환에서 CanIon의 역할에 기초한 본원 발명의 분석 방법은 본원 발명의 분석 캐스케이드에 개입할 수 있는 화합물을 동정하는데 활용될 수 있을 것으로 기대된다. 적절한 구체예에서, 정신분열증 또는 양극성 장애로 고생하는 환자는 CanIon 길항물질의 치료요법적 효과량을 함유하는 제약학적 조성물을 환자에 투여하여 치료한다.

징후

CanIon이 정신분열증 및 양극성 장애와 연관성을 보였다는 점에서, CanIon이 관련된 징후에는 중추신경계 질환이 포함될 수 있다. 신경계 질환은 복합적인 유전적 기초를 갖고 특정 증상을 공유할 것으로 예상된다. 특히, 본원에서 밝힌 바와 같이 징후에는 정신분열증과 다른 정신 질환, 신경퇴행성 질환, 기분 장애, 자폐증, 약물 의존과 알코올 중독, 간질, 통증 질환, 정신 지체 및 인식 장애, 정신 장애, 불안 장애, 섭식 장애, 충동 조절 장애, 인격 장애를 비롯한 다른 정신 질환이 포함될 수 있다[Diagnosis and Statistical Manual of Mental Disorders(DSM-IV)]. 이에 더하여, 협심증, 고혈압, 부정맥을 비롯한 다수의 심혈관 질환 역시 CanIon 조절물질, 바람직하게는 길항물질로 치료될 수 있다.

제약학적 제형 및 투여 경로

본원 발명의 방법을 이용하여 확인된 화합물은 정신분열증 또는 양극성 장애 관련된 질환을 치료 또는 완화하는 치료요법적 효과량으로 단독으로 또는 적합한 담체나 부형제와 혼합된 제약학적 조성물 형태로 사람 환자를 비롯한 포유동물에 투여할 수 있다. 치료요법적 효과량은 본원에서 밝힌 방법으로 증상의 완화를 결과하는데 충분한 화합물 함량을 의미한다. 적절하게는, 치료요법적 효과량은 지속적인 주기적 복용이나 투여에 적합하다. 본원 발명에 따른 화합물의 제형과 투여를 위한 기술은 'Remington's Pharmaceutical Sciences,'(Mack Publishing Co. Easton, PA) 최신판에서 알 수 있다.

투여 경로

적합한 투여 경로에는 지주막하, 뇌실내, 정맥내, 복강내, 비강내 또는 안구내 주사를 비롯하여 경구, 직장, 경점막 또는 장 투여; 근육내, 피하, 골수내 주사를 비롯한 장관외 전달이 포함된다. 중추신경계 질환을 치료하기 위한 약물을 투여하는데 특히 적합한 방법에는 이식된 펌프(Medtronic, Inc., Minneapolis, MN)를 통한 척수액으로 주입을 비롯한 지주막하 전달에서처럼 연장된 기간동안 화합물을 전달하는 장치의 외과적 이식이 포함된다.

조성물/제형

본원 발명에 사용되는 제약학적 조성물과 약물은 부형제와 보조제을 포함하는 하나 또는 복수의 생리학적으로 수용가능한 담체를 이용하여 통상적 방식으로 만들 수 있다.

주사용으로, 본원 발명의 약물은 수용액, 바람직하게는 생리학적 적합성 완충액(예, 한크 용액, 링거 용액) 또는 생리식염수 완충액(인산염 또는 중탄산염 완충액) 형태로 만들 수 있다. 경점막 투여를 위하여, 침투할 장벽에 적합한 침투탐상제(penetrant)가 제형으로 사용된다. 이런 침투탐상제는 당분야에 공지된 것이다.

경구로 복용할 수 있는 제약학적 제형에는 젤라틴으로 만들어진 푸시-핏 캡슐, 젤라틴으로 만들어진 연한 밀봉된 캡슐 및 가소제, 예를 들면 글리세롤이나 소비톨이 포함된다. 푸시-핏 캡슐은 락토오스와 같은 충진제, 전분과 같은 결합제 및/또는 활석이나 스테아르산마그네슘과 같은 윤활제, 선택적으로 안정제와 혼합된 활성 성분을 함유할 수 있다. 연한 캡슐에서, 활성 성분은 적합한 액체, 예를 들면 지방산, 액체 파라 혹은 액체 폴리에틸렌 글리콜에 용해 또는 부유시킬 수 있다. 이에 더하여, 안정제를 첨가할 수도 있다. 경구 투여를 위한 모든 제형은 이런 투여에 적합한 용량으로 제공한다.

협측 투여를 위하여, 조성물은 통상적 방식으로 만들어진 정제 또는 마름모꼴 정제 형태를 취할 수 있다.

흡입 투여를 위하여, 본원 발명에 따른 조성물은 적절한 가스 추진제, 예를 들면 이산화탄소를 이용하여 가압된 팩 또는 흡입치료기로부터 에어로졸 스프레이 형태로 전달된다. 흡입기 또는 취입기에 사용되는 예로써 젤라틴의 캡슐과 카트리지는 상기화합물 및 적합한 분말 염기, 예를 들면 락토오스 또는 전분의 분말 혼합물을 함유하도록 만들 수 있다.

화합물은 주사, 예를 들면 거환 주사 또는 연속 주입에 의한 장관외 투여용으로 만들 수 있다. 주사용 제형은 단위 약형, 예를 들면 방부제가 첨가된 앰플 또는 다중-용도 용기로 제공할 수 있다. 이런 조성물은 현탁액, 용액 또는 수성 소포에 용해된 에멸젼과 같은 형태를 취하고, 부유제, 안정제 및/또는 분산제와 같은 제형화 약품을 함유할 수 있다.

장관외 투여용 제약학적 제형에는 가수성 형태의 활성 화합물의 수용액이 포함된다. 수성 현탁액은 현탁액을 점도를 증가시키는 물질, 예를 들면 소디움 카복시메틸 셀룰로오스, 소비톨 또는 덱스트란을 함유할 수 있다. 선택적으로, 현탁액은 고도로 농축된 용액의 제조를 가능하게 하는 화합물의 용해도를 증가시키는 적합한 안정제 또는 약품을 함유할 수 있다.

대안으로, 활성 성분은 사용에 앞서 적합한 소포, 예를 들면 무-발열원 멸균수로 재구성되는 분말이나 동결건조된 형태일 수도 있다.

전술한 제형에 더하여, 화합물은 저장소 제형으로 만들 수도 있다. 이런 장기 제형은 이식(예, 피하 또는 근육내) 또는 근육내 주사로 투여할 수 있다. 따라서, 이런 화합물은 중합성이나 소수성 물질(예, 오일에 용해된 에멸젼으로) 혹은 이온 교환 수지와 혼합하여, 또는 거의 불용성의 유도체, 예를 들면 거의 불용성 염으로 만들 수 있다.

또한, 화합물은 지속-방출 시스템, 예를 들면 치료 약물을 함유하는 고형 소수성 중합체의 반투과성 매트릭스를 이용하여 전달할 수 있다. 다양한 지속 방출 물질은 당분야에 공지된 것이다. 지속-방출 캡슐은 화학적 특성에 따라, 수주 내지 최대 100일동안 화합물을 방출할 수 있다.

치료 약물의 화학적 특성과 생물학적 안정도에 따라, 단백질 안정화의 추가적인 전략을 강구할 수도 있다.

제약학적 조성물은 적합한 고체나 겔 상 담체 또는 부형제를 함유할 수도 있다. 이런 담체 또는 부형제의 예에는 탄산칼슘, 인산칼슘, 다양한 당, 전분, 셀룰로오스 유도체, 젤라틴, 폴리에틸렌 글리콜과 같은 중합체가 포함되지만 이들에 국한되지 않는다.

효과량

본원 발명에 사용하기 적합한 제약학적 조성물에는 소요의 목적을 달성하기 위하여 효과량의 활성 성분을 함유하는 조성물이 포함된다. 좀더 구체적으로, 치료요법적 효과량은 치료된 개체의 기존 증상의 발생을 예방하거나 또는 이를 완화시키는데 효과적인 함량을 의미한다. 효과량의 결정은 당업자의 능력에 속한다.

본원 발명의 방법에 사용된 화합물에서, 치료요법적 효과량은 세포 배양 분석으로부터 초기에 추정하고, 동물 모델에 1회 분량을 투여할 수 있다. 이런 정보는 사람에서 정확한 용량을 결정하는데 활용할 수 있다.

치료요법적 효과량은 환자에서 증상의 완화를 결과하는 화합물의 함량을 의 미한다. 이런 화합물의 독성과 치료 효능은 세포 배양액 또는 실험 동물에서 표준적인 제약학적 과정, 예를 들면 LD50(실험 개체군의 50%를 치사시키는 분량)과 ED50(개체군의 50%에 효과적인 분량)으로 결정할 수 있다. 독성과 치료 효능간 분량 비율은 치료지수로서, LD50과 ED50간 비율로 표시할 수 있다. 높은 치료 지수를 보이는 화합물이 바람직하다.

이들 세포 배양 분석 및 동물 연구로부터 얻은 데이터는 사람에서 사용되는 일정 범위의 용량을 확정하는데 활용할 수 있다. 적절하게는, 이런 화합물의 용량은 부작용이 거의 없으면서 ED50을 비롯한 순환 농도의 범위 이내가 된다. 이런 화합물의 용량은 이용된 약형 밑 투여 루트에 따라 상기 범위에서 변할 수 있다. 정확한 제형, 투여 경로, 용량은 환자의 조건에 비추어 개별 의사가 선택할 수 있다(Fingl et al., 1975, in 'The Pharmacological Basis of Therapeutics', Ch.1)

컴퓨터-관련된 구체예

'본원 발명의 핵산 코드'에는 a) SEQ ID No 1 내지 3의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000, 2000, 3000, 4000, 5000 또는 10000개 뉴클레오티드의 연속 스팬; b) SEQ ID No 4 또는 이의 보체의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000, 2000, 3000, 4000, 5000 또는 10000개 뉴클레오티드의 연속 스팬; c) SEQ ID No 1 내지 3의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000, 2000, 3000, 4000, 5000 또는 10000개 뉴클레오티드의 연속 스팬, 여기서 상기 연속 스팬은 A1 내지 A17에서 선택되는 이중대립 인자 마커를 포함한다; d) SEQ ID No 6의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300 또는 400개 뉴클레오티드의 연속 스팬; e) SEQ ID No 6의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300 또는 400개 뉴클레오티드의 연속 스팬, 여기서 상기 연속 스팬은 이중대립인자 마커 A18을 포함한다; f) 전술한 뉴클레오티드 서열중 한가지에 상보적인 뉴클레오티드 서열.

'본원 발명의 핵산 코드'에는 또한, a) SEQ ID No 1 내지 3의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000, 2000, 3000, 4000, 5000 또는 10000개 뉴클레오티드의 연속 스팬; b) SEQ ID No 4의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, 1000, 2000, 3000, 4000, 5000 또는 10000개 뉴클레오티드의 연속 스팬; c) SEQ ID No 6의 적어도 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300 또는 400개 뉴클레오티드의 연속 스팬; f) 전술한 뉴클레오티드 서열중 한가지에 상보적인 뉴클레오티드 서열에 상보적인 뉴클레오티드 서열이 포함된다. 상동성 서열은 이들 연속 스팬과 적어도 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80% 또는 75% 상동성을 보유하는 서열을 의미한다. 상동성은 디폴트 파라미터 또는 변형된 파라미터를 갖는 BLASTN을 비롯한 본원에서 밝힌 임의의 방법으로 측정할 수 있다. 상동성 서열에는 또한, 본원 발명의 핵산 코드에서 우리딘이 티미딘을 대체하는 RNA 서열이 포함된다. 본원 발명의 핵산 코드는 전통적인 단일 문자 형식(Stryer, Lubert. Biochemistry, 3 rd edition), 또는 서열상의 뉴클레오티드를 기록하는 임의의 다른 형식이나 코드로 나타낼 수 있다.

'본원 발명의 폴리펩티드 코드'에는 SEQ ID No 5의 적어도 6, 8, 10, 12, 15, 20, 25, 30, 40, 50, 100, 150, 200, 300, 400, 500, 700, 1000, 1200, 1400, 1600 또는 1700개 아미노산의 연속 스팬이 포함된다. 적절한 구체예에서, 상기 연속 스팬은 SEQ ID No 5의 아미노산 위치 277, 338, 574, 678, 680, 683, 691, 692, 695, 696, 894, 1480, 1481, 1483, 1484, 1485, 1630, 1631, 1632, 1636, 1660, 1667, 1707, 1709중 적어도 1, 2, 3, 5 또는 10개를 보유한다. 본원 발명의 폴리펩티드 코드는 전통적인 단일 문자 형식 또는 3문자 형식(Stryer, Lubert. Biochemistry, 3 rd edition), 또는 서열상의 폴리펩티드를 기록하는 임의의 다른 형식이나 코드로 나타낼 수 있다.

당업자가 인지하는 바와 같이, 본원 발명의 핵산 코드와 폴리펩티드 코드는 컴퓨터로 판독하고 접근가능한 임의의 매체에서 저장, 기록, 처리할 수 있다. 본원에서, '기록'과 '저장'은 컴퓨터 매체에 정보를 저장하는 과정을 의미한다. 당업자는 본원 발명의 한가지이상 핵산 또는 한가지이상 폴리펩티드를 포함하는 제품을 산출하기 위하여 컴퓨터 판독가능 매체에 정보를 기록하는 공지된 방법중 한가지를 용이하게 채택할 수 있다. 본원 발명의 다른 측면은 본원 발명의 적어도 2, 5, 10, 15, 20, 25, 30 또는 50개의 핵산 코드가 기록된 컴퓨터 판독가능 매체이다. 본원 발명의 다른 측면은 본원 발명의 적어도 2, 5, 10, 15, 20, 25, 30 또는 50개의 폴리펩티드 코드가 기록된 컴퓨터 판독가능 매체이다.

컴퓨터 판독가능 매체에는 자성 판독가능 매체, 광학 판독가능 매체, 전기 판독가능 매체, 자성/광학 매체가 포함된다. 가령, 컴퓨터 판독가능 매체는 하드디스크, 플로피디스크, 자성 테이프, CD-ROM, 디지털 다기능 디스크(DVD), Random Access Memory(RAM), Read Only Memory(ROM) 및 당업자에게 공지된 다른 유형의 매체이다.

본원 발명의 구체예에는 본원에서 밝힌 서열 정보를 저장하고 처리하는 시스템, 특히 컴퓨터 시스템이 포함된다. 컴퓨터 시스템(100)의 한가지 예는 도 2의 블록 다이어그램에 도해한다. 본원에서 '컴퓨터 시스템'은 본원 발명에 따른 핵산 코드의 뉴클레오티드 서열 또는 본원 발명에 따른 폴리펩티드 코드의 아미노산 서열을 분석하는데 사용되는 하

드웨어 성분, 소프트웨어 성분, 데이터 저장 성분을 의미한다. 한 구체예에서, 컴퓨터 시스템(100)은 Sun Enterprise 1000 서버(Sun Microsystems, Palo Alto, CA)이다. 적절하게는, 컴퓨터 시스템(100)은 서열 데이터의 처리, 접근, 조작을 위한 프로세스를 보유한다. 프로세스(105)는 임의의 공지된 중앙 처리 장치, 예를 들면 펜티엄 III(Intel Corporation), 또는 유사한 프로세스(Sun, Motorola, Compaq 또는 International Business Machines)일 수 있다.

적절하게는, 컴퓨터 시스템(100)은 프로세스(105), 데이터를 저장하는 하나이상의 내부 데이터 저장 성분(110), 데이터 저장 성분에 저장된 데이터를 회수하는 하나이상의 데이터 회수 장치로 구성되는 일반 목적의 시스템이다. 현재 가용한 컴퓨터 시스템중 한가지가 적합하다.

특정 구체예에서, 컴퓨터 시스템(100)은 주기억장치(115)(바람직하게는, RAM)에 연결된 버스에 연결되는 프로세스(105) 및 하나이상의 내부 데이터 저장 장치(110), 예를 들면 하드드라이브 및/또는 데이터를 기록하는 컴퓨터 판독가능 매체를 포함한다. 일부 구체예에서, 컴퓨터 시스템(100)은 데이터 저장 장치(110)에 저장된 데이터를 읽어들이는 하나이상의 데이터 회수 장치(118)를 추가로 포함한다.

데이터 회수 장치(118)는 예로써 플로피디스크, 컴팩 디스크 드라이브, 자성 테이프 드라이브 등일 수 있다. 일부 구체예에서, 내부 데이터 저장 장치(110)는 컨트롤 로직(control logic) 및/또는 여기에 기록된 데이터를 보유하는 플로피디스크, 컴팩 디스크, 자성 테이프 등과 같은 이동가능한 컴퓨터 판독가능 매체이다. 컴퓨터 시스템(100)은 데이터 회수 장치에 삽입되면 데이터 저장 성분으로부터 컨트롤 로직 및/또는 데이터를 읽어들이는 적합한 소프트웨어에 의해 프로그램될 수 있다.

컴퓨터 시스템(100)은 결과를 컴퓨터 사용자에게 전시하는데 사용되는 출력장치(120)를 포함한다. 컴퓨터 시스템(100)은 네트워크 또는 광역 네트워크에서 다른 컴퓨터 시스템(125a-c)과 연결하여 컴퓨터 시스템(100)에 집중된 접근을 제공할 수 있다.

본원 발명에 따른 핵산 코드의 뉴클레오티드 서열 또는 본원 발명에 따른 폴리펩티드 코드의 아미노산 서열에 접근하고 이를 처리하는 소프트웨어(예, 검색 도구, 비교 도구, 모델링 도구 등)는 실행동안 주메모리(115)에 존재할 수 있다.

일부 구체예에서, 컴퓨터 시스템(100)은 컴퓨터 판독가능 매체에 저장된 참고 뉴클레오티드 또는 폴리펩티드 서열과 컴퓨터 판독가능 매체에 저장된 본원 발명의 전술한 핵산 코드 또는 폴리펩티드 코드를 비교하는 서열비교기를 추가로 포함할 수 있다. '서열 비교기'는 뉴클레오티드 또는 폴리펩티드 서열과 데이터 저장 수단내에 저장된 펩티드, 펩티드모사체(peptidomimetics), 화학약물을 포함하지만 이들에 국한되지 않는 다른 뉴클레오티드나 폴리펩티드 서열 및/또는 화합물을 비교하는 컴퓨터 시스템(100)에서 실행되는 한가지이상 프로그램을 의미한다. 가령, 서열 비교기는 컴퓨터 판독가능 매체에 저장된 본원 발명에 따른 핵산 코드의 뉴클레오티드 서열 또는 본원 발명에 따른 폴리펩티드 코드의 아미노산 서열과 컴퓨터 판독가능 매체에 저장된 참고 서열을 비교하여 상동성, 생물학적 기능에 관여하는 모티프, 또는 구조적 모티프를 동정할 수 있다. 본원 발명의 다양한 서열 비교 프로그램은 본원 발명에 구체적으로 정관한다.

도 3은 새로운 뉴클레오티드 또는 단백질과 데이터베이스상의 서열을 비교하여 새로운 뉴클레오티드와 데이터베이스상의 서열간 상동성 수준을 측정하는 프로세스(200)의 구체예를 도해하는 흐름도이다. 서열의 데이터베이스는 컴퓨터 시스템(100)에 저장된 개인 데이터베이스, 또는 인터넷을 통하여 입수가능한 GENBANK, PIR 또는 SWISSPROT와 같은 공공 데이터베이스일 수 있다.

프로세스(200)는 출발 단계(201)에 시작하여 단계(202)로 진행하는데, 여기서 비교되는 새로운 서열은 컴퓨터 시스템(100)의 기억장치에 저장된다. 전술한 바와 같이, 기억장치는 RAM 또는 내부 저장 장치를 비롯한 임의 유형의 기억장치일 수 있다.

이후, 프로세스(200)는 단계(204)로 진행되는데, 여기서 데이터베이스에 저장된 첫 번째 서열이 컴퓨터의 기억장치로 판독된다. 그 다음, 단계(210)에서 비교가 실행되어 첫 번째 서열과 두 번째 서열이 동일한 지를 확인한다. 동일하지 않은 경우에도 2개의 뉴클레오티드 또는 단백질 서열을 비교하는 방법은 당업자에게 공지된 것이다. 가령, 2개의 검사 서열간 상동성 수준을 높이기 위하여 한 서열에 공백을 도입할 수 있다. 공백이나 다른 지형(feature)이 비교동안 서열에 도입되는 지를 조절하는 파라미터는 통상적으로, 컴퓨터 시스템의 사용자에 의해 입력된다.

두 서열의 비교가 단계(210)에서 실행되면, 결정 단계(210)에서 두 서열이 동일한 지에 대한 결정이 이루어진다. 물론, '동일'이라는 의미는 절대적으로 동일한 서열에 한정되지 않는다. 사용자에 의해 입력된 상동성 파라미터이내의 서열은 프로세스(200)에서 '동일'한 것으로 표시된다.

두 서열이 동일하다는 결정이 내려지면, 프로세스(200) 단계(214)로 진행하는데, 여기서 데이터베이스로부터 서열의

이름이 사용자에게 전시된다. 이 단계는 사용자에게 전시된 이름을 갖는 서열이 입력된 상동성 제한을 충족한다는 것을 알려준다. 저장된 서열의 이름이 사용자에게 전시되면, 프로세스(200)는 결정 단계(218)로 진행하는데, 여기서 더 많은 서열이 데이터베이스에 존재하는 지에 대한 결정이 이루어진다. 더 이상의 서열이 데이터베이스에 존재하지 않으면, 프로세스(200)는 최종 단계(220)에서 종결된다. 하지만, 더 많은 서열이 데이터베이스 에 존재하면, 프로세스(200)는 단계(224)로 진행하고, 여기서 포인터가 데이터베이스상의 다음 서열로 이동하여 이를 상기 서열과 비교한다. 이런 방식으로, 새로운 서열은 정렬되고 데이터베이스상의 모든 서열과 비교된다.

결정 단계(212)에서 서열이 상동하지 않다는 결정이 내려지면, 프로세스(200)는 결정단계(218)로 즉시 진행하여 비교 데이터베이스에서 임의의 다른 서열이 가용한 지를 확인한다.

따라서, 본원 발명의 한 측면은 프로세스, 본원 발명의 핵산 코드 또는 폴리펩티드 코드가 저장된 데이터 저장 장치, 본원 발명의 핵산 코드 또는 폴리펩티드 코드와 비교되는 참고 뉴클레오티드 서열이나 폴리펩티드 서열이 회수가능하게 저장된 데이터 저장 장치, 비교를 실행하는 서열 비교기를 포함하는 컴퓨터 시스템이다. 서열 비교기는 비교된 서열간 상동성 수준을 지시하거나, 본원 발명의 핵산 코드와 폴리펩티드에서 생물학적 기능과 구조 모티프에 관여하는 모티프를 동정하거나, 또는 이들 핵산 코드와 폴리펩티드 코드와 비교된 서열에서 구조적 모티프를 동정할 수 있다. 일부 구체예에서, 데이터 저장 장치는 본원 발명의 핵산 코드 또는 폴리펩티드 코드중 적어도 2, 5, 10, 15, 20, 25, 30 또는 50개의 서열을 저장할 수 있다.

본원 발명의 다른 측면은 본원 발명의 핵산 코드와 참고 뉴클레오티드 서열간 상동성 수준을 측정하는 방법인데, 상기 방법은 상동성 수준을 측정하는 컴퓨터 프로그램의 이용을 통하여 핵산 코드와 참고 뉴클레오티드 서열을 판독하고 상기 컴퓨터 프로그램으로 핵산 코드와 참고 뉴클레오티드 서열간 상동성을 측정하는 단 계로 구성된다. 컴퓨터 프로그램은 디폴트 파라미터 또는 변형된 파라미터를 갖는 BLAST2N과 같은 본원에 명시된 프로그램을 비롯하여 상동성 수준을 측정하는 다수의 컴퓨터 프로그램중한가지일 수 있다. 상기 방법은 전술한 컴퓨터 시스템을 이용하여 실행할 수 있다. 또한, 상기 방법은 컴퓨터 프로그램의 이용을 통하여 본원 발명의 전술한 핵산 코드중 적어도 2, 5, 10, 15, 20, 25, 30 또는 50개를 판독하고 핵산 코드와 참고 뉴클레오티드 서열간 상동성을 측정하여 실행할 수 있다.

도 4는 두 서열이 상동한 지를 확인하기 위한 프로세스(250)의 구체예를 예시하는 흐름도이다. 프로세스(250)는 출발 단계(252)에서 시작하여 단계(254)로 진행하고, 여기서 비교되는 첫 번째 서열은 기억장치에 저장된다. 이후, 비교되는 두 번째 서열이 단계(256)에서 기억장치에 저장된다. 그 다음, 프로세스(250)는 단계(260)로 진행하고, 여기서 첫 번째 서열에서 첫 번째 문자가 판독되고, 이후 단계(262)로 진행하고, 여기서 두 번째 서열에서 첫 번째 문자가 판독된다. 서열이 뉴클레오티드 서열이면 문자는 A, T, C, G(또는 U)가 된다. 서열이 단백질 서열이면 이는 단일 문자 아미노산 코드가 되고, 따라서 첫 번째 서열과 두 번째 서열은 용이하게 비교할 수 있다.

결정 단계(264)에서 두 문자가 동일한 지에 대한 결정이 이루어진다. 이들이 동일하면 프로세스(250)는 단계(268)로 진행하는데, 여기서 첫 번째 서열과 두 번째 서열에서 다음 문자가 판독된다. 다음 문자가 동일한 지에 대한 결정이 이루어진다. 이들이 동일하면 프로세스(250)는 두 문자가 동일하지 않을 때까지 이런 순환(loop)을 지속한다. 다음의 두 문자가 동일하지 않다는 결정이 내려지면, 프 로세스(250)는 결정 단계(274)로 진행하고 판독할 더 많은 문자 또는 서열이 존재하는 지를 확인한다.

판독할 문자가 더 이상 존재하지 않으면 프로세스(250)는 단계(276)로 진행하는데, 여기서 첫 번째 서열과 두 번째 서열간 상동성 수준이 사용자에게 전시된다. 상동성 수준은 첫 번째 서열에서 전체 서열중 동일한 서열간 문자의 비율을 계산하여 측정한다. 따라서, 첫 100개의 뉴클레오티드 서열에서 모든 문자가 두 번째 서열에서 모든 문자와 정렬되면, 상동성 수준은 100%가 된다.

대안으로, 컴퓨터 프로그램은 본원 발명에 따른 핵산 코드의 뉴클레오티드 서열과 참고 뉴클레오티드 서열을 비교하여 하나이상의 위치에서 본원 발명의 핵산 코드가 참고핵산 서열과 구별되는 지를 확인하는 컴퓨터 프로그램일 수 있다. 선택적으로, 이런 프로그램은 참고 폴리뉴클레오티드의 서열 또는 본원 발명의 핵산 코드에 대하여 삽입된, 결실된 또는 치환된 뉴클레오티드의 길이와 실체를 기록한다. 한 구체예에서, 컴퓨터 프로그램은 본원 발명에 따른 핵산 코드의 뉴클레오티드 서열이 참고 뉴클레오티드 서열과 비교하여 한가지이상의 단일 뉴클레오티드 다형성(SNP)을 보유하는 지를 확인하는 프로그램일 수 있다. 이들 단일 뉴클레오티드 다형성은 각각 단일 염기 치환, 삽입 또는 결실을 포함할 수 있다.

본원 발명의 다른 측면은 본원 발명의 폴리펩티드 코드와 참고 폴리펩티드 서열간 상동성 수준을 측정하는 방법인데, 상기 방법은 상동성 수준을 측정하는 컴퓨터 프로그램의 이용을 통하여 본원 발명의 폴리펩티드 코드와 참고 폴리펩티드 서열을 판독하고 상기 컴퓨터 프로그램으로 폴리펩티드 코드와 참고 폴리펩티드 서 열간 상동성을 측정하는 단계로 구성된다.

따라서, 본원 발명의 다른 측면은 본원 발명의 핵산 코드가 한 개이상의 뉴클레오티드에서 참고 뉴클레오티드 서열과

구별되는 지를 확인하는 방법인데, 상기 방법은 핵산 서열간 차이를 확인하는 컴퓨터 프로그램의 이용을 통하여 핵산 코드와 참고 뉴클레오티드 서열을 판독하고 상기 컴퓨터 프로그램으로 핵산 코드와 참고 뉴클레오티드 서열간 차이를 확인하는 단계로 구성된다. 일부 구체예에서, 컴퓨터 프로그램은 단일 뉴클레오티드 다형성을 동정하는 프로그램이다. 상기 방법은 전술한 컴퓨터 시스템 및 도 4에 예시된 방법으로 실행할 수 있다. 또한, 상기 방법은 컴퓨터 프로그램의 이용을 통하여 본원 발명의 전술한 핵산 코드중 적어도 2, 5, 10, 15, 20, 25, 30 또는 50개 및 참고 뉴클레오티드 서열을 판독하고 상기 컴퓨터 프로그램으로 핵산 코드와 참고 뉴클레오티드 서열간 차이를 동정하여 실행할 수 있다.

다른 구체예에서, 컴퓨터 기초한 시스템은 본원 발명에 따른 핵산 코드의 뉴클레오티드 서열 또는 본원 발명에 따른 폴리펩티드 코드의 아미노산 서열내에서 지형(feature)를 동정하는 식별기(identifier)를 추가로 포함한다.

'식별기'는 본원 발명에 따른 핵산 코드의 뉴클레오티드 서열 또는 본원 발명에 따른 폴리펩티드 코드의 아미노산 서열내에서 지형을 동정하는 한가지이상의 프로그램을 의미한다. 한 구체예에서, 식별기는 본원 발명의 cDNA 코드에서 개방해독틀을 동정하는 프로그램을 포함할 수 있다.

도 5는 서열에서 지형의 존재를 검출하기 위한 식별기 프로세스(300)의 구체 예를 도해하는 흐름도이다. 프로세스(300)는 출발 단계(300)에서 시작하고 단계(304)로 진행되는데, 여기서 지형을 조사받은 첫 번째 서열은 컴퓨터 시스템(100)에서 기억장치(115)에 저장된다. 이후, 프로세스(300)는 단계(306)로 진행하는데, 여기서 서열 지형의 데이터베이스가 개방된다. 이런 데이터베이스는 각 지형의 이름과 속성의 목록을 포함한다. 가령, 지형 이름은 '개시 코돈'이고 속성은 'ATG'일 수 있다. 다른 예에서 지형 이름은 'TAATAA Box'이고 속성은 'TAATAA'이다. 이런 데이터베이스의 예는 University of Winsconsin Genetics Computer Group(www.gcg.com)에 의해 만들어진다.

지형의 데이터베이스가 단계(306)에서 개방되면 프로세스(300)는 단계(308)로 진행하는데, 여기서 첫 번째 지형이 데이터베이스로부터 판독된다. 이후, 첫 번째 지형의 속성과 첫 번째 서열의 비교는 단계(310)에서 진행된다. 그 다음, 결정 단계(316)에서 지형의 속성이 첫 번째 서열에서 발견되는 지에 대한 결정이 이루어진다. 속성이 발견되면 프로세스(300)는 단계(318)로 진행되는데, 여기서 상기 지형의 이름이 사용자에게 전시된다.

이후, 프로세스(300)는 결정 단계(320)로 진행하는데, 여기서 이동 지형(move feature)이 데이터베이스에 존재하는 지에 대한 결정이 이루어진다. 더 이상의 지형이 존재하지 않으면 프로세스(300)는 최종 단계(324)에서 종결된다. 하지만, 데이터베이스에 더 많은 지형이 존재하면 프로세스(300)는 단계(236)에서 다음 서열 지형을 판독하고 단계(310)로 복귀하는데, 여기서 다음 지형의 속성이 첫 번째 서열과 비교된다.

결정 단계(316)에서 지형 속성이 첫 번째 서열에서 발견되지 않으면 프로세스(300)는 결정 단계(320)로 직접 진행되고 데이터베이스에 더 많은 지형이 존재하는 지를 확인한다.

다른 구체예에서, 식별기는 본원 발명에 따른 폴리펩티드 코드의 3차원 구조를 측정하는 분자 모델링 프로그램을 포함할 수 있다. 일부 구체예에서, 분자 모델링 프로그램은 공지된 3차원 단백질 구조에서 잔기의 구조적 환경을 나타내는 프로필과 가장 일치하는 표적서열을 동정한다(미국 특허 5,436,850). 다른 기술에서, 일정한 가계에서 단백질의 공지된 3차원 구조는 중첩되고, 따라서 상기 가계에서 구조적으로 보존된 영역을 정의한다. 또한, 이런 단백질 모델링 기술에서는 본원 발명에 따른 폴리펩티드 코드의 구조를 근사하는 상동성 단백질의 공지된 3차원 구조를 이용한다(미국 특허 5,557,535). 전통적인 상동성 모델링 기술은 통상적으로, 프로테아제와 항체의 모델을 확립하는데 사용되고 있다(Sowdhamini et al.,(1997)). 또한, 관심있는 단백질이 주형 단백질과 빈약한 서열 상동성을 갖는 경우에는 비교 방식으로 3차원 단백질 모델을 개발할 수 있다. 일부 경우에, 단백질은 매우 약한 서열 상동성에도 불구하고 유사한 3차원 구조로 접힌다. 가령, 많은 나선 사이토킨의 3차원 구조는 약한 서열 상동성에도 불구하고 유사한 3차원 형상(topology)으로 접힌다.

최근 스레딩(threading) 방법의 개발로 표적과 주형간 구조적 관계성(relatedness)이 서열 수준에서는 전혀 검출되지 않는 많은 경우에 유사한 접힙(folding) 패턴의 동정이 가능해졌다. Multiple Sequence Threading(MST)으로 폴드(fold) 인식을 실시하는 하이브리드 방법에서 구조적 등가는 낮은 분석능 모델을 구성하는 거리 기하(distance geometry) 프로그램 DRAGON을 이용하여 스레딩 출력으로 추론하고, 전체-원자 표현은 분자 모델링 패키지, 예를 들면 QUANTA를 이용하여 구성한다.

3단계 방식에 따라, 후보 주형은 먼저 새로운 폴드 인식 프로그램 MST로 동정하는데, 상기 프로그램은 하나이상의 3-D 구조에서 다중 정렬된 서열의 동시 스레딩을 실행할 수 있다. 두 번째 단계에서, MST 출력으로부터 얻은 구조적 등가는 잔기간 거리 제한(interresidue distance restraint)으로 전환시키고 이차 구조 예측으로부터 얻은 부가적인 정보와 함께 거리 기하 프로그램 DRAGON에 넣는다. 상기 프로그램은 공평한 방식으로 제한을 통합하고 다수의 저 분석능 모델 확증을 산출한다. 세 번째 단계에서, 이들 저 분석능 모델 확증은 전체-원자 모델로 전환시키고 분자 모델링 패키지 QUANTA를 이용하여 에너지 최소화시킨다(Aszodi et al.,(1997)).

이후, 분자 모델링 분석의 결과는 본원 발명에 따른 폴리펩티드 코드의 활성을 조절하는 작용제를 동정하는 합리적 약물 설계 기술에 활용할 수 있다.

따라서, 본원 발명의 다른 측면은 본원 발명의 핵산 코드 또는 폴리펩티드 코드내에서 지형(feature)을 동정하는 방법인데, 상기 방법은 지형을 동정하는 컴퓨터 프로그램의 이용을 통하여 핵산 코드 또는 폴리펩티드 코드를 판독하고 상기 컴퓨터 프로그램으로 핵산 코드와 폴리펩티드 코드내에서 지형을 동정하는 단계로 구성된다. 한 구체예에서, 컴퓨터 프로그램은 개방해독틀을 동정하는 컴퓨터 프로 그램이다. 다른 구체예에서, 컴퓨터 프로그램은 폴리펩티드 서열에서 구조 모티프를 동정한다. 또 다른 구체예에서, 컴퓨터 프로그램은 분자 모델링 프로그램이다. 상기 방법은 컴퓨터 프로그램의 이용을 통하여 단일 서열, 또는 본원 발명의 전술한 핵산 코드 또는 폴리펩티드 코드중 적어도 2, 5, 10, 15, 20, 25, 30, 50개를 판독하고 상기 컴퓨터 프로그램으로 핵산 코드와 폴리펩티드 코드내에서 지형을 동정하여 실행할 수 있다.

본원 발명의 핵산 코드 또는 폴리펩티드 코드는 다양한 데이터 프로세서 프로그램에서 다양한 포맷으로 저장 및 조작할 수 있다. 가령, 이들은 워드 프로세서, 예를 들면 마이크로소프트 워드나 워드퍼펙트에서 문서 파일로, 또는 당업자에게 익숙한 다양한 프로그램, 예를 들면 DB2, SYBASE 혹은 ORACLE에서 ASCII 파일로 저장될 수 있다. 이에 더하여, 많은 컴퓨터 프로그램과 데이터베이스가 서열 비교기, 식별기, 또는 본원 발명에 따른 핵산 코드나 폴리펩티드 코드와 비교되는 참고 뉴클레오티드나 폴리펩티드 서열의 출처로 이용될 수 있다. 다음의 리스트는 본원 발명에 따른 핵산 코드 또는 폴리펩티드 코드에 유용한 프로그램과 데이터베이스에 대한 안내를 제공하지만 본원 발명을 제한하지 않는다: MacPattern(EMBL), DiscoveryBase(Molecular Applications Group), GeneMine(Molecular Applications Group), Look(Molecular Applications Group), MacLook(Molecular Applications Group), BLAST와 BLAST2(NCBI), BLASTN와 BLASTX(Altschul et al, 1990), FASTA(Pearson and Lipman, 1988), FASTDB(Brutlag et al., 1990), Catalyst(Molecular Simulations Inc.), Catalyst/SHAPE(Molecular Simulations Inc.), Cerius ^{2} .DBAccess(Molecular Simulations Inc.), HypoGen(Molecular Simulations Inc.), Insight II(Molecular Simulations Inc.), Discover(Molecular Simulations Inc.), CHARMm(Molecular Simulations Inc.), Felix(Molecular Simulations Inc.), DelPhi(Molecular Simulations Inc.), QuanteMM(Molecular Simulations Inc.), Homology(Molecular Simulations Inc.), Modeler(Molecular Simulations Inc.), ISIS(Molecular Simulations Inc.), Quanta/Protein Design(Molecular Simulations Inc.), WebLab(Molecular Simulations Inc.), WebLab Diversity Explorer(Molecular Simulations Inc.), Gene Explorer(Molecular Simulations Inc.), SeqFold(Molecular Simulations Inc.), the EMBL/Swissprotein 데이터베이스, MDL Available Chemicals Directory 데이터베이스, MDL Drug Data Report 데이터베이스, Comprehensive Medicinal Chemistry 데이터베이스, Derwents's World Drug Index 데이터베이스, BioByteMasterFile 데이터베이스, Genbank 데이터베이스, Genseqn 데이터베이스. 본원의 개시를 고려하면, 많은 다른 프로그램과 데이터베이스는 당업자에게 자명하다.

상기 프로그램으로 검출할 수 있는 모티프에는 루이신 지퍼(zoppers), 나선-회전-나선 모티프, 당화 부위, 유비퀴틴화 부위, 알파 나선, 베타 시트를 인코드하는 서열; 인코드된 단백질의 분비를 지시하는 신호 펩티드를 인코드하는 신호 서열; 호메오박스(homeobox), 산성 스트레치(acidic stretch), 효소 활성 부위, 기질 결합 부위, 효소 절단 부위와 같은 전사 조절에 관여하는 서열이 포함된다.

본원에서, 다양한 간행물, 문헌, 공개된 특허 출원이 언급되었다. 이들 간행물, 특허, 공개된 특허 출원은 본원에 순전히 참고로 한다.

실시예

실시예 1: 이중대립인자 마커의 동정 - DNA 추출

공여자는 무관하고 건강하였다. 이들은 프랑스의 이질성 개체군을 대표할 만큼 충분한 다양성을 제공하였다. 100명의 개체로부터 DNA를 추출하고 이중대립인자 마커의 검출을 검사하였다.

30 ㎖ 말초 정맥혈은 EDTA의 존재하에 각 공여자로부터 채취한다. 세포(펠렛)는 2000 rpm에서 10분동안 원심분리후 수거한다. 적혈구 세포는 세포용해 용액(50 ㎖ 최종 부피: 10 mM Tris pH 7.6; 5 mM $MgCl_2$; 10 mM NaCl)으로 용해시킨다. 세포용해 용액에 펠렛의 재부유후, 상기 용액은 상층액에 존재하는 잔류 적혈구를 제거하는데 필요한 회수로 원심분리(10 분, 2000 rpm)한다.

백혈구의 펠렛은 다음과 같이 구성되는 3.7 ㎖ 세포용해 용액으로 42℃에서 하룻밤동안 용해시킨다:

- 3 ㎖ TE 10-2(Tris-HCl 10 mM, EDTA 2 mM)/NaCl 0.4 M

- 200 ㎕ SDS 10%

- 500 ㎕ K-단백질분해효소(2 mg K-단백질분해효소/TE 10-2/NaCl 0.4 M)

단백질 추출을 위하여, 1 ㎖ 포화된 NaCl(6M)(1/3.5 v/v)를 첨가한다. 활발한 교반후, 용액은 10000 rpm에서 20분동안 원심분리한다.

DNA의 침전을 위하여, 2 내지 부피의 100% 에탄올을 상기 상층액에 첨가하고, 용액은 2000 rpm에서 30분동안 원심분리한다. DNA 용액은 70% 에탄올로 3회 세척하여 염을 제거하고 2000 rpm에서 20분동안 원심분리한다. 펠렛은 37℃에서 건조시키고 1 ㎖ TE 10-1 또는 1 ㎖ 물에 재부유시킨다. DNA 농도는 260 nm에서 OD를 측정하여 평가한다(1 단위 OD = 50 ㎍/㎖ DNA).

DNA 용액에서 단백질의 존재를 확인하기 위하여, OD 260/OD 280 비율을 측정한다. 1.8 내지 2의 OD 260/OD 280 비율을 갖는 DNA 제형만 하기 실시예에 사용한다.

풀(pooi)은 각 개체로부터 얻은 DNA의 등량을 혼합하여 구성한다.

실시예 2: 이중대립인자 마커의 동정: PCR에 의한 게놈 DNA의 증폭

실시예 1의 DNA 시료에서 특정 게놈 서열의 증폭은 앞서 수득된 DNA 풀(pool)에서 실시하였다. 이에 더하여, 50개의 개별 시료를 유사하게 증폭하였다.

PCR 분석은 다음의 프로토콜을 이용하여 실시한다:

최종 부피 25㎕

DNA 2 ng/㎕

MgCl2 2 mM

dNTP(각각) 200 μM

프라이머(각각) 2.9 ng/㎕

Ampli Taq Gold DNA 폴리머라제 0.05 unit/㎕

PCR 완충액(10x = 0.1 M TrisHCl pH 8.3 0.5M KCl) 1x

첫 번째 프라이머의 각 쌍은 본원에 개시된 CanIon 유전자의 서열 정보 및 OSP 소프트웨어(Hilier amp; Green, 1991)를 이용하여 설계되었다. 프라이머의 첫 번째 쌍은 길이가 대략 20개 뉴클레오티드이고 표 1에서 PU와 RP로 분류된 칼럼에 개시된 서열을 갖는다.

표 1:

앰플리콘	SEQ ID 1에서 앰플리콘의 위치 범위		프라이머 이름	SEQ ID 1에서 앰플리콘의 위치 범위		프라이머 이름	상보적인 위치 range of amplification 프라이머 in SEQ ID No 1	
99-62626	12343	12810	B1	12343	12363	C1	12793	12810
99-62632	13814	14296	B2	13814	13832	C2	14279	14296
99-62633	24863	25396	B3	24863	24881	C3	25378	25396
99-62611	69198	69650	B4	69198	69218	C4	69632	69650
99-62605	73005	73483	B5	73005	73022	C5	73466	73483
99-62635	79808	80334	B6	79808	79826	C6	80314	80334
앰플리콘	SEQ ID 2에서 앰플리콘의 위치 범위		프라이머 이름	SEQ ID 2에서 증 프라이머의 위치 범위		프라이머 이름	SEQ ID 2에서 증폭 프라이머의 상대적 위치 범위	
99-79335	51031	51559		51031	51051		51539	51559
99-79336	60925	61374		60925	60945		61354	61374
99-79338	80271	80720		80271	80290		80700	80720
99-79339	91037	91486		91037	91056		91466	91486
99-79314	100285	100784		100285	100305		100764	100784
99-79316	106568	107020		106568	106585		107000	107020
99-79322	165864	166401		165864	165884		166381	166401
99-79306	235713	236210		235713	235732		236190	236210
앰플리콘	SEQ ID 3에서 앰플리콘의 위치 범위		프라이머 이름	SEQ ID 3에서 증 프라이머의 위치 범위		프라이머 이름	SEQ ID 3에서 증폭 프라이머의 상대적 위치 범위	
99-79310	31618	32100	B15	31618	31635	C15	32080	32100
99-79311	42324	42723	B16	42324	42344	C16	42704	42723
앰플리콘			프라이머 이름	SEQ ID 6에서 증 프라이머의 위치 범위		프라이머 이름	SEQ ID 6에서 증폭 프라이머의 상보적 위치 범위	
99-62617			B17	1	20	C17	435	453

적절하게는, 프라이머는 서열분석에 사용되고 증폭을 목표로 하는 특정 염기의 상류에서 공통 올리고뉴클레오티드 꼬리를 보유한다.

프라이머 PU는 다음의 추가적인 PU 5' 서열: TGTAAAACGACGGCCAGT를 보유한다; 프라이머 RP는 다음의 RP 5' 서열: CAGGAAACAGCTATGACC를 보유한다. 추가적인 PU 5' 서열을 보유하는 프라이머는 SEQ ID No 7에 기술한다. 추가적인 RP 5' 서열을 보유하는 프라이머는 SEQ ID No 8에 기술한다.

이들 프라이머의 합성은 다음의 GENSET UFPS 24.1 서열합성장치에서 포스포르아미디트(phosphoramidite) 방법에 따라 실시한다.

DNA 증폭은 Genius II 서열증폭기에서 실시한다. 95°C에서 10분동안 가열한 이후, 40회 실시한다. 각 사이클은 다음과 같이 구성된다:95°C에서 30초, 54°C에서 1분, 72°C에서 30초. 최종 신장을 위하여, 72°C에서 10분으로 증폭을 종결시킨다. 수득된 증폭 산물의 함량은 형광계 및 개재제(intercalant agent) Picogreen(Molecular Probes)을 이용하여 96-웰 마이크로역가 평판에서 측정한다.

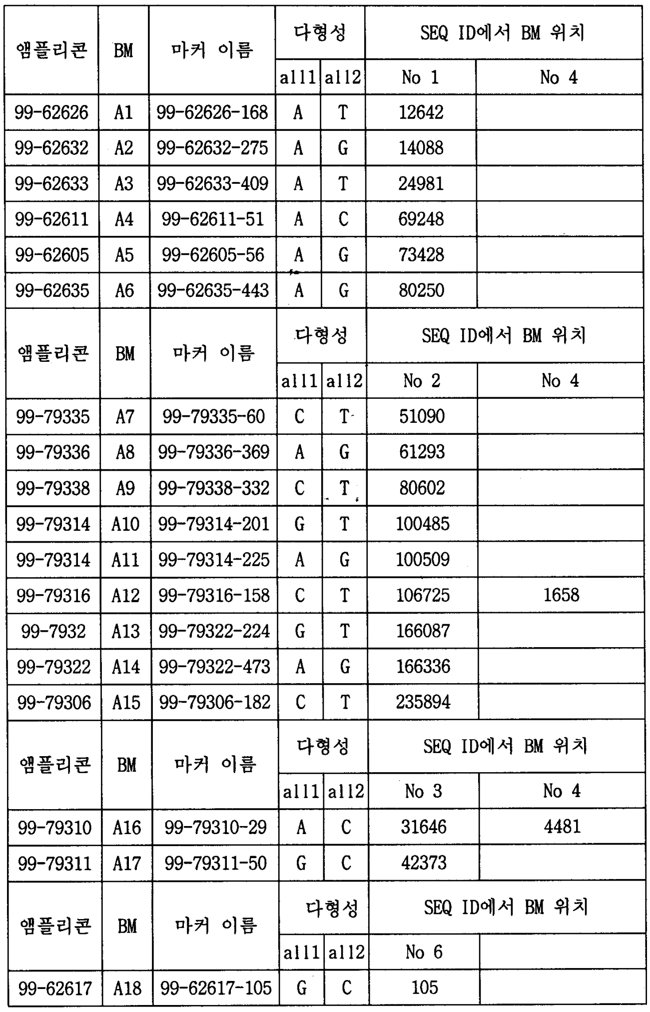
실시예 3: 이중대립인자 마커의 동정 - 증폭된 게놈 DNA의 서열분석 및 다형성의 확인

실시예 2에서 증폭된 DNA의 서열분석은 ABI 서열분석기에서 실시한다. 증폭 산물의 서열은 염료 종결기 사이클 서열분석 프로토콜(dye terminator cycle sequencing protocol)을 이용한 자동화 디데옥시 종결기 염기서열 반응으로 측정한다. 이런 염기서열 반응의 산물은 서열분석 겔에서 이동시키고, 서열은 겔 영상 분석(ABI Prism DNA 서열분석 소프트웨어(2.1.2))으로 확인한다.

서열 데이터는 추가로 평가하여 증폭된 단편내에 이중대립인자 마커의 존재를 검출한다. 다형성 검색은 전술한 바와 같이 동일한 위치에서 발생하는 상이한 염기에 기인하는 전기영동 패턴에서 중첩된 피크의 존재의 기초한다.

17개의 증폭 단편에서, 18개의 이중대립인자 마커가 검출되었다. 이들 이중대립인자 마커의 위치는 표 2에 도시한다.

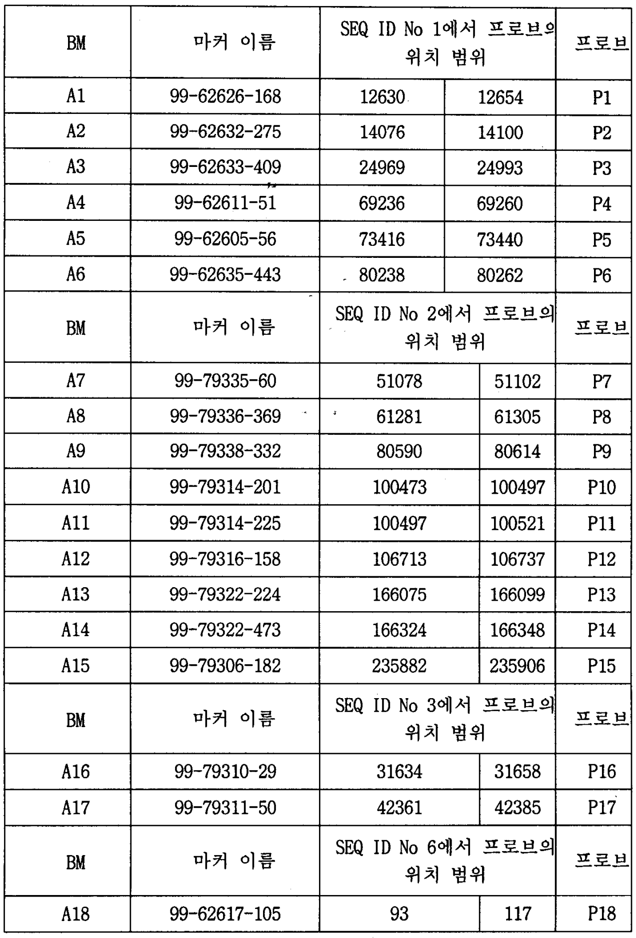
표 2:



앰플리콘	BM	마커 이름	다형성		SEQ ID에서 BM 위치	
			all1	all2	No 1	No 4
99-62626	A1	99-62626-168	A	T	12642	
99-62632	A2	99-62632-275	A	G	14088	
99-62633	A3	99-62633-409	A	T	24981	
99-62611	A4	99-62611-51	A	C	69248	
99-62605	A5	99-62605-56	A	G	73428	
99-62635	A6	99-62635-443	A	G	80250	
앰플리콘	BM	마커 이름	다형성		SEQ ID에서 BM 위치	
			all1	all2	No 2	No 4
99-79335	A7	99-79335-60	C	T	51090	
99-79336	A8	99-79336-369	A	G	61293	
99-79338	A9	99-79338-332	C	T	80602	
99-79314	A10	99-79314-201	G	T	100485	
99-79314	A11	99-79314-225	A	G	100509	
99-79316	A12	99-79316-158	C	T	106725	1658
99-7932	A13	99-79322-224	G	T	166087	
99-79322	A14	99-79322-473	A	G	166336	
99-79306	A15	99-79306-182	C	T	235894	
앰플리콘	BM	마커 이름	다형성		SEQ ID에서 BM 위치	
			all1	all2	No 3	No 4
99-79310	A16	99-79310-29	A	C	31646	4481
99-79311	A17	99-79311-50	G	C	42373	
앰플리콘	BM	마커 이름	다형성		SEQ ID에서 BM 위치	
			all1	all2	No 6	
99-62617	A18	99-62617-105	G	C	105	

BM은 '이중대립인자 마커'를 의미한다. all1과 all2는 각각, 이중대립인자 마커의 대립유전자 1과 대립유전자 2를 의미한다.

표 3:



BM	마커 이름	SEQ ID No 1에서 프로브의 위치 범위		프로브
A1	99-62626-168	12630	12654	P1
A2	99-62632-275	14076	14100	P2
A3	99-62633-409	24969	24993	P3
A4	99-62611-51	69236	69260	P4
A5	99-62605-56	73416	73440	P5
A6	99-62635-443	80238	80262	P6
BM	마커 이름	SEQ ID No 2에서 프로브의 위치 범위		프로브
A7	99-79335-60	51078	51102	P7
A8	99-79336-369	61281	61305	P8
A9	99-79338-332	80590	80614	P9
A10	99-79314-201	100473	100497	P10
A11	99-79314-225	100497	100521	P11
A12	99-79316-158	106713	106737	P12
A13	99-79322-224	166075	166099	P13
A14	99-79322-473	166324	166348	P14
A15	99-79306-182	235882	235906	P15
BM	마커 이름	SEQ ID No 3에서 프로브의 위치 범위		프로브
A16	99-79310-29	31634	31658	P16
A17	99-79311-50	42361	42385	P17
BM	마커 이름	SEQ ID No 6에서 프로브의 위치 범위		프로브
A18	99-62617-105	93	117	P18

실시예 4: 마이크로서열분석을 통한 다형성의 검증

실시예 3에서 확인된 이중대립인자 마커는 추가로 확인하고, 이들 각각의 빈도는 마이크로서열분석으로 측정한다. 마이크로서열분석은 실시예 1에서 밝힌 각 개별 DNA 시료에 대하여 실시한다.

개체의 게놈 DNA로부터 증폭은 동일한 세트의 PCR 프라이머(표 1)를 이용하여 이중대립인자 마커의 검출에서 전술한 바와 같은 PCR로 실시한다.

마이크로서열분석에 적합한 프라이머는 길이가 대략 19개 뉴클레오티드이고 고려된 다형적 염기의 바로 상류와 혼성화된다. 본원 발명에 따라 마이크로서열분석에 사용되는 프라이머는 표 4에 상세하게 기술한다.

표 4:

마커 이름	대립 유전자 마커	Mis. 1	SEQ ID No 1에서 마이크로서열분석 프라이머 mis 1의 위치 범위		Mis. 2	SEQ ID No 1에서 마이크로서열분석 프라이머 mis 2의 위치 범위	
99-62626-168			12623	12641		12643	12661
99-62632-275			14069	14087		14089	14107
99-62633-409			24962	24980		24982	25000
99-62611-51			69229	69247		69249	69267
99-62605-56			73409	73427		73429	73447
99-62635-443			80231	80249		80251	80269
마커 이름	대립 유전자 마커	Mis. 1	SEQ ID No 2에서 마이크로서열분석 프라이머 mis 1의 위치 범위		Mis. 2	SEQ ID No 2에서 마이크로서열분석 프라이머 mis 2의 위치 범위	
99-79335-60			51071	51089		51091	51109
99-79336-369			61274	61292		61294	61312
99-79338-332			80583	80601		80603	80621
99-79314-201			100466	100484		100486	100504
99-79314-225			100490	100508		100510	100528
99-79316-158			106706	106724		106726	106744
99-79322-224			166068	166086		166088	166106
99-79322-473			166317	166335		166337	166355
99-79306-182			235875	235893		235895	235913
마커 이름	대립 유전자 마커	Mis. 1	SEQ ID No 3에서 마이크로서열분석 프라이머 mis 1의 위치 범위		Mis. 2	SEQ ID No 3에서 마이크로서열분석 프라이머 mis 2의 위치 범위	
99-79310-29			31627	31645		31647	31665
99-79311-50			42354	42372		42374	42392
마커 이름	대립 유전자 마커	Mis. 1	SEQ ID No 6에서 마이크로서열분석 프라이머 mis 1의 위치 범위		Mis. 2	SEQ ID No 6에서 마이크로서열분석 프라이머 mis 2의 위치 범위	
99-62617-105			86	104		106	124

Mis1과 Mis2는 각각, CanIon 유전자의 비-코딩 가닥 또는 CanIon 유전자의 코딩 가닥과 혼성화되는 마이크로서열분석 프라이머를 의미한다.

마이크로염기서열 반응은 다음과 같이 실시한다:

증폭 산물의 정제이후, 마이크로서열분석 반응 혼합물은 제조업자의 권고에 따라, 10 pmol 형광서열분석 올리고뉴클레오티드, 1U Thermosequenase(Amersham E79000G), 1.25μℓ Thermosequenase 완충액(260 mM Tris HCl pH 9.5, 65 mM $MgCl_2$) 및 사용된 각 이중대립인자 마커의 다형적 위치에서 뉴클레오티드에 상보적인 2개의 적합한 형광 ddNTP(Perkin Elmer, Dye Terminator Set 401095)을 20 μℓ 최종 부피로 첨가한다. 94°C에서 4분후, Tetrad PTC-225 서열증폭기(MJ Research)에서 55°C에서 15초, 72°C에서 5초, 94°C에서 10초의 PCR 사이클을 20회 실시한다. 이후, 통합되지 않은 염료 종료기는 에탄올 침전으로 제거한다. 시료는 포름아마이드-EDTA 적하 완충액에서 최종적으로 재부유시키고, 95°C에서 2분동안 가열한 이후 폴리아크릴아마이드 서열분석 겔에 적하시킨다. 데이터는 ABI PRISM 377 DNA 서열분석기로 수집하고 CanIonScan 소프트웨어(Perkin Elmer)로 처리한다.

겔 분석이후, 데이터는 각 증폭된 단편에 존재하는 이중대립인자 마커의 대립유전자 확인을 가능하게 하는 소프트웨어로 자동 처리한다.

상기 소프트웨어는 상기 마이크로서열분석 과정으로부터 기인한 신호의 강도가 약한지, 정상인지 또는 충만한지, 또는 이런 신호가 불명확한지와 같은 요인을 평가한다. 이에 더하여, 상기 소프트웨어는 현저한 피크를 확인한다(형태와 높이 기준에 따라). 현저한 피크중에서, 표적 부위에 상응하는 피크는 위치에 기초하여 확인한다. 2개의 현저한 피크가 동일한 위치에서 검출되면, 각 시료는 높이비(height ratio)에 기초하여 동질성 또는 이질성 유형으로 분류한다.

실시예 5

CanIon 단백질에 대한 항체 조성물의 제조

실질적으로 순수한 단백질 또는 펩티드는 CanIon 단백질이나 이의 일부를 인코드하는 발현 벡터를 포함하는 트랜스펙션 또는 형질전환된 세포로부터 분리된다. 최종 제형에서 단백질의 농도는 예로써 Amicon 필터 장치에서 농축하여 수 ㎎/㎖ 수준까지 조정한다. 이후, 상기 단백질에 대한 단클론이나 다클론 항체는 다음과 같이 만들 수 있다.

A. 하이브리도마 융합체에 의한 단클론 항체 생산

CanIon 단백질이나 이의 일부에 대한 단클론 항체는 Kohler, G. and Milstein, C.,(1975)의 전형적 방법 또는 이의 변형 방법에 따라 뮤린 하이브리도마로부터 만들 수 있다(Harlow, E. and D. Lane. 1988 참조).

간단히 말하면, 생쥐는 수주동안 수 ㎎의 CanIon 단백질이나 이의 일부분을 반복적으로 접종한다. 이후, 생쥐는 죽이고, 비장의 항체 생산 세포를 분리한다. 비장 세포는 폴리에틸렌 글리콜로 골수종 세포와 융합시키고, 과량의 융합되지 않은 세포는 아미노프테린을 함유하는 선택 배지(HAT 배지)에 성장시켜 파괴시킨다. 성공적으로 융합된 세포는 희석하고, 희석액 일부는 마이크로역가 평판의 웰에 도말하는데, 여기서 배양액의 성장이 지속된다. 항체-생산 클론은 면역분석 과정, 예를 들면 ELISA(Engvall,(1980)) 및 이의 변형 방법으로 웰의 상등액에서 항체의 검출로 동정한다. 선별된 양성 클론은 증폭시키고, 이들의 단클론 항체 산물은 수거한다. 단클론 항체 생산의 상세한 과정은 Davis, L. et al. Basic Methods in Molecular Biology Elsevier, New York. Section 21-2에서 기술한다.

B. 면역화에 의한 다클론 항체 생산

CanIon 단백질이나 이의 일부분에서 이질성 에피토프에 대한 항체를 포함하는 다클론 항혈청은 적합한 사람제외 동물을 CanIon 단백질이나 이의 일부분으로 면역화시켜 만들 수 있는데, 이는 면역원성을 강화시키기 위하여 변화시키거나 또는 변화시키지 않을 수 있다. 적절하게는, 사람제외 동물은 생쥐, 쥐, 토끼, 염소 또는 말에서 선택되는 사람제외 포유동물이다. 대안으로, CanIon 농도가 짙은 정제되지 않은 제형이 항체를 만드는데 사용될 수 있다. 이런 단백질, 단편 또는 제형은 당분야에 공지된 적합한 어쥬번트(예, 수산화알루미늄, RIBI 등)의 존재하에 사람제외 포유동물에 도입한다. 이에 더하여, 이런 단백질, 단편 또는 제형은 항원성을 증가시키는 약품으로 사전처리할 수 있는데, 이런 약품은 당업자에게 공지된 것으로, 여기에는 메틸화된 소 혈청 알부민(mBSA), 소 혈청 알부민(BSA), B형 간염 표면 항원, 키홀 림펫 헤모시아닌(KLH) 등이 포함된다. 면역된 동물로부터 얻은 혈청은 공지된 과정에 따라 수거하고 처리하며 검사한다. 이런 혈청이 원치 않는 에피토프에 대한 다클론 항체를 함유하면, 다클론 항체는 면역친화성 크로마토그래피로 정제할 수 있다.

효과적인 다클론 항체 생산은 항원과 숙주 종류 모두에 관련된 다수 인자의 영향을 받는다. 또한, 숙주 동물은 접종 부위와 용량에 대하여 상이한 반응을 보이는데, 부족한 또는 과도한 용량의 항원은 저 역가 항혈청을 초래한다. 다수의 피내 부위에 투여된 소량(ng 수준)의 항원이 가장 안전한 것으로 생각된다. 다클론 항혈청을 생산하고 처리하는 기술은 당분야에 공지된 것이다(Mayer and Walker(1987)). 토끼에 대하여 효과적인 면역화 프로토콜은 Vaitukaitis, J. et al.(1971)에서 확인할 수 있다.

효능 촉진 주사는 정기적인 간격으로 제공하고, 항혈청은 예로써 공지된 농도의 항원에 대한 아가에서 이중 면역확산으로 반-정량적으로 측정하는 경우에 항체 역가 감소하기 시작하면 수거한다(Ouchterlony, O. et al.(1973)). 항체의 고원기 농도는 일반적으로 혈청 ㎖당 0.1 내지 0.2㎎(대략 12μM)이다. 항원에 대한 항혈청의 친화성은 경합 결합 곡선을 작성하여 확인한다(Fisher, D.,(1980))

단클론이나 다클론 프로토콜에 따라 만들어진 항체 제형은 생물 시료에서 항원-보유 물질의 농도를 측정하는 정량적 면역분석에 유용하다; 또한, 이들은 생물 시료에서 항원의 존재를 정성적으로 또는 정량적으로 확인하는데 사용된다. 이들 항체는 상기 단백질을 발현하는 세포를 죽이거나 또는 체내에서 상기 단백질의 수준을 감소시키는 치료 조성물에 사용할 수도 있다.

본원 발명의 적절한 구체예를 예시하였지만, 본원 발명의 기술적 사상과 범주를 벗어나지 않는 다양한 개변이 당업자에 의해 진행될 수 있다.

참고문헌

Abbondanzo SJ et al., 1993, Methods in Enzymology, Academic Press, New York, pp 803-823

Ajioka R.S. et al., *Am. J. Hum. genet*., 60:1439-1447, 1997

Altschul et al., 1990, J. Mol. Biol. 215(3):403-410

Altschul et al., 1993, Nature genetics 3:266-272

Altschul et al., 1997, Nuc. Acids Res. 25:3389-3402

Ames, R.S. et al. J. Immunol. Methods 184:177-186(1995)

anderson et al., Symp. Soc. Exp. Biol. 48: 85-97 (1994)

Anton M. et al., 1995, J. Virol., 69 : 4600-4606

Araki K et al. (1995) *Proc. Natl. Acad. Sci. USA* . 92(1):160-4.

Ashkenazi, A. et al. PNAS 88:10535-10539(1991)

Asz di et al., Proteins: Structure, Function, and genetics, Supplement 1:38-42 (1997)

Ausubel et al. (1989)Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Bartunek, P. et al. Cytokine 8(1):14-20 (1996)

Baubonis W. (1993) *Nucleic Acids Res* . 21(9):2025-9.

Beaucage et al., *Tetrahedron Lett* 1981, 22: 1859-1862

Better, M. et al. Science 240:1041-1043(1988)

Blouin JL et al. Nature genetics, 20: 70-73 (1998)

Bradley A., 1987, Production and analysis of chimaeric mice. *In* : E.J. Robertson (Ed.), Teratocarcinomas and embryonic stem cells: A practical approach. IRL Press, Oxford, pp.113.

Bram RJ et al., 1993, Mol. Cell Biol., 13 : 4760-4769

Brinkman U. et al. J. Immunol. Methods 182:41-50(1995)

Brown EL, Belagaje R, Ryan MJ, Khorana HG, *Methods Enzymol* 1979;68:109-151

Brutlag et al. Comp. App. Biosci. 6:237-245, 1990

Brzustowicz et al., Am. J. Hum. genet. 65:1096-1103 (1999)

Bulman et al., Hum. Mol. Gen. 6(10) 1679-1685 (1997)

Burton, D.R. et al. Advances in Immunology 57:191-280(1994)

Bush et al., 1997, J. Chromatogr., 777 : 311-328.

Carlson, N.G. et al. J. Biol. Chem. 272(17):11295-11301(1997)

Chai H. et al. (1993) *Biotechnol. Appl. Biochem* .18:259-273.

Chee et al. (1996) *Science.* 274:610-614.

Chen 및 Kwok *Nucleic Acids Research* 25:347-353 1997

Chen et al. (1987) *Mol. Cell. Biol.* 7:2745-2752.

Chen et al. *Proc. Natl. Acad. Sci. USA* 94/20 10756-10761,1997

Chen, Z. et al. Cancer Res. 58(16): 3668-3678(1998)

Cho RJ et al., 1998, Proc. Natl. Acad. Sci. USA, 95(7) : 3752-3757.

Chou J.Y., 1989, Mol. Endocrinol., 3: 1511-1514.

Clark A.G. (1990) *Mol. Biol. Evol.* 7:111-122.

Coles R, Caswell R, Rubinsztein DC, *Hum Mol genet* 1998;7:791-800

Compton J. (1991) *Nature* . 350(6313):91-92.

Davis L.G., M.D. Dibner, and J.F. Battey, Basic Methods in Molecular Biology, ed., Elsevier Press, NY, 1986

Dempster et al., (1977) *J. R. Stat. Soc.,* 39B:1-38.

Deng, B. et al. Blood 92(6): 1981-1988(1998)

Dent DS amp; Latchman DS (1993) The DNA mobility shift assay. In: *Transcription Factors: A Practical Approach* (Latchman DS, ed.) pp1-26. Oxford: IRL Press

Denyer et al.,Drug Disc. Today 3(7): 323-332 (1998

Eckner R. et al. (1991) *EMBO J.* 10:3513-3522.

Edwards et Leatherbarrow, *Analytical Biochemistry* , 246, 1-6 (1997)

Engvall, E., Meth. Enzymol. 70:419 (1980)

Excoffier L. and Slatkin M. (1995) *Mol. Biol. Evol.,* 12(5): 921-927.

Feldman and Steg, 1996, Medecine/Sciences, synthese, 12:47-55

Felici F., 1991, J. Mol. Biol., Vol. 222:301-310

Fell, H.P. et al. J. Immunol. 146:2446-2452(1991)

Fields and Song, 1989, Nature, 340 : 245-246

Fisher, D., Chap. 42 in: Manual of Clinical Immunology, 2d Ed. (Rose and Friedman, Eds.) Amer. Soc. For Microbiol., Washington, D.C. (1980)

Flotte et al. (1992) *Am. J. Respir. Cell Mol. Biol* . 7:349-356.

Fodor et al. (1991) *Science* 251:767-777.

Fraley et al. (1979) *Proc. Natl. Acad. Sci. USA.* 76:3348-3352.

Fried M, Crothers DM, *Nucleic Acids Res* 1981;9:6505-6525

Fromont-Racine M. et al., 1997, Nature genetics, 16(3) : 277-282.

Fuller S. A. et al. (1996) *Immunology in Current Protocols in Molecular Biology* , Ausubel et al.Eds, John Wiley amp; Sons, Inc., USA.

Furth P.A. et al. (1994) *Proc. Natl. Acad. Sci USA.* 91:9302-9306.

Garner MM, Revzin A, *Nucleic Acids Res* 1981;9:3047-3060

Geysen H. Mario et al. 1984. Proc. Natl. Acad. Sci. U.S.A. 81:3998-4002

Ghosh and Bacchawat, 1991, *Targeting of liposomes to hepatocytes* , IN: *Liver Diseases, Targeted diagnosis 및 therapy using specific receptors and ligands* . Wu et al. Eds., Marcel Dekeker, New York, pp. 87-104.

Gillies, S.D. et al. J. Immunol. Methods 125:191-202(1989)

Gillies, S.O. et al. PNAS 89:1428-1432(1992)

Gonnet et al., 1992, Science 256:1443-1445

Gonzalez et al., Drug Disc. Today 4(9): 431:439 (1999)

Gopal (1985) *Mol. Cell. Biol.* , 5:1188-1190.

Gossen M. et al. (1992) *Proc. Natl. Acad. Sci. USA.* 89:5547-5551.

Gossen M. et al. (1995) *Science* . 268:1766-1769.

Graham et al. (1973) *Virology* 52:456-457.

Green et al., *Ann. Rev. Biochem* . 55:569-597 (1986)

Greenspan and Bona, FASEB J. 7(5):437-444 (1989)

Griffin et al. *Science* 245:967-971 (1989)

Grompe, M. (1993) *Nature genetics.* 5:111-117.

Grompe, M. et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86:5855-5892.

Gu H. et al. (1993) *Cell* 73:1155-1164.

Gu H. et al. (1994) *Science* 265:103-106.

Guatelli J C et al. *Proc. Natl. Acad. Sci. USA.* 35:273-286.

Hacia JG, Brody LC, Chee MS, Fodor SP, Collins FS, *Nat genet* 1996;14(4):441-447

Hall L. A. and Smirnov I. P. (1997) *Genome Research* , 7:378-388.

Hames B.D. and Higgins S.J. (1985 *) Nucleic Acid hybridization: A Practical Approach* . Hames and Hamill, Pfugers Arch. 391, 85-100 (1981).

Hammerling, et al., in: MONOCLONAL ANTIBODIES and TCELL HYBRIDOMAS 563681 (Elsevier, N.Y., 1981)

Harju L, Weber T, Alexandrova L, Lukin M, Ranki M, Jalanko A, *Clin Chem* 1993;39(11Pt 1):2282-2287

Harland et al. (1985) *J. Cell. Biol.* 101:1094-1095.

Harlow, E., and D. Lane. 1988. Antibodies A Laboratory Manual. Cold Spring Harbor Laboratory. pp. 53-242

Harper JW et al., 1993, Cell, 75 : 805-816

Harrop, J.A. et al. J. Immunol. 161(4): 1786-1794(1998)

Hawley M.E. et al. (1994) *Am. J. Phys. Anthropol.* 18:104.

Hell et al., Annals N.Y. Acad. Sci. 747:282-293 (1994)

Henikoff and Henikoff, 1993, Proteins 17:49-61

Higgins et al., 1996, Methods Enzymol. 266:383-402

Higgins Ed., IRL Press, Oxford.

Hillier L. and Green P. *Methods Appl.,* 1991, 1: 124-8.

Hoess et al. (1986) *Nucleic Acids Res* . 14:2287-2300.

Huang L. et al. (1996) *Cancer Res* 56(5):1137-1141.

Huston et al. Methods in Enzymology 203:46-88(1991)

Huygen et al. (1996) *Nature Medicine.* 2(8):893-898.

Izant JG, Weintraub H, *Cell* 1984 Apr;36(4):1007-15

Julan et al. (1992) *J. Gen. Virol* . 73:3251-3255.

Kanegae Y. et al., *Nucl. Acids Res.* 23:3816-3821 (1995).

Karlin and Altschul, 1990, Proc. Natl. Acad. Sci. USA 87:2267-2268

Kettleborough, C.A. et al. Eur. J. Immunol. 24:952-958(1994)

Khoury J. et al *., Fundamentals of genetic Epidemiology, Oxford University Press, NY* , 1993

Kim U-J. et al. (1996) *genomics* 34:213-218.

Klein et al. (1987) *Nature* . 327:70-73.

Koch, et al., J.Biol.Chem. 265 (29): 17786-17791 (1990)

Kohler, G. and Milstein, C., Nature 256:495 (1975)

Koller et al. (1992) *Annu. Rev. Immunol* . 10:705-730.

Kozal MJ, Shah N, Shen N, Yang R, Fucini R, Merigan TC, Richman DD, Morris D, Hubbell E, Chee M, Gingeras TR, Nat Med 1996;2(7):753-759

Lander and Schork, *Science,* 265, 2037-2048, 1994

Landegren U. et al. (1998) *Genome Research* , 8:769-776.

Lange K. (1997) *Mathematical and Statistical Methods for genetic Analysis* . Springer, New York.

Lee et al., FEBS Lett. 445 231:236 (1999).

Lenhard T. et al. (1996) *gene.* 169:187-190.

Liautard, J. et al. Cytokinde 9(4):233-241(1997)

Lin MW et al. Hum. genet., 99(3): 417-420 (1997)

Linton M.F. et al. (1993) *J. Clin. Invest.* 92:3029-3037.

Liu Z. et al. (1994) *Proc. Natl. Acad. Sci. USA.* 91: 4528-4262.

Livak et al., *Nature genetics* , 9:341-342, 1995

Livak KJ, Hainer JW, *Hum Mutat* 1994;3(4):379-385

Lockhart et al. *Nature Biotechnology* 14: 1675-1680, 1996

Lucas A.H., 1994, In : Development and Clinical Uses of Haempophilus b Conjugate;

Manger et al., Anal. Biochem. 214: 190-194 (1993).

Mansour S.L. et al. (1988) *Nature.* 336:348-352.

Marshall R. L. et al. (1994) *PCR Methods 및 Applications.* 4:80-84.

McCormick et al. (1994) *genet. Anal. Tech. Appl.* 11:158-164.

McLaughlin B.A. et al. (1996) *Am. J. Hum. genet* . 59:561-569.

Morena et al., Annals NY Acad. Sci. 102-117 (1999)

Mori, et al., Nature 350: 398-402 (1991)

Morrison, Science 229:1202 (1985)

Morton N.E., *Am.J. Hum.genet* ., 7:277-318, 1955

Muller, Y.A. et al. Structure 6(9):1153-1167(1998)

Mullinax, R.L. et al. BioTechniques 12(6):864-869(1992)

Muzyczka et al. (1992) *Curr. Topics in Micro. and Immunol* . 158:97-129.

Nada S. et al. (1993) *Cell* 73:1125-1135.

Nagy A. et al., 1993, Proc. Natl. Acad. Sci. USA, 90: 8424-8428.

Naramura, M. et al. Immunol. Lett. 39:91-99(1994)

Narang SA, Hsiung HM, Brousseau R, *Methods Enzymol* 1979;68:90-98

Neda et al. (1991) *J. Biol. Chem* . 266:14143-14146.

Newton et al. (1989) *Nucleic Acids Res.* 17:2503-2516.

Nickerson D.A. et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:8923-8927.

Nicolau C. et al., 1987, Methods Enzymol., 149:157-76.

Nicolau et al. (1982) *Biochim. Biophys. Acta* . 721:185-190.

Nissinoff, J. Immunol. 147(8): 2429-2438 (1991).

Nyren P, Pettersson B, Uhlen M, *Anal Biochem* 1993;208(1):171-175

O'Reilly et al. (1992) *Baculovirus Expression Vectors: A Laboratory Manual* . W. H. Freeman and Co., New York.

Ohno et al. (1994) *Science.* 265:781-784.

Oi et al., BioTechniques 4:214 (1986)

Oldenburg K.R. et al., 1992, Proc. Natl. Acad. Sci., 89:5393-5397.

Olesen et al., *Voltage gated ion channel modulators* , 7-8 December , Philadelphia PA, USA (1995)

Orita et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86: 2776-2770.

Ott J., *Analysis of Human genetic Linkage* , *John Hopkins University Press* , *Baltimore,* 1991

Ouchterlony, O. et al., Chap. 19 in: Handbook of Experimental Immunology D. Wier (ed) Blackwell (1973)

Padlan E.A., Molecular Immunology 28(4/5): 489-498(1991)

Parmley and Smith, gene, 1988, 73:305-318

Pastinen et al., *Genome Research* 1997; 7:606-614

Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85(8):2444-2448

Pease S. ans William R.S., 1990, Exp. Cell. Res., 190: 209-211.

Perlin et al. (1994) *Am. J. Hum. genet.* 55:777-787.

Persic, L. et al. gene 187 9-18(1997)

Peterson et al., 1993, Proc. Natl. Acad. Sci. USA, 90 : 7593-7597.

Pietu et al. *Genome Research* 6:492-503, 1996

Pitard, V. et al. J. Immunol. Methods 205(2):177-190(1997)

Potter et al. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81(22):7161-7165.

Prat, M. et al. J. Cell. Sci. 111(Pt2): 237-247(1998)

Ramunsen et al., 1997, Electrophoresis, 18 : 588-598.

Reid L.H. et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:4299-4303.

Risch, N. and Merikangas, K. (*Science* , 273:1516-1517, 1996

Robertson E., 1987, Embryo-derived stem cell lines. In: E.J. Robertson Ed. *Teratocarcinomas and embrionic stem cells: a practical approach* . IRL Press, Oxford, pp. 71.

Roguska M.A. et al. PNAS 91:969-973), (1994) Rossi et al., *Pharmacol. Ther* . 50:245-254, (1991)

Roth J.A. et al. (1996) *Nature Medicine.* 2(9):985-991.

Roux et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86:9079-9083.

Ruano et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:6296-6300.

Sambrook, J., Fritsch, E.F., and T. Maniatis. (1989) *Molecular Cloning: A Laboratory Manual.* 2ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Samson M, et al. (1996) *Nature* , 382(6593):722-725.

Samulski et al. (1989) *J. Virol* . 63:3822-3828.

Sanchez-Pescador R. (1988) *J. Clin. Microbiol* . 26(10):1934-1938.

Sarkar, G. and Sommer S.S. (1991) *Biotechniques* .

Sauer B. et al. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85:5166-5170.

Sawai, H. et al. AJRI 34:26-34(1995)

Schedl A. et al., 1993a, Nature, 362: 258-261.

Schedl et al., 1993b, Nucleic Acids Res., 21: 4783-4787.

Schena et al. *Science* 270:467-470, 1995

Schena et al., 1996, Proc Natl Acad Sci U S A,.93(20):10614-10619.

Schneider et al.(1997) *Arlequin: A Software For Population genetics Data Analysis.* University of geneva.

Schwartz and Dayhoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of Protein sequence and Structure, Washington: National Biomedical Research Foundation

Sczakiel G. et al. (1995) *Trends Microbiol* . 3(6):213-217.

Shay J.W. et al., 1991, Biochem. Biophys. Acta, 1072: 1-7.

Sheffield, V.C. et al. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 49:699-706.

Shizuya et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:8794-8797.

Shoemaker DD, et al., *Nat genet* 1996;14(4):450-456

Shu, L. et al. PNAS 90:7995-7999(1993)

Skerra, A. et al. Science 240:1038-1040. (1988)

Smith (1957) *Ann. Hum. genet.* 21:254-276.

Smith et al. (1983) *Mol. Cell. Biol* . 3:2156-2165.

Sosnowski RG, et al., *Proc Natl Acad Sci U S A* 1997;94:1119-1123

Sowdhamini et al., Protein Engineering 10:207, 215 (1997)

Spielmann S. and Ewens W.J., *Am. J. Hum. genet* ., 62:450-458, 1998

Spielmann S. et al., *Am. J. Hum. genet* ., 52:506-516, 1993

Sternberg N.L. (1992) Trends genet. 8:1-16.

Sternberg N.L. (1994) *Mamm. Genome.* 5:397-404.

Stryer, L., *Biochemistry,* 4th edition, 1995

Studnicka G.M. et al. Protein Engineering 7(6):805-814(1994)

Syvanen AC, *Clin Chim Acta* 1994;226(2):225-236

Szabo A. et al. *Curr Opin Struct Biol* 5, 699-705 (1995)

Tacson et al. (1996) *Nature Medicine* . 2(8):888-892.

Taryman, R.E. et al. Neuron 14(4):755-762(1995)Schaid D.J. et al., *genet. Epidemiol* .,13:423-450, 1996

Te Riele et al. (1990) Nature. 348:649-651.

Terlau et al., Naturwissenschaften 85: 437-444 (1998)

Terwilliger J.D. and Ott J., *Handbook of Human genetic Linkage* , *John Hopkins University Press* , *London* , 1994

Thomas K.R. et al. (1986) *Cell* . 44:419-428.

Thomas K.R. et al. (1987) *Cell* . 51:503-512.

Thompson et al., 1994, Nucleic Acids Res. 22(2):4673-4680

Tur-Kaspa et al. (1986) *Mol. Cell. Biol.* 6:716-718.

Tyagi et al. (1998) *Nature Biotechnology.* 16:49-53.

Urdea M.S. (1988) *Nucleic Acids Research.* 11:4937-4957.

Urdea M.S. et al.(1991) *Nucleic Acids Symp. Ser* . 24:197-200.

Vaitukaitis, J. et al. J. Clin. Endocrinol. Metab. 33:988-991 (1971)

Valadon P., et al., 1996, J. Mol. Biol., 261:11-22.

Van der Lugt et al. (1991) *gene* . 105:263-267.

Vil, H. et al. PNAS 89:11337-11341(1992)

Vlasak R. et al. (1983) *Eur. J. Biochem* . 135:123-126.

Wabiko et al. (1986) *DNA.* 5(4):305-314.

Walker et al. (1996) *Clin. Chem.* 42:9-13.

Wang et al., 1997, Chromatographia, 44 : 205-208.

Weir, B.S. (1996) *genetic data Analysis II: Methods for Discrete population genetic Data* , *Sinauer Assoc., Inc., Sunderl및, MA, U.S.A* .

Westerink M.A.J., 1995, Proc. Natl. Acad. Sci., 92:4021-4025

White, M.B. et al. (1992) *genomics.* 12:301-306.

White, M.B. et al. (1997) *genomics.* 12:301-306.

Williams, et al. Science 257: 389-395 (1992)

Wong et al. *gene* . 10:87-94 (1980)

Wood S.A. et al., 1993, Proc. Natl. Acad. Sci. USA, 90: 4582-4585.

Wu and Wu (1987) *J. Biol. Chem.* 262:4429-4432.

Wu and Wu (1988) *Biochemistry* . 27:887-892.

Wu et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86:2757.

Yagi T. et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:9918-9922.

Yoon, D.Y. et al. J. Immunol. 160(7): 3170-3179(1998)

Zhang et al., Neuropharmacology 32 (11): 1075-1088 (1993)

Zhao et al., *Am. J. Hum. genet* ., 63:225-240, 1998

Zheng, X.X. et al. J. Immunol. 154:5590-5600(1995)

Zhu, Z. et al. Cancer Res. 58(15): 3209-3214(1998)

Zou Y. R. et al. (1994) *Curr. Biol.* 4:1099-1103.

(57) 청구의 범위

청구항 1.

SEQ ID NO 1 내지 4 또는 6으로 도시된 뉴클레오티드 서열중 한가지, 또는 이들 서열에 상보적인 서열을 포함하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드.

청구항 2.

SEQ ID NO 4의 적어도 50개 뉴클레오티드의 연속 스팬(span)을 포함하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드에 있어서, 상기 폴리뉴클레오티드는 생물학적 활성 CanIon 폴리펩티드를 인코드하는 것을 특징으로 하는 폴리뉴클레오티드.

청구항 3.

SEQ ID NO 5의 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함하는 사람 CanIon 폴리펩티드를 인코드하는 분리된, 정제된 또는 재조합된 폴리뉴클레오티드.

청구항 4.

제 1 항 내지 3 항중 어느 한 항에 있어서, 고체 서포트에 부착되는 것을 특징으로 하는 폴리뉴클레오티드.

청구항 5.

제 4 항에 따른 적어도 한가지 폴리뉴클레오티드를 포함하는 폴리뉴클레오티드의 어레이.

청구항 6.
제 5 항에 있어서, 어드레스(address)가능한 것을 특징으로 하는 어레이.

청구항 7.
제 1 내지 4 항중 어느 항에 있어서, 라벨을 추가로 포함하는 것을 특징으로 하는 폴리뉴클레오티드.

청구항 8.
제 1 항 내지 3 항중 어느 한 항의 폴리뉴클레오티드를 포함하는 재조합 벡터.

청구항 9.
프로모터에 동작가능하게 연결되고 제 2 항 또는 3 항의 폴리뉴클레오티드를 포함하는 재조합 벡터.

청구항 10.
제 8 항 또는 9 항의 재조합 벡터를 포함하는 숙주 세포.

청구항 11.
제 8 항 또는 9 항의 재조합 벡터를 포함하는 사람제외 숙주동물 또는 포유동물.

청구항 12.
적중(knock out) 벡터와의 상동성 재조합에 의해 파괴된 *CanIon* 유전자를 포함하는 포유동물 숙주 세포 또는 사람제외 숙주 포유동물.

청구항 13.
SEQ ID NO 5에 도시된 아미노산 서열, 또는 이의 생물학적 활성 단편을 포함하는 분리된, 정제된 또는 재조합된 폴리펩티드.

청구항 14.
폴리펩티드를 생산하는 방법에 있어서, 상기 방법은 다음의 단계로 구성되는 것을 특징으로 하는 방법:

a) 프로모터에 동작가능하게 연결된 제 13 항의 폴리펩티드를 인코드하는 폴리뉴클레오티드를 포함하는 세포 개체군을 제공하고;

b) 세포내에서 상기 폴리펩티드의 생산에 도움이 되는 조건하에 세포 개체군을 배양하며;

c) 상기 세포 개체군으로부터 폴리펩티드를 정제한다.

청구항 15.
제 13 항의 폴리펩티드에 항-CanIon 항체를 결합시키는 방법에 있어서, 항체와 폴리펩티드가 특이적으로 결합할 수 있는 조건하에 상기 항체와 상기 폴리펩티드를 접촉시키는 단계로 구성되는 것을 특징으로 하는 방법.

청구항 16.
세포내에서 CanIon 유전자의 발현을 검출하는 방법에 있어서, 상기 방법은 다음의 단계로 구성되는 것을 특징으로 하는 방법:

a) i) 엄격한 조건하에 제 1 항, 2 항 또는 3 항의 폴리뉴클레오티드와 혼성화되는 폴리뉴클레오티드; 또는 ii) 제 13 항의 폴리펩티드에 특이적으로 결합하는 폴리펩티드와 세포 혹은 세포 추출물을 접촉시키고;

b) 상기 세포 혹은 세포 추출물에서 폴리뉴클레오티드와 RNA 종간 혼성화의 존부, 또는 상기 세포 혹은 세포 추출물에서 폴리펩티드와 단백질 결합의 존부를 확인하며;

여기서, 혼성화 또는 결합의 존재는 상기 CanIon 유전자가 세포내에서 발현된다는 것을 암시한다.

청구항 17.
제 16 항에 있어서, 폴리뉴클레오티드는 올리고뉴클레오티드 프라이머이고, 혼성화는 상기 프라이머 서열을 포함하는 증폭 산물의 존재를 검출하여 확인하는 것을 특징으로 하는 방법.

청구항 18.

제 16 항에 있어서, 폴리펩티드는 항-CanIon 항체인 것을 특징으로 하는 방법.

청구항 19.
CanIon 폴리펩티드의 후보 조절물질을 동정하는 방법에 있어서, 다음의 단계로 구성되는 것을 특징으로 하는 방법:

a) 제 13 항의 폴리펩티드와 검사 화합물을 접촉시키고;

b) 상기 화합물이 상기 폴리펩티드에 특이적으로 결합하는 지를 확인하며;

여기서, 상기 화합물이 상기 폴리펩티드에 특이적으로 결합한다는 확인은 상기 화합물이 CanIon 폴리펩티드의 후보 조절물질이라는 것을 암시한다.

청구항 20.
제 19 항에 있어서, 후보 조절물질의 존재하의 CanIon 폴리펩티드의 생물활성을 측정하는 단계가 추가로 포함되고, 후보 조절물질 부재하의 활성과 비교하여 후보 조절물질 존재하의 CanIon 폴리펩티드의 생물활성 변화는 후보 조절물질이 CanIon 폴리펩티드의 조절물질이라는 것을 암시하는 것을 특징으로 하는 방법.

청구항 21.
CanIon 폴리펩티드의 조절물질을 동정하는 방법에 있어서, 다음의 단계로 구성되는 것을 특징으로 하는 방법:

a) 제 13 항의 폴리펩티드와 검사 화합물을 접촉시키고;

b) 상기 화합물의 존부하에 상기 폴리펩티드의 활성을 확인하며;

여기서, 상기 화합물 부재하의 활성과 비교하여 상기 화합물 존재하의 활성에서 차이의 확인은 상기 화합물이 CanIon 폴리펩티드의 조절물질이라는 것을 암시한다.

청구항 22.
제 20 항 또는 21 항에 있어서, 폴리펩티드는 세포 혹은 세포막에 존재하고, 생물활성은 전압-개폐 이온 채널 활성인 것을 특징으로 하는 방법.

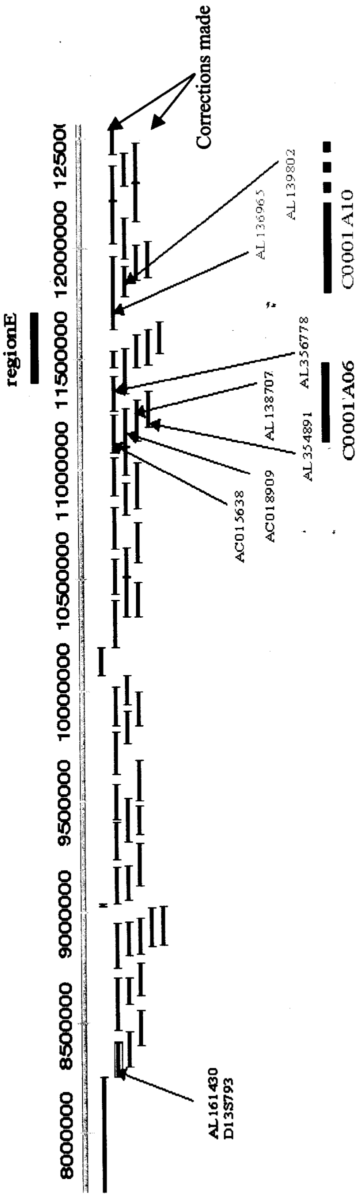
청구항 23.
제약학적 조성물을 제조하는 방법에 있어서, 다음의 단계로 구성되는 것을 특징으로 하는 방법:

a) 제 19 항 내지 22 항중 어느 한 항의 방법을 이용하여 CanIon 폴리펩티드의 조절물질을 동정하고;

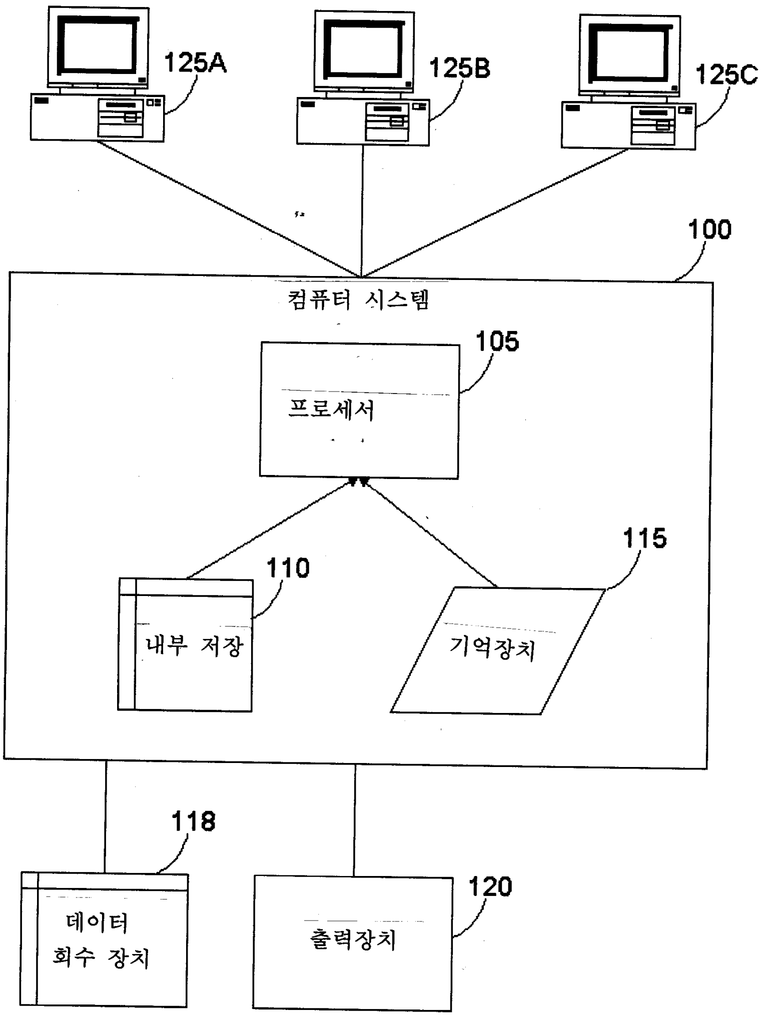
b) 상기 조절물질과 생리학적으로 수용가능한 담체를 혼합한다.

도면

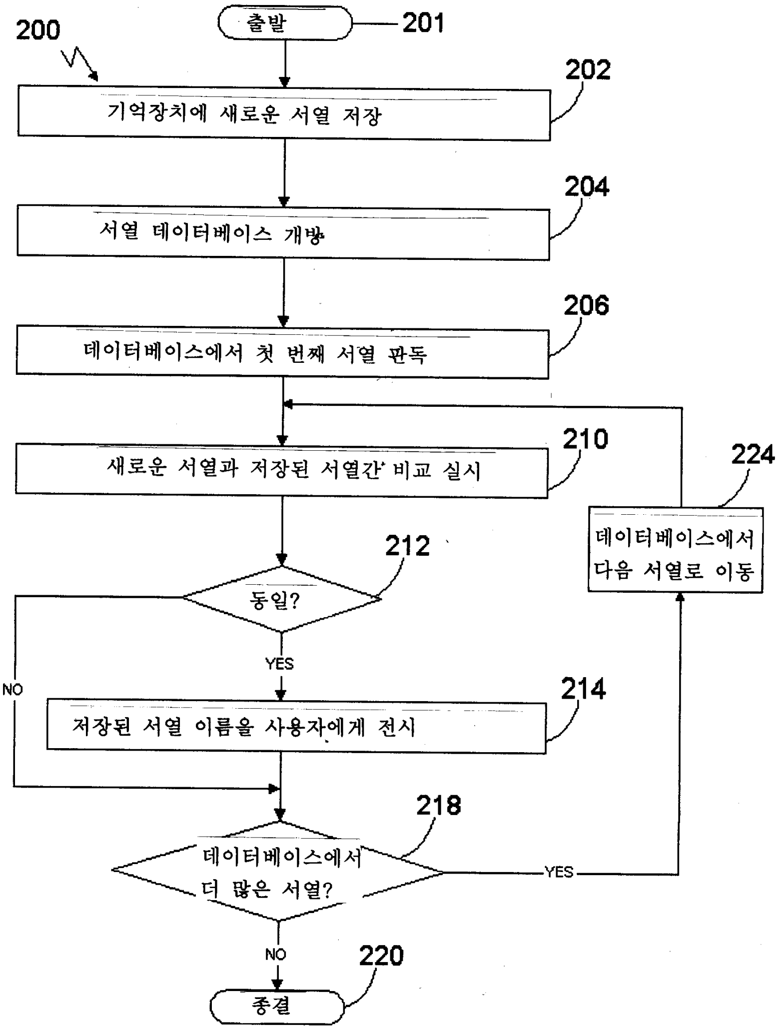
도면1


Corrections made
regionE
8000000
8500000
9000000
9500000
10000000
10500000
11000000
11500000
12000000
12500(
AL161430
D13S793
AC015638
AC018909
AL354891
AL138707
AL356778
AL136965
AL139802
C0001A06
C0001A10

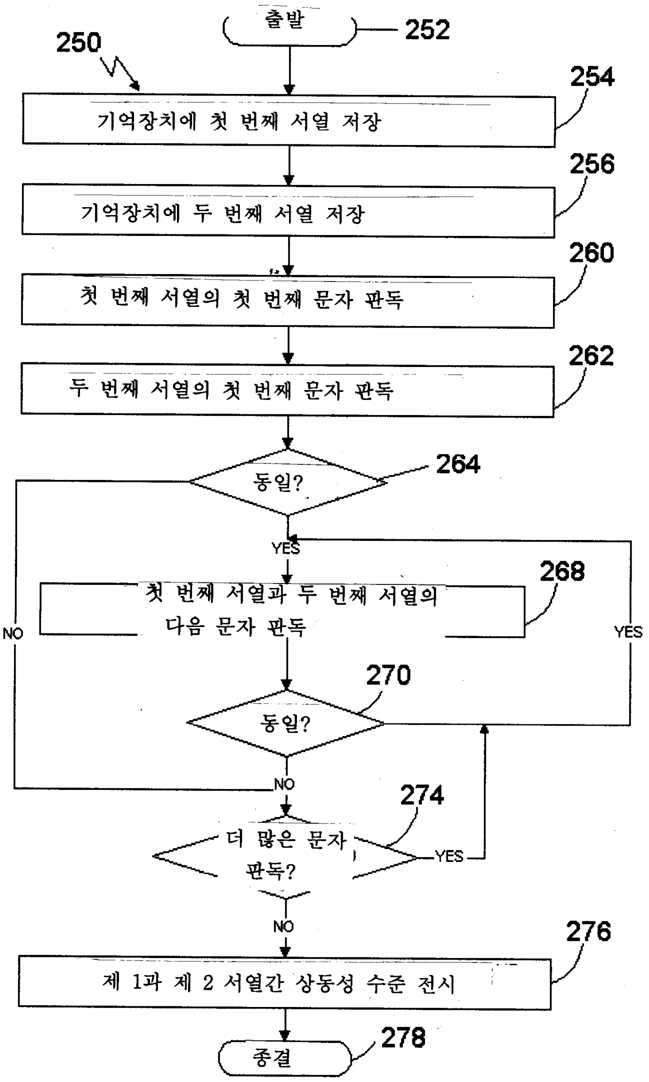
도면2


125A
125B
125C
100
컴퓨터 시스템
105
프로세서
110
내부 저장
115
기억장치
118
데이터
회수 장치
120
출력장치

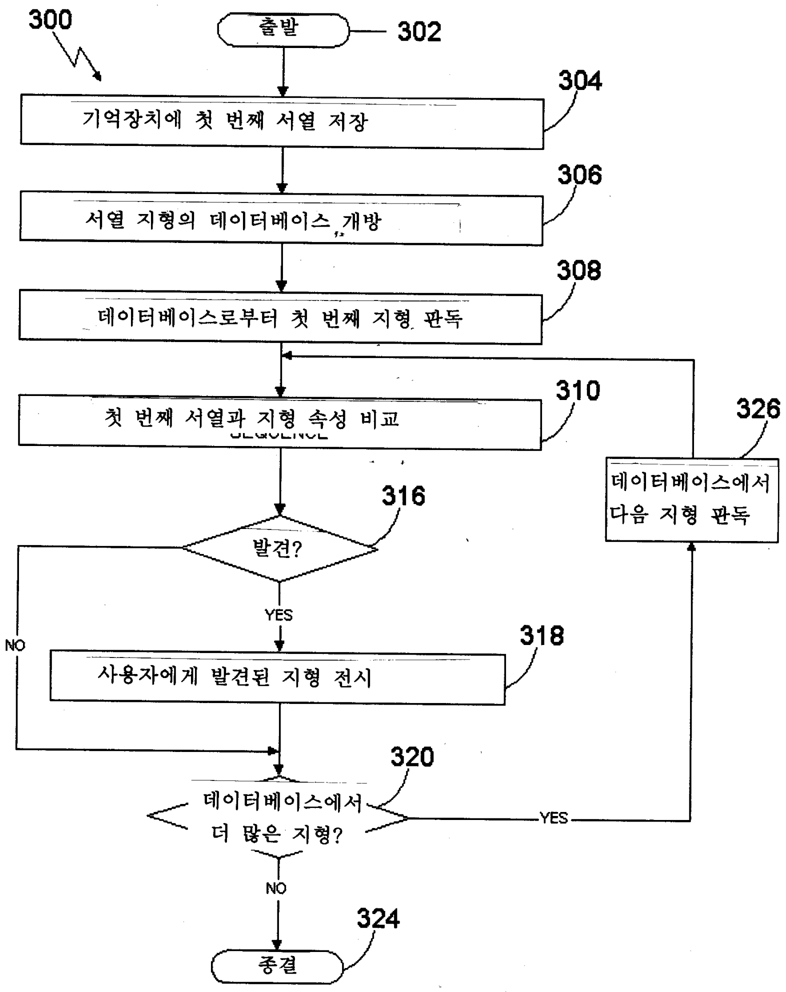
도면3


200
출발
201
기억장치에 새로운 서열 저장
202
서열 데이터베이스 개방
204
데이터베이스에서 첫 번째 서열 판독
206
새로운 서열과 저장된 서열간 비교 실시
210
224
데이터베이스에서 다음 서열로 이동
212
동일?
YES
NO
저장된 서열 이름을 사용자에게 전시
214
218
데이터베이스에서 더 많은 서열?
YES
NO
220
종결

도면4


출발
252
250
254
기억장치에 첫 번째 서열 저장
256
기억장치에 두 번째 서열 저장
260
첫 번째 서열의 첫 번째 문자 판독
262
두 번째 서열의 첫 번째 문자 판독
264
동일?
YES
268
첫 번째 서열과 두 번째 서열의 다음 문자 판독
NO
YES
270
동일?
NO
274
더 많은 문자 판독?
YES
NO
276
제 1과 제 2 서열간 상동성 수준 전시
278
종결

도면5



SEQUENCE LISTING

<110> GENSET

<120> SCHIZOPHRENIA RELATED VOLTAGE-GATED ION CHANNEL GENE AND PROTEIN

<130> 93.W01

<141> 2001-12-04

<150> US 60/251,317

<151> 2000-12-05

<160> 8

<170> Patent.pm

<210> 1

<211> 82293

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 1..2000

<223> 5'regulatory region

<220>

<221> exon

<222> 2001..2026

<223> exon 1

<220>

<221> exon

<222> 19188..19334

<223> exon 2

<220>

<221> exon

<222> 22996..23178

<223> exon 3

<220>

<221> exon

<222> 39731..39814

<223> exon 4

<220>

<221> exon

<222> 41337..41476

<223> exon 5

<220>

<221> exon

<222> 41565..41693

<223> exon 6

<220>

<221> exon

<222> 73013..73167

<223> exon 7

<220>

<221> allele

<222> 12642

<223> 99-62626-168 : polymorphic base A or T

<220>

<221> allele

<222> 14088

<223> 99-62632-275 : polymorphic base A or G

<220>

<221> allele

<222> 24981

<223> 99-62633-409 : polymorphic base A or T

<220>

<221> allele

<222> 69248

<223> 99-62611-51 : polymorphic base A or C

<220>

<221> allele

<222> 73428

<223> 99-62605-56 : polymorphic base A or G

<220>

<221> allele

<222> 80250

<223> 99-62635-443 : polymorphic base A or G

<220>

<221> primer_bind

<222> 12343..12363

<223> 99-62626.rp

<220>

<221> primer_bind

<222> 12793..12810

<223> 99-62626.pu complement

<220>

<221> primer_bind

<222> 13814..13832

<223> 99-62632.pu

<220>

<221> primer_bind

<222> 14279..14296

<223> 99-62632.rp complement

<220>

<221> primer_bind

<222> 24863..24881

<223> 99-62633.rp

<220>

<221> primer_bind

<222> 25378..25396

<223> 99-62633.pu complement

<220>

<221> primer_bind

<222> 69198..69218

<223> 99-62611.pu

<220>

<221> primer_bind

<222> 69632..69650

<223> 99-62611.rp complement

<220>

<221> primer_bind

<222> 73005..73022

<223> 99-62605.rp

<220>

<221> primer_bind

<222> 73466..73483

<223> 99-62605.pu complement

<220>

<221> primer_bind

<222> 79808..79826

<223> 99-62635.pu

<220>

<221> primer_bind

<222> 80314..80334

<223> 99-62635.rp complement

<220>

<221> primer_bind

<222> 12623..12641

<223> 99-62626-168.mis

<220>

<221> primer_bind

<222> 12643..12661

<223> 99-62626-168.mis complement

<220>

<221> primer_bind

<222> 14069..14087

<223> 99-62632-275.mis

<220>

<221> primer_bind

<222> 14089..14107

<223> 99-62632-275.mis complement

<220>

<221> primer_bind

<222> 24962..24980

<223> 99-62633-409.mis

<220>

<221> primer_bind

<222> 24982..25000

<223> 99-62633-409.mis complement

<220>

<221> primer_bind

<222> 69229..69247

<223> 99-62611-51.mis

<220>

<221> primer_bind

<222> 69249..69267

<223> 99-62611-51.mis complement

<220>

<221> primer_bind

<222> 73409..73427

<223> 99-62605-56.mis

<220>

<221> primer_bind

<222> 73429..73447

<223> 99-62605-56.mis complement

<220>

<221> primer_bind

<222> 80231..80249

<223> 99-62635-443.mis

<220>

<221> primer_bind

<222> 80251..80269

<223> 99-62635-443.mis complement

<220>

<221> misc_binding

<222> 12630..12654

<223> 99-62626-168.probe

<220>

<221> misc_binding

<222> 14076..14100

<223> 99-62632-275.probe

<220>

<221> misc_binding

<222> 24969..24993

<223> 99-62633-409.probe

<220>

<221> misc_binding

<222> 69236..69260

<223> 99-62611-51.probe

<220>

<221> misc_binding

<222> 73416..73440

<223> 99-62605-56.probe

<220>

<221> misc_binding

<222> 80238..80262

<223> 99-62635-443.probe

<220>

<221> misc_feature

<222> 2375,2377,4041..4042,5180,7138,7174,10118,12577,12637,12648,13556
15322,15341,15357,15383,15409,15994,16064..16065,16077,16237
16264,17751,19182,19424,20735,20840,21865,21885,21900,22814,23954
24969,24985,25001,25005..25006,25017,25040,25057,26646,27478
27496,29278,30793,31122,31125,34515,34541,34563,34572,34884,34915
36809,36834,36944,37011,39862,39881,39895,39903,39921,39948
39954..39955,41108,42572,43423,44122,53546,53634,55777,56140
56160,56182,56185,56200,56206,56219,56234,56258,56268,56280,56290
56324,56561,57259,58256,58297,58910,59321,60724,60743,68789,68803
68871,69936,77190,77484,78780,78891,79343,80394,81773,81795,82232

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

<400> 1

catcaatgca ctccaggcta ggcaacggag ataggccttg tctctaaaaa taacaaagaa 60

aacaaaagaa gagagatact gagcaggtag aaggggtact atgatatttt ggatgtacct 120

gtttttaaa atgattgtct actccaatac acagtgaatc cctatgcatt ataggatgag 180

tgtgctggat tccatgactt ccttgaagcc catgaattct ctgagactgt ctccatcatc 240

attgggattg tttggctcat gagcatgaat tatagagact taacagaggt agttcaacca 300

aggacagaga aaggtctgaa atgcctgttg ggtgaggagt ggcactgaga ttagtaaaaa 360

atcactgcct tgaccacttc aaaacattat aaataaatgt ttaaatattt ccttaatata 420

atttactaac agataattcg aacgaatgtc aaggaataac acactatcat tcagctacca 480

tatatatata gctatatatt tcagtttgtg aaataaacta gtatttagaa caagcaattg 540

aaagccgcaa ataaagatga tgaaacaagt tggtaaatac tttttctaag caaactggtt 600

gtagatttag cattgtttaa attactgtca attttttctt atgaggagtg atttcatcat 660

ttgtttaaaa cctggctttt gacacagttt ttctgaagtc tttagaaatt tgaaaaatta 720

tctcgcatga ctctgtattt catgaaatac caatatttta acatttccct gtaaacaata 780

aaacgtacta atatggatta tgaatgtgaa caaattgtga atggactttt taagaaagta 840

gattaattaa atatctgtaa ccagaaaata catttttaaa atgtaggttt gtgtagtttt 900

acttctatac cttgccacat aacattttgt gatagacaaa tgctctttgt ttagaatttc 960

```
tcttgtagag aaagcttcat tcagttgttt ctttcccttt atttatatgc ctgtgtgcat   1020
gatgtctaaa tcactggtca ggagttccct ggggattttc gtatctgact acatggcaag   1080
attccagacc caaagtaaac ctttcagaga gtgggaatga gtgggaggat gtcgccctgg   1140
gcatggtgct taggaggcca gatttcccta aacctgctct gctgggagcc ctaggaagcc   1200
tcagatgcag tctgctgtcc aggtatactc acagggtatg gggtggggga ctgtcaattt   1260
ctgaatagac tttgtctcat ccacacgaca acatatttta attttaaagg cgaacgtcaa   1320
cctatttaga aaactataag tgtgactgat gcattcatta tttaataact ccggggtgaa   1380
acttgaggtg aacggtttca ccctttcata ctcaggaatg gaactagtcc tcggtaccgt   1440
gcacagtctg acctcacgct cgcacgctgg gctcccggcc tcctggcagt gcacttaagt   1500
gtaccccgcg tgtgcccggc ggactcccgt gcacaccgcc gccgtgctgt tgagtacgaa   1560
agacccccga gaggagctcg cagatgagag cctagctgtt gctgtgagac tgactgcggg   1620
tggcagccgg caagaaggaa ctgccagtac cgaccctggc gctccccacc aactgcggcg   1680
cagcctcctc tgcggtgccc gcccttcggc atcactgcag tccctgtagc aaagctaact   1740
ttaaaaattt aaataaaata ataatttttt aaaggcccga aaactccggc tcccacccct   1800
cccccagcct gcccggcggc ggcggcgctc agtgagtgac acggcgagcg tgagccgcgg   1860
ccccagccgg gccgagcgcg ctgcctgagc tgagccgccg taggtgaggg gcccgcgtcc   1920
ccgcccgccc tgggcgccgc gcctggcact gatcctgccg gtcgcccact gtcgccgccg   1980
ccgccgcccg cgggcaccat gacagctctg agcgctgggg ttacaggttg gtcaccgccg   2040
ccggtgcccg ggaggctgcg gcgcggccgg ctccgtgccc tgcgcccgag gcccgcgatg   2100
cgcgctcgtt agttccttgg ccggtccggg ccaggacggc tcaggggcgc tggggctgcg   2160
agaggggctg ctagaggggc tgcgggcggg ggccccgcgc gggggagggg tggggagggg   2220
cgcctccacc cgagagcggg gcggtggggt ccttgccggc ctggccgggg cgcgggagcc   2280
gcgatgcgcc ctccctattc ctgcgcgtgg gtgtcggggg cgcctgccgc gggatgcccc   2340
tgggctggct tgggaagccc agggccggcg acacntnggg cctgccccgc ctgggcatct   2400
agcaggtgtt ggggggcagg tggctgagga cgaggccact gggcatcact acagcgctgc   2460
acctgccccc tcggcaccag ggccgctgcc tgggtaaacc tctgggccaa cgcgcgggtg   2520
gggaggtggg ggtgggcagc ctggctccag caggtctccg cgggggctcc gcggccttgc   2580
cccacccccca tgctcctgtg gattgctcct tgctaaatgt cccttatgat attcattggc   2640
cttgttaagc agccgtcatt agggtggcta caccgaacgc aaaacagtgt cctgccaagc   2700
aacaccattt tgtcactgaa tctttaagac accagcaagg atggcctggc tgggatccct   2760
ctgggggccc agatcgcagt cattttgctg acttaatgtt cctctttcag gatggtgtga   2820
gaatgctgaa ctcagaccct tagtgcttga aataccttag aataaatctt tcagtttcag   2880
gtacaatttt tcagacactg tattattaag gatgtgtgaa tacacgctca ctttcaggcc   2940
tcagaaggtt ttactcttaa tctggtgttc gtgtggtcat cagctaggcc cgactgccct   3000
```

ggtggtgaaa tgtgggttga ggaccagtga ccttcgtgtt agtagttggt ggataacacc 3060

tctaataggg gatcaaaaat taacaggtgg cgattttgaa atatatatgt atgtatatat 3120

atatatatat gtgtgtgatt tgtatgttca taactacaac tgtgatttga tatggcctgg 3180

gagaggagcg gtttcactat ggtagttacc cctgaggcag atgcatctca ggactggctt 3240

tttaaattca aagataaggt ttccacagtt ctaaataaaa atcatggctt cccaaacatc 3300

tgcaacgaga aacatttact ccctaactgg ttggctgtca ttctgcgtgc ttttctagag 3360

tcgatttttc gttcttctgt agtatacttt gacctagcca ctgctaaaac cagacagtgt 3420

gaagacaaaa tccagctttg gagaaaaact taaaaagtga aaaggttcag ttggatcctc 3480

ttctgtgttc atagaagaca taatctgttt gcagaataaa tgtgccgtca ctgataagga 3540

ctatgctaaa atctctacag aactctagga gaaagcatgt cttataaaat agtggttgta 3600

cacttataac tttggatagg acttgcaaga tattcaaagg agaaatgtac ttaaaaaatg 3660

tacttaaaaa aagtctactc aaaatagtgt catttagttg caatttagtg aactgaagat 3720

ttgagtagca ttttcttctt cttgggtatg gcatttcttt gaaatggttt ccaggtgatt 3780

ttatcattcc ggcatgcttt gtgtgtgcat gtatgtgtat acacaaatgt acacatttgt 3840

tttatcattt atcgatgatg ttcagataat ttcagaattt ccatagctgt ataaatttat 3900

cccaggtcag agtcacatgt ggatttcaag tttgtgaatg aaaattcaca tctacacctc 3960

tatatcctaa gataagctaa gcactgggcc agaagctggg tggtgagtga gaaaaagaga 4020

taaactgaga gcctaagaat nnttcaggaa agacaaatat gtaagcaggt ggatatacag 4080

tgagatgtct tattttctag ttttgctgtg gaaactatga tagtgagctt aaacatatta 4140

agaatggtcc ccaaggccac acacggtggc tcatgcctat aatcccagca ctttgggagg 4200

ccaaggcagg cctattgctt gaacccaaga gtccagacca gcctgggcaa catggtgaaa 4260

ccccgtctct acaaaaaaaa aaaaaattca ccaagtgtag tggtgtacac ctgtagcccc 4320

agttacttgg gaggctgagg tggaaggatc gcgtgagccc agagatgtca aggctgcagc 4380

aggctatgat ggcaccactg tactctagaa tgggcgactg ggcacagagt gagactctgt 4440

ttaaaaataa aaaaagaatg cccctgtccc aaaatgttat caacatttta gagcttcact 4500

catgttatac aattctcgcc tgtctcaagg accagcacag attttcatac cttcaaactc 4560

tatgcaaatg cttaaatttt tttaattgat tttttgaatt ttaaagtaac caaagttgat 4620

agagatgata gaatgataga gaaaaaagag taaacagcta gttttaataa ccgaagacag 4680

atgaccacaa tatctgtgtg ttttttgtag tctttttaag ggcttgactt caggtagttg 4740

tcatcgtggc attggacaac ttttatactg tttttaccac tgaatattta ttataaatat 4800

ttaaaatatt tctatataat tttcttactt aaaaattgta aatggctgta taacaccata 4860

cccagtaaat acagcctcat atatttaacc atacctctat cattggacat ttgctgtttc 4920

cagtttttca tcattataaa cagtgtcctg atagttataa tgtagagatt tatttctatt 4980

ttatgttttt cagtataaat atccaaatat gtccatatct gtcacctatt tttccatatc 5040 ttcaaattgg tttcctgaaa aattcagatt ataatttatt tatgataaat ttccagtaat 5100

tatgtctttg aaaaaagaag tatctgaatt attatagaga tatattgtat gtctttaaat 5160

tgatttccta aagaattcan gattataatt ttgaataaat tgcttaatct gcttttgagg 5220

taaatctacc cttttacaaa aagtgttggt ggatcacatt ccctctttta cagaaataaa 5280

agaaggtaac ataataattt ggatgtttga cctttgtgct atataaaaac ttgaccttta 5340

caaaaatcac ctgaaaacca gatgctaaaa tatacctctt agagtaaagc caaaatagga 5400

atactataaa agttcaaagc tggaaaaaga aaaaaatcaa tgcaagttcc tttagaaaat 5460

tggctgttgt attcaggagt ctaattctgc tttttgcttt cctactcaga cttcaagtaa 5520

atgatcaaaa gggcttacat agtgagaaat ctgtaccccc acaggtaaag gaagaagagg 5580

atctttcttt cctaggctgc agaatccaag gcttttctgg tagattcagt atagatggaa 5640

caggactgat aggaaaagct aaagttcagt tcacaatttc ctttctaaga tcagtgtaca 5700

aaaaatacgg aagagagacc ctggaagagc cccttagtat gaaggagcga agttgaaagt 5760

tggaagcagg cttttgtaaa acagaattgc ttgtttctcc atgagatccc agggcaaggc 5820

agccataccg gaatatgttc tgaagacagg gactgacccc acaggtccag ggtgggggca 5880

ggggtaggcc tccgcagcag gagtggaaga attagaagta ggatagacca gaacactgaa 5940

gccaagagaa tcatttatga agagatcaag agtgagaatt atttccaatt atgcaaagaa 6000

aagagatgaa aagaggatgt tgtttggaga gagaatgata tagatggaag acagccaaca 6060

gagacacacc aataaatgtt tttcctgaca gagaggatca cagtgcatgg aacacagcca 6120

tgttttaaag tgagataaga ttcttttttt cttttagatt aaaaacaatc caaatcttca 6180

gatgaaaata agttatcatg taccaaaaca attaataaaa cttagccaac atctacatat 6240

agcctgataa aaactttgga tttgaaaacc aataaaaaga aaaatatttc accagcattg 6300

tgatcaaatg cttcagaaat tacttatcca caaagagaaa atttagatta atgttatact 6360

tcctctctaa aattatcagc aaacataaaa taacgtcctg aaaatctgta gtcctgaaaa 6420

atatcagtag acaattctct tacgttgtaa atgagtcaaa agtgaaaaat aaataggtga 6480

agaaatacat actaggcaag tgatatcagt aaaaaaaaat gaggtagaag tgtgtaaata 6540

tcagaaaaca gttcaaaaaa ttgtttgatt aaatgagaaa aaagaccgta aagtcataaa 6600

agttgctatc tttaatacat ttataactac aggcttccca tatataccaa cgatcagagc 6660

attatagtat gcacagtcaa aagcataatg ctggccgggc gcggtggctc acgcctgtaa 6720

tcccagcact ttgggaggcc gagatgggtg gatcacgagg tcaggagttt gagaccagcg 6780

tggccaacat ggtgaaaccc cgtccctgct aaaaatacaa aaaaattagc taggtgtgat 6840

ggtgggcacc tgtaatctca gctactcggg aggggaggct gaggcaggcg aatcatttga 6900

acccaggagg cagaggttgc agcgagccga gatcacgcat tacactctag cctgggcgac 6960

agggcaagac tccgtctcaa aaaaaaaaaa aaagtctgat gcttatttta ggggaaatgg 7020

aagcggaagc atgcccacag actcaatatc tttttaaaga gttttgtgac ctaaaatggt 7080 attgggagtc agcatttcat cttaagaagt acaaaaaaaa aagaaaaaaa aaagaaanga 7140

ggataaagtt ttaagaatga aaacatgaat gaanttcact atgacataga taaactctaa 7200

atctgatagc aaaaactgag agctacagca ttccaatctg ggaggaatat tgatctacag 7260

gtatatttaa ggtaaattag aatgccaaca aattggaatg atgacaagga ggttagtgaa 7320

gaaatgcact ggagaaattg gatatacaat tccaagattg aagaactgga atcagattgc 7380

agtaatggaa atgaagaggt ggatggtaaa gatatatcca caaaaactaa tgctaccacg 7440

tcttagtgat ggaaggaaga gaatgaaagg gagtcctagg ttaaaaataa tttccgatca 7500

tcgagcttgg gtgattgtgc tattgtattc tattaagtct agttaaaaga aattgttacg 7560

gctttatctt tcagttaata gtatctacag ggtaaatcct ttgcttcatt cccttttcta 7620

gttgttttat tgttattttg aaattcacat gagattactt ccagaggcaa tatcgttttg 7680

cttcctttct ttgaaggcat tgaatggcaa tttaaaggtg tagtctagac tagaaaatca 7740

taacagaaac cccaacaaaa actcaaatat cttagttatc cttttataaa acaaaaattt 7800

aattcctgta tacaatgcaa gttagttatt ttaaaagtaa attctgaaga tttctgagga 7860

aagaatagaa tatcctgtat atacatacaa ttataaatca aaatttttgt ctcagacata 7920

ttttgcctgt gttgcttata attagactct cttattcttt tccacacaga gccatgacta 7980

tgcaatgtga actctctaag tgttctcttc tgagcatcta ggctagtgat gcctcattaa 8040

tttggaaaca gtgatgggag aggtgtgcag ggatcacctg aagatgtttt taaaaaaata 8100

caatgtcctg ttttccaccc ttaaagattc tgatttagta tgtctgaggt gaaggaggcc 8160

tcagcttgag tgtcttagtt agcttgggct gctgtaacaa atagcacaga ctaggtggct 8220

taaactccag actttggttt tatacagtct ggagactaag aagtccaaga tcaaggtgct 8280

ggctgattcc tgtgagggct gtcttcctgg tccacagacc attcagctgc ctccgttctt 8340

cttatagatt cgctagtccc attcatgggg tctccacctt catggcctca tgacctccca 8400

aaggccccac ctccaaatac attagagata agagcttcaa catatgactt gaggaaggca 8460

gaaccattca gtttatagtg ttgagaatgt ttctgaacct cccttgataa ttcatatgtt 8520

aagcagtctt taggaattac tattttaaac catccccttt tgagcaaatt tgtggattca 8580

tattgcaaaa tttaacaaac acaaaaacat aatctccctt tgatcttctg ttatgtttaa 8640

atatagtagc aaattttgaa aaccactgtt ttaaaccatc caccttcttt atctaacctt 8700

tacagattca tattataaga ttaaacaacc acaatcaaaa taagtatccc ctttctggtt 8760

tttcataata tataaatgat agtagcagct catcaaaata aaatatccaa cactggataa 8820

ttaatataga gtaggatata tactcagcac agggtaggac ggtggtggcc aggggctggg 8880

aggaggggaa aatggggagt tgttaattac tgggtagaac gtttcagtta tgcaagatga 8940

gtgagctcga gagatctgtg ttatgacatt gtgcctgtag ttaacatttt atagtgaatg 9000

taaaaatctg ttaagacagt tctcgtgtta aatgctccta ctacaatgaa acaaaaataa 9060

aaattactgt gatgtaggtg acaaaaatac gacccagttc cctctcatac tgcttgacag 9120 ttttagtttg ggcatgtagg tgcaatgaca gttttagttt gggcgtgcag ggtgcaatga 9180

caaaagggca ttggggcatt gtggtgtgta ctgaatccct ttagcaaacc caatcacttt 9240

aaaggccaga gaaaattcaa ttttcacaat taaaaaataa aacaaaacag aatccttgga 9300

gaagtaatga cagttttgga gtttttctag aaataaccag agaaaaataa cctgctgttt 9360

tactttggga tttaagaaat agacttataa aaagtgggga accatttttt atttctaatc 9420

agtaccttga caaaaggaaa acagaagagt ctgtttttgg agatggttct ttgaaacatt 9480

tgccttcccc tatgtggtca ccaccttggg tagcccttga agtctgctgg gactggacca 9540

tgcatgcaaa attctgtttt ttgcagagct acattattta gtgagggcca agacacagat 9600

gaatgagtac cccaaaagaa aagtggcagg tggcatttgg aggagagaga gagaaggaga 9660

gaggcagaaa gaaaaagaga aggaagaaca aagaaaacca ggcagccaaa tcagggtctt 9720

gacttttgta caaaattaga ttgaaaatca accacaacaa tccactacct gcaggggcag 9780

gggagtgggc cgctgttctg ttgtggagac acgcattgcc gctgtgatgc ccagggtgtc 9840

agtttttga tgctgtctta gttctcctgc taacatgttc tctgctgtgt tgctggggga 9900

gacgcaggac tccaggctaa gtttaagaga gaaaacatag agctgctgcc tttcctctga 9960

gcagaggcca gattacaggg aatgatgcat ccagctcatt gaagttcaca agctaaggat 10020

acgcggagga gaggcttgct tgagagggcc cggaaggact ggcattgtaa gtctagtggc 10080

aggcatgctc taatttggtt ttattactta gacgtggnaa taggcaggaa cttggagatt 10140

cccgtctctg ttctctggga aagaagactg ttctatttac aaagacttga aagggtcaga 10200

ctgaaagaag acatggacag caaagtattc ctggggtggc aatcagaaag acctggctcc 10260

ctcgtggttt atcacagcta accatgtgac ctcagacaag gcacttatct tctccaaaac 10320

tcagtttcac ctgtaaaatc agaataacag agagagagac acttacctca ttgaggcatg 10380

gtgagaaata ggtaacaatt agcctcacac ttaataaacc ctcagcaaat gttaggtgta 10440

acttgtttaa actgaaaatt tttttaatgt tttacttgtt tcaaagaaat tataaataaa 10500

acacatcatc tttgtgaaca tgtacagaag ttcgtgtttg tgaaatgtac aaaagttctt 10560

aaggttaaaa agtgaaagtc aaatgcttca ccatttctct ccctccctct gttctcccaa 10620

caaatagact acttggattc taaacttgct tccaattcat gtgggtctgg gcatgtttgc 10680

ctccaatttt taatctctgt agtgggaata atactaaaaa ctacttcaga gaggtgtcaa 10740

atgcttagaa tagtgcctgg catggagtaa gctttcaata atatttgttc tcatctttca 10800

agttaaaaaa atgagtacat aatgtctaag cttgcttcta ttctataatt atgcacagta 10860

gtttatggac tttcattcta ccctgagaaa tgatcaaatg cataaaaatc tgagcagtgt 10920

caagagtatt aatcactttt atcattggaa gtcactttgg atatcatcta ttgcaaagtc 10980

ccagacagag gagaaatctt ttctgcaaac tctaaaatgt gtagtcattc cctctgagag 11040

tatccataaa ggagtttacc acctttggg atattccaac atattatgat tggaaactta 11100

tagagtagat gttcttttta tcaaataaac agtcacctct tgataaactt ttgttagaga 11160 atttctattt tgtctgcatt ttccttgcta gaagaacaga gagcactaca aaacaaatca 11220

taacaactgt gttctgggct ttggattttg ggtatgacag aaaaactttc tcaatacata 11280

ccttttaatt gctcttgcac aatctctgta gaaaattaca tacattcttt tagccatgta 11340

tcataggtgg aatctatttt ttagttttca tattcttagc ctagtcagtc aagaaattac 11400

gagactaaaa ctaactctag gaagtgaaga aatctgtcat tttaggaaat ggagagatct 11460

gtctcttcac tacatggggt gagattatca aatgggagtt agtaaatatg tttttctcat 11520

tctttaccaa atttagagaa aaaaataatg aaggtataat agaatatttc taggattcat 11580

cctcaaggaa aaaggaataa taaatggata ataccttcag aaacttgaaa gggttattta 11640

gatagattta atgagaggtt tgatgatttc taccagaaaa aaattgacat ttttataagt 11700

gtgtggtgag aaaaatcata aagtgagtca gataataaaa atggattcaa tttcaggaga 11760

tgagaaatgt ctgtactatt tgcacaaaaa ataattgatg ccaagagttt gtaaacacca 11820

ggctatatga tttaaatata ttgaaaaagt cattttctca tttcctgtcc attctgtaat 11880

atacttaatt tttgattcaa tagttctgtc aggcacattg tcccgtcttg tctctgtgcc 11940

taagtttcct catctgtaaa atggttcttg tgaaaatgga atgacatagt gcatattctt 12000

ttttcttttt ttctttttga gacagagtct cactctgtca cccaggctga ggggcagtag 12060

tatgatcttg gctcattgca gcctctactt cccaggctca agtgagcctc ccacctcaga 12120

ctcctgagta gctgggacta caggtgcatg ccagtactcc cagctaattt tttgtatttt 12180

tttttttttt tgtagagatg ggttttctcc atgttgccca ggctggtctc gaactcctgg 12240

actcaagcga tccactgcct cagcttccct aaatgctagg attacaagtg ggagccacca 12300

cacatggcct gaaaatgtat gatattatta aacagactgg cacataataa cccctcaacc 12360

aacattttttg ttagaaggac tctgtttttg catgacagta aacgtaattt gggttgaaat 12420

caaccagtga gctcactatt ttaacatgga aactctcatg tatcccaaag cactttttca 12480

tggggtatat gtgtgtgtgc aagtgtgtgt gagctactga aggctatttt cctttggaga 12540

aagcattctg agtaggagat aacattcatt tttacangca acttctgttt ttaacacttt 12600

cccttaaatt gtcgtgatga ttgaaatctc aaggcangtc awtgagcnag cccctaaatg 12660

ggagacagca gagttgctgg attctggtct ggaaagtgcc cagtgcgtgt gagaggctgt 12720

ggaggccgtg agaggtgacc tcatggaggg agaggaagag cgtgagcatg caggacacag 12780

atgatctggg ctcagagaaa gaggggaagt tcatacactc aagatgcaca gagaggacac 12840

tcagctgggc agagggaggt gcaggtatca gtagctctga ggctcaaatg ggccctggga 12900

attggaggac ggcagagtgg gcagggatgg tgacacatga ggcaggtgga tggatgttgg 12960

ataacagata ataagacaaa aatatggggt attttttcct agttgaaacc atagcataaa 13020

atgtatgatg ctacattcat ttgctctgaa tggacatgtc ttctgtctga gaaggaactt 13080

agcaagcaaa attgacagtc aaactttaca gaaacaaggc catcttgctg ttaagcatcg 13140

ccttttgtcg tcaagttttc actgtggtgt attgaggtaa aactgaagat gtaaatgtgc 13200 atatagggaa aggacaggaa gctcttgttc aaatgtagat tttagtggtt tgttgttttg 13260

agttctgtgg ctattggaac taaaaagagg aagaaataaa aacatttgaa ataagaatta 13320

atattgtaat gcctacatta acaaattggt aacatcaaat acatgtgtct tcaaaatatg 13380

gggtaatgcc tttctcagat acaaatttaa cactgggaac ctaatctgga aagctgtgct 13440

tatatcagct caccttgtaa ggccttgcac gtataagcat gtaaatgagc tatgggacaa 13500

gaaagacaag gaaggggcac agatgtttga aggtagtgtt tacttacagt taggantttt 13560

ctgcacatgt gtatttattg acagaggatg atgttcacag aatatggcta gatttaaaaa 13620

atctggttac aaaaccatgt gtaagcagcc atttaaaaat aaatgtatgc atggaaaaat 13680

atatgtggca aagaaagtcc taaaagagat acaattcata tcctcttgga gagtaagcac 13740

cctaaagaaa attccacaag agaccataac taattacttc atctgcagcc ccttgttcgg 13800

tggcagacac aatgtatgtg ctcaacaaat tctttgcaaa aatgattgag tctctagaag 13860

tggaatccct gatggctttc atttttgttt cttattttgt ttatatagtt tttctaattt 13920

ttctacaaca aacctgtatt acttgtgtta atgaaaataa aataagaata gaattatttt 13980

cttctaactt aatttaaggt ttaagtattt gagatctagg caaagaaaga ccagtcagct 14040

ctcctggatt tttgcatgaa tctcattatg cccgatagca cagaaacrtg ttgggaatcc 14100

actgctcaat attccttgcc tagggaggta gtatggagcc agcagtaagg tactgaagga 14160

ataagaaaac cactagatgg acgttgaaga gaaagtgcgc cacactagag gctctctttt 14220

aattaatcag ctactctgcc tttaaataca gtatgttgtc ttccatgtca ggaacatgga 14280

gggaaaagac ctagtatgtt ttgttcaatg ttaaatccac gtcacttaac aaaatggtca 14340

ggatagtcaa aaaacaaaac aattatgcaa acatgaatga atgagaaact tctgctagca 14400

atgctggttg gtttcccaaa gacatgaaag tcagagtttg aggtgaagag tcccggagat 14460

gttattgaac attttctttt cttaatttcc ttctcaaact gtgccaggga tactaagcaa 14520

agagataaaa acccaccatt aagtgtgtat gtgtgtgtct gtgtctgcaa tagcgcggtt 14580

gataaacaat ttcacttgca ctcctaaaga ccccacagtt tatgcacaaa agttatcaga 14640

agagtggttg ttggtattga tgtattgaca accttaataa cacctaatga ttaaaaatga 14700

gtcaatgact tctgtcttct ttggggcttt ttagaatgaa atagatgctt tcattttata 14760

caagttgtat gtttatcttc tcacatcttt attcccaaat attaagtttt tcagttataa 14820

caaatcacac acatagacaa agtcttcatg agttaagtga gggaattagg aaagactgat 14880

gtgggattgg cattggataa caaagaggac atggcttgga taaagggttt gaattagttg 14940

agtaaaggtt gcaggatggt gaataatggt atgtttagcc aacctgagtg cagcatgtgg 15000

tatgagtgtg aaaggaaggg caggtacaaa tcagtgggac tcagggtcct ggaaacataa 15060

ggttaaataa agtaggcagt tcgtaagctt caagaatgcc ttcatattca gtgctgtata 15120

cccagcacat accgcagtgc ctgataccac agtggatgca taatatttgc tgtgtgaatg 15180

aatgaatgaa caaatgagtg agtgaatgag aaataagaag atacactgag gttttctggg 15240 ctacaaagga gacctacagt gtgtttggaa gctgattttt ttttatataa aagagggaac 15300

attttgaaag gcatcccaga cntgagggcc catggcctac ntagaaagga gtctagnaaa 15360

tggggatagg tgaaaaaggc cgnttgtaca tatatggaat gaatgagcna agaaaacaat 15420

ccactgctga gagtgggtgc aaagagaacc tgctgggcct tcccgctgca catggggaac 15480

ctgaagagtt tagagcagta gggtctggga catgagagga aggagaaaag caggatgcct 15540

ggctcatgtc gctctctcag gctgaatctt cattcatgat ggtgttggca ccgtcttcct 15600

acatctgcgg gggaagcagc accgatacct ctctaaatta gtagctggac acagatggta 15660

atggaaactg taacaacagg tgatcactta aggggttaag tagagaagac aggcaagaag 15720

ggcatgctgt gaaccataat ttggaagaca ggagaagatt caaaagactt taatactgaa 15780

cccaagagag ctgcaaggga aagaaatgac ccatagtgtt gcaagggcgc tgcagttaaa 15840

gaaaatggaa tattttaatt ttaatactgt gaagacgggg ctggaatgaa gtagggaaag 15900

tatataccaa tagaaatgaa attgctatca cagaatacaa aaactgggtc ttaaccattt 15960

aaaacagcaa tgaacatggg atgaatgagt ttcntgagtt gctataagct gtttctgagt 16020

aaaagattta ttaatggata cattgtgaac tgtgtagcct ctcnnccgat attgaantat 16080

ttaaatacca actactttta ataaatgtcg gaaacctttt gaaagccatg tttttcaaac 16140

tttgtgtact gtcttatgaa ggacactttg gtaaacagaa aaaagagata ttggcttata 16200

aaaacataat aatgataagg ctcagataca caggaancaa tttcctatca aaataataaa 16260

cttnacctag gatacttaaa tatccctact ggtttttatc agagaaataa taggtctgct 16320

gtatgcagtt ttattgattc aacatgaatg aaggcataaa acaatatttc tgtttgcaaa 16380

ggaaaattat tacaatcaac cacctcaaat taattttgaa ttcctcatgt gtttgtaact 16440

taaatttatt ttactacttg aatgctttat ataatattta atagtattaa attaaatcta 16500

gagatccagg agctagttgt tgagttttta tgtatgtgtg gattagccag gtatctttgg 16560

agcttagatc atttcttggg cactgtgaga tttagtggca cctgggagca ggtaaagtga 16620

aggtaaatgt gagtgacagc taatgtcatt cagttataat tgcactattt tggttggtca 16680

gatataacca aagtgggaat tttttaaaaa acaataaatt ctggacatgt atttttctag 16740

ataaacaaat ccataatcaa aaatgatttt tgctttaggt ggcgtattga taataaaggt 16800

gcatggaatt catatgccca ttgtctccac tgctcagtga aagaggaatt gtggacagat 16860

ctctaaaatt ccttatataa tctttgaagt gaaaccactc ttttacttgc tcctttgtat 16920

gtgtcatctt gaaggatttc taaaagattg tatctacact cagtatttgg gtggatagga 16980

gactgtcatt gaagaggcat ctcaaattga ggccaccaga tttccttccc cactaccagt 17040

acatcagaga aaaaaaaaac actttaaata ctctgtatac attaacatgg gatatcaaag 17100

agtcccaaac tccaattata taacataaat taattagtgc agcctactgc aaagcccata 17160

ttcaaactca atattcagag caataacagc ctggaaatac ccagcatgtc ttcaagccag 17220

ttagcttgtt ttcttctggg aataggaggt gaatccgtcc gttttcacac ttctataaag 17280 aaatacctga cgacgctggc taatttataa aggaaagagg tttaattgac tccattctgc 17340

atggctgggg aggcctcagg aaacttacaa tcatggtcga agacaccttc ttcacaaggc 17400

ggcaggagaa agagtgtgtg aagtaggaac tgtcaaacac ttataaaacc atcagatctt 17460

gtgagaactc actccctatc atgagaacag catgggggaa accgccacca taattcaacc 17520

cctctctcca cacgtgggga ttgtggggat tataattcaa gatgaaattt tggggggtgg 17580

gaggacacag ccaacccata tcaggagagt taggcagaga tacctgccca ggtagataaa 17640

ggagcacata gcgatgggga taggtgctca atacaggcaa tcagattcgg agctgctttg 17700

ttctgagttt ctgtgtttta ttctcttttc cagtaaaaac tgcatctgtg ngcgcgcgca 17760

cacacgcgcg cgcacacaca cacacacaca cacacacaca cacacactgc aaacatgttc 17820

ttgtactcca gtgaggacta agtggcaccc attgtgcttg gttaggtttc tttctctaga 17880

cccatggagc accagatagg agagccagtg aacccgcctg atgttaaaac tgcagtgcta 17940

cgttttcctg actcagcctc tttttttgc ctgaagcgct cttttttttc tgttacccgc 18000

cttgagaaat tctcctaata gtctgcagct tcagcatggt caactctaaa ccccttacac 18060

attgtagtca aacagagctt attttaaatt tcaagtttgc ttggatggcc ttggtcaagt 18120

tattacctgc cttgaggcta gcttttctca tttttctcat gcatgtgtaa gaagatgagt 18180

tttttgtttt ttgttttttt ttggttaggg ctgttgtaaa gatcaaataa aatatcattg 18240

ctaagtgtcc caaacatagt atatgcccat agctgttagt tcctcacctt cctttaagac 18300

tcaatccagg tgctaatcaa atgggaagtc tcacagtcga ctcagggatc atcaggtacc 18360

acctgccttt ataacaattc ttctcacatt gtattgcact tatttgctat ttgtgtcctg 18420

agggctgaca tataatagcc ccccttcaag tgactgctga gttgtggaag gcagtagcaa 18480

tggccatagg gaggtaacgt ctatcacagg atgtagatgt ttcagaattg tggaaattta 18540

agcagttagc aggcatctca tctcccacct caatgtaatt tcatgaacaa tttacttcag 18600

agtagctgga agaatttaga taatcactgc attcataaaa caacatctcc atttagggga 18660

ggttcttccc ccattctttc tgacctccct cccctcccaa gttgggagat caacttgggg 18720

ctactgttgg aagtaagaga cgtctgtgaa tgtcgtatta aattaacact ggggaaacca 18780

tggggaagat ctgtgcattt catctgtatt actcagcagt gttctagcac aaagcacctc 18840

tatgtggaga cagctttagg tataagagca tacatctaaa aagattttcc caagaatatg 18900

tctgccctaa gaagcaaccc attttactct agctattcta aatagctagc ctctaaactg 18960

cttttgaagt acataaaata atccattgct ttgggagaaa aaaatcaagc ttctgttggt 19020

attaatttat ctcatccttt taaagtttct attttcatat gtttattata cactatggag 19080

tagaacacat atatgcctta caaataagta gatatatgtg gggaactcaa ttttttgatg 19140

agggcaagat ggtttagatg acaaatacaa cttttccttt tncccagact gtggttttgt 19200

gcttgctcac caaagctaac ctcagcatgc tcaaaaggaa gcagagttcc agggtggaag 19260

cccagccagt cactgacttt ggtcctgatg agtctctgtc ggataatgct gacatcctct 19320

```
ggattaacaa accagtaagt ttcttctttg agtatggaga agcatatgtg gatgtgagat  19380
aaatgtgatt gacaaattca aatatcatgt acctttggct tcgnaaatat ctgagctgca  19440
gtagaatatg gatgaggttc agagtctgga agtagagaac aatccaatga aatagtgaga  19500
tgtcgctatg agttactcac ataggaagat gcactgggta gtctgatccc ctcccctccc  19560
cacgctgtgc accctgtcct cagagtcaca tggtgtagag atgataccac ttctcaggct  19620
gatccaaagc tggggtccaa gatgagcaac ctgttctgaa tgaacctcct ttacttgttg  19680
tcaacgttcc tgaatccgtg agacaatatg cacatgtttg gaattaaagg tctccttctg  19740
atttgcaaaa tttgagtaaa aggtcaatga gataatggat ggtatctctg aagaactttg  19800
tcaccttcac cgcctaatta agttgatgag gatttgagaa gaactgctag gccaggcacg  19860
actgtgtgcc agaaagcata gtaaccagaa ctaagcctga caggcatgca tatccatgag  19920
tgctggaatc tcagctctgc aaaaccaagc aggatcaact tcttactcat ctcatggttg  19980
atgctgaagt aacaaaagga aatggcctgt tccctcttgc agtgcttgtc cttcccccaa  20040
ccatgataca cacatgatct gcttggactc tttgctgtcc ttcccagtct ctcttcctct  20100
aatctgccct ccaggttact ttccctgcct gctcacattt ctttctatat ttctattttt  20160
attcacatca cgcttctgct ttataatcat ggaagcatcc ctattacctt cattgcaaat  20220
taaaatacgt tagtgtgaca ttcctgcttc ttttcattca ggacctaacc tgtccagcaa  20280
ctgctccttt cctgtgtcag cacactgtcc acttgagccc catcagcagc ctcctactcc  20340
tctaggctct ctggattgtt cacattgctt tccctgctcg tactgctccc tggccccagt  20400
gaccctttcc cagctcctct gatgagtcaa atcctcctca tccattagga ctggatgaga  20460
tgtcacctct ggcatcctcc tctgaaaggc ttatttccgc cacagacaga atgtacctct  20520
cttctctttc cacttagcct gttttatatt tgtatgtaaa tataaagtga tctgtaaata  20580
actgtatttt tatatgtatc atatattatg tgtatattta cattatatac attgtaggtc  20640
gtgtgtatat ttatataata tacatatatt tatggatcta ttgtcttagg tatatataca  20700
tgttatatac acatacatat aaacatataa caggnttata tgtatatcac atatatacac  20760
attacacata cacatatatg ttatatgtgt taacatatat gcatattacc cacacatata  20820
aaatatcatg tgtatataan catatataca tgtcgcatat gcatatacat atatattata  20880
tgtatgtata atgtgtacat atatttatat gtgcatgcat ataacatgta tatgtaatat  20940
gtatatatgt tatatacgtg agactataca tgatatatgt atattatgtc ttatctagaa  21000
catgttttat aatgttatat aatgtgacat gtttttgcct cttgcagtgc ttgtctttcc  21060
cctaaccatg atacatgcag gatctgcatg gactctttgt tgtccttccc agtctgtctt  21120
tctttctcta atctgccttc cagactacat gtgtttacat gtatgtgtat ttacatgaaa  21180
cataatatat acatacatat aacatgctat gtgtatatgc acacattata tatatgtatg  21240
tgtgtgtata tatatatata tatatatata tatatatata tatatatata atacattcat  21300
aggtgtttta cttcttggac tgcaaagcag gagtcagaag tgtgttgctt ccatattttg  21360
```

ttctccattc ccagcattct gacacatagt aggccttcat tcatacttag agaattgaat 21420
ggaatagaga atgaagcacc cctaactttt tcttagtgtc catcttatta cagagctttc 21480
tacaaactta tgaagaaggt agtataaaag atgtggcctt ggaggcactt actaaaatta 21540
ataattcttc attaaaatta tgtctttaaa ctaaaaacac ttttgttgtt gaagaactta 21600
tatttggaag ctcttttgta taattagatt ctttatgttg aaagtccaga tggggctaaa 21660
ttcacacacc aatgtttatt gaagtgcttc caaattgcat gaaaaataca tcaaccttta 21720
aagtaaactg aatcatcagt ggcttatgtg gtatggaatt agcagcaaaa tatcttcatg 21780
taaatattat aactgtcaga taaatattat aactcctcag ctgcaagtgt ctgagtctcc 21840
tgaaagttct gacagtcttt tctanttttt catcttgagt gtttngctag tatatatagn 21900
ctgcagaatc catctactac aagactaaaa tttttttcag agtattatat taaagaatta 21960
tcataatttc atatatgtcc cagatagtca ataaaaacgt atttggatta tgtttaagaa 22020
ttgtatacta attgattaac tggtaatatt aattgtggtg ggtttagcat ggagatgcag 22080
ttaagtattt ttttaatcct agatactagt tttcataaaa aaatattgag tatgagtgat 22140
tgtatgtcct gtatatatat tcatgttaac acacacttat ctttaattct ttcttcttat 22200
tttggggtta aagaaaatct tcagacagca gacatgcctg tagtgttgca cagaaagtta 22260
ttttgtggaa ttttgaaaga tgtctgggct aaatttgtat aaaatgtgtt ctccaaaata 22320
aagagaaaaa tctatgcatt tttttcaaat gtaggatgag accagatttt taaaaagttt 22380
ctttgtattt ttatacaaag gcattatata tcagatacaa atagaaagga aattctaaat 22440
atgggatgtg ttatatatca attcatttaa aaagcatttc ttaactacta tgataggccc 22500
ttgtgctaga tgctagggaa gttatggtaa atatagtatg cccctaattc taggtactta 22560
tagtacaaaa atgaaaaaaa agtaatggca catcaagttc ctatgtgact aaattacaaa 22620
cacatatttt aattgaacca gaatcttcct atattgtagt ctgatgccca ggaggttatc 22680
agatgtttgg gagagcctga tcaggttccg tctatgcttt gggtctttgt aacgcatggc 22740
cacagaaatg cctccacctt tctgtcttag tgtgtttatg tggaaaatag gtgccttgag 22800
aattaaatga tatnaaaaac gtgaaaatat ttgtagcaca ttatatgcac ttagatgtta 22860
gatttcttct tcttcacctc catattagta aactatttta acctcagtat tatgttttcc 22920
acttagtgtg gttcataata caagtttgag atatcagtta tggtgccgtg tagtaactct 22980
ggaccatcat tgcagtgggt tcactctttg ctgcgcatct gtgccatcat cagcgtcatt 23040
tctgtttgta tgaatacgcc aatgaccttc gagcactatc ctccacttca gtatgtgacc 23100
ttcactttgg atacattatt gatgtttctc tacacggcag agatgatagc aaaaatgcac 23160
atccggggca ttgtcaaggt gagcacttcc atgtcattta aactaagaac ctaaatgtgt 23220
tgaaagtctt aatcccttta atcatatttt cttgttcatt tcagtaaaac aaaacatcat 23280
tttatcttat acatatggaa gtccttttta aaaaggcata actccagaga ctgggagaaa 23340
ttgttgcatg agactatttt aaaagttaaa taaattttgt aagtcacaaa ctcttggaat 23400

atctgagaag gaacaagatg gtttgactaa tgactgtgtg aagtctcgag tagggttatt 23460

catgccactt agggctgagc ccacaaccca gcacaggggc tggcccagag aagcacctta 23520

ttatttgttg acttagggag tgcagaatca gagtcccatg atagagttag agcatgatat 23580

tattgcagca cgatgtcatt tacttgatgt gttttgaaaa tacaaaagga taaaaaaaaa 23640

aggacatgag atggaaaaaa tgtctggggc caaaaggaaa attatgtgtc atgtgaaatt 23700

agtaccaatt ggaaacaaaa tagatataaa atgttttaaa agggtaaaga aagggaaaga 23760

aaaggaagaa aatttctgtc actggatatc ttttccttcc atagcactaa tctttacata 23820

ggttaaacat atgtagatga tgctgtccct agagtattta aattgtacta tttggggaag 23880

cagtaatagg agtgcagatt ttagagtttg acaaatgaca gacgactcag atcttagttc 23940

tttcacttac tggnttgtat cagcatctca agttagctaa ttttggtgag aaagtgtttt 24000

attttgtaca aggattttga gaatctactg agacaatata taacacataa ttatttgtca 24060

atttcattgc ttccctttta tacataatta attgtgtgcc atccttgaaa gacagagatc 24120

acagacaatg cattaatagg ttcttgttta actttttatt gtggacttta aaaaatttac 24180

cccaaagtat attacagtta taatgtgaat ccctgtaact acacatcaac aattatcaat 24240

atttagccaa ttctgtttca ttttaaagca aatctgagaa accaagtcat ttcagaaata 24300

actccttttg tattggtgac tgataaaatt ttctcttttt ataaccacca taccattatc 24360

aaatataata atattttacg tatgataaaa tatttgtaac aagacttagc aataattcct 24420

taatgtaggc taatacttaa cccaccgttc atatttccta atattgtcac aaattttttt 24480

tatggttggt ttcttcaact caagactcag acaaggttta catagtatac ttgattgttt 24540

ttgttttcgt ttttgggagg gaggttcact tttgttgccc aggctggagt gcaatggcac 24600

gatctcagct cactgcaacc tctgcctcct gcattcaagc aattctcctg cctcagcctc 24660

ccaagtagct gggattacaa ggcctaagcc accaccccgg gctaattttg tattttcagt 24720

agagatgggg cttctccatg ttggtcttga actcccgacc tcaggtgatc tgcccgcctc 24780

ggcctcccac ttatgtttct taagtctcct ttaagtctct gtaagtccac taagtttcag 24840

cttttcacat cctgcaaaaa attttcctat gtcattgatc cattggaaga catgcgtctt 24900

ttgtcttgtg gaagattcta catttgagat ttggctgatt gcttccagat ggtgctcttg 24960

aatagtttnc tgtgtactcc wtgtncctct gaactgcttt ngtanntttt gaccagncaa 25020

tgcacacagg tgcttctttn cccacatggt cacctcnaaa atatgtcaga cttctggata 25080

tttgccagtc tgaagggtgg gaaacggtaa tccactgtag ttttaatttg catttccctt 25140

ttgagagtga ggttagatgg gattttatgt atcaagagcc atttttaagt tactattttt 25200

tcattcaatt ttggccactc tgtgtgtgtt agaagtatca gctgtttgtg atataagttg 25260

caattgtttt tcccagtttg ccattgtcat tggctttgtt tgtggttatt tgtaccatac 25320

acaattctta tccaaatcat ggacaaactt tctattgtct gcattttatg tcatacagga 25380

gagttttgta ggtttcctca tacagatttt gcacattctt ggttatattt atgctttggt 25440 attttattta atcattttgt tatttgtatg aggtcttcta ttacatgttc taacttgttc 25500

ttttaaatga aggctgttgg tttctataca ttaatattat attctgctac tttactaaat 25560

gttctagatt cttgtattag tcataatagt tttcctgcta attcttttga tgcttccata 25620

tgatcatatc atattcaaat aaagaaagat ttgtcttttt ctttgtaatt tttatgccct 25680

agttgctttt acttacttaa agtaaaataa atgaattgac taacactgcc ggtaaaattt 25740

taaatactag tagaggtagt gaacattttt attttgtttt gtttttggtt tatctagaac 25800

atctctagtg tttcttcatt aactagttca aaaaattggg atttgaagta tgtgtacatt 25860

attggattaa gtgaataaac tttaatttca cattttgaga gttgtagcac aactagcttt 25920

ttaatttgtc aaatgtttat ttttaatttg tcaaattatt attttttcct tagtgcttgt 25980

aagagaagag atttttaaat attgagttat ccttccattc cgataaagaa aaatccctac 26040

tgggtcaaca tacattttta ttctatttta atgtgttgtt ttgttatata tgttattatt 26100

ttccttattt taaatttggt ctaatctttg tcaggttttg atattagatt tttcatgtgt 26160

gtgggtggcg ggcatgtatg cacactataa aatagtctga gcactctcag aattgttttc 26220

actttaaaag tttggtatca gttcccggtg aaatcatctg gtcctgttgc gtttttgtgg 26280

gggttccttt gcagtatttc ctatttcttc tatggaaatt gtctgtttta tttttatttt 26340

tatttttat tttttgtttt ttttttgttt ttgtttttgt tttttcgaga tggagtctca 26400

ctcttgtcac ccaggctgga gtgcagtggt tcattttcgg ctcactgcaa cctctgtccc 26460

ccaggttcaa gcgattctcc tgccttagcc tcccgagtag ctgggactac aagcgtgtgc 26520

caccacgcct ggctaacttt ttgtgtgttt agtagagatg cggtttcact gtgttggcca 26580

ggatggtctc aatctcctga ccttgtgatc cacctgtctt ggcttcccaa agtgctggga 26640

ttacanggcg tgagccactg tggcagacct atgtcttcat tttttagtag agacaagatc 26700

ttgctctgtt gcccaggctg gagtgtagtg gtgtgatcat agctcactgc aacctccagc 26760

tcttgggatc aagagatctc ccacctcatc ctcccatgta gctgggacta caggtgtgtg 26820

ccaccatgtg tggctgattc ttaaaacttt ttttagagac agggtcttgc tatgttgccc 26880

aggctggtct ccaactccca gactcaagcg atcctcccac ctcggcctcc caaagtgctg 26940

ggattacagg cgtgagtcac agtgttcagt caaaactgtc tgtttcagac ttgcttggtc 27000

tattggggtc tatttggtga agtatatttt cttaggaaat tatcccctat ttattttgtt 27060

taattatttg gtaacataga cctgtgtaaa tagcttttct ttcctctgtt tgattttctc 27120

ctaattttgt attttttgt tattttggtt agcattttgt tgtcatcagg ttggctagca 27180

gtattttttg tttgttccat tttgttttag gttagccaat tggtttgctc attttgttga 27240

ttttttttaa aaagaatcag cttaattaac ttattaattt ttgtttaatt tttttccttt 27300

aaaaatattt ttattatact ttaagttcta gggtacatgt gcataacatg caggtttgtt 27360

acatatgtat acatgtgcca tgttggtgtg ctgcacccat taacttgtca tttacattag 27420

gtatatttcc taatgctatc cctcctccct cccccgaccc catgacaggc cccggttntg 27480 tgatgttccc cttccntgtg tccaagtgca tttattagtt tctttctttg ctttatttgg 27540

tttactttat gggtctgttt ctaacttttc gaaggtgttt gtctcacatt ttcatttctt 27600

gagactctta aatctgctgc ttccttgcct cgtggttaaa atgttaacca caaggataat 27660

gttaaccagc agcataaaat gttactgctg agaatgtttt ttttccctcg aagtggcttg 27720

ctcttttctc ctgaatacgg aaagagactt tccctttatt gttcaagtca gcagctctat 27780

gagaatgtat gttatctttt tgttgatagt tacggatata ttttcctagt ttcatcatgt 27840

gatatttatt ttattgcatt gttttcagga ttccttttgt attctaatta taccaggatc 27900

tatggtttgg tctttccttt cctttgtttt ctttccatgt atttccctct gtttttatgt 27960

tttctacatt ctctttccct tgcttcactc ctcatgctgt cccctcactc ctctgtcccc 28020

tgtatttcac tcctctttct caaatacttt tttcatttct ttttcatcat ttcctgaaca 28080

gtttttgttc tttcttttt cagtctctta tcatttagac atcatttctg tgttttctta 28140

tttctagttt tgtgattctt ttggagcttc tatttgttta aatttctttt ggtcatgttg 28200

tcatattcga ttatagtctt catctgcttt gagtgtttat tttctgttta tacgtttatg 28260

tattatttta tttttacttt tattattttt tgagacagaa cctcgctctg tcacccaggc 28320

cggagtacag tggtgcaatc ttggctcact gcaacctcca ccttctgggt tcaagcaatt 28380

ctcctgcctc agcctcctga gtagctgaga ttacaggcac gcaccaccac acccggctaa 28440

tttttgtatt tttagtagag atggggtttc accacgttgg ccaggccagt ctcagactcc 28500

tgacctcagg tgattcgcct gccttggcct cccaaggtgc tgggattcca ggcgtgagcc 28560

accatgcccg gcctatgtat tatttcattc atcttattct gtaatgtttt tacaagggct 28620

tcatcatgat ccgtttgttt ggttgttctt cttgttgttc acatgaaatg ggagtttttc 28680

ccacacattg ggtagagtag cctctgcata ttgtcagtac gaggggtccc tccactgttg 28740

ttatcctgat gggtaaacat ggagcttttc tgtgtaccga ccgtcaggtt ctgagctctg 28800

ctgcctctgg actgctggcc caattatcga atatttctta cccactggcc ccagtgtccc 28860

tgccctactg aattgccatt ctaaacatgg ggctgtggtc tggaaagggt ttttgttggg 28920

tgtcctttgg agtttctaat attccattgc ctctgatctt tctgcagttt cttggggaca 28980

ttggtagtca tttgtttttg ttttgttcct ttgttttaca ttaacctgca acttgcagcc 29040

taccagttta catttctaca aatgttgcca tttattttaa ggttcacctt ctgagatatc 29100

tgctcagcat tccacacgta tttattagtg attttggggt tttggaaatc tactggttta 29160

tggaagaggt gggcatcctc catagcttca atgaacagat ttcaggatga ttccaaaatt 29220

gtctgtatct tccacaatgt ggctaggcta gtttttataa gcactaggtt tagataangc 29280

ataggttcag aaaactggaa cctgtctcag gaggccatct tgggagaaag ttggtcttgg 29340

gagaatgcta acatctctgt tattcacagt actcccctgg acgttgcccg gccctagctc 29400

acattctgtg tagttgcccg gccctagctc actagagggc catgaatggc cctgtcactt 29460

gtgtgtagtg gttattgggg ccatggcaac ctttgcccat gcattgaaag atctgaatat 29520 tctgggtgaa gcattttatg attgaaattg ccatgtgttg acttcttaag ggttccattt 29580

gaatgaaaca aataagtgac atgttcctac aagaaaactt ttgttttttt aagtcattaa 29640

ataaagactt agatttcaca tcactaaaat gatagtaggc attctgctag ttgtcacagt 29700

ctaatgataa aaagaacaga ggcatactta atatataaat tatatactca ttccaagtga 29760

acagagcagt gagagaacac aagttttaat aatgtaaatg ttagaagttg aatcaatagt 29820

taaaagagaa agaaatgact taggaccaaa tcgttttcaa aatgctgtaa attagccatt 29880

tccagtattt tcacatttcc atcaacattt gggtgatgag ctatttacta caaccaaaaa 29940

tttttttttt tattttcaat ctgtaagagg tgaaatttaa aatgaaaatc aaataaggtt 30000

ttctggtttt cattcagaaa taattggatc aaatacaact gtaatgcaat tttagtaata 30060

ttgcactgtt aacattatct tgaagtagtc agctcaactt catagctcag taaaactggt 30120

attctatgtt taaatgctag ctaacgtttt agttaaatat agttaaatat gaaatctgtt 30180

tatcctgaac caagctttaa atgtatcaac tcactaaatt tttaccatca tggcctctat 30240

tattgttaga gggtaataat tatatcagta agtactgtca ttttcatgat attaaatcta 30300

agccacagtt ttattatttt tatactacag acttcttcaa acatgcatgt cataacagat 30360

acagctcaaa gagtatattg ggcttgatat tattttctta tattcaaaac ttaaactttt 30420

aataagactt tattttagca aaatgtcctt tgtcacatta aatgtatgtt cctaatttat 30480

attttattag ctttgcattt ttgtatttgt aaatttgttt tgcttttata cttacatgag 30540

aaatttgtac atattgtagg tgtttaagtt acatcataat aaaaacacac tgtcacatcg 30600

aagttctata cctttcaggg gccctctcat attactcaag tttgagttac tatgctttct 30660

tcaactctac cataattttt aaaacttttg ttaagcactt gtctgctgaa aacctcatca 30720

tttttcatac tttttctacc ttaatacatt tgccaatagc actgcagtta tctctcaatt 30780

ctttagtccc ttncacccag aataatttgc aaatatgtgc tgcattttct cactaacatt 30840

taagcttgta ttttttcata cataaatagg aacgttatga atatacatgt tctattttaa 30900

tcttcaggtt acacttttca tagtaagttt tcacattgtt tggaacagag ttgtgcaaat 30960

atgaaatgtt taatattcac tgatttccaa ctaaaatatt tctaaggaat agtcaaattt 31020

agttattgag tgagcagtac ctgttagcca tgtgttctgg gtggaaaagt gtgtgtgagt 31080

tgatgtacac atatgccagt gatttgcgaa atgctaacct gntanttttt tttttttttt 31140

ttttttttga gacggagtct cgctctgtcg cccaggccgg actgcggact gcagtggcgc 31200

aatctcggct cactgcaagc tccgcttccc gggttcacgc cattctcctg cctcagcctc 31260

ccgagtagct gggactacag gcgcccgcca ccgcgcccgg ctaatttttt gtatttttag 31320

tagagacggg gtttcacctt gttagccagg atggtctcga tctcctgacc tcatgatcca 31380

cccgcctcgg cctcccaaag tgctgggatt acaggcgtga gccaccgcgc ctggccaacc 31440

tgttattttt taaagagtct tttccatgca taaaactatg ttagtccata agaaaatatt 31500

tatagaagca aaccaatgta ttttttgccc agctgtcatg cacagttctt taaaattata 31560 cctagccaat tttggcaatg ttatggtgta cctaagagtc atacctaatc gttaaagtag 31620

tgttatccaa tgatacctag tttataaaat tatgcacgac aaaccaaacc cagaaacttt 31680

attttttcaa aacaaacgtg tcaaagctgt ggcttcctag cattgcctac taagtttatt 31740

gacctagtaa tttgccaaat atattaaaat atgtataaat ataaatataa aatatgtaag 31800

atatgaaaat gtttaagata tttaagatat aaatgtaata taagaatgag ttatattctt 31860

aaacccaagt acagtttcca agttttcctt tagtcatcca gcaaccattt attgaacact 31920

tacaatgagc tcatctctgt ccttggagaa gggaaacaga caaaaccaag cagagctcct 31980

gtcctcagag aaaacctggt tcactctgtg ggacagactt agcacataca ggcactcttc 32040

attttcaaca tccaatatat ttctgaaaac atggtggata gcatattttc agaaaatagg 32100

cacttctctt tcactaattt atatgggaaa actaataatg tattagggga gctcgactga 32160

tttctccaaa tatcactgca ttaccatcta atacctatac aagtcaacct gcatatttct 32220

gcaaaataaa aagcatttaa aacatccctt acccaattat ttgcactaat caaaccacag 32280

tggatgtttt atatgtgccc aacaataaca ggaagtttct gttattaagt aaatttaaaa 32340

aatacattaa agtaccatcc accatatttt attgaattgt tgatatcatc aattgtaaaa 32400

aaaaatggca tcacttcaca gattcacggg tctctaagaa ataataaagg cttctaatta 32460

tactttgact gtaagatgta agattttaat tgaagcatca attagaagat gcatgctgaa 32520

tccagaattg tgcaactgtg ttggagcagg gtagggtgtg ttttagaatt ggtgaaatat 32580

gataaaatgg gcctgggttg gtaaactcaa tttttggatc ttttgaaagg ctaactttca 32640

ggggaaaggt aagacagagg aagatatcaa tgcactagta ttaaatagtt aagtaggtaa 32700

acaatgctgg tacaagattg ggcccaaaag gtgaacaaca gagatacacc tgcattgcac 32760

actatacctt tggcagaatc cactggcaac tttaatgtct ccattggcac acgtgcaaag 32820

tcattatggt ggttcttcta ggttcatagt gattctgggg caaacatttt ttgcagatat 32880

aaatagcaaa tattagtata atgagaatct acatagagag ggaccttctt tttttttatg 32940

aaagggcttt aatagtggtc tgtaaaagct atgggggagg caggcataaa aaagaaaaac 33000

accaaaactg tctgcttggg gatatcagag aagtcttgaa gaaagaggaa cgtgaaacaa 33060

tgccttgaag gatgagtaag tgtggtctgt ggcaggggtg atggaagtaa gatagaagta 33120

agggatctca cggagaaaga aaaggaatat gaggcacaaa ggagtggaac agcaaggtac 33180

atgccgagct caaagtagtc agaagtggca gatgggctgt gcacaggagg cagctggaga 33240

ggagtttgtg aaactagacc tggctgcatc atggagggcc cagcaggctc tgcaaggagt 33300

tctgactttg ttctgtggtc agtaaagatg catggaggat tttcatctga gagtggtaca 33360

atctgatttg tgtttgaata gttcactctg gaagttacat gctggatgca cagctgggga 33420

gagaaactga agcagaaagt ccttccagca agaaatgtct ttaacaattt ttgttggagg 33480

aatgcatttg ttttaagtgt tttccatctt ttataatatt cccaaagctt gaaggtgctt 33540

gcccatcaat tcccactcct caacagcttt cttctaattg cctccctaat tcctcttgct 33600 ggaagcattt tcactcacca aatccctaga ggtgaaaaga aactatttgg atgatacatg 33660

tctgtgttgt tgcagagcac ttcatgcatt gcagactgtg agttctttag accattctca 33720

gttgtctatt tcacatctgg caatagatac aagtaaaagt agcatgcatc ctttaagcac 33780

tgttgaagta atatataatg atgctaagac atagtgacag atgggccctg gaaacttggg 33840

acaataaatc ctgttctcat gaaactgtac cagtgcatat cattttgtca ctgcaaaata 33900

tatccctata aatgtgtttc ctttttctgc ttgttaaata cttgcatgat atgttgttac 33960

tttcttacct tatcgttgat tggcagtctt caagtgttta ccgttataaa aattgacata 34020

tacagctctt gaagcactcc tctcctcagc caccaaagtc ctgatttttg cgcatagctc 34080

aagcactctc tgccctccat ttcttgttga tggaacaaac aaattaaaac attttttcat 34140

atttaaggtg ttaaaatatc ttgagtacca ctgggcattc atatgtgtta gcttctggtt 34200

gcatatttta gaattgggtt gcaaatatta taacttgagc ctcaatacag tagaatttta 34260

aaggatctgg ccatgcttat aaatggaatc aaagcaactt tgactactat ggctagggct 34320

tctgacgata atttgataaa tgtgccagat acaaaacata atccaagatt gtattgcaaa 34380

gtgtcagatt tcaaggtctg cattctactt atttaaaagt attataaact tattgtgcct 34440

tagcaattat tgataatgat taatatgata gtaattttaa ttcctaccag ttcaattgta 34500

tttacaatca tggtnaatgt ataaaatatg taaaccgtaa nctgtatttg catggcatgt 34560

ttnagttctg cnttatgcaa actatgctaa tccaggaagt atccataagt catctcataa 34620

tctcatctgt gatctttcat ggttacaaag gagccagaaa ggaataattt ctttttagaa 34680

gcatataaat aatttcagtg aatagatgga tggatggatg cattcattaa gaaatattaa 34740

ctgacatctt gctgtataca ttgtatatta actgacatct tgctgtatac actgtatcta 34800

agacatggtt catgtaccta tctcatcatt catttaatat ctgctatatg tcacccactg 34860

tactgggtgc tgcaaacaca accnttaaat tatccgactt aaagtctcag ggtgnttact 34920

catcttcaga atacactcaa acatttacca tgagagatag acttcttttt tttttctaa 34980

tgtaactcca gagaatacgt tcagagaaga tttaagtagg aaactcaaac ttcaccatca 35040

gatccaataa gttggagaga ccaaatataa attcatgaaa taaccaatat ggattacaaa 35100

ggcatattct gcgagtaaca caagatgcac attatttata tacatgagct gagttgtcag 35160

ttgttgcatg acgcctgagt taaggtgatc agaggaggct gtagtcagat cactctggaa 35220

ggaaaggttg gcacagactt ggcagaattt ttcgggaaga tggacagcat gaaaaacaat 35280

gaacattaaa aagccaaaga ttttttacaa atggaaaaat gctgagggtt gctggccttt 35340

gtatcttggg tcataactct gcttataaaa gatgtttatg aatattgaat gactcttgga 35400

aacatgagcc aaaaagtaat acatgaaact tatgcccaga ggattatttc catacttcac 35460

tgggaaaata aaaagaggtc aattctgatt gaccactgca taccatctag catgtgctga 35520

gcttgttggt gacacaggaa tgccactata tgtttctgaa aagtccaagg cagaagacat 35580

tccaaacaaa acaattaaat tacttcctat ggaaaagaaa gttttctaca ttacatggta 35640 gaatttttgc taagccagtt tttataactg aggtcgttcg aaattaagat gtttgtatgc 35700

ttctaatttc tgatcaccac tggtgcttct aaaatgattt aaaataattc attattttat 35760

atagtttgaa atgtagtgtg ggccagtttt acaacattaa tttccttagc aaagaaatca 35820

gaaacaaaaa ttcagtcaca tctgttattg atagcttttc attctgaatt ggtgaggccc 35880

caatttttaa atttaaattt atttttgcca actggagaag tgattttgtt aaagctaaaa 35940

tgtccaactt aaaaaaaaaa ttgaaaagct atttttaaaa atttaacatg taagcattgg 36000

tccaaaaaca tatcagaata cattgtgaaa agcatgtctc tcctccacct ttggctcctg 36060

gccattcact tgcccgcccc cctccccata acacaaaggg aagtatcact gtttatcctc 36120

atttgttaac attgtaaatg tacctcattg taatggtgaa ctctcattca atgttccacc 36180

gagtttaatt catttctaaa gcaaattatg agcaattaaa aagagggtaa ttgtagtgct 36240

gctacatata acattgttaa tgacaaaaat gaaataaaat agtaatactg ttttgaaaaa 36300

gtacaatata aattcatcct atttgggtgc ccgtgtactc aacgatctat aattactctt 36360

ttcctgccta atatgtactc attcattgct cccaatttta acatagttct ctgctgagtc 36420

cccataaaat ttcagttttc gtttttggaa ttctttgcct ttttaaaaac aagtgacaac 36480

tctgcctagt tcttaaaata aagatgaagt gatacttttc tttttttttc cctaacacac 36540

aagtaaaaaa catttgttca tttaataaaa gtccatttga aaatatacac ttttattttt 36600

cttcatgaca cttgtcaaca tataacataa tgaatatttt atattttatt tcattcagaa 36660

tgtaaattgc ttaaagacag aagcgtttgt tttgttcacc ctcgtatccc tagttcttga 36720

acactgccta ggttaaaggg atgcttatga gctgagggaa agaaggaagt aagggaggga 36780

gagagggagg aaggaatgaa gggcacaanc tttaaaataat tcatcttcta ttanccttct 36840

tgggtattaa tcaaatttag gctcagaaaa agtcaatttt atatgggcta tgttatctgt 36900

atttagatta taattaattt acattttcaa ctgtgcaata tacnaggcaa gtatgatgct 36960

tatagttaaa gatcctctgt gagatcttta acacaactat taaaaaatgc nactataaaa 37020

aaagttgaag atgactataa aaaatcctta cgttatttgt tatatatgta acctttgtga 37080

tgcgtttatg tgtagcacat aaacaccagt agcatcggta gcatcctaag atgcagcttt 37140

tctcttcctt gggtataaat cttcatatga tgataatcac acaatttaat aaggccattg 37200

tccttcttgt cacccagagc caagcttcag agtttgcttt catatctctt ttgtgagtct 37260

tcccaaaccc aggagtcacc aagtttggtt gattaccttg cccctgaaaa cattatttgt 37320

ccgttcattt gggaattaac attggtttct ctaaactccc catagaatta atcttaccat 37380

ctttcccacc atgtttatag ctttaagtta ttttgcttta tatactagct ggatgtgttc 37440

acatctgttt ttctacaaga ctgttctttt ggttgttggt ggtggtgatg gtggtgtgtg 37500

tgtgtgtgtg tgtgtgtgtg tgttagctag gcattatata tcctctaatg taaaacaggc 37560

acttacccag tgttcattaa gttgagttta tacttcattt ttgggtcaaa taacttccat 37620

taagtgtttt atcatagtgt tagttttgca tacaccataa actatttcca gaaacctctc 37680 aaggagatat ttggatactt atttgttact aaaggtagag gcttatgtca atatcctgca 37740

acatgtttta tcatttgatt tgaaatcttg acattttatt tctattttca tttccctagc 37800

aggtttaatc agaggaggta atgatatatg tgactgaatt cagaattgac ttacatggat 37860

gattacagaa ttcatgattt aatgggatag tcctaaaaaa gcaaaggctt ttctgtttaa 37920

agtttctaat ttttaacctc atgaagtttt ctgttttatt attccaaagg catttttttaa 37980

attcccatgt tattaaagca ttctaatgtc agaatagata ccctattaaa aaatgaaagt 38040

ggtacatctt tatctttctt ttacatatgc aatcccaatt aaaattaaag tcaatatgtg 38100

ttttccttgc aaattaagac aagtttcttc ttcacttttt tttgttagtc ttgaattcta 38160

tttttattac actgtggtcc aagagaatgg ttgttttgat ttcagttatt ttgcattttc 38220

tgagagatgt tttatgtctg attaagtgtt gatgttagaa tatgtgctat gtgacaatga 38280

gaaaaaatgt atatatatat atgtgtatat atacatccac atatatacag tggaaagttc 38340

tgtagatgtc tatcaggtgc atttgatctt tgggggtttt tgccttaatg atctaatact 38400

gtcaaggtgt tgaaatctcc caatattatg cagaactgta agtctctttg aacatctcta 38460

agaacttgct ttataaattt gggtgttcct gtgtaaagtg catatatagt taggatagtt 38520

gggtcttttt ttattattat aaaaaatttg cagaatgtgc agttttgtta cacaggtata 38580

cacgtgccat ggtggtttgc tgcaccatca acctgtcatc tacgttaggt gtttctccta 38640

atgctatccc tctcctagcc ccccacctct tgacaggccc gtgtgtgatg ttcccctccc 38700

tgtgtccata tgttctcatt gttcgactcc cacttatgag tgagaacatg cagtgtttgg 38760

ttttctgttc ctgtgttagt ttggctgaga gtgatggttt ccagcttcat ccatgtccct 38820

tcaagaaaca tgaactcatc ctttttttatg gctgcatagt attccatagt atattgggcc 38880

acattttctt taaccagttt atcattgatg ggcatttgga ttgattccaa gtctttgcta 38940

ttgtgaatag tgctgcaata aacatacgtg tgcatgtgtc cttgtactag aatgatttgt 39000

aatcctttgg gtatatatcc agtaatggga ttgctaggtc aaatggtatt tctggttcta 39060

gatccttgag gaatcgccac actgtctttc acaatggttg aactaattta cactcccacc 39120

aacagtgtaa aagtgttcct agttctctac atcctctcca gcatctgttg tttcctgact 39180

ttttaatgat caccattaaa ctggcgtgag atggtatctc attgtggttt tgatttgcat 39240

ttctctaatg accagtgatg atgagctttt tttcttcttc actttttaat cagcttgaag 39300

tgaatgcctt gttagttttc ctactaacgg tgctaaacct agcagtttta aagaccttcc 39360

tgttgaccac agtgcactct tggttcttgt aggcagggag ctggcttagg acctacatag 39420

catatttgga acaatggctt ctggttcttt aagtgttaga atttgtgtgg ctgagaaaaa 39480

aaaaatagtg atattttccc tttgcatgaa ttaaatattt tacctgcaaa gttgcttgtc 39540

ttttataaat tatttttctt ataagtaatg ctacttaaaa tgctaaactt attccttgtg 39600

aaagtctgtt ctattatgat acaaaaattg tagatagctt ttattgaaat aatatttgac 39660

attttcctcc acagattgtt ctttaaaata tttgctttgc ctaataatgc catgtaaaat 39720

```
ttatttttag ggggatagtt cctatgtgaa agatcgctgg tgtgtttttg atggatttat  39780
ggtcttttgc ctttgggttt ctttggtgct acaggtaatt aaatattttt taaattacaa  39840
gcaggtattc tccaaaatgg gnaattattt tgcattaaaa ngtctaaata gaaantatta  39900
canaaatata tatcaaatgt ngattgtgta attgcctaca acattttnat tttnnatttt  39960
tattttaaca gatgtgatta gtctgtttac tctgacatat ttgaaatttg aaaaattaat  40020
tttaattgat acgtaatact tgtatgtatt tatgggggac tgagtgatat ttgatgtata  40080
caaggtgtaa tgatcaaatc tgcataatta gcaattttat atcacctcaa acattttatt  40140
gttagaaaag tcaatagctg taatatctta attattgtca gttatgattg tgaattattt  40200
ctcaatcaga aacgaaaatc aatactatag aaaattatta gctgttaatc agatataaaa  40260
tatattccct gaaatgaagg aaacaatgtg gagcaatatt tctgtatttg taaataaaaa  40320
attgaactta aaaatttaaa gtctagtctc cccatatttg gataggtatt agccaacata  40380
aaatactata ttgcaaatat taaatgcttt attataagtg atggggaacg tacattaaaa  40440
agacttacag gacagatttg attggaaaat aagtagaatt aatttctgtg cattcattca  40500
gacatgattt tgccttttaa ttcaactaaa gttactatat tctattgatt tgatttgatc  40560
tgttttgatt tggttgtggt aagaatgcct aacatgaaat ctactgtttt aatgaatttt  40620
aattgtacaa taaattattg ttaactgtag gtacaatgtt gtacaggtct ctagagttca  40680
atcatctatt ctttatttaa aaattcagaa ttcaaaactt cacagcccat tctgccaact  40740
catcatagtg tttgttgata agctttcctt tcatattttt ataaagttca cctctcccac  40800
aatttaattt catcactcat atattttttt aatcagcttt gtcaggttga attgtttctc  40860
acattaatgt gttcttttcc atttaataac tattatgtga ttgacaaaca tttcaaatta  40920
ttattttgcc tgttagtaaa tagtgcaatt ctttttttaat attttaaccc tttctctatt  40980
actttatttt ctagatattc aaaagcctcc actaaagaac gttattaaca agaaagacag  41040
catacttact atatattcta agttggcact tcttgtaact aaaaactgtc ttaacgaacg  41100
agacaggnaa atgaataaga ttaatgttgt tcttattagt ttatgtattt gtttctgtta  41160
agatagctat agcttcaaaa tatactcaaa agaaagtgga ttttggtgga aaataatgta  41220
tggctaaatt cagagtagag gtgcacctac ttatgctgct aacatctgta tgttccatga  41280
ctataagcaa tgctaaggga aaatccaacc tcatttaccc atcttttaat ttgtaggtgt  41340
ttgaaattgc tgatatagtt gatcagatgt caccttgggg catgttgcgg attccacggc  41400
cactgattat gatccgagca ttccggattt atttccgatt tgaactgcca aggaccagaa  41460
ttacaaatat tttaaagtga gcgctgcttt acaataagtt aagaggaaaa gtttatgaag  41520
tttttttgtaa gttttatgta ctttacctac ttttgtttct ctaggcgatc gggagaacaa  41580
atatggagtg tttccatttt tctacttttc tttctacttc tttatggaat tttaggagtt  41640
cagatgtttg gaacatttac ttatcactgt gttgtaaatg acacaaagcc agggtaagtt  41700
tatctattaa ctgcatttta atctagttga ttaacaaaga tatctctgta attttcataa  41760
```

gaataacctc tatgtcagat tttttaaaca gtgcttatgt cctctttgtt ctgtgtctct 41820

gcttcctcgg tattgttttt gtccaatttt tattttattt tattttattt tatttttgag 41880

atggagtctc gctctgtcat gaggctgggg tgcagtggcg cgatctcggt tcactgcaac 41940

ctccgcatcc cgggttcaag cgtttctcct tcctcagcct gccaagtagc tgggactaca 42000

ggcgtgcacc accatgcccg actaattttt gttttgtcca atttttaaag acaagtattc 42060

aagacaggac agattggaaa atgtaacctg gtggaaaagt agattaagca aatgcaaagt 42120

gtacatttag aaaccttagt gtggcaaaat ggggagtgtg ctaacataat aacataaacc 42180

tagagcttca ctgtggaagc ttttcatatt agaatgagtt acatgttcgg gaaaaaaagc 42240

tcatttttat tattatttac ttttgtcatc attccctaaa atagtttcct aaatgattat 42300

aacatcattt ctagactaaa cactattatc tgttgcattt gactgttgta gggttggtgg 42360

gtaagataga atgtatgtac tttaattacc tattatgtat tatgcaatag agcacctgcc 42420

tcttctactt ggctgatcat tgttaatata caactccagg gatgtgatgg caggagagag 42480

cagagtggga tgttggcccc tggatggtcc agaagtgtcc cattaggtgt tagctagata 42540

tagagtttgg ggttccaagc acaacccaga angtgcaaca gatatgacta tgccagattt 42600

cagggcagca gggattgggg ataatacaat tgattttata ccttatattc aaatcttcac 42660

ataggaattt gtctatttat aaatgaagaa ctcatccata aattaccgat tgaaaaatct 42720

tgattcaaat cttagaaatt agatttaaag cactttgtaa aattcaaaaa gagattacaa 42780

gcagccattc atcttcaaaa tatttttaat atatcaatat ctacttacaa accaattggg 42840

tatataagga tcattgagtt taaaaacatt gaaagggtta tgttgtgttt tctgcagttg 42900

gaggatttct ttttcataaa gtctatatac ttgtatatat acaaagtaaa taaaccacaa 42960

ataagtgaat atctctgtga gtatagacat catgataagc ttggctcaat tccaacactt 43020

tattttatga gtttgcataa attatttagg ttactcctca atgaagtaga ttaataacag 43080

gacatacttt ctagaggcat caggaagagc aaatgaggga atgaatgtaa agacaagcaa 43140

agtaaccagc acatagtaaa ggcattcact gtatgccagc tgtcattatt ttgttgtcat 43200

aaagagatac gtaaacacgc ttacatgctg aggtggtagg cataagtgta tcctccaaaa 43260

tagagataac atggttaaaa aaaaggaaag agcaataaag aagaccaaaa tttaaacttg 43320

taagagtgtg aaaggaaaac atctatgagt aactcatggc tatgttattt tactgacaaa 43380

gtaactaaat taattctgag tttatttcac atttataatg gtncatttaa ttattagtct 43440

ctcacagaaa taaatttaat aattgaattt ctaattcatg cacaaacagt atatgtatat 43500

tatgcaactc ctaataatta tataatgttc taaaaactaa aagtgaattt agttgtttgc 43560

tgttgcacca ccccaaagaa tgtcaactcc aaaaaggtag agattttgtt tccacgtgtt 43620

ttcagaccag gaggcatgtg caagtttcta ggaaaatatg aatgtattca gtaacttcaa 43680

aacatcagta taatttcact ggttaaatgt taaacagtaa agcataatct attagtcaag 43740

gccttaatta attttatttt tatttagaag gcatgcatct catacctact tagaaattta 43800 ccagaggggg ctgtgcacag tggctcacac ctgtaatctc agcactttag gaggctgagg 43860

agggtagata acctgaggtc aggagttcca gaccagcctg accaatacgg tgaaacccc 43920

tctctactaa aaatatgaaa attagccggg catggtggca caggcacctg taatgccagg 43980

aggcacctgt gatcgggagg ctgagacaag agaattgctt gaatccagaa ggcggaggtt 44040

gcagtgagct gagatcgcgc cactgcactc cagcctggac cacagagcaa gactccgtct 44100

caaaaaaaaa aaaaaaaaaa angaaagaaa tttaccttag gttcacagaa gtgtggacag 44160

ataccatgcc ttccttgatt atccatagta ttatgttgaa gtaaagatct catttttct 44220

tgtttttaa tataactttt atttttgcat gatgtgagtg aataatctat tcaaaattat 44280

ttttaaaaga aaaatactgt tgctactaac cccactccta atcttcatta ttattgttca 44340

aaatagcctt ggaaaataac tttttgataa ccattttctg gacatattat tattttcccc 44400

tattattgaa attatgtctc tactttaacc actttattga gattttggtg aatataaaaa 44460

tctttacata tttattatat atgtcttgag ggaactattt ccttttctg tcagaattaa 44520

atccagagag taagttattg catttattat tctacttcct gaaaattatg atccttctca 44580

aattaacttg tatatttctt atgaggtact tcctgaaaat aatgatcctc ctcaaattaa 44640

tttgtgtatt tcttatgagg tacgggtcaa gatgtagtct taacttagtg gttggaatgg 44700

acatcggtat cttaggagaa gctctcagta tttccgttaa gtgggatgtg agctgtagag 44760

ttttcaaagt gaggaagtta cctttaattc ctatagtgcc aagagttttt gatcattgat 44820

gtatgctaaa ttttgtccag tgcttttccg ttttgttcag ttaatgtaat gaattacatg 44880

attgactttt taaatagaaa ccaatcttgc atttctggta tattgctaat ttggtcatat 44940

tgtattcctc tttatacatt ggaagatttg atttgctaat gtttttgtta caaattattg 45000

catctttgtt cacaaggagc attattttta ggttttttaa attgaaatat ctttaccggg 45060

tcttgaaatc cgggtaatgc tgtcctcata aaacaagtgg aacatgtttc caagggttaa 45120

tataagattc atattctttc tttcttaaat gtttgataga attcatgagt aaaactcagg 45180

gcctgaaatt ttctttgtga gaacgttttt cattacaaat taaatgtatt caacttacag 45240

agctatggag ttgttatctt tcttgttctg ttaagttggt aaattttgct tttcaaagag 45300

cttacctatt tttataattt gtcaatattg cataatgttt atacaaatgc cctctgatta 45360

tcttttgaat gtctgattat ctttttaaca tctgtagaat agtggtaata agccattttt 45420

tcctgacatt gataatttgt acctttctct ccattttttt atgtcttctc aattgtatta 45480

ttcttttcac acacacacaa gaatctttgc cctatttaat gttctctatt gttagtctgt 45540

ttttcatgtt gatttttat cttcattatt tccttatttc tgctttttag tttttgtttt 45600

ctttttcatc tttcataaaa tgaaggtata tagacatctt ttaacatttt tctaatacaa 45660

taatttaaaa atataatttt cctctaagca ctgttttagc ctctaagcac tgttttagct 45720

gcatttccaa aatgttgata tgttgtattt ttattattat tcaattcaaa tatcttctaa 45780

tatatgtgat ttctcctttg acctgtaggt tatttagaag tgcattttta gaaggatgag 45840 attttctaaa aatgttattt gggtgcataa ttttattctg ttgtggtcag acaatagtcc 45900

ctgtgaaatt tcagcctttt gaaattcatt aggaatcatt ttaaggacca gtatatggtc 45960

tgaattggtg aaaattccat gagaatttga aaagaaaaaa gtgaaatctg cagtttttga 46020

gtataatatc tataaatgtc accaaagtca agttggctga taatttgttc tgggtatttc 46080

tatccttctt ggtttttaaa atcagggttg ttctagcaat tgctgagaga gtactattca 46140

attttctaga catgacagaa attttcaatc tctcttattt ttgtcagttt ttgctttcta 46200

taattttcaa ctttaattag gcgcataacg tgtcattgta tctttcagat gggctgacct 46260

ttttatcgtc atgaattttg aatttttttc ctaattagct tcgggatgcc agggtatatg 46320

atgggtcaag aacatgacgt ttcaaatttt tcagtaatca taactctaaa gtaatgtgta 46380

actctaaaat aatttatttt atttaatcat cttttatctt attaatcaat tgatttgttc 46440

ccacaattac ttacgtgtac tttagaccat tcctggattg acagtaaaag ggagcacatg 46500

acaaattctt agttttagag catgggctgc acaatcctga gcccagcctg gtggtgatga 46560

aattaagccc agtactaaga gcgtaaatga agaagaaact cagtagtaag agagtgtgga 46620

ccaactggtc agatgctcta atcagcaact ctcattccac tgtcaaaacc ttggagattt 46680

tgtatgtttt ttaaataggt gaagtgaatc aagtggtgcc taccaaattt attttgtcct 46740

ctgcatcagt gcgtggcata caacattaac aatgagaagc aactaactcc tataagctat 46800

tttgggggaa tagaggatgg tacaaatatc acaactttac agcaaatatt acaactcctt 46860

caaaacacag tttggcagtt tttaaaaaaa taaaaataaa agagttgatc aagctgggta 46920

cagtggcaca cacctgtaat ctcagcactt tgggagactg aggtgggatg atcacttgag 46980

cccaggagtt tgagactaga gtgagctatg attgcgtcac tgcactccag cctgggtgat 47040

agagtgagac ccagtctcta agaagtaaaa atgaaagaaa gaaaaaaagt tggacaggca 47100

cctgccacat ctcatttatt ccactccagg tatttaacca aaataaataa aaatgtatgt 47160

ccatagaaag atttacacat gaatgcccac agcaacttta ttcatagtac tccaaactgg 47220

tgacaaccca aatttccatc aacagataga taaactaatg atggtatatc catataataa 47280

atactgttta ataataataa ataatgaact attgatgtat acagcatctt ggataaatct 47340

caaaataatt atgatgactg aataaagcca gatttaaaaa gaggacatgt tctattattc 47400

catttacata aaattctaga aaatgcaaac ttatctttaa tgatagaaaa gaaatcagca 47460

gttgcctagt gataaaggtt gtggtgggaa ttaagggatt acaaaggggc ctcaagaaac 47520

ctttggggtt tattgacatg ttcattatct tgactgtgct tctgtacatg tgccaaaact 47580

tatcaaactg tacactttcg tatgtgcagt ttatgttatg tcagttacac ccactgaaac 47640

tgcttaaaat gtctcagtgt gaagtagaaa tctaattttt catatggtta agcaattgtg 47700

cctgcactat ttattgtgtg ctattaccca ctgagtttag tttcacctgt ctcattcact 47760

aattccctgt ctatacgtgg gtctgttcct gtggggctct gttgtgcttc actggtctat 47820

tggactacct ctttgccaat atcacactct gaaagtcttt cagattacaa tctggcttta 47880 cacatggggt catatttcca gctttcagct ttacatgggc tcaagacctg gtcttccgcc 47940

tcacatgggc attaaaaccc aggcattagc tcccattcat aacactatgt caactcatgt 48000

tccttctgtt cctactcatt tactgctttc cttcttatgt tttgacattt tggaatgcag 48060

gttttgtatt gtgagcctaa atgcggagtt tttaatgcta tctatttgtt tgtagtggaa 48120

ggagttgaat ttggtgtagt ccatcttgcc acaaccagga ctcacatcca tagctatttt 48180

tttttctatt tttgtctaag gtccttgagg acacgtatga gaaaggtaaa gaggatattg 48240

tgttatattc atctttttat ccttactacc tagtatccta acacagaata aataattagt 48300

aactatttat tgaaggcatg gttgagaaca tattctattt agacagataa tattaacatg 48360

atgacaaatt gataaagaag atgaaataaa aaagaaacag agaaaagtaa gttgctaaaa 48420

gcaggaaagg agaaatggct aaaaaataga aaaacaaagt ttaagcagaa agaagcagtg 48480

aattagttca tatacggagt aagaaaattg aagcaagtgg tgagtaaggg agttgactgt 48540

gtgggtgagg gaagagacaa atgcatggaa gaaaatctgt aggtagttgg atttattttt 48600

ttagtgtatg taggctgtag aatatgccta gattgttcta tgagtgaata tgaaaatgga 48660

aaaagatatt gtgtgggttg gagacaatca atatggagat gaagaacagg gaaatcttct 48720

ggataagaaa gcagctgtgt agtactgcag tgggtacaga ggctggcttc aaattaccgt 48780

tttattaaaa ccaagtcaga atgtgagagt tataaaactt gtaaatgcag ctagaagatc 48840

aggcattatg ctggcaaacg aaagctcata ttggcacata aatgatcgca gaaattgaga 48900

acagatacac gaatgtcttc tgaactcatg accctagagt tgcttcagta agtattcttt 48960

gctagttcca cgctgtattt ggtaccaaga aaagaatgat actgaaaaaa tgcatgaaaa 49020

catgccttgt gatctaaata ttagctccag ctcattactt tattctaaaa atgaagttca 49080

ccatgaaaaa tttcataaat gttgacagaa tgcttagtta cctgagacat agccattatg 49140

gtgaaatggc tgcatttttc tctttcccat atgtgtaaaa cctgcctagt agcatacaca 49200

cacacacaca cacacacaca cacacacatt tattttgtct tctttttatg tgactctttg 49260

tttactattt aagaagaaaa tgcttaacta gccacatggt agtaatgttg ctcataatca 49320

caactgttat agctttgata attcatttca agaatctatc aataatatct tttttgttgt 49380

tgttttttttg agacagactc tcactctgac acccaggctg gagtgcagtg gcatgatctc 49440

agcttactga aacatccatc tcccaggttc aagagattct cctgcttcag ccccccaagt 49500

agctgggact acaggcgtgt gccaccacac ccgggtaatt tttgtatttt tagcagagac 49560

agggcttcat catgttggcc aggctggtct cgaactcctg acctcaagag atctgcctgc 49620

ctcagcctcc caaagtgctg ggattaaagg cctgagccac catgtccagc cttattttta 49680

aaaaaaaatt tcagttcaat attactgcat tttaaagccc cacacataag cagaaaaact 49740

gtcagcacca ttttaccgtt actattagcc cagacataaa tgtagtttta ctcttagaga 49800

ctcaagaaat atccaaactc aaaataaaac agaaaatgtt aattgatttg aaaattttta 49860

aacttttctg tttcctttttg tagcctctaa actatgactt tttttgctgc tataaagaca 49920 catgcacacg tatgtttatt gcagcactat tcacaatagc aaagacttgg aaccaaccca 49980

aatgtccaac aatgatagac tggattaaga aaatgtggca catatacacc atggaatact 50040

atgcagccat aaaaaatgat gagttcatgt ccttggtagg gacgtggatg aaattggaaa 50100

tcatcattct cagtaaacta tcgcaagaac aaaaaaccaa acaccgcata ttctcactca 50160

taggtgggaa ttgaacaatg agaacacatg gatacaggaa ggagaacatc acactctggg 50220

gactgttgtg gggtgggggg aggggggagg gatagcttta ggagatatac ctaatgctaa 50280

atgatgagtt aatgggtaca gcacaccagc atggcacatg tatacatatg taactaacct 50340

gcacattgtg cacatgtacc ctaaaactta aagtataata ataataaaat aaaataaaat 50400

aaaataaaat aaatttcaaa aaagtcaaaa aaagaattcc ctgtgtaaga acattcaaag 50460

tgaaaggagc atcataggga aaagacaaac cgtgaaagca gagaaatagg agacgccaca 50520

tcaagacacc accatgctta caagttggca ttttattctg aagaccataa tgaactgcca 50580

aagaagctga aacctaaaat aaaattattg tttttacata taaagatgat ttttcacatc 50640

tatatttatg atctcacctt ttaaatcagt atctactttt gaaaatgtac atgtaacctt 50700

caacatatac acagaaaata agataaacta atttatatag ttttgtttca atcctacaca 50760

attagagttg atttttactc catatttaga tttccagaat cccagctcat taatttaata 50820

tttattaagc ataaaagctt taagacctag aatgcatggg acgtggaatt aagcactagg 50880

gataaaagag gaagaaaaaa ataggtcatg tcctggagga gctcacactt gaggcagcca 50940

aagattaatt accttaggca ttgattagta ataataataa taattagtta gaaagtgtta 51000

tcagatgtca aaaaagagag aaaatgtgag ttttatagtc tttgtaaaga atatgacctt 51060

tgagcaaact ctagaattag gaatggagtg tttgaattgg aaattaaggc aggataatac 51120

ctttcagatg aaacaatgta atagagtact aggatgtata atagaatgct ggccccaaac 51180

tgggatggta gatttaagtt gtgagtttct ttctacaact gcagagtcca ttgaataaat 51240

acaactttat atattaaaat taaaaaaata aaaaaaaata aaaaaagaaa ttctgtcatt 51300

tggaataaca ggaatgaatc tagaggacat tatgctaagt gaaataagcc aggcacagaa 51360

agatcaatac tgcatgatct cacgcatatg tggaatccta accagttaat tcattgaagt 51420

agagaggaca atggtggtta tcaggggctg cggcgagaag gagatggagc tgcaagggaa 51480

tgaggagctg ctgatcaaag ggtacaaagt ttcagctagg cagtaagagt aagttttgag 51540

atccatgcac aacagggtga caatggtcaa taataatgcg ttatatatgt caacataact 51600

aagaaagtaa atttcaaatg tatcatcaca aaaaagttaa gtgatatgat ggatatatta 51660

attggcccga tgtaaccatt atgtcttgta tacatacatc aaaacatcac atggtgctgc 51720

attaaagtaa tacaattatg atttgtcaat taaaatacta ttatttaaaa aatgtattat 51780

acacatctcc gagtcacata ttaagctaag gtatgatggc tgtgttttca gcaatcagtg 51840

ttaccagcta ctgccgtgac ctcatttctc ttaatcattt aagagaaacg attaagagaa 51900

gttgttaaca attgttccaa gttctgccta attgatttgc atgtattata gcctaatcat 51960 gacagagatt tcagaccaga aacttgcata tcttgggttt catagttgac ctccagaact 52020

gattgtgttt aatatgtgag gcaaacaaga taaagaagca acacaatata tacagttttg 52080

cttttaaatg gaacttagtt cctgtaaggc ccaatagcag gaaatcttag ttatgtcttc 52140

agtagcagta ctagtaaagc tttagattct tctctgttaa tgaatggaga cctaaggcaa 52200

ttatgaaaag agtaatgaga tttgactaag caaagcagct catttcctaa agaaagttga 52260

ggagagcaat ggcttgtagc agaaacaggt gttgtgaaat accaactgtg atggcagaaa 52320

tctccccgaa gctggaacag gaatgtttaa aaatgagctg gcaggttaaa tagaaaagat 52380

gaagacttac ctgacacctt attatcattc ctgaaaaaga agggaaacat attcatagct 52440

cggtatcaaa taatgaaatc ttatttccaa gagtgagata attagaaaaa ctttaaggta 52500

agtaagtgta accattaaat atatgaaagt ttaatataga gagagcagga ttcaaacagg 52560

aactcaggaa catattcgcg atggctggac cactagcttt aagcccagag gttgggtgat 52620

cacacatcac agatgcccag gacaatcttg gcttatccct actgtttcaa agtaatcact 52680

aacagcctgt tattcgattt taaggaatct caatgtagac aatgcatttt atggtcacct 52740

tgcccagggg aagtttatag aacttggtat tttataaaat atcaccaaga gttgaattag 52800

agataatata aagttggatt tttttgcagt ttgctcttgc tggtccaaat acattttaca 52860

tttattaaaa taaatacatt aaaatttatc aactaaatta aaatttattt gtcacgtatt 52920

ctttagggta cttggcaaat aagtgaagta ttgtcttata aattttgttg acaaaataaa 52980

gaggtttttt catggaatgc tatgcagcca tgagaaagaa tgagactatg tcctttgtag 53040

caacatgaat gaagctggag gccgttatcc taagcaaact aacacaggaa gagaaaacca 53100

aataccgctt gttctcactt atgaatgaga gctaaacact gagtacatac agacacaaag 53160

aaggcaacaa cagacactgg agcctacctg agggtggagg atgagaggaa ggtgaggact 53220

gaaaaactac ctatcgggtg ctatgcttat tagctgagtg gcgaaacaat ctgtacgcca 53280

cgctcgcatg acacgcagtt tacctatata acaaacctgc acatgtatcc ctgaatctaa 53340

aataaaagtt aaaagaaaaa gaaatacaaa tggaaacata ctggaattta gagcattaat 53400

ttttaatttc taaggacttt atcaagagac agttttgtag aatatgtaaa atcataattg 53460

ccccaaattc atgcaacggt ctatgacgtc ttcttgattc ctatgtattt atttatttac 53520

atatgtttat tgacacaata gatgtnaagt atttttggag tacatatgat atttcgatac 53580

attgatgtaa ggcataataa tcaaatcagg gcaattggca tatccatcac ctgnaaacat 53640

ttttcttaat actgggaaca tttgaattat tctctactag ctattttgaa atatgtaata 53700

aattactgtt tactatagtc accctactga tttatcaaac gcttaggtct tatttatttt 53760

atctgactgt atttttgtac ccattcacca acctttcttc atccaaccct ggcctttacc 53820

cttccccttc tctggtaaca accaatctac tttctatctt cataagatcc acttacttaa 53880

ggcctaaata tgagtgagat catgtgatat ttgtctttct gtgcttggct tatttctttt 53940

aacataatga cctccagttc tgtccatgca gctgcaaacg tcaggatttc actcattttt 54000 atggctgaat agtattaaca ctgtgtgtgt gtatatatac atatcacatt ttctttatcc 54060

atttatccat tggtgggcag ttaggttgat tccgtatttg gctatttgga atagtgctgc 54120

agtcaacatg tgagtgctgc tatctcttaa atatatagat ttcctccctt ttggatatat 54180

acccaatagt gaaattgctg gatcatagag ttgttccatg tttagttttt ttcagtccgt 54240

actgtttttc acagtggctg tcagaatttg cattaccacc aacagtgtat gagtgttcct 54300

gtttctctac attctcacca gcactcactg tctttttat aaaagccatt ttaacttggg 54360

gtgacatgat atttcactgt agcttttatt tgcatttccc tgatgattag tgatgttaac 54420

catttttata tacctacttg ccatttgtat gtcttctttt gagaaatgtc tattcagatt 54480

ttttgcccat ttcttaatca gattattaga tttttcccat tgagttattt gagttcttta 54540

tatattctgg ttattaatcc tttgtcagtt gaatcatttg cagacatttt cccttattct 54600

gtgggctgtc acttcaattt gttgattgtt tcctttgctg tgccgaagtt ttgagcttaa 54660

tgcaattcca tttatctatt tttgctttag ttgcctgtgc ttttcagatt ttgccccccc 54720

aaaaatttgc ccagtccagt gtcctggagc atttacccaa tgttttcttc taggagtttc 54780

atagattcag ttcttagatt taagtcttta atcctttatt atttgatttt tatatgatga 54840

gaggtaggag cataatttta ttcttctgca catggatatc catttttccc agcaccattt 54900

gttgaagaga ctgtcctttc cttaatgttc ttgatacctt tgtcaagaat gaattggctg 54960

taaaagcatg tatttatttc tgggttttct attccattcc attgggatat tttctgtttt 55020

tttaccagta ccatgctggt ttggttgcta tagctttgta gcatattttg aagtcagtta 55080

gtgtgatgcc tccagctttg ttctttttgc acgagattgc tttggctatt cagaatcttt 55140

tgtgttttaa tataaatttt aggatagttt tttatattcc tgtgaagaat gtcattggta 55200

ttttgataga gattgcattg aaactgtaga ttgcttcggg tagtgttgac attttaacaa 55260

tattgattct tccaattcat gagcatcgaa tatctttcca ttttttgtgt ccacttcaat 55320

ttctttcatc agagtgttat agttttcatt gtagagatct ttcacttctt tagtcacaat 55380

tattcctagg catttttaat tttttaataa atattgtaaa tggcattgct ttcttgattt 55440

atttttcaga ttgttcactg ttggtgtata taaatgctat taatttttgt atgttggttt 55500

tgtatcctgc aatttactga gtttgcttat catttctaac agtttttggt ggagtcttta 55560

ggattttcta agtgtaagat catgtcgtct gcaaatgagg ataattttat ttcttccttc 55620

ccaatttgga tgtcgtttat ttctttctct tatctgattg ctctggctag aacttccagt 55680

agtatgttga ataaatctga tgaaaatggg tgtcctgatc ttgttccaga tcttagcaga 55740

aaggctttca atttttccgc ctttggtatg atgttangcc atgggtttgt caaatatggt 55800

ctttattgtt ttgaggtata ccccttttat actcaatttt ttagtgtttt gataataaag 55860

acatgttgaa ttttaccaaa tgctttttct gcatctattg aaataataac atggtttttg 55920

ttcctggttt atttatgtaa tgtatcatgt ttattgattt ccatatgtta agccatcctt 55980

tcatttctgg gaggttaaat aatatgtcca gatggtaagt aagaaagcca ggattttaac 56040 tcaagtctaa cttccaagcc tatgtacgtt ttaccatttg agaaacttct gtattttaat 56100

tctacattca taaacctata gacatataga gatccacatn ccatatgtat aaactgtttn 56160

ataatatatg tagtataatt tnctnacatt aaagtttttn attaangtct ggtttctant 56220

ccctatgtat atgnttatgt ttcatgggat tcgacttnca tctaaatnct ataaactagn 56280

tgatatattn atatattttt agtactatct atatatctct tttnaaaatt gttctgcaca 56340

ttgatctcca agatgtgaac cacctgaggt gccagaaacc agcctttcat atgataataa 56400

tattaataaa gtaaatataa gaatatattt cagaataaat tgaactttca gatgaaaatt 56460

aataccagca tgatttttac cagaatagca aaaatgtttg tgagagtctt cagtatacca 56520

aaagaaaacc atgcagtttg aatgggggtc taaaaggaaa ngtaacttat tttgatccat 56580

cttttgatca tatgtttact tcccaaaaag aagataatgt aatattatag gttataagat 56640

tcttattttt agtggtaaat aacttttatt gaacaattgc tgtatgttat tttctcactt 56700

aatacttata acaacccgat aagttatgtt ccatttttac actgtatatt acagccaagg 56760

aaattgaagc tgagtaaagt taaacaaagt ggctgagatg aaacagcttg tatcagaagc 56820

aagactagac ccatttctcc ctcacactgc tttctatgct ttcagttgta acttgttcag 56880

tgtatgagat aaccttcctt caacatgtgt gtccaaaatg tttgaataaa atatttagtt 56940

ttcctttagt tacaaggagc ctattcttaa aagacaagtt ctcagtggga cgcatgagta 57000

agtggttaat tacgattgct tattttaaga ctgaattcta aaggcatgca ctggaacttc 57060

tgaacatcct ctgtgagtgt ggaatgagga gctctaccta ttctgtccct cacaaatagt 57120

gtttactgta gccgcctcca cagaatcgat atttactatc acttttgtat ctcaggttcc 57180

acccaaccag cttggttttc cccatgatag aaaccataat aagtatgcat aactacagca 57240

gcaccaagag taaattttng aaaattaatt tttaggatga aatctctttg gtaaagaaaa 57300

tcacctggca aaagttgtaa aaaatactga ttccaagcca catttgtgat tcgtattttt 57360

aggcaaggct gagattctca gaattggcca ggttagaaat caatcgttta aaattttagt 57420

tttgcaatta aactccaaaa gatatgttat agagtagaac acaaaattac atggcttttt 57480

aaaaaaatag accttgagtc ttcaaagaaa tgttagtctt ctagctagct gcttactgcc 57540

ctgcccaagg gtagcatgcc tgtcctcttc ccttaagact atttccccag aaatgtccaa 57600

ggagcaccag aggtaattag agcctcctgt ctttcctgca aagaggcagg tatgtctttg 57660

ttttctggga cggacaacca tctcaatttt atttatttag tgactcttac atgttcttat 57720

ttctagagat tgaattgggc tggtgtactt tgtggttgag tgtggttgga gtcgtcaggg 57780

gcattggcac ttttaaaatc cctcactgac tttacttttc gtgttcacat cattggggag 57840

tttgtaaatc cattctttca ttgctataaa gaaatacctg aaactggcag ggcacggtgg 57900

ctcatgcctg taatcccagc actttgagag gtcaagacgg gcagatgact tgagactagg 57960

agttcaagac caacctgggc aatatggtgg aaccccatct ctactgaaaa tacaagaaat 58020

agctgggaat ggtactgtgc acctgtgatc ccagctactt gggaggctga ggcacaagaa 58080 tcgttggagc cgggaggtga gggaggttgc agtgagccga gatcatgcca ctgcattcta 58140

gcctgggtca cagagtgaaa ctccatcgaa ggaaggaagg aaggaaggag agagagagag 58200

agagagagag agagagagag aggaagagag gaagagagga agaaagagaa agaaangaaa 58260

gaaaaaaaga gagaaagaaa gaaaaaagaa aggaaangaa aaagagaaag agaaagaaaa 58320

gaaagaaaga gaaagaaaga aagaaaaaga aagagaaaaa gaaaaactga gactgggtat 58380

cttatttctt ttattacatt tttgggtcct atatttgtat catacagaat gagttctgta 58440

tatttttcag atggaaaaga aatattgact taaaaggtac ctctcaaatc ttaatgtgta 58500

caataaactg gagatcgtgt taatgtagat tatttttca gaagatctgg aagaaagcct 58560

cagtttctga atttttaaca agctcaccag taatgtcaat gtttttggcc aaagaagaat 58620

attttgtcaa aaattcagta agtggaagag cctgtgtttc tttttctagt ttttctggct 58680

tgtaaacaaa gatgagagtt ttcatttacc aaagacagat atatgcaaag aaaacctaag 58740

aaagaagggc agctgccaga ttaaatgtga tggtttatgc acatcagctg cataaagata 58800

cctaataatt aagagaaaaa taattaactc tgtagtgcga cagtctgtca gagagctctt 58860

taaccaagtg aatcataaac attaactgca ctaggacaaa tttttatgan tgtgctgatg 58920

cacacagaag gacgcagcat cactgctgtg atattgcccc ctcaaaagta cattatagtc 58980

taaatttaac cataaagaga catcagtttt atgcgaagct caagatacac gtatctttca 59040

tgttctgtaa tttgtaatgt tgattttaag tagtctttag caccctaggg agcaagcatc 59100

tcctaataat ttttttcaga actttctggg taataaatgc catcctgtta aataagtatt 59160

ttctactttt tttcctgctt tgtatccatt tattcacttg tttgcttttt aaaaaatttt 59220

tactttaagt tgtgggatac atgggcagaa tgtgcaggtt tgctacctag gtatacatgt 59280

gccatggtgg tttgctgcac ctgtcaaccc atcatctatg nttttcagcc ccgcatgcat 59340

tagctatttg tcctaatgct ctccctcccc ctgcccctca ccccgcgaca gaccccagta 59400

tgtgttgtat gttctcattg ttcaactccc acttatgagt gagaacatgc gctgtttggt 59460

tttctgatcc tgtgttacgt ttgctgaagg tgatggcttc cagcttcatc catgtccctg 59520

caaaggatat gatctcattt ctttttatgg ctgcatagta ttccatggtg tatatgtccc 59580

acattttctt agtccagtct atccttgatg ggtatttggg tggttccgtg tctttgctat 59640

tgtaaatagt gttgcaagaa acatacgtgt acatgtatct ttatagtaga atgatgtcta 59700

ttcttttggg tgtataccca gtaatgggat tgctgtgtca aattttattt ctggttctag 59760

atccttgatg aatcatcaca ttgtcttcca cagtggttga actaatttac attcctacca 59820

acagtgtaaa agcgttccta tttctccaca tcctcgccag catttattgt ttcttgactt 59880

tttaataatc gccgttctga ctggtgtgag atggtactgc attgtggttt tggtttgcat 59940

ttccctagtg atcagtggtg ttgagctttt tttcatgttt tttggctgca taaatatctt 60000

cttttgagaa gagtctgttc atatcctttg ttcacttttt gatggggttg tttctttttt 60060

tcttgcaaat ttgtttacat ttcttatagg ttgtggatat tagacctttg tcagatgggt 60120 agattgcaaa aattttctcc tattctattc tgtaggttgc ctgttcactc tgatgatagt 60180

ttcttttgct gtgcataagc tctttagtac ttctagttta attagatccc atttgtctgt 60240

tttggctttt cttgccattg cttttagcat tttcatcatg aaatcactgt ccatgtctat 60300

gtcctgaata gtattgccta agttttcttc tagggttttt atggttatgg gtttaacatt 60360

taagtcttta atccatctcg atttaatttt tgtctaaggt ataaggggtc cagttttagt 60420

tttctgcaca tggctagcaa gttttctcag caccatttat taaataggga atcctttccc 60480

cactgcttgt ttttgtcagg tttgtcaaag atcagatggt tgtagatgcg tgatgtgatt 60540

tctgaggtct ctgttctgtt ccattggtct atatggctgt tttggtacca gtgccatgct 60600

gttttggtta ctctagcctt gtagtatagt ttgaagtcag gtagcatgat gcctccagct 60660

ttgttctttt tcctaaggat tgtcttggtt atacaggctc ttttttggtt ccatatgaaa 60720

tttnagggta gtttttttct aanttctgtg aagaatgtca atggtagttt gatgggaata 60780

gcattgaatc tataatttat tttgggcagt atggccattt tcatgatatt gattcttcct 60840

atccatgagg atggaatgtt tttccatttg tttgtgtcgt ctcttatttc cttgagcagt 60900

gatttgtagt tctccttgaa gaggtccttc acgtcctttg taagttgtat tcctaagtat 60960

tttattctct ttgtagcaat tgtgaatggg agttcattca tgatttggct ctctgcttgt 61020

gtatttttgg tgtataggaa tacttgtgat ttttgcacat ccattttgta tcctgagatt 61080

tgctgaagtt gcttatcagg ttaaggagtt ttcgggctga gatgatgggg ttttctaact 61140

atagaatcac gacatctgga aacagaaaat ttcacttcct ctcttcctat ttaaataccc 61200

tttatttctt tctcttgact gattgctctg gccagaacat ccaatactat gttgaatagg 61260

agtggtgaga gagggcatcc ttgtcttgtg ccagttttca aaaagaatgc ttccagcttt 61320

ttcccattca gtatgatact ggctgtgggt ttgtcacaaa tagcacttat tattttgaga 61380

tatgttccat caatacctag tttattgaaa gcttttaaca tgaagagttg ttgaatttta 61440

tcaaaggcct tttctgcatc tattgagaat catgtggctt ttgtctctgg ttctgtttat 61500

gtgatgaatt acgtttattg atttgtgtat gctgaaccag ccttacactt cagggatgaa 61560

gccgacttga ttgtggtggg taagcttttt gatgtgctgt tggattcggt ttgccagtat 61620

tttattgagg atttttgcat tgatgttcat cagggatatt ggcctgaagt tttatttttg 61680

ttgttgttgt gtctctgcca ggttttcgta tcgggatgat gctgacctca taaaatgagt 61740

taggaaggag tccctccttt tcaattcttt tggaatagtt tcagaaggaa tgataccagc 61800

tcctctttgt acctctggta gaatttggct gtgaatctgt ctggtcctgg gctttttttg 61860

gttggtaggc tattaattac tacctcaatt tcagaacttg ttattggttt atttagggag 61920

ccaacttctt gggagggtgc gtgtgtccaa gaatttatcc atttcttcta gattttctag 61980

tttatttagg tagacatgtt tatagtattc tgatggtatt ttgtatttct gtggggtcag 62040

tggtgatatc cccgttatca ttttttattg tgtctatttg attcttctct cttttcttct 62100

ttatcagttt agctagcagt ctgtctatct attttgttaa tttttcaaa aactagctcc 62160 tgaattcatt aattttttga atggcttctt gtgtcactgt ctccttcaat tctgctctga 62220

tcttagttat ttcttgtctt ctgctagctt ttagattttt ttgctcttgc ttctctagtt 62280

cttttcattg tgatgttagg gtatccattt gagatatttc tagctttctg atgtgggcat 62340

ttagtactat aaatttcctt cttaacactg ctttagttcc cagatattct ggtacgttgt 62400

ctctttgttc tcattggttt caaagaactt cttgatttgt gccttaactt cattatttac 62460

tcaggagtca ttcagcagca ggttgttcaa tttccttgta attgtgttgt tttgagtgag 62520

tttcttaatc ctgagttcta atttgattgc actgtggtct gagagactgt ttgttatgag 62580

ttcagttctt ttgcatttgc tgagcagtgt tttacttcca attatgtggt tgattttaga 62640

ataagtgcca tgtcgcactg agaagaatgt atattctgtt gatttggggt tgagagttct 62700

atagatgtcc attaggtcca cttgatccag agctgagttc aagtcctgaa tatccttgtt 62760

aattttctgt ctcgttaatc tgtctaatat tgacagtggg gtgttaaatt ctcccactat 62820

tattttgtgg gagtctaagt ctctttgtgg atctctaaga acttgtttta tgaatctggg 62880

tgctcctgta ttgggtgcat atatatttag gatagttagc tcttcttgtt gaattgatcc 62940

ctttaccttt aagtaatgcc attctttgtc tttttttatc tttgtttgtt taaagtccgt 63000

tttctcagag gctaggattg caacccctga tttttttttt ttctttccat ttgcttggta 63060

ggtcttcctc catcccttta ttttgagcct atgtgtgtct ttgcacatga gattggtccc 63120

ctgaatacag cataccaatg gatcttaact ctttatccaa gttggagact gggtatttta 63180

taaagaagaa aggtttaatt ggctcatggt tctccaggct gtacaggaag catggcacca 63240

ggcacctgct cagcttctga tgaggcctca ggaagcttac aatcatggca gaaggcagac 63300

ggggagcagg cacatcacgt gacaagaatg aacacgagag agagatagtg gaggggaggt 63360

gccctacact tgtaaatgac cagatcccct gtgaactcag agcaggagct cattgatcac 63420

ctaggagttg gctcaagcca ttcatgaggg ttctgcccct gtgatccaat cacctcccac 63480

caggccccac ctccagcgct gaggagtata cttcgatgag atttgcgcaa ggacaaatat 63540

ccaaactata tcagggagtc gttagtatgt ctggccacct tgaaaggatg aaagcttatt 63600

ggagtggaaa aaaaattgcc tctctccagc catacaactc agtcaaacca cattcaatta 63660

acataaatat attttgacct ccaagtatgt aagaggtgat ttatattcac aatttctaat 63720

ttataataca gtttcataaa tttgatcttc acaagaaaga tgagcaaatt gctccttgcc 63780

aaagttaatt ggctgaatta ccaactgaag ttctgttatg cttggcccca aattctggtg 63840

gccccttttcc attatgtcaa ctggctctct ttacatatag attctcaccc tgtagtttcc 63900

agttcctaag gcaaaaggag aaaagttgct gcgtgaggat ctcttggccc agattttaaa 63960

agtcaaagac agtggactgc cttataacct caaacggttt cacatttcag tttggagctg 64020

atcagtgcat cctctgtgtc ctgatgctta acttttcttt tcttttcttt ttgaggtgga 64080

atctcactct gttgcccagg ctgtagttca gtggcatgat gttggctcag tgcaacctct 64140

acccgctggg ttcaagtgat tctcctgcct cagcctccca agtagctggg attacaggca 64200 tacgctacca aacctggcta atttttgtat ttttagtaga gatgggtttt actgtgttgg 64260

ccaggctggt ctcgaattcc tgaccttagg tgatccgccc gcctcggcct cccaaagtgc 64320

tgggattgca ggcgtgagcc actgcacccg accaactttt ctgatattat tttgaatatg 64380

ttcatgccct ctcattgcct tatcctaacc caactgcagt cttctggctt ctgtcaggtg 64440

cactgtccct tatccatact ggcaaagtag ttgagggagt caacttgggg aacatatatt 64500

attttaaaaa gctattcaga tataacaaga taactattta aaatattcct ttaaaataaa 64560

atcatgtact gcattgtcac ttaaatttca tagtcacaat atttgaaaaa atacttaatt 64620

ttaccatttt tatgcatgat ggctacaatt ctttacaact ttgcctacca acccttttcag 64680

gaatgggtga aaggagtggg aaaatcatat taaaatactc agttttacaa cttatttcct 64740

agaaaaaatg tgccaggtta acttgtttat taaaagaggt gtgatgtcaa tattatttcc 64800

tgcttcagaa tatttatctt gttttgcagt tccactaaaa gtccaggcgg agccattgaa 64860

agttttccaa aatcaataaa taatttatct ctgcaaataa gactcttgct ttccagggtg 64920

ggattagtat tccaaagcat ccagcattag aattagctag accaattcat aggagtttat 64980

aattaatcct gaccagagga agctaaataa gagtgaaaat gaagtcagat gattgtctga 65040

agtgaagctg gaaggcggta acagccagac taagttttct gtaagtacat gcaatgacca 65100

ggaaaacctc tcaatctagg ggtcagcttt tagcaagata atgtgtcttg ctaaattttt 65160

atatctgcca caaaaggcaa gacaagcacc acgtaaggta tcaaaggcta tattcattat 65220

tctcttgagt gcaaaagagt ctgtgtctgt gtctatgtat gattgtgtgt gggtgtaatt 65280

ttctccatgg gtttatttag aattgtgttc ctctattttt gatgtatttc tagaacatac 65340

tatataaata aaaatgaagc cttacagatg caatcaaaaa caaaatgctc aagttgtata 65400

gtgttatctg ttgcatgtgg atttgttggg ctacttattt atttacatca ctgacaattt 65460

cttatggaat gccgacttca gaaatgtatt gtagcttgga ctctgcaaat gtcttttaca 65520

acaaatgggt caaatttcga taccaagaag tcatccttag cctgtttgtc tttagctttt 65580

taacaaaaaa aggtaataac aatgtaaaca aaattgctat gtataaattc ctatcatagg 65640

tatatgtttt gaaaatctag caagtaaatt tagtaataat ttattttgtg agtatctttt 65700

gacctttttt gtcctttctc acctactagg aacagtttat aatcagctca tgctgcatca 65760

gtcacagtat ccatttaggg tagcacgaat ccttgtcttt taccctaatc tataaatctc 65820

taacagtgta gtttaaaaat tgcttttgta tgataacaac tgcttgttaa caatccttgt 65880

atctcactaa acagctacac tcaccagaga aaaggtgaaa aggtctctct ctttctctct 65940

ctctttcttt ctgtcactct ctctctctca tatgcaaatg caggctctga cattttcact 66000

ttccagaatg ttatgctgtt tattccatcc ctagaactag tacgtgtgca tcctttaata 66060

atagtcactc caagcaacac cctattgcaa ggttttccca gtacccacca tctgccaccc 66120

gtagtcatag ttgatgtttc cgctttgaac ttcagttaac atcaggagca tatagatagc 66180

tcttaaagca ttcatcacac tctcctgtgg tcattccttt tttattctct ctcatactaa 66240 gcagtaaagt ccctgagtcc attttattcc ttttattcat ctccattcct agcatgctgt 66300

gtggcataga gtaggcttgt caaaaataat tattatgaat taataaacag acattcaaat 66360

acaattaagt gcccccctac gaatacagga aaatttttct gaaaagaggc taagtcatca 66420

taagcatagt catattcatt aatgacagtg attacagaag gccaacacta agtgcagtgc 66480

cagaagccat gatggtagtt ccatatctgt agagaaagac gatattgtga ataacgaaga 66540

gaggaatctt aacatgaagt cttaagctat tgcctttgtc gttggtgagt cttgtggtga 66600

gaggagaact atgcacagtg aagcatatgt tgaattatgg ggcacatgct tttgcttaag 66660

tatatccctt ttcttttata tgaaagcgtt atgaggtcct tctgggaatt gtttttttaaa 66720

cagctggaaa gtttctcatt tttcagatga ggaaactgaa gctcaaatgt ttcaagtgaa 66780

tctcccaagg taattaagtt ggttaatcac cgaactgaga tgatcatctt tgtgttgtga 66840

tatacagctc agtgctctgc accgctagga gtggtgaata acttaaaacc acaggactct 66900

gaggtcctga ggtattgttt gtggccatta tcaacaattt tagtaacaag ttcagaacac 66960

attttcttag atctcaaact cctacaaaat gttatttgta aatagaatgt tttgttaaga 67020

attaaaaaca aacaagcaaa catatgctag gaggataatt gaaggctttt aaaaaaacaa 67080

ttagctgctc ccagggattc tagtactcat tttgatacat ggaagctact aaagccagaa 67140

atttaaaaca ttgaaaatca gtgccaattt gtggattatt acacttaaaa gtgcataaaa 67200

caacaggatc cacttattaa agagtttctg tagtttacaa ataaataatt caccatagat 67260

ttaaataatg gtagaaagac acattcattc attcattcat ttattcagaa aataattatt 67320

aaacatctac tatatattag gtttctccag agagaatata gaatagaata tattatatat 67380

atgtatgtgt atatacatat atgtatgttg gcttagctca aaagatatgt ctatagatac 67440

agatacagat ataaagagat ttactttaag gaattggctc atgtaatttt gtaggctgtt 67500

aagaccaaaa tcttcagggc agcctggaaa tttcagcaga agctgatgtg gcaggccaaa 67560

ggctctgcca aagatgtggt aggccagtga cctctccctg cagagaagtc aaaatctttc 67620

ctattaataa gatgagtccc actcacatta tggagggtaa catgctttac tcaaagtcta 67680

tcaatttaag tgttaatcac aactgtaaaa taccttcaca gcaacatcta gactggtgtt 67740

caaccaaaca gccaggcacc atactttagt caagttgata cataaattta actatcatag 67800

acgctgtgtc agggacttca tggaattgga aattacacta aggaacagtt gagtttggaa 67860

gacatccaat gaacagatga ccacacaaac atgaattgcc tttgtctcaa gtgccagaaa 67920

caaaaagtga gggtaccttg agaagttata gcagggaaac ctggcccagt ggaaggcact 67980

tggagtggct cccccgagaa aatgatagtt aagctggggc ctgaaggcta caaaaatggg 68040

gttaaccaaa gaaaacagga gagacttcca agtggaggag aagaccctgc attaggcagg 68100

agctgggaga aggccaacgt ggtcagacac tgggatactt gggggaaagg gggtacaaag 68160

gggtgtttga agtgtgggag gagccagatc cttcagtatc ttgcagacca tgttacaggt 68220

ttgccactgt taaggtagga cagtgattat gaagggtctc tccatatgct ttgtgccaaa 68280 tgtatcagaa aggctaagag tgggaggcag ggtggccagt gaggaggcca gataagagtg 68340

tgtgtagaga gggagatgga gagaatggaa ttagctcaaa agatattttg agaatatgtt 68400

ggagaacgtg acgattactt atgagggttg agggagaagt aggtatcaag catgattcta 68460

agacttcaag tttgctccac ggtttggaag ctggtcctgt ttgttatgac agggaactct 68520

ggtttaggat ttcaaattga ataagggacc tattgagctg gaggtacttg tggtgtagcc 68580

aagtggagaa atctaataag cagttggaca tgtagctcta agactctaca acgatgtaaa 68640

ggatgacaga gaattgagaa atcatgaaag tggatgagat tgcctgggag agttatcatg 68700

aaaagagaaa ggtcaaagac agagctcaga gcagctcatg actaaaggtc agctcacgac 68760

tagaggagat ggaaaagcag aatggcctna agaaatatgg acnaaaccag caccttcaaa 68820

gccagcagag ttgcatctgt gtcttcctcc catactcaca tctctctcca naccacagcc 68880

atgatcactg ttagtgattc atggagttag gttgggtcca cctggataat tcaggatcat 68940

ctcccaaatt taatgtcttt agccttaatc acacctataa ctttagttcc cttttgctat 69000

gcaacattat ctactcacag gttccatatg ttagaatggg gacatcttta ggggaccagt 69060

tttttctcta ccacaaaaac aacaattcag agcactgctg agaggtcaca gaagataaaa 69120

ctaaaatttt tctgattggg tatatcatat ggaagtgaac aaaagtaatt tcattgtctt 69180

ggagtcacag aaatgggcta aaaagttaaa agagggagat tacaagtaga tactttattc 69240

aagaaatmtg tgagtaagaa atgctggtct tgaaacccgt ttaaacatga gtggatattt 69300

attaccactg aaagaatcat tgtctatgtc caggattttc ctggtggggc aagaattcca 69360

tccaccccac ccccaactcc agagagtcac ttgttctcta tccaaagtac ctgagggaaa 69420

atgactacgg aaataattaa ggttaatgaa taaaaataaa ctctggaaga ttcttgttaa 69480

ctgaaaaaaa tgaaacaaag attaacaaat atatattctg aaactcttga cattcatttc 69540

aatttttaaa aacagtaatt accttgtggt tttagataca tgtgttccag ggcaattaat 69600

cttctaataa atgtgtttct aataaacatt tgctaataac ttgatgttgg taaaaatgtt 69660

tcaagcctgc tgcctctgaa gccaaagact tcaaaataat atctagtgag agttgtgctt 69720

aaaatgagta tcagcactaa tatgaaagaa aaagcaaata taataggaag aaaaaaggaa 69780

aaaatactct aatatttata aattaattta aaaaatcata atcattaaca gttgagtgat 69840

taacttattt ttggcaccat tctgagtgtt gcacattatc taatttaatc attaaaatgc 69900

ccctacaaag aatgcaataa ctatcctcat tttggnaagt gggataactg aggctataca 69960

cacagggctg gagtttcaat ctaggaaatc tacaactttc aaactgtcaa taacagtagg 70020

taacatctag cattgattgt ctatcaagta cctgttttat acaatttaca gatgtaaata 70080

cttacgtctc accccaggct tatgaggtag acgttttcac accattatgg attggaaaac 70140

taaagcatgc ttaagtcaca catgtaataa gtaggtcaga actcacagcc aagaagattg 70200

gctccaaagt ccaatttctt ctccaaatat ctgaaatatg attttttgg aaactttttc 70260

atgtagacag agtgttatct tagacatttc tgctattgta tttttaaaag tatttggaat 70320 taaaaatata tgtaattttg atgggctagt agacccttcc ctctctgaac ttcaccttgc 70380

cttcaagagt aaatgaagaa ctaggtgatc catatagtat ctcttagctg catagttcta 70440

tgaacatgct aaagtttctt ctcctctttt aacaagtctt tattgttctg tttttacatg 70500

ctgggattta tgctacattg tattattgta atgattaata atcatttaga tcaagcatta 70560

tttgagtaca gtacttgcaa cacattaggt gctatgcaaa gtgcattact tggtttactt 70620

tacttaaccc tcacatgagc tttgctttta attatgatta ctctctgtac cttacagatg 70680

aggaagctga gcttagaggc ccttagagcc ttatttaagt cacacaacca gtaaatggca 70740

gagctgggat caaacccaga aagcctcact ctgaaactca gaagcccagt ctccttacca 70800

tcttactgtt atattgtagg taccaaatca ttgatgacaa gcatcgttga tgaggttaat 70860

agtatttgta ccacacattc tgaaacagag tcagatcaga ctgaagcaaa gatgaatcat 70920

ttggcctaca tgtaagaatg agaaatgggg agatatacac gatgtgtctg tctaaaataa 70980

tatctttctg acccacccaa ttgtggtaat ttccacaatt agagcctttc ccttgggttt 71040

tacagatggg gtccaaagga ccagcttaaa atatgactag agggagcaat gatttaacca 71100

cttatttcct gagaccagca gacagatgag ctcatctgct gctgacaaag ggcagaaccc 71160

taaaggaata gcttagggtg tttgaatcat gggtgaagtt tctggagctt tgtttatggc 71220

tgagaagtgg gcaaagtatc accagccatg gggtgtgtgt gtgtgtgtgt gtgtgtgtgt 71280

gtgtgtaaaa ctatgtacgt gtgtgtatgc gtggtggcag caaatagcta agggattgca 71340

caattaggtg tcagacttta gagaaagcat attataaaag ggatggaagc agcaactgaa 71400

cataagaagg tattaaagaa aaagagtcat ctggtaagtc tacctcaatt tggtcagaaa 71460

ctcacgtgac tgaggctgtc ttagggttgt tctccattta agaaattcct agaaacattg 71520

atgtagtggg aaattagtgt aacacaatta tggatttaga aggcttctga aatcatcaag 71580

cccagacttt ggcgtctgct caatctgtgt gggatacctc agatagcagc tatttggggt 71640

ccatttttaa ggtcaaccca gaatgagagt ctacagtcag cagtcctatg ttatctatac 71700

ttgctgtcag gaaagccatt acatatctaa atcctttata tcaagttgat atccaaccta 71760

ttcttaagaa ggataaaaat gggttatgtc attatccacc taaaggattc ttatattcct 71820

agagctatca ctaaatcacc tcataatctc tctttcctat gggatacctt gagttccttc 71880

catttctcct tatttgtctg gttatttcta aactagacac cttccaacac tcagaaggac 71940

aagtatcatt gatgagtttc gtagtattcg tactgcaaag aaaatgattt tcatcatgca 72000

acattgctta gaagtgggga agtaaaatgt gtgcccacag tgagacaaga attttgtgaa 72060

tcacaatgac tcctcagaca agttaacgtt ctttgcaaat ctgtagaaaa attccttcct 72120

ttgcttatct gggtctcagt ctgtaatggc aaaaaccgta atggtgatag tattcttata 72180

aagagctatt gtggggatga atcagatatc atttgagaat gtttaatatg taattaatgt 72240

tatgaatatc atttatggat tatgaatctt aaaactatat tcagatacac atacacaatt 72300

ttgcaaataa catcatatat atatatatat atatatatag agagagagag agagagagag 72360 agagagagag agagagagag agagagaaag tctctcaagg tctcatatat atattaccaa 72420

ttaccacagt tgggtgggcc agaaagatat gattttagac agacacgtca cgtatatctc 72480

ctcatttctc attcttatat gtaagccaaa tgatccatct tcgcttcagt cttaactgac 72540

tctatatttc aaagtgtgtg gtacaaatat tattaacctc atcaacaatg cttgtcatca 72600

gtgatttggt atctacaaca taagagcaag aaggtctctc aaggtctctc tttctctctc 72660

tctccctctc taaatatata tatatatata tatatatatt ctctttgctg tactcttaga 72720

tagatagata gatagataga tagatagata gatagaaaga gcaaaggtct ctcaaggtct 72780

tgctctgttg cccaggctag aatgcagtgg catgatcata gttcactgca acctcgaacc 72840

cctgggctca agggattctc ctgcctatgc ctcccaaagt gcaacgatta caggcatgag 72900

caacctcact tggccagata catagatatt aatgaagggt aattacatta ttcattgtaa 72960

tcaacattta tgattgttga aattgtctaa cactttcatg tgttgttcac aggaatgtaa 73020

cctggaatag tttagctatt ccagacacac actgctcacc agagctagaa gaaggctacc 73080

agtgcccacc tggatttaaa tgcatggacc ttgaagatct gggacttagc aggcaagagc 73140

tgggctacag tggctttaat gagataggtc tgtcttactg gtaaaatacg ttttaagttt 73200

ctattttgtg atgtttatga agtggacaga ataattaaat aattgatact aatattttga 73260

ctggctgtag tatagaagaa atacaaatat aaattttaat tctatgccta ttttataata 73320

tattctaccc gacagtgcgt agcatacaag ctgtgtatgt atatgggcag cagtgtttgt 73380

atggagttac agatctctca tttttgacta tttcgattct caataacrgc atttaagcat 73440

ttaaaaatat ttgtttgttg aacagggttc tgggttgttg actcaaagat aaacaaaatt 73500

tggtttcctt tcatatgtaa ctatgtaact atacccctaat aggcaaggta agtcttaaaa 73560

actgtaatac taggtattga tgagtgcatt agagctgaaa agtgctcctt gattagaata 73620

caattaatga actacttagt gaatatgcat caactaagta ataaatgtat gggaaatgct 73680

tgtgcatgag aactatgtgc tcaaatgtgt acattcttct tggcttgcct tattaagtcc 73740

taaaatgtgc cattacatac ataatgaaag aatttagcat ctgccattta gttgatatca 73800

ttgtattttt ataattccat gcaagttgcg catattatat tattaaagca aaaaatgtat 73860

aattccaaac aagtatttta tgccttgaga attgttatca gagtaatttt ttaaactaaa 73920

atttcttcaa gtgtcaggcc tgcgtgctaa accataaagt tacagctttt atatagaata 73980

tagtactgaa cgtgggctga tggtatcttt ttaaagtcag ttttgcaaat aacatctcag 74040

gttttcatct ttatgtttta gtatttcctc atattttatt tcaagcatta tagttagatg 74100

gcattcttcc atctcttttg accagtaata ttgtctccta tttcacaaaa atccagaagc 74160

caactgatgt gaagcaattt caattttcta tactctctta aacttatcca actttgtaat 74220

tagtcttaaa tccttccctc tgtctcaaaa gtacagaggc atcatgaatt ataccaataa 74280

cttttaggca tgaggaagaa atcacagatg tggtatggga ggcgagatgt atataggctg 74340

ataggctgtg cggaataagc ttcaggagtt ccaacagctg ggcatgacta gaatgcatgc 74400 cttatcagag tctgcagctc tttcttttag ggtctgagtg tatccatgtc ttctatggag 74460

agtggtcacc attccagggt taaggttgcc cttacagaaa ataacctcca attttttaaa 74520

atcatgtagc taaagcacct cataccctca ccgtttttaa tgcaacaatt gtacccaaat 74580

tcaaaatgca gcattatatt caagttgatt catctttctt cactctgact tacaatggct 74640

ttgcttaata tagggtagaa tctctttttg agagaaaaag tctccttgta ctgtgctgac 74700

ttgagaggac cagttctagg aggaaattgc cttggaggca gtcgcccccct actaaagcaa 74760

ttatcttact atagctgtac tgttaattta tctcccaagg atttaagctc cttgaaaaca 74820

ggaagtgcat cttttcgtct tcaagtctcc agggcacagg gctatgcctg gctcacaggg 74880

agccctccat aaatatttac taataggctc aacaaagcgt atgtcttttt aattcattgg 74940

taaccaggca tgtttgtatg aggaggcaga gccacaatag agaatttagt tgttttcttg 75000

ttgttctttt gcctgatttt gtctgttttt gttggcaata ttcattattc ctcatgactc 75060

cttcctttct atacaggctt aatttcatcc gtttgcttat ttttcccaaa ttcatcatca 75120

ctgtatgact aattgttaga taatgaaatg ttgatcttgt accataacca aaacattttt 75180

aaatggtcgt tataatttag atcgaaattg atagctgctt cctgatatta tgtatctctg 75240

tagatacaca aactactata tttaaaagtt cttatggaaa caaaagacaa tttttgccc 75300

aaaatgagga aaataaaata ttttattaga cacatgtgat catctcttca tgcatgcatt 75360

tgaatttaaa acataattag gtgcctaatt atgcatatga atttggctaa tatacttgaa 75420

tatgtaagta attatggtta gttaaaaagg agcccatttc cttttcttac taatcaaatt 75480

accattcaaa acatgaattt caattcactt ataaaatttc tggcagaagt agctttatgt 75540

ctcattaaga tttaacactg atgatgttga aaaattcaaa ttttaaaata tattaaaagc 75600

cagagtcatg gagtagaaaa aattacaagg aacccctgtt gtaagattga agagtaaaga 75660

catctttgcc tctttgtgtc cccaaaatca tccaaatatc aagaagattg agaaacagaa 75720

atgaaatttt catgtttaag aactctgaaa agctgtaact ataaaccaca tcacaggact 75780

gggcatctaa cagaaaactt cctgagaaga ttctgcctac tgtggatctc agacagcacc 75840

agcaggatag caatattcta agataatgca tgttcctgga ggaacacaac tgtctcatat 75900

ttgagtcagg gaatcagagc aaaggctgat tgtggaggct tactaggacc ccatgcagca 75960

tcacaaggca gaggccctca tttctgcagg cagttggtgg ggagatagag tggaggggac 76020

acaactgctt tgttgctgct gatctgggtt ggctagaatg ggataaaagc taaacggaag 76080

gttttcatgg ccactataac ttaggatgac ctggggaacc ttttaaaaca aacagattct 76140

caagcaccat cctcatgtca attaaatcac aaccttagag tgatccaagg aatagatggc 76200

ttttaaaggc tctccaggtg attaatttat gtgcatgcag agttatgaac cattggacta 76260

aagaaataca cacatagaca cacacacacc tgtattataa aaaccttcaa ttccatcttg 76320

atccattttg gttgttgtaa caaaatgtca cagaatcagt gtctcataac caagagaaat 76380

ttatttctca cagttctaga gtctgggaag tccaagagca agacagcagc agatttcatg 76440 cctggtgaag gctcacttcc tggttcatag aatggcacct tttcactgtg tcctcacgta 76500

gtagaagcag taagggttct ctcttgaacc tccttcataa ggttactaat accattaatg 76560

tgggctccac cttcatgatc taatcacacc caaaggtcct accttctggt agcatcaccc 76620

ccatgggggt tagaaattca acaaatgaat ttgggggaaa catcagcatt ctgtccattg 76680

cactatccta gaacattacc ctgttcctcc ccaaagtgaa acagtgacag acatctggtc 76740

taacaaaaat ttatttcaaa gatgagtcat aattgaaaga atgaggacca catctataaa 76800

aatgtacttt gaggtacttt taaaactgaa aaaaaatcat gatatataac tcataaaagc 76860

ttagtaggaa acagagcact gaattacaga ttcgaaggat gcagcttttc tcagtaagtg 76920

aacatagaac agtcaacact gagacacatc ttagtaaaat tgttggaaat caaagctaaa 76980

gaaaacattc attaggcatc cagaaaaaat aaggtcaagt tagctataag gaggtacatc 77040

agagaggcct gaaacttcat gtcaatatgg aactctagaa tgtaatggca caatgcttac 77100

aaattcttta gggaaagaat gggtcaagct ctttctgaat atttttgaag aggcaggaac 77160

atagagacta tattttgcat gaaatcttcn ttgaaaacat tcagttaatc aagaatttct 77220

tatcatgagc tagtctctag tatctatttg ttttaagaac aaagctaata taactgtagg 77280

atttacaaaa tacaatgtaa atattataaa ttatgaatat ctagaagaga taagatgaaa 77340

accgggggag atagtgtgat tatgttgata ttcttaacat tcacatttta gcagttaaaa 77400

aagtaacatt tatacaatat atgaaatgta taatagagat atgtgcatac ttgtcatgag 77460

atgagaatgc atgataaact cagnaaatta ctttctttag aaataactca aaaatccaaa 77520

taatattatc tatgaccacg tcccaaaaaa tcatttcaat atgagtaata atattggaat 77580

caattctttc taaatcttga taatagttct caaaagaaaa gtaacaaaca gtatattttt 77640

taacctctga atagagttat gctatttgct gtagcatgtt taaaattatt aggggcacat 77700

gactcctttg tggtaaagtt tgcttgaagc acagttgttc aaaagaatgg ccctggggtt 77760

tagatcatga ctccacttta tattagtgtg acagccttag gaaagtaatt aactctcttg 77820

cctcaatttc cccatctagg aaatggggta agtgatagac ctaccttata cagttaatgt 77880

aaagagcaac tgagaaacta cacttagagt acgtgtagcg tagaacaaac gcagaatgaa 77940

ctagaaattg ggatgcacgg gcaaaatatg aatgtatctc tgtgttattt gagggaaaat 78000

ggtgtaataa taagggaggg caggacatag atattttcac actcacacac tgcatctgat 78060

atgattccct ttgtttacag acattaacta tgcaggtttt ccagtatctt attcaaggga 78120

gaggcattgg ggcaggggca cggtgttttt tgtttgtttg tttgtttgtt tgagatggag 78180

tctcggtctg tcgcccaagc tggagtgcag tggcacaatt ttggctcact gcaaactttg 78240

cctcctgggt tcaagcagtt ctcctgcctc tgcctcccag ttagctggga ctacaggcgc 78300

ccatcaccat gcctggctaa ttttttttgta tttttagtag agacggagtt tcaccgtgtt 78360

agccaggagg gtctcgatct cctgacctcg tgatccaccc gcctcggcct ccaaaagtgc 78420

tgggattaca ggcgtgagcc actgcacccg ccggggcacg gtgttcttaa ccttgcatct 78480

```
tgcctatgtt aactcatatc tcccatttca ttttattcca atcaatgact agcattctag  78540
aaggcaagtt tttggtgatg tttggtcttt attcatcaat tattcccctc ccctctcaca  78600
aatgtcctga cttctgtgct tgtaccaatc tctctagaat ctgatcaaca tttagcccgt  78660
gtcattgaat tttctagttt gatgaagcat catgcatatg tttaatttta tttatttatc  78720
catatacatt cgaatctgtg acccctggat tggagcctct ccctttcagc gcctttcaan  78780
caggtgctcc ttccacagag cctaagtaaa gcaccagctg gggcattgag cagaaggcac  78840
agccatgctg tgatgtgcag gggctccagc actaagtgat ttattttccg nacctcctct  78900
aattccctga gccacgtgac aactgagata ctagttattt ctgtctcttt cttccattta  78960
tgcacctttc catctttatt tctggatagc tttcctgtgg ctccttccta aggaccccca  79020
actgactcat ccctgctaag acattagccc gtgcactcaa gatatcacac actccccatg  79080
ttgcaagctt cctgagtagt agacataatc taaaggacat cttgcaaaat cgcaaaattt  79140
atttagcaga cattgagatt ccgtgattgc agtatgggtg gcaaggtggg taggaatcat  79200
acagtggtaa ggtttgggga ttgacaacat cctaaaagcc acactgtcga agtgacacaa  79260
aatcacagca acttacccta tttagaaaat gtgtcttccc ttttctgaat acttacccca  79320
ctttcagttc ctacttccgc tgnttcccac ttgctcataa acatatgaga cataaaagct  79380
gagacataaa agaggcgaca ctgtgatttt ctgtaataac actggaaatg aaattacaga  79440
agtttagagc tggaacagat aatgcagtca tacattttac ccacacaggt ggggatctga  79500
gagtcagagt gagtgggatt agaattacag agaactccgg gcagaggcag gctgaactca  79560
cttctccaga tttgcagatg tccaccatac ttgctgtttc caccaaaccg tcttacaact  79620
caaaatagtg tttaagaaga aaattcaaaa acaatttgca aagcttagtt ttattcctaa  79680
ttttctagtg gaatgttagc attatttccc tttagtgaaa tgctgatggg gaagaccgtt  79740
atccaaacag aggtgatgtg ccagttagaa agaaagcaac acaaaatata ccatatatgc  79800
gaaataaatc acagcaaaaa aacaccatcg aactggtgat tccttgtcat gaccttttaa  79860
aagaaagtaa aacatttggt acaagagatg agaacaaaac tattcaggtg gttatttgaa  79920
cccttcattc attgtacagt caggtaaaaa ataaaaaatg ttcagtttgg actaaacaac  79980
ttaatgaatg tcattggatg ataaaactca tcctgaaaag tcacttagga agtaaatcgc  80040
tatcatgcat tttttattag tctttagaat catcagaata aatgacaaaa gcccattttc  80100
aaaaaaatcc agaataaatg aaattcaact gaaacagctt atagcagagg aagcccacaa  80160
gcaagataaa ctatgaatca agacattctg aaaataacta atagaaaaga catacaacat  80220
atgcccttaa cactacacga gctgtatatr gtggatcact ttgcctgttc tatatattca  80280
aaagtcgaag gtaaagccag gaaataataa acactgtatt cttatggtga ctaattgtcc  80340
atcagacagt cgttactttt gtaattactg tcattactta ggacaactgt tggntatcat  80400
ttatcgattg catattatgt gcctggcacc agtgaaaagg ttttgcatcc attgtgtttt  80460
taatcctcaa aatcagccca tggggtatca ttatccctca gctgatgagt ggggagaatg  80520
```

agataagcaa gaatgaactc tcccagagcc tcttaccagc cgcatttgga agccatgtct 80580

tgctgccccg tgtgctgagg ctgaagcctg tgctcttgat ctacaaacta ctacgttacc 80640

agagtagcca gccatgagta accttcattc ctgtggcttg tcaggacacc tggtccatta 80700

taaagtctca ttatgatgat ggaagaaaca catcttccaa aagagtgacc tgccatagtg 80760

gcatatcctg caacatggtg acctccctct tgtcctgaag tctttagaca agtcttagaa 80820

aaaagaataa atgtgggctg ggcgtggtgg ctcaagcctg taatccaagc actttgggag 80880

gctgaggcag gcagatcacc tgaggtcagg aattcaagac cagcctgggc aacatagtga 80940

aatcctgtct ctacagaaat acaaaaatta gctgggcatg atggcgggtg ccagtaatcc 81000

cagctactcg ggaggctgag gcaggagaat cgcttgaatc tgggaggcag aggttacagt 81060

aagccgagat tgagccactg aactccagcc tgggcaacag agcaagactc tgtctcaaaa 81120

aataaataaa taaataaata aataaataaa taaataaata aataataaag agtaaatgtg 81180

aaggaatgtg acgtccactg caaactgatg gccccttaag gggttggtct ctggctttaa 81240

gtctagtccc ttgtattctg tattctgtga gacagagagt tggtgcataa tatatgttag 81300

ttgaatgaaa gaatataaat attaaacaat gtacctggga ggcaatgcag ttttataagc 81360

tactgagctg taaacatctg ggcaaggatt agagataatg ccagaaaaaa ttagtaaata 81420

ctacttttga tgtaaaatta ggtggtagac acatcatggc tagagatgtg ggagagaact 81480

tagcaagcaa aagaaaatat gctatttggt atattttagt aagttttgtg tgtatttgaa 81540

ctaaattcaa aatgccgtga taaggcatgt gtttgctata tgtaatgtat agctacatgc 81600

attaatttct gttgtgtatt tgaatttact ttctttgaga tatgaatatt taaatttgct 81660

ttctaaatta tcacaaatta accaacaaat ttatacaagg tgattttgga aatccacatt 81720

aattaatttc tttattagaa tgtgacataa acagtattta ccttattagt ttntctttac 81780

gaaaaaatta taaangcttt tctatatata tcctccttaa aatattagat aagaaaacta 81840

caagtcgtca gctccacagg agttaacatg tggattttgg atgctaagaa caaatttaag 81900

atgatgactt gtaatcttca tgaattgctt ttggcaacaa taaattaatt gggcctttca 81960

agaaaatgca gaaaactaag gtaaagggtg attgagtctt ccagatttcc tatattccca 82020

actcacttct gaagaacgta attatgagta ttgaaatttg ggactggtat tttttacaag 82080

aagaaatata ctgtagctgt tatttacttt tgtggtattt ttcttctttg aatttgggta 82140

gtattgacta taagctatgg actcagctcc ctgtgtgggt cataaatcag ttttattaaa 82200

tattgattta gaaatatcag ttattatttt cntcactggt ggagttctgt gcttcggact 82260

acccctgtat cttgctgaaa taggattgac aat 82293

<210> 2

<211> 237961

<212> DNA

<213> Homo sapiens

<220>

<221> exon

<222> 43726..43868

<223> exon 8

<220>

<221> exon

<222> 43998..44102

<223> exon 9

<220>

<221> exon

<222> 52093..52179

<223> exon 10

<220>

<221> exon

<222> 77568..77699

<223> exon 11

<220>

<221> exon

<222> 98226..98393

<223> exon 12

<220>

<221> exon

<222> 106567..106758

<223> exon 13

<220>

<221> exon

<222> 144109..144246

<223> exon 14

<220>

<221> exon

<222> 159794..159868

<223> exon 15

<220>

<221> exon

<222> 191292..191428

<223> exon 16

<220>

<221> exon

<222> 192967..193108

<223> exon 17

<220>

<221> exon

<222> 211540..211613

<223> exon 18

<220>

<221> exon

<222> 225006..225107

<223> exon 19

<220>

<221> exon

<222> 225544..225613

<223> exon 20

<220>

<221> exon

<222> 228450..228541

<223> exon 21

<220>

<221> exon

<222> 228630..228752

<223> exon 22

<220>

<221> exon

<222> 231289..231345

<223> exon 23

<220>

<221> exon

<222> 231589..231709

<223> exon 24

<220>

<221> exon

<222> 231813..231944

<223> exon 25

<220>

<221> exon

<222> 232900..233067

<223> exon 26

<220>

<221> exon

<222> 235355..235459

<223> exon 27

<220>

<221> allele

<222> 51090

<223> 99-79335-60 : polymorphic base C or T

<220>

<221> allele

<222> 61293

<223> 99-79336-369 : polymorphic base A or G

<220>

<221> allele

<222> 80602

<223> 99-79338-332 : polymorphic base C or T

<220>

<221> allele

<222> 100485

<223> 99-79314-201 : polymorphic base G or T

<220>

<221> allele

<222> 100509

<223> 99-79314-225 : polymorphic base A or G

<220>

<221> allele

```
<222> 106725
<223> 99-79316-158  : polymorphic base C or T
<220>
<221> allele
<222> 166087
<223> 99-79322-224  : polymorphic base G or T
<220>
<221> allele
<222> 166336
<223> 99-79322-473  : polymorphic base A or G
<220>
<221> allele
<222> 235894
<223> 99-79306-182  : polymorphic base C or T
<220>
<221> primer_bind
<222> 51031..51051
<223> 99-79335.pu
<220>
<221> primer_bind
<222> 51539..51559
<223> 99-79335.rp  complement
<220>
<221> primer_bind
<222> 60925..60945
<223> 99-79336.pu
<220>
<221> primer_bind
<222> 61354..61374
<223> 99-79336.rp  complement
<220>
<221> primer_bind
<222> 80271..80290
<223> 99-79338.pu
```

<220>

<221> primer_bind

<222> 80700..80720

<223> 99-79338.rp complement

<220>

<221> primer_bind

<222> 91037..91056

<223> 99-79339.pu

<220>

<221> primer_bind

<222> 91466..91486

<223> 99-79339.rp complement

<220>

<221> primer_bind

<222> 100285..100305

<223> 99-79314.pu

<220>

<221> primer_bind

<222> 100764..100784

<223> 99-79314.rp complement

<220>

<221> primer_bind

<222> 106568..106585

<223> 99-79316.pu

<220>

<221> primer_bind

<222> 107000..107020

<223> 99-79316.rp complement

<220>

<221> primer_bind

<222> 165864..165884

<223> 99-79322.pu

<220>

<221> primer_bind

<222> 166381..166401

<223> 99-79322.rp complement

<220>

<221> primer_bind

<222> 235713..235732

<223> 99-79306.pu

<220>

<221> primer_bind

<222> 236190..236210

<223> 99-79306.rp complement

<220>

<221> primer_bind

<222> 51071..51089

<223> 99-79335-60.mis

<220>

<221> primer_bind

<222> 51091..51109

<223> 99-79335-60.mis complement

<220>

<221> primer_bind

<222> 61274..61292

<223> 99-79336-369.mis

<220>

<221> primer_bind

<222> 61294..61312

<223> 99-79336-369.mis complement

<220>

<221> primer_bind

<222> 80583..80601

<223> 99-79338-332.mis

<220>

<221> primer_bind

<222> 80603..80621

<223> 99-79338-332.mis complement

<220>

<221> primer_bind

<222> 100466..100484

<223> 99-79314-201.mis

<220>

<221> primer_bind

<222> 100486..100504

<223> 99-79314-201.mis complement

<220>

<221> primer_bind

<222> 100490..100508

<223> 99-79314-225.mis

<220>

<221> primer_bind

<222> 100510..100528

<223> 99-79314-225.mis complement

<220>

<221> primer_bind

<222> 106706..106724

<223> 99-79316-158.mis

<220>

<221> primer_bind

<222> 106726..106744

<223> 99-79316-158.mis complement

<220>

<221> primer_bind

<222> 166068..166086

<223> 99-79322-224.mis

<220>

<221> primer_bind

<222> 166088..166106

<223> 99-79322-224.mis complement

<220>

<221> primer_bind

<222> 166317..166335

<223> 99-79322-473.mis

<220>

<221> primer_bind

<222> 166337..166355

<223> 99-79322-473.mis complement

<220>

<221> primer_bind

<222> 235875..235893

<223> 99-79306-182.mis

<220>

<221> primer_bind

<222> 235895..235913

<223> 99-79306-182.mis complement

<220>

<221> misc_binding

<222> 51078..51102

<223> 99-79335-60.probe

<220>

<221> misc_binding

<222> 61281..61305

<223> 99-79336-369.probe

<220>

<221> misc_binding

<222> 80590..80614

<223> 99-79338-332.probe

<220>

<221> misc_binding

<222> 100473..100497

<223> 99-79314-201.probe

<220>

<221> misc_binding

<222> 100497..100521

<223> 99-79314-225.probe

<220>

<221> misc_binding

<222> 106713..106737

<223> 99-79316-158.probe

<220>

<221> misc_binding

<222> 166075..166099

<223> 99-79322-224.probe

<220>

<221> misc_binding

<222> 166324..166348

<223> 99-79322-473.probe

<220>

<221> misc_binding

<222> 235882..235906

<223> 99-79306-182.probe

<220>

<221> misc_feature

<222> 6485,14246,14248,14250,14252..14254,14256..14262,14969..14980
17712,18786,19477,19496,21043..21064,21547,24607,24946,24974
25072,25124,25136,25142,28843,31239,31372,37750..37751,40488
46542..46557,49309,49326,49357,49395..49398,49400..49402
49408..49409,56275..56280,61048,61050,61059,61856..61857
61860..61861,64842,66307,66326..66327,69835..69840,70893,70905
73279,76154,76354..76355,76371..76372,76407..76408,77095
78933..78934,79363,89491..89492,95836,95838..95839,95841..95842
95958,96966,97049,97067,97078,97292,98642,98712..98713,98716
98740..98741,101368,101373..101374,101376,101378,101401..101403
101406,110014,129061,129066,129973..129974,135336,135345,135359
139924,147368..147369,147469,147507..147510,147767,151694,151729
163111,164480,169764,173340,173390..173391,173398..173401,173407
173414,173419,173426,174800,174894,175041,175486,176256,177240
178864,181149,181154,181230,181287,183433,186606,186648,191065
191267,193117,193203,193223,193380,194051,194053,194057,196577

198388,198655,198828,198849,198862,199667,199693,199701,199704
199747..199748,201117,201224,201226,201314,201410,202524,203887
204396,204939,205891,205930,205998,206908,207363..207364,210938
212018,214253..214254,217022,217333,219027,219909..219912,220924
222620,222882,223870,224902..224903,224911,224914,226899,226915
227232..227233,227415,228441,228446,229373,229471,233305,233405
233512,233601,233629,236096,237468,237471..237472,237478..237482
237486,237488..237490,237492,237495..237496,237498..237499
237504..237507,237509

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

<400> 2

gatgttaaat aaggcgataa aatatgagga agaaaaatac ttttattcaa attataatgg 60
gatgtaaatt gaattttgtt gatgctgctg ctaaaaattg attcatcaga gttaaatgat 120
attaaagcat atttcatgct accatatgtg atatcagatg agaactttac aaataaatag 180
aaaacagaaa gttatatatc tacaaccatg atttattaca gaacgtatat ttttactgca 240
gtcataaaag aaaaagccca gtaaagacta atttgtaaaa tatgtaattg agatttataa 300
tgctttcaaa aaaaaaaaag caccaagggg tttttctgat gcatttatct tgatgtttac 360
tgttgccata ggaattactt attgattttt atgaaaacac atatggaaga gagaactaag 420
gcaatagaat tcaactaaaa gggcttcatc tatgataatt ctgtttgctc tgataaaaat 480
tcctttatca tggaacaagg aaaacctgac atatttttgt tgagtcgagg acaagatcaa 540
gaaacatcct atgaggaata tttgtgaaaa ctcttagcca aatgcaagga gttttatcct 600
tgtatcttga atcataagtg cccctaggtt acatgaaatg tgtgcttgag gatttttaaa 660
gggactgtct ttaggtttct caggttctca aatagggcca cagacttact taaactgaaa 720
ttaggtgtcc agatattctt cagaatgtta ggaagaaagg agtttaaagg cagaaattgc 780
aggcatgact ttttaaaaat aaggttgtcc attgtctgct gaaagatgtt taggagttct 840
tgttgaatgg atgagaaaag atagtgacta gaggagtttt tcatcagaca gtgaaaggaa 900
attttaaaag taaattttgg agagcataaa gtgatgattt atttcatgtt atatgtagaa 960
gacagtgttc attggaaata tctagtggat tccaagaaat tagcattttt atggttgtat 1020
tttaaacatt tataagttaa aaaagtaata cttcaaataa tgtaagagat ttaatgattt 1080
aagaaagtaa taaattctgc catatataat tccctccctc cttccctctc ttcctttctt 1140
ttctcatttc ttctcccctc tctgtaagtt tcatggtatg ttataacatg tttttcattc 1200
ttcccttact gatagacatt taagtggttt tcagtctttc cctgcaataa gtaatgctgt 1260
gttgaataca ggcatatata cattctgtgc tcatgtgtga ttatgcttta gaacagattc 1320
ctagaaacaa tttctgggtg aaatgatatg cacattttta atatgagata gctggcggtt 1380

tactttttta aaagataagg attcatagta ctgcctttat tgcgtgagtg aatttttata 1440

aatacaaggt attctttttc aatatgcata aatattgaag aaaaaaaata atgtaggcat 1500

ttcgtattcc gtactctgta ctcttctgaa cactggcagt ccattctctt cttgaatgta 1560

tgtattcatc ccattcttcc tggaatgcct ctctccattc tcttcttgaa tgtatgtgtt 1620

catcccattc ttcctggatt gcctctctcc taccgctccc cctacccccca tctcctgtga 1680

ctcacctgtc aggacctact caggtaagtt tcccctaatc tgccagcagc actcatctcc 1740

cttctgtctg aactgtgata gcaaggtttg tgcatcaaag tggacacttc gttacttcgt 1800

tctacatgtg tcccaccgtc tgggagtgtg tgtgtgtgtg tgcacacgtg cgtgcctgtg 1860

tgtgttttac ccttatgttc tggattcgtt cattagaccc agttgtctca ggtgggagca 1920

gagcctcagg cacttctttt gatgtctttt ctcgctctca ttgaactcgg tatatacttg 1980

aggactggct gatcacattt atgagaattt tatatttgat agattgcctt ttgaaaatca 2040

tcattctgca gaaatttagt aattgaaaat ttaaatcaaa ttcaaataca actctaggct 2100

ctaatttcac tgttcgtttg aatgtatgac ttctctgtta agccagatac tcacgtgaaa 2160

caaatgtgaa tagaatagca tggaaatcgt gtatggacag tcaaggaaac tttctctgct 2220

ggcgacagca gcctgcagaa cataattccc tggacactct atgaaatcaa atccgtttta 2280

tcaactaatc agttggagca aaaaataata atagtgttat ttaggaaagc cgtaacacag 2340

ctcttacctc aaggctaatg aagacctgac ccttgaactc taggggcttc catgttccag 2400

cttgtgagcc ctgctttgcc aaaaccctgc tgaagaggta agagcctcct tctctggaga 2460

gaaagccccc agcctttgac aaagtgcttc tagcctgagc catgtgggct atgaaaaaga 2520

agttggtcat cccgaggagt cactggacag tggagagagc ccatccattt ttctccttta 2580

ttcttcccta cccctcagct ctcactccct taggacttgg gattctggaa tcataaaagc 2640

aaaatattct aaatatacta aaacatgcta aataacttaa atatgctaaa taaaattaaa 2700

actaaaataa acgtattaga ataaatcatg aatggctaat aaatataagc tgccaccatt 2760

aaaacaaaaa taacaaccac taccacaatt gctactataa agattattat tgatcagctt 2820

tatttttagg taagtagaaa agtgagatgt atttctattt ttctgagaat gtaaccattc 2880

tgaaagaata ggagatactt tctcacagta agcagaagaa tgagatgtac tcataaataa 2940

tatattcaac aaacagttat tttgcactta ctctatatag gtactgaaaa tagaatgatg 3000

ttaataaata tgaacagccc tcaaacagag atattaaaat aaaacaagaa atgaaggtga 3060

aataaattta gaagaaacga tggatcatta acttgtagta gttcttacag ccatgtgctg 3120

gatgcaggca tccctgtctg agtgtgtgcg tgtgtacaga cattgtttcc aaactttttt 3180

actataactt tgttaaagtg aagcgtgagc actccacatt tgctgcaagt ctcacttccc 3240

ctatttcctt tatttctcac atttacgtaa cttgtctaac ctctgaaaca aatgttttta 3300

ttagttctgg taatgcaaaa taatctttag agaattacgt ttttattttt gcttttgtta 3360

aataagatct acagtagcat aagttaagtg tgccaccttc attttgtagc tggctaaata 3420

```
tttctgcttc ctgtgtactt ttctgaaaat ggctttaatt tgtcttattt ttgataatta   3480
gcttcaatta gtgcagaatt ctcctttggt attgattgga cataaaacct gagatatgtt   3540
cgttgtattc tggaatctat tactgccaat aacaagccta ccgtgagtcg aattttaagg   3600
ctttcctttt tttttttttt tttttttttt tggctttgtg aatttgggtt ttttttaaca   3660
tcgtttttct ctggtagttt ttaagattta ccctttgtct ttgctgtctt atggtcttag   3720
tatcctgtgc ctaggcatgg atttttattt tttatctact gtgaactcat tctctgcata   3780
tgagactcat gcctttcttc aattctggaa ggtttttagg tattatcttc tcaaaacagc   3840
tctcttctcc attctctcta ttctttgttt tctggaatac tgattaatat ttgtttttat   3900
catatttttt attctatcct ccatgtcttt aaatcgtgct ttctaacttc cctctatttt   3960
cttgacttct cttgttacca cacaatgggc gggttcagtt gcttggcagg tgggagtcca   4020
atgaccacaa ccaaggagga tttaaccaag ggatttcatt actggcaaca agtaagtaga   4080
acactgggga tagttcccaa agcagtgcct ccccagacac aggtggaaac agggctttgt   4140
gggggctgct tagctgagtc actgtgtgca gaggtggagt gcaggcaggc acagctgttg   4200
ttcatgcttc ttcatatgtt gcccatacag aaaatggcac ataagctctt ccctgggcag   4260
ggcatttagt atggtaaggt gaagaggtct ccaaagttca tctccagctc aggcatcccc   4320
ggatgcctgg gaccagtttt ttgtgtttgt tctttgtttt gcaggggctc ggcttcttcc   4380
tagaacactt ttaaacaacg gaaactcaag gtgcaacagt taaaagtgga tagttcttca   4440
cagtgtgtac ccaaaaaccc agggaccctg ggttacactc ttgatccttg attagttgca   4500
caaggaatct tctatcattc tccgtccctc caagagttgt ttacgtattt tcaagctcaa   4560
attctttgct gaactcttaa ttttctctca gtgtgtctag actgctattt atttatcgaa   4620
ttttcaattc caacatctat aatttttata cctaatattt ctatctaatt ctgtttcaaa   4680
cctgcctgtt cttttatctt agattcctgt tcttgctgtt ttctttctaa tcttttcttc   4740
atctcaccaa acattttaaa catgcttatt ttagaatttg tttctaatgc ttccatattt   4800
tctaactctc ctttacaata gttcattttc ttgtgtggtt tataatattg gctatgagct   4860
catttccact ggctatgagc tcatttttca ggggtactct ggcctgaact gtgactgggg   4920
ttgtgtttgc tgcagtttca tggatatcaa ccgttcttac tgatactcgt gttacttttg   4980
tggcttaaag aatatattaa tgtgttctca tgctgctaat aaagacatac ttgagactgg   5040
gtaatttata aagaaaaaga tttaatagat tcacagttcc acatggctgg ggaggcctca   5100
caatcatggc acaaggtgaa gaagtaaagg tatgtcttac ttggtggcag gcacatctta   5160
catgggggca ggcaagagag cctgtgcagg gaaactgccc tttataaaac cgtcagatct   5220
cgtgaggctt attcaccatc atgagaacag cacaggaaag actcgccccc atgattcaat   5280
tacctcccac tgaatccctc ccaggacaca tggggattat tatgggagct acaattcaag   5340
atgacatttg ggtggaaaca cagctaaacc atatcaggga aggagaggtc cttgaactat   5400
ttgcattatg ggggcactgt aaaattagac tacacacttc agttggccca ggtctggatt   5460
```

ctgatttctc tgaggagact cttttgcccc gcatggatct ggacagaagt caagcaacat 5520

tctgcttttt tctgagctgg aggccagggt tcttccagtc cagataaaat gcagaacaat 5580

cctttgtaac acccagcttt gtagagacac ctcagctcca acttactgtc ttatattagg 5640

tgcctcccaa ctcctctctc tattgggacc cagaagccct gaagatttat attgactcca 5700

agtacatctg tgaaccattt ggcttcagct gctgctcaca gctctgattt tatattctcc 5760

caaactctct gaccgctgga tagtttgctc atgtagcttt gagataattt atacattcaa 5820

aacttactct cttttatcta tttccttgtg ttgatttggc atagggtgag gagggtaagg 5880

gaaaaaggtg tctaaataag catagtccaa cacattgacc taaagcccat atatatatat 5940

gctgctaata acattatata tatgtatata tatattatat ataatttttt ttgaaacagt 6000

ctctctctgt cacccaggct ggagtgcagt ggcatgatct cagctcactg cagcctcctc 6060

ctcccgggtt caagcaattc tcatgcctca gcctcccaag tagctgggac tacaggcgtg 6120

agccaccacg cccggctaat ttttatattt ttagtagaga cagggtttca ccatttcggc 6180

caggctggtt tagaactcct gacctcaggt gatctgccca cctcagcctc ccaaagtgca 6240

gagattacag gtgtgagcca ctgtgtccag cctatgtttt tatattttta ggattcccat 6300

gtaatgatgc aattttcctt tctttgatat gatgcatcta tgattttata ttctaaatgt 6360

atttttatac ttatatataa aatgcatttt attaattaag cattagttct taaaagttgg 6420

tatttccttg aacttttaat tttacaaata ttccatttcg gtgttattat ttttagtatt 6480

ttttngctgc tgaatagtta tgccaacctc tgtgcagcct gggaagacaa cacatcagca 6540

gtcaatttgt gtaaagacaa aataatatta gagagcaaga gaatgagaca agactgcaaa 6600

tcagcaatga taattattga gtactgagaa caattccagt ggattaagat ttgggatgct 6660

taagcttctt aatattaatt gcagcaaata ttcaaatgct ttcactttcc atataagtac 6720

ttcagcaaat aagcaaaata tgaaataaaa accctaattc agaattacct aaatataaaa 6780

caagtaaatt gtcctaggca attcttacca tcttatgtcc agtaaaacaa taacctatat 6840

ttgtgcatac agatatatta tctcatttaa aaactagcgg tgattgtaaa tcttgacacc 6900

aaattgaagt tgaactgaaa gaggctttcc tgtaaagata cttaggatcc ttacattgta 6960

tgtaaggaaa gtcacctagt cctcacatca accccaggtg tgacatgtta ttaacttatt 7020

ctaatggtac agatgttgag atgcagagaa cagaagcagc tcttagggga cagagacagg 7080

gctcagtgat gattatattc acctgtcatg gggtagagct acaatagtga gtattctgtg 7140

gatggtccta tccagataaa agaaggaatt ggtttcagga catagtctgc atcttactag 7200

gtggggaatg ttggacaagt tacttaacct cttctgccct gtaaaatagg gataaaaatg 7260

gtacatatca cagagtgtca taggaagact aaagtagact attctcatcc tcacacctaa 7320

tgagttctca atatttaact gattataatt atctttacta gatacataat taattaaata 7380

aattaaggaa ttatagagca aatagtcata gggtcacaca gatagaaagt gggaaaaata 7440

ggagacaaac ctatgtctat attcattcca tggcattaga tttagattac ctccttttag 7500 tctggcaaaa cactcatgat tatacaatgt aatttaaaaa tatacagtac tttgctgggt 7560

gtggtggtgc atgcctgtgg tcccagctac ttgggaggct gaggtgggag gatggcttaa 7620

gcctgggaag cagaggttgc aatgagccga gatcacgcca ccgcactcta gcctgggctg 7680

tagaacgaga tactgtctca aaaaagaaaa aaaaatacag tactcacatt ttactttgct 7740

agtatatgct gtattcactg gtgtatgctc tcttgcatgg agactttcat tgggaataaa 7800

atgtttgaaa ctattttact atttcttgca ctgctttgag atactaccag caaagctatt 7860

tctaagctaa gacgctatgt cctggcagga cttctaaaga cttaaagcaa tgctgtagct 7920

gtgtatttac aactgactag agcagggatc agcaaaccat ggccagtggg tcaaatcctg 7980

cttactgcca gtttttgtaa acaaagtttt atcgcaactc cctgcccatt tgtttagaca 8040

tggtctgtgg ctgcttttct ccaacagcag aaaagttcag tagtttcaac agaaaccctg 8100

aggcccataa agcctaaaat gtttaccatc tggccattta cggaaagtta gccaaccccc 8160

agactagaat aaaggtagga aaaaccactc agcccaagct tctttttga ttctatcttc 8220

ccacagctgc ctagaatcct ttgcagtctt tatcctagga ggctcagagg ctataacaca 8280

tctctaattt tgtctagatt tgtgagaatc catttctcac aaatccagtt ctggccctct 8340

ttccctgact cgatgttcct tgaactctaa acatgaattt ttaaaagaag gattacctgg 8400

gatgggtata tccctgaaag tcacctcctt tctgctgcct ctcacaaata cacaggctat 8460

gtgccttaaa cctgtcattc actttaagga atttttgaaa agtggataat aggttagaca 8520

taaaattcac tgaatctggg ttaaatttaa tcagtgagta ttctttagaa tggaagcatg 8580

agcctgaatg attaaagaaa ttgcttaaaa aattagaatc aaaatactaa atatcatccc 8640

ccccccaatg ttttagctag aaaaattact tgaaaatgag ccagagaagg agaaaaataa 8700

ttatcttaca gaaaactaat aaaaatacaa taaactatca tttatttaaa tacaataaac 8760

tatcatttgt taaaatacaa taaacatgca aaagcgctat cattgatagc acttttaatg 8820

tttaaataac ccactgatta ctttctctag tggattattt gcaagggtag agtaagaatc 8880

gttgcttctc tggacacata acccatttgc catgtgacat tgcctccctt ccaccaagaa 8940

gtggaatgta tttttcctcc ccttaaatct gatgcctata actttctttc cccttgagat 9000

aatggcagag tggcattgtg caagttttgg aactaaggtc ccacgaggcc ctgcggctct 9060

ctctcttgtt tatctgagta ctgccctaag atgaccatgg agagaagccc agtctagcct 9120

agaaaaggag aagaggccac atggagaact gaggtcagcc agccccagct gcttgacttc 9180

tgaacaaagc cttctccgac tatttggccc tgttagaccc agctgagatc agcagaagaa 9240

tcacccatct gaccctagcc caaattgcct tctgcagaac cacgggccaa taaatagcag 9300

atgctttaac ccaccatatt ttgcagtgat ttgttatatg gcaacatatc attgaaaaac 9360

cctcatttta ccaaagaagg ggggtcatga cattaagcaa ctacctaaac ccacaatgta 9420

aatggcaagt cagggattta gatgcaaatc attccaacta aaaatgtata ctctaaatca 9480

atgccctgta ctatactctg tattttccaa aaactaatta gattgtgtca tcagcatttc 9540 tgaagacaga ggtagtcgtc tgggaagtag agttttctca accagtattt tgcacttgtt 9600

gggatttttg agatcaggaa tgttttggaa aacctgagat gtgcagtcac atctctgagc 9660

agtgttttaa gtttattcat tcactcattc attcattcag caaatattta ttgagatcta 9720

ctatatgtta ggcactatgt ttaccactgg aaggcaggca ccaggtaagt acaaacacag 9780

aaactgagat ggatggcatt gcattttcat ttgattacaa atttattctt gcttataagt 9840

tgattaccta ccttattaag agctgatttg tttcaagttt tagaactcca attgtgtcct 9900

taaatttttg tctagcttaa aacatagtgc tagatttcaa gaagtaaatt gtcaaacagg 9960

tttgctgttt agtttttgta aaaaaggaaa cttgtaggca gaaacaaaaa ccaaaataaa 10020

taaaatctaa attatggtaa taactaaaga tagttgaggc attaacagag tttgactgaa 10080

ctctcaagtc tcaagcaaat ttctaaacta aactgttttc ttctgttaat cagtttaaaa 10140

gggaatacct tgtgctaatt tttcttggtg ttggtaacat taataaagct ctgtgaaatg 10200

aggccagaag gaacagaaat tgagtcaaaa taatgataac acctcttaca tattcttacc 10260

atcttgtctc ctagccacac acaccatgtt taaacctaaa aggagagaca cctggagggg 10320

ttaattagct tgcagctgtc tgtgactagt ctgcagagga taaaacggat atcagccgta 10380

agaaagcagc tagcactagc ctggatgatg aacaaagaaa agaagcagct gctgaggaaa 10440

agcacatttt gctgaccttt gaggagggct taagaggccc tttgccccct gaaaatggga 10500

cctgggaaga aagtcttgat ggcagacagt acaaaatcaa gggacagcag ctggctgttt 10560

tgtgtaaggg gcaaaattca agaggatttg tgaatcattg tgatttacag tgaccattaa 10620

ctattgtaat aaggagagaa aaatcaacat ccattcagtg cccacattgc aaattgccaa 10680

gcactggagg ggtactttca cctaggccaa caaacttaat aagcagagaa acccttcagt 10740

gttttggctc ctacagctca agtgtactag gttggcaagt ggaagttcct ttggtcatct 10800

ctgtttatac acttgtttga gatatttcat aaagtcagcc catttgttac tttatctgta 10860

gtatgtgctg tttcatgtgg aagatgcttg ccttttgata gtattccaaa ctgaccaacc 10920

ctgtacggat ttaagccact gagaagctat gaacagacag gtgaagtttt ggtattgcca 10980

ggaaccagct tgtccagaag tttaatcctt taaggaaaag gatagcaagc ccaatgcaac 11040

agcccttcat gctaaaaact ctcaataaat taagtattga tgggacgtat ctcaaaataa 11100

taagagctat ttgtgacaaa accacagcca atatcatact gaatgggcaa aaactggaag 11160

cattcgcttt gaaaactggc acaagacagg gatggcctct ctcaccactc ctattcaaca 11220

tagtgttgga agttctggcc agggcagtca ggcaggagaa agaaataaag ggtattcagt 11280

taggaagaga ggaagtcaaa ttgtccctgt ttgcagatga catgattgta tatctagaaa 11340

accccatcgt ctcagcccaa aatctcctta agctgatagg caacttcagc aaagtctcag 11400

gatacaaaat caatgtgcaa aaatcacaag cattcctata caccaataac agacagagag 11460

ccaaatcatg agtgaactcc cattcacaat tgcttcaaag agagtaaaat acctaggaat 11520

ctaacttaca agggatgtga aggagctctt caaggagaac tacaaaccac tgctagacaa 11580 ttgttaaaaa ctaaagttgt tggggaacaa attgtactta aaatatattt attgtgtaat 11640

ctgataaaat attttaaatt ctggtgttac aaagatttag tgtttctttt gttcttttta 11700

caaaacagaa acccactaaa atgttaactg attgacatta tatttttttgc tttcgtcagt 11760

gttaaatact tgttgcacag tagtatgagt tgtttctatc atttactctt ctacaagcca 11820

gttgagacgt gttacgagga ctcaaaatat tctaattggc ggtgaccaaa tgtcatacaa 11880

ttatgaagta tacaagccat tctaaccttc acagtaaatt tttttaccct ctggttgcat 11940

acttaaatag aattatacta ataatgcata ttttctgagt gttaatgtca tgtgaaattg 12000

ctactaagtt aaacgttgct tcccatggtt ttctgggtga tgggtgttta ttctgcttat 12060

tcagtgcctc tctacataga ggaacctatc tcagtaactc ctatactact ctacaacaaa 12120

atggcagaag aaaagttacc catttgaaaa aataccttat tttccaaaca gaagtcttca 12180

gacaggcagt gcacaatatt tgccttccag aaaaggagat tgcataaatg cataacttac 12240

atgttagtac atgtgactgt ttccctgtct tatcctgaat gcatggaaag tcatgagact 12300

tattttttttc ttcctactgg attccattat aaaacctcta tttgtatagg ttataaattc 12360

taaccaataa ttgagcttac tatgttgagg ctgtaggcag aggataagat ttacatgtca 12420

gactttatgt tctgctcttt ggattaggct aattttcaaa agttattgac tgtaaaataa 12480

taatcatgca ctgaggaaat acaatattaa aataaatttt gaagaagatg gaattatagc 12540

aggggacctg tggaaataga atctaagaac atattgtaca tatgctgaat taatagagtt 12600

gcattgtgtc gttcacatag caaattcagt acagattgtt tgttaatgtc aacagattat 12660

gccgtaaaat ttgtaatttg tttggcaaat cttgacaatg ctgagacatt tattctagtt 12720

aatacaaatg gctaaggtct gtaattatga gttatttaag ctaagaaatt tggcaagttc 12780

attattcatc agtctctacc tactattcca aattaattga ctgttttata tactgataca 12840

attaatagta aactttttat attaagtata catatgtgaa ataaaatcat attattattg 12900

gattataaga tgttaataat ggtaagtcat gactgttatt ttgttgaagt gattttttag 12960

aatttttaaa ggtttattca acaaaataga gttatcttac ttaaaaatgg atgcttaggg 13020

ccaggcgcag tggctcacac ctgtaatccc agcactttgg gaggccaagg cgggcagtat 13080

cacctgaggt cacaagtttg agaccagcct ggccaacatg gtgaaaccct gtctctacta 13140

aaaatacaaa attagccggg catggtggtg ggcacctgta attccagcta cttgggaggc 13200

tgaggtagga gaattgtttg aacccgggag gcggaggttg cagtgagccg agatcatgcc 13260

actaaatttc aacctgggcg acagagtgag actctgtctc aaaaaaataa aaaataaaaa 13320

aacaacgctt agagattttt aaattaaatt agtttcaata cttctataat agttaaaaca 13380

agtggtttct atttaaatca caaatgatag aggttagaga tgaaactaaa catatgagat 13440

catcaagtcc aaccttctgc tacgatatta gacttcaaat ggaaagcaca tcaaacgtta 13500

accagatgat ttctttgata gtctgtttaa aaatattgaa ttaataaact ggaaacattt 13560

taaaaattgc tgttcttttg gaattaatat aaaaacatat actttaaagt aaggaatttt 13620 atactgatta agttatacac tgcttagcct agcatgtgta tgttatactt aaaaatgtgc 13680

attgtacatt gtcaaataat ttgcacacga tcatttacac atctgtattg ttcacattga 13740

cgtactttgt gtcatatata tttgtgcatt atcagttgat ttgcacattg cacaccagag 13800

tcacacagtt atagatctgt gtgaattccc tatattccac atgcatggac atgttcatat 13860

atacccacac aaatacaccc catgaaatga tatccagtta atgaatacaa aggtttttaa 13920

aggagaaatt cattagaaaa aaatcctata agcctctagt tttcccaatg aattaatcta 13980

agaatggata ttttattttt tccatgtttg acttgtaaag taaaacttta gtaataaaga 14040

gtgtctgcct tacagttcat atatttcttg aatttaaaaa actcaattac ataagaaata 14100

cattaattct tttctttaaa ctattaagac agcaaagtct tttttatgat tatctacaat 14160

tccggtcctc tctccggtat ctccaagata acatgattat gagtttaatg tatgacattt 14220

aaaactttga aactcacttt catatngntn tnnncnnnnn nntacataga atatatggag 14280

tgttgtttac atattattta tatgaataaa tatacaccta ctccaggtaa tttagtggtg 14340

actatcacgt attttatttg gtcattctac aattgattga tgccttttgt gtgtgtgtga 14400

ttactaatac tggagtatta atatacttgc atgagtctcc ttgtgtacca tgggaatagt 14460

tctgtatttg aaaatagaat tactgggtat gcacattttt aatttttaat atatacttcc 14520

aaattgttcc tgcatactcc caccagtact aagaatgtat ctgttttcct cttagcatgt 14580

caatatttga tataatgaaa cttgtaaatt ttagccaact aaatgagtgt gatttctatt 14640

cctttatagt aactgatgag gttataatca agttaagcat ttttttatat gtttattggc 14700

cactcaggtt tttttcttct acgaattgct ggttcacatg cttcgcccat tttccttctg 14760

gacttgtgtg tgtgtctttt tcttacatat attggtatat ttgacttttg tctattattt 14820

attttaaggt tataggattt tagagttaat gaaaggataa aatgctggag gcctagaatt 14880

cagaatcgta cataggattc taaacagaag ccatgttcta aataaggatt tggcttgtgg 14940

gacaaaagaa aaatgtggaa tttcacagnn nnnnnnnnnn acatttatgg gcaatagaga 15000

atcaagaatg actttgtcat aacatcggca tcattgaatg gtattgtcac ctataccatg 15060

atcaagaaat taacggtgtt gctctgcaga ctggcaaatc tcattgccct ttcttctaac 15120

cctgggaaag tattattaca atggaaaaag atatgcgaca agatattgta ttaatagggt 15180

gttagaatca aataaaagga gggatagatg tagtaagaga agcagatttt atagctatcc 15240

ccattccaag aataaaaaac aaaaaaataa taggcgagtt cattgtgggc atctgtcagg 15300

aacaaggaag gatagagctc tgaaagctga ctgcagtcac taacaaaaac ccttttgaga 15360

catattccaa gaacctgaaa aactggtaat ggacccttcc ttatactcgg tgtcttgaaa 15420

ggcagtaaac tgtctaatgg gatataaaaa tagagtcttt ttttcctccc ccactctaat 15480

attctcctat tttaaaaacc ttctttatat ggcactgcaa ataatagcac acagcaagaa 15540

ctgtactttg cagtattatt gagcatctca agtgtctgaa aggtacaaag ctaacaacca 15600

aatgacttca gaatggaaaa tgtcccagaa tttttaattt tccttctaaa acagaaatat 15660 aggggaagca aaaacaacct taatctaaaa tccagcttaa aatgacaacc agaggttttc 15720

tagcttatat tttatttcag aatgcttagc aaatttcttg actttgtatt aattttagtg 15780

agttgtaatt tgaggaaact catggttttc taagatcctt tatacctgaa aagttgtcgc 15840

atacgtgtaa gacttgcttt gttattgttg ttgtaaattg aatacctgaa ggtagatttt 15900

agtacactaa ataatcatga tgttcattct gtgctaatag gcaggaaaga tggacaatag 15960

aacttactgg gatgccagga agtgctctga tcacccagga gtgagacatg agatgggcag 16020

aaagaacaag attagttaac tatcatggga aacataagaa gaatttggct ttagggattg 16080

tgtgtttgag attgattttt aaaagaaaaa gacctattta aaaatccaca gatgcttagg 16140

gaatagaata gctatttttt taattcataa attttatatt ggttagtttt cacagcagtg 16200

aattatcaac tttatataaa agtaaaatga taatttgata tatttcctga tatgagtcca 16260

cattttcaga gccatctaaa ttcagccttg tttgttcatt atacatatat aagaaaatcc 16320

ttcttctatg caacaaatgg ctgtatataa aatatacatg acttgttgat ttaaaatctt 16380

ttaccttttt ttctcatcat aaaggtaaaa cacatgcatt tcataaatgt ttgaaaatgg 16440

gaacgtatgt accaaaaaaa tcacttcaaa atcatggttt acgtatcttc taaaaacatg 16500

taatcaaata taacaaagta tatttaggcc aatattttga ctatgataaa atcgctttga 16560

ttcagaaggt attgtgtttt ctgagtattt cattatttaa agcagaaggt ttcatctttc 16620

agaagtgtca ctccagcctg ggtgtcctgt gtggttgagt caaatcagtg cttccactat 16680

cagtaggttt ggcacaggat gaagcaagaa catcaaacta aaggggccaa gaacaatagg 16740

tgggagaggc tgcaggagga gcagagaggt cctgataatg aggttccctg actatagcag 16800

gttaggacta gacaagccct gtgaccaaca gaggaaagag tttctgagca ctttctatat 16860

ctggctgtac cttgggggac ttcctccatt ttgtctgggc tcatccacag tcctgctaag 16920

agattctgca aaatatactg cagttctcat tcatgtgccc cttgagatgc atttgagctg 16980

caggtatttt ccctccagtg gtacccagtt tattatttct agtccattca actcataaca 17040

acatccagca gcactttacc cacaggtggg tgtttattag gtttttattt cttttgtcca 17100

ggtttattga gtttgaagag taggcaggga tttctgtttt tatgagtcac ttgtcagctg 17160

taatgaagtc ttcttcacta aaaaaaagat gtgaggacaa tcttgcgcag tgtcataatt 17220

attctaaaaa tacatcatgc ttaattgtat ccacaaaggt tttaaaaatt taaacatcag 17280

ttttgagttt tagggctttc ttattggtga attatcaata attgtagtgc tttcttgaaa 17340

agttccctgg ggtggatttc cttagttgca tataattcta gcctcttcag cctactaagg 17400

ataaaaggga catgaagtac acatatctgg gaattaatta catttttctg ctcaagagcc 17460

aagtcatttt cagaaattga tgtcaaatag gtcttccata gctgaaagtg ctaggtgtct 17520

cttcacatag ttaagatttc cccctattca gaaatattag tttctgatga caatggaaaa 17580

ttccaaatgc agcagcttca ctgtgttcta ataatgttca tgcttaggga tcctgaaacc 17640

acttgcagta tgcctggtct agccacagca gggagccggc gcccatgcca gttcccagag 17700

ctgaccaccc cnaccgcagc cagcgtgtct ggctgtgcac catggccaga ctccatgctc 17760
actcactcat gcaccccgca ctgctcccca cctggctcaa ccttggtggg catgggatcc 17820
aggctggtag tgaaagtcaa atgctgcctg ccaggccggg tgggcagaac gagcccagcg 17880
ggtgagaaaa acttgggcaa acgtgccact ggctacagag gtttccggct ggtgaagtga 17940
caccctaagg attctgtgac aatgggaaaa ggggtaatga agggatagag cagattaagg 18000
aatttaaaac ccagttgagc aagtatatct gtaaataatg tacttttgta tactacaaaa 18060
ccacaaagaa aggatctccc tcctctatta tcacagggtc atcccggaag aactgacatt 18120
tcaatacgtt aaaaaataag tagataatgt ttttagcttt aaaaaaaagt ggaataatct 18180
atgcatatat aaagaataaa tgcacataac atttatgtgg gttatgtaag cagaggtaag 18240
cattataata aagcaaacac tgatgagcca ctggtgagga aacataacct tcctaaatct 18300
ctgaggcttt cagagtggct cttggcccat gtagaggtaa gggtgtgaaa gaacaggaag 18360
agagggcgtc catgtttgag aaaccacaaa atcacaggag tcgctaaagg agtagtttaa 18420
ttggttatat ggaggaaatc tggagtgatg ttagaagatg caattgacca gatcaggtgg 18480
atcacagagg ccatcctgag cagtacaata ggaaattaat cagggatccc ctatggttac 18540
cataccttat gggtttaaac caatgatgct tacgagggag aagttagctt ttctggacaa 18600
tgtccttggt gtgcactttt cagatggaga ttggcaagcc aatggctaag gagggtgaat 18660
cagatcaaaa ttcttccttg actttcatag catttccaga ggatgtctag tttatgagta 18720
aaatgtcttg ccttcatctc ataaaattat ctctatgtcc tcacctgtaa agatgttcta 18780
tatcanaaca ttattattta tacaaataat tatattttct taaaaagact tttttgtaca 18840
gtaaaataag tggaagcaaa cttaaatgat tcagaacaac tttctaaata tttgactaaa 18900
caatttcagg aaatagccat ttaatccctc gaaaacaaaa tttggcttcg gttaattttt 18960
agaaaatcct aagatgagga gatgaagcta aaaaaaatat attgaacact gtgcatttca 19020
ataaagaatt tacagaatag ctggaaaatt aaggttaagt atacaaggta attatatatt 19080
tatttttaaa atgaaaaggc atttaacaat ctaaaaacta catcaaatac attatattta 19140
gcatcatgaa actacttaat tatctgcacc tatcagtgaa gcattttttt tctcataaat 19200
ctgttctatg ataatctaca ggaagacata attttaaaat ttgtaaggat tagagtggat 19260
ggtcacctgg tcccaacctt tggaacattt gaaaagcaag gaacatataa actgaacatt 19320
aattttctcc ttgcaatgta tacaagaaat ttaaaaaaga attttccaga gtacaggtat 19380
aaaagggtat attcttgaca gggaatcaga gtgctgacag atggctaaga tttccatcca 19440
agaaaaattt gatgtttgaa cttaagaaaa tggaaanttt ctttgctgga cctcanataa 19500
tattttcagt aatgacagaa agtcaggaaa gatgaatgtg gaaacaaaat gttcttgaac 19560
acatgaaata aataggtaat ttagatacag tgggtggtgc cagtatgctc tgtaggtgaa 19620
gaataaatct ggaatcctag gaggtcgccc taaccagtga ggtctctctc acttatgaac 19680
ataaaatgtc tgggtactta ggactcctct aggcaataaa tgaccttcag ttagattgag 19740

ggctgatcga actgagaggg attagtacat aatttaatct atttgctaat tagcatgatg 19800

aattagacca ttgtcatctg ttgctgtgtg ggagaggaag aggacaaggc atttaagttt 19860

tgatgagaag gacatttatt gaatagcatt gtacacttta agagtggatt ctcatatatg 19920

atctaagtat cttagtctat agactaagtt tgattctttc ctttcattag tgacagagct 19980

acaactgtca cttgagtatt gggtgacttg acttctcttc caagtcgaac acactgcata 20040

acatggaaag aaggcaagga agaggagaat atatcttgtc ccatcactgg tcatttcagg 20100

ggagattttg gcactgaaga cagatcagga gcatatcaga taggaacaaa aaaaacaggc 20160

ttggtagctt aaaaggagga gtcaagtctt accaactctg aacactcacc tcagtagtga 20220

tgtttcattt ttaccctctt agggtaaaag agtctggtat ttctaagtat atgctttcta 20280

gtgtttgtgt ttactatatt taagaaacac cattagatca attcttatta atgattactt 20340

gatattaatc attatgatta gacactagtt taaaagcata gattagtttc attaaagact 20400

tttcttgaga gattatatat taaatgggtg gtgtacacat tcttactaag aaagttttct 20460

gctgtttata tcttaaatta atgtagaatt tctcccttaa tttctacata gaaacttagt 20520

atagataagc ttgtgcagtc agctgtgaat tcatctaggc ctactaccat ttcttcatct 20580

cctctaaaga ataaaatatg aactatcttg ttgaaattaa gtgctactag atctatactg 20640

ctagtgatcc tgtaatatca atacatgcat gcatacatac ataaatattt taagaaggaa 20700

aagaaatctg gttggaactg atgccttccc agtgagcctt tgctggctct tgacagtttt 20760

cgctctcttc tttaagggtg ttcagattat ctagactttc tagaattaat gattatttta 20820

tcgttttgac aatttttctt tctaagttaa aaccacattg gcattttgtg aatatgatca 20880

tatttctcat tttgaccaat ttttccctag cattacagtt cgttaattta ctggtaatta 20940

aataatgatt aaacatatta attaaacaaa taatgattat ttatttttgt attacgtttt 21000

attccttaaa gcctctctcc atcgttaacc cttttccttt gtnnnnnnnn nnnnnnnnnn 21060

nnnnaagccc ttccttagca tctcactgga ggtacgttca ggcagcatca ctttagtatc 21120

tcgccagcat tgtcctcttc ccgccagtgg gaaccaacta acaagtacac tctcaggtat 21180

gaacttaagt acaggcaagt caacagattt catatgcatc caaatgacac gttaaataca 21240

aagctttgct cacctgcctt tgtacaggag tctgctgctg ctgggtttta atgaaactta 21300

aaggctgaat caaagcctta ttagtgggag gatctttgac atgaaattct ttttgttcag 21360

tataatttgc attttgtatt catttgaaaa attgtcctgt gcatttgttc gaatatcgag 21420

aatagagaac tgagtgattc ctgggtttat gtgggtgata ttgtggaaat aagaggtgct 21480

gtggtatgta cagctagtcc ccaaatatga ttgtatttta ctagagatag agaattcaca 21540

gtgactngaa ctattcaggc tgtaaaatat aggaaccatc tgccatagaa taagcatatt 21600

aggaaaatgc cacctattcc tacattagtc tggcccttgc tccaaaattt cacatgtact 21660

catatgtttg ttttccgtgt ctctgattga tgcagctagc taatgcctaa tgacagtttt 21720

tttaatcatt ccctatatac attttttat tttattatta tactttaagt tttagggtac 21780 atgtgcacaa tgtgcaggtt tgttacatat gtatacatgt gccatgttag tgtgctgcac 21840

ccattaactc gtcatttagc attaggccta tctccaaatg ctatccctcc cccctccccc 21900

caccccacaa cagtccccag tgtgtgatgt tcgccttcct gtgtccatgt gttctcattg 21960

ttcaattccc acctatgagt gagaacatgc ggtgtttgat tttttgtcct tgcgatagtt 22020

tgctgagaat gatggtttcc agcttcatcc atgtccctac aaaggacatg aactcttcat 22080

tttttatggc tgcatagtat tccatggtgt atatgtgcca cattttccta atccagtctc 22140

tcgttgttgg acatttgggt tggttccaag tctttgctat tgtgaatagt gccacaataa 22200

acatacatgt gcatgtgtct ttatagcagc atgatttata atcctttggg tatataccca 22260

gtaatgggat ggctgggtca aatggtattt ctagttctag atccctgagg aatcgccaca 22320

ccaacttcca caagggttga actagtttat agtcctacca acagtgtaaa agtgttcctg 22380

tttctccaca tcctctccag cacctgttct ttcctgacat tttaatgatt ggtgtgagat 22440

ggtatctcat tgtggttttg atttgcagtt ctctgatggc cagtgatgat gagcattttt 22500

tcatgtgttt tttggctgca taaatgtctt cttttgagaa gtgtctgttc atatccttca 22560

cccacttttt gatggggttg tttgtttttt tcttgtaaat ttgtttgagt tcattgtaga 22620

ttctggatat tagccgtttg tcagatgaaa aattttctcc cattttatag gttccctatt 22680

cactctgatg acggtttctt ttgctgtgca gaagctcttt agtttaatta aatcccattt 22740

gtcaattttg gcttttgttg ccattgcttt tggtgtttta gacatgaagt ctttgcccat 22800

gcctatgtcc tgaatggtat tgcctaggtt ttcctctagg gtttttatgg ttttaggtct 22860

aacacgtaag tctttaatcc atcttgaatt aatttttgta taaggtgtaa ggaagggatc 22920

cagtttcagc tttctacata tggctagcca gttttcccag caccatttat taaataggga 22980

atcctttccc catttcttgt ttttgtcagg tttgtcaaag aacagatggt tgtagatatg 23040

tggcattatt tctgagggct ctgttctgtt ccattgatct atatctctgt tttggtacca 23100

gtaccatgct gttttggtta ctgtagcctt gtagtatagt ttgaagtcag gtagcttgat 23160

gcctgcagct ttgctctttt ggcttaggat tgacttggtg atgcgggctc tttttttggtt 23220

ccatatgaac tttaaagtag tttttttccag ttctgtgaag aaagtcattg gtagcttgac 23280

cacatagttg gaagtaaagc actcctcagc aaatgtaaaa gaacagaaat tataacaaac 23340

tgtctctcag accacagtgc aatcaaacta gaactcagga ttaagaaact cactcaaaac 23400

cgctcaacta catggaaact gaacaacctg ctcctgaatg aatactgggt acataatgaa 23460

atgaaggcag aaataaagat gttctttgaa accaacgaga acaaagacac aacataccag 23520

aatctctggg acacattcaa agcagtgtgt agagggaaat ttatagcact aaatgcccac 23580

aagagaaagc aggaaagatc caaaattgac accctaacat cacaattaaa agaactagaa 23640

aagcaagagc aaacacattc aaaagctagc agaaggcaag aaatagctaa aatcagagca 23700

gaactgaagg aaatagagac acaaaaaacc tttcaaaaaa ttaatgaatc caggagctgg 23760

tttttgaaa agatcaacaa aattgataga cctctaacaa gactaataaa gaagaaaaga 23820

```
gagaagaatc aaatagacgc aataaaaaat gataaagggg atatcaccac cgatcccaca  23880
gaaatgcaaa ctaccatcag agaatattac aaacacctct acgcaaataa actagaaaat  23940
ctagaagaaa tggataaatt cctcgacacg tacaccctcc caagactaaa ccaggaagaa  24000
gttgaatctc tgaatagacc aataacaggc tctgaaattg tggcaataat caatagctta  24060
ccaaccaaaa aaagtccggg accagatgga ttcacagccg aattctacca gaggtacaag  24120
gaggaactgg taccattcct tctgaaacta ttccaatcaa tagaaaaaga gggaatcctc  24180
cctaactcat tttatgaggc cagcatcatc ctgataccaa agccgggcag agacacaacc  24240
aaaaaagaga attttagacc aatatctttg atgaacattg atgcaaaaat cctcaataaa  24300
atactggcaa accgaatcca gcagcacatc aaaaagctta tccaccatga tcaggtgggc  24360
ttcatccctg ggatgcaagg ctggttcaac atacgcaaat caataaatgt aatccagcat  24420
ataaacagaa ccaaagacaa aaaccacatg attatctcaa tacatgcaga aaaggccttt  24480
gacaaaattc aacagccctt catgctaaaa actctcaata aattaggtat tgatgggacg  24540
tatctcaaaa taataagagc tatctatgac aaacccacag ccaatatcat actgaatggg  24600
caaaaantgg aagcattccc tttgaaaact ggcacaagac agggatgccc tctctcacca  24660
ctcctattca acataatgtt ggaagtcctg gccagggcaa ttaggcagga gaaggaaata  24720
aagggtattc aattaggaaa agaggaagtc aaattgtccc tgtttgcaga tgacatgatt  24780
atatatctag aaaaccccat tgtctcagcc caaaatctcc ttaagctgat aagcaacttc  24840
agcaaagtct caggatacaa aatcaatgtg caaaaatcac aagcattctt atacaccaat  24900
aacagacaaa cagagagcca aatcatgagt gaactcccat tcacanttgc ttcaaagaga  24960
ataaaatacc tagnaatcca acttacgagg gatgtgaagg acctcttcaa ggagaactac  25020
aaaccactgc tcaatgaaat aaaagaggat acaaacaaat ggaagaacat tncatgctca  25080
tgggtaggaa gaatcaatat cgtgaaaatg gccacactgc ccanggtaat ttacanattc  25140
antgccatcc ccatcaagag atttttttac acattccata tgtgaatatc ctattttctg  25200
attcagttct gactttattt ccattatatg tgaatgtgtt tctttactat ttttctatca  25260
gtttccttcc tggtcaattg actatatggg aaaatatttt atccattcaa tattcaaaaa  25320
tagtttattg aggacccact atgtgctagg catttgatca caacagtgaa aaacagagaa  25380
aaaatccttg ctttcatgga gtttacaaat tctaccaaag gaaacaggaa ataatgaaaa  25440
taagtaagta aaataaatag catgctacat agtaataagt gtcaaggaga aaaggaaagc  25500
agacaaagag aggggaaggt gtgagcagga attgcagttt cagataaggt ggccatgaag  25560
gtgtcccttc aaggaaaatg tagtagttgg gaaggggcaa gtcacatggc tctctgggga  25620
agagtgtgca gggcatcggg agcagccctt gcaaaggcct tgagaaggaa gcatgcccca  25680
tgctctcctc ctcatatgag gaggccaaca ggtcttgaat gaagctatgg agaaagagaa  25740
agggaatacg attggaaagg taacaaggaa tggatcacgt agagccttcc agggcaacgt  25800
aatgaccttg gctattcctt ttaaatgaga tggaaagccc gtggagcatt ttgagtagag  25860
```

gcgagacaaa actgatcagg acaactctgc ttgctgggca gagaatagaa tcaagggaag 25920

caatgacaga agagactgtc ctaagagtcc aggcgagaga tggcagtggc ttggacaaag 25980

gtggtagcag tggaggtcac agaagatgct gaatatagtg ttcatgttta ctttatggga 26040

tttactaaga gattggctgt gtgttaaaag gaaaagatca ttctaaagct tctggcctca 26100

gtgattagat agacttgcca ttaacttcca gaagagcatg tttgagggga aatatgagga 26160

tctcggtttc attacatgtt aagttccagt aatgtgatta gattgaaaat gctatagctg 26220

gtagaatttt ttcacttttt atgtcaaaac agaaaacatg attgaataat gaaaactaca 26280

catgctcaat ttttttaaaa ttgttttctt tgacacagta aaggatcaat gagacttgct 26340

taattcttgt caagaataca aagtcaacaa tggctgagat gaaaatattt tccataaagt 26400

aaaatgatca aatttgcaaa agcttggtta catgatgtga atttcactca aaccacaagg 26460

cattacacta gagtcttgct gaaatttaag acagactgtc ttaggaatta tttgccaata 26520

gacaaaatga atatggggtg gttgatgtgc tgaagtggcc aatgagcatg ctttgattga 26580

cagatgagga aatgtatccc tacattttat caaactatgt ggtgtctgat ttatatcact 26640

gaggtggcca ttcattatgg tcatgaactg tgaggtcaaa gcacttgcaa aaacttgcca 26700

aaatacttcc tattccagta ctaaatattt ctcatatttg tttgcctaga ataacaagat 26760

gcaggcttgg tagtcctttc aagaatgtga atttctttgg tatgtctgag tttagtaatg 26820

ttgaattgtg tgtgtgtgtg tattattctt aagactttgc ctgtatttta acattttatt 26880

tggtaaaaat agcattattt tggcaagatt tttaaagcct tgaaatggga caaattaaag 26940

tataatgctt tttcttgaga aataatttat tgctccatgt tgttgaatca taccattagt 27000

catgaaacca aaaccctata tttaaaaagt aacaatctca tattcaaagt atctcctttg 27060

aacatttcaa agaatcttaa gcaaatcttt atgtgacatt aaattgtaca ttttaccttg 27120

ttattttaga tctatgggtg gagttcattc ttcgggctaa gtgtaggtat gaagtgtgaa 27180

tttctgtgtc ttggtatgcc agtaacttag ggcaaggtta gtagagtggc ttttgtgaga 27240

aaaaaagagc tgttttatca gacatcagtt ggaactttgc cttgttttgt ttccctgaac 27300

cgtccaatct gcttcatata cttgtgtgaa atcctaaaat gtctgtatgt caataatttt 27360

tagttagcat attttttcca aacagacatc cttgtgaaca tcagccgttt ctgagaactg 27420

ggcccacgga ccatagctac tttgagtctc agtctatgta atagttgggg taaagtaagc 27480

tttaaagaga gaactcacaa cacaatggct caagcaagac acaagttgat ttctttctga 27540

agctgccatc ctgggtatgt gggtatttgg atgagctggg ccccatgaag tcatcccaaa 27600

gtgcaggcct ctgtggcatc tctgccaaac tcaacactta ctctccaagt ctgtccttgg 27660

cttctctatt tctagctggt ggaagggaag ggaaaaatga gtccagggaa ggagtttttg 27720

tagatggctt gggagttgca tacatcagga agaactttgt cacaaggcca caactccttg 27780

caggaaattc taggaaacat attccctcac tgagcagcca tgcgcccagg aagaagagaa 27840

gaatgagtct agagtcgcaa gcagcaatat ctactacagg cttaaataac caaaggacaa 27900 aatgcccagt ccattataaa tatattccaa aaatgaagcc atgaaatttg acttcattta 27960

ttgaagtcaa aacaaattta ttgttatgaa ctaatttctc tattttatat ttctcaaatg 28020

attacatttt taacacagaa caaagccata gttagaattt cagtaatcgt atggaatatc 28080

caaaaggtac acccaataca cccaaaagct aaaccagagt cactgctaaa ttttatattg 28140

ttctgattta cctaacagat tagtgaaaat aattaaaatc ttgctgcttt ttaaatttta 28200

aatatcggat tttggttaga actctaagaa tatttaaaac tatctttacc tactcttgtt 28260

ggacaatcat ttgaggagct aagagtgcag agaaggcaga gaaatgttct cctcaaccca 28320

agcaaaagca cagaagaact attgctaatc caaaatgcag gaaggaaatt tcacccctcc 28380

tcccttccct tttcttgggc tagtctgttt aactttaaaa gcagaatttt tgctctcaca 28440

aagtatacca gacaggaatt tgttcaaatt cctcaaaacc aagccaagaa aagtttcaga 28500

aagttcagac atgcatggca gagtggaaag aacatagtct gtggaggcga gcacggattc 28560

attggaatgc cgctccagac atccactctg ggtgaactca gggatgtgtt tagtctcatt 28620

aagccacagt cttttcattt ttaatgtaag agtgataaga acgacttgcc agtgtgttgg 28680

tggagaactg aaatagagaa aggcctggtg tgtaacagga tgtttctctc catctgtgta 28740

atctgtgaca gtcctaggaa gttcaaagtt atgactgttc tgggtttgtg gccaacagtt 28800

gtctgtggct tccatgggtt ggagctgtct aatggggccc agnccaccat cgctctgctc 28860

agaagctaat cataccaaaa catgaggttt ttctccaagc gacaaagagg ggctaagcac 28920

ccaaggatct ccataatagt agcaggaaaa tgaagctgag accttaccct gacttagaag 28980

ttcataacaa taatggcccc aaacaagatg ctcatgtcta ttggttgtgg ttaataatca 29040

tgcaacaaca tgaacttcaa acccagccta tgcttggctc ttccagggac tgaaggcatc 29100

tcagagcccc accccttacc tcagaaaccg caggaaggag cctgggttct cagagtccaa 29160

gaaaatgcac ctgataatag cacctgaact tcctccacac aaagaaagcc tgatagggac 29220

aggggattaa gcttgcatgg agggcaggat ccaacccaca gtagcacttc cttgccatgc 29280

gaccttagca aatggcttaa ctcgctttct gtatcagtaa aatggtgttt gtagtactaa 29340

ggtgtcagtg aggattaaat gagccaagcc tgacccacag caattcatga caagtctcag 29400

ccactttaat gaatgtgatg atattctaac gtttgggatg ctgaggtgga ggacgatggt 29460

gtcccacctc ctgcaaatgc catgtgaaca ctgaagctac acaagcaatg gctctaaagt 29520

caaagccagc atcaggactg gtataaaaga atcaccagga caagggcagg gcagagggga 29580

ggccaaagca gctttgtcag tgcctgcacg gctggaccta gaggaatttc tcgacataga 29640

gggtaaaaca acccattggc ccacatttct gaggaccgac gtgggagaag ggtcacatga 29700

ggccatgaaa tgggcatctg cagaggaaat ggaagccagt gtctggagac aagtggagcc 29760

attgagctag caatgcgtga tagatactgt gggaagggat gtggggaatg cagcacctct 29820

caagggaggt ggctgagtct ggcaatgagg gcctggggca tccttagtca gcaaagcagg 29880

tctgcagcct cgggagctct ggcagctggt gtgggtcacc aactggcctc gggcaccatt 29940 catggaaatt aatctggata gtcccctcct ggagaccctc caggatggca gatgagcaga 30000

ctggccggct ttcaaagtct gctgacctct ataattgtgg ctcttcacat ggtccaacac 30060

actcttccct tttacatgtg tttcctcttt gaatatttcc accagcccct ctccctgacc 30120

tctctctgct ttgttacttt cctttgccca tttaatctac attttatgtt ttgtccagat 30180

taacatacct gtgtttcatc tgccaggcac tggcccaggc cctaggaata tggtggtgac 30240

caacacagag tggctgcttt catggagttt acaatctaca aagggacaaa aaggataaac 30300

aactttaaac tttaaaattt ggacaaactt ttataaatac tgcatttttc caataaaatg 30360

agataaacaa tggtataacc tacattgtac agaatctgtg tctccctaaa attaaagctt 30420

tgagataaaa ggagacaaga ggtattgttg gacttgaatt tgtcgtgcaa ataagattaa 30480

aaccttgaat attagccttt tgattgttgt ctctttttgc agtatgaggt agttgtgaat 30540

tctaggtcaa aacataaggg ttaaaaggct ggaaaaatta aataggccaa atgaaagaat 30600

tgtcttgcta aaatgtctat gtagatgata aaatgtcctt gtgcctcttt aacttcaatg 30660

atttgtgctg gatattttat aagcgtccat taggcaatag caaatttatt caactataag 30720

agaattacga aacacaacta tcataggaga tacttatttc cttccaacca ataagtaacc 30780

accagtgatg agaaggcaga aagtgacaag attaaccaag taagagagtt ggatacaaca 30840

ttcaactcca tttcttcatg tcccaataag agacaacgtt tgacctgccc tctcattcct 30900

aatcttgttg aatgaatcta gtacagaggt tcataaaaac cgtcaggtat ataatgccgg 30960

attttgtcaa caaacaataa caataatttt ttctttaaaa attattccag gccaggccgg 31020

gtgcggtggc tcatgcctgt aatcccagca ctttgggagg ctgaggcggg cggatcacga 31080

ggtcaggaga tcaagaccag tctggccaac atggtgaaac cccatctct actaaaaata 31140

caaaaatcag ctgggcgtgg tggcagggac ctgtagtcgc agctactcgg gaggctgagg 31200

caggagaatc actttaacac gggaggcgga gattgcagng agctgagatc cagccaccgc 31260

accctagcct ggcaacagag caagattccg tctcaaaaaa aaaaaaaaaa gaaaagaaaa 31320

aaattattcc aggccaggtg tggtggctca cacctgtaat cccaacactt tnggtggcca 31380

aggtgggtgg accgcttgag ctcaggactt caaggccagc ctgagcaaca tggcgaaacc 31440

tggtctctac aagaaataca aaaattaacc aggcgtggta gtgagtgcct gtagtcccag 31500

ctcctcatcc gcctgtaatc ccaggcggag attgaagtga gccgagatgg caccactgca 31560

ttccatcctg ggtgacagag ccaaaccctg tctcaaaaaa taacaaaata cacaccaggg 31620

cctgtcaggg ggtgggaggg aaggggaggg agagcatgag gacaaatacc tagtgcatgt 31680

ggggcttaaa acctagatga tgggtcaata ggtgcagcaa accatcatgc cacatgcata 31740

cctatgtaac aaacctgcac attctgcaca tgtatcccag aacttaaatt taaaaataaa 31800

ataaaaataa taacataata aaataacaaa aaagtcttta catctgaaat aagtgtactt 31860

aaaattaaaa ataaaattaa agtacaaaaa cttaaaataa aaatcatttt agatatttcc 31920

tcagatttaa atggaggtca ttgaagctca aacacctggc tgcaaagtga tttgtaagac 31980

```
gaacacctac actctacaaa aagacttatg aataccctgc attggttgat gaacaagaac  32040
agtgagttat accttctgaa tagtgaatta aataaaatac tattaaagtt ttaaataatt  32100
gccaattgct ccatatttct ctgcatgtta gagtttagct agtgtttcca gggatgattt  32160
ttctgtattc ctcccagtca ggaatagcca gtcagtatta ctccatcagc cacagcacaa  32220
atgtctgatt gctctcttga ggtcgttagc ccagatactt gctagaaagg tatgattgct  32280
ggccaaatga gctctggcct ggtgcagccg tcttgccaat gtgcctaatg agagcgcgga  32340
aagacatttt attccaggca taagcaggcc agatcacagc tgcccaaagc tctctggctt  32400
tgattaagtt tgttgtgcta ttatgtaatg ggtacttgaa ccatattccg tgaatgaccg  32460
ttttccttcc tctttcgcca ctattctaag tatgttgtct gtatctaatg ttgccttgca  32520
gttagttgtt aaaatcaagc agaaaagaaa acaggcaaat attaaaacgt aaatagaaaa  32580
agttgttggt ggggaaaatg acttataatg cattattctg acataatgtc ttggggtcca  32640
ccaggagaat cgccccatca caattcagcc tctctctgtg ctaatgacgg ccttttcatc  32700
ctgtgctgtt agtgatgcat aagccaaatt gtgaccaggc agactagacc acatgtaagc  32760
agagaaccac gtctttgtag acattaattt aaaactcgtt ctaatcaaat atgttcattt  32820
atctgtacta tttttgatct atatcttaaa tcaaaacgaa gctttttga cttctggttt  32880
ctgtttctgc ttgtaaagag cttggaagtt atcaatcctg tcttcataac aagaaaaatg  32940
ctgagcacac tgaaagtcag taactggaaa tcaacagtgt tttaaaagtt aattggttac  33000
aattttctca tattcttctc tcacctctgt agagttctgg tctacattgg aaatgctgga  33060
aaagtaggaa tatacagtac agagctttct tggcagcata acttaataca tgaaaaatga  33120
tttttaattt tagtttaacc tttaaactgg aatctgtagc acattggaaa aaaaaaacct  33180
aaaccagttg aacaaatttc aattttaaac atgttttcag gctttggaag cttactgaag  33240
aatgtggata gcaaatacag gaaaaattct gtgaatattt ttattcacat atgtaaacgt  33300
gaagttataa gacagaacaa tttccccatg tacaagcttc aaatttttag cagttagtta  33360
acattttcta ttcagtttta accattaagt tgtaatatta ttacaaatga tattagaaca  33420
catctttgct tttagtgctg tattataaat atttaaatta atgacattca gtgtaaattt  33480
ggtatttatc gtaatacaaa taagaacttt atttttagtt aattccatgt gttatttgta  33540
tttcttatct cacatttatt taaaattaca tttggtccct atcactcagg tatgtttctt  33600
ccctcccacg tagacctaat cggatttttt aaaattggat tttaatctgc ctttataaac  33660
aaacgtcaat tatttatttc caatcttggt tgatcttttt attccgagat tatttactag  33720
ttcattcagc cttcctatta aaaataaata aataaataaa taattgctct tgtagctctt  33780
ctgaggtaga ctgaccaggt ctaccttatt gttaattttt gcattttcat ttttattgaa  33840
tatgtatgat gatgcagata tgaaccatgt gcatatctca gaataggagg aaaaatatct  33900
tagtactttt gccaaatgct tggttctaat ttttggacta ttatgaattt tttatgtcat  33960
aactattatc ttaacattat gcgaacttct agccctttttg cttcaatttc ctttgaaaag  34020
```

atttccttcc aataaatctt taaaatatgg gtgacaaaac taaaagaccc tcagaggaaa 34080
catatttctg gtatgacata ttttgtgtac tctttaggac aattgaaaag cttctaatta 34140
acatttcttt agcgccttag tggtgatcag cagtatcatc actgccctag tgcatccatg 34200
gaagttgagg tccagggtgt actgactcac acaggttacc taggtaagaa gtaatggatg 34260
actggaatgt taggaactgt aaatgacagc taactgtggc aaagcgtagg gttttgtttt 34320
ttttttcttt gtcttgtctt ttttttttga gacagagtct tgctctgtcc ccgaggctgg 34380
agtgcggtgg tgtgatctcg gctcactgca ccctctgcct cccaggttca agcaattctc 34440
ctgcctcagc ctcccaagta gctgggatta caggtacatg ccaccaggcc tggataattt 34500
ttgtattttt agtagagaca aggtttcacc atgttggcca ggttggtctc aagctcctga 34560
cctcatgatc cacctgcctt ggcctcccga agtgctggaa ttacaggcat gagccactgc 34620
gcccggcctt tcttttcttc ttaagacaac taatggaaaa taaacttcca agaagctctt 34680
ttgacctgac ctcactgctc ctcaaaggca gaatgctgat tcacaggctg aggggtatct 34740
gaccaggttt ctttgctacc cttactctca gcctacctgg aaattgccga gcgcaggggg 34800
cgcagggagt gctgtagttg gcgtgtggac tgagcgtcct gggaactgtt ctgctgtagt 34860
acctacactt tcctgccaat gtcagcatcc ggcttgtcaa cacttctttt ccatttgtga 34920
atactatagt gaaagccttc acctctcctt tctttgaagt attctctgtt cagaacttgg 34980
aactttagaa attttctaca tttaaggaaa catatttcaa ctatgattct ttattactta 35040
taaaaaatta acttagaaga tatatgtaaa catgtgtatg gtttaaggta acaaagtgag 35100
tttaccaata tggcaagttg tgccaaaaga gtgcttaagt gtaaaaaaca aacaaagaaa 35160
ccaaaatcaa aattaaagta gctcaaaaac ataaagttta taaatatttt taaggcattt 35220
ataattttaa aaaagtcctt ttatgtttgt cccatgttct ggaatagcaa ataaaaaatg 35280
tggtttttgc tttattttgt tttattgcct tgatatgcca cgtggcaagc ctcaggctgg 35340
gtgtgtggag cacatgactt gagtgtgtgg acaagttgtg ttgaagatta ttattacagg 35400
cacaatgtta cagcacagtc tatcagggaa aaatacttaa gtttattttt ttgaggctga 35460
gcgtggtggc ttatgcctgt aatcccagct ccttgggagg ctgaggcaga aaattggttg 35520
aacccaggag gtgaaggctg cagtgagcca agactgtacc actgcactcc agcctgggtg 35580
acagggcgag actccatctc agaaaaagaa aaaaaggaaa ggaagaaaga aagagagaga 35640
aaaaaaagag aaagaaagga aggaaggagg gaaggaagga aggaaggaag gaaggaagga 35700
aggaaggaag gaaggaagga aggaaagaaa gaaagaagaa aaagaaaaag aaggaaggaa 35760
ggaaaagaag aggaagaaga agaaggaggg aaggaaggaa ggaggatcgg agggagaaaa 35820
catggtttac agaggtttta ttcactagga tggtaacatc ctcccaagta atttagttct 35880
agctgcttta tcgtttagta aatctctgca attcaggatg cacctttctt gctgttgctg 35940
gtctgctgat ccacatgatc ttatttcatg tatgtggatc atccagccac cgtgacaaat 36000
gttcattatg ttgagtccag atgtccacct taagagtaca cacccctcgc tgctgctttt 36060

ttttttaat tatgtcacac taaatacatt tactttatga taactactgt tggggtacaa 36120

tttactctta ccacagaacc agctctcagt taagaagaaa gtgctgagat ggtgtgggga 36180

aatttggtga agaaaaacat tatagtaagt tacagtcaat ctcactactg acctttggct 36240

aaaatttttt ttaatatgtg gctaaaatct gcataacaga tggtatatct gagtgaagag 36300

attattagag gatgagagaa actgagagtt caccaaaaat aaaaactaat aattgcaata 36360

tagattttgc catcaaaagt gacctcaaat gcaagtttgt agctgaattg aagaggtggc 36420

cactttgttc ttaatctggt ggctgctgct taggtttcat tgtagaggct actgagctgt 36480

tttgcaagtt tgtattatct gtctcatgat ccagacagaa ccagagtcta ccaacgtgga 36540

ttaagaatta gtctgttttc ttctgtgttt ttctatgaat atactgaatg tctcatcttg 36600

attcagaaga aaaacctgaa agcttaccca aaaacaaaca aacaaacaaa caaacaaaca 36660

aaaaaaaccc tgtgtgtggt ttgtcttggg acagttcaga aggcagtcat ttaatttcgt 36720

ttgggaacat tttgcactat ttatttcatt caaaatatgt taatttattt tcgattttaa 36780

ttaatgtact tttgtaatgg cttattccgc aagcagagat agcctccctt cccccaaatt 36840

ctaagatatg catgtataca gtcatcttac tacatagcag gaattttaaa aagaaataag 36900

cttttctttc cagtgcctag ttgttaaaac agcaaaatga taaaaattta tgtaatggta 36960

ctattgagct gtctactgtg tggcaaatac tttatacaaa taagcaatgt cttaggatct 37020

aggcttttag ccttcaaaaa tctttaaaga agaatgaaat gtcattttta agtcttgtat 37080

ttgatctcaa gtgaagagat tcaattcatt tcaaaaatac ttaatataaa ttatatgcta 37140

gaccctgtgc taggcactgc agggcaaaag ctgaagaatc agagaggctg tcaagttaga 37200

aaagtgtttc tcaattggca atttttccaa caagggccac tggtgaatgt ctacaagtat 37260

ttttggttgg gggttgcagg gactgttgga atttagtggg tagaggtcag gggtgtggct 37320

acacatcttg caatgcatag gacagacctt cccattaaag catcatccac ctgcagatgt 37380

cactagtacc aaacttacat tttgtgtagc ttcagcaagt tctctcttaa actattcatc 37440

gctctctttt ttcattagtt ctgcagttta ataagaatat tgccctatgg agcagttcta 37500

gcactttggg aggccgaggt gagtggatca cctgaggtca ggagttcgag gccagcctgg 37560

ccaacatgac gaaaccctgt ctctactaaa aatacaaaac ttagccaggc gtggtggcat 37620

gtgcctgtag tcccagctac ttggggaggc tgaggcagga gaatcacttg aatctggagg 37680

cagaggttgc agtgcgttga gatcaagcca ctgcactcca gcctgggtga cagagcaaga 37740

caccgtctcn naaaaaaaaa aaatagctgg agattaggca agttgtatat atgtaaacat 37800

aaaaagacag aaacaaataa aaaccaatat agcacaaccc accccatctt gaacaagttg 37860

tacttctaat catgcacatg tgaaaacatc aagctcttat ttgaccacca ttcaggattg 37920

aatatttagt cattgcaaat gtagttgagc aaataagtat ctatttggag tttaggatta 37980

ggatcataca cttccacagt taaatctact gactgacagt agacagttca ttgagacaca 38040

gatacagcat agtcatccat gtatagtggt tgagtaacgt gtaagtttct tttcccttcc 38100 aactgaaggg gatgctcaaa ctgagattta accaggacaa agtcaatcag cttctctctg 38160

aaaagggtga tgtctgatca gagtcagctg gaggatcaga aagagtcatt tagaaaagtc 38220

tgagggaaac tgatgatctt gagtaacaag aaagttgcca ttttgttcac cacagatgaa 38280

agtctattcc cagacagtaa atgaagccgg ataatcaggc caatttcata aacagtagat 38340

cagattggaa gagttaaagg accagataag aaaacaagca cctgtagcag gatgacccga 38400

gtaacacaaa ggatgattca cgcaaagggg gtgcaaggac taattctgca gactcctata 38460

gcttgttgtg ccttgcaagc ccaagagagg ttcccaaagg gactgaagcc tggcagcatc 38520

cagtgaatgg gctctagcat acctggtaag gagctgtctc agaagatggg gtgttccata 38580

tccctggagg ggttaaagtt caagaacatt gctgaggatc caaacactgg gagagagaaa 38640

gaaatggatg gtcataaaat aaatcctctt tgttctttta aaaattctga gtgtgggagt 38700

gaaggtgatt tttcagacca agcccttggg tactgactca ttcttattcc actgtctcgg 38760

gcttacatct ccaattctaa cttcccaact tctgtcatgg cagatcaggt attaccattg 38820

ttagaatagt actataggta ttttgtgttt ggagagaggt ttagtggtag gaccccaggt 38880

ttgtagtgtg ttgctagtta tcatcacaat tttatgaccc tctcacttga gagtggaatg 38940

acaatgtggt tatgtaacct aagactagaa gaaacaaatg ggagtactcc caaaacacaa 39000

atatctcacc tctttttca tgggaaaaga gtcagcagaa gagggcaaag cggtaaataa 39060

aatcctaaag gtaatcaaaa cagatgggaa tataattttg ttcagagcta taaaaaatat 39120

cctctactat tctcaataca ggcagtgtgt agaaagtttc aaggttagta tttatcttct 39180

aaacatagac ttgtggtctt cttatttaat caattcaaac caccagtgga agaggtcaga 39240

aagaatgaat gaaatcagga gaactgatta tcatgagtat cattgttgtg attgcatttc 39300

ttgcacagct ttaatttgtg ccttgagagt ttggggtctg agaaagtagt gctattccaa 39360

gagctctcca gggatatttt tggaacaggc cttattgttc tctaaaaaat taatgcgtta 39420

tattcccagg tgctgttgta gcatcatttt tcatttaaca aaaatatggt cgtttagcca 39480

gaaattttag ttctataatt aaaatggact aaaaatattc tgtagcacat ctagaaaaaa 39540

tattttaaat gtatcaatgt tcttcctaga aattttccat gttggatcaa acagtagatg 39600

aagacttttc tgtgtgtata gcatgggttt ttatgtttat atatgagtgt gcaggtgttt 39660

taagagaatt ttgaaaatca gcaaagcaaa gtcaattcac tacaaaatag tttccaaatt 39720

atctgtgcag tattaatatt gaggaagaat taatcaaaaa gcaaatattt gagaaatgct 39780

ttcctatgtc tgtaggctac aaaatcattg tttaatttat caacaatctt tgagtgctcc 39840

tcagtgtctg gtacagaaat gaaagataca gttttaactt tagagtgctc acacctactg 39900

agggaaacaa acaaggtaga gaattcagta tgacgagtga catgagacag ttgaatggga 39960

ggctctctgg gagtactaag ggccattgcc tagtgtagct gagggcacag aggagctgag 40020

ctgggcagcc caggagtctg tcaagagggc cagagggaag cagtgcaggg ccagcagcga 40080

gagcgggact gtgcaaaaga agtttcagta gctacagttt ctttctcagg gtcctccact 40140 tcccaaagaa aatgaaagga acatacattt gatttttgtc agtttaaata ttggtcgacc 40200

attggaaaac catttttgaa aagtgatttc ttttgctatg tttggcaaag ccaggagtag 40260

taaagagatg aggaaggacc atcttcgtcc attgcctttg cctttgtaac tgttacttct 40320

tcacagcctt tctacccagt gtagttcacg tatgggaaac atttctttcc atcgtcagcg 40380

tcttcctggg gagtgttttg gttatacttt cacgaatgtc tcacctcttt gtgtcaacaa 40440

acccttctct ttgagctgta ctgtgggttt ctgtttcttt cctttttntc ctttgtccaa 40500

ttagaaaaaa aaatggagta gaagtttcaa ggctactagt tcaatgtgct ctttttttact 40560

tgactccaaa tgacatcttg cttgcagatg tgattaagaa agatctggac agggagaggg 40620

aggataattt tcaatgttcc tctcaaattt aatcttatag tgacaagctc catattctac 40680

atgaaaatgc tcagttaata cagctattaa acttgaatgg tagataatag tggatttctg 40740

tagccactgt ttggtttcat tcaaaaacct agaggaaaat aagggatata cattataaat 40800

tgtttgttta atgttaaaaa tatgtttcat attcagtgca cataaagcaa aggtattgtc 40860

cagtgttagt acctatgaat catcaataaa aacagttgct tctgtttgct gttgacattt 40920

tatgtatgtg ttgttcctct ctgaaccttt ttcttaaaag tcaactttta agaaacttat 40980

aagtttactg tggagagcct atgaacagat gcatgagggg gcgcctgttc atatggataa 41040

gatagggcta taaatgccct tatcttgcca gggctcttct aggcctcttt agggttaagg 41100

catactccct tctgagaatt tgtgttgtaa ccggttgtct agcttcacgt cctgtttcta 41160

tggattgttt gtaaccagct tttgctgcaa ctgttactgc tgattaatac cttgctaatc 41220

ataggttatg gaaagactgt tttctgtttt aaggctctgt tagaaattgc tgatgcacac 41280

aatattgtaa attcttatct ctgtatactg tacttctgca tacagatatt atgttaaaga 41340

attacttcat ccccatgtga ccatctcact tcataatcaa atgactctaa atccctcact 41400

aaactacccc caccctcact aaacttaaca ataaatgctg gtatatccag tacgttggcg 41460

gcatcacagg accagaaggc agtgatcccc ctggacccag ctttcactaa aaaaaaaaaa 41520

aaaaaaaaaa agaaacaata aacttagcca ttataatccc tcagaaaaac attaactata 41580

tttaaggata ataagagaaa ataatctctt tttaacatac ttattgatac ataataattg 41640

tacatattta tggggtacct gtgatatttt gttttatgga taaaatgtgt aattatcaaa 41700

tcaagatatt taggatatcc ataacctcaa gcatttatca tttctatgtg ttgggaacat 41760

ttcaaggcct ctctgctagc tattttgaaa tacacaatac attgttgtta actataatca 41820

ccctcgtctg ccatggaaca ttacaactta ttccttctat ctaactctat gtttgtaccc 41880

attaataatt tctttttaaa gccccagtct caatcaagag aaatataaat agcatggtaa 41940

aaaaaaattt tttttgacct gatgatatgg gttcaatttc tggttctacc actttccacc 42000

ttggagacct tcattaacca gtgaccttca cttttctttt ctgtaagatg gtattgataa 42060

tatctgccct catccccact tcgccatttg cctagctaat aagacagaga tttccttagc 42120

atgttgtgtt aaagaggcta gaaactcaac atgctttaca aataccagct gtatgacaag 42180 gagttacagt accaatgtac aaaatatgat ttttaaaatg aaaggcagga catttatgaa 42240

gttatattct atagaaagta tattagcaca cattaatgca taataggcga taggactgca 42300

ctttgcttct ttctcttagc ttttctttcc agcctctttt tctcccttgt catcaagtct 42360

gcggctggac ttgccttctg ttctgcatgc cttggtgttt gctgttgctc tgattctgct 42420

ttctcatgca caggctcaaa gaaacagtga ggctgcccca acctgtgcaa ctctcctctt 42480

tgtttggata cttggtcata tgcacattct agtttaaatt tctgtgtaac taaaacatgt 42540

atgagtggag attggtttta tattctgtaa taataaaact tatttataga agtagaattt 42600

attgaggtac aaatgtttac cagttttatg aataacaact tatatttatt ttataaacaa 42660

cttcaatggt atatgtgtgt gtgttttctc acaaaagcat attgagcaat tattttcaat 42720

tttatagtca taaatgatac ctgaaatatc ttaagaacct accctgtgat ttacaggaat 42780

taagtcatta aatccttaca acacagtaaa gtctgtgtta tttcccattt tacagatgga 42840

gaaactaact acagaaagac tattaaagag gcaagatatc atcacagttg gatgtgtggg 42900

gctttgtacc agaatgactg agttaattcc agttccctta attattatct ttgcatcctc 42960

agacaagtca ctttatcttt atatgcttta gtttccctgt ctgtaaaaca tcttatagtg 43020

ttgttaggat taaacacatc ttttccctac atggcatttt aggttctctt aaccaatgtt 43080

aaaaaatcta tccacatatt tttagctttc ctttctttcc attactagaa gtactttcta 43140

tgaggtaatt ctcccagaaa atacaaagtc ctatttggtg attcactgcc ttacattctg 43200

gtgttcggta acaccatgga atccaaagac aagctccatc atatcaattg tggagcctga 43260

gatactttac tctccctgtg agaaagtgat ttcaaatgaa actcagttct cccatttatt 43320

tcattttggt taaaactata gcaggtgctt ttcagcttat tgaatgtgtt aggctccaag 43380

tggcaagctt aaatggtaca tatccttttt ttaattatac atatatatac ttggtttgac 43440

tgtttatttc cttttccaca ttattatttg catttaacag ctaaaggtat aatatcaaat 43500

tattttattg ttagatgttg aaattaccct ggcttgaagt cagaaacaac tttattaaaa 43560

ttttaagaga acagtaaaat gtgttgcatg tgattctaat caatttttga agcacttttg 43620

tcattgagaa aataagtatg gataaatatt gattgttttt ctggaaaatg ctttcttatt 43680

tcaaaagtca gctgtgactc agtctgtcct cctctgtcat gacaggaact agtatattca 43740

ccgtctatga ggccgcctca caggaaggct gggtgttcct catgtacaga gcaattgaca 43800

gctttccccg ttggcgttcc tacttctatt tcatcactct cattttcttc ctcgcctggc 43860

ttgtgaaggt actgcgaaga attgtactaa gaaagtctat gttctgtgat agttctgcag 43920

atcttttgtt tggttttggt tgccctctgc tttgtcattc ttagactttg gtaaatgtgg 43980

tttgtgattt ttttcagaac gtgtttattg ctgttatcat tgaaacattt gcagaaatca 44040

gagtacagtt tcaacaaatg tggggatcga gaagcagcac tacctcaaca gccaccaccc 44100

aggtacgcta tcagcaagag gatgttataa cccattcgag catgcacatg tctcaccctc 44160

accatcagtg gaccttgggg aagctcttgc atgatgggct tctatagtct gactcaggaa 44220 gctgtttaaa aatacagttg ccagtcaggc acagtggctc atgcttgtaa ttccagcact 44280

ttgggaggct gagggagagg gattgcttga gcccaggagt ttgagatcag cctaggtaat 44340

gtagcaagac cctgtctcta caaaaaaatt ttaaaaatta gccaagagta gtggcatgag 44400

cttgtagtcc cagctactca gtaggctgag atgggaggat cacttgagct caggaggtca 44460

aaggtacagt gagctgcaat cacaacactg caccccagcc tgggtgacag agcaagatcc 44520

tatccccaaa aaacataaaa gaaaataaaa taaaacaggt tcaaaggccc aaacacagaa 44580

caatcaaatc agagcttcca ggcagtgagg taatgacacc tatattttgt aacagctccg 44640

tgggtggttt tgatacaaaa ccaagatcgt caaccacatt tattacatac ctttgattct 44700

gacatcacaa atcctcaaat ggcttgtgga taaaatgcaa ccactatgga atttcttact 44760

tttatgggta tttattcatg tacttgcaag cctatgtaat gtctatgtgt aggattagaa 44820

ttttgtaatg gaaaatatat ttcattcaga acaaaccaca gtacaataaa ctatgaaaaa 44880

gatccctcaa ccccaaatgc atgtttgtga ttttttaaag aatccctata tattcagaag 44940

gattcccata agttccatag caatattgtt tccataaaac cataaggctc ctgagaattt 45000

gaaggtgggc tacagatttg ctaaaagcag cagggcagtt attacaaaag agggcaattc 45060

agatccaatg cctttgtttt tatttttggt tttgtttcac gtgtgtgttt ttgtggtgtt 45120

aaaaaaagga agggcagatt tgtgtatctt cttctagcgt gaatctttat aaaagagaga 45180

gaaaatataa attccagaat caggaataaa atgatataat gccttgtaat tattataata 45240

ttaaattcta aaggatattt aacaactttg taaacactta caaatactgt taagacaaag 45300

gtatctatca aggagccgag gtgtacaaat tttacttttg aaaaccctca agttgagctt 45360

ttaataatta tttgttaaaa tgcaattaaa cacactaatt tgggaaaagc ctggtggtgt 45420

tgaggcatct ttatgatgct gtaaaagtct gaattgtcca gtcaattcat tatagatgaa 45480

atagtgtcag tcttattcag tttttacaaa caactgttca aacttccatg aaaggttttt 45540

tgaatatgct atccaggtta tatcagaaat gccatgtaga aaagccttca agaaatcaac 45600

catagcaagt tctcatctcc tctctcctct ctttttgcaa actatgagag ataagaaagt 45660

tcccggtcca ttccgaggtg gcagcctcta gaaggacatg ttttcagtta tgactttgag 45720

acctaaatgg aatatggaac tatgatctca gcccttaact cacatgcagt gttcctggag 45780

atgacctaag attcctttaa aaccatagta tgtttatgga gtaactctca cgtataaaat 45840

cacagctgga aaaaatagtg tttggaagca ccaccaagtc aggcaatggt tctaatagaa 45900

agggaaatgt gaagagactt ggaacaaagt tgtacagaat cattttgatg caaaattgta 45960

cttaatcaag aggaataaat tggtaattca gaggtaatgc gtctagatat tcagcagtgg 46020

taatcataca gagaattgtg aaaaaccttc attaagaaaa atgctatagc tgaggacctg 46080

gttgctaagc aaccaaagct cctgctgtca gcctctgggt gaaacaccac tcaggcagcc 46140

tagcaacgag cttcaggccc cgcccaccct ggctagccac ttgcgagtgg gccctctgat 46200

gcttttatgt gtagagaagg gagagacata atttgccatg atgctattat gtttgtaaaa 46260 tatgaccaaa ttccgtgaac agggacttct cattcagtgg tttaaaacac tagatcacaa 46320

gttactaatt tgaaggaaat aaaataaaga tgattattat aatgcacaca gcttttcaaa 46380

catggaagta cattcttgcc tttctctagc atatggaaaa tttgcagttg ttaaatatag 46440

gctattctca tcaaaggact aaagggtcct gggcttgttg gacttgttga tgtctaagtc 46500

tgcagacttc ccataacagc catgtgtata gaaaatatgt gnnnnnnnnn nnnnnnntat 46560

atatatatat atatatatat atgcacattt tcagaacaaa gaaggggctt tctgttatat 46620

taggcagctc actaggaact ggctctgaca gggaccattt caagggaggc gaaaggaaat 46680

ctgtgatgag aactgtttct ttaaaataac gaagcaatat ttttcaattg aatggtcacc 46740

accaaatgcc tccaagacac agcataggaa caaattacag agacattaat gccaagggtc 46800

tttaggagac cataatataa catgtaactg gaaaacagaa ggaaaatgat gtagaaaatt 46860

tcaaattact aaatattgac tctggcctct gagagccttg tgatttcata cctgcctgat 46920

tcttagacgt gggtaactat tgccttgtgc tcagttgtgg attgggttga aggataagtg 46980

attcatacac ccatgatata ataccctgct ttgttaatcc actatagggt tacccaggcc 47040

acaacttttt agtgatgtcc catgactgga tgttgggatg tgcaggtgag ctgaaagacg 47100

aagtcctgag taccgtatct cactaaccca gacctttact cccattattt agatgcttta 47160

cttcagatac ctgttatgag ccccatggtc ctctgatcta gactagctca tgatggagat 47220

gagcttatcc caactttgcc ttgagtagaa ctcaaagaac tgctaaattt tagagccaca 47280

caaggctact tttctgttaa aatggtagga gacagaaagc aggcttctga cagtcaatac 47340

atatctaact tttataaatg aatcataatc caataggatc atattgaagg tttggtaaat 47400

aattaaatat tctgattgca caaagacagc aatcatttct gggtctgtta ttatctctga 47460

ttgagcactc ttgctgacat gtgttttctg aattaatatc ccctaattaa cagagaagac 47520

tgataaaggt agaactggtg tgggaagaat gggcactttt gagtttttct cctttgggtt 47580

tccatcaaca ttcttagaat atgggttata caatgcctaa agcaacttgt gcataaatgt 47640

tcagtgctgg gttaaaaagt agagaaaaaa tggatgaaaa gacacatgat ccaaaatgga 47700

aagaatttct cctccaaaaa gaaagctaca tataacttaa tgctcaaaga aaatacttta 47760

gcaaaaatgt gtttactgtg ttccaagacc atttatttcc tcatgaatat acatcattaa 47820

aatatgtgag aaaacaattc aattaatgca tttcagggaa gggtgcttct agcatggtgt 47880

aataatgagt gccttctaat aagaaaccat aataatagaa tagtacaatc atcataggaa 47940

accataaatg agaattaaat tttgtgaggt tgcatcattg aatatattat ttcgtaatga 48000

aatcaagcca ggtatccact ttcactattt tctgtcatag tgactattag ctattttttg 48060

aaaaaccata aattttcaaa gaatccaaag agcatgcaaa ttgcaaagta aattttagat 48120

tttctctgca ttactgtgta tgtggatgtg tgtgtgcctg tcttgagttt ggaattctaa 48180

ttcagctgct agtgattctc cattaccaag attctttctc ttaccactat ccattagata 48240

acctgttatt catgttgttg atgatgatga tattatgcct ataatggaac aatcccagaa 48300 gcatttttag ctactttaca aaccaaccga ccagttataa aactgtagca gttgcaaatg 48360

gacaggcctg agagaattct caaagttgaa atctccatgc gcttccaaat gattgtgcta 48420

agaggagagt gccatggaaa ttcatttgtg ttattggtga gattctacta attcagaaca 48480

tcttcttaca gtcactccca tatattgaga tctgcactag tctaatggct ttacaaactg 48540

tcttatttaa ttttcataat aatactgtga agtgggcact attaatgtct ccactttgca 48600

tgtaaaaaat gacaatgatt ggaaagttaa ttatctaacc caatgttata taactggcta 48660

gtggtaaaac tgggactcaa aaccaggttg ctcagctgca agagcctttt tattcagcat 48720

gagatgctac cacctttcaa tttagccatt acaagccctg ctgtcaacat gctgtctttg 48780

ttctggcctc atgttaactt ttcttttctt aaaaaagaca gtatttattg ctggctgaaa 48840

gggacattcg catacattgt tgccaagagt atgaattaaa acatcttttc cagaaaaccg 48900

tttggcaata tgtatccaaa agtaaaacat acatactatc tgaccaagaa agtccgcttt 48960

tagaaatgta tcctaagaaa aattgggggc agtagccaag ggcatatgtt tgaggatgtt 49020

tatgaggcat tatggcaatt ttttaaaaat ggaaacaatc ttagtctcta gcaataagtt 49080

atcgattata aattatgaaa ttggaataaa tgcaaaccat taaaataaga attgaacttt 49140

attgatgtga agatatgctg atgacatatt agtacaaaaa aagatcctat gtttgtctgt 49200

gggatgcagg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtataatac ctggaatgtt 49260

aactgtaaaa tgttagtggt ggtgtctctg gggaatgggt ttaagctgnt tcttttgcaa 49320

ctcatncctc attttctaag cctttaatag taaaatnatt taattataga atgagaaagg 49380

aaaattacat cagtnnnntn nnaaaaanna aattaatgga cattctaaca gaaaccacgt 49440

atactagtta agttacattc agttgcaagg gacaaaaaag cccaaaataa gttgcttagg 49500

tgacgtatgc gatgaggaat cccattctct ctcatataaa gttggaggtg ctcattcagc 49560

atggtatggt tgtctgtgtt gccacggatc cagacacctt ccttcctgtt ccactttgca 49620

cagcccccat taccacagtc accttgtaga ctagcatggc tgctgtagct ctaaccctca 49680

catccatatc ccagagagca ggaagaagaa ggaaatacag gaaaagaggg tttctgcttg 49740

catcctatag tcctgtatgt agttacatgg tcaagctcag cttcaaaaaa agggcaagaa 49800

ctggcctttt ttcttggtag cctagtggcc ccatacagat tctctttgta aaacagggaa 49860

aagagataca aaggaaagac tggcagtccc tgccacgctc cccgcaagga tttgagtctc 49920

tgaggattct tggcttctgc tagctagccc agcttctgtc tcaaagatct actgtaatat 49980

taaccaatct aaatgttgta taattcaggg gattaaagag aaattgaacg tttttgcatt 50040

tcctttactt atttcataca aataagtgac cttgggagaa cacttttact gaaaattaaa 50100

ttgattcttt tgattttggt ctattgaaaa aaaaatttaa gagagaaaat tagtctgcta 50160

ctgatcaaag tcaatatatg ccccagggat cctttatgcc acaattagag agtataggat 50220

gtatcaacag agggttttta ttaataactg gtgactggca atgctggttt tgtcttttat 50280

ccttaaaatg aggcttccag aattttgaca gaaactttac aaggtcctat aacaggtttg 50340

```
catctccact gatgtctata ttctcagaat gagctattca ggaagaatat gcttgattca  50400
ccccaaaaag agtttttgca gtttattcta aagaaaatca gaatcctgta aatggacagc  50460
tgagatgagc aatttgcagt tggcatcttt gtcaaaggag ctttatgtaa atgtgcacct  50520
taattttgtc atttaaaaat acatatagga ggcagaagct gggggattgc ctgagctcag  50580
gagttcaaaa ccaccctgga caacaagctg aaaccctctg tctagtaaaa aaaaaaatac  50640
aaaaattagt tgggcattgt ggcgggtgcc tgtagtccca gctactcggg aggctgagac  50700
aggaaaattg cttgaacccg ggaggtggag gttggagtga gtggagatca cgcccctaca  50760
ctccagcctg ggcgacagag caggactcta tctcccccca aaaataaata aataataaaa  50820
atacatatag ggactgtgat ttttggcaac tgtctctcac cacctcactc attttataga  50880
aggcaagttt attttccaca aaggatctac caccactact accagggatt tgaaactcta  50940
attacaaatc ctaggtaatg ggatctttgc atcatctcat tcgtatttcc atcagcagtt  51000
acccccttcga atacagggtc agactaacag gaggacataa ataagggtga caaaggaaac  51060
atttatttct gtaaaacagc gacctcaggy aggaatgaaa cagtgagcac ttgtagttag  51120
tatcattatt tagaaaaaca aaaatgttta attttgccct attcctacca tacacacata  51180
cacacacact cgctcctaag aactagaaag aaaaatatgt tagtgggaaa atacgtgaaa  51240
gtgggaaatg tcttccagct gggtgtaata ttgacacata ggtaatatgt gatgatacct  51300
aggagataat aggtacctag gagataataa cctaggtatc taattcaatg gatttcagat  51360
ttagtgggcc tgtagcatct tgaatttcat cacttgggtc tttaatccaa aagaaactac  51420
ttgttccctt cataatacgt acataaaatt tctgctgagc aggttaattg ctgcctagtt  51480
aaaataatta cacaaagtgt attaagcaat gtggcaatct tctaccatta aaataacacc  51540
agttctttcc ctatgccaca cgcacacttc ttattttatg gatcctgaga atattttttat  51600
taggtcagtg ctcctgttca gccttgaggg taagatacca tggtagtcca agccagggag  51660
tggaagactg tccactaaat aaaaataaaa atttgcaacc tatagtaaaa acacaccaga  51720
cacaatgctt ttaaattctg attttcattt tctcccagat ctaatagata tttatcaata  51780
taatggggat aaaatactaa agtatttctc ttgtggagaa gaatatttga gcaattatat  51840
gcaaagcata tgcaagacag aaagtgcctg acatagcaag tagtcaataa actcaagtta  51900
ttatcattgt tcttttttca ttcatttaag ctttcaatgt taaagaacta taattgaaat  51960
gacttcattc caaatataaa tccataagaa gacaaatttt tgataacata aaaatataca  52020
aagctggaga catagagttc tcaatgttcc atattaaatg tctctgactc atctcccttc  52080
tttgtcttga agatgtttca tgaagatgct gctggaggtt ggcagctggt agctgtggat  52140
gtcaacaagc cccagggacg cgccccagcc tgcctccagg tgcagtacaa tgacattttt  52200
aaaaatcgcc cagcaaaggt ctttgaattt tatttcatcc aagaaaatcc acagctcttt  52260
aagctctaga tttgtccaaa tttaaaatcc tgaagttaga gatggtattt cactccttcc  52320
tctattccca ggacctagct tttttttttt taacatacac aatagggatt tgataagttt  52380
```

ctgatggctg caggcatgta agagcatttc agtggtattg aatcaatgaa gaattttgtt 52440

gacatgtgaa atcttataaa aatattcttt accgaaggac tgagttatgt ggcagtgggc 52500

aaattcattg tttcatacct ccccgagtaa ctgggaaaaa tatgttaata catagtctct 52560

ctgtttttct gcatttggaa gctttcagag gaacataatg tagaggtgtt tctttagcaa 52620

agtgcactga tagcaaacat aaggattgca ggtggggcct gagagtcctc atgagataga 52680

ttctcacagt gattagaaga tggagtctca cgtccctgcc tgtgaacttt ctggaaaaac 52740

catcttctcc aagctgccat tgacaacaat atggataaca ataataacaa taaggcccaa 52800

taaactcctt tatctcttct tcaggggggcc atactgacat cttctcttcc ttggtttccc 52860

ctccttgccc cctaaatatc agtaactcat tcaaaataat gtcaccttac caagagcagc 52920

acccctaact ttccataata ttttcacttt cattttccct ccaagcagcc cactcgtagg 52980

accggagaat tgattcttcc acctggagaa ttttattttc tttagccttt ttggttttca 53040

gtgacaaatc ctcttctcgc aaggggtggt ttccatagtt gtttatatcc tgccctcata 53100

atttggagaa gtgttcacat ctgccgtggg atgagactgt atctcttttc tttcttttgg 53160

gtctttctcc agatagggac ttcttatgca actcaaggat gggtacatga aaaataaaat 53220

tgtactctga gccattactg tgggctatgt ttatatggcc attttaccat agagttattt 53280

acttcttttt gtttctattt gtattgaggt gtgattaaca aataaaattg taaatactta 53340

aggtatacaa tgtgatgatt tgatatatgt acacagtgta aaacgattgt tgcaaccaag 53400

ctaattaaca cacccatcac cgcatatatt cagtagttac ctctgtgtgt gtggggtggg 53460

gggtaacact taagatctaa tctcctagca aatctcatgt atacaataca gtattattaa 53520

ccattaatga atcactttgt ttattagatc ctgagaactt ggccatcttt tgactgataa 53580

ctaatatctc cccatttcta ccccctgtcc ccagcccctg gtaaccacaa tcctactctc 53640

tgcttcaatg aatttgactt tttgagattc catatataaa tgagattatg cggtgtttgt 53700

ctttctgtgt ctagttcatt tcacttagca taacgtcctc caggcttatc catatagttg 53760

tcaatggcag gatttttttc tttattatgg ctgagtaata tttcattgta tgtctatacc 53820

acattttttt tatccattca tccatcagta gacacttagg ttgttccctg tcttgaccat 53880

tgtgaataac gctgcagtaa acacggggct gcagatgtct ctttgagata ctgatttcat 53940

ttcctttggg catatatcca gaagtgggat tgtgggatca tatggtagtt ctatttttaa 54000

tttttggga aacctccata ctgcgtttca gaatggctct accaatttaa atccccacca 54060

acagtgtaaa agggttatct tttctttatg cctttgccaa atagttatct tttgtctttt 54120

tgatggtgac cattgtaaca gatatgaggt gatatctcaa tgtggtttaa atttgtattt 54180

ccctgataat tagggatcgt atgcactttt tcatatacct attggccatt catatgtctt 54240

ctttggaaaa aaatctactc agaatttgct catttataat cagattattt gcattttgtg 54300

tttgttttg ttttttagtt tgggggcttt ttgttattga gttgtatgag ttctttgcat 54360

cttttgaata ttatctcctt ttcagaaata tggtgttcca attttcccct gttccataga 54420 ctgccttttc attttgttta ttgttagtta tttatttctg tcagtttaaa aaaatggagg 54480

gcaataagta gaaaagctta caatgaatat tgataagtat tattttgtgg tcctccttcc 54540

atcctatatt tgatgttgac tatgctgcta atataaaact actacaacta cctttactcc 54600

tactatcact actaatagta ataacaataa tgacaaagat tatagtacta gcttcttacc 54660

taaaaagcac ttttattgac attatttttt aatccatcaa ctgatattta ctgagtacca 54720

actatgtgct agacactaag atagtcttga aagaaacatg cacagtctcc acactcatag 54780

agcttacagt atgcttgaga gtgggaggga aaaaggcagc attgagctag caatcacaag 54840

agtgatgtgg gttttgaaag aagtaatgca agatgcaatg agaatatata aattggggtc 54900

cgctttgaga ggtcagggaa ggttttagtg cataatctta aaatcaaaac cttgagtaat 54960

tccagcattt tgggaggctg aggcaggcag atcacctgag gtcatgagtt tgagaccagc 55020

ctggccaaca tggtgaaacc ccatctctac taaaaatcca aatattaccc aggtgttgtg 55080

gtgcacgcct gtaatcccag ctacttggga ggctaaggca caagaatcac ttgaacccgg 55140

gaggtggagg ttgcagtgag ctgagattgt gccactgcaa tgtagcctgg gcaacagagt 55200

gagactctgt ctcaaaaaaa aaaaaaaaga gagagagaga atgaggaaaa gaatgactaa 55260

ggatgaggtt tactacaacc ctgtgatgta gaaagaccta catgcctggt cagccattta 55320

taggagaaag tgctgattca gagaacatag ggatgtgaat tacattcaac cagagggaga 55380

acccaggact aatggctggg tcttatggct cctagcccca tgctttatcc agctgaaaaa 55440

aatatgcaag taatcataag ttcgtgatag agaaataaac tatcagggct actctgagtg 55500

ccaaatcaga gtcgtatggt atccagagct ttaaaatcta ttaaatcctt gttaatagat 55560

gcctcaaggt ggccatctga aattaggagg caagtcagcc tccttatcac aaagatttga 55620

gaataacaag ggattaatga cagggtgaag tgtgatgaca aaatcgagca aggagatagt 55680

aaaataaaga ttggaatggt gtggggatag atgttgggct tgaagtgtat agatgtggca 55740

tctcagagac ccctatatga gatatgtgtt atgaatatag cagcatgctt atgacatctt 55800

atgaatgtga ctcatatgct tagaaactgg ggggaaaggg gtcaggtaat tcagtgtgac 55860

acaagtttca ttccactcaa cttcatgtgt ttggccctga ataaagcttt ccagaggcag 55920

caatgaagag tgcatggaag aaaagtggga tggccaagaa gttatctgtt tgtaccgggc 55980

aggagctgtg attagagaga gaatacattt ctatgtctca gcaagaaaat aattaattaa 56040

ttaaaagaaa agaattaagt aaaaatatat tgaagataaa atggtcctga aatacttgat 56100

ggctttcact aatgattaat gccatctaaa actgggtaac aggaaaactt tcatctcagt 56160

cttccactcc ctggcttgga ctaccatggt atcttaccct caaggctgaa caggagcata 56220

taccacttct ttttatttat ttatttattt atttatttat ttatttattt atttnnnnnn 56280

tttatttatt tattttttga gacggagtct cgctctgtcg cccaggctgg agtgcagtgg 56340

cgggatctcg gctcactgca agctccgcct cccgggttca cgccattctc ctgcctcagc 56400

ctcccaagta gctgggacta caggcgcccg ccactacgcc cggctaattt ttttgtattt 56460 ttagtagaga cggggtttca ccgttttagc cgggatggtc tcgatctcct gacctcgtga 56520

tccgcccgcc tcggcctccc aaagtgctgg gattacaggc gtgagccacc gcgcccggcc 56580

ctttttattt ttaaaacaag caaaatgtga acatgttctc ataaaaaatt aatatggaaa 56640

taaaattaaa aggaagagtt tacttgaaat cttctttcaa tgtatccttc ctcattcttc 56700

ttccaatgag taacttggag gttaatgttg agtcgattct tctcagactc ctgcatttat 56760

gcattttatc tcacatgaaa aagcatttct gcctccagga agcactgaca ttttcacaga 56820

atgtttcctc ttcctacacc acttccttct cactcatcct tcagaactca gctcaaacat 56880

cactttctta tcctcaaaaa ataagccagg tacctctgtt acatgttttc atatcatcct 56940

acacttttct tttatgtctg tcataatttt aataaaataa attaaggata taattagttg 57000

ttgatgtctg cttctccaac caggggaaat tccattactt tagagaccat accagtcatg 57060

gtcacttctg caactgtgct ttagacatga gagatgatca attatgttta ttaaccgaga 57120

acagtccagt gaaagaagga ggcagtacct tcttggctta tgaaagtact tatttgacct 57180

ctcgcttaga agctgctctg tttgtcctcc tggtcttgtc tgtaagaact agaaattgaa 57240

tagactgtgg caataacacc cttttccata cagaacactt atttctatat gcatgtcgtt 57300

ggagtctaaa gcaagcagta cacttggagt ttgcttccag aaagtgcctc gcagaggttg 57360

taggactttg gatatctctg gatagttttg ccaggatgtc tcctctaata tgtggtggca 57420

gagccttggc aggacataac cctgacttct gttatgtttc acctctttca gaactctttt 57480

catcttgtaa gcatataaga taaatttgtc ttatgtttag aaatatagtg ttaggaatag 57540

catcattgaa aagaacccaa actcattctt tttttttttt tttttttttt ttgagacaga 57600

gtctcactct cctgcccagg ctggaatcca gtggcacaat cttggctcat tgcaacctcc 57660

gcctcccagg ttcaagtgat tctcctgcct cagcctccag agtagctggg actacaggca 57720

tgcgccacca cgcctggctg atttttgtat ttttagtaga aaccgggttt caccatgttg 57780

gtcaggctgg tcttgaactc ctgacctcat gatccgccca cctcggcctc ccaaagtgct 57840

gggattacag gtgtgagcca ctgtgcccag ctgaacccaa actaattgta tgttaatgag 57900

aaaacagagt ttattggaat gaaacatttg aacggctgta ctttgtaatt ttgaaactgt 57960

tgaaataaat tagaaaaaat gaggtatcaa taatatctga gagacagcaa actgatgttt 58020

gagttcccat ttaactactg ggaactgttg attttgttca aagtgggact ttgtttttaa 58080

atttgaccca aagtttaaaa accagagtat tttacataaa aacccacacg tcctacccat 58140

cttgtaaaag agaatgatct ggcaataaag agctctcatt tccacctggc tagagttggc 58200

cagagctgag tggtggctgc cacatgagca gaggcagaga ggctttattg tgcctcagtt 58260

tcccaactct ccatattgtc tccacagcac tggaggtgag tgttttacct acccaagaat 58320

tcatacattt atttttctaa gagtagagaa agatttctct ttgtccgtga ctctatcgtg 58380

tggaaataaa taaacaagta ataaattcat aaatatataa atgtataaag atacacaaat 58440

gcatggatga atgaataaag aataaataca aggcatttta aaagcaccct gtgattacag 58500

```
aaaacaatga gaaatgaacc tagtatttta agaaaaaagt tattccgatg tgttaaatat  58560
gtaaccctaa gcctagtttt ctcattatgt tacttgctag gtctcaggag gcacttggtt  58620
gggtgactca cggaataaat gaaaataatg taactttaga tcacaaagat tatataataa  58680
caattagtcc acacatttca gtgattccaa aatagtattt tcatatattt tatgttttta  58740
acattttaaa atgttaaaat cataactcag cataattgga agctggctca agttctctta  58800
taaatgaaat gtactaaaaa atcaaaactg tctccattgg tattatgtat tatacttttc  58860
accatatttt ctactctgta tcctcatcca gttactataa gtaattaggt aatttaataa  58920
aacaagattg attaaggtgg cttaattttc ctcaccatta ttaaattaaa tacatgaagt  58980
aaccaatgct tctgtgatgc tgagtttttt acgagtaata atttattctc tctttctccc  59040
tctctctcta tgtgtgtgta tatatacaca aacatatata tacacacgtt tttatatata  59100
catatataag tgcatatgtg tgtatatgta tatgtgtata tacatatata tagaaacatt  59160
gctggcataa aactattagg attttatata tttcaatggg cagtgattct gatcacacat  59220
tagcaactaa ttttaaattg ttttcctaca ttcccacatg tgcacataca tattttacct  59280
tatatcaaat atttccatta ttatggtatt tgtttcctct tttcaaatat tgcttatgga  59340
gatagtcaat gaattagggc acataaaaga tcttcttact cttctacaga aaaaataggc  59400
tctattgagc catttgtgga ggagaatata ggcatctttt gtctgcagtt tgaaaccaca  59460
aatttatttt gtgttctgaa acacaaatta caattacaat catagctata attataatta  59520
gaataggaac catatgcagg catgcaagga aacgattcaa tctggtgccc ttgaagtgat  59580
ctttctgttt ttataataaa tgaaaatgca catacttgta aatttaactg caagaccaca  59640
agtttggagc aggaggacta gcagaaaggt attgttagct tggtaagtta tatgtggtga  59700
gctctttaac gccacttgtt aaaatttttt gttaatcaac ttaattatat taatttgtct  59760
aggacatttc tcatgtttat ccagttatta atgaaattta taccatgctt aataaggctg  59820
agtttaataa gatactgagc atggtaaatg tgtgcaattc ctaataaatt tatcctcacg  59880
ttaaaacgag taagcgaatg tcctagaacc cctgaaaaac cccctcccat cctctgtcat  59940
gcctgcaatg tctttctggc tctttctcat ctctgtgcat tgctgatgtg cttgtcactt  60000
aggggcccct cgtgcctgtt cctttttttt tttttttttt tttttttttt ttgagaagga  60060
gtctcgctct gtcatcaggc tggagtgcat tggtgcgatc tcggctcact gcaacctcta  60120
cctcccaggt tcaagcgatt ctcctgcctc agcctcctga gtagctgaga ctacaggcgc  60180
ccgccaccat gcccagctaa tttttgtat tttagtagag acggggtttc accacgttgg  60240
ccaggatgat ctcgatcttc tgacctcgtg atctgcctgc cttttcctgt ttttggtcca  60300
tactgtgggg acaagatgac tcgtggtgaa aaagaaaagc ctaggaacat agataagcca  60360
agccccatac ctggtttctt ttcctttttg tgctacaact gcctaaatgc gcatctccca  60420
acatgacctc ctcacctctc ccttcttcct ccccagctac agatctcctt atcttcacct  60480
gatgcataga cacccaacac aacatcttgt tttgagacta cacgctattg tggtctgtta  60540
```

gagataagac tctggaaaca tatttcagct acactaccta ccagtggtac gcgacaggac 60600

acgtttatgg aaaatgggga tgagaatgca gttttcctgg gtttgagaga attattgagt 60660

atttatgact tagtaaagaa ttatttggaa ggaacagacc atctgctcaa agggaacatt 60720

gattgcaagt gtcattcggc ctttctctgc ctgtctctca ctctccgttt gagccacatg 60780

atttctgtct gtgagcctgc ctctgttttc tcaccttgaa cactggcttt attctgtttt 60840

gttttggttt gtttctctcc ttgcatatta cccaagtatg gctgattgtg actgggtgta 60900

atgacctccc aatcaaaagc cctaacgaga ctccccagaa aacacccaaa ggtgaaggtg 60960

cactgggaaa aaatgaagtt gcgtgggggc atcagaaaca ggaggcgtgg ccgggagcag 61020

caggtcagga gagtgcaaat ggaaggaran gagcaggtnc tgctgaagcc aagtattgga 61080

tttccctgca gagcaaaccc aaagtcgttt gaaaccgccc tttttaatcc tctctgaaag 61140

ccttgcccca tctctgctgg aaccactact caaccacccct ccttagcttc atggaatggg 61200

aggagagagg acacaactct ttggccatac taagtggtga gacttgactg cagtacttac 61260

aactcaggga tcatttttcc tgtcaaactt ctraatctta taggacttga gaagtttatt 61320

aagagctaaa aaaataggct attacagtct cacctgaagc tcaaatcatg ctgtataaga 61380

tagtttcttt agagaaacac cacgtagtta ctataaactt acaaatggac aatttgtaaa 61440

ttggtgctat tttacaacat aatattaagt cgatttcctt gcatttaaaa aatacatatt 61500

ctgaaataat ttagaaacac agataaaaaa tacctatgtg tctgtcagaa ataataatag 61560

ctaatattta tggaccactc actgtgtgtt ttattaggta ttaaattatt taatgctcac 61620

agcaccataa aggtaggtac taatacgatc tgttttacaa aagaggaggt tgaggcacag 61680

acagcataag gaatttgatg aaggtcacat agttactgaa tggggatttc atctcaggca 61740

gtccacgtta tacttttaac tacttcattt tcacccctgt gtattaccct actgtgaaaa 61800

taacccagtt tcttcattca tttattcatt ccactatcga tggacattga gcttgnnttn 61860

nattgttttg ctattgcaga ccttctaaga cctttcttat ccgtgccttt gtatgcacac 61920

gtgagagagg tgttctaaca tatatagcta ggacagcaat ttctgggtta aaggatatga 61980

gcacctttga ctctctgaga tatggcagat ggctctccgg atgatagggc tattttgtat 62040

gccttctaga aacgtgtaag ggtgctcgtt gtctacatct tcatctgcac ttgatgttat 62100

caggctttta caccagatga cacgaaatgg tgtcctgtaa tgtaaatttg catctccctg 62160

atctccctga ttaatggatg ttgaacattt ttaggttcaa gggccacaca ttttcttttt 62220

tttttttttt ttttttttga gatggagtct tgctctgtcg cccaggctgg agtgcggtgg 62280

cgtgatcttg gctcactgaa agctccgcct cctgggttca cgccattctc ctgccccagc 62340

ctcccgagtt gctgggtcta caggtgcctg ccaccacgcc cggctaattt tttgtatttt 62400

tagaagagac ggggtttcac tgtgttagcc aggatggtct cgatctcctg actgtagatt 62460

gcttgtttat ctcatttgcc catttatctg ttgatttttc taagtgatgt tcaggaattc 62520

tgagtgtttc taaagggtat attgtttctt tctaaataaa aatttaagat attagatata 62580 aagacaaatc taaaatatta aatgcttgac acatgtgcct ttttcatctc tcccatgttt 62640

agtcaataag gcaattcttt tatgtggtat taaagaaacc gattagaaag ttgcttcgtt 62700

gtgtgcttaa gtgtgagtct tagaaaggtg aaaagttata attgactctt ttagcaccaa 62760

accaagcagg tatatagctg tttaaatact attttagcat aagccatagc tgtgggaagg 62820

gtcactccag tgcttgataa atagaagtct attcaattgg accctcattc ctcagcggct 62880

gctgtgacaa gccgttttcc cgctgctgct caggtaggag ttttcatctc ttgacacttt 62940

gacagatgtg gctgcacact tttgaaccat cttccaacac acccccgtgc tctattacca 63000

tgaggaaagt ctctgacaag agaatgaaag tacatatgca tgagacgtgg tacaatggat 63060

gtcagtttcc ctctggtagc atattccacc gcgtgcctcc tgtgaggaac agagggaggt 63120

aacaggtgct tgctgaaagt aggtcaaacg gtggcgaaat tattccttca gtgttttaat 63180

gctctttgaa aacatttcca ttgaaggcta agattgaggg gaggtggcct gtgtgatgtg 63240

agcatttcaa gtgacaaagt gacatcctag gcactcctga aactgtcttc ttcgtgaacc 63300

ttcccacaaa ctggcagatc atttcaagaa ttgatggagc ccatgagaat gtaggttccc 63360

tccagaacgc agctctctcc ccagcccatt cccacaaaga aactaggttt tccctgcatt 63420

tccagaaacc tccaaaccaa cccattatgc ttctcctttt gtcctctcca aattcaatct 63480

gatccttctt caattctgtc attacttcag tattttgata ttttttgact cattctctcc 63540

actccatctt catattcctg ccctcaaaat caggtcagtg cttcctcaac agagtgatct 63600

ctgattcctt gattatattt ggtcattttc tttttgttt ctctttcttc tctattgctt 63660

cattcacgat aggcatttgc caaaaatttt ctgaattata cgtgattgta catatccagc 63720

tgcgttcacg tgcagaaaac atagcctcac acacagcaga cccaaattct tactctcccc 63780

ttaggcaggt gtcaaacgat acttccccca aatgctcacg catgcacgca aacagagaca 63840

cacagtcaaa aacacagcat gcatgtgtac acacacatac acactcatac aggctcacac 63900

atgcatgcag agacacacag tcaaaaacac agcatgcatg tgtacacaca cacactcata 63960

caggctcaca catgctcaca cgcacacaca cctcaaacac acacagtgac acagacacac 64020

acatgcacac tcgtgcacat acacatgcct acacacttat acatgctcat gcacgctcac 64080

acaactcaca tgtacataca cacacagtga cacacaatcg cagacacaca catgcacaca 64140

ggcacacaca catatctaca cgcacatgct catgcacaca cacttcacac gtacatacac 64200

acagtgacac acaatcacac gcacacacat atgcacacat gcatatggct acacacacgt 64260

gcatgctcac acacacctca caatgcacac tcacaaacgg tgacacactg acacacacac 64320

acatgcatgc acacagacat gcatacacac gcttatacac acacacacct cacgcatact 64380

cagtgacaca gtgacacaca gacatttaca tgcttatgca cacatgcaca cacacatgca 64440

caaacacact tgtacataat catgcatgct cacacataca cacgcacata cctcacacac 64500

acacatacac tcacacacag gcgcacacac attcattttc aatttctcag gggaatataa 64560

aacaactgct tggaggaaaa gtttgtagga aaaagtgaat ataaaaatgg ttatgcagta 64620 attacctgga ataaatatta aaaccaggaa aaataaaata gctacatgta gaaaggggta 64680

gctaaaatgc tacttgaaaa aggcctattg taattttgaa gtcagtggcc aaataaacta 64740

aacagagagg ccaattgtat aaatcctcag ttctaaatat cccatatata tttatttaaa 64800

tatatattca atatgtatct aatatatcct gattatattt cnaaataata attatattta 64860

gaatagacac cagatgacca ggcataggat aataaaaaat aagattaaaa tatcagtgtt 64920

atttttattt atttgtttat ctgtcttact ggcaattaaa cagttttgct tctttgattt 64980

ttaaacctca ctggaagaaa tatgcagaac attttaaaat tatttgcatg ccacattatc 65040

tcaaactaaa ttcctgtctc ctattatatt ttcatcctcc agtttgaatc agtcctataa 65100

ctcttgaatg tttacaatta tattaattag gagacatgtg ctaacttatg ccaagcattt 65160

ttaaaatgtc cattggagac aaatttccat caccaaagat ttgttgcatt gtcacctttg 65220

gcaaaaatga aggatgagtt tttctgtttg ttcgtttggt ttggttttgc aagtcttgat 65280

ctccaggtta aatacacagc cttctttaag acttaaaccc agggtttgaa agtgctgtgt 65340

attttctcag tgttcaacac aatctgcagt agaagggttt ggagcagcaa tatgatgcct 65400

tatgcactct aaagctgctg acaggtttag acataatcta aaatgtaagt cactttgtgt 65460

atcaggcagt tcagtccttc gcatttgttt taatgttttt cctaaatact ctaaatattt 65520

caaaatgagt ctttccttaa agattctaca tagctcaaaa ttattttatt tatgtatgca 65580

tttattaaat atcacaagtg aggtcctgca ttgcattaga aatgcagtgt attagaaagg 65640

acagctagct tcaattttga attcataatt tgtctcttga gtttaaaaca aaagttttgt 65700

gttcaactgc ccaccgttta atactaaaaa taaaagatga acataaaagt caatttgaag 65760

tgattctaga tgtaggaatt taattagctg tgggaatgtc aatgacatct caaaaccatc 65820

atggcgtcgt gtcacctgtc tagcagccct attgtagtcc ttcctgactt ctcactaact 65880

tacccatgaa gacaaagata aagaaaaact cacaacaata caggtaacta agaaagcacc 65940

ttttgtcaaa gtctgttgag attattcagt aactatgtgg ttggcggagt actagtgaga 66000

aaaatataaa acatcatgta tattcaaatt gtttaatcta atgcgtccat ccttttagga 66060

actcagtact gaagtctagg aacatcaggc taaaaaaaat aatggttgga tttttctact 66120

ctttctagac accagggaac atgagagaaa attattgtgg aaggactaat attcttattt 66180

ttattattac ctatgtccag ccaagtaagc aaggtcaaca cttcaccgtt gaaaagccaa 66240

gattttagca ttttcattct taatttttac tctagtattt cactgacact tgagtgtatg 66300

tcagaantca gagggtcatg aagaanngga cagccctggt gactgccaca ggctgagact 66360

tactgctgga acatgtatag tgaggcctcc tttgttgacc atcaccctgg gtcttcttgt 66420

gcccaaagca gaatggtgtg gcttcctgga ataggagatc atattacttg aagcaacaag 66480

attgtattgc ttgaactcta gtgtcctaat ttgttttact cactttgtat taatacatgc 66540

caaagatagg gaaggggatg aaagggtctt gggaccacca tctaggttac agcatccctc 66600

ttttaggagt ccttgctaag acacctcagg ggaaactgcc tttctgaagc ctggctcatg 66660 gtccagttaa ctatgtcaac agtaccaaca ttactttcat ttaatacagg agatagagaa 66720

atgtttcacc cccgaggcta ctttcctcag aaggaggcag aagatgctcc atggccaatt 66780

gagtacatgg aggaaattaa gttatatcaa aaattattgt tgaatctact atgaatccag 66840

ccttgcagga actgagtact gaagcttagg aacaccaggc taaaatatat atatatatat 66900

atatatatat atatgagctt tttctactct ttctggacac caaggaacat gaaagaaaag 66960

tgttgcggaa taactaatat tttttatttt tcttacccat gtccagccaa gtaagcaaag 67020

tgaacacttc attgttgaaa agccaagatt ttagcatgtt tattcttaat ttttcttcca 67080

agtcaataat gaagtcatgc aggggtattg agttcttaat ccataaatta taggttagtt 67140

aattgaaagg tttttgtaca tgtcatcatt tacttttcct ttctcctact aatccataaa 67200

tgtaggccac aaagtgtgag gaaaccagga gattttaagt agtgttctat aagatagttt 67260

aaattgtgaa tatccattat ggcattgtac tctaggaaat acatgttgtc ctaaattgaa 67320

gctttcataa agaaaatttc ccatgtttta aatcaattca agacagatga ttttattggg 67380

caggatatag gaaaagttct gttttaaata tttcttactt taactgatat atttgtatct 67440

ctgtagtctt agtggcagta tcccaatgcc cagaaatggc ttttgtgtgc tcctagtgct 67500

tagtaaggca tcatagctat gtgactgcca tggcattatg attgataagg agtctctaaa 67560

cagcagaaaa ggactatatg taatatatat gacccagact agagtaacaa agctaaacac 67620

cttagaaaag cccagaaaat agtctcccct tcccacaagg gttattattt atgatcaatg 67680

tatgagtttt caaataccct acagcagact cagagatgcc ataagagctg ggggagctgg 67740

gtggtgaaca atcaagttga agtccaggag tcctgactgt tagcattttc tgctttccat 67800

aatgtaaatg ctcccaccat ggccaatttc aagtcaccaa catgacatca gtaaacatga 67860

aggtggagag agatgtgatt gcagctccag cacaccacag ctggggcaag ctgcactccc 67920

actgcagtca gagatggcat atacatacat acatgtcata catgtctatg taaacacgtt 67980

gcaaaatttt gtttgcagaa agtggaattt ttgatgcctc catccaaagt acctagtttc 68040

caacatcact gtagagggaa agagagaggg agactttcga gaaggcacac cctttctcaa 68100

agtgacacat ccaaagtgat gcatcacttc catttctgtt ctggttgtga gaaccaatca 68160

cattgctcca cttcagtaat agtaagctga gtgatgtaga tgagctatat gccaggaaga 68220

ggaaaccagc ttggtaggta taatagcttt tgctaaacca caaggcatcc taaaacttaa 68280

tggcctggac agacactaac tcttgctcat gaatttgcgg atcagttgaa cagccctgct 68340

agtctgagcc aatctcagct gactacagct gagctcaccc ctgcctctgc ggtcagccac 68400

tcgtcagctg gcatctctga ttttgggtgt ggctcgccat taattggaat gaagagaatg 68460

gcagggccca tggccacctg gccaccatgc aacaggctag gctgcagggt tgatgtggtc 68520

gctcggggtt ccaagagcag caagagtaat ttctaataca caagtatttt ccaagtctct 68580

gcttacttca cctttgctac catctcactg accaaaataa gtcacatggc tgagtctgga 68640

gttggtgtgg aggatactgc caaggatgtg tggatacagg gaggtgaatt caattggagg 68700 ccactgggtt ttttcacaca ggctaagagg aatttgaaaa tcactggctt ggcatcataa 68760

tcaatcctgc ttagaggtga tttctttcac accagccatt actggtgctt tttccttctt 68820

gttaaaatct agtttgtgtt gtaaagtcaa gaggagctgg tgattgaaaa aaataacatc 68880

cacatgtgtt tgttgatttg tgatttcaca tctaacctta ctcttacaga gtagagtgaa 68940

agctatactc acatggccta gagataaagg gtctgtccag gattccctta ctcggctctc 69000

cattccaata aagagagcca gccactgata gggcgtggca gctggcaggg tgtccactga 69060

aatcataggc tcttggtacc tttgtggaga tggacatgtg gattcccttc tgtcctgact 69120

ctctaaggac gcatcaggaa cagacagatg ggaaaggcag tctctaagga tatccactgg 69180

ttatgttttc agtgttgata taagtgttac ctttttcata taaattcttt gtattggaca 69240

gaataattgt gtacattgat agtagaagta atgaatatcc agatgtcttc taaataaggg 69300

aagcaaaaga tgtcccactt gcaaatggga gatttatttt aggttttatg catttggtct 69360

tttacattgt gtccttgttc ttatgttcat ggctcatatg caagatggtt tctgagaatt 69420

tcatcatgta attgcattgt tctccaaata agtttatttc aaaatccaat taaaacacaa 69480

gttttagtac atggctgatg cagtattcag tggaggagac tttgttttac taagttgagt 69540

tgttttactc actcaagcag caatttaaag acagatttca tttctgagct gtactattca 69600

ttatcctaat ttatatttat tgcaacctct gagaccacta ttttgtctct ttgcctacag 69660

tcttcagaat ctttaagtta ctgtttcata ggaaaaatta ttggagaaat ttaatagaaa 69720

gagcttttat gaagtactta agcatttata ccacaaaaaa agataagata gctttagatc 69780

atttgaaag tttataaatg ttcacggtcc atttcaagat tactaatctc ttccnnnnnn 69840

acctgatttc ttatgttaca ccttcaggcc ttccctgtaa gcaactgact gcatctctac 69900

tttcaaatcc ttagagtgct ttccaagact gaaataatga gagttacaga gcacacggtg 69960

tccactgcag gctagactct acactgaagt ctcatttatt ttcgttttac agccaaaaga 70020

tgtagatgtc atcatactca tttcctctgt gagaaaactg agtcacagaa aggtgaaatc 70080

cagcacagag acaagattta aactcagttc atcctgtctt caaaattgat attccattaa 70140

accatggcaa tcttaaatct ttctgtgaaa aacagattgt cttcagcttt cagatttagg 70200

gttcatgtac ttataccttc tatccaacca caaatacttg gtgaggaagc tgctgtcact 70260

cttttttaat atatagagca atgttgactg atgaattttg actgtctgat acaaggcaag 70320

agactacctg gtttgggctc acaatgtcta agatgtttct atctctgtgg gcctagaaat 70380

ctacttagac tttctgagcc tgtttctcca tatgtaaaag gggatgcaaa agtatctacc 70440

tcacagagtt gttgaggatt aaagatgtta ttacatgagc aagcaacatc tttggaatga 70500

atcttggcac aaaccaagca ctccagaaaa cagtagctct tatttcaata atccatcctg 70560

ttggtgtcca ttttaattgc gtgagaagcc ctccccgtgt ttggcttaga cttccacagt 70620

ctcctgggag gtttgctttc atggcaggac tcagaatcaa tgcgggaaaa agtgaccttt 70680

cttgttgttc tgcttgaggt tggactcagt cccattgttc cagactcatt gtctcattac 70740 acttttggac acaaagcaag atgaaagcgt tggttgctgt tatgagtttt atatctgtag 70800

ttgtcaaagt tacagagacc ccacgttggc aagtcacctt gtgtggaacc ctatggtcag 70860

gtaccatttt gtagccctgg gcacttttat acnaatcatg tttcnataga aggtgatttg 70920

agctggtata caagacagtt tgagctggtg tctgtcaaaa tcgttgacaa tgctgtgtta 70980

gtttcctgtt gctgctgtaa caaataacca caaattcagt gacttaaaac cacacacatg 71040

tattctctta gaactctgca gttagaagtc caaaatggac ttccttaggc taaacacaag 71100

gtatcaacaa ggctgcattg ctttgggaga ctctggagaa gaaatggctt ccttgctttt 71160

ttcacatcct ggaggccacc tgcatttgca tttcttaact ttatgagagt ctgtaatgat 71220

gtctttataa tgcctccaaa tggttgcaaa actaacacat gcttttgttt ttatgaagcc 71280

aaacagattt agttgatttt tgaccatgga gttggggtat ggggacaagg ttggcagcca 71340

tttttatgtt agtgtagtcc aacatattac attccagcca gcatctgtgg tattctattc 71400

tagtcattat agacttttag aatcatgaaa ctgtgaagtt gcaaataatc ttataggtca 71460

cccagtgcag cctgctatta aggagtatcc cctacaacag tgcttctcaa aattcaacat 71520

gctcacgaat cacctgggaa ccgtgttaaa gatacagatt ctgattcagt aggcctggga 71580

tggggcctga gatcctgcat tcttaataat tgcttccagg ttgctgctgg tccagggacc 71640

acactttgag taacaatgct cattcctgac agataatgct gccttgggga ttaatctagt 71700

atagactctt tgtattttga aaattccagc tgtcatatac acagataagt aatttttgga 71760

tcaaagccaa actaaaaaca tatccaaaac tcctctgctt ttaaaaaatg taaaccaaaa 71820

tgacattgca attaaaggaa actttaaaaa atcaaatgtg tatgtggagt ctgaaggacc 71880

ttatattaaa atatataaat aaaaaagaat atgcacacac atatgtgaat ccacatatgt 71940

atgtactcac atacacatat actagagatg tatgtgttca tctgtctaca tagataggta 72000

gagggataga tcatgcatcc atattaaata atgcaattgt gaagacctca agtcatataa 72060

tattttgtgg tactgactca gctttaacat tttcctttat attaattttt attagaaaat 72120

aaaaaattga aaagtataaa tttggacttt cctgtagata caaggatgaa tctaaatgta 72180

atgtcatcag cctttttaca ttaaagaaac tgatagccag tttcaagcac tggctgaagt 72240

gaattgagtt gaagtgggcc atgccactag ctctaacttt aaagctggag ttactgattt 72300

taatatttaa atgtaggtac tgcacatatt tatcacctta attcttattt aatatttgaa 72360

agttactgct ttggagaaag tcagctaagc aatggtgatt atagtgtatc aaaatccatt 72420

gtaagtgatt attttaataa ctctgaacca acagggacat gtagacattc agacaaaaaa 72480

caacctaatt ataatgacag caccaaagtg aagctgtgac agttacaaaa ccaatttaaa 72540

agtttcctga cagccatggc taaaagcttt gtttggatct ttcagtaata acagtttact 72600

taagacataa atccacagct aaggaaaata gaaaaagaaa gttgtggctt ttctgccatg 72660

cagatgaaat ggaaatcagc agaaatattt ctaatagaac attatgctga aaggagtata 72720

gaaaaggaat aatagtaatg gaatttaact ctcacaatgc aagaaagact aaataacaat 72780 aaaaatttta atgagcatag aaatgtgcca gatgatttgc ctatctgtgt tatgaaggtc 72840

agtcatttct gcaagtaaat tgatgtgcat ataaacatgt tgattggcag ctctggaatt 72900

tgatttttgt attacctgta aaataaagtc aaagggaaac ggatatgcat caatagccac 72960

ggtgacagga agccttaata aaagtctact ttcccgatgt cactggcatg tctcaaatta 73020

ttagatcctt atcaatgtta cttacctcac tagactatgg gctgagtctg atatgcaaca 73080

taattaattc ttcaacctcg ggtaattaat ttctcattcc cgaacagaaa aatgcttgat 73140

tcacaaaatt gcattccttt tttctgtaaa cctttgtaga atatttgtaa gtgtaatcac 73200

tgttctaagc cctttacata tattgactca tttaaaatta ataacaagtc tttggagtag 73260

gtactatttt ctccacttna tcaatgagaa aacagaggtt taagtaactc ccctggggcc 73320

acagtgccag taagtggtgg agttacggct ccaggcagtt tggcccagag cacctgatgt 73380

aggataagga agaccccatg aaacacatac atagagtcat ggagacattg gcagtgctca 73440

gaagaggcag ttcttgttct cttaccccat tttagtaata acagtttgct taagacataa 73500

atccacagct aaggaaaata gctctttgag ctgcaaggac taaaggattc tcaggtcaga 73560

ccagacagta agggcatgtg ttaataaata cccagaagaa atgaaacctc agaaagagga 73620

tgtgtggtgg cgctagtccc tgggcagatg tggctcagtg gtgtgttgtt tgtgctcagt 73680

tcctttccat aggtctgctt ctctctcctc tttctgctgc ccagggcctt cagcctccac 73740

tctctccatt taagtattct ctaccccaca gtgggcagcg ggccccagaa tggactttaa 73800

atagactagg gtaaaatgac ttagtctctc tgaactgtgt gaaatgtgag taatatatct 73860

caccatatag agctatggtg gatattaaag tagataaaac aaataacatt tgaataatgt 73920

ggattatgac caagtcggat ttatccctgg gatacaagga tgattcaaca tatgcaaata 73980

aatgaatgtg atacatccta tcaacagaat gaaggataaa agctatatga tcatttcaat 74040

tcatgctaaa aaagcatttg gcaaaattca acatctcttc atggtaaaac ccctcagcaa 74100

aactggggat agaaggaaca tacatcaaca taataaaagc tatatatgac agacccacag 74160

ctagtatcat actgaatggg gaaaatctga aaagcctttc ctctaatacc tgcaacatga 74220

caaggatgcc cactatcacc tctgttattc aacatagtac tggaagtccc agctagaaaa 74280

ataggacaag agaaagatta aagggcatcc aaattggaac ggaagaagtc aaattatcct 74340

tgtttgcaga tgatatgatc ttatatttgg aaaaacctaa agactccaca aggaaactat 74400

aagaactgat gaacaaacct agtaatgttg caggatacaa aatcaacata caaaaattag 74460

tagcatttct ataagccaac agcaaacagt gtgaaaaaga aatcaagaaa gtaatcccat 74520

ttacggtagc tacaaataaa ataaaattcc ttgcaattaa ccaaaaaatt gaaggatctc 74580

cataatgaac actatcaaac agtgatgaaa gaaatcgaag aggacaccaa aaatggaaaa 74640

atattccata ttcatggatt ggaagaatca atattgttaa aatgttcata ctacccaaag 74700

cagtctacag atgcaattga atccttatca aaataccaat gatattcttc acagaaatag 74760

aaaaaaattc taaaattata tgaaaccaca gaagagccaa atctatccta aacaaattga 74820 acaaagctgg gggaatcaca ttacctgact tcaatttata attcaaagct acagtaatca 74880

aaacgccgtg gtactgccat aaaaacagac acatagacca acgaaacaga atagagaacc 74940

cagaaggaaa tccacacacc tacagtgaac tcatttttaa caaaggtgcc aagaacatac 75000

actggggaaa agagagttcc ttcaataaat ggtcctggac atccatatgc agaagaatga 75060

aactaacccc tatctctgta aggtacaaaa atcacatcaa aatggattaa agacttaaag 75120

ctatgacctc aaactatgaa attactacaa gaaaacactg gggaaaatct ccaggacatt 75180

ggactgggca aagattttat gagcaatacc cccataagca ccggcaacca aagcaaaaat 75240

ggacaaatgg gattacatca agttacaaag cttctgcaca gcaaaagata caattaacaa 75300

agtgaagaga caacacacag aattagagaa aatatttgca aactacccct ctgacaaagg 75360

attaataacc agaatatata aggacctgaa acaactctat aggaaaaatc taataatctg 75420

atcaaaaaat gggcaaaaga tttgaataga catttctcaa aagaagacat acaaatgaca 75480

caaacaggca tatgaaaagg tgttcaacat cattgatcat cagagaaatg caaatgaaaa 75540

ctacaatgag acatcatatc accccagtta aaatggttta tattcaaaag tcaggcaata 75600

ataaatgctg gtgaggatgc gaagaaaagg gaacccttct acactgttgg tgggaaggta 75660

aattagtaca atcactatgg agaacagttt gaacattcct caaaaaacta aaaattgagc 75720

tgctgtaatt tagcaatccc actgctgggc atatacccaa aagaaaggaa ataagtatat 75780

ctaagagaaa tcttcactcc catgtttgct gcagcactgt ttgcaataag atttggaagc 75840

aacctaagtg tccatcaaca gacgaataga taaagaaaat acagtacata tgcgtaatgg 75900

agcactattc aaccatagta aagaacaaga tccagtcatt tgtaacaaca tggatggaac 75960

tggagatcat tatgttaagc gaaataagcc aggcacagaa agacaaacat cacatgttct 76020

cacttatttg tgggatcaca aaatcaagat aatcaaactc atggacatag agagtagaga 76080

ttggttgcca gaggctgaga tgggtactgg ggggacgggg ggaggcaggg atggttaatg 76140

ggaacaaaaa aaantagaaa gaatagataa gacctactat ttgctagcat aatacagtga 76200

ctgcagtcaa taataataac tgtacatttt aaaataactt aaagagtgta acataatcat 76260

ttgtaactca aaggataccc cattctccat gatatgctta ttccccattg catgcctgta 76320

tcaaaacatc ttgtgtactc cataaatata tatnngtgta tatatatata nntttatgtg 76380

tgtgtgtata tatatatata tatatannca cttactatgt acccacaaat atttttaatg 76440

tgaataatat ggcctggaac atagtaaccg tttaatacat gataaccttt ataataatac 76500

tcattatggc caggcatggt ggctcatgcc tgtaatccca acactttggg aggccgaggc 76560

gggcggatca tgaggtcaga agatcgagac catcctggct aacgtggtga aaccctgtct 76620

ctactaaaaa tacaaaaaaa ttagccagac gtggtggcgg gcgcctgtag tcccagctac 76680

tcgggaggct gaagccggag aatgacgtga atccaggaag cagagcttgc agtgagccaa 76740

gatcacgcca ctgcactcta gcctgggcga gagagcgaga ctctttctca aaataataat 76800

aataataata aataatactc attatttttt aatgtttaat tatctttatc ttaatgtagg 76860 cagtagactt ggaccagact tgggttccag tattgagtcc atcttcaatc cctgtttgac 76920

agtgggaaag ttacttaacc tctttgagcc tcagtttcct tattagaata ggttcctagg 76980

gctgttatga gttaatttat gtaaagcacc tagaaaaact ccaggtagag agcaagtatg 77040

atgtacatgt cacttactag aagttgtagt aaaattttca acaaatattt atggnaagta 77100

ataagtaaaa aataaatctc aggtttctct tacttttcca gttgtataat aatcatctct 77160

aaaatatttt catctgcagt attctcccta aacacagaat aaagaaatac ttgttttgtg 77220

ttactagctt tttaaaagcc atttatccat acagtcagaa ttaaggataa atactaccca 77280

accaaccaat taaaaattaa gatgcaagcc atacacaaat gatatatgtt ttccattttt 77340

agtcaaaagt ctctgaaacc atcttcctag tcactgttta gctgcctgaa agaaaggttc 77400

ccttgcctgc ccagttcagc agtgtgggac atggatggct ttatcttgga agccctgttt 77460

tgcatagcag cacatcacaa aggtgctgtc agccaaggac taagtcaatt acacttaatc 77520

atgaggttta tttctttgtt aagagtctgt ctgtttcgcc ccctcagaaa atgatgcggt 77580

catccgtttt ccacatgttc atcctgagca tggtgaccgt ggacgtgatc gtggcggcta 77640

gcaactacta caaaggagaa aacttcagga ggcagtacga cgagttctac ctggcggagg 77700

taagcagctc tccaccgtgc agatcgctgg gcaaggagca ggtgaaccgc ggacagtgcc 77760

tcttagagat gcaaaagtga agtgtagcct gttcattact tgactttcac caaagctgaa 77820

tcattgtctg cttcttaagt ggttcctctg gtcacagtgt ggcctgattc atttcctgtc 77880

accgcgggcc tgagtcatct tcctcacagt atgtcacgtt gccagggtaa ccctttttccc 77940

ctgggtctta ctctagaatc ttgctgtttt cctatccacc tcttattatc tctccctttc 78000

tcagcgtaca ttagtttgtt tcctactaag tacaagcagg cacagattca gaagatcata 78060

aataagctgg tcagctcagc aatacgtaga agagggctgg aaggacatat agagagatgg 78120

taatgataaa ttactcctct agaaacgggt gttacaacct acattaaaga caccaattaa 78180

aggtaggttt catgattaaa taagcatctc cctagcactc tacatatcat attacagttt 78240

tggaatagta cttcctattg ataaggtaat tttgttaata cctgattctt aaaatatttt 78300

aatttttatt tctttcaata agcattttgc atttatatta aagcatttta ataattttcc 78360

aagatctttg ttttttttaat gtaaaaactt acctaaaata tatctcagaa ttcaatggtt 78420

acatatttac aaaaaaattg attgacaacg tagaaaccca ttcattgtgt gagtgaggta 78480

tacatacaca gttgtacaac gcttcaataa atccaaccta tgaaaaagta aactgaaaag 78540

acagaaaaaa gtcacaattt tattattttt accaaagcat ttttcacata aaaaaaatca 78600

atttgccata taattaagtg ttattacatt cctagaattc tctcagtaca tttaaaaatt 78660

gatttggtta ataaaaattt agtctaccca cgcacaagct tttctgaaaa agaattgatt 78720

gcatgtacct aacgagcata gacatctacg tggatatcaa gaaactttta agtagtcgct 78780

tctgaattct cctcttcatg gaagtatgtc tcctgaaaca aaataccgtt aatacaaaac 78840

ttttactggg ctacagtctc tgtgatgcaa catttgccca atggccacct ggtatatgga 78900 gcttacaatc tagcaaaaaa aaaaaaaaaa aanncaataa acaataagag ttttgtgcta 78960

tggagaacta atgacagggg ctgatgaaga aattatatcc agcccagcct ctgaaaaata 79020

ggtataattt tgctgaaaat ccagcacacg atgaatggat tgtcaataaa agcagagaga 79080

aggtcgggta tggtggctta cgcccgtaat cccagaactt tgggaggcca aggcgggtgg 79140

atcacgaggt caggagttcg agacaagcct ggccaacatg gtgaaacccc atctatacta 79200

aaaatacaaa aaattagctg ggtgtagtgt cgggcacctg taatcccagc tactctggag 79260

gctgaggcag gagaatcact tgaacctggg aggcggaggt tgcagtgagc caagagcgca 79320

ccactgaatt ccagccaggg tgacagagtg agactccatc tcnaaaaaaa aaaaaaaaag 79380

cagaaagaag cctagactga tactaatatt tgcaggtaaa tttctaacac agggaaagaa 79440

caatgaagga agcctcatca agttatttct gtttttccca ggaagtagga agcaaggtct 79500

gcatccagcc agatgatgag acagtgaagt gggaagttgg caaagaaagg ggagagggaa 79560

gaggtgcctg gggttaggca aaaaattaac aatacagcaa atgcaacatt aaagatccat 79620

ttaggactat ggatttacag tgataccaat tcagatagct ctgcgacttt ttagccctgt 79680

ataagaagag ggaaaagcaa catttaaatc aatccacctt tagggatttt tttcaggtta 79740

ggggaactga gaaacaaggg gaaaagatgg attgctggca aaggagttga aatgatgacc 79800

ctgtggtttc tgttggatgt ggaaatggga ggagaatgga ggagtgaagg tgaaggatca 79860

atggtctcta tatggttaaa ggacgggtac caccaaagtg atggaataat agaaccagaa 79920

gaggttatga tcagagagtg gagtccagga ttttacaatt ccagggtaat tctaagtaat 79980

gaaatataaa atctagggtg tcacgtgaat gtgggtggct taaacatatt gatggagtca 80040

aggcagaaat ggaaacgact gttgaagcta aggcgctggc tgagtcttca catggacatt 80100

gaagtcactc agagtgaggc cagttcttgg agagaagagt ggagtcaggg agcaaagtca 80160

cggggaatga gcgagtgact gatgaacaga ttgacagaag gacagaagac agctacaggg 80220

agagtgggtg acagcatggt ttaacctcat ggaatacgtt gcaaaggagt atcaaggagt 80280

aatggatggg cacctcttcc tcgtctacat gttctactag aaagggctgg aaagggctgc 80340

aggggtaact catgtatttg agagagaatc agttgaggca cagacagaaa gtagattctg 80400

ggaacagaga ttgataatgt agaggacatt atttaacgta ggagagactc aaagattagt 80460

tcaaggctct gaggagggag aagtgggaag atcgatgaga cagaagaatg atacaagagt 80520

aatagcatca gtaggtaaag atttgctggc aaaaatgcca ggaaacaaaa gcagaaatga 80580

ccacaagaga ttacagttat gyttatctct ctgggaggag gaatctgcat ctgctcctga 80640

gggaatcttg acattagaga ttacagtaac attgcttttg tgatctagca cagggtgggc 80700

aggtggtcac actaaatcag tcctggagca taagtgaata ggccctccac actggcaaca 80760

gcagggaagg gaattcttta gtaaatttta cctaactttt tcatgtacag aattgatcct 80820

ttcctaatgt gtttatgtat ggtagatgat ttatcctcat accttaagaa attccaacaa 80880

aatgtcaaat aattgaatgt tgtagagttg actggttact atttaagtgt attgtagagt 80940 ccaggggtag ttggtaaatc cttgtctttt ttcttttta tataatatta aaaataaaaa 81000

caacctcaac aacagacaga caggattgga tggtggaatt cagtatatat tatgtgtcca 81060

ataattgttg tcccctgaca ataataataa gaaagggtgt gggaggaaac tttgagaggt 81120

ggtggatatg tcaatggcat tgatggtaat ggtttcacag gggtacactt attcccaagt 81180

tcatcacatt ggtacattac atgaatgcag cttttaacat gtcattcatt cttttttttt 81240

tttttttttt ttttgagaca gagtctcgct ctgtcgccca ggctggagtg cagtggcggg 81300

atctcggctc actgcaagct ccgcctcccg ggttcacgcc attctcctgc ctcagcctcc 81360

caagtagctg ggactacagg cgcccgccac tacgcccggc taattttttg tatttttagt 81420

agagacgggg tttcaccgtt ttagccggga tggtctcgat ctcctgacct cgtgatccgc 81480

ccgcctcggc ctcccaaagt gctgggatta caggcgtgag ccaccgcgcc cggccacatg 81540

tcattcattc ttaacaaagc ggcaaatcag ttaaccagct aaaaactaac ttgaaatcct 81600

ggaaaaataa ataaataaat aaactaggta gtggtgattt tgttgactta aacaaccaga 81660

tagaccccat cagtgtactc ttgaactact taagcattct gaatttgttc cactctaggt 81720

tttggggttt ttttgttgtt gtttttgtct cctttttttt tttttttttt ttttgagaca 81780

gtcccactct gttgctcagg ctagagtgca gtggcaccat ctcggctccc tgcaacctct 81840

gcctcccaag ttccagtggt tctcttgcct tagcatccca agaagctggg ataacaggca 81900

cacccagcta atttttgtat ttctagtaga gatggggttt cgccatgtta gccaggctgg 81960

tctcaaactc ctgggctcga gcgatccatc catcttgacc tcccagagtg ctggaattac 82020

aggcgtgagc catcactcct ggccgttcca ctctgttttt gattttgtgt gtttgacttt 82080

ccccaagtat gtttcgtttt tattcagatt acattgccat ttctctttgt atcagaaact 82140

tatacatgag acaaagtgat ttgtttcaaa tgaatgtgaa aaagtcactt taattaaact 82200

tactttttca taatatattc tatttttata aagtttataa agatgtcttg agtttaggcc 82260

agggtttggc aaactctttc ttgcagaccc agatagtaaa cattttgcct tttgtagacc 82320

accctctctc tgtctcaact actcaactct gaaattgtac aaaagtgagt gtgtctcaga 82380

atggttgttt tccaataaaa ctttatttac aaaaatggat cttggtttaa gctatagaaa 82440

cagagagcaa tgtcaaataa ttgaatattg gattatattc aattatttaa aatacaaata 82500

aataataaca ttttgtgtta aagtcaaata ttttaatact gacaacaata aagctatgat 82560

ttttgggggg cccattatgt tctaagtgct ttaaaagaat tcattgtctc atttaattct 82620

caaaatagtt ataataacag caattcaagt ttatagacta taatgatgca gtatgtaagg 82680

caatagcatt atgtaatatt ttgaaaagtt tattttaaaa atggatcaca acaggtcatg 82740

agcttttaaa gaagttattt ttttaagttcc cacatattcc agaataaatt atttagaaaa 82800

tctaaaataa agcactttac ccattaagag ccaactcaca gtctaaaaat caccctttaa 82860

gaataaatta agatacttag cttctgttgt gaagctgggg acttgttttg ttactttgga 82920

ttaattaacc tacatttaga attagttaag aaacaggaaa cagtttacct cttttgtagc 82980

ctatgcataa acagaaaaag gtaaagttta ttttcctctg taagaatttt aacttttcat 83040
tgtgatctaa gagaatatat tattattgca cattaaatca tatgttataa tgaacatgat 83100
tacattttaa aatatatatt attttcttaa gaaataaaat ataagcaata tttttcttat 83160
tattatatga atgtatttgt atccattaga ctgtacattt aacaaaattt atcagcacag 83220
cacaggagat taaaagcttt atcttctttt atctcaattc ttcttttgag aaactcacaa 83280
acagatggga gagatacaac aaggaaggca acggacctac agagaggccc aaagttagct 83340
ctgtgtcacc ccctagcatt atccccagct atgttctctg atttcctcta gacctatctg 83400
aatggggagc tcaatcctct tccacctcct ggtgttaagt attaatgagg acagtcatgt 83460
tttccaatga agccagaggg gagaaagaag tcctaggact ctgtgctaca taccccacag 83520
taaattaact agaactcctc ttaccctcaa ctctgtctat ccctcctacc cagcctactc 83580
ccatatctcc ccaacttgtc atcagccagg caggctcttc ccctgctact catgcacagc 83640
aattccactc caccccccaga cttgcatctt cacttgcttt ttgttcagca aaacaaaaag 83700
ctctgctatt aagaggatct agattcctca gaaaagaaag tatcatctat tacttaagcc 83760
ctcttagtcc cccacaggaa atatttattt taaatacatt cagagtttct ctgcaaagat 83820
ttcacatgga tataaaattt ttcatctcat ttcaattata aagtcagaga acagagagga 83880
agaagggcct agtatccttt cttacatggc agttgtgctt ggtgatgggg aatcagaggt 83940
aaataaagta aataagatat ggccatcgat gtgtgactct aaggacctct ttctactctt 84000
agaagctata gttctgtgag attatctggt gtaatcttga gtgaggtgta tacctaaaat 84060
ttgtggcaat aacatgattt atcaaagaaa ttttagtgac tctacaagaa gaaattaaga 84120
acaatttcca agtcagtgtt tgtctcaaca catcatatcc atatgagcag ttttagttat 84180
taaatccgtg tttccaatag atctaaaaat cattctctat tgtggtagac taaatatgta 84240
ttaagttgtc ttgtcggagg gaatgagtca taaactgcca aggtgctagt tcgctgagag 84300
aagttacagc ttgagagaat gttgtttcta aaacagatta ttgatgaagc aaaacgtgct 84360
gagactaaat taattttaca caaataagat ttgggcagga ttagacttat ttatgtcaga 84420
tacctatgat gtcagaagca accattagcc tttgtgtatt aatatgatta ttttgcagga 84480
ttctttttct atggttattt ttccttcttt ttgaaatttc tatcctttga aaatctgtga 84540
ctttctttta tttccttctt aattagaaag tatttcataa attcaacaca tagagaaaat 84600
attataacca atatttgtgt gtccgtcacc cagattttta aaaaattaat agtttttcat 84660
gtttgggtca ggtttttcaa tcagaaataa aacatcatta gtagagtaga agatccctgt 84720
gtaccacttc ttgatctcct ccttctccct aatccacaga gaactatatt catcatatca 84780
atgcatgctt tataatagta tgaagtatta aggtgccact ttttcaatcc tcattatatt 84840
tgggaaattg attcatgtcg attgaagcat gaatttattt gttctgcatg taagggcaat 84900
gggtatttca attgcttcca actttttgct gttgaaaaca atacttcagt aaacattgtc 84960
attagaattt ctgtgccaca tttgggagag ctctttctaa atatatacct agacttaaaa 85020

ttgtgaattt gtcagatatg ctcactttca actttgctaa atattgttaa attgaacttc 85080

aaagtgcttg tgtcaattta cattcctacc aggatatatg agcagctcta cttcatacta 85140

cttttttttg ttcaggattt ttaagtttct tacaatctca tataatcatt atagtcttaa 85200

tatacatttt cttggttacc atagagattt gaccttttttt tcgtattctg tcagccctga 85260

atttcctctt ctataatttt gcctgtttaa tctgtaccct aattgatttt tcagttatct 85320

agtagtatgc agtgcagaca tctcccagcc ttcagcttgt cttttaactt ttatttgctg 85380

tcttttcttc aatagaagtt attaatttta acactgtaag atatagtcat ctttaatttt 85440

tgattcagtg ctttttgtgt cttatctata ccagagtcat taaagacatt ctcttatagt 85500

aaaggctaga gttttcaagt tgtgcttttt caattttaga ctccaataca tttggaatga 85560

gataaggact tgatttttttc cccatatata aacaattgtc cttcgaagat tcattgaaaa 85620

ttcattcttt cctcactcct gtagtgcaat ttctgtcata tatcaggttt acacatgtgc 85680

acaattgtgt ttctggacac ctcagtctgt gtctatacat caaagcttta gttgataaca 85740

tgttgtcata agtgctatcc tttataataa accttgagat ctattagggc aagctactga 85800

tgttgatatt cttcttcaga tttgttttat ctgttcttgg tcttgctatt cttgctctac 85860

attatgaaaa tttaattcat tttagtaatg agattttaat tgatgttaag aacatctcta 85920

ttcagttatt ctaaccttcc ttagtgtttt ccaacaatgc tttataattt tctccaaagg 85980

tcttccacat tttttgttac attcatgcct acatatccta cagtttcttt tgtgaatgaa 86040

ataatttcaa ttacattttt attgatgtag gaataatatt aattttttat attgacctaa 86100

ttcatagcca gcaactttac tacaattggt tgtttgtgat catccataga cttgattgga 86160

atttctatgt agataattgt atcttctgaa attaaggttt tgtctttcca cgtctaatta 86220

tcatatctct gattttttttc ttcttttatt atccgggtgt gggtctctaa taaaatgttg 86280

aatagaggca gtgatagtgg acatgctggg tgtttttcta tattttgata cgaatgctta 86340

tgatttcatc attaatgtca tttgtaggtt tttatagatg cctttcatca tgttaataaa 86400

gctcacttct atttctaggt tgccaagagc atttctcatt atgagttgag tattcaatct 86460

tcttctacct atactgagct gactacacag actttccttt ttaatttctt agtgcagtat 86520

attagaataa tagatatttt gatactgaaa ctattttaga ttcctgaaac acttgatttt 86580

atagcattac ttttaataaa caaatccaat aagctgacat tatattagga attttcaatc 86640

ataaggatag tagttctgtg attttatttt cttttacatt gctagtctga ttttgttagg 86700

gaagttacac ttcctttcac aaatggaata gacatatttt tctcttttttc tcttctgtga 86760

aagaatgaat aaagtccctc tggaacacca tctagggcta ttgtttgcag tgggatccct 86820

ggtggagaaa gggagaagag agatcttgat taatgactta ttttatggtt tgattattat 86880

ggtttgtcag tttcatggtt tattgcttgc tatttataga gtcattaatt ttaactaatg 86940

tctcacattt ccacagtttt ctctcaggat ttttataatc tctattgaaa ctgaatgtat 87000

aatccttttt gctgccaata ttgtttattt ataactttct tttttataga agtaaatgac 87060 ttctgatgct aactacctgg attttagacc aaacttcaca gcacagaggc acagtcctct 87120

acaagattac ccttactcga gatactaact gtaagctcag ggggttccag ggcctccctg 87180

acatatgaca tatcatatgg ctaatatggt tgatatgaca tatgacatat catatgacat 87240

atcaacaggc taaaaactcg gggggtttcca ctagccttct caagtttgat aatttgctag 87300

aatgactcac agaaatcagg aaagtgctat acttctgatt acatttttat taaagcaaaa 87360

aggttgcaaa tcagagccag ccaaaagaag agatgcagag agaaaggtct ggagagtcct 87420

acatgtgaaa cttgctgttt cctcttttca tggagtgagg acatatcacc ctctcagcac 87480

atggatgtga aacaatattc agagcgctga gaaccaggga agctcacctg agctttgtgc 87540

tgaggcttct gattgggatt tccttaggta ggcatgatcg ctggaatcac tggccatgca 87600

cacaaggttg gagtttctga catggccagt gctcatgaat gatctcattg ccataaacta 87660

ccagggtcac cataaataac aaagatacct atcaattggc aaattccaag gatgtagagg 87720

ctacttccca ggaacttgag gcaaaagcca gtaaaattct ttatttttca ttcaaagaac 87780

cagcttttag tttggttgac actaacagtt tctaatgtac ctgaaattat tggttcgttc 87840

ccaatatttt gaactattaa gtccctttat ttctttgttt tttaaatttc tcctctcacc 87900

tttattattt tctcttttc atttatctct ttggtcatct taaacttctt tgaaaattct 87960

tactggggtt ctctgggtgg aatataaatc attctattat aaagatacat gcatgcatat 88020

gttcatcgca gcactgttca caatagcaaa gacatggaat caacccaaat gtccatcagg 88080

gatagactgg ataaagaaaa tgtggtacgt atacactatg gaatactagg cagcgataaa 88140

aaggaacgag atcacatcct ttgcagggac atggatagat cgggaagcca tcatcctcag 88200

caaactaaca caggaacaga aaaccgaaca ctgcatgttc tcacttataa gtgggagctg 88260

gataatgaga acactgaata gctgaacatg gacacaggga gaacaacaac acacactggg 88320

gcctatcagg aggtgggggg aggggtagca tcaggataaa tagctaatgc atgtggggct 88380

taatacctag gtcatgggtt tataggtgca gcaaaccgcc atggcacaca tttacttatg 88440

taacaaacct gcatatcctg cacatatatc ctggaacttt aaatttaatt gaatttaatt 88500

taaaaaaata aaataatgaa aacaactaaa tgatggtgaa aaaaaagttt gacaaaatta 88560

attttactga gagtaaattc atagtttgaa ttttagaatt ttgttttctt tctcattaga 88620

tttatttaaa atctttgcat gctgcctcat ttggcttggg agattttctt atttctctct 88680

ccctccctct atgatcactt ctactttctt agggatattg cagttgcctt aaccttgacc 88740

ctggacccta ccatccagag ccaggttttt aagtgtcatt tggagaatgc tagccccttg 88800

atatgggtga tatcacagtt cttgtcacta gtctatattc cataaaatat cttgttctga 88860

tttttcagat aagctgataa tttctcctcc tgcctctgga ggcctgcagc tgcctataag 88920

gtcttagcct tgggtagttg gtctctgtcc tttccttagc agagaaaccc attctagtcc 88980

ctgtcttttc actatgaagg ttctcatctt taccttatct aaggtgcctt agtctcctca 89040

ccctatagga agcaattgag tccaacattg cctgattctt gacctagcat atgtttctgt 89100

tgcatttctg gtcaacaaag aggttaacct cgtttagaat tccttcctgc ctctttggtt 89160
tcctgttttt tcctttgatc acagtatgtt tgaaacagaa ggggtttcat aaatcatgaa 89220
ctcattttgc cgtcttgcag ggagtcctct tttctagaaa agtgatttct gcattcctct 89280
gaagacattt ggttttatga tatttgttaa aatgaagagt catttatttt ggccccaaca 89340
aaatctgaca cccacgtttg agttcattaa tcataaaaat tggccaaatg aagaaaataa 89400
tcaatcttgc catagcagca gtctttgaca ttattggtga gtttaatttc attttctggg 89460
aaacttcaag taaataggtt tcttctgaca nntacttttt ggggggtgtt tctaaaagat 89520
ctggacatta tcttgttagg taaaaatata atacaataat gacaatctgg acttctcaac 89580
catccatctc aaaatttgct ctttataaca tactgatata cactataagc taagagtagg 89640
tgatttacag acagctccag actgatacag gtaatcagta atcacttttt gttgttattg 89700
aatctgaatt tcatgtaaga gaatgggatt gcctgaactg attttcttgt gagctttgtc 89760
ctgtaatctt cctttgtgat aaatatgcat ttttttctca tgatatccat catgattctt 89820
tgaaattaga aaacagagac tgcaacacaa atattagata acacatttca aatatctttc 89880
cactcatggg caatgtccat ctctaccatt ttatctcata ggcctgtgtt tacgagatta 89940
gttaaactct gctgttgtta ctggtagagg gtctcgactg caagttgtcc aggttcttga 90000
cgttttgaat aaagaactgg acaaaacgca cagcaaagca aggaaagaaa gaagcaagga 90060
aagcagagat ttattgaaag tgaaagtaca cttcacagtg tgggagtgag cccaagcagc 90120
agctcaaggg ccccagatac agaatcttct caggtccaaa taccccctag aggtttccca 90180
ttggccacct acttggtgtt caccccatgt aaatgaagtg gttgcccgca atcagaggct 90240
gaggtgaagt tacgaagtta cattcctagg caagtgtctg gttgggaaaa gcaaccgatc 90300
agaggtactt tcaattatcc atctgccgca cagaaaaggt gggggtttgc aaagggagta 90360
gcccctggtc cttttgctac tcaggcatgg aaagttaggg tttttctttc aatttagttc 90420
taagaagtca gtgtgaaatg gacttaggtt ccctgcctcc agaccctatt ctcctgcctc 90480
attatggcga catcaaacca ggtatatcgg ttgagcatgc attctcaaca gggagtaaaa 90540
tcacctccaa aggggtgaaa atatgttctt agttgggggg tagaaaattc tctcttttat 90600
aatgcagatt gccttgtaga ccaaataacc aaagtagagt atgtaaaaag atatataata 90660
tggtagaggt atcaaaactc catgaggtgg tgattaagaa caaaatctct aaaagacttc 90720
ttagaagagg tgatttaaaa aaagattggg aaacactgct ttggggtatt tgtaagaaga 90780
aataaatatt tctcttcttt tctatgcatt gtttctaagt tttcttaaag taatgcaaaa 90840
aaaaagaaaa gtttattata ctcttacaaa ttttgtctta actcttaaaa agttaagttt 90900
taacttgtag gtactcacaa aggtggaatg cactgcagaa tatgaaacag aagaagggtc 90960
ccaatgagcc cttcttggtg caggggctgc agactgagga gggccaaagc agtggagggc 91020
tgccaggcta agagtgacaa atggaagcca ccgttccact tttccacagg ctgcctgctg 91080
aatgcctggg atttgcagaa agccttgaga ttggacacct gcacaggtga ctttggtggg 91140

agccctccag gcaactaggt tctgcagaga ccccagttgc cttctccctt ttagatggct 91200

cccagcatgc ccctggcaaa ttccaccctg tccagctacc atcaaggttg agttcactgc 91260

aataatctgg agctgcacat tgagaagtgg ttactagagc ttgcctccac tgcttctatt 91320

ttggttcttc tgaactatta tagtcctctg aatttaagtc ctgtaccaaa ccccataacc 91380

tacactgaaa tacttagtga tcatttaaac taatttgatt ttcctaatca ccttctttaa 91440

atcctgacac agttcttctc tatggcagga taacccatac cttcaagtta agttaggcat 91500

atttttatgt gcattgattt tccataaaat gttatcactt tggcatgaat atttcagaga 91560

gctagagtgt tggcttgtaa aagttaacaa attatgttca aagcaagcca tttgattcca 91620

agagaatctg gaatcaggaa ttctagggga aaatcagtga gactatagat atttaaatat 91680

aatcacataa caatttgcta aatcttttag tgccttttag tgtctttcaa aagaatatca 91740

ttttcccaca gagaaaattg taatgcaaac agatccagct gcatatttta ctataaatta 91800

tgctcattta aacagttcat tccatctgtc cttttgataa tatatgcagt tagtaaggtg 91860

attaattgta tgtatacata attagacaat ctattggaat atctttcagg cattcattag 91920

tagttttctt ttaaaaaaac actttaaaat ggatcaccac tctttttttt tggaatgaaa 91980

atctcagtct gaaaatacag gtgaagggag agatgatgta atgaatgcct attactgctt 92040

cctctgctaa tacagcaacc aggtcaggca tgtggatggg accttcatga cacatcaatt 92100

tcagtttctc ctgagattcc ctagtatcca taggagagaa aacaggttca agtcttcccc 92160

atatcattgc aaggccagaa ctaagaagcc tgtgctagag aggaagtacc tgtcgcagga 92220

cttgcagcaa agtctgagtc ctctggtccc ctgaggccac tgctcagggc agagagcctt 92280

tcatttcccg cagaggagct ggacctgcct gtatgctgct tgtggtccag gagcagtcca 92340

cagagtcctg tgggtctcag tgggccagga agggaagctc actagagact gaattactcc 92400

aggtggagaa gggggtcaat aaatgactcc actcatttct gggaaatcct gtcaccaaaa 92460

agccttcttc agatgttcag caaaggagct atggcttttc aggtctttcg cctccgtaaa 92520

ctatggtgcc ctctactacc cctcatttgc cactgcttaa gatagcataa aagacaaacg 92580

acaattacat caaaaacatg tatccatggt gaaggaaaga aaagttaact cacacacaaa 92640

gggagactca gaaactgagg atgaaactca gagaatcaag acaaatgtct agtctagaga 92700

tttcccaagc aagtagaatt gatgggcgag agaatccgtt gttgttgttg ttgttgttgt 92760

ctcaagaaaa agtaagcagg gcacaaaata caatttctgt aacttcaact aaaaatgtcc 92820

aaaaaatgag ttaaaaaagg gcatagtcct tgagagctag tttttaatat gggagatgaa 92880

gcaggaaaaa ataactcaaa aagtggtaga gatataataa taagggcaaa gtaaaatagc 92940

ctatccacac agctatgaat aataggagtc cctgaaggaa aaagcgaaca gatagaagaa 93000

aagttgtcat aaacaacgga agaaagctta cataaatgaa aaaaaaaaaa aaacataaat 93060

ctgcaaactg aaagagctca ctaaatttta gtgagattaa atgaaaatag atgcttactc 93120

atgtcctggt gaccttccta ggttgcaagc ataaagaagt aaccttgttc agcccctaaa 93180 caggctgaac ctaaacctga aaaagtaaaa taaatgtatt gtgcaatgga tgactcatct 93240

acaatactgg aaagtggaag gtaatgaatc aatatacata gactactgag agaatagaac 93300

tgcaactaca gaatctgaac acagacaaca tgttcatctg tcagtgttaa aaatgcattt 93360

tatatacaaa aggactaaga gagcataaaa ccctttgtat ttatttgaga aaactacttt 93420

aaagtgtgct aatcacaaga taagtaaatt gcaatctcaa tgactagaca aaagaagaga 93480

gaattgatgg gcattggatc tcaatagata ataataaata tgtcaaatag ggatctttga 93540

aaaaagagaa atgataacat taaagaaact agtaataaat ttgtaaatga tttcagcaaa 93600

accaggactt aggaaagggt gtgatatcaa aggagagtgg gatggtagaa gtgtactatg 93660

aagtttgcag atgtgtgggg tggagggatg agtgagccag caatttagtt gtgaagcaat 93720

aagaaaaaag cattggtttg atgtctaacc attgagaatt aagccaggag attagaaaaa 93780

ccaagtcgag cgtccaaatt aatgaatagg gaaagaaatt agcaaatagt tgaccaagtt 93840

aaaaaaagac aagataaggt taaaaatgaa acagcaatga cataaaatac acaggaagcg 93900

taagaccaag tataactttg gcacattttg tttgagcagg ttgaattttc ttatcaaaaa 93960

ccctagaaaa tctgagtatg tttaaacaca aaataaaaat tgtacactgc ttataggaag 94020

cacatacaaa gtggccctgc aaggtaacta aaaatcattt ctaccaaaca cttaaaatat 94080

taagacgttt taaatgttca gcagagtgtg caaaaaatta aggaaaattc ttgaggccaa 94140

aaattaaatg tgacgtgagt ccagaatacc aggtgggccc caaggcagtg gctgccctga 94200

ggacatttgc aaatccttgt gaccagaact tgcagttttc aggagtgtgt ccaggatgga 94260

agtgactaga ggcctgtgca aagtgaagag gcagaagaga ggccccacat aacatcagga 94320

aggaaaaatc ccatccctca gagggaaagg atgggatcat cattctggac cttgacccct 94380

ggcagatcaa aaatgttatc tctggccggg catggtggct cacgcctgta atcccagcac 94440

tttgggaggc cgaggtgggt ggatcacttg aggtcaggag ttcgagacca acctggccaa 94500

catggtgaaa ccccatctct actaaacata caaaaattag ccaggcatgg tggcatgtgc 94560

ctgtaattcc agctactcgg gaggctgagg caggagaatc acttgaacct gggaggcaga 94620

ggtcccagtg agctaagact gatcatgcca ctgcactcca gactgggtga cagagtaaga 94680

ctctgtctca gaaaaagaaa aaaaaaggtt atcatgttat ctgtactgaa ttgtgtaaga 94740

cagcctgccc actcgtgggg ttggagctca aaatcccatt acgtccagaa tctgagaaac 94800

agcagccaaa agttgattta agggggaccc aaatttctaa gatgacccag gttccctgga 94860

agaaccaatt ctttctggaa aaatatgtac cctaaatgca ggtcttgaaa atctagagac 94920

aatcaatgat ttcctggcaa aatacaaact gccaaattga tggatgagat aagaaacctg 94980

atctgaacag cactggacag ctttgctcca ccatcaccaa aaggcctgtg cccatgtggg 95040

tttcctggtg ctttctgtct acttttaaag gtgtgctaat tcttatttta tttcagaaat 95100

tttcatacaa aagataaaaa ctgtacagaa tgaatgataa agctagcata atcttagcca 95160

aaacctgata atgatataca agtaaaacta taggccaatt tcacataaaa cagtaactgt 95220 gaaaattata aataaaatat attacccaat aaaatccaac atgttagcat tatcatctag 95280

agttttttc caaaaataaa taaatggtta acatcaggag cttttgtgtt atatttcatt 95340

gcctcagtga attaaaggtt aaatcctttg aaggagaata ggaagaaaag gaaagtcttt 95400

aaacatgata aaaattaatc accaagactc atttcaacca ctatatatga acgctgaaga 95460

aactttcatt gaaaacaagg agcaagttat ggattcctat tattaccatt ttgtatttag 95520

caaccattat ataactgtac ctttataact aatgcattca tatgattaac tccaaaataa 95580

atcttaaaaa tgaaataatc ttattttatt tccttatgaa attgtgcaca tttctataaa 95640

aatcgtgaga ctttatgaaa aatattagaa ttaatagaag aatctagtaa agtgggtaca 95700

gaaggtcaat atagagaaat caatagcttt tcttcccact ggaactaact gggtagaaat 95760

agtaaaaaag aaaaaataca tagatagata gagagagaga gagagaaaga aagataggat 95820

atttaggata tttatntnna nnaaacttat taaagaacgt aaaatccagt tcaaataaat 95880

aggaaaagag tgcattcctg gatatgaaga ctagtattaa tcttcttaaa gttctatttt 95940

atatgtaaaa ggtaaagnaa aaaaatcata aagttgtaga catatgtgaa catgtaaatt 96000

ttatgtttgt agcactaaca atatcaaagt tagaaaccag actgttggct agggaaaaaa 96060

tgtttgcagt atatatagca aagattattg tcctcaatac caacaactcc caaaacgcat 96120

aaggaacaca acataacacc atataatagt gtcaaggtta cggacagtaa catttaatta 96180

ctttaaaaag aaattaaagt gattaataaa catataaaca gatgagcaac aataagtttg 96240

attacactga tactggcatg gatttgagac aagataactt tcctgtattg ctgaagggca 96300

aatgatttct tataatattt tgaaagtaat cactttaata ttttctaaat ttaaaacaca 96360

ctaatctgtc tttgccccag caataccaac tttggaaatc tatttataga tatgaaagcc 96420

tcaataaata ctggtattga gggaggatgc ttagtgcagc ctttttattg aagtagcaat 96480

gtgcttaacc accttcctga tgattcctgc aaataccacc catcaatgaa ttagtagaga 96540

aaatgttgag ttacataaac attgaacact atgggaaatt atgcagctga taaataaaac 96600

ataggaccta tatcccttaa tctggaggga tatccacgat tgttagtgac tcagaaaaca 96660

agttgcagag tgatatataa tattcctgca tacccaaatc tttatcattc acagttcacg 96720

taactttagg taataccaag aatgcctatt tgtctttaga atttttttc tttccagagt 96780

aaatacctgt cacattattt ttaacaactt catagaactg tatgtaatat ttatacataa 96840

taagtgtata gaaagtatga acttgacaga ggacaatgac atctttaaaa ttagaaatag 96900

tgctttcata aattgtaaca attgttttta catctcatgt aaaaacatgt aattgcaagc 96960

ctgttnctgt gattgcatat gtgtggatga agaagcagat gatcaaggaa agaatataca 97020

tgacattaac atgaaatacc tgagaaacnt gggcaaagga gttcatncca cgtgtttngt 97080

ggacttaaga tcatttcaaa tttaaaatat attaccttgt tatgtttttt aaggagggaa 97140

atggcaatta aaattaaaaa gcaactagaa tctgcagaaa tggtagattt gagaagtgat 97200

gctaatttcc tcttcacatc taaattaact atctgcagtg cctttaatgt ttcatttctt 97260 aaaacactat tattaaaact gcattttccc cncatgttcc atatatgagt ggaagaaaat 97320

cccttctgat actccacatt agtctagaag aagagaagca aggtctgcta tattcctatc 97380

ttcaataaat catgtagtgt cttctcagag ggacatatta tgactaatcc cagtagtcac 97440

cctcatagtc tttaattaga agttaatctt ctattgcaaa ttttaattcc agctgcaatc 97500

tcctttacaa aatgaaaatt cttgttcata ggacatgtaa aggaaataac atgctgtctt 97560

tatgtgtgtg ttttataata ttgttatttg gaatagtaaa ttcatagaat gcatgtgaga 97620

gattaaaaga tcaaataaat gtgttaatat actaatattg atcattaatt ttattaatat 97680

taaagtaaca tattgattta tttgcataag cacacataga tcaattgctt ggtatgaaat 97740

ctcaatttct cttttgccat gtggcctaaa gcataaagga ataaatgata ttttaacttt 97800

taaattttag tttaattagg ttaataatat gtaagttgtt attttataat atgccatgta 97860

cattaaataa attttatctt aaaaaataca acatacaaaa aaattcagcc ttttgtgtac 97920

atagaaagtc tttaaataat aatggtttta aaataatatc aaatatagat aataaagaaa 97980

taatgccact cacaaacatt caaaatcata acatgttttt ataattgttt aagattgctt 98040

tggtctactg tgttctagtt ggtatgtttc aaatgcattc ttaaagtaaa atgtcattct 98100

tgaccttatg gggcacctta gtgtatatga tagtggcaaa aattcataca ctgacaaata 98160

tgtgaaataa tttagaataa attatagttc tgaatttcaa tttcaatgtt tctttgtttc 98220

tctaggtggc ttttacagta ctttttgatt tggaagcact tctgaagata tggtgtttgg 98280

gatttactgg atatattagc tcatctctcc acaaattcga actactactc gtaattggaa 98340

ctactcttca tgtataccca gatctttatc attcacaatt cacgtacttt caggtaataa 98400

aaataatgcc tatttgtctt tagaaattta tttatttcca gagtaaatac ctgtcacatt 98460

attttaacaa ccacatagaa ttatatataa tatttatata taataactat tatggtaaat 98520

atgaacttga gaaaggcaca atgacatctt acaaattaga aaagtgcatt cacaaattgt 98580

aacaatcgct ttttaaaact cctgaaaatt cacagaaaca atgtagatgt tttacttact 98640

tnttcctaaa ttagcagtat atagatacac atatttttat atatggataa catatatcca 98700

tacatgcgta tnnaantttt aaactacata agatgttggn ncaaacttta aagactagga 98760

aacattaatt taatcccaag tgtttctctt ccttgtcgat aaggctaagt gtttgctcaa 98820

caacggttta taatgaaaat attaaaggtt tataaaactt acctttacag attgtaatta 98880

ctcatacagt attatttata aatgcctgtt acttcatgat ggcacttata attcttattt 98940

aacttgtaat aattggatat cttttttttaa atgtgatata ttggatcaca ttgctttcta 99000

attatgcaat aaacagtaat tattagttac attctttgtt attattgaat atagcaacac 99060

gctcactgtc ctggtaaata cttagtcaaa atgataaaat atgtctacat ggtattctag 99120

tcaaaaacac ctcaaatgta gtcattttcc tgattcaaac ttggtcattt atttttttatt 99180

ttattttatt ttattttata ttattttatt ttttatttta ttttatttta ttattattat 99240

actttaagtt ttagggtaca tgtgcacaat gtgcaggtta gttacatatg tatacatgtg 99300 ccatgctggt gtgctgcacc cactaactcg tcatctagca ttaggtatat ctcctaaagc 99360

tatccctgcc ccctcccccc accccacaac aggccccaga gtgtgatgtt ccccttcctg 99420

tgtccatgtg ttctcattgt tcaattccca cctatgagtg agaatatgcg gtgtttggtt 99480

ttttgttctt gcgatagttt actgagaatg atgatttcca atttcatcca cgtcccctac 99540

aaaggacatg aactcatcat tttttatggc tgcatagtat tccatggtgt atatgcgcca 99600

cattttctta atccagtcta tcattgttgg acatttgggt tggttccaag tctttcctat 99660

tgtgaagagt gccgcaataa acatacgtgt gcatgtgtct ttatagcagc atgatttata 99720

gtcctttggg tatataccca gtaatgggat ggctgggtca aatggtattt ctagttctag 99780

atccctgagg aatcgccaca ctgacttcca caagagttga actagtttac agtcccacca 99840

acagtgtaaa agtgttccta tttctccaca tcctctccag cacctgttgt ttcctgactt 99900

tttaatgatt gccattctaa ctggtgtgag atggtatctc attgtggttt tcatttgcgt 99960

ttctctgatg gccagtgatg gtgagcattt ttccatgtgt tttttggctg cataaatgtc 100020

ttcttttgag aagtgtctgt tcatgtcctt cgcatttata taatttatca tggaaacact 100080

agaatgtgca ttgattgttt tctaaataga tattgttaaa tatttgggaa ataaaatgac 100140

agtagaattt taaagaagta attttgagaa acaggactct aaccccttag gaaaaagaac 100200

atgcattgca actttattga ttttatttat tagtaatacc aaaatattat tttcagaaca 100260

atattcactt taactgccag atcacaggta actagacact tttagtacct gggggaaagg 100320

acatgcaaat ataattcaca aaaagtaaaa agcaaaaccc cctacaaacc caggcaaaaa 100380

gctgagccct tgtaatcaca gctaagcata ttgaaaccac agagggtttt catagttcaa 100440

gcgctcactt ccctcagcat ccgtggatgt aaatctctcc tcatktgttg agttatatcc 100500

tgaagcttrg ctatcataag tgacatggct ggggcaagag atatacatgt ttaaaactct 100560

tgatacaaaa tgtgctttcc aaacattcct gtgattataa gctcttgtgc actgctctct 100620

agattcacaa agttatcatg gccattgagt acctacctta gggcatcatt gcacataatt 100680

atagagaata tcctttatat caaaacaaca gcataccttt tttacaaagc tattaaagta 100740

gccaaattaa gaaaattggt gttgcctatt tcagcaaagg tacagtataa aatatattca 100800

gggccttaaa attgctatag tctctgacat gccatttttc agtaaaaagt tagaaaataa 100860

catttgtgtg caaaatattt attatagaat taatgagaga aaaaagaaaa ctgtttaaat 100920

gacaaacatt tgaggaatag cttaataagt cataacgcga atttaaccat aaaattataa 100980

ttgtgatgag ttttttataa cttagaaata tgactgtttt atttaaacga taaacagcat 101040

aaattatatt atatacataa ctacatatat gtatgtatat atttattata tacatatgta 101100

aaatacctat aaaatggaaa aaaatgacta gaaagaaacc aatttattac aaattattaa 101160

tcatctctag tttcctactc ttgtcatcat gtctgatttt cctactagtc tgttttttta 101220

caaaatatct aatattagcc tgcaaggctg taaccatttt atagatagac agatgataga 101280

tagatagata ggtagatgat agatagatag atagtagata gatagatagt agatagatag 101340

```
tcaaatcaca gtgaagtttg caatttcnag ttnngncntt gattttggat tattctgtaa 101400
nnngtnctgt ttaatttatg ttaatgatta ggtgattttc atttgtatct ctctatacat 101460
ggccatgctt ctctttttct aatgtattaa taatgaataa aaatgatctc agttaatttg 101520
tggctagaaa gtaaatctct tgagaaactt aggtgtcaca agggtttttg caaatttgca 101580
tgtggagaat ccactgtacc agacagagcg gctctttgta taaaggccta cattggtctt 101640
ccctattggt cctcacactt ctcatttcaa tcaacctttc tttactgtgt tggcaacaag 101700
tctttcctta gacaatttca tatttctcac atttaaatga agtaatgttt agtgcaaagt 101760
ggcttccttt aaaatattag ttatctgtta gactaactac tgaagttatt catgtattta 101820
ttcaagtgat atttctgagc ttctactatg tgtcagccac tgttccagat actgaagatg 101880
gagcagtgga caaagcagag aggattccca ttttatgaag ctgaaattct aaaggcaaac 101940
agttcacaaa gaagtacaca aaggaatgtg aagtacaagt actggaaatg tcttaaggaa 102000
gaagtgtgca gggtaaaaag catagagatt gtgtagtagg ggctatttta gattagttgt 102060
tcatagaagc cctctctgaa gaggtgaggt ttgagcagag actgaatgaa gtgaaggagc 102120
aaactgtaag gaggtctgag gggaagcatt ccaggtggag agaatcacac tgagcatgtc 102180
ctgaggtggg tgcatgattg atgtgttcaa ggaatagcac aaaggccagc gtggctgggc 102240
tgcaaccaat gcagggcagc atggcctgtg ggaaatcatc aggtcttgca aaacgctgtg 102300
gttcttacat atcggtatca tcctgatgat attgactggt gcgtcttgtg aatttcctgt 102360
gggtcgggca tcttgctgag tgcattgtct catttgtttc tcataacaag ccctaagata 102420
aggactgtat gtgtaaattg aggtctcaag aaatctgttg acttgcccaa gaccgcatac 102480
ctagcagatg atggagctaa atcttcttcc cggctgggcg cagtggctca cacctgtaat 102540
cccagcactt tgggaggccg aggcgggtgg atcacgaggt caggagatca agactatccc 102600
ggctaacatg gtgaaacccc gtctctacta aaaatacaaa aaattagcca ggcgtggtgg 102660
tgggcgcctg tagtcccagc tacttgggag gctgaggcag gagaatggtg tgaacccggg 102720
aggtggagct tgcagtgagc tgagattgcg ccactgctct ccagcctggg ccacagaggg 102780
agactccgtc tcacaaaaaa aaaaaaaaaa attcttccca agaacatgaa tttcagagtt 102840
catattctta actatgcaat ttctacagca tctcactaca aatcaacaga gtgaccttta 102900
agagaaggac tcagaaatgt gaccaggaaa atgtgtttca gagacaaaat gttctgaagt 102960
gcttatcttt gagagtatgg tttacttcat gttgtatctg taggaacaca aaatcaaaaa 103020
tctttatagg gtagatagag tgaaaaacct gctgatccct gaataaatga aaaggcacaa 103080
ttagactcac tgacagggat ttttcctttc taggaatgca ggtttaaccc tttgaaaatg 103140
ctgatcaaat caccaatgtg gaagtggtat taaatagttc cattttttc aagctgttta 103200
cagttaggta ctgaattatt tatgtccgaa aatagacaaa tctgataaaa atgtgaaatc 103260
tttccagatt attattctgg aaattaatac gaatgttaat actttaatag ttcttatgat 103320
aaagctgttt gcttgttctt tgtaaagatt atctgagagt aatatgtcct gatttgggaa 103380
```

tgtatgttaa gaattacagc aggatggccg ggtgcagtgg ctcacacctg taatcccagc 103440

actttgggag gccgaggcag gcggatcacg aggtcaggag atcgagacca tcctggctaa 103500

catggtgaaa ccccgtctct actaaaatta caaaaaaatt agctgggtgt ggtggtgggc 103560

gcctgtagtc ccagctactc gggaggctga ggcaggagaa tggtgtgaac ccgggaggtg 103620

gagcttgcag tgagccgaga tcgctccact acactccagc ctgggcgaca gagtgagact 103680

ccgtcacaaa aaaaaaaaaa gaaaagaaaa gaattacagc aggaaaaatt ggtagctgca 103740

ccatgtggta tgatgcttaa tttgaacaac ataatttctt acatatgttt aaaaattgca 103800

aattagcatg aaggaaacat gctttagtca ttttttttca attggtttac ttacatgctg 103860

taaaccatat catttgcatt ctgcagacaa taggaataag actgaccact gatacttcct 103920

gttaaatcaa agtaaatgca gtaaacgtgg tttcaggtat ggcaaaattc agaggccgat 103980

atggtgtcag gggctcagtg tatgtcttca cctacctttc ttctgttggt gtgtcggtgt 104040

cattcttaag ctggttttct ctatgtgaag gcaaagatgg ctcccaggag ctctgtaaga 104100

gtttgtgtga ctcttacagc agtacccgtt ggggattaaa aaaaaaaaaa aaagaaatag 104160

aacatttctt attaaaggca tccagcaaaa gccctgagga tgactcagat gactgtgaat 104220

actctggcca gatctggatc atgacgggtg gggagtcctt cttgcacctt tgtactaagg 104280

ctggggtagg aatgatttgc taaaggaaaa ccaggcagag ttgaagaagc cagaggcact 104340

ggaaagccaa aaagaaacaa atgttatttt agacatgtaa aaattaaaat ctaattttta 104400

aaaatgaatt ttctgtttct tactgtgttc taaatacact attcatcttt aacagacaac 104460

tcaatctctt attgacatca aaaaatcttg gatgatcatt gtataaacta tatttcattt 104520

atgccaagaa tgcaattttc ctatgttaat atctctgaat ttgggagctg ttataatcag 104580

tggtgtgtta tactttaatt ggcagcattt ttttatttct tagagctaca tacaaaattg 104640

gtgacatctt agacttaatg aaaagttgta tttttcacac tatatacttc agcttttaat 104700

tatgtactaa tttatacata ctaatcatat aaatatgtcc ttaaaagctg tgtctttgtg 104760

ttatactcgc tttgtaatag tagccattga atcctatttg tccaacacaa agtttggcac 104820

atagtaatca ataaaggttt gttaaacaaa tgagtgaata aatttgtcaa ttaattactt 104880

attttttttt gcaaatgtga gtatgcaaaa catctaaata tatctgtata tatttagatg 104940

tgtgtgtgtg tgtgtataca cagtcatgca tcatttaaca gcagagatac attttgagaa 105000

atgcattgtt aggcaatttt gttgtcgtgt gatcatagag tttacttaca caaacctaga 105060

tgtatagccc actatatacc taggctatat ggtatggcct attgctccta ggctacaaac 105120

ctgtacagtg tatgactaca ttgaatactg taggcaattg tcacacaaag gtaagtattt 105180

gtatatctaa acatagctaa atatagaaaa ggtaagtctt ataggaccac catcctatac 105240

gtggcccata cttgactcaa aagttatgcg atgcatgact aaaatatatg taaatatata 105300

cacacacaca tatacatgcg tgtatatata gagagagaat agttgagtac atataacaaa 105360

gatccaaaaa tacaaaactt acataagaca gaagtttatt tctctgcagc agttggtaag 105420 tatgcaggga aagcaggtgg ctgtgccccc catggtcacc cagcctgtcg ggttggacag 105480

gacctgtcgt attctatgca agccttctat ttctggtcca aaggggctac actatttgtt 105540

tccatttcca ggcaagagaa aggtatttta aaaaggattc agagaaagtt taacttaatg 105600

caccaaaact tgtacaaatc atttcctctc acattccgta ggtctgaacc cagtccactg 105660

gagacagcta gctgcaaggg agactggaat gtgtaggcta tagccccagg caaccacttg 105720

ctaggttaga attcggaaag ttctgttttt taaaggagga aagagaaaat ggatactggg 105780

ctgataatct caataagaag taaagattct tggcaaaagt ttgcttctct ttccatttcc 105840

ttgtgatttt ctaataatac ctaggaaaat attttatgag acagtgaata taactaatac 105900

aggttgagta tctgtatccc aaatgcttgg gaacagaagt gttttggctt tcgatgtttg 105960

tcagattttg gaatatttgt atgtatgtgc ataatgagat atcctgggaa tgggacccaa 106020

gtctaaacac aaaaattgtt tatgttttat aaacacctca aatacattcc ctgaaggtag 106080

ttgtatgcaa tattcttaat agttttttta tgcatgaaac aaagtgtgta tacattgaac 106140

catcataaag caaaggtgtc aggtgtagaa tttttcactt atggcctcat atcagcactc 106200

aaaaacttta aatttttgag cattttatgt tttaaattag gaatgttcaa cctgtattga 106260

attaaaggtt tccttgcata tttgaacgtt agaaaaatat taattttatt tcaacaatac 106320

gtactagtga atatcatttt gtataatcat ataatatgtt aaaacactta gttcatccat 106380

tcatttggga tttattgagc tagaccaatg gttattctat tgttaatttt agtcactaga 106440

tattcaaata gagaggtatg attttcataa agcactataa aataatattt tagtatgaat 106500

atatttaaga ttcagtgtaa ataaatgatc atatgtgtaa ataaatgttt ctctctccct 106560

tgataggttc tccgagtagt tcggctgatt aagatttcac ctgcattaga agactttgtg 106620

tacaagatat ttggtcctgg aaaaaagctt gggagtttgg ttgtatttac tgccagcctc 106680

ttgattgtta tgtcagcaat tagtttgcag atgttctgct ttgtygaaga actggacaga 106740

tttactacgt ttccgagggt aagagtttta aaaatgcagt aagttaaatt cattgttctt 106800

attttagtaa taatgattaa catcaaagta atttcacttt agtcattcag aagtattatc 106860

ctatttttga taacttggga taattagatt ttaatattta aaattgtctc ttctgcaggg 106920

tgcaggtgtt tgagaccatt ttcagcattg tcagattata tatcaattct aaaattctgg 106980

aaataatcta atagagttag atgctgaaac ctgaatgcta gggttcttag cacctataaa 107040

ttaaaataca gcagaatttt aaaataaagc ttgcattttt gtgttcaaga gagttaacat 107100

atttttaatg ttttataatt tatatattgt tatgttagtc agaatctaga aggaagcttt 107160

tgattttcaa taaaataaaa ttcataggct cttcagattt taaatgtctc tcagctttct 107220

cagtagaatt cacttaagct aaaaaggaag ttagaaatca atggatccat aaacaaaagt 107280

catgtaagag gtacagctct taaccagaag gaccttgaaa tttaactcta tgtggactta 107340

gaaacataaa aaaatacaat gaccagacaa ataatgtata tatttttacc ctaagttgaa 107400

atagtaaggc ctattatatg gtgtattagt ccattcttac actgatgata aagacatact 107460

cgatgaaact gggtaattta taaagaaaaa gaggtttaat ggactcacag ttccacgtgg 107520
ctggagaggc ctcacagtca tggtggaagg tgaaaggcac atcttacatg gtggcaggca 107580
agagagaatg agagccaagc taaaggggga aacctcttat aaaaccatca gatctcgtga 107640
gacttattca ctaccacaag agcagcaggg gggaaaccgc cccatgatt caattatctc 107700
acacctggtc ccttccacaa cacatgggaa ttatgggagc tacaattcaa gatgacattt 107760
gggtggggac acagccaaac catatcatat gggaagaata atattggggt aaaaacaaca 107820
tatattttta aaagttaagc tttgacagca aaagggtctg tacttctaat ggatagctta 107880
gctttatttt agggattttt ttaaaaatcc tatgctaaat ttattatttt agctaccagt 107940
gttggaattc atccttttat aatgcagatt cttcattcat gtgccgccaa ggatttgtcg 108000
agtgccttta aagcatctgg tcatttaggt aatcatgggg acataagtcc tcaagaaatt 108060
ctcagttgag gggaggaacc agtaggtaag cacagggcac cacaggcctt gatgaaggca 108120
gacaccaggt gctctgggga cacagaggca aggccgattc ccaggcatgg aaaagtccag 108180
aaaggcttat gggaacaggt gactccaatg gctggatgta gccaaaggaa gtgcaaggga 108240
acggattccc atcacaggaa cagcatgtac taaatgtctg aggtgagcaa atccttcttc 108300
cagaacttag caagaaccat caatggtaag aggaggacaa aaggccaagg aactctgaag 108360
tcggttgaaa gtgtgatcgt atctgtcatg ccaaactcct atcaactcta gtagggaagg 108420
caccaggttc aagaggctga agaagagacc cagagccagc agatgagaca tggggtttca 108480
ttgagggctt acatacaggg gagggagtcc aatgctggct ggctggtcag gagaacaaga 108540
cctgcttgca aaatgcatgc aatttacata gcatttttac ttagcacact cccccagcaa 108600
cctccacatg gaaatcttaa ttcaccccag acttggggcc tcaattccct gtgtatagcc 108660
agtgttccac gggaggagat ggggactcag atgttcctca tcaacaagga gtcagtctct 108720
gggttggcca ctcctagatt ccctagctca gaacacacat tcaggtgtgt ctggcaagaa 108780
gggtcattct cggggtatgc ttgagttatt gctttcaggt gcacttacat acgatgttta 108840
agtgctgatg agagagggca agtagctagt gaggttgcag atgggtgaga taaggaggat 108900
gatggcaaag gttcctgcag aggtgggaga aggcagtatc cacagcaaga ggacaaagat 108960
tggccttaga caggtgttga gctgacaggt gtcgaggtga tcctcaggac agcaagccaa 109020
gctgaaagtt acaatcaaga ctttgttccc tcactgcaac agtacaggca ggaggcaaag 109080
gtaggcaccg ctctccagct ctccagcgtc cggaggtgtt ggttaagggt caggtggatc 109140
agcaggcatg ggggagggag ggctcttttc cttgctaagc agtcctgaac aaaaggttcc 109200
agccattcct gcacctaggg aggggaagca taaggaagga agggatagga gaagaaagga 109260
acggagccag catggaggaa agtgcctcat ctgggccctg atcaggaggc ctctgcacag 109320
aggcacctgg tctgggactg cagctatggc catgatccca gctggcctgg cagggcagag 109380
ctgagtcata aattgcaagt ccctccagaa atggcagtgc cctggtcatg caccatgtct 109440
agggtgggga tggcagtggt ccccgggagg tctcccaggc aaggccttgc catcaggtct 109500

gaatggatca cacataggat tttggccagg agcaggactc ctggaacgga ggctggaggg 109560

aggcaagcca aaatgatagc cacatcatta ggcccacctt ctctgggtct gccccttcac 109620

tgggtaaata aaaagatcag ctggccatct tttccatggg aacatgggga tggttattgt 109680

aagagttaaa gaggtaagaa acacaaaaag cagctcaaca gtaaaagaca gatttatttt 109740

ggagaataaa ctttagaggg gcttctggcc aattttgctc aggagcattc tctcttacag 109800

actaagggta tttaagggtg taggaagtgg ggagcttatc gcaggttcgg aatgtttcta 109860

tgtgagggaa agtttattgt ggggttggga ttgggatgtc tggttggagg ggaggctgtc 109920

tcagggttgg cgtgtttctg gtcagagggg gtttatctta gggttggaat gtttctggtg 109980

atgctgacat taggcattag gctgatgttt gggngcttga tttgggtgtt tttttaatca 110040

aagggaactt aaaatggtgg tgtttctcca agacgacgat ctcctgctct gtcaatccag 110100

tccctacagt tatagaaagg aagaggggtg gtatattctt tctggctact tcctgctgag 110160

tggagggtca gagggttctt tggtcttggg ttgattgtgg gagtaatgcc atctgtagat 110220

gtttttaggt aattgtgaga tgtccatgat cctgtcagtt aaaaaatctt tgaaaaaggt 110280

gaattaagcc aggtaagcac attagtccta ggcatattat taggagaggg cccaagaatg 110340

ggatggtcca tgttaggatt ttgtttctaa accaagaatg tatttggttg ttttggtatt 110400

cctttcacgt tttagccctt tcttcaagtt tttcagcagc atctcttact aggcctgatt 110460

ggttgagata gaagcaacat tcctcactca gtgagaggca gagacccccct ttttcagcca 110520

ttgtaagatc taatcccagt ctattttgga ggactgctcc aggcaaggag tctaattggt 110580

cttcgactct tataaggctt tgggctatat cttttaataa atcctgtagt tctgttgaaa 110640

gaactttaaa gtatgttaag gagttggcca gtctgcctgc tcccaatcca agcccggagg 110700

ttatacataa ggcagccatt aaaggaatga cttggatgct cctccttttt ctaacatact 110760

ggatgaatgg aacaggtgaa gattgattag gaggaactag tccaacggag ggagaaagat 110820

aaactagggt acaggttctg gtctagttgg tggggagaaa aagatatgtg ttggtaccgc 110880

ataaaaagaa cgagcctttg tcataatgca ggcggagata gggaaagaag ataagcaaat 110940

taaagatggg gtgtttttttc tttcctgcgg ttttttactc caggtggaca aggaggaggc 111000

cagggagaca cccactatag tagagatata ggaagagtta ttttttacac atttaatgca 111060

atcggatgtt gcaccattga tattgagcca tggaaaaagt tgggcaccag gggcagcgca 111120

agttaggcag gaggcattgg ttgacggaga ggctgtaaac tagctggttg cagaaatgct 111180

ccatttgttt caggtgatac ttggtagtca acgtggctac tttgatggtc agaggggtgt 111240

ttaacaatga tagtgtgatt gcattggtta gatgttaagg accctacagg gaaatttccc 111300

tccgaggctg aaaagcaaag tgaggcttgt tgtaaaaggg gggtgtgtgt agttatgggc 111360

ccttcaatgg gagggtcttg gtgcttaagg cattgtagag agttgtaata ggtttgaaat 111420

attttgttgg cccgattggc cctggaagtg aaataatctc tgaacagggt gtcggctctc 111480

tcaaaaaagg aagctccttt tgggagttta taagttaggg ttatatttcc tgttcaaagg 111540 tcatgaagag gtgtaggaag ggctgtgtaa gctgaggaag acagcgataa gcatatccag 111600

cactctggag caaaggaaga gttagctttg cagcaagagg gattgtgtaa ggtttataga 111660

gcattctagt ttgagagcag tgaggagtgg ttttattggt gaggttgata tgattaggga 111720

atttatattg aaagaaacaa tcaaaaagag aaaaaagtta tttgggtagt agtaaagtcc 111780

tggaaagttc cttgaagcca taggtaccat agaacagtca gaaactgatt agcatgtata 111840

atgggagctt tgtaagggta ccactgaagg tctgtagcaa ggttggttag gaagcagtga 111900

aagtttggga gagaaagaga cataagcagc ttatgtggat ttctcttcct tttcctccag 111960

gatttaggtg aggcgaaggg aagtaggtcc tgtggagaca caagagaagg ctgaaggggt 112020

tttagatttg gttgtttgtg tatgtggtga agggaagtct gttttcttga gagaggtata 112080

atgaaaccag ttggggagtc cctggagttt tgctgccatg ggtgtagtaa ggatgatctg 112140

gtaaggtcct ttccatttag gtgtaaggag ggaattgggg ctaggactgg gatcttttat 112200

cagaacccag tctcctggct gtaggaaggg attagaggag ttggtgacaa gttgtggcag 112260

gtgtttgtca gcatattccc aaatgaaatg gcagatggta tgtacaagag gggaaatgag 112320

aggggttggc agaggtgggg cttgaccctg aggtggacca aaaggggcaa gtggtctccc 112380

atatgtgagt tcaaaggggc tgagcattaa aggtttatgt gagagcacct gaatttttag 112440

aagggctaaa ggcaaaagtg taaccgagtc tttatgtgtc tggagtgagt acctggtgag 112500

ggtgtttttt agaatgccat tcatttttc aaccttttct gaagattgag gtcgataggg 112560

gatgtgtagc ttccaggtaa tttgtagggc ttgtgaaagt gtttgagtaa tctgagaaat 112620

gaattcaggg ccattatcag attgaaaaga aagaggcacc ccaaacctgg ggatgatttc 112680

tgttattaat ttggaggtga cattagaggc tcgtttgttg gttgtgggaa aagcctcaag 112740

ccattccaaa aaggtatcaa ccagaactaa aagaaatcaa acccttttta ctgggggtat 112800

atgggtaaaa tcaatttgcc aatcctgtcc tggaaggtgt cccctggctt gatggattgg 112860

gaaagaagag tgtctagtgt tggaatgggg tgaagctttc tggcaaatag agcattgatg 112920

ggaaatggct tttaactgtt cctttatatc tggggttatg tgtttatggg aatttaagaa 112980

gtgttgtaga ggtgaatggc tagcatggaa gaggttgtga atgtcccgta aaagagttgt 113040

tttttcagg gtcaggtagg actaatttgt tttgtatgaa tcagtttggg gggtttgaat 113100

tgtgcccctg ccatgattag ttgttgtgtt tggtgttctg gataaaagaa ggggatatgt 113160

tgtatgaagg gaaataagta ttggggaatt gggtgattgg ttgagacatg ttttgcccag 113220

tagtcagcct catggttccc taaagaaatg tggcttttgt ctgatgttct ttgccatgaa 113280

taactgcaac cttttctgga agtagaagtg cctttaatag aggatgaatg agctttccat 113340

taatgatagg ggttcctata actgtgagat agccccgctc cctccaaatt taggcattgg 113400

aatggatgat gttataagca tatttagaat tggtgtatta ttaacttgtg tgtttttggg 113460

tagggttagt gctcttatta gggcaactaa ttctgcttgt tgggcggatg tgcccaaagg 113520

caagggggca gcctctatca ctcttctaga tgggagagag tgggtatcat aacgctatcc 113580 ctcaatgatg gcatatcctg cttggagagg agggtttttt gatgcactgc catctataaa 113640

ccagtctggg gctcccttta tgtgagtgga agtaagatgg tgaaacatgg taggaggact 113700

ttcaattaga tcagagcatg agtattggtc aggatctaaa atcggtattg agggtaaaag 113760

agaggcggga ttaggtgggg agcatctatg aagagagata gcgggttgaa ggagggttaa 113820

atgtaaggct tgcatatgag aggatgagat gggggtgagc gcctatggct gaccatatct 113880

tgtagactgt gagaagaaaa tacctgaagg ggtttgtaga acatgagttt ttgtgtctca 113940

gagataatta gagaggctgt ggccaaaatt tttaagaaaa gggaccagat tttataaatg 114000

gggtctagtt gttttgaaaa atatgccact gttgagaagg agtctcccat aagttgggct 114060

aatagtccaa gggcctgatt atgagaactg tgtaaatata gatgaaaagg tcttaggagg 114120

tttggaatgt ctaaagcatg ggcctgcaat aaggcacgtt ttagatgaga aaaagcatga 114180

taaaggtctg gggtgggagt gagtggttgg tcaagatttc cttttgtgtg ttcataaagg 114240

ggtttagcaa taatggcaaa atttgctacc cataatcgga aatatcccac taatccaagg 114300

aaggaaagta agtccctttt cgttttagga aatgggattt gttgaacacc ttgctttcat 114360

gccaacagca tttctcgggt atttggagtt atgatcagtc ctaggtatga aactttcggt 114420

gaggttaatt gggccttcct tttggacacc aggtacccac attcagccaa aacatttaaa 114480

atcttggtgg tgtgttgaat acaatggtca agaaaggggc tgcaaaggag taaactgtcc 114540

acatattgca gcaaaatcct gggttgtaga ggtagttggg agaggtctaa ttgaagtgac 114600

tgaaatagtg atggctgtct ctaaaccccc aggggaggac agtccagatg agttgttggg 114660

agtagcctgt gtcagggttg gtccaagtga aagcataaag gttttgagag gagggatgta 114720

gagggatagt aaaaaaggtg tctttggcca ggtgcagggc tcatgcctgt aatcccagca 114780

ctttgagagg ctgaggtggg cggatcatga ggtcaggaga tggagaccat cctggccaac 114840

atggtggaac cctgtctcta ctaaaaacac aaaaattagc tgggtgtggt ggcgtgtgcc 114900

tgtaatccca gctactcagg aggctgaggc aggagaatca tttgaaccag ggagtcagag 114960

gttgcagtag attttgctac taaatgtact ataagaaaaa aaaaaaattc aaaaggacct 115020

tcaagtattt tcagaaattt ttaaattttg gagttgtgaa taaggactgt gggcctctat 115080

ttttaaaaaa tgcatttaag agtctgtact tctgtcccta caagtccgat tgtgcatcaa 115140

tgccccattg ctttttagct atgtgatact gggcaaatca ttctttacat ttttattctt 115200

cacaaccaag atgtaagccg aaatgtgttt tttttcttc tgcattttgt gtattaaata 115260

atatacaaag atgacacaac tgggcttaga aagtatttct taaaagacta gattacagcc 115320

tatgtaatat atttgaaatt gataatatta tctgtataaa tggcaaggct gctttttaga 115380

tccattcctg acttagcccc tctatacttg ctctattttc cagctataat agtaatcaag 115440

tagtattaca ttttcagatc tccaattact caattgcacc attttgcctg attattttga 115500

ccatcataat gaaagcagct cggcctctcc caggcccccct aatcagttat tcagcagctg 115560

acaatggtga cagcctgaac gtgtgaagtg actagaagga gtatattttt cctttcacat 115620 aaatgttact aaaggtgaaa tccatgaaga gtaataaaat tttacaagga aatgttaaaa 115680

aatattacaa ttacaaaagg ggtcattatc atccctaatc aaataaagca tcaccaattg 115740

gtcttacttg ccagatagaa aatcattcag gagagatgat taaaaacata ttttactcaa 115800

ggctagattt gtcgtcagac gcacatccct gggtttcttt tccctccctt ctctagacat 115860

tatccagtgg ttgagtgctt aaggtaggtt ggttaattac agtattagcc aaggcacagg 115920

acatgaccag aacaaatttt ctgcaagaaa aatcacatga tgtatcgcaa agtgaaatac 115980

agactttaaa aatcagcaat tgtctatgtt atcttgactt gggtccaggc cgtcttggag 116040

gctgcctccc agcaatgttt atagggctgc agaatggtgt gggaaaataa agtgtggaag 116100

cacatgggag gcttggggag ggcactcacc ctgaatctgc ttcagcacta tttaccaggt 116160

gtgtagtctt aggtgacact taatttgtct gagcttcatt catttttaaa atggtaatga 116220

tgtctatcaa acaggcatga taactaagag ccctacttac acatagtaag actttgaagc 116280

ttggagagct tgcataaccc acacaaagtt acacagctgg acagctggat ttgaagttgt 116340

ctactgctcc taccaactat tatccattag gtcacttcta gccaagcaaa cacactagaa 116400

agtgttcagt taggaaatac ttattgagtg ccaagaatat aaatattttg tcaatctgta 116460

cctggaagga aaggcctctg gaatagttca tcacaaactt cccagaggac ttcaggaaat 116520

caagaaaggt atagaggtat taacattttg gggcacagcc tgtgagtgtt aggaatgatc 116580

atgtcaggat cccaaaccag cacctgatag tctctgtgcc acagcagaaa cagcatttgt 116640

ctaagaagat cccacacatg agacaggtga aaggatgttc caagtgacag cttgtgaaga 116700

ggtgtttgaa tggagaatga gcaacaatgc tgaaccaaca caaacatctg tcatccttac 116760

atactgggtg aggctcagca tgccggattc cttttcccag tattgttcat tcctccagga 116820

tcctcattac cctccatcat tccccttacc ctatccaatc agaagcgtct gcataattcc 116880

gttgcaaaca ctttggctgc aagtaacaga aaacttgatt ccagttgact ttaagaataa 116940

ggaaatttat tccataaaat tcaagagatc aggtagcttt caggaagtca ggactggatc 117000

tttcactgaa attctcttaa ttgttgtttg aattgtcagt gttcccccaa aaatgtatgt 117060

tgatatccca actcccagaa cctcaaaata taactatatt tgaaatacag gtaatcaagt 117120

taaaatgaag taatcagggt ggaccctaat ccagtatgcc tgatgctttt atgaaaaggg 117180

aaagaacaga cacaaatatg cataagagga agatgatata gacccacagt atattagcct 117240

gttctcacat tgctataaag aaatacctga gactgggtaa ttacaaagaa gagaggttta 117300

attgctcccg gttctgcagg ttgtatagga agcatggcag catctgcttg gctcctgggg 117360

aaaccacaag aaacttacaa tcacagcaga aggcaaaggg gaagcaggca agtcttatat 117420

ggctggagca ggaggaaaag aaagagagag gaggtgctac acacttttaa gcaacctgat 117480

ttcatgatga ctcactcagt gtcatgagaa cagcactgag aagatggtgc tcaaccattc 117540

ctgagaaacc acccctataa tctaatcact cccatcaggc cccacctcca atattgggga 117600

ttacaattga acatgagatt tgggtgggga cacagatcta aactgtatga cacggggata 117660 tggccatggg actggaatgg catctacaag ccaaggaatg tcaaagaatg ctggcaaaca 117720

ccagaagccg taagatacaa gggaagattt ttttccctag agctgtcaga gcacagccct 117780

gctaacacct taattttaga cttctggcct ccatacctgt gagacaatac atttctgttg 117840

ttttaagcca ccaggtttgt ggagcttttt taccacagcc ctagcaaact catacagtga 117900

ccttctccat gggttgtctt cctatgttag caaaatggat gtgacatttc caggcttcct 117960

gatgacatct agaggaatat gtgggtcctc tgttcccagg cctcctcctt aggaccagga 118020

aaccttttct atcatccctt tccttccagc aaacttgctg tcatctctca ttggcaggat 118080

tgagtaacag ggtcattcct agaattaaaa ttactggctt tggttaatca catgggaagg 118140

agtagatgtt gtgaagacaa ctacttgggt cagacccaaa acctggacaa gatttctctc 118200

agctctgtct ccaaagctgc tcctatattc cccatgcaaa caaggtttgt aaatagacgt 118260

ggtagaaaga aagctggtgt ttaattcctt tcaccagaaa catctggacc ccagatcctt 118320

cccaagatgc ccaaataggt aggtacctgc aaagtagcta gcatagggct ggacacagag 118380

tagatcagtt atggattcct gctctccctg tttttatggt actggtcagt ttctatcaca 118440

taccattgac attacacatg ggtcttactg cctttgctag tctacaggta ccttgaggtt 118500

gagagcatgt gcattcctta cacctccaag agcacttaac actgtgcagt gggatctgta 118560

agcatcagta agtagagtca gaaaggaaga cagctaggac tctatcagta tataagccca 118620

agagatggag aagggatgct cgaattctac tcgatggaaa aacaaagaat gctatattga 118680

tagaatttaa taaaacaact ttgtcatctc aagccaataa ctttctgttc tagtttggac 118740

caaacagcta cactgctaaa tagaaataac aagatgaaat cctgatgtaa ggcaatatgc 118800

aacagttcct gttttgttgg gatgaatata ttgtatgagt acaatttaga tgaactattc 118860

attctggaac agcctgtaca ggtcctaaga cagattgcca taagcaagaa caaccacagt 118920

ctttctgtat caaaaacaaa ttgcatgttt ttccataatt gcctaacaaa agctcataag 118980

cagcttcagt aagattttac tcagtttttc ctccatttat tatagaaagc aactgcctct 119040

aaattactga aataaatgca gtaaacattc actgtccatc tttaaccagc tctgtatctt 119100

tattccttct gatttcccac taaaataata aacagacctc atatgtttag cggttgtata 119160

aatggggata gtttttattt ctaagtaata tattatcctc tataacacaa agtacttaaa 119220

ttttagtatg atttatttga tataggtttt cttaattttc attctctaaa tgttaattac 119280

ttgctctatt gagttttcac ttataaaata tgtccaaaac agttattgca cgatagaagc 119340

aacaaatgct ataaaacaca taattaaaaa ttaactttcc tttcctcttg ccaccaaaat 119400

gactagtcgt gcagttagtt gtattaaagg aattaccaaa aaatttgaag aaacaatttg 119460

ctccttaatt ttttaaatta tttgtcagat gtagtgagag tgtgtgacat ttttaataga 119520

ttcaaggttt attttaatat ttaaaactat tagttcaatt aatattgctg aaatctttaa 119580

tttttgttct gtctttcatg tatttttaat tttatcatga aatatttgat gtgcaaaaag 119640

atatgaataa taataaaaca tactcctctg tggccatact taacaaacac taaattgcaa 119700 tagacctgaa gtccctgggt gcctcttcct gatctctgtc ctctcactct cccatggaga 119760

atcacttttc tgattttgat gtgtgttttg ttaataaatt tctttatact tttaatgcat 119820

atatggacat caactgaaaa tcatatatct tttgtgtaat ttaaacttgc cttatatggt 119880

ctcattctgt actatttctc ctgctaattg cttttattgc tcaatattat gtttatgaga 119940

gttacctaaa tgagtacatg gagctgtatt ttatttattg tcattactgt ataaaaagca 120000

actacatatt atattgcata tataatacat tctttataca gacagaacac aatttatttt 120060

ttcattttcc tgtgcagtag gcagattcta agctgtcccc catgatatct ccgcacgagt 120120

gtccataccc tgcatcatcc cctacccttc gatactggta ggatttgtga atatgatggg 120180

ctgtcattcc tgtgaatagg ttgctaagca tttaatgttg agtcagtcat aaatagagat 120240

tgtcatgtgt gggcctgtcc taatcagagg tactggcctt ttcttagatt taaagcaaga 120300

aagggtcttc tgttggcttg cattaggaga tatacctaat gtaaatgatg agttaacggg 120360

tgcagcacac caacgtggca catgtataca tatgtaacaa acctgcacgt tgtgcacatg 120420

taccctagaa cttaaagtat aataaattaa taaataaatt aataaatagg aaaaaaaaag 120480

taagcctcca tgttgtgaag tgctgtgtgg ccaagaccct gggcaacagc catcaagaaa 120540

gcaggactct cagtcctgca accctaagga acaaccctct gataataacc caaatgaagt 120600

tagaagaaga ccgcaagctc cagatgagac caaagcccag ccagcatctt gatttccacc 120660

ttttgagact ctgagctgag agcacagtta tattacacca tgcccacatt ctgaccttca 120720

aaaactgtga agtaacaaat gagtgttaag ctgcttcgtt tgtggtaatt tgttacacat 120780

caatagaaaa gtaatgcatc caggtgttgg atgtttctaa atattgtaaa tccaaacagt 120840

actgctctga acatttttgt acattgattt gtgcacatgt gaaataattc tttttagggt 120900

atgtatctgg gaatagagtt gctgggtctt agaataaaca tatcctcatt tctcattggt 120960

attgataaac tttcattcca aagtagtttt accaatttat cctaagctat cagcagtata 121020

tgagagtatt ctctacttta ccactcagga acacttggtc tggtcaggct ttattttgct 121080

tgaaacactt agctcttcat ttagcctgaa ttgatttttg tggctgctgt gagagaggga 121140

acaattttat ttccatttgg ataaacaatt gccacagtac catttattga ccaatctatt 121200

catgcccccct agttctgcat catatgccag ctgtgtcata acgtggttcc atatgtgagt 121260

gcccaagtca agaggcacag gctcccttgt catccccttt tgccattgaa ttttttttac 121320

tggatctgga tcaacctcct ctctgaggtt taaaaccaat aactatatag ctttgctcaa 121380

gtatctcagc agtctcctcc tttcaaaggt tcatgacttt gtaccttatc tctgacacaa 121440

cattaaaaac caaagtctag cagagatcgg aaattctcta gagaaactga cttttcttga 121500

ttcaagctac ttcttagctt ttggtccctg gagtttttaa ttattttctt ggactattca 121560

cctttaaaag gatgcttttt gcattttatc cagaatttcc tggggttttt gaatcatttt 121620

tttccataca ggatgagaaa ctgaggttca ctaataaatt tttgagaaga agaaacagta 121680

ctccatggac tctcagatgg atagactttt ttaaaaatta tactttaagt tctagggtac 121740 atgtgcataa cgtgcaggtt tgttacatat gtatacatgt gccatgttgg tgtgctgcac 121800

ccattaactc atcatttaca ttaggtatat ctcctaatgc tatccctccc acctcccccc 121860

accccatgac aggccctgga gtgtgatatt ccccaccctg tgtccaaatg ttctcattgt 121920

tcaattccca cctatcagtg agaacatgca gtgtttggtt ttctgtcctt gcgatagttt 121980

gctcagaatg atggtttcca gcttcatcca tgtccctaca aaggacatga actcatccat 122040

ttttatggct gcatagtatt ccatggtgta tacatgtgcc acattttctt aatccagtct 122100

atcattgttg gacacttggg ttggttccaa gtctttgcta ttgtgaatag tgccacaata 122160

aacatacgtg tgcatgtgtc tttatagcag catgatttat aatcctttgg gtatataccc 122220

agtaatggga gggctgggtc aaatggtatt tctagttcta gatccttgag gaatcgccac 122280

actgtcttcc acaatgattg aactagttta cagtcccacc aacagtgtaa aagtgttcct 122340

atttctctat atgcctccag cacctgttgt ttcctgactt tttaatgatc gccattctaa 122400

ctggtgtgag atggcatctc attgtggttt tgatttgcat ttctctgatg gccagtgatg 122460

atgagcattt tttcatgtgt cttttggctg cataaatgtc ttcttttgag aagtgtctgt 122520

tcatatcctt cgcccacttt ttgaaggagt tgtttgattt tttcttgtaa atttaagttc 122580

tttgtagatt ctggatatta gccctttgtc agatgggtag attgtaaaaa ttttctctca 122640

ttctgtaggt tgcctgttca ctctgatggt agtttctttt gctgtgcaga ggctctttag 122700

cttaattaga tcccatttgt ctattttggc tcctgttgcc attgcgttgg gtgttttact 122760

catgaagtcc ttgcccatgc ctatatcctg aatggtattg cctagatttt cttctagggt 122820

ttttatggtt ttaggtctaa catttaagtt tttaatccat cttgaattaa tttttgtata 122880

aggtgtaagg aagggatcca gtttcagctt tctacatatg gctagccagt tttcccagca 122940

ccatttatta aatagggaat cctttcccca tttcttgttt ttgtcaggtt tgtcaaagat 123000

cagatggttg tagatgtgtg gtattatttc tgagggctct gttctattcc attggtctat 123060

atctctgttt tggtaccagt accatgctgt tttggttact gtagccttat agtatagttt 123120

gaagtcaggt agcgtgatgc ctccagcttt gttcttttgg cttaggattg tcttggcaat 123180

gcgggctctt ttttggttcc atatgaactt taaagtagtt ttttccaatt ctgtgaagaa 123240

agtcattggt agcttgatgg cgatggcatt gaatctataa attaccttgg gcagtatggc 123300

cattttcatg gtattgatcc ttcctatcag atggataaga cttcagatgg atagactttt 123360

taaaatgagc atgatctagg gcaggccatc ccagtgtttg cctttactgt ccgcagcctc 123420

acagtctttt gcctttgttt gcattgagga tattgtgtcc tctttgaatg tcatagcctc 123480

agctttaccc tccctcagtc tgccacacca ctgaaaatca tctttcacaa aatatctgat 123540

atacccattt atgatgaaga taaataaatc atgttcttgg tttttcattc ataattctat 123600

caacattctt tatgcaacac agctaggaag catactgaca tgaacaaatg tactcactgc 123660

cctctttaac tatagcattg tcaagaaccc aaggaaggac tgaaacaact catcccatag 123720

tagaacacat aagaaagatc tcaagtaagc caaaggtgac atatgctaca atagatgtac 123780 tcagttctaa agtaatagca aagacatttg ggagagtatt ttaaaattct gggaatttac 123840

tgtgtgccta gcacttccaa atatcagctc atttaatgag agaaaaatat tgcttccaga 123900

gaaggaggtg tgcagagctt tataaagatt tgtctttagg ggtgaatata atttcaacag 123960

gtagagattt gggtcatggc attagatgaa ggcattccat ataagaaaaa ccaaatgagt 124020

aaattagaga catggagaat cacaggcatg gagggacata tgaccaaagg atggcagtag 124080

caatgggaaa ccagacatgg agggtcttca gtggttggaa ctgtgacaca agggagcttt 124140

tgaagatttt tactttattt tttttaattt tttttttctt ggagaatagc agaagagcta 124200

ttttgaggaa agttacactg gaagaaatgt tatttctatg aggaggtatt ttcattgaac 124260

aaaattcatt tttattgatg aaagacaaga atttatgatc agttgataaa tttcaggcat 124320

ctttggagat attttatatt tttttctgtg aggctgttac tgtgtcaaat ctttcacagt 124380

ttcataaaat tgagtctcat ctgaagcttt aggccagatt gttatccatg gttgagcatg 124440

actctgcaat agaaactaac aatttggctt attttaaagt gataccatga gatatatata 124500

tatatatata tagtttattg tcattgtctc cccccccaaa aaaaactgta aaattaattt 124560

tcattatttc ctttctacta taacattagg gttttttttt ccttgctcct gttggaattt 124620

atttttgtgg tgactcatct gtggcataat gatgaaatga taaaataata tcctgtgtac 124680

ttccttttag aggatgattc acataaaatt ggccatcttt tcctgctgtt gtatattaat 124740

atcattggtc cactgaactg caatgatcag aattatgact ttccttgaat aaagcatcct 124800

gtttcctacc cctaacccta gaagagtact ctttcaggga gaacctattt ttgcagctca 124860

tcttattttg gtagtttttc accaatcatc ttatttgcaa tttcccaact atttaattgt 124920

catcttacat taatattatt ttaataatca gactgttcca ttgttcaaat tttaaaaata 124980

aatgaaagtc attaatgaaa agtattgtct cattccatta taatgagctt cctcaagtat 125040

gttttcagtg atgaaaacat gaccgtgtta ttaaatattt tgtggattct tgcctacaaa 125100

aagaggacag ttcatacatt tgtgtctgta tgccttttag ctctactgca taaagaatgt 125160

cggtggaatt ttgaacaaca aaacaggcac atgatttatg tagtcttgcc tacaaaaaat 125220

acagtctgta atgagactct attgtcagtc tctcatggct acccacaatt tcctatcaat 125280

aaaacaggat aataatacat aacttagagg attggtagga agattaaagg aagtaacagg 125340

tgcaaagtag ttagcactgc gcctgccatg taaacaggca cttaacaagt agtagcagtt 125400

gtccttctta aatgcatgga aagcccatgt cttagcatgg catgctgaat tcctagcagc 125460

actgaggtag cccagcatcc aagggtaaat gactggagaa ttaatggaat ggtgggcact 125520

cagttttctg tgttgtatag aaagtttaaa actttaaaaa tagtccccaa ttccagttta 125580

ggagattatt tttaatcagt cttttgcgac tttccagaca aatgcaaaag gtaaaaaaaa 125640

atcatattac agaaaaataa gacaaaccga caagttattt ccagacgcgg aaatagctta 125700

tcattaacta ttacatatac ggaactaaac agagaccatt cacccaatga tatgtatatt 125760

cattctttct cttgggtata cccttcttgg atgtaacacc cagagtctta gtgtgccaga 125820 agactaaaat gcctacttct gcagaggact gtctctaaca ttttttctag gaatcaatgt 125880

tgaaatctta atttccagga agcatgatgg gataaagcat aatgtgtgta cgtgtgtgag 125940

catgcatctg gctgagcaca tgtgcatgca gatgggggtg gggaaggaga gagggacccc 126000

ccctaagcaa atggaagcta atgggctgca tcctataatc ttaaattgtg caatacaaaa 126060

ggaaaagcag agcaggtggg accagatgca gaaagaacct tagtagctag tttagatgga 126120

ggtctgcagt attgctgctg tggcttttgc tgttgcttct ttgtgttgtt attttagttt 126180

tcagaattcc ggagaaaaag gcaacaggca aaagataaaa atcctgcctc aatccaacag 126240

aaatgtactt gctaaattcc tgtgagaact ctgcaaagtt ggcaataaag aatgcatatt 126300

aatatgaacc accaaaaagt gaattctgtt aacattctag acaaatgtca catttgctgt 126360

ctcttgggat aaaccaagtc atttcaagtt gagtaaaatg aagtaaagcc agtgggatta 126420

tgcaaaggta cagcaaatga aaggaaaggc aacactctct tgaagagcga gggagaactt 126480

tcctgtatta ctcactagac aaagattatt atcttcatct tatagttgct gccaattcaa 126540

caataaataa aaagaaatgt ttattatatt tggccttgtt ctaacagaat ataatgtgtc 126600

cggcaataat tatatccagg tcaaataaat ttaatcaaaa tttaaagtaa ttaaaaataa 126660

aacaaattaa gatggaaaca aaaaatggca ctccaaaacc taaaaggctc atcatgagga 126720

cactcatgtt tcccagaatc ccggaccaag cgtgactaat ttttctagca aatatacact 126780

ttgatagaca aacatgttta tacaagttct acagtgatta tgaggtacca actagttcct 126840

agagaagcat aattattcct gatcaagtcc agaaagacat ttttctggat gttattaatc 126900

ttaatggtta ctatatgtga ttgcttttta tgtgtgagag tttacaagag ccttgaagaa 126960

aaatggagag atggatgggt atacacagag agagagaggt gcacacacac atagatatcc 127020

acatatatct aatataattc tcaaaacgac ttcattaagg aaaattatgg acgttataaa 127080

aaagaagcag tgaggaacag caccagaaat tgatgagaac caacaagaaa acatataagg 127140

ctacaataac tagaagtgca tgacttaggg ttcgcagaca gatgaatggg caaacacgca 127200

ttgctataaa gaaatacctg agactgagca atttctaaag caaagaggtt taattggctc 127260

accattctgc aggctgtaca ggaagtgtag caccagcgtc tgctccttgt gaggcctcag 127320

caagcttcca gtcatggcag aaggtgaaag gggagcaagc ggtttcacat ggcgagaaca 127380

gaaacaaaag agagaggagg cgccaactct ttcaaacaac cagatctcct gtgaatttag 127440

agcttagagc gagagctcac tcgttatcat agacagcacc aagccatttg tgagacagct 127500

gccccgtgac ctaaacacct cccagcaggc ccccacctcca acactgggga tcacatttca 127560

acatgggatt tggaaaggac acacatgcaa accatatcac catatattca ctcaaacaac 127620

atgtaacaaa gtgtcacaat gaatgattaa aagacatatt aaacaagtta agttgggaat 127680

actaaagccc tctcatactg tgtccggaat tgttgggttc ttggtcttgc tgacttcaag 127740

aatgaagccg cagacccttg cggtgagtgt tacagttctt aaagatggtg tgtctagagt 127800

ttgttccttc tgatatttgg acgtgtccgg agtttctgcc ttctggtggg ttcctggtct 127860

```
cgctggcctc aggagtgaag ctgcagacct ctgtgatgag tgttacagtt cataaaggtg 127920
gcacatctgg agttgtttgt tcctcccgtc gggagttctt tgtccctcct ggtgggttcg 127980
tggtcttgct ggcttcagga gtgaagctgc agaccttcgt ggtgagtgtt acagttcata 128040
aaggtggcgc atctggagtt gttcgttcct cctgttggga gttgtttgtc cctcctggtg 128100
ggttcgttgt cttgctggct tcaggagtga cgctgcagac cttcgtggtg agtgttacag 128160
ttcataaagg cagcgcctct ggagttgttc attccttccg gtgggttcgt ggtctcgctg 128220
gcttcaggag tgaagccgca gacctttgcg gtgagtgttg cagctcataa agacagtgcg 128280
gacccaaaga gtgagcagca gcaagatgta ttgcaaagag caaaagaaca aagcttccac 128340
agcacggaag gggacccaag caggttgctg ctgctggctt gggtggtctg cttttattcc 128400
cttatctggc cccacccaca tcctgctgat tggtccattt tacagagcgc tgattggtcc 128460
attttacaga gtgctgattg gtgtgtttac aaatctttag ctagacacag agtgctgatt 128520
ggtgcattta caatcctttg gctagacaga aaagttctcc aagtccccac tcaacccaga 128580
agcccagcca gcttcatgtc tcaatactat gtatcaaaac attttcttga tgaattagag 128640
caaggatatt taaagaatga aaaatgtaga gaacaaaaag aaaataaaat cttaatcttg 128700
gatagtcaaa ggccacctaa gcctaacaga ataaaatcaa ttacaaagaa gaatgaattt 128760
gaatttttaa acaagtcaaa aataaccacc aaaaatattt tctaaacaaa atatgaaata 128820
caaccttaac ttaccatttc agtaagaaaa tgtcaaatgg acaagattta tttgctatat 128880
ttggtttttg ttatcaacaa tgcacactgc tggcaaggaa gcacagatac tgtgatatat 128940
tgcatgtggc agtgaaaatt tgtgtaactt ttctgagaat tggtttttta ttatataaaa 129000
tgacactaaa aagcatctta tcctttgatc ttctaattgc atttataaaa atctctccta 129060
nagaanagta acctgagatg catacaaaga ttgatcttca aaaaatgttt tttgttgcat 129120
tatttactgt tttgaaaatt ttgaaacagt ttttctgtcc aaccatagta gaatgactaa 129180
gtaaattttg aaatatagtg gtttatcatg caattatcaa aactatattt ttatagacta 129240
attagtgatg taagtcatga gtgagagaaa tggtgatgat aaaatatttt taaaaatttt 129300
accaaactat atttgcagct gaattcaatt ttgattcaag tacatttgta tgcctagaaa 129360
caaactgata gaatgtttaa attacaaaac aaaatatata catttttttgt tatatgtgta 129420
tatatatata tatatattta aaaaacctgt tgtttgaatt attttatgct tcatatattt 129480
tggtattttg agatttttta ctatgaaaat cttcaaaaat ggaaacaggt ggactgataa 129540
acaaatagat ataaaacctt actttatagt taaacatcag tgagaaaaaa actgtttttt 129600
aaatggcaag tacatgaaca ggtaaagttc taaaagaaag aaaatacaca cacacacgga 129660
catgctagtt atatttacaa taaaactaat ttaggttttt gacttgtaaa atattttatt 129720
aaagtaaaat aatttcaatt aagtgtttat tcataaacag ctatcaagtt tatttttaag 129780
gtattaaagt actaaataag cactattaga agtgtggctt tgaaatttat atctgcaaaa 129840
gaccacatca tttttgataa agctctttat ttgtgagctt agtgaagaga aatgagaaat 129900
```

caaagttgat tacacaggtt actttttcac atgataagag aaagatacag ggagggagga 129960

aggcaggaat tgnntgagat tttagaaaat ctcagattta aaacttttaa aaattaagta 130020

ttgggcatga ggaaatgagc agagaaaata gtagtcatat ttccacattt gtaaagagaa 130080

gaacaatagc caggcaaact ccatgtctga agtgtggttt attttaatat gcaaaacata 130140

agcatttctc tgaccaaagc tgctttctat ttaaagctaa atataaaaat attatttttg 130200

agcatagtat tgcgagaaca tagtgaaatt tcctatgtag attctgtgct gttttagtta 130260

cggaaatatt tttaattaat cctccattta acaatggggc atacatctgt gtttttcgaa 130320

gcactgtttt gtatatatta gtaattatca ttgcactgaa atgaaaatat caatttcatg 130380

tacaatttta tagcacatat ttgattgaca atgaaatgca tttaacttat aaaaatgtaa 130440

tatggctgac attagcaaaa tctcttatta agaaataggt cataagtact ccaagaaact 130500

tgacagtctg ttctatgaag agcattaatt caaaagtccc ctatttgcat aaatctccag 130560

gtctcatttt cacagctgtt tgttgtcagc tagacggttg cagttcaaca ttattttaac 130620

ataatcggtg ggaagtatac ctttcaacta aaaataccat gctctgcaag aaagattgtc 130680

aggggtaaaa tataagttca ttcagttgat cttgatattt ttccttaaaa taaatatttc 130740

taatacagga actttaagtc tcaaattact gaaatatttt tattattaaa atatattttg 130800

taattttata ttttcataaa aataaaaagg ttatgtgttc atacataaaa atgttaatca 130860

gtagagctta ttttagcaaa gagtaaaaaa tactgaggat atctattttt cattaatcag 130920

ccaagcatta tatggagtgc aatataggtg tgctggaaaa aagaggatta aaatttttag 130980

atattgctga gaatatgcat gtatgttctt tgttttatta attttaataa tggaaaggga 131040

tcagcaatta ttagatctta aatgtctttt acctgtaaca cccaagagca cagaaggaat 131100

caaagtcaag agatctcagc ctggatactg agcaacacca tttgctacct gtaagctgga 131160

gttcttgatt taaaccctat gtacttcagc tctccatctg cagtaggggc cattatatct 131220

acagtaatca cttcacaggg tcattgctaa gattatatta aataatggac aattaactac 131280

tgtgggaatt gcaaagcacc atacaaatgc tagtcctggt tcagtgcctg ttctcatatc 131340

atttcatctt agctgcaatg aaatcctagt acaaattcat gaccttgtcc agctgtcatt 131400

tatggtataa tacataaaag ttaacccaag catagccaag gaaagagaca attggacata 131460

cctacatata ttttcataac tagcaaaaag attcgctggt ttttgccaat gtgcttctgt 131520

tacttttttt ttcttttact cttgaaatta tatttaagtg tagactccac ttaaagtaat 131580

ttgctgctgg aataggcata atgttctgat gtctgctttg cttctaatta tgttcaagct 131640

cattttgaca tttaaggcca attacaaaag tctagattca gaagactttt ccttagtcat 131700

ctaaaatttg aacaataaac agtttttttc tcctctctct ttattcattt aagtggtatg 131760

cttggcaaca agcatatttt gggaattgta aaagcaaaat aaataatgtg ggttcatttt 131820

gaaaaaagaa ctacttggtc tttcttgcca acataaaagg cagtcacaga attaactgaa 131880

accaccagtt tcaactttat ttcctcctaa atcacaagga ataaggaatt attttacaga 131940 tacagtgata ttttaaacgc ctaagttcta ttaacataaa ggtaattgac ttcaacagat 132000

ctctaattcc aaaagagttc atacatagag agatatgtat aaaaatagct acaaaatctt 132060

tatcatttaa aagtcactca taattaaact ccccagacta ggaccctaca ccaaaattca 132120

gttagtgttc cattgcgttt taacatccag tcatgtgatg acttacgaca tagagcaaag 132180

cacttaattt acctgcctcc ttattttcag tcaatttgcc aactgcaatt tttactgttg 132240

attcagtaga agaccctgat gagttgacag caccaaaagg caattcatag catctattag 132300

taagaagtca ctgtagcttc tttgtttaac ttttataggt aagatctggc ttggaaaaat 132360

agactggaaa gtatattgta aaaaaaacaa aattttccaa ttgactgtca ttatcagtta 132420

atagattata gaatagtaga cctttatagc tggaagatat gtatcgtgta cttcaaagat 132480

gaaaggaaat aatgtctttg aaaacactgt caaattttca aaattatgta tctaaaagaa 132540

attccttaaa aaatagttta tcctttataa tttttgaact ttataaaaat ggaatcacac 132600

catatatgac caaacatttt aagtcataaa gaatacacag gaattctggg ccgggcacag 132660

tggctcacac ctgtaatccc agcactttgg gaagctgagg tgggtggatc acctgaggtc 132720

aggagttcaa gaccagcctg gccaacatgg tgaaaccccg tctctactaa aaaatacaaa 132780

agtcagccag gtgtggtggt aggcacctgt aatcccagct acttaggaga ctgaggcagg 132840

agaatcgttt gaacctggaa ggtggaggtt gcagtgagcc aagatcgtgc cattgcactc 132900

cagcctgggc gacaagagcg aaactctgtc tcaaaaaaaa gaaaacatag gaattccgga 132960

atttaaccac tgtgatagaa gttgcttttc cttgggacaa tatttttaag gcagatttca 133020

aaattattat ttcaactatt ctttatataa tgctttatta tttttaaagt attcacatat 133080

atgttatttt gattttcaaa acaatcttcg agtgacaatg ggattgcata tttgagaaag 133140

gctggtctca ggacggctga atgtttttcc caaagccgtt aagtaagtcg gtgacttaga 133200

ctcagacgct gtgaaagcag aggctgagta cttggttctt gacctcactg atctgctggc 133260

ttacagagca tcctggtctg catactcagc gattcaatct ctatcttaaa cgcagtccct 133320

cagccattcc ttctagaaag gtgatctctt tagagtgaga gcagtgatgg aggtggctga 133380

agttacacca gggggagcag ctttgaaaag ctcctggctg ataggcatgt actaatggga 133440

gcaggagctt tgttggcagt gtcgaagtcc ctttggaggc tgtgtgagga gggaacccat 133500

ttccctggga tatagctctc tccacagggc agctggatgg ctaggggcag ccatggagta 133560

tcagattcac ctaggaatcg atatttttcc tgagggtttc acaaaaaaag acagagagat 133620

ggaatttgaa gaactgggaa gagtgtgatt gaattggagg gatctgaaat aaagaaatag 133680

ctatggtcaa tggctctaac caacagtact atatttatac tccaaattcc aaatccacac 133740

tatgtgtatt taggtacaat ttagtttagc catagtaagt tgtggcttct ggagctttta 133800

tttaccatag gattttaaat acatttcatg tcatcttcat tggttttcct taaccсctaa 133860

ttctccacaa gatcttcatc ccgtcagtta tgactttgcc aaccaatata ctttttagaa 133920

atgtaacccc aacattaggt acgaccagca tttactgtat atcaatatta tcctaaatat 133980 tgctgtagct ttgaatagct gaattaagtt ttaaataaaa gcaaaggaat tataaaacaa 134040

aatattaacg gattatgagg tcagtgggtt tttctgaaga taaaactatt ctactaatta 134100

ttgctgtttg ggaaaaatag aaacaaggag actgagtaaa tgaatcatta aggcaattaa 134160

aaacagaaga gaagtcactt ttgcactcag tgaataccat caaggcaagt cttctatttt 134220

agggctgtct tatagtaaga tataacaggg actctctcaa taaaaaggtt gagctactat 134280

ataaatcatc aagtagattt tattttgaaa gcaacttttt ttgtagagaa agagttcttg 134340

aaaccagtta aagatttaaa aagttgatac aagttatgtt gtaaaatcaa agttatatgt 134400

ggatgatacg agttgtattc cccaccacat tttaacattt ttagtgtata tatttttaaa 134460

tgaaaaaatg tacattttac acgtatttta gtacaaacag taaaaaaaaa gtcagaagta 134520

atttatcaaa acatatgttt acacaaaaac ttgtatataa atgttcataa cagcattatt 134580

cataatagcc aaaaggtgga aacaacccaa atatcaactg ataaataggt aaaatgtggt 134640

atatccatta ataaaatatt atttggcaac aaaagggaat gaagttccga tgcatactac 134700

aacatggatg acccctggcc acatattggg tgattccatt tatagaaaat atccagagta 134760

caaaaagcca tagagacaga aagtagatgt ctttgatttc caaaggctga gggagggagg 134820

aatggagaat gacagctaac aagtatgtag cttcttttgg ggataaagag aatgttctta 134880

aattagatag tggtaatggc tgcacaacca tgtgaataaa ctaaaaccca cttaatcgta 134940

cactttaaaa ggtgaacttt atggtatgca aattgtatct caaaaaccta tttcttttaa 135000

aagtcaaaaa aaatagtata gaccccttcc atagttcctt agtcctttcc attctagtcc 135060

ttagagttac aggtttattg tatacacttt caagcattta tatgcacaca aaagcatatc 135120

tatgtgcttg ggatatagca aagggatggt actatatatt ctacttactc tactttgacc 135180

aaaacccaag atgtttaacc attctacttt ttcttaatat ttgactatgt ttgaacttaa 135240

attgatatcc acacatacta aaataccaaa atatgcattg agtcatcaaa tttcaaaggt 135300

catccatgat attttttaag ggcttaaaac tccttngcat gatcnttcag cctgtttant 135360

tcagcatagg aaaataccct tgctttgcct tcacctaact ctaaggaagt ctattagaga 135420

tatttagggt cagaaatggg ttcttggggc ttgagagaga aaaggatctt ctgctattga 135480

tgatctggct ctgccatgac cctatggcat gggtcctcag tggccctgtg ggaagcctag 135540

tgtaggcaac atggcctagc ccacctatgt tgatgtttaa ggggtgctag gaacaaaaga 135600

gaagctgtgg aaacttttgc ttatccccct ttcttccctc tttggtggtc ttatttttat 135660

ccgacgctct taatctcaac ttataaccgg aaaatacagc accaaagcgt gagggcattt 135720

aggggaaaaa tgttatctgt tttggaagta atatcaagga agatcaggga caaatgcaga 135780

actggaagca aggactaagg ttggatttag caggatttcc tgaaagattg gatgtaggat 135840

gtgaggaaat gagatgaacc aaagatgtct tcagggtcat tggcctgaat aacttgaaga 135900

atagatttgc cattttcttg gctgggaagg agtgcagaaa gaacaggttt gaaagtgggg 135960

catgaggagt ttcatatgac acatcaaata cttgagattc ctgttatcta acatctgtcc 136020 aagtaaaaat gtggagttga ttctggaggt ggggctggaa atataactgt ggaagtcatc 136080

agaacataca agggatgtga aggcattaca atggatgaga tcaagcaggt agtgagtatt 136140

gagatataag aaaagatctg aggtcaagta ctggagctct gtaacattta gacattagga 136200

aattgaagag aaaccagcaa aggagcatgc aaagggctgg cattggattt tgtggctctg 136260

tagtgcagtg ctctggctct tttggtgact gaaagatgct tgttcactgc ttttctctgt 136320

gttttcaccc ggagcatccc tgtgacattt ctatatatat atatttattt atactctaag 136380

ttctagggta catgtgcaca acgtgcaggt ttgtacatat gtatacatgt gccatgttag 136440

tgtgctgcac ccattaactc gtcatttaca ttagctatat ctcctaatgc tatccctccc 136500

ccctccccc accccacaac aggccccagt gtgtgatgtt ccccttcctg tgtccaagtg 136560

ttctcattgt tcaattccca cctatgagtg agaacatgcg gtgtttggtt ttttgtcctt 136620

gcgatagttt gctgagaatg atggtttcca gcttcatcca tgtccctaca aaggacatga 136680

actcatcatt ttttatggct gcatagtatt ccattgtgta tatgtgccac attttcttaa 136740

tccagtctat aattgttgga catttgggtt ggttccaagt ctttgctatt gtgaatagtg 136800

ccacagtaaa catatgtgtg catgtgtctt tatagcagca tgatttattt atttattttt 136860

tccaactctt ttttttatta tactttaaat tttagggtac atgtgcacaa tgtgcaggtt 136920

tgttacatat gtatatatat gccatgttgg tgtgctgtac ccattaactt gtcatttaca 136980

ttagctatat ctcctaatgc tatacctccc acctccccc accccacaac aggccccagt 137040

gtgtgatgtt ccccttcctg tgtccatgtg ttctcattat tcagttccca cctatgagtg 137100

agaaagtgca gtgtttgact tttgtccttg tgatagtttg ctgagaatga tagtttccag 137160

cttcatccat gtccctacaa gggacatgaa ctcatcaatt ttatggctgc atagtattct 137220

atggtgtata tgtgccacat tttcttaatc cagtctatca ttgttggaca tttgggttgg 137280

ttccaagtct ttgctattgt gaatagtgcc acaataaaca tacatatgca tgtgtcttta 137340

tagcagaatg atttataatc ctgtgagtat atacccagta atgggatggc tgggtcaaat 137400

ggtaattcta gttctagatc cctgaggaat tgccacactg acttccacaa tggttgaact 137460

agtttgcagt cccaccaaca gtgtaaaagt gttcgtattt ctccacatcc tctccagcac 137520

ctgttgtttc ctgacttttt aatgatcgcc attctaactg gtgtgagata gtatctcatt 137580

atggttttga tttgcatttc tctgatggcc agtgatgatg agcatttttt catgtgtctg 137640

ttggctgcat aaatgtcttc ttttgaaaag tgtctgttca aatcctttgc ccactttttg 137700

atggggttgt ttgttttctt cttataattt gtttgagttc attgtagatt ctggatatta 137760

gccctttgtc agatgagtag attgcaaaaa ttttctccca ttctgtaagt tgcctgttca 137820

ctctgatggt agtttctttt gctgtgcaga agctctttag tttaattaga tcccatttgt 137880

caattttggg ttttgttgcc attccttttg gtgttttaga catgaagtct ttgcccatgc 137940

ctatgtcctg aatggtattg cctaggtttt cttctagggt tttcatggtt ttatgtctga 138000

catttaagtc tttattccat cttgaattaa tttttgcata agattcctca tacattcctg 138060 ttgtgctttc accctctact ctctggttct caccctgttc tcatatatcc aaggaagatg 138120

atacactcta attagaaaca acaactgaac tcagctcatt attttcaaaa tctttcttct 138180

ccattaggta aaaatcattg cagtctgttt attcaatatg ttttcttttc cacttaaatt 138240

ctgttcatgc agagactgat tcctgggaac atcaatgtgc tggcaacttt gactttgtaa 138300

tttccctttt gttgtcccac tcaactgcat ggctgatctc cattcatttg ttcacttcca 138360

gagcctactg ttggcaatct gatttttctt catagtttcc cattaaaatt cactagagat 138420

gcttttgatt tacttatttc ttccctggca gggtagagtt gcttgttcca gatcacctca 138480

tagcaaactg gaccattctc aagtaaactg ctcatccctg gtgcagttat cagcagggag 138540

gcatctgcag cattaacagt gattgacccg tatagtgatg tttaatatgt gataatgtta 138600

ctagcagcaa acaaaatagc acttgtaaac acagttgttt actcatactt acatgtatag 138660

gtttggactt gaggggtaaa atagcaaaat tacccaactc ccctgctaat aaaacctaag 138720

gaaatgaata aaatgacctg agaaaattat accagcataa ctacagtaag gaagagctac 138780

aatttcatgc cattttctcc gaatagcagc taggaggaaa gaaaagaaga gaaagatgaa 138840

ggaggaggaa aaggaaaggg tcgcggagga ggagaagaga atgagagaga acatgaatat 138900

tctaggctca tcccttaaga ctgtcttctg gaagaagtta agaagaaata aacttggaaa 138960

aacaacgaaa ccgtatactc aaattcccca atctgtgaag tctgtggcca aagggctatg 139020

cagagccctg gaaagtgata agaaccactc atgtatagaa gccacttgtt ccagaggcaa 139080

aataatccac tagatctcac catttttatt tttttcagct gaggaggctt gagagtgaag 139140

gcaaaatgtc ataccttctg aaaatgaatc atctcccatt ctgctacaaa aatgtagcat 139200

aatgaaaata tgaagagaca aaaacagaag cttcaagagt acaggaagct caggataaag 139260

aggggtgtcg aaagccccag gagttgtcct attcgcccat acagagaacc aaaaccaaca 139320

ggaggacatt gacttctccc tcacctggag tcagtggagt gggaatgatg gctggagaga 139380

ggccagtgaa tttggggaag tgaacactta ggactgatta gttctggctg caataaagat 139440

ttcctgaacc tacctgaatt gaagaaatat ctgaagcctg atatttcaag cataattgcc 139500

catacacagc ctcagctaat acttctctcc ctgtacccct ctcctcatac tgaaacagta 139560

aaacacagag aagaaaacct cgttgataaa ttgaaaaagg atgctaatga gtgggatgag 139620

atatctggaa taattctctg cccctgtatt agaaagaatg aaagatctag ataataatgt 139680

taaaagtgga aattttggca actgtggaac cagtctagaa agaaattagt tcaaattaaa 139740

acaggtagtg agaggacccc aagaataata tctatgagag taagaatcaa atggtttctg 139800

ggcaattgat ggtttaaaga aagaacatct gcaggatagt aagaagggtt tattctaaat 139860

gataagtgta atgtgtgaga atctggaagc tacttgagat agaagaaaga aagaatatat 139920

ttanttttcc caatgagaca tgtgagaaat agtcagaaaa tctcagcaaa aacaaaacca 139980

aaccaaagat tgtattaaaa aaaaatggct caaataacgt gagttaaatt ggcccagatt 140040

ctgagcaatt aatggaaagt aaggaaggtt ccatttaacc ttgaccattt actgtctaat 140100 agagtcttgg agggggaagt ataattttag aacattacct ggcccaagag tgagcaatgt 140160

tttcatagga aacattgact atcaacatat tgattatcaa cctgctgaga agcacggatg 140220

acccaattat gatagaaaaa caaaatgtaa atattgtcaa acttgaggat ataaatgtat 140280

acatgacacg agatggaaaa tgaaagcaga gagaagaggt gacctgacaa gtggcaatgg 140340

taataccttc atctcacact ggaaggggtc aagtctttct agtttgatgg aaaaagaaat 140400

cgagatttac atacaaattt gggagtggtt aagggaagtg gggagtgatg ttgtgcctga 140460

atgaaatctt cacatatcat agcagaagtc tatgtctaaa gtggataaat caagaaatcc 140520

taaataggga taggcataac gccatcagaa gtgaaaacaa aaaattcctg tctcattact 140580

tataccaaaa taaatttcag atttaaactg atttaaacag ccaattcagg ttcaacatgt 140640

ttaaaataat aaagcaagga aagtgtcaaa ctactgcata aaaccatggg taaatttatg 140700

tctaatcatg cagtgagaga agtcttttta agcaaaataa aaacacttaa aaggaatgac 140760

aaattagacc acactgaaat ttaaaactga tctttgcaaa aatataagaa atgtataggg 140820

gaaaatgcac tgaacataca tggcagataa aagatgggtt ttcttatttg taaagagttg 140880

gtggaaacca ataaagaaaa taaaatggtc aacagggcag aaaaatggaa taacatgcga 140940

ataggtagtt cagagaaaca acttcctctc tgccttcttt ccactgaaac tctactcctg 141000

tgcattactg aggtcagttg cttcattttc ccaaaactgt tcttgattca cccaagagga 141060

ctaaattggc tcttctctgt tgcccctcag tactttaacc atattttgct tacacaccaa 141120

agttactatt cttccctttc tttttatttt ctttttctt ttgagacagg atctcgctct 141180

gttgtccagg ctggagtgca gtggcacaat ctcagctcac tgaaacctcc gcctcccagg 141240

ttcaagcaat tcttcttcct cagcctcccg agtagctggg attacaggtg accatcacca 141300

caactggcta atttttgtat ttttagtaga gacggggttt caccatgttg gccaggctgg 141360

tctggaactc ctgacctcaa gtgatctgcc cgccttggcc tcccaaagtg ctgggattac 141420

aggcgtgagc catcatgccc agccagtatt cttccttttt gagaattgat cctgttcatc 141480

tttttcttcc cctctcaatt atacaattct taagggaaga aactatgtct tactccactc 141540

aatccccaat gcatcaaacc tttagtaacc actaataatc cttgtttgaa taaataaaga 141600

gcacaggccc tacggacaat gtaagtcagt aagtgcactt gagtcactat ctcaagcaca 141660

ctaccatttg aatttaaatt tctggataaa aatctccaca aaactgataa ggcacccttc 141720

aagcaagaca caagttaatt ttcagcctct aatagaagta tttgttggca ctacctttca 141780

ttggtttatt catttcagca gccataaaca cacagatccc attgtcacag aaagccttgt 141840

tcgcaaatat aatttagtgc atctcacaat atgagctttc aggtggttga gctgtgtgga 141900

aattcgattt ttttttctcc cgagactctg aaatgaaaag aggataaccc cagtgctctt 141960

gcccacatcg cttatctcac aacaacagat tacaatccca taactccacg ttcccaagca 142020

gctactgtgt tgatcaaatg taatgcaaca gatgagaaaa catcagctgg ttcatgccca 142080

agacagctga aaggtgcact tacattttta gtctccagat accaaaataa atctgttggg 142140 cttaatcatg agttgtcctg ctacttttgg ggaaagaaag ccatgtccag tgaaatttct 142200

tttatactgg caggcatttt gactcttcac ctattgctgt tgaatctcca ggatatgtac 142260

atgttttcac accgatacct tcattgaaga aaatacagca atctgatatt tttatgtgaa 142320

tcatctttgg aattcgtgtt cctcaccttt ggatacaata aatattcaag agaaaaaaaa 142380

tacaggaaaa aaataattgc cctcaatgat gagttctgaa atacttataa ggataagtga 142440

tgtgttgaaa aattctcaaa agcaatttta gaaggaataa tttttgcaga gagtggctgt 142500

ctcttgggca ccataacaag aaaaatagta tattgccctt ggtgatggtg cggtcaagta 142560

tcaataacaa cacgcttcat attttcatga acaaaacact tcttcataag gtaatgagat 142620

tatagaaaaa tcaactcaag gtaaaaataa ttactttgca catggacaaa ggataagtgg 142680

aagaacttat cttctagttc gtggttcatt taacaattaa ttgctccttt ttcaaatagg 142740

tccactgtgt tctgggcctt gaacagtgta cctctaagaa agtttgttat ccttactgaa 142800

ggaaattaca gttaggtcgc ttctcagcct tttggctaag atcaagtgtg aaaattatac 142860

ttagattact atgtaattag gcctggtaga tcctgggtag acttcactga atgcaggatg 142920

caatatttgg aaaagtactt actttggagt cctaatgcct gggtgctact tctattagtg 142980

ccagaaatta atagtgtgac tttgggataa tcacttaaca tctttgagcc tcattttcct 143040

tatcttcaaa aaaaaaaaaa aaagaaggaa tcagatgaaa tcatctttac cagattaggt 143100

ctatctttag gtcccccagt gtaactgggt ttctgtcttg gtctcagtcc atccttgcat 143160

tatcctggac tttatttatt ggaatcccaa tatctgagaa ttcaatttat gtctttatcg 143220

tactgaagtc tataaagatg tttttaaaga gtcactagac tccagtgggg caacttgctc 143280

cagatatcat catagtctca ttaaaaagta tttattgact actaatgtgt ataggataca 143340

tgagttggag actcacagat acacatatat ttaattagaa aaatatgtaa gaatgaaact 143400

agccaaagat attggcagaa agtgtgtgaa atttttagct tggctggcct ttgttggaat 143460

gctctatttt ttattttttt gtttgtttgt ttgttttttt ggtcttggct tcttccatga 143520

tcttgaccct gtgccttgaa acttttttta aatcatgaaa aatgtgtttg tcttcaagct 143580

gtccagtgtg gcaatatcac ttagaaatat ttcttttta aagagtccca tgtttttaat 143640

atgatagcta aaataatatg atagaaaggt accataaaaa taaaattata tgcctttata 143700

tgccttagac aagtgtaatt cactataaaa ataagtttct cttagcagca ggaacaaaga 143760

aaaagtaatc tgttgaaaat gttagatagc ttgctgaaag gaaaaaaatt aacggcaaaa 143820

accaaagatc ttgaatttgt ttctaatttg ctcaacagca tttactgcga acaaaactaa 143880

ttttaattca gggtaattat tttcttattc cacatcaaat agttaaatat cacattttta 143940

agatctttat tcttatatct ttcctatgtg ttgaatcaag actgttcttg ccctgatcag 144000

ataaaaatca ataatattga gatcaaagtc ttcaaacatc acatgctaat ttgaagttgt 144060

atgcaagcct aaagctagtg tgtaatacct ttgttctttt cctcttaggc atttatgtcc 144120

atgttccaga tcctcaccca ggaaggatgg gtggacgtaa tggaccaaac tctaaatgct 144180 gtgggacata tgtgggcacc cgtggttgcc atctatttca ttctctatca tctttttgcc 144240

actctggtga gttcagcatt ttttttttcaa ttatatcttg gaatttataa tgggatgaac 144300

atcaacagat atttggccta ttgtcaattt caaaatggga gccaaaggac taatgcagac 144360

taatctggtt ggatgagatc cctccttcca ttctaagttt ctacacatca aagaactgct 144420

tgatttgtta agttagatat tttctccatc aggtaccagt gtttgtaaga gggaaaacaa 144480

gacagaacaa aatgaaaatc tcacatttac ataccaattt aagtatcccc aagcattcac 144540

acctagatga tctcaggatc aaaccggaac tccattatta ttttttttct catgtagaga 144600

cactgtcttc agtttcactg catcctcttc acagctgtaa ttaaactgaa gacagaagct 144660

atagacctca gggtgcagtc agttaattag aaatcagcac agtggttttg tgctatgtgc 144720

gactggaaaa aaacacatgg aaaaaaataa aacacattgc tgcaagtatt gtaaaagcac 144780

aagaaaactg aataccaaag tccagtaaga accgacttct gtcacacaag cagatttttc 144840

tcttcaatat acttccatta aaaacaaaga aaaaatgtta tgaaatgcag cctgtgactt 144900

actaaggaat gtaataccaa atgctcatct ctatgtaata aatttgttca tatttaagaa 144960

gcaaacctaa gttctaacat cacgttagaa gttccaagtt gaggactcaa atgtttccct 145020

ttatcaggaa tacaagaaat aaatacaact gctactattg tttttggcaa tatgtcttac 145080

agttacaaaa tatcttacaa tattattgct aatctttgga acacccgatt agatagtatt 145140

cccagttaac agatgataaa actgatactc agaagagttt agtaatttat ctgatgtcac 145200

atggaatgta aatatgacca gagtgttccc aggtctgtct gatctagatt tggtccttag 145260

tcttgcagct aattattatt gcttccctgg ataatgaagt ttaaagttgt ctgaaaaaat 145320

gtctttgtca aatcaaactt ctgccaaaga atcactcaca tcaaacaaat tacatataat 145380

ctctaaggtc agttttaatt ctaaatatct atgaatgaaa taagtccttg ataaactaga 145440

tataacttag aagttgtgtt tgctatagac agacagaggc tttcaaaact aagacatgtc 145500

caccaaaaag ttaatgagaa ttcattgtta catggtgtga agtcatgggc tcaaacagta 145560

gaactacata gacatagttt tggccttttt atacatggag ttgactttta aaagttacgc 145620

gaaagatagg aacattgtgc aataggaatg gaagcctggc ttgcaaagct gaatctggct 145680

gtgctggtgt ttgcacaaca cttagccctg gggtctttgc atcttccctt ccaggacctt 145740

taagcctgtg ccatttattc attcatccat tcatttacac atcatttatg cgttcaacta 145800

agtttattaa gctacttgtg tttaagttac tctaaaacta tggaattgct cagactgttg 145860

attgtgggtt aaatgcagga tagatttcat tgaaatcatc tatgtcaggg aatatcaatc 145920

taattatttt atagtaatca acctaatcat tttcctaata aagcaacaaa atggtttact 145980

ttcctcactg aaaatgtttc acactcagat atatgtttga ttcttaaatc agggacacaa 146040

aaaatgaaag ctggaagaaa aaactagcct attacattat ttagtaaagt aacatttttt 146100

caagtgaaaa ttctactggg gtcactaata taataaaata aaatattaaa actggtatat 146160

agccttcttg ggcaatctaa atcttcttta tgggctaaaa tatggcattg gttatagaac 146220 acttaaagtt catccaagga ctcctggggt tcacctgagg ctttaggaat ccacagagca 146280

caagatgaaa accaatctct aattctactc tttccactta ctgatttaag tatctaccca 146340

aaacatctgt tctaactttt ttattgcaat tttcaaacaa cagttataag ttaaaactga 146400

aaattgaaat aacttgttct tttattcaca ctttttaaaa gtctgtaagg caattgagaa 146460

ttcagaagtc tacatgttaa cgagaacagt ttggtaaatg tggtgtcgta caagactacg 146520

aaattggagg gcaaaaatcc agattcaggt tctagctctg ctggtcactc gatatctgat 146580

ttctgaagac tcacataact tccctgatct ctctgttgtt tgtatcggaa agtgaggata 146640

ataattcaca atacacccag tttccaaggt tgtcatgtat ctgcgagtta accctgtatt 146700

tctgaggatg ttttagagca ccagcactga gcccctttaa atgtagataa gacggtttgt 146760

gcatttaaag ccatatatta ataaatgtta aatttatatt atgttcagat atttttaatt 146820

tatgcaagta tcatttatta gggtatcatt tatgcccctt cctgaattaa tttccattct 146880

tcttctttct ctccttccct atacctcttc tcccaggcaa tgactctgat aattttaata 146940

taaattcatt ttccaatttt tagaacatat ttcactaaat ttaacattct gtcattcagt 147000

agaagactca gtttgtgtaa cactaaggag aaaaacaatt aaactataca catagaggat 147060

aaggtgcatc cctacttcag agtttttaaa aggtgaaagt aaatgtgcat cttaaaatca 147120

atgaaatgca gaggtttccc acacatatat gcatccatta aaaatatagt atttagggcc 147180

agcacggtgg ctcacgcctg taatcccagc actttgggag gctgaggcag gtggatcacc 147240

tgaggtcagg agttaaagat cagcctggcc aacatggtga aaccctatct ctactaaaaa 147300

tacaaaatat tagccaggca tggtggtgag cacctgtaag cccagctact caggatactg 147360

aggcaggnna gaatcacttg aatctgggag gtgggggttg cagtgagccg agatcatgct 147420

gttgtacccc agcctgggca acaacagtga aactccatct caaaaaaant atatatatat 147480

acacacacac agacacacac acacacnnnn gtgtgtgtgt atatatgtgt atatatgtat 147540

atatgtgtgt atatatgtat atatatgtgt gtgtatatat gtatatatat gtgtgtgtat 147600

atatatgtat acatatagta ttttaaactt aaggttacca tctaaataca ttttatgccc 147660

aaaattctga agattaatat taagttttct tagtttgcaa agcactgctc cagtcaactg 147720

atggccatcc ctaggggagt ggtctggtaa tacatacttc tctcccntat aagtccacgc 147780

tggggccaag ttatacatgg tcatgtgttt ccctccttat caattatttg tagcagcaat 147840

ttttagaaca aatggtatca aactcaacat atattctgca gttaaatttt tcactggcta 147900

tttttttctg agatcttttt atcttactcc acatgaagct aatttggtca ttgttataga 147960

gagccgtgtg gtactccgcg gactgtgaat atgccatatt tgtttatcag ttccctttct 148020

gatggctatt taggctgttt tccaaagttg tacttttata atgctgcaat acatttatgt 148080

atcgtgccca cttatgtgaa tttctctagg atatatgtcc aaaagtagaa tgcttagatc 148140

ttaggtactt cccaaattgc tctataccag ggacctgcaa gcccagatac caaatccagc 148200

ccaccacctg tattatacaa cctgtgagct aaaaatggtt tctataattt agtggttata 148260 caggtaccca cagaatatcc tcagtattgc ctcatgaccc tcaacgtcta aaatgtttac 148320

tgtctggacc tttaagaaaa cactgtggac tcctgctcta tagataaatt gtgctaattt 148380

accctcccaa gtgcaatagt tgagaattcc tgtgttctta cactctcact tagacgtcac 148440

attgtttgac atgtgaggtt ttggcaatct tatgagtgaa cactggtatc aattgggctt 148500

gcatttttgc atttccataa ttattaataa gggagcacct cttactatgt ttattgacca 148560

tctttttgt gactcaattg tttgctttac ctattttcct aaaaacaaat ttgtcttaac 148620

tcatttattg gatatttgtt ggatattaat ctcttgacta ttacatgtgt tgcgaatatc 148680

ctctcttagg cagtattttg tcttccagtt tgtctgtgat gtcttttgac attctatata 148740

tttttatttc atttttagtt aacaaataat aattgtatat atagtagtcc tcctttgtcc 148800

ccaagggaaa ctttccaaga ttcccagtag atacctgaaa ccatggatag tacagaaccc 148860

tatatatact atgtgtttta ttatacatac ataactttga taaagtttaa tttgtaaatg 148920

aggcatagta agaaattaac aacaataata acaaaataga acagttatga caatatactg 148980

taataaaagt tatgtgaata cgcattctct ctctgtctcc atctctcctt ttcaaaatgt 149040

cttaatattt tcaggctgtg gctgaccaca gggaactgaa actacagaaa gtgagacagt 149100

ggataagggg ggggcctact gtatttatgg ggtacaacgt gatgttttga tacatgcata 149160

cattgtggaa tgatcaaatt aggtcaaaac tatcggtttt tcactgatgg tctgtattta 149220

ttttgacctc atgcttaagg aattctttcc cacctgaggt aatacaaata ttctcctata 149280

tcttcttcta attgtttaca atttttgctt ttcaccttta agtctttagc tcatctctac 149340

ttactagtta gatatagcat aagctagatg aaatatactg tttaaataaa actcaaatca 149400

atttttctca gatttatatg agaaactgca aggtttagaa tatttatttg atttatctga 149460

ttatgaagtc taacacaaaa aaggtaaact gttcaaactt ttaaatttct aaattatttt 149520

ttaaatgata tataatttaa taaaaggaaa cactatataa ctaaattaaa acaaagatta 149580

tcacttctat actacccaaa aaaagtctaa tactgaagtc acttccccag ttcacaaaaa 149640

attttgctta aaacatggtc agcagttgtt tctcttttcc tttgtaccct tagactaaca 149700

gttactgtaa tgggctttat gaagtgagtt caatttgatt tgatacagtt ttactgaacc 149760

tctactacca cgtgcaggta ctaacctacg tcctggccaa agtgggaaag agcttaattc 149820

atgtcctctt aggggctaca aataattgga aattagcaat aacgaacaca gacaggctaa 149880

ccaacaaaca aataaaacca gtggttccaa ggcctaaaga tgcaaatggt aagagaggca 149940

tgtgtacatt gctacttgct atgtgatcaa taaacatctg cttcccagtt gttcccaatc 150000

tgccttatga gaaatatcaa aagtagtggc agtgaaacag ctggagatag tctgtgggct 150060

aacatataaa cacatgcgtg agggtgacag gtctgtgtta cacacacatt gcatgcccct 150120

aaaaactaaa cttcacaagt taaaatattc cataagcctg tcttcagagc tgaacttacc 150180

ttaaaatgag catttttgtt taattgtatg tttcctttta caaaatgcta caacgcattt 150240

taacccaaat gggtgagaag ggaaaaaaaa agtagtctgt gtaaacatac aagctctgat 150300 gtgtcttttg ttcacaccga acagtggtgc catacaactt gatgagttcc tgatttagtc 150360

ttcccaggga caggcatgtg ccaagcagcg aaaaataaaa tgtgtggttg cagcttgttc 150420

tggaataaag aacagcaggc tatttcatca cccctggggg caccaatttt cagcccacaa 150480

tgaaaatcaa gctgaagtat atagcaaatg catattcata acctccatgt attgtccctt 150540

tcaaacatct ctctcttccg ccaggtatgt attctgtatt tctgtcctga ctgattgcag 150600

tttattgaat gcttgttaaa ttcatctttt tccccatctt tgccaatgat ttggggatag 150660

ggactatgga agacctaaca agttgaaaga ctccattttt cctgcaatgt ttttctgtct 150720

attactgttt caaaggaaag ccctgtgatg tgaatactga aaagaaagaa aatgtaatta 150780

agcgttcatt gtgggtagtc tttttgttac ctttgaagtg ctgaaattaa cacagcaatc 150840

tcttgatttt gcctcggtag tggcagcttg ctgagaatag tcatattttg ctgacacaaa 150900

tgggcattcc aaagtgtaaa ttaaataact tgattgtata gaattcattt aaaagtgcat 150960

ctgtgttgaa cgtatcccat tttacactgc gctgattaca tttgactttt tgttgaccca 151020

tataaaaagt gaaattgttt cataccacag gcggcaagac agcttaggat agggtgatta 151080

taaatggtaa acaggaaaag aaaaacacaa gagtgtggga cattgattcc tgctgtgttt 151140

ctagtgtata tttcaaaata atttaaagat aagaaattat taataaagtt cacagcaaag 151200

acgaaaataa gattaaaggt gaaatttcag ctaaagggac atctttaagc tatataaaag 151260

gaaaaaaatc ataaaattaa acttcataaa aagaaagaaa aatgaaaagg gaaagctata 151320

tagggtaagg tcatggctat ctgcagagtc agagagaatg taaaatagga tatgttgctg 151380

cagggagaga aattttcaac cagaaaatcc attgtgaatg acaaggctgc cctcttgcaa 151440

aaataaatca tacttttcag ctgtttgaat catgctaatc caaaagtctt attccgaggg 151500

aatgagaact acattatgtt ggacaaaatg atctgcatca caacaaaatt ggaagaaata 151560

attatacttg caaatgcaac tatttcagta atttattgaa tttgactgaa atttattgta 151620

gttaaataga aaaaaaataa atatatctgt ggtcacatgt ataaaactta aacagagcaa 151680

ataactacca cagncattaa ctgaggtgaa tctgggctgt ttcagaggna gaaaaaggaa 151740

gaggaggtat tagagaaaga ggagcgtcat aaagaatatt agagaaatca aacgcctgtg 151800

agaccctggt ccaggtacca gtttgaacct ggaaaattac gcttttgaaa tccactaagc 151860

tgaaaggtca aaatcctatt atatcattca taaatattct gagagagact tcccagactt 151920

cagcaggcct gataatgtgg cattttccac actgaatgca ttgtcaatgc taaacatacc 151980

ttagcaatgg gcattataat ctgagggaaa taaagaatta gatatctgaa atacattgtt 152040

attctccact ttttcattga catcactgat tcataatgta tcaatatcgg tcataaaatt 152100

aaaaagtgga tacatacatc agctgtatat tacacagtag gacaaatgct gtgttgcata 152160

ttaatatgat caaacgctat agacctgtta aaagtgccaa atagtaaatg catgaaaaaa 152220

tagatatttg gctttctttt ttacagcatt gtacaaacta aacaccactg tgttgaaagc 152280

atacatgtac ttagcagggt tcaagaggag atgaaaggtg gtgtcattag ccttttttccc 152340 ctgaaattat atcattaaaa aaaaataggg acatagtata tcctgctgca gtataaattg 152400

aaatatcggg ccaggcgtgg tggctccccc tataatccca gcactttggg aggccgaggc 152460

gggtggatca cctgaggtcg ggaattccag accagcctga tcaacatggt aaaaccctgt 152520

gtctactaaa cacacaaaaa ttagcctggt gtggtggcag gcacctaaaa tcccagctac 152580

ttcggaggct gaagcaggag aaccacttga acccgggagg tagaggttgc agtgagctga 152640

gatcgcacca ctgcactcca gcctggggga cagagagaga ctctgcctca atcaatcaat 152700

caatcactca atagaaatat cacataaatc taaatattaa gaaataaata atgcaaagca 152760

aattagttgc ataatttaga acagtcaaag cacaaaggta aaagtgtttt atacatatga 152820

tgacactcct gtgatttaat gtggcaagtc cctgtaaaat tcctttatta catcaatacc 152880

gttttaatat gaacctgtaa atcacattta gtggctatag aaaaaccttt tggcacagca 152940

acatgtggta gtgtaggtaa acattagctg aaggtaatat gcagacatcg ctgtaatggg 153000

ttctgattac taatgagcat tacctcagct ttctgtaaaa acagaataaa tgagaacact 153060

cagagaactg aaaaatggtt cttcctttca gaatgagggc ttgtcaccca ctcccatgac 153120

tttccatacc atgtgttggc aagtgtctga cacccaggtc gacctgctct ttcagtccaa 153180

atccacattg cactgggttt gtaggaaaat actgtatcat tgataacata aattaaatat 153240

gatggacacc aagctgccat tcttccttca cttctaccct ttatcgcagt gcttattgtc 153300

accttattag agtctgtgag aaaagaaccg caaagtcatt ctcacctcct ggcaatcagc 153360

tacagactca ccctgcacct cctgaatctt ctgtttccat tgcaaagaca cagccaaagc 153420

cttgggtcag acttcagtca tctctcaccc ggagtgccag ttgccacctt atcacagtgc 153480

cctacatgcc agggacagtt tgatgaggat caatctgaaa gtccacattc aaatagggct 153540

tcagcggaaa cccaagcccc ctaccacaca catacatctt gaaccagttt ctgttttgta 153600

agtttgtgct tttgtcatgc cactgtcaaa ttaattgaaa aatgatgcct aaacaaagat 153660

atttgctgaa taaaattagg tttatgctac tttcaagcaa acttcaaact ttattatatt 153720

atttttaaat ttgttgcctc cccaagagtt tgttctaatt ctttaagttt gggccttcct 153780

gttgtgaagt cacgcagcgt tttcccagag aagccaaagc tcttccttta ggagtcttgc 153840

cttctcctat ccacagcaaa tatcaaatgc atctgccttg aggagtcagt gcagctggca 153900

tgccgtggct ctctctgtcc tgctctgagg gtggcttttc tctgtttatc ccaaggaaaa 153960

agtatatacc agtttcttgt ttaaccaatg taagaaaaca ttacatatct tttttttttt 154020

tttttttttt tttgagatgg agtctcgctc tgttgcccag gctggagtgc agtggcacga 154080

tctcggctca ctgcaagctc cgcctcccgg gttcacgcca ttctcctgcc tcagcctcct 154140

gagtaactgg gactacaggc acccgccacc acgcccggct aattttttttg tacttttagt 154200

agaaacgggg tttcaccatg ttagccagga tggtctcgat ctcctgacct cgtgatctgc 154260

cggccttggc ctcccaaagt gctgggatta caggcgtgag ccactgcacc cggctagaaa 154320

acattatata tctaataaca ttttgtcaaa ctgcacttaa ccttgtctag ttactatttc 154380 agtgagccag tttatttttt tgttttgttt gtttttagtt tttggagaca gggacttgct 154440

ctgccatcca ggcctagagt gcagtggcac attcacagct cactgcagcc tcgacctccc 154500

aggctcaaga aatcctccca cctcagcctc ctcccaagta gctggggcca caggcatgcg 154560

ccaccatgcc tgactaattt ttgttttttt tttttttga gagacagggt tgcccaggct 154620

ggtctcaaat tcctggggtc aagcaatcct gcctcagctt cccaaactgc taggattaca 154680

cgcatgagtc accatgccca gctgagccag tttaaacttg ctgcattaca tgcattaaaa 154740

acaaaaacac acacacaaaa agccccagaa cttttaaaag gaaaaattta gcagaaagaa 154800

actgatgtca tcactttgtg attattccag cctccctctc ttatccgtga cctctgtggc 154860

actcaagacg cctgtccatc ctgccttcct ctccagcctc caaaccctct gcctggaagc 154920

cagactctgg ccagacttac tcaccaggcc ctcaccacac cacctggaat gcgttcctca 154980

cagctctttt ccaaaatcag tcttctcctt ggcatctgcg tggcttggga gagcatcctt 155040

ggctctctgt cattgtctca caatcttcga taagtagtgt tcggaaagtg ccccccctggc 155100

ccatgtgccc agatgggctc tccaagtcta aatgcttcaa aatggtacag aactaaatgt 155160

gtgaatttat gatggggctt catttctggt cctccttctt cacacttgcg tggccgccct 155220

tcccatccgc tctccaattt tctttcacgt ttctgctttc tgagaagaaa taagactgat 155280

ttctgggatg gagcatgaga gagacagatg cagagagaga gagagacaaa gcgtgtttcc 155340

gtgggagatg cccacacacc tcacttggga gcatcggcag tgttggtacc cacacaggtg 155400

tgttttgaca tccaagactt gaccctaatt cttcttggaa cctgatgggg agagaacctt 155460

gctgggggcc ccacatgggc caggggcagc cagggcagag ttgggcagct gcctgagagt 155520

tctcggtgga gaggcccaga atatcccaca agaacaagtc tccgttagtt ccaagtggaa 155580

tgacagccct ttctgagttt ccctgagttg agagctaata catgtgttgt gggaacaaca 155640

gcccaagcca aaggccacgc ggaggccagg cgtggtgact cacacctgta atcctagcac 155700

tttgggaggc caagatccac ggatctcatg agctcaggag ttcaagacca gcctggtcaa 155760

ccttgtctct acaaaaaata caaaaaaaaa aaaaagtagc caggcatggt ggtaggtgcc 155820

tgtagtccca gctacttggg aggctgaggt gggaggatcg cttgagccca tgaggtggag 155880

gctgcatgag ttgagatcac accactgcac tccagcctgg gtgacaatgt gaggccctgt 155940

ctcaaaaaaa aaaaaaaaaa aaagaccagg tggagagtat agtggattaa atgggcccct 156000

acaaaagata ggaccatggc ttaacccaca gaacctatga acacgacctt attggaaaaa 156060

gagtctttgc agatgtaatt aaattaagga tcttgtgatg agatcatcct ggattatctg 156120

ggtgggccct aagtccaatg acaagcgtcc ttgtaagaca caaaagagca caagacacag 156180

agagaaagcc ttgtgactgg tgttagataa aggccaagtg actggtgtta acaaagccac 156240

aagccgaggg acagaggagg ccaccagcag ctggaagaga caagggagta ttctcccccg 156300

tggcctttgg agagagtgtg gtcttgccga caactcaatt ttggactctg gcctccagaa 156360

ttgtgatcaa ataaatgtct gttgtttgaa gctaccaact ccgtgcacgg tgttacagta 156420 ctctaggaag cgcatatgga ggacatcact gggaatcaga gctcttagct ttgggagctg 156480

agatggtggt caagtcagca gaaagaaagc cccaaacccc ctgtacagcc tcatcagttc 156540

ccctgatggg ggatggagga catattaccc agcttgtagt gggtaacctt gaccttaaac 156600

taccctaaaa agaaaaggag catttattca aatgtaactg agcttgaaag gactgaatgg 156660

aactgatgta gcaagataat gtttccagcc cccaacagaa atggagggat tgtgcaccaa 156720

attcaattca tttgtagaaa aaccaagaaa tttattttag ctttgcataa ctaattttat 156780

gtctgtatat tttcagacat tgtatagcta tactacatgt cagggttgct cggggagaat 156840

taaatgaaaa aaaagctgca atttgttgag ctattattac atttgtaata ttttacaaac 156900

gagagcacca aaactcactg tgggctggca gttgccacgt ggctcatagc tggcacagcc 156960

tgtttgctac ctgagctctc tgactccaag cacagtgctc tcttctatgt ctgcttaatt 157020

cagtgataaa catcactgca gtgagtgcaa tattgtattt atgtatcaga ccaggcagtg 157080

cctcacacat ccatggagtc cagtcaactg tttgtcaaat ttaatagaat aacatagcct 157140

tgctgtcctt aagtagcctg ccctctgtcc agtttggggg atataattaa gaataaaatc 157200

tcctgccaac ccacgaaacc tctccttaaa ggtagaggag agagagccca gttttattat 157260

tgaaaaagct ctaacgcagg ctgcaatgca catcactggc aatctgctaa agagatcacg 157320

gaggcagaaa atacctccct cttttatatc accgctttcc atacgtgttc tcaagataga 157380

aataactagt cctcaagtaa gaaggcttga cagcaccact tgttaaatat agtttatcct 157440

aaattcacct ggtaagtggg gtggccatct gtgggagctg attgccttta tccaaaggaa 157500

aattaaaact tctcatattt ccgtgacaag taggtaggtt tacaacttgg agccagatgt 157560

ctaaactccc accctctacg gagaagggga gacggaggct cttgtcttcc ttgatgttta 157620

cattacttca aggagacagc cctcaaggaa aacattcttt taggttgtag aactggtgag 157680

aagcttattt ggctcttaaa agtatttata tacatctcaa acagaaaggg aaagaattta 157740

caattccaag tcttataaag gaaatgctct aagaaaaggg aggcgaacag gatcccttcc 157800

gacccccctt tttttagatt catgtttgta ctttcactag agagagagga aacatgatta 157860

aattaggaca cagaacaatg cctgctccag tgcccaggtg tgagagaagg ggtataagat 157920

atcccctcct gcagaaggaa accagcaggt gaacatctca gaattcgggg tgagaaatca 157980

agacagggga aggtagactc tgcctctgat attactttct gtcgtcgagg tattgggaaa 158040

tgggctgaac atggagcaaa ctttaaaatg tccagtttgg aatgatgtct ggttcagaat 158100

aagcctgatg cccagttacc agtggtgtac cagtcatggt gatgttaata aaaaccctct 158160

ggtctagatc attctgtatc ttcttcaagt tatggtaaga ctccaaaggg catggctgtt 158220

gaaaagaagg gtgctggcat ttgtcccaag ccaagcatat cccatcataa tttaatcatg 158280

ttgatatact cattaaaatt gaaaaagagc tagccattaa cttctgcata ataacttctc 158340

tagcactggc tgaattatct caactttctg accttcatgc ttaaccacct aatattaaca 158400

aatgggaata attcattaag aagaaaactg aaaagaaaaa tatgagttag aaatgaaaaa 158460

```
gatgatcaag tgaatatggg acttaaaaaa attcaaagtt gaagagggac tacagcgtaa 158520
aatattttta aaaattatac tattttgttt tattttattt acctgttcta aaagaagcat 158580
acatttttta aaaactgaga aaatcataac attttttaaa agaggaaaaa ttgtattact 158640
ccaatacaca aaagcaagta ccattaatat ttctaaagta cctccttcaa gcattttctc 158700
tattactccg ttatatgcat agttgagctt ctaccatata tatatatata tatatatact 158760
cgtatttact gttataatgt gttttcaaca ccataaaaat tcaccagtgt attaggtatc 158820
tgttgctgtg taataaatta cccacacatt agaagcctac caggaaatag atgtttatta 158880
tctcaattag tttctcaagg ttgggaatct gagacaggct tagctgagtg tttctagttc 158940
aaagtgtcat gaggttgcag ttgagatgtt gtctacagct gcagtcacct tgaagtttta 159000
ctagggctgg agtgcctctt tccaagatgg gtcactccag tggctgttgg caggaggcct 159060
cagttcctca ccacatggac ctctccagag ggctgcttga gcatcctcac aacatagcac 159120
ctgcttccac agagcaggtg atctaggaga gaccaagggt aaagacacag tgtattttat 159180
tgcctagcct cataagtcac acacaattat ttctgcaacc tgctgatgac acaggccaga 159240
catattcagt gtggaaggga actacacaaa ggcatgaacc aggagtctag aacaacgctg 159300
gaggccatct tgaagagtag ctaccacatg ggaatttgta atggctgcat gtatcttatc 159360
acataaatct ctataactgt attgactatt catgcaaagt tggatgttat ttctaatatt 159420
tccttattac aaacaaccct ctaataagca cctgtgtata ttaagctttt catcaatagt 159480
tcagatcact tttttagaat attatcacac aagtgggatt gttgagttca aaaaaacctt 159540
tttcttaatt tcctttttaa gtaagttggt ttcataaagc aagttaacat attatttcaa 159600
tgaacgtatg cattttattt catgtatttg tgataaatac cttgaaatgc ctatgcaatg 159660
attcatcact gcatgtttat tttgcattta atgaattata tactcaatat attttaggta 159720
gctatattat acatataata ttttgatact tatggcatgt gggttttttc attgttattt 159780
caccatttta tagatcctcc tgagtttgtt tgttgctgtt attttggaca acttagaact 159840
tgatgaagac ctaaagaagc ttaaacaagt aagtagtgtg tgggggattt tttttaaaaa 159900
ggacaaaaaa ccaggggcaa accatttgct tataggcaat attaaaaatt tgagcttcaa 159960
agtaaaaata aaatgtaata aggctatcga aggtggtaaa atggtgttaa gtttaacttt 160020
tatctaatat taacaaggga aactttctat tactctataa atatcctagg aatagtttta 160080
ttatttttga ataaaagtaa aatgggcaga gtatagttaa ttggatttgg agactgttct 160140
ccttctgcca ttaaatgaag gttgtttgga aatgagatgc ctcaaggcca gattttgcat 160200
aattatatgt aatagtcttt atcatcatca taaaaatact acaatcaacc tgcctggtct 160260
ctagacacta aaggacaatt tagtctgggt gcccatagtg agtataagaa acagaataaa 160320
tgatcaaatg tggacccagt tatgtatgaa aacagacggg atgtcatcat ggtgaaagtt 160380
ataccccaga gaatgaaact agaaataggt gaatgctgtg ttcttaatgg caggaggatt 160440
acagctatcg tcaatcattc ctatagcttc tgttgcatgg aggatatttg gcttgaatat 160500
```

```
ttggctctat caaagagtac agaccctttg agaagaagag ctcgcaagag agtgggacaa 160560
tgggtggaac gttcttgtat caccctcccc aaaactgaga gaatagttga gatgttaact 160620
accaaactct acccaactgt tagaatgtta taatgttgac agaaccacac ctgttataat 160680
gttaacagaa ccatggctct gaaactttcc atttggaatt ataagcaggc cctccatatt 160740
cagaatttcc caaagctctc tatttttaac aaagcaaaaa gtataaccta tttaaatttt 160800
ttaatttttt ttaatctatc tcatgatcag gttaggcatt tcattttatt actcttacct 160860
ttaaagccta tgcttttaac agagggtatc aacaaggaaa aaaaaaaaca gaacaaacaa 160920
acaaaaaaaa attgttatga agattcctac ttggagcata ccaagaaagg ttttgagttc 160980
tttaccatcc tacagctacc caggcaaatg ggcatcttcc cggctctgct agagatgatt 161040
tcttccccgt cttttcactc tgtggcagca gtgatagatt cttcaaatga gagttgggac 161100
attcctttat gccatcaatg aaaatacttt gtgtggcttc tagttgcaga aaaaaatatc 161160
taataggtga attcaagatt gaaagtgacc agccatataa ttagacatga tcaataaaaa 161220
tctgctgctg tagaaccact ttagtccatg ctttacaaat ctggtattga gctgtcactt 161280
ttccttcatg ccatggttat tatgcacaga tatgtactat cctctgagac agtatgttac 161340
aatgtgtatt ctatagataa aatttaaaat atccaatttt ttttagctag tgctaccgcc 161400
aggttccata tgcttttaaa taaccatata attccagccc tcatttgtaa aataattttt 161460
ctgctggaaa gcaacaaaat aaacttgtat gtagcaattt tttcaggatt gtgaattcaa 161520
acctaaggat actttgcagt gttcattttt ttcaacattt taacttattt atatatacac 161580
gtatatgcat agcctactta aattttttgt catctttata tctagtggtc acaataggca 161640
tttcatatct gacccaaagt gaagaagcaa caataaaaag tcaggacttt aattggcatt 161700
gacattcttc ctgggctata aagttttaat aaactatatg gcccttcact ttgggtaaat 161760
gtcattattt ataaaatggc ttttaatcac ttgctagata ttagtgaagc tttgtgtttc 161820
tggaaattct gtatatcaga actggttatc caattgcact agcctacaat ttggaccata 161880
tactctgggt atttgggcat tgaaaactcc aatcataacg cagcaattcc atttttgtta 161940
gtacttttta cactattctt tattcatggt tggactgaaa ctatgtatct attaaattct 162000
tctattaatt tcatacttga aactcgttga acctgaaatt cattatagca aatattatta 162060
aactaagtta aagtgttcct ggttaacttt ctctcatttt tgtatgatta tgttttgttt 162120
tcagtagttt ttaacaacgg tgtcattttt tatgcccctg atttttctta ttttcagtga 162180
tcttggcctt taaggtaggg ctcatttgca tatttttaaa aacacacaca ttctcatgca 162240
ttcactttca tatacacatg tggtcatgtg gatacctgaa actttagact gtcttttaca 162300
gattctttta aaaaacaatg aaattggtac cctaaatatt tgtttcaatt gtccatgtac 162360
aaatccagtc cttaagattt aaacttatgt actctatttt tactactccc agatagagag 162420
tagagtggag gaaatgttca aaaaaagtat ttgattttct taatgtttca agtttcagaa 162480
agcgctatgc attccaaatt ccatctttgc tccgttcaaa ctttttgctt cttttaccac 162540
```

atctgcttgt atgtagattc tcatagctat ggtgacattg gttgaattct tgctcagtat 162600
caagtgtttt atgttttcac ttatttttca caacgaccca ttcaatgcag gtgttattat 162660
tatgtccatc ttacagatga ggaaactaag acatagagag gttaagtaac ttgctcaagg 162720
ttacacagcc aacaaatgtc ggagctagaa tttgaaccta ggcaatttgg accgaaagcc 162780
cacactctta acctctaaac tgtattgatt caaactaaag ctatataaag gccaaacatg 162840
agtcatgcat actggggagg gcactcatag gactatgaat tatttgttta taaataatga 162900
tgttgtagat aacaaaggca tcctcatatt caagtaactg ttagaatgaa ggacagtttc 162960
tgagaatagc ggtgaatgca gcaacccaat cgttaaaact attcacatta gttctttgtt 163020
tcctctcaaa actcagagga tctaatcaga gccactggag ttacagtctt gctgcgcgat 163080
ggcaaacgcc actcaaggtg ccactttgca natgatcttc ttggagtgat gatcctcctc 163140
agacgcctac ataacaggct tgcagcatag cagcatttta gtgactttaa tgtcttttat 163200
tatcgatatg tttttaactg tgttaatatg taaaagttgg ttcatagatt tttttcaaga 163260
gtaataaaaa ccagggcctg ttcgctctaa aatgggggtg ttggccaggc gcggtggctc 163320
acacctgtaa tcccagcact tggggaggct gaggtgggcg gatcacgagg tcaggagatc 163380
gagaccatcc tggctaacac ggtgaaatcc cgtctctact aaaaatacga aaaacaaaaa 163440
aattagctgc gcgtggtggc gggcacctgt agtcccagct acttgggagg ctgaggcagg 163500
agaatggcgt aaacccggga ggcggagctt gcagtgagcc aagatggcgc cactgtactc 163560
cagcctgggt gacagagcga gactccgtct caaataataa taacaataaa tgggggtgtt 163620
aacatgcaac actcagtagg gctcaacaga ggaagaggaa gcagccctca tacaggggag 163680
aggagaggcc gcccttaagc taagggtgtg agatggtggt agaaagtgaa aaggtactag 163740
tgaagaaatt ataatggaaa agaaggaaga agtgaagtag gaaaacacca gatatgaaat 163800
aaagagtatg actataactc agggataaag gtgcaggaat gtggccaaat ccagagttag 163860
gaatgtgagc aagtcagtgg ccaaaagtca aagcccacca gggcgttccc ctcagaagct 163920
ccctgagggc aggcatgggc tgcatgatgg attgtcggtg ggcccgtgtt ctcctggaag 163980
agaagcatgg cgctgtatgt cttttcccat tatttttcta tgcaacctcc gtgtgcatca 164040
gaggtggttt ctagccaagg actgctgagc tctaggcgtc ctcatgagaa gctatggtgt 164100
catttttaat acttcaggaa gctggaagag aattgtgcat tcacttaata taagcccgca 164160
cctgctaaat aaaatgtcaa gtttaatact tttagctgga aacctgaaaa ttctaagtgg 164220
cagcactgtt tccttaatac tatagttctg aagtgttttc atatatagtt tcctttaagc 164280
aaatttgttt aaagctacaa tgaatttctt attgcttaat cgaggcattt tggcaggaaa 164340
gaagtttggg ttttagaatc cagaggagaa agtaaaagga ctttttgcac agaaattttg 164400
agaggctctg gtgtgtttat tcccatcacc gcctctccaa gacatcacca gtgtgtgccg 164460
tgatggtcac ctctatggcn ttttcctttg tcacttccca caatcatgtt ccaaaaactc 164520
acctcaatca tataaccata tgaaattcct ctattgcctt ccctgtgtgt tatatatatg 164580

acacagctga ttgctaatag caatagggca gcgtaatctt ttcatagctg ccatcattta 164640

catgaccccc atattgaagt acgtagaaac catatttctc ttttcagcat ccccaaaagc 164700

caatacagag acctcggaaa aaaatatata tgtgtgtgtg tgtatatcta ttattacata 164760

tatgtaatat atgtaacata tatgtatatg tcatacatta aatacaaaaa tatattataa 164820

ttttgtatta catatatgta atatctatac atatatgtat atattaaata taaagattat 164880

atttataaat atatgtaata catgacataa attatacata atgtgtgtaa tatatgtaac 164940

agtatataaa aattatattt tttatattta gtatttatac acacatatat ataaaacatt 165000

gtttctgcat gacatttaaa aattggctaa aaatgagatt gtggtgagct gaggagattc 165060

ctgataaatc agttggtata gaggctacaa caaaaccctt gacctaaaac ttctctgaac 165120

agaagagttg atgactgtga agctcaagtg atgcaggaat gacactgggg aatgtacaca 165180

gagttttgca aggtttgaat atgaagagaa ttcttgcagc tttaatgatt tttgcaaatt 165240

tggtgaaaga cagatataaa gggttttctt ttaagtggta taaaattcat taagctttga 165300

gggaagaaaa aacccacaga aaagcctgag gggaggaatc atcgaaaagg gaagtatagg 165360

tagtctgttt agattaaatt cagagctgtg aaattgaaaa ctgtttaccc agccctgtat 165420

ccagtctcac ttgaatcctg gttccaattt aatgatgttt ggtctttgct agagtccact 165480

gcatagtgtt ggatagatat tatatattaa gaagctttcc tttcttattg actgtgactt 165540

ttttaaacaa aaaccataaa tatcttatat tctcttttgt taaagtatga gctctaggtg 165600

aatgtttttc tttagcctta agctgcatga tagttgattt tatgtgtcaa cttgactggg 165660

ctaagggatg cccagatagc tggtaaaaca ttgtttctgg gtgcatctgt cgggtggttc 165720

tggaaagatt tgcatttgaa ttggtagcct gagtaaagca gacctccttc cccaatgtcg 165780

ttaaaccact tccaaagctt tgaggtccca gatagaacaa aaaggtggag gaaagctgaa 165840

tttgctgtct gtctcattgc ctgcaccatg taaacaacct tctcctgccc tgaggaatcc 165900

tggttcttag gccttcaaaa tgttgccctg gcttccaaag gtctccagct tgaagagagc 165960

agatagtggg acttatcatc agcctccata attacatgag ccaatacctt aaaataaata 166020

ctgatgtata ttttattggt cctgtttctc tggagaatct tgcctaacac acctgaagct 166080

ctcattkgca tgtccatttt gtcgcctata ttgagttcat tccagattga gggcggtagc 166140

aaggtggcag ggatactgct ggaagatgcc attcggctga gcactggttg catctcctga 166200

tgttagttga atcttcaatc ttctcagcat tttgtcagtc tctcatgaat gaactcaaac 166260

taatgtaatt ttaaaagact gatataactg ctgcaggagc aaatgctatc tattgagaac 166320

taccttgtgt aggctrcttt atattaaatt atcacattta atcctaaata ttaaagacgt 166380

ggcaactaac attcagggga gtgaatttct cacccaagtc cacacaattc atatgtgcct 166440

gggtcagagc cacaaaggct gtctgaccct aaaccaatgt ccttgctgaa aaaggccatg 166500

ctgcctgcag agaaccagct ccttcattta taatgaaaag catatctagg tgtggattgg 166560

atgagatgct tatccatttt ttagatcagc atattttctc cctgtgggta ttcaaagaag 166620 aagccaaatt ataatcagac tatagcttag tgagccagga atcaggatgg agagggaggc 166680

gagaaaggtt gtgtgattta aatacataga tgatggtttg atagatgcag caaaccacca 166740

tggcacatgt aagcctgtgt aacaaacctg cacattctgc acatgtatcc cagtatttaa 166800

agtaaaatct aaaaataaat aaaaatattt ttaaaaatac ataatggcat aggctaagat 166860

aggctgagat caactgagtt atctctttta cagttgagat attagtgacc tagaaactaa 166920

atctactcat tctgtgaaac tggggtctgg acactcatct actatgctat agatattgtg 166980

ttggggctag ggacatgtga tggtcaggac actgtttcta ccctcaaaaa taactcctga 167040

gccagtggga gaaaaagaca agtaaacagg cagtttgggt tcagtggtga tgattataag 167100

gataatgttt cggaagcaca cagaaaacat ctaaagcaga ctgtgaaatc agggacgttt 167160

tctcacaagc aggtgcagac tgaaatctgg aggttgagta gtgctaatcc aggtgaagag 167220

ggttggagga agatggcatg aaggagggta gtgaagacca agggaactgc agagtctttt 167280

ttaggtggga aaaaaacaga gtggctttct cgacactaag agtttgctgt ggctaatgtg 167340

tagagcaaga gagaaaattg ttaactggaa aaataaggag ataccagcac aggaaatgct 167400

tgagcagcac taagcatttt gatattattc tgatggtatt gatgccacag ggtgtcataa 167460

tcatacctac attttggtgt cagtgtgaaa tgcactggag ggcacccaga ctggagatgg 167520

aagaatccgt tcttgaggca agtggtgatg gcagcctgaa ccaaaggaaa taggaatgag 167580

agaatatcga aagcggtgga gtacagagaa gaggatttag tgacaaattg gatgtaggaa 167640

gtgagaaaaa ccaggacaca gggatgtcac ccagatttct agatttcttt aagtgcgctg 167700

atgcagaaga aatacaagag aaagaagaaa tgagagtttt gtttatttgg gttttttttc 167760

gggggaagaa gtttggtcat attttccctg agatgcctgg gaaatatcta tgagacaatg 167820

tctatggaga gatacaaagt ttaaagctta ggtggaatat ctaggctgaa atagatactg 167880

aaaaagagca aaatcaacat gcagatggta atttaagcct ttagagagca ggaacttacc 167940

aagggaagga tggttcccta aggggtaata aataaatggt atgaaccagt gaggaggagc 168000

caaaatgtgt atgagaagtc acagcacaag aggtaagagg aaagcctaga atgtgtttgt 168060

caccaagatc aaaaggtgaa agagtttcaa gagagatcct tgttcaacga tctcaaagac 168120

ttcagagagg tcaaaaaaat agtcaccatg gaaagcaatc ccatttacag tagcaccaaa 168180

aagaataaaa tacttggaat aaacttagcc aaggaggcaa agacacatac cctgaaaact 168240

ataaaatact gaggaaagaa cttaagccac aaataaagat gaattctgac ttacgatgat 168300

tcgacttaac agtttttgtt tttgtttttt tttgacagag tctcaacctt ttccttaggc 168360

tggagtgcag tagcatgatc tttggctcac tgcaacctct gcctcccagg ttcaagtgat 168420

tctcctgcct cagtctacca agtagctggg attacaggca tgcaccacca aatccaacta 168480

atttttgtat tttcagtaga gacagggttt caccatgttg gccagggctg gtctcgaact 168540

cctgacctca agtgactcac ctgcctcagc ctcccaaagt gctgggatta caggcgtgaa 168600

ccacctaacc tggcaactta caattttctg attttatgat ggtgcaaaac agatacacgc 168660

```
tcagtatgct cctcaactta tgatggaatt atatccagaa acactcatag taacttgaaa 168720
atatcataag ttaaaaaatg cacttttaca gtattttcaa cttataatga atttataaac 168780
cctattgtaa gttgaggaac atctgcaagt gaaaagacat tctgtgttga agaactggaa 168840
ggcttaatat tgttaaaatg ttcataccat ccaaagcaat ctacagattc aatgcaatct 168900
ttcgtaaaat tccaatgacg ctttttgcag atatagaaaa gaaatcttaa aattcatacg 168960
aaatcttaag ggataccaaa tagccaaaac aattctgaaa aagaacaaac ttgaagatgt 169020
catacttcct gatttcaaaa catattacaa agataaaata atcaaaacag tgtgatactg 169080
gcataaaggc acacatatag accagtggaa cagaatagag aggccacaag taaaccctca 169140
catatatagt cacatgatct tcaacagtgg tgtcaagacc acactatggg gaaaggacag 169200
tctcttcaac aaatgatgtt gaaaaattgg ataccacaca caaaagaatg aagtttggcc 169260
cttaccttat atcatatata aaaattagct caaaatggat taaggacctt aacaaaagac 169320
ctaaaaccat aaagctccta gaagaaaaca aggcaaaagc ttcatgacat tggaattgtc 169380
aatgatttat tggatatggc acagacagca aaagcaaaaa taggagactg cattaaactt 169440
aaaaacactt gcacgtcaaa ggaaacaatc aatagagtga aaagccaacc tatgggattg 169500
tagaaaatat ttgcaaatcg tgtgtctgat gaggagttaa tattgagaat atgaaagagc 169560
tcctatagtc aacaagaaaa atacaaataa tctgttttaa aaataggcaa aatacttaag 169620
ctgacattta tctaaaaaag gtatacaaat ggccaagtag atgaaaagat gttcaacatc 169680
actaatcatt agagaaatgc aaatagaaac cacaatgaga taccacttta tacctattag 169740
aatggccact atcaaaaaaa aaancagttt ggcaagaatg tgaaaaaatt gcaacatttg 169800
tactttgttg atgagaatgt aaaatggtgc tgctgctgcg gaaaacagta tggcagttcc 169860
tcaaaaaatt aaaagtagca tgatcatatg ataaaatata cttctggata tgtatatcca 169920
aaagaaatga aaatagaatc tcaaagacat atttgcacac tcatgttcat tgcagtattt 169980
ttcacaataa cagaaggtgg aagcaaccct agtgtccacc agtggatgaa tagataaaca 170040
aaatgtggta tctacacgca aggaaatatc attcagcctt aaacatgaat aaaactctga 170100
cacatgctac acatggatga accttgagga catcatgcta agtgaaataa gccagtcaca 170160
aaaagacaaa ttctgtatgg tttctcttat gtgagtttat ctaaagtagt caaattcaga 170220
gaaacagaag acagaatgag agttgccagg ggctagggag gggagtaaac ccattgttgt 170280
tcaatgggta cagagtttcc actttgaaag atgaaaaagt tctggagatc tgttgtacaa 170340
caatgtgagt atgtttaaca ctgctgaact gtacacttaa aatggttaag atggggctag 170400
gcgtggtggc tcacgcctgt aatcccaaaa ctttaggagg ccgaggtaag cacattgttt 170460
gagcccagga atttgagatc agcctgcgca acgtggcaac accctgtctc tactaaaaat 170520
acaaaaaatt agcctagccg gacatggtgg cgcatgcctg tagtcacagc tacgtaggtg 170580
gccaaggtga aaggattgct tgagcccaag agatcaaggc ttcagtgagc tgtgatcatg 170640
ccattacatt tcggcctggg caagagtgag accctgtctc aaataacaac aacaacgaca 170700
```

aaaatgatca agatggcaaa atttatgttg tatgtttttt accacaatta atttctttta 170760

tttaaaactg atcatattaa caataaggag gtggttagtg aaggcaataa gagcaagctc 170820

agggccttcg atgatacaga gctgggttgc agaggatagg gaatcaatga cgggtaagga 170880

agtgaaggtg gatcaccaac cttttcctca aacgcctcat actctgtttc cagatagacc 170940

acctggttgc tggctgattc ttaattttgt tagatatgtt aacaccattt gctacagtgc 171000

tacccaacct tttgccattg tggtatgcat agagaatgat atttgtacaa cacgctgaag 171060

tcaataggta ggaagcaaaa gcagctggaa atatcagata tcaaggattc ctgatgctcc 171120

agggaacagc tgcccaggag cgagaggacc gtaccctgaa gaacctataa atcattagtg 171180

gctgagagtt gaaaaactct gcatttcact caaatctgtt gatgctacct aactccactt 171240

gaagccattc tttcaatctt ttggtcagag atgactcttt tgaggtcagt ttattattgc 171300

ccataacttt ccatgatgta aatctcaaaa gatcagctta tagagcattt gatctgctat 171360

ttctgaaata gctacttact tttataataa attcttactt ttataataaa tgttagactt 171420

tttataataa attctctgtc acactaagac ataaacagca ataaaattgg tttgaagaaa 171480

tgaaatggct aagacagaac agtttagatc aaggtttaag ccaagaagtc acaacagagg 171540

ctcagaattt gggaatatca attatcaaca taaaaacaac agcctaggaa tctagaagtc 171600

attcaaaaag tagaaatatg ggaggccaag gtgggaggat cgcttgagcc cagtagttgg 171660

agaccaacct gggcaatata gtgagaccct catctctaca aaaaataaaa attatcgcag 171720

tgtggtggtc catacctgtg gtcccagctg cttgtgaggc tgaggtagga gaatcacttg 171780

agcctgggag gtcaaggctg cagtgagctg tgattgtgcc actgcattct agcctgggca 171840

acagagaaag accctgtccc cccaaccccc caagaaagat aaagatagaa acagccaact 171900

ggatttttac agtcaaataa gtataggtgg gagattcaaa aagtcagttc aaaacagaag 171960

tacctgttag aattgatgtt ctcctgatat gagtggaatt ctgtccctca aataggtacg 172020

tggaagccct acgcctcagt ccctgtgaat gtaacttgat tggaaatagg gtctttgtag 172080

atgcaatcag gttaagatga ggtcattagg gtggactcaa tatgactgtt gtcctcataa 172140

aaagagaaga gacacagaca cagacacaca gagagagcac tgtgtgttga tgaagccaca 172200

gaccagagtg atgtggtcca caaggaacac ccaggattgc tggcatcaca agcagctgga 172260

aaagtcccat ggagtagacc cccctcagaa cctccagaag gagcttgcct tgctgacatc 172320

ttgatctcag acttccagcc cccagaaata tgagacaata catttctgtt ttcagccacc 172380

cagtttttttg gaactttgtg gcagcggccc tagacaacta acacacctcc agagccttgt 172440

tctctggact caagaagaaa cgatcccttc taaagctggc aagtagcctg gctgagagca 172500

tggggcagca gggaacaatg gacgtaccct tggaatccta cgccagcctg cgccaccctg 172560

gagcagcttt ctcagcagtt ggagttattt tcttctgcct ttctttgatt ataggaatca 172620

aaccccatca tagtaaaatt ttatagactt gagactttct tagaggaaag gtttctgtaa 172680

gcccttccaa aatttcatgt catgtgtgat taactcaaac cagaaggaca ttgttttaga 172740 gaccttttct ttgcctcatc catgggtact cgggactgtt ttgttagcat gatctctctt 172800

accatggaac aaacatgaaa ctacacacct aagattgtcc actcgagtct acaatggtgg 172860

gttgtcggtt gtatcattct ttaattgtgc taattcagaa ttattttcat tctagtttcc 172920

acagcactct gcaaggtaac ccttcaatta aaatatatat atatataggc caggcatggt 172980

gactcacacc tgtaatccca gcactttggg aagccgaggc aggtggatca cctgtggtca 173040

ggagtttgag actagcctgg ccaatatggt gaaaccctat ctctactaaa aatacaaaaa 173100

attagccagg cgtagtggcg ggcgcctgta gtcctagcta cttgggagac tgagacagga 173160

gaatcacttg aacccaggag gcagaggttg cagtgagctg agatcgcgtc tgcgaatctc 173220

catctcaaca agagtgaaac tccatctcaa aaaaataaat aaaataaaat atagtgggat 173280

ttggagggaa gatgtgccca aacaggagag aagctcacac ttcaactcca actgaacacn 173340

cacccactgc ttaactgtaa ggatgctatt ttattttgtt tcctttaatn ncaagagnnn 173400

nacaaanatt gttnaagana gcatgnagct tactgcactg taccattgaa aacagcaagg 173460

atgataagga agagcaacct cgcagcactt tttctttcag atcctgtcac aggcagtggg 173520

aaactagatt gcaatggttc taagagaaaa gcatgacagt tttcagtttt atgcggtgcc 173580

tggcagttgt tgaattacaa aaaggcactc ttatagttgt gttaagattt gttgtgagga 173640

gttagaatag tgaggtttac cctgtgtgga cctacacatt aaataggaac atgctggagc 173700

tggagtcatt gccctgattc cttttggact caactttcac acagtccgtt tggatctttt 173760

ttcaagtaat tcatgtactt ttccgttctt ctggaagcat cctcacaatc tgctttatca 173820

tccattcttc aaatatttga gtacctacca tgtgtacaac gctaggccag gtacaataat 173880

aagtaagttc tttgctccca tggagccttc tggtctaata ggggagttgg atgtaaatca 173940

agtaattaaa ataattccta actgtgtaaa gattgagatg agctttcttc agaaaataac 174000

atcatttaag gagagacact cgtagaagac cccgacacaa atggtggtca ggggtgagga 174060

tgaagaaagt tacaggagag ctaagttttg caagatgagc cttgtggact tcacggaggt 174120

gtggtgggtc aggggagggg agcacaactc cagggagaag ggaaagcctc tgtgcaccca 174180

gaatgcagga agaccagggt ggagtggtat agagtaagtg tggaaaagtc agcagggcat 174240

gaaccatgca gtgccttgag aaacacggta aggtcattta ccatttgttt tttgtttgtt 174300

tgtttgtttt gagacagtct cactctgtca cccaggctgg agtgcagtgg tacagtcttg 174360

gctcactgca acctctacct ccaaggttca agcgattctc ctgcctcagc ctcccgagta 174420

gctgggatta caggcgtgcg ccaccactcc tggctaattt ttatattttt agaagagacg 174480

gggtttcacc atgttggcca ggctggtctt gaactcctga cctcaggtga tccgcctgcc 174540

ttggcctccc aaagtgctgg gattacaggc atgagccacc ataccagcca gaacatttac 174600

ctttaatgtt tacctttaac ctgagccact gaaagcttat aagatgggag gcagtggtga 174660

accagagagc tgtttggaaa ataaaatcaa gagggacttg gtgataggtt ggatttaaag 174720

acatgagtga gagagtctca aagatgattc ctggtttcag gtttctacag ttgcatgaat 174780 gaaatgtcat taactgagtn aagaaacact agcatacaac cgtatttggg ggaaacgttc 174840

atcaactcag ccttgggcat gttgagttgg tgaatctttc aagatgtctg aacngagatg 174900

cctaatagtt gtgtataggt ctggagttca ggggagaagg atggctgcct agagatataa 174960

atgtggacat tcaatgtaat tctcaattat ttattcaggt gctgttctag acatgtggcg 175020

taggtcaatg aacagaacag ncaaaaatga ctgcctcctt tgagcttcag tcctagtaaa 175080

gttctccatg ttttcataat actggaaggt gtaaatctgg atgagattgt atagggtggg 175140

aagaaaaaat atgctggagc cttgaggaag agcaacattt aataactgct ggtgtttgcc 175200

tttatggtaa ctcaacctgt ctctgttttc ctccttgtgt attttcactt tcagttccat 175260

tgcaaatagt caacattttc tgtatctcag gttcaaaact tgatcaataa gaacagattt 175320

ggtcgatcca gctaatcaat atgattcctt ttaggcagga atttcatctc tgattgcctc 175380

gcaatgctgg ccagtcttat agaaagtccc tgtctttggc ccggtttctg gctcaaccag 175440

ctttgacaag ggtgatgagt tactttgtac cataaggcaa tttatnagga agaaacacct 175500

tcataaaaaa aaacaaaaac aaaacaaaac aaaaaacaga acaggcagac atttggagaa 175560

ctcacctgca tattacagag agatttaata tgtttatcta ttgcatgcca ctgtaattta 175620

acagtctagg caaaggtaac agaaaagagg agctgagata atccagctcc tccccaacca 175680

ctcattcagg aggaagcaga tgtgagtgcc attgtagggt cctgagggag tattggttcc 175740

catttcctag cctgcccctc ctccccacgc ttggtcagtt acaccttcat atggctcctg 175800

ttctcacttt acctctctgc tttatgataa gcgcagataa aaggcaagag acacagacaa 175860

agcgaaaaag aaggaagagg gaaggaaatt tagaagggac cttgcaaaca ttttccttcc 175920

actgagtggt ccacaggagg aaacaatgaa aacttctctc tcctttggga ggccaaggcg 175980

ggcagatcac aaggtcagga gatcgagacc atcctggcta acatggtgaa accctgtctc 176040

tactaaaaat acaaaaaatt agtcaggcat ggtggtgtgt acctgtagtc ccagctactc 176100

aggaggctga ggcagtagaa tcgcttgaac ccaggaggca gaggttgcag tgagccgaga 176160

tcgcaccgct gcacttcagc ctgggcgaca gaatgagact ccgtctcaaa ataaaataaa 176220

ataaaaataa caaaaaataa caaaaaacaa aaaaanacct tctctctcac taacaaagag 176280

atgtgatttt ggagaaatgt cttccctctc tttcagtttt accaaatcta acttgggtat 176340

ctttagatcc ggtgtgccct cttttctttg atgacacgag tcctaactcc tatatattgt 176400

cttatatgat tgttttatgt catccaagct ggcacttata tcatctccct ggtctcatag 176460

gcatttaaat ctgggcccct gcactagacc ctgacaagat ctcttctaga tctaaaattc 176520

tcatccggga aagagtcgag agctgtcatt actgggacac ctgtagctgt atacttcatt 176580

atgacagatc tgttgtttat gtccagtgtc tcttctgatt ttggggtgtt catgctgcat 176640

ggttttctaa tatacgtagt tctaatatgt tctctgttca ttactgcgaa tttctgctcc 176700

gacccctggc cacctacaca aaccgcaatc cccagcatag gcaacatgtt agtaaccacc 176760

cggagggggc gactgcaagc catgatgaag agctgggtct accatgtacc tgacctgtgc 176820 tgcaagctct ggcactagag ctggggtcac agggccctag ataagtaatc tcctctgaga 176880

atagtcagat tttgtcaata catacaattt tctcttagca gaaatcttgg ttgtattaat 176940

aaatgctgtc attgacagca ttgctttgta tttgtacatg gatcaaatga aagaataaaa 177000

gtcaggaaag tgatgatatt cattccttgc catttactta atatgccatc aaactttccc 177060

actttcctat tatttttca ctctgggagt ttaggaaagg gaattatatg actacctgga 177120

ttctcatatg ccttcaacat atttttaact ggatcatttc atcttactct ggatttgtgc 177180

ctaaaatctc agtgctaagt gctctatctt ctaaattgag attccacctt ccaaagatgn 177240

actggtggtc ctcttattcc tgattatctt cccagctcac actcccccag atgaaataag 177300

gagagtgctg tgggaggaaa gttgtggtag aatcatcagt agcagtgttt acggagagag 177360

attattcctc ctttgccaaa gcgcaagaga ggacttccag agaatcactt gtaggagtga 177420

ggagtaccag tcaaacagaa agagactggc tcttctatct ggaagcttgg aaacccgtag 177480

ctctactcag ttcagccact agattacagc aagctgagtt ggttgtgaaa ggaagctggc 177540

atcttccttc tgaagaatgg tgcagatggg ggtagagtga acctgccctt gcgtgacata 177600

gacagttctg aaagtcagag gccagccaga gtacagctag cagtagcttg tcaagaaagg 177660

actgtcaaga agaatatttt ctaagattaa tgctgggctg gggcaacccc accccttttct 177720

ccttctctgc cacctcagac agtgcagagg ggagctcagc ggtgcccatg acagatgtta 177780

gaggaaggtt ggggaggcag agggctgcac aggcagaggc caagccaaga catgctcatt 177840

caggagagga ttaaaggaag gacctaagga agacaagatt ctctttctct taggacgtct 177900

ttggctgcgg aagaccagag taaaactctc cttatacaaa aaagaaatca cacaaattat 177960

atgttaaccc tatcaacaaa ggacataagt gctgaatggg tgtagatcca aaccaaacta 178020

gaagggaggt cagtggtttc ctaagataat accttctgct cctaactctc cactctgacc 178080

ccaactttgg aggagctaca cccaaaagaa gaagatgtgg agcaatgaaa gaaaagacag 178140

ttggttgctc accctgtaat ctcagcactt tgggaggctg aggtgggtgg atcacaaggt 178200

caagagtttg agaccagcct ggccaacgtg gtgaaaccct gtctctacta aaaatacaaa 178260

aattagctgg gtgtggtggc gcctgcctat aatcccagct actcaggaag ctgaggcagg 178320

agaattgctt gaacccagga ggcagaggtc gtagtgagct gagatcacgc cattgcactc 178380

cagctctggg caacagagca agactccatc tcaggaaaaa aaaaaaaaaa agaccatgtc 178440

ccctccctag ctcaaaacca tatgtcatac tgtcatgtaa tggggtgagg aagaagggag 178500

gtttgaatca gatatgagac tagactgcac tgaagtaaat ttttaataac caaagttgac 178560

caaaaagtta gggaattaac tagggataac aaccagtctt ttatcctcat ttataatcca 178620

actttcctat aattgaggat gactgatttc actgttggcc ttcccttaca ttccacttgg 178680

agcattactc agtgagaggt agggcatccc caggcacagt gagaggtgag acattcctgg 178740

cattctgtca ccgtgaaatt tccagtcctc atcttgtgtt ttctgtaatt cagctctaaa 178800

gccagccatt tcttcaaaga gtgctgattt gttttattgg agaatggtat ttagggacca 178860 aaancctagc actaggtttc tcattgtttc cagtatatca gtgtttctag gctctctcag 178920

caagaagagc agggaaatac atgtatgtgt actaactcaa gtacacacat acatgtattt 178980

agctagtttt gtatatattt acctatctat gcatgtgtct tacataacca tagaaccatc 179040

tgccttgata acttgtactc aggtattaac aatatactaa ttttgttaga ttgttactgg 179100

gacatgagtg cagtatatac atgttatttt ctttaatgtt atccctattt ccacttgttt 179160

ttaataaaat atgtggatta atgtgaattt tcatagtgta tctaataagc ccttttatca 179220

acataggaga catttaaaaa ctattcattt gaccacagac atacacaaaa atgataattg 179280

gaaatgcatg attgtggcca gaaaagaaat tacttaaggg aataggacac tgctggggtc 179340

caagacccca atgttaagac tgcattccaa agtacgtatt ttccatgcct ctaatttaaa 179400

gtgtcatatt tatttgatag taaataccat tactgataca acgttacatg cacactaagt 179460

gcaacaaata agggaatttt ctatagaaaa tgagagccag agttcatacc atcagagaaa 179520

gaagttataa atgagcagag gggagatgcc acaggagaac ctgtggggct agattaggag 179580

tatcagtaag aactcagctt attttaagag agaggatagg taggtagtta catgcatata 179640

tatatatata cctgcataca tatggttgga gggttatgta agatacatac gtaggtaaat 179700

atatacaaaa ctaggtaaat gtgagagagt gtgtgtacat gagttagtat acatacatgt 179760

atttccccgc tctttctgct gagggagcct agaagtaagg atatactaga aacaatgaga 179820

aacctagtgc tgggattttg gtccctaaat aacattctcc aataaaacaa gtcggcactc 179880

tgaagaaatg gctggcttta gagctgaatt aaagagaaca caagatgagg ataggacatt 179940

ttgtggtgcc agaaaataaa gaagacctaa aaaggaaaaa acaaacaagc aaacaaacaa 180000

acaaacaaaa tagatagaag tatgtctgaa gactagagaa gccgactgaa agtgcccccca 180060

gtggccaaag caggaagact ttgagcaatg aaatgaataa ggaaagtact gggttataac 180120

ccaaagaata caataaatag tcacaaatac atatttatac agatacatga ttgaataaat 180180

aaataaagga ggagggacaa ctgattctct cagaataatt ccaagtaaga aatgtagaga 180240

aatagaaata atagaataat aaatgaagta aacagaaaat tgccattagg aaactattgt 180300

tgtaattgcc acagacaaga gctgatgaat accaaaatta gtatgtgaaa ctttaaggag 180360

aaataaaaca tttgcatggc ctcaaagtgt cctctccaag tcttttttta attattgtag 180420

cagctttaac atccacccac aaattctttg atactcctcc ttcctaagag ataaagcttt 180480

aattcttcta agtgtcagtt ggagtcagtg acttgcttca agccaatgga gtaggaaaag 180540

agaaacactg aatgtaactt tataatgcag aaacctggca gacatcacct tacccaagta 180600

aggacagtta acctcaatgt catatacaca ctgatatgac gagaagggcg tatcatctcc 180660

atggtatttg tcccaaaaat gtataagctc agtctaatca tgagaaaaga tcagacaacc 180720

caaaattggg ggacactctg caaattacct gactgatact cctcgaaacc atcaaagtca 180780

ttaacaacaa gaagaattag aggaattgtc acagccaaga gaaacttcag atgtgacaat 180840

gaaatgcaat gtggtttcca ggatgagctc ctggatcaga cgaaagtcat tcgtgcaaaa 180900 acctgggaat tcccaataac ttctttagtt aaaagtattt gaccaaggtt aattttgtag 180960

ttttgataaa tgttctatgg tatataacat gctgggtcaa ggtatacagg aactctctgt 181020

actatctttg caactctcat gtgagtttta aattacttca aaatagtgtt tttgtaaaca 181080

ggatgagaat tttctatggt aagctgtaat aatgcaacat agaaattagc tcctgtgttg 181140

ccttgttcna aagnatatgt acatctgtgt agctttctgc catgaatacg ttgtatcaag 181200

tttatacttt attattcttt tttgatcttn aaaaagctgg gaacacacac tttatcaaga 181260

gccaaatgct ttttgtgatt ttgtttncta atagaaaatc tgatgtaact tgtcattact 181320

cctctttcat ctttgccagt aatataactc atctgaaatc ggcagacttg cttgtgctgg 181380

ataatataga tgattccata ttgtaaaggc tctaagaaaa caataataga atgaagagaa 181440

atattacaaa tgatcctctt ctgcctgttg atgggactcc gttagtgttg ccttcttgtt 181500

ctcatctcat caccattgga ttatcattgt attgtgtgca caatttgagg ccacagggcc 181560

acagagagtt gtgcccttaa ttaaaaatgc atcatgctag taggggaaaa gcacacgtaa 181620

ttagaaagtt atagtgcaaa cccctgagca aggtgcaaca gagggagatg tgactaaatg 181680

caaggtgcac agaggaagat gtgcctaatt tggtttagag gcaggtcctg gaagatgtgt 181740

ctaaaaatga tgcataaaat tttcttagga aaacgagaga aaatatccca tataagagga 181800

ataacaggca caagattcac aagtaaaagt gtaccaattg cttgaaagtg ttgttttggt 181860

gtggcagtat acaacagttt tagaagctca agctctggaa tcaggaaaat tgttgtcaaa 181920

ttccagttct gccacttcta gctatgtgac ttcaggcagc taattgaagc tccatggacc 181980

tgtttcctca tcgatgacat tgcataatat ggagctaatt atggtaccga catcagggag 182040

ctaattttta gggttaaata atattagcat aaagtgttca gcatcatatc cagcacgcgt 182100

gtttactgct tcacctacca gttgttattc tgagaaatat tcctggagaa gatggttttg 182160

gaaggattta taagagttaa ataagtgtat tgcaggcatg gggagcagta tcttcatgtg 182220

acatccctgt gtgagggggg cacagaccta caggttagag catagagggc aaagcactgg 182280

aagacccgct tatgccatgc taagaagatt agaatttctt ctgtacataa ttgatagtct 182340

ttgcattact ttaaggacca gagggacata aaggtatttc ttttctagat gttctaaaag 182400

attggtttaa agggaaaaaa ctaagacaga aagccagtta ggaaggtttc aggagtctca 182460

acgagatgaa atttggatga ccaattgacc attggattca gcaaaatggg aaatacgctg 182520

actttgccaa gagcagcccc actgaggcag ggggggacac ctgagacagg agaggctgga 182580

agaggggtgg gaagcgatat aatggtgctg taagcatctc tttgaagaag caaccttaaa 182640

tagcaaagga ggcatgatca ggagaaatgg tttaactgcc tattttcgtt ttcattttta 182700

caaccaagag tcttgagaaa acaatccaat caacatagca aacgagaaga aattggggaa 182760

agagggagag gttgagttgt ctcagattag ccagagggaa gggactcaag cacacatatg 182820

tgtcacccca ttcatatggg tgaagtcctc ctcctgtaga gcaagcagtc actgtggcat 182880

ggtgatagct ttggtttgta tagcacattt cccttgagta attcaaatta caggcatatt 182940 gctatctcac ttacacgtct ggcatttata tgaaagacta aagatcctca ttttttattt 183000

gagaacactg aaatccaaag gggtgaagag attataaaac tgaacaaaag tttaaatcca 183060

tgtcttctta ctttgttttt tactgttttt acttatgcta tctctgatgt tttataaatt 183120

aaaaaaaaga gaaaaattac atgggagtgt atgtgcgtgt ttgtgtgtgt gtaagtgtgt 183180

gtgtgtgtat cctgaattcc ctttggtgtg acatttttcc ctagcactta acgaagaatc 183240

atctcacaat aagaaccaag tatctagttg acatgaaact ctaatataat ttcctctgaa 183300

gtggagcaaa aaagatatat actttttttt atcacactga gcatatttct atgactcagt 183360

aaattctctg atggaaacaa ctgtgtctca gtggagtaac aaaggacaga tataatataa 183420

aatagagatg ttntgggacc tagaagaata aacctacaaa atcaatcttt attcatcacc 183480

actgtctcta agaaattatt aattacatta atgggatatt aaaggacaat aatatgagca 183540

aaatgtgctc attcaacata attaggtatt ctggatttga atagtttttt ttttctttta 183600

ctaattcatt agagttgttg gttagcttag ccttgctgca aatattatct atgttgttgc 183660

acaaagtaac cttctgaagg acatttaaag cttgtaagat cttaagggtt gtgataacag 183720

catgaaagta aacattttat gcacacttca gggatatcat aagtggcagg aatgcaacag 183780

tctgttataa gtgctctgct gagctcccaa accaacaaac acaaaacagc cacttaattg 183840

cccagtgcat ccgcccctcc gaaaacagag cttctcttca tacttactta agggacatgc 183900

tgcttggacg aatgagtaac cagccaggat acagcacatg cagttaatca gctaatgtta 183960

tgaacatttt ggatgtcagc tatcacataa aaatgtagct tttatggccc ataggctaac 184020

caagaaatga gccctccaat ttctatctgt ggcttcagtc gatgcgtcca aataattgaa 184080

aaccccactt tgaccttgga tatactgggt gtattaaata tttttgtaaa acggggaaaa 184140

attagtgact taaacatgtt ttttgatgaa acgtttgaaa actgtacagt atcgctgtct 184200

tctatgcact aatttgactt ggttgccttt tgtccacagc agaaaagaca gtgcccaccc 184260

acctgaagaa attctgtatt ggatttcaca ggccagtgag ggatattctt tgtatttgca 184320

tattaaggtt ttataattac acagaggtat tattaaaaac caacatgtgt gatcattatt 184380

ccagatgccc catctgtttt gactgtgcga tttgggttgt aaaatcacat atacaaatta 184440

gcaagagaaa ctcccttgaa aaataatagt ttacgcctta gcaagactct atgttctgaa 184500

ttgctatgtc atatatatag atacagaaaa aatagctaaa gccactgtgt aaagaagcaa 184560

agagacttca gccttaattt attcacataa caaagaggag ggactccttt cttctgattg 184620

aagataatgt ggttattttt cctatattga acaaccaaaa aatattctac agttacatca 184680

actcacaaat ctatgcacag ttttcaaact aacacggact agatttcaaa atgagaatct 184740

cccaagtctg aatgtttttt aaattgtttt cttttgggaa atcataccaa aaaattacaa 184800

aaatttttat ttaatgtaat tcttcaattc agttattcct cacagcaatc agaaaggaac 184860

tgaaactgaa taatgtttga atattttaac agaagaaagc acaggctgaa atagggaaga 184920

tttaagcgat gaggcagaac acagtatgtg tcgtgcagag caattacacc cgaagtctct 184980 tgagaagtca gggaaagcag aggatttcag ccgccccgatg tgtaatgata tctcagccag 185040

gcagatgctg agccgtagaa tagcactgag cttagacgac ttctagtttt aaagctcaag 185100

gtttgaagtt tgtttctttg attctaattt cattcttgca aatgtgcttt tctttaccat 185160

gtgaatgtgt gacaaagcta tcatttgatc tctcggttgc agccagactt catgctcagc 185220

tgaggccatt ataaacttcc tttgtactgt tctttacatt agatttttca atatgagaaa 185280

taaaatccta ggattctcta cactgtgttc ttttccttca agaatcatga tcttaaaaac 185340

actttttaaa tggacattta caattataat tactttatgt taaaagatct tcttttcttg 185400

gtatagtata atgatcttaa aaattcataa ccacttctta taatatatat tattcattca 185460

tttgcatcaa agtaaattta atcttgcaaa atctttgctt aaatgcaaaa aaggggaaag 185520

aatttggata aaaggtagaa taccatttca ttacatttta aagaaactta atagtggctt 185580

caaaataaca tgtaatactt tgtgtatcat ggtcacctca tcccccacca ccaacacaca 185640

ccctgtctac tcccatctta ctctttcttc ctttttctgt gtctgtagct tggtcaaaat 185700

catgggttgg gaaagacgtc aaagttttgg tggctgatag tgccgaagta agaatggaat 185760

aattttgctt aactgcatca tcttcccttt tgaaaaactt ttaactgaac tctaatatac 185820

atatagaaag aggtatagta tacatttttg taaactgaac acacccagca ccctgctcag 185880

gaaatattga ctgtgcctta gaagctcctt tccatactcc ctcccgctcc ccgccacccc 185940

cacccaggac gaccactctg cagacacccg gcagcacacg tcagtgccac ctgctggttt 186000

aacttcctcc acgtggaatg gagtacatac aacagtatgt actctgttgt tgcgtctggc 186060

tcccttcact caatgttatg ttgtgaaatc cacatttctt acactgagat gtagttcact 186120

cattcaaata tttaaataaa cccaaatgta ttttcccact caactgctag tgggcatttg 186180

gacagttcac attttgggac ttttaacaac tgacactgct gtggccaagg gagggcttgt 186240

gccttgtggg tgtgcacgca gttcttgcag gcacgcacct aggagaggat ccacagggct 186300

gctggattta cctgtgctca tttttagttg atgctgcttc tctcttaata acagcaaaga 186360

gtgcatggga actgctgtcc agaagaacaa gtccttgctt ttaggaggtt tgattcagtg 186420

cctcagccat gagattcaaa atactcaccc ttcaagcagc cgctcatcct tctctgtgtt 186480

aactggccct taatgcacaa ctgtagggca atgatgccag gtttgcacaa tcgtgaagaa 186540

ccttctgagt ctccaaaaag agcttcccgt ctgttctttg tagtttcaca gcgcaatgac 186600

caagtngtgc tgagttctaa ttcatttatt tacaaagagt tctatttnct actccctaga 186660

atgtactata gaagtaaatt taaaatggta tcttgttcct aaggagcgta caagctttgg 186720

ggagtgtaaa tatctctaat accaattaga atgcgggtga aacccatttt taaaatgtgt 186780

aaacaaggtg actgttggta tcctacatat gttaaacaat atattcaagg atattcgata 186840

aaatataact tgccatttct ccaacatgtc cctttataat tttttttttt ttttttgag 186900

acagtctcgc tgtgtccccc aggctggagt gcagtggcac aatcttggct cactgtaagc 186960

tccacctccc gggttcacgc cattctcctg cctcagcctc cctagtagct gggactacag 187020 gcgcctgcca ccacgcccag ctaatttttt gtatttttag cagagacggg gtttcaccgt 187080

gttagccagg atggtctcga tctccttacc tcgtgatccg cccacctcgg cctcccaaag 187140

tactgggatt acaggcgtga gccactgcgc ccagcctaca aaatttttat atctgtactt 187200

tgtctaaaca ttcattagtt tgtcaacgac agacatcagt tcaggttaaa cgacacagga 187260

gtctaatttt attctcttct gtgaatctaa ggacacgcat actccacact aacaaattta 187320

ttcaatcaat caatagatat tttatgctta tctctctacc cttcaggata tatctaaaat 187380

cttttctttt atttcccagt tgtctgctgg agagtttagc ttagagttga gttccgccac 187440

ctgatcaata ccattattcc ttcactgaat atttgctcag ctcaccttgt gtccagtgcc 187500

atgccagcca ctgcccgtgg aataatgagc agagccctgc atggttcctg tcctcaccaa 187560

gtttacagtc tcagggaaaa gagaactgtt gtgcaagggg tcacacaata tagtgtcaaa 187620

cgacagtacg agtagggggt ttaaggcatt cccccaaaat ttttgtccag ctggatcctc 187680

agaatgagac catctttgga aattgggtct ttgcagaggt aaccagttaa gttaattgtt 187740

ctggatggtc cagtgtttgg catctataag gaaaaggaga gggggattca gattcagaga 187800

cacagaaaca cgcaggccca gaaggccatg tggtgacaga agcagggatg gggtgttaca 187860

gctacaagcc agggaactcc gagcatggct gagaatcatg agaaacaaga gaaaggcaag 187920

gaaggactct cgcctggagc cttcagagag agcatggccc tgctggcgcc ttggtttttg 187980

acttctagcc tccagaatgc tgagagagta aatttctatt gtgttaagct aatgatctag 188040

tggtaatgtg tgacaggagt cctaggaaac taatgtattg tggattcaat tgacccgtaa 188100

gaaattgcag atgttcaact gctttgagcc acacaaacag caatattata ggattcatcc 188160

taccatgtta tgagcttgta agaggaaaac tttgtcagag aaagcttctg tgaggaattt 188220

ctagtggaac tgaaatctag aaaacaagta gaatttgatg gagttcaaga gaaaaagagg 188280

aagaaagggg aacccagcag agggccagca gagggccggc agaggtgggg gaggcacaga 188340

ggttctgagg cagagggtgc agatgtttcc aagcaactga aggatgtgac tggaacccgg 188400

tgtggggata gagctagctg aggctggaca ggcagggagg ggccaaaaaa aggcctctga 188460

gcctagattg aagattgaga ttttgaccag gcatggcagc tcgcacctgt agtcccggct 188520

actcaggagg ctgaggcaag aggatccctt gagcccagga gtttgaggcc agcctgggca 188580

acattgcaag atcccatatc tttaaaaaag aaaagtttat atttttgttt ttgtttgtta 188640

aaacaaaagt ttttttgttt gtttgttaaa agcagtgggt agccaggtgt ggtggtgcat 188700

gcttataatt ccagcgactc aggaaactga ggcttgaacc agaaggcaga ggttgcggtg 188760

agccaagatc gcaccactgc actccagcct gggagacaga gcaagactcc atctcaaaaa 188820

aaaaaaaaaa aaagagagag agaggaagaa aaaggcagtg gaatgccatt gaaaagtttt 188880

aagacagagc gtgagaaatg gcaataaccg cagaatgaag agcagattag aaagagctga 188940

agcagatgga gaaagcactg caggaaatgc aggaacccag tggtgcattg gccagggagg 189000

cactggtggg gatgcaggca gtgggcaggc ctagaggctt ttaggaaaga actctgtgtt 189060 ctcatttcaa tgatatgtcc caattaatat accatttgcc cttgtgtctg tttgaagaca 189120

ttgaattctc ataaaatatg cataccagtc ctgatacatc tagagccccc aagtcaactc 189180

ttctgaccac gtgtgtctat tctgctgccc gtgaagtctc gtggagtaga agtagtgcct 189240

tcattttttt tcttatttct ctctttgctc tataccttgt gtttccgagc actctgttcc 189300

tatatctgcc agagtggcca ttctatttta ttatcattgt ctctttatag aactatctac 189360

ccagttatac tgtgggctcc gtgagggcga gtgcagtgac ttatttcttt ttatatcccc 189420

tgagcttagc acagaacctg acctgtaaga tttgatttat aagtgacagt atcaatgtat 189480

aaataaatga gtcagcaaat gcatgaatta acaatacttt atttctgctc tgtcttatca 189540

cctttggcct gatatacctg ccttacactc tgatctgatc taatagatta actgcttttt 189600

ctacttctgc tctatttacc actgacaagt atttttctac tagttgtatc aacttttaga 189660

aaacaagact cttcctggta aatctcttta ctcgtgtgtg gaatcatttt tctaactgtg 189720

tacctctcac agagatgcta ctacttttttg ttttgtgtct ttgggagcta attccttacc 189780

attgtcatca tcataaccat ggaactgttg agaacgtgcc attttgcaca cattactgcg 189840

ttgagcactc cagttaggaa gggaggaagg caacgaacag cagtgactcc ttctccaaat 189900

aagattacca tctaagacag aaaaaccaaa ggctacccac aaaatcaaat cagacttaca 189960

aatggctaag gtaaatgttt atgttacaca aaacaaaaac ataagagcga tttctgcagc 190020

ctaataagaa aaccagaaat caggatgtat aatctcatca tgagctcaca accaatagga 190080

tagttgagtc atccatcata ttgaaatggg aatgtttccc caaaaagaag ccctcctgtt 190140

ctccttagcc aatttgtatt tccacatata ggagggatgc attgccatct ctcttcctgt 190200

tttaaaagct ctcttggcct ctaattctta tctttccata attcctaaaa caaattatag 190260

agtaaaacca tacccaacaa ggaggtaaat gagtggatga tttttgaaac ccacccccca 190320

aatccctcca aaaaggcagc ctgaattaca tgtactcttt gaaaggcatc agtgttaaaa 190380

gatgcagttt gggaagccgg gacagaagac agggcagacg tggggaaggg tgttccactc 190440

tgcagaaggt cctggagaag gggagaggta aacagaagta aagagggact ccaagagacg 190500

tcggcactgg gagagaaaaa gagagcatag gaaagtatag catgaattct ggaagccaca 190560

gtccaccttt gtagtatctt agttcagcct tgttccagtt tatgatgtta aacctgagct 190620

tccctgaaac tcaatttgta gaaatataaa tatacactct taaataaatc catttgattg 190680

tcattcctca caaagatatc atatggcatt taagttcgtc gtgtccaagc atgtttgcag 190740

ccaaatggaa aactcgctgc atggaccaaa tgtggtattt gtgttttgag cttttctgtt 190800

ttaatctgct aggcgacata accaattttc ctgccattgt ctctctttac aaatgctact 190860

acttagtatg attctgtagt taattttttt taaatttttc atggttttta aaaattttta 190920

attctagcgt taggcatgtg taaacaaatt ttattgttta aatatgcaag taatagtaaa 190980

ccattaaaaa tgtctgagat gaaaagagac ttgtctttgg gaataaaatg acataaagtt 191040

ttccttgaat ttttttctgg tttantttt ttgttgttaa gttttgtttt tgttttttct 191100

gttaactcaa aagttttagc tcatagagca atttatggcg ggcaagagag gtatacttgg 191160
taggcctttg ctgccatcta ctggcatttc atgtaattac ttattgcttc cattctaaaa 191220
ctaattttgt gtaacgtata taatagtatt ataggttccc ccttttnttc ttctttccaa 191280
tttttcttta gttaaagcaa agtgaagcaa atgcggacac caaagaaaag ctccctttac 191340
gcctgcgaat ctttgaaaaa tttccaaaca gacctcaaat ggtgaaaatc tcaaagcttc 191400
cttcagattt tacagttcct aaaatcaggt accaaaatac ttctttaatt gcatatttca 191460
catatattgt aaaagattca tcatcttaaa tctgatttaa cccttctttt aagtatatgt 191520
gggttgggtt tctttctatt ttattgtcat ctttctacta cctctacctc tatttaaaat 191580
aaataaataa ataaacaact ttgactcttt taagcaatgt attcacagaa gagtcacttt 191640
ataactgcag tagcctaagc aagaaggctg cttcccatcg cttctctctg ggatgactcc 191700
tcgggaatgg gacatcccag gaaactggag aagctgctca gtgttgcagg gatggcaggt 191760
agacatcaat ctcagacccc aaccgtatga gactaaaatt gcctctgtgg cccttttatt 191820
ttcaagcaaa cagggaaacc acagctggcc cagtttgtgt actactacag tgtttatgga 191880
gcctaccatg tcaggttacc atatataatt tgccttccac tttttaaaag aagaaccgtg 191940
gcgtttttcc actcatactg atttgatttt ctaccacgtt tgccctccag ttaagcaaat 192000
gcataaaaga tgctttaaac acctattttt aataataaaa ggaaattctt ctagttatag 192060
ctcaagaaat tacaaacaga attacagtaa tatgttacta gctttaaaaa taattctgtg 192120
gggggttata aaagtatgag actatatatg gtgccataat atttaaaatt cgggtagaaa 192180
tatatttgca tttctaattt taaaaaacaa aattaattga atttggtaat ttgaagagac 192240
gttacttgaa ttttcgttat gctttgaaga tcagatgaac gcacttgctt ttgagtctta 192300
aaatgtaaaa ttctctctta gtcattctta tgataaatgg acatttagag agagcaacta 192360
attcttccaa gtcttagagg gggaaaaaaa gtgttttatt accacaataa gattcaaaaa 192420
actgggctgg gtgcggtggc tcacacctgt aatcccagca ctttgggagg ccgaaccggg 192480
cagatcacct gaggtcggga gttcgagacc agcctgacca acatggagaa atcccatctc 192540
tagtaaaaat acaaaattag ctgggcgtgg tggcgcatgc ctgtaatccc agctactctg 192600
gaggctgagg caggagaatt gcttaaactc gggaggcgga ggttgtggtg agctgagatt 192660
gcaccattgc actccagcct gggcaacaag agtgaaactc cgtctcaaaa aaaaaaaaga 192720
ttcaaaaaac taccatcggt tactgagtaa tattgcttga aagttatttc aaggatattg 192780
ctttttaagc aggagtcact gagtttcagg aagctgttgc taatgaaata ttttagtcat 192840
gatctagttt tctatgatct tttatcactg ctttatcttt gctcaagcaa actcatcatt 192900
gcaaaatgct gtcacttgct cccttccact aataatgcct gatgcacata tacttcattt 192960
gcaaagggag agttttatga agcagtttat tgaccgccag caacaggaca catgttgcct 193020
cctgagaagc ctcccgacca cctcttcctc ctcctgcgac cactccaaac gctcagcaat 193080
tgaggacaac aaatacatcg accaaaaagt aagatcngct gtttcgaata accaggcttc 193140

taaaaagcat ctatgtttga gaaaagaacc gcagagttct ttagggtcct ggaatccaat 193200
ttnctcaaga gaataatgaa aanaaaatga tttcatttgt aaaagatatc atgaaaccta 193260
acatttttta ctctatgcat tgaaattttg catttgcccg tgtaagcttt tgagctgcgc 193320
cacgtggcag ccatgcgact ttgtaaacat caagagcctc ccaaatctgc attttctcan 193380
gtgcggaata aagggctggg atctccaagg ctcagtccag gtgtaatgac ctgccatctc 193440
ctgcaagtac aactgcttat gtctacttat tacatcagag aggatacgag agcatgtttc 193500
tctaaaatcg gtcgaaagaa ttgccattta ttcaatgctt agtatgtggc aggaacttag 193560
gtgcaagagc ttattcatgt tctagtttac ctatttcttt atcatttgaa aagtagaagg 193620
gaaaattttt ttcagattcc acagcaaagt atcagtaggg aaccacatat aaataaagca 193680
aatttaggta ataattttaa agaaatgtac tttttcattt attcactttt aaaaaaaata 193740
atataagggc tgtgtgccgt ggctcacccc tgtaatccca gcactttggg aggccgaggc 193800
gggtggatca cttgaggtca gaagttcgag accagcttgg ccaacatggt gtaaccccat 193860
ctctactaaa aatacaaaag tcagccgggc gtggtggtgg atgcctgtaa tcccagctac 193920
tcgggaggct gaggcaggag aatcactcga acctaggagg cagaggttgc agtgagctga 193980
gattgtgcca ctgtactcca gcctgggcaa cagagtgaga ctccatctca aaaaaaaaaa 194040
aaaaaaaaaa nangacntag aatgtattat ctaggtataa taagtaaaca tatttatata 194100
tacacacata gacatataaa tgtatatgtg tgtatatatg tatgcatagt atcaaactat 194160
atatatcagc ttggtcacat aatcaatgtc agttcagtta gaaaatgaaa gattaataat 194220
tgtcctttaa tattaaagag attaaacagt agagtttgca aaaacatcta gttcacttgc 194280
aagtattcca ctttggcatt cctagagtgt tttttattcc tttacattgt tactgccctt 194340
tttcttttgc ttatagtgaa atccatattt ctctatgatc atgtgacaat aaaatgccaa 194400
ctgtttcatc ctccaggatc ttctttctct ccgctttctt tcccttccca cagggatgaa 194460
ccctttggaa ttccctcacc tttctcttta tctcgcctcc cctttttctc ctctttcttg 194520
cctccccttt ttcctccttt ttcacctccc ccttttctcc tcctctttct cacctcccct 194580
cccctaatgt gatactactc ccgtttacct ctccccagc tcccaggttc tctgctgtag 194640
ttcagtcatg ttgctcttgc tgaaaatgat gatgattttc tttgcaggat gtctcttctt 194700
tttcaaggat ttccatgtgt tgattttcct tcagccatct agagatgcta ttctactatc 194760
tttgcagaat ccaagattgt aaatgacaag tctgatgcca atgtcatttt ttttttaagt 194820
aattagcctt ttggttggta gacttggttt gcttgtttct tcaggaaaac ttgtaagatt 194880
ttctgtgtat ccttggagtt cagaagcttt cttagaatat accttatatg tcttttttca 194940
tcagaactgc tccaactaga gtcttttcag cttgccaact aagatatttg ctcagggaaa 195000
tttccattta ttttgcgttc aattgttact cttccatctg cttctctctc ccctgggatt 195060
gccatctctc tatttttgtt ctgtgcttgg tcttccagga cactaattag atttcaactg 195120
aggtcatcct ttgcttctaa atttctcgtg aaaaattatg atatggacat atgaccctta 195180

actatcctta aaggcttgca tattaccctt agctgtcctt agaggcttta cttttatctt 195240

ttatttttat ttttttagag atgagatctt gctgtgttgc ccaggctggg gtgcagtggc 195300

tattcacagg cacggccata gtgcactata gccttgaact cctgagctca agtgatcctc 195360

ctgcctcagc ctcccgagta gccgggacta taggcatggc caccatgcct ggcttttttg 195420

atttttttaa attaaactta tcacctctct tcttcacagt tataatccac tcaacactgt 195480

gcatggggca tggaagcaga aggtcatatg tcagtggata aacaagtggg aagatagtgc 195540

cttggcttca gccagtcagt ctgaaattgc agtttgcgtg gactagaaag caggttgttt 195600

tgttttgttt tgtttccaaa gcagattgta agccagaagt ctgggctgtc cttgtgacca 195660

cttatctcca ggaagaggat agagagctga tttccattat ggccaaagtc acaaaaccat 195720

gacatgcttt taaaaatcac tggtaacact accccactgg aaacacaaaa ttgaaatatg 195780

agtttatttt ctctctaaaa ggaaatgtga ttgtttctca catttcctga acacagtatt 195840

ttgatttggt gaggacaaag ataaaaagag tcttgagagg catgtgccgg gctgaggagg 195900

agtctaaaga gaatgtggcc cgcccctctg tgcgcccgcc cgaagcgaac cctccccagg 195960

tacagttgga ataacacgtt ttatcccata caacatgtct tattgctccc acttggaaag 196020

tggaatgtca aaaacatact ttttcttcat cagaggaccc tattatggga gtggcaagac 196080

aactgctcct gttactatcc acgctttctt aaagctttga ttcgatctag gttgtcagaa 196140

aagtagactt aattcttttc aggatccacc attttcagca atagttccca ccagttcttg 196200

gtgttcctca taagacctta ggatttgtca accttgttga gtgcccctgg gccagggaaa 196260

tctaaagcag aaaatccctc ttcaaaggat ttcagtgata aggatttaat ggaagcacaa 196320

attgcggaga cagcattctg gcatttgtat tatgaatacg gcaagccttt ttcattttta 196380

ttaagaaacc cgaaagccac ttaaaaacaa gagcgcataa acacagaaca aaaatccccg 196440

tcacagaccc taccaccatt agttcccctt ataaagtatt tttcagcgag cacttagaaa 196500

tttccattaa aatcccaatt gcctttggaa atcctagtca cagtaaagtg acaggcaggg 196560

ttcttcactc tacgtgnaaa aaaaaaacca cccagccctc tgcaaattgt gcaaattgag 196620

gctgtgtttc ctcagtgact ggtatgaaag aaaaaggaca ttagaggctg gtttatcggt 196680

gattctcata acatgaacaa aggaacaata aaataaaact gttctcttaa tggttatctt 196740

gtaccaggag aaacttctga cttcactatg ctagagggaa tcacctttct ctttccaggg 196800

aaaatagtga agccccaggt gcctgaagac tgcaagttca gggagaaaaa aacaacaact 196860

gacagcaacc ccagaagggg gctggggttt gatgaagcat tgcagaaaac acaaaagcac 196920

attaggcaag attttctccc agtaattacc caagctttga acagggatga gtaattccca 196980

tggcaacctc actttgcaga aaccatgttt ttaaaaatag ccctggtgga tcacctttat 197040

tatacagttt aaaatacctg tccttctccc agaaaacaac cagcaaacaa aaatctttaa 197100

aatcttgggg ctgacccaga tcctaaaagg aaaagacagg aaagagtgtt ttcctgcctg 197160

tcgtttgaga ctccatagaa acgaccttcc caagtggtca tttccccatc cctttatcct 197220 gctacctacc gttcctgtga gcaattgtga gaggttcctt ttcttcctga ttcaaggatt 197280

ttgacagggg cttttccagt ttccaagcag aaagaaatca agcaacaatt tctacttgta 197340

tttttcctcc atattattga aattaaactc tagtttcctt aagcagagtg gtgattaata 197400

aagtgagatc tgaacttgac tttccagtaa ccctgtaata aaagcagccc gctctgttct 197460

ccgaagtgag acatccagtc aaggtcaagg tctcacctgt ggttgagtct acagcgtctc 197520

actgccctgt ggcctgagga tggttctgtg ctgctgtctt ggctgatgca ccatttctta 197580

tgtcagaacc tgcaatagcc tgggaagcat gagtgaacct cctggagcca tgctaggatg 197640

agggctgtcc cagcacctcc acctcgtgcc cctgtacttc ccagcataga atcagagccc 197700

acataataat gaacagtgaa tattttctct gacagtaggg caatatctgc ccaccctctt 197760

gtcagcctcc tttctgctga agcatatgat agttctttaa taagccttca caatgaacag 197820

ctaacttatg tacagtaaag aggaaaatga tcattggaat gactatgttt tgtttaccca 197880

tttgatttta taactcatta acattcattc agtcattcaa caagcactta tagagcaccc 197940

gttattcatc tgctattctc tgggctgtgg gaattcaaca gtgagcaaga tgcgaactgt 198000

accctttaga agctcatagt ccaatgggag gaaatgcaca tacatgcatg aaagtaaaaa 198060

gtattactgg aaccatgaga cagaaccccc taagttcccc ttacagaaga agtgtgattg 198120

attgataaag gaagaaagca ccccaggcag agggggacat atatgcaaag gtatgcaaag 198180

catcaggtgc tcgatgtttg agagaaattg ggtagaagag tgattaagtg gtgcaggttc 198240

caaccctcat cagtaaagtg gagataattc tcattttcaa ggtttgttgt gagaatcaaa 198300

taatgtaata atgttaaata cctagcattg cacctgggct agcatatgtt ctgtgcctct 198360

caaaccagac aagaaaggtt actgttanta aactgtacca tgaagtcaga ctcagcaaac 198420

agatttctat tcatcagtta ggacaaaaga cttttacata aatttaacct tcagatgtaa 198480

cagtctagtt gagtgcaata tattaattac tgatcaattt ccttatggta cctaagtctg 198540

atttttttag agtttgattt ttgttgatat tttagttgct atccctattg tctacacggt 198600

gtcttcgaat ctgaagaagt aaattaggtc attctggtaa gatttactct tccgnttttt 198660

ttattgccat acagcatttt ataatgttct atatgattat agagtgcttg ataatttgtt 198720

cttgtaactc ctcacatagc aagttactca gcctgatcta taatttccag gatgaatgaa 198780

tgtttgttca ccttaaaaga gagagagaga gagagagaat gaacaaanga acgaacgaac 198840

acagcctgnt ttcattttta gngcttttgg tcttttgaaa gaattttttt tttaaataat 198900

ggcttagctt ctctgcagta ccattcaaga tgtatctcca ggaaaagagt gacagatgac 198960

atgatgctac catgtaacca tgggttgtat ctcaaaggga agaattggtg ttcctcatgt 199020

caaatatgaa aactctgtac acttctctct agcttaaaaa tgctttagta gataatggat 199080

tcccaattta gtcacaataa tggaaagaaa taaaaagatg cgtaagattt catcatcact 199140

ggccatcaga gaaatgcaaa tcaaagccac aatgagatac catctcacac cagttagaat 199200

ggcgatcatt aaaaagtcag gaaacaacag gtgctggaga ggatgtggag aaataggaac 199260 acttttacac tgttggtggg accgtaaact agttcaacca ttgtggaaga cagcgtgcta 199320

ttcctcaagg atctagaact agaaatacat ttgacccagc catcccatta ctgggtacat 199380

gcccaaagga ttataaatca tgctgctata aagacacatg cacatgtatg tttattgcag 199440

cactgttcac aatagcaaag acctggaacc aacccaaatg tccatcaatg atagactgga 199500

ttaagaaaat gtggcacata tacaccatgg aatactatgc agtcataaaa aaggatgagt 199560

tcatgtcctt tgtagggacg tggatgaagc tggaaaccat cattctcagc aaactatcac 199620

aaggacaaaa aaccaaatac tgcatgttct cactcatagg tgggaanctg aacaatgaga 199680

acacttggac acnaggaagg nggnacatca cacactggga cctgtcatgg ggcaggggga 199740

ggggggnnag ggatagtatt aggagatata cctaatgtaa atgatgagtt aatgggtgca 199800

gcacacaaac atggcacatg catacctatg taacaaacct gcacgttgtg cacatgcacc 199860

ctagaaccta aagtataata taaataaata aataaataaa taaataaata aataaataaa 199920

taaaacgatg cgtactataa aatcaaacca ttttttcttc ttactggtct ttgcagtaaa 199980

ttctcaagct aaaaccaaat atctgtgtgt gtgagcactc accctttcat ccaacaatca 200040

tttttgagga ccattatatg caaacagttt gctgggccct gagtaaacaa acacatcttt 200100

ctatattcaa ctaacattag agtaggaatt ccctctcaat gattattttt gcagaatatt 200160

taattgtgta accattttta ggttggaact atgtatcagc tgtttgtatt ccttagtaaa 200220

tgcgcattga gagtcaaaag taaatcagtc taggattaag aatttttttat aggtaattta 200280

aaagattttt aatcgcattt tctgtatcat gcacaaaaga ggtatcaagg agttaaggcc 200340

gctgctaaat ctatcttgga aatttacatt ttgctcttct ttttaaactt tttcctcttc 200400

tgactacagt tttacttata aattgtctct tctatcaagt gtcatgctag tctcagtgtg 200460

ttgtatttgt tgctgatgat ggtgttgagt ccaagtcctg cttccttgtg ttgccacaga 200520

ccttaaattt cgaatttcta cacccataat cccagcactt tgggaggcca aggctggcgg 200580

atcacctgag ctcaggagtt cgagaccatc ctgggcaaca tggtgaaacc ccatctctac 200640

taaaaaaaag tagctgagca tggtggcaca caactatagt cccagctact cgggaggctg 200700

aggcatgaga attgcttgag cccagggggc ggaggttgca gtgagccaag atcatgccac 200760

tgcactccag cttgggcgat acagcgagac tccgtctcaa aacaaaaaaa aattttggaa 200820

tttctaaatt cacaaacagg ttaaaagcat ctgacattta ctctagaatc acaagcccac 200880

tcaccttctt ttgtctctgc agatttccaa tctttgcctt cctgtttata ttagatataa 200940

tcataaagtc attttgtatt ttacttttgt aattaaaatc tatctagcgt atatttatag 201000

ttttgtttgt ttgtttgttt gagacggagt ctcgctctgt cacccagcct ggggtgcagc 201060

ggcacgatct cggctcactg caagctccgc ctcccgggtt cacgccattc tcctgcntca 201120

gcctcctgag tagctgggac tacaggcgcc cgcaaccacg cccggctaat ttttttgtat 201180

ttttagtaga gacagggttt caccgtgtta gccaggatgg tctncnattt cctgacctcg 201240

tgatccaccc gcctcggcct cccaaagtgc tgggactaca ggcgtgagtc tccgcacccg 201300 ggcatttata gttncttatc attaagtact gtatattaaa cctatctaaa tttgtttcag 201360

tactcgaaga tattgctttt ccctcgattt ttgcacagtt tcattatttn atgtgtaaga 201420

tgttggagat ttgaccaagt gaccccaggc tactgtgtta gtggccgaaa cattgtcata 201480

acattcacgt gcaaaaaaaa atcagatcca ttaagtttta aattgtttag tttctaatat 201540

tacaagctgt tcctatatat caagttagtt gattcatttc aaaataagca gtccctcttt 201600

gtaattttta cgtcattcat atttcttatt ttcaaggact aactacatta aaatatataa 201660

gatagaaaaa gtaatacttt aatattatga aattaattca attttcaaat atatttcaag 201720

caccaacaat gcacagtgct agacaatgaa agaaaatgta ataaacaaag aatacattat 201780

acatatccta aggagtttac aatctattgg aatttacaga aagataaaaa gatgtttcac 201840

acagggcagg gatcaacttg tgctgggtgc ctaggagtcc aaaggaagag cgaagatttg 201900

gtgtatgcat ggcatgcaca cacgtgcaca tgcatgttta tatgaatatt ttaactccta 201960

cacatagtgt ttttacaggg ttcagaacac atgatattga aataagttac aataggaaat 202020

cttttttttt gtttgtttgt ttttttgaga cagagtctcg ctctgtggcc catgctggag 202080

tgcagtggca cgatctcagc tcactgcaag ctctgcctcc caggttcacg ccattctcct 202140

gcctcagcct cccgagcagc tgggactaca ggcgcccgcc accatgcccg gcttattttt 202200

tctatttttt agtagagacg gggtttcacc atgttagcca ggatggtctc gatctcctga 202260

cctcgtgatc cacccacctc ggcctcccaa agtgctggga ttacaggcgt gagccaccgc 202320

acctggccta caataagaaa tcttgaccca tcatctaaaa ttctcttcac aaaattcaca 202380

ttttaagata ctgttcaaag cttcataaca tgatattata atggtattca cctttatata 202440

catccataat caacatttta atcttgctga tgctcacatt aatcaccttt taaaattttt 202500

tttctaagca taatgtcaat tccnagagaa agtttatttt atagaaacat cttgccacac 202560

acatgatctg tcgtatgcac tttgggatta attatgaatt cacgggtgca tgaaaattca 202620

aatacagtca aatgagatac caccttcgac ccatgcaatt ttaagtaatt gcttattccc 202680

aattactata ttttatttgg gagacttatt caataaagtg gctcgttaca ataattacac 202740

tacaactttc tagcagaaat gacattcatc aagccatatt cattggaccg tggttttgtg 202800

tcttatggtc atatgttttt ataaccactt tcccatttat tattaacctg ttaatatttg 202860

taaagtgaga aaagaaaaat ggagccactc aaactttttc ccagttagaa acagaaatgt 202920

attgttaaca ttgtaaatca aaaaagctat tgaatattta taatacttgt ggtgtgcttt 202980

gtcaaataaa tattgatggc ttcattctga aggtggtttc tagggtcaag ctatggttct 203040

ttaggtctaa gcattgccgc atcctaaagg aatctatatt caagataggt ctgtggttac 203100

ttctgggtga ttgaattaat caaattaata ttctgattga ttgagagttt tgatggatcc 203160

cagattttaa ggaatcagta agtgtaaaat ataaaaacta tgatcatact agctttaatg 203220

agttaggaaa tgtattgatc tctttcagtg cttcagagtc attattctga aaatcaatta 203280

tgcatatgta gaacaaacaa aaagttggtt agtttcgatt ttgcattagg atgtgagaat 203340 agcaagagtt tataatatgt gtaaacaaaa tatttcaagt caagcagata atattttgtt 203400

atgatatagt ttagaaaatg ctttttttata ctaagcctag tgacatttta catagacata 203460

atgcagtgaa aagatttgaa cactgaagct gcatgatcca gagttctagt aataattttg 203520

caatttatta tctgagtgaa cttagacaac ttaattttta tgagccttag atctttaatc 203580

tggaaaaaag agataataat accttaccta cctatgagga ttagtttgaa gattaagttt 203640

aaaatatgta tatatcacct cagcacagat caaggagcca agaaatacag atgtagaagt 203700

tataaaatgt acatttttaa atgtgcaatg ggcatctata aaattcttca ttcaatgact 203760

agagaataac attctaattc aagaacacat gtatgatttt aaatgtaacc atttaccaca 203820

gacttctaaa tatatacaca tcagtgtcct agaaaagata ttatttgacc aaaatgtaac 203880

aaaatcngaa acctaataag aagaaaataa ccacaaaaaa acctacacat tcagaaattt 203940

aaaactattt tcaaataagc acagaagaaa caccaatgaa aataagaaaa tatctacaat 204000

taaaaaataa tgaacaaaag atgcatcaaa cttgtgggaa gtagctaatg gcatttttag 204060

aggaaaaaat ttagccctga atgaatatct agtaaagaaa taaaggctta aattaatcct 204120

ccaagcattc atttcaagaa tttagaaaaa ttatcaacca aggaaacact aagatagcag 204180

aaagaaagaa ataataaaga gaaaagcagt tattaatcaa atggggggaa agaggtgata 204240

aataatatca acataattga aattttgttc tttggaagca gtaataaaat tgataaatct 204300

actttaaaat tcaacaagaa aaaattaagg agccttcttc tataaagatg acagaggcat 204360

tgacattgtt ttttaacctt gaattcccac atgaanaaca aaaagaatac caagatatca 204420

aaaccccaaa tccatagtta atatttacca taaaactcgg taatccacaa accctaaata 204480

aaagcaagtg ttgacaaaca gccacatgat ttgcatgata ttagtgtctg tgggaggaaa 204540

ccaagcttag ttacagagcg tccaatgttc atggaatggt agaacatctt ctttgctcca 204600

ttctgaaata aaaagtccag gagtggaatc aaaattgaac aggccaggaa caagagtgat 204660

aaggataaaa gcagatccag acaaaatttg tggggggagg aggattggga aacagagctt 204720

ggaaattcca gaaagaaagc taccatattt gatcgatata ttaaaacaac agaagaggaa 204780

actctgaagt tagaagagaa atactaaact gtttgggaaa actaattttg tataaaaatg 204840

agtaaggaaa attatcaaag ccaaattccc tgcaaacgtt attataagaa aaaaacagca 204900

gaatactgtc tctcactgta gacaaggaaa gcaacctana aaaacgtgct ctaaccagct 204960

caaactgtag catactattg aaaaattagt taaaagacat taaaaattac ataaggatga 205020

aatagcaaca taaataaaag ttagaaaaac tcagaaatga ggtgccagaa ctgagaaaag 205080

ttataaatta aagcaatgaa aactaaacga gaaggaacat aaagagacct cagatgattc 205140

aaaagtgaga aatagaaaat ggagaagagg gcatttttta aaatcaaaga gaaatgaaga 205200

aagagatgaa aatgatttgc aagaaaaatg acaaattgaa gaaagtcaaa aatccgacat 205260

ctgcattata ggtgtcccta aggaagaaaa accacagcaa gagaataaaa cgtatactaa 205320

acttttctga aacataaaaa gatatgaaac tatatattga cagcacacac ctaaaaatat 205380 tggagtaaat gattaaaaac aatacataca ccagacaaat gacaggtctt taaatttttt 205440

aaaaatatat tatctagccc aaataagaac aagtgattta taaaggaaag aacgttagat 205500

tatcatgaga tttttcagca gcaatgcttt atgtcagaag aaaatagagt aacatatata 205560

ggatgcccag gcaaagaata tctgagccag atttttttcc atcctgaaaa ctgactttca 205620

actttcaagg tcagagaaac tgctttcaag atctaagaat atgagtaatg taagatacaa 205680

gtaatcttgt ttcaatgaac ctgtaatcta ctaaaaaaca tgcctcagat aacctaaatg 205740

aactccaaaa gggactctta aaccatttgg cataaaaatt cacattcata tgtaacaact 205800

aactacataa ggacaacatg aagagtgaaa tatgtaataa ttatttgttc taatatggat 205860

agagtacaga tatacaaatg aggaggtgaa ngggaagaag atatgcaaac ccttaagata 205920

ttcagtagtn ttcttggttg tgtattttca ttctgagact attgtctatg aaatgtgaaa 205980

taaaactgag tcactgtnga gacaatctaa ttctggcatc cctggtgctc ttgataatca 206040

agagtgctct ttgtggagaa aaaaggagat atagatatca tatagaaaag cttaggtaaa 206100

aataaaaatc ctaaagaaaa cgtggcatgt atacaccatg gaatactatg cagccatata 206160

aaagaatgag ctcatgtcct ttgcagggac acggatgaag ctggaaacca tcattctcag 206220

caaactaaca caggaacaga aaaccaaaca cagcatgttc tcagtcataa gtaggaattg 206280

aacaatgaga acatatgggc acagggaggg gaacatcaca cactgggtgc tatcggggtg 206340

gggggcaagg ggagggatag cattaggaga aatacctaat gtagatgacg ggttgatggg 206400

tgcagccaac cactgtggga catgcatacc tacgtaacaa acctgcacgt tctgctcatg 206460

taccccagaa cttaaagtat aactaaaaaa aaaaaaaaaa aatcctgggt catgaattta 206520

aattgaaaac atcagaataa acccaggcaa aatatttgaa acagtgacca actccatagc 206580

agtgagcatc tatgggctct agatatgttc ttgaaacacc atttttccaa ttttccaata 206640

agtaaaccag ggcttcttga agaaaatggt tgattttgaa taggaagtgt acaagctgac 206700

cttggagcat cttgtcatac taggaacaag gaagctttgt caaagactac gagggtcata 206760

taaaaagaac ctacagactg accaacttga aaaactccca atgagcacag tagagataat 206820

ttgactaaca gtaagaaatg gtaattgaca tggactgaaa tccattaaac attaaataag 206880

tttaaatctt gagttcataa tgacattntt ttaaatagta attggttacc ttgagcgatg 206940

atagaaaatc aaccatgtaa aattaataag taaaagtcga taatcaagca tttatcttgg 207000

cttttccata ccagttttgc tactgggtaa ccaaataata caggagggaa agcacctcat 207060

tataaaacta tcccaactaa taaatgaaaa agaaatgaca gttttgcaaa cggaatgaag 207120

tcaagggatc taagaaatga acatcaagac ccgtaacatc acaaaaagaa acacacacac 207180

tctgttcctt tcgatggaag aatctgacac cacctatggc cttgccaaag ggatgaaacc 207240

agagtctgta tgatgaaacc tctggatcca gctgccgatt ttcagaaaat agaagacaga 207300

agaacgtgtt gaactgtacc aagagtgtgc tttcagcaaa atccagactc tctcttggct 207360

ggnngtgcgg tggcttactc ctacaatccc agcatattgg gaggccacgg caggaggatc 207420 acttgagccc aggcatttca agaccagcct gtgcaacatg aagaaacccc atctcttaaa 207480

aaaaaaaaaa tgttacttag ctatatgtga tggcacacgc ctgtagtccc agctactcag 207540

aaggctgaga caggaggatc gattcagccc aggaggtcaa gactgcagtg agtcatgata 207600

attccactgc atgcactcca gcctgggtga cagaacgaaa ccctgtcaaa aaaagaaaaa 207660

aaaaaatcca gactcaaaaa ctataggtca aatggcccag gtttctctat tttaagggga 207720

aaaaagggta acatgtagat taattagaag acatcgccaa aaaatcagca agattcatac 207780

attggaatgt ctttcaggtc caaattattc aatcacaaat aactctatgg gaattgagtt 207840

gaagcttttg aattatgtgt taaagtatgt aattctggaa acgtgtttta acttagaagt 207900

tcaagactta aaatatctct atattgttga taatttaaat ttaatttact tgttttgagt 207960

tattaattct gtcaaagaat tgttatctga ggcagttgga tgttttcagt ttccaaacaa 208020

aaccgaagaa aatcttcttc aacgaatctt tgactttttt caaagtaaac tccttgtaat 208080

tccatcattc atccccttcc tttttattcc cacatctctg atatgtttac atccagaatg 208140

agcaattact ggctcctttc atagctgtaa tttacaattt cagatctgaa attactcatc 208200

atgaaatgga tctctgtatt ctttgttttg ctaattattt caggtgattg tcagtatctt 208260

ctagagccag aagagggcac tgcagagcag tgtagctctg agctctttat gtttagcaca 208320

gtctaggaat gccctttgcc aaaaaaaatt cgcaaagctt tactatagaa gtaagaaaaa 208380

agaagggaat aactagataa tcactatagg ccttctttat ttagtagaaa aaagaaagta 208440

tataacgatc taaatacata tctaggacat caaatgtaat atcatatata tatgtataag 208500

tagatgaaag aagacatata tatgtgtgtg tgagtccttt ccttctttta agaacatccc 208560

aaaggccagg tgtggaggct cacacctgta atcccagcac tttgggaggc caaggtcggt 208620

ggatcacttg aggtcaggag tttgaggcca gcctggccaa catggtgaaa cctcgtctct 208680

actaaaaata caaaaattag ccggtgtggt agtgcatgcc tgtaatccca gctactaggg 208740

aggctgaggc aggagtattg cttgaaccca gggggcggag gttgcagtga gccgggatca 208800

caccactgcc ctccaacctg ggcaacagag tgagactctt tctcaaaaaa caaaacaaaa 208860

caaaagaaca tcccaaagct aaatctgatg tttttctatt ataatattga ttctcctgct 208920

acctctggca ttagttataa aaacagttct caagtactct cctgtatgaa ttaatactgt 208980

attttgtctt agttcaaatt caggacaaag gccattcagt agccttttag caatcttgtg 209040

cagggcagaa atcatttagg caatagtgta ccagatatta tcccaaaagt ttttcctaac 209100

gtgaaattgc aggggcagta gataagtgat tggattcaga ctctcatctt ccactgaagg 209160

aaagacaact aacacatgtc attgtccttt caagatgaaa taagagttgt ggaatctgga 209220

gtcctctctt aaaactcact gtcacaggac ctctgtgcct gggtctgtct caggactatg 209280

tgtgttctta ggtgtctaaa gagcagatgc tccacgatga tggggccacc agctgagtct 209340

tccatctcca aaagcggctc caagcagctt tccttaggag ctgagctgtc aactcaagag 209400

aggttaaatc catgtcaaaa gtcaactttt ggccgggtgt ggtggctcat gcctgtaatc 209460 tcagcacttt gggaggctga ggcaggtgga tcacaaggtc aggagttcaa gaccagcctg 209520

gccaatatgg tgaaacccca tctctactaa aaatacaaaa atcagtcggg catggtggcg 209580

ggcacctgta gtcccagcta ctcgggtggc tgaggcagga gaatcgcttg aacccaggag 209640

gtggaagttg cggtgagctg agattacgcc actgcactcc agcctggatg gcagagcaag 209700

actatcttaa aaaaaaaaaa aaaattatgc caaaagtcaa ctctttcagg aggggagaag 209760

taggaaagtt tcaagattta tcattggaaa tagcatgtat gtgctaccaa gaaaatggaa 209820

gttccaaata gttggtatcc actcttgtct ccttgtccat acattaattt ttctcctaag 209880

agttctagtg aataatgaaa ccaattccgg tttttcattt ttgagcagaa aaatgaacaa 209940

accacctttc ctcatgaggg ggaaatcacc agtatgttca gtgatgctga tctttactac 210000

atatctgctg taaaatatga aaagtggatt gctagttttg atgttatcta ggttatcatt 210060

tgtgactaaa aagacctgtt tgttactctg gttcttctcc actatggctt tatggacaca 210120

ctgactgaca ttatagcatt tgttcatagc tttgtttgta gaatggggtg ctgtgagttt 210180

gtacctcaac atacctaacc cttgttgaaa agaaactaac taggtgatct atttttactg 210240

aatacagagt tttctgtctt ctgagggttt tctttccccc ctcaactttg tttgacacac 210300

agtatgccct gtagggcttc ttgcctggtt tcaagatctg gctagcacag cacttataca 210360

ggtagtttgt accaaattaa cttgccctgt gcacacatgt ccacttcagt ccttcaccca 210420

gtaccgcctt tgcctctagg aatcttagcc ttccttcaag gcccagtata acctccactt 210480

tttcattgaa gcctttgccg actagtccag acttgctgtt ccaccgcaga gagctccagc 210540

ccctctgatc ttgtgttttc tcctttacgt atcatctagc actgtgccat ctactgcact 210600

gtacttgtca ctgatgatct catgtgtcct acccttgtct tcctatctag gctattcact 210660

tctagaaacc aagaactcca aattctattt accctaaatc tcctatggtc actaatatca 210720

tgctaagtgc tcagtaaatc taatctattg atcaaaatga acaggtctca attactatct 210780

tcagtctgac tactggagtt acactttcct gttatttatt atttattcta gttatttatt 210840

attgtgaatc tagttattta tttattatta tgagaacata gtggcccatt tgcctgtttt 210900

ctctttgccc ttcttctgcc tgctttaact ccttgtanac cttccaccaa tctaaaatct 210960

gcaactagca agccacacag aactgagctt ggggagtata tctttgaatt tacagtcagc 211020

atccaggatg gttttctcct tagttctctt ctctggagta attttatcat cgccaactag 211080

gaagagatct tgaatgccca tttcattcat tccccacacc tgcaagcaaa acaactccta 211140

aactcattgt tcagattagt gttaggttac aggccaagaa cctgcagcag acataatgtg 211200

gattagagaa cattgcgcta atcaaagaca aataaggagt ccagtatgag ggtcaacgtt 211260

aaagtctttg gaaactgaaa attttattca gacgaaggcc attatcagag tgaggcacag 211320

ccaatgcacc atccgtggct gggacgtggg gccccacaag tttgttcatg gtgctgccca 211380

aataaaagta aaaatagaca agaattgtca attgtacgat gatgctttag cgaaacatat 211440

tcatgtttca tgttttaaca atgtcatcta ttgaatgatt tgacagtata ttatttctta 211500 acattccaaa actaaaagtc atactcttat cattttcagc ttcgcaagtc tgttttcagc 211560

atcagggcaa ggaaccttct ggaaaaggag accgcagtca ctaaaatctt aaggtaacac 211620

caattaatgt tgacacatat ttaaacacaa gtgggattat attaatcgaa aaatagcttt 211680

ttgaaaaatg aattaatgtt ctgggcatca gcatatactt tctgccagaa agtagctttg 211740

agtatctatg tacttataaa caatgtaaaa gacatattga tgtcatttct aaaaaaaaat 211800

tttagtagct cagtaggcct ttctttcaag tatttaaaca taaatgataa gtataaacat 211860

tgaccaaaag gtgaaaataa cttactgtta tttttaacag ttttatatct tttgtgatat 211920

gtaaaggaga aattgcaatc taaaggataa atggtaaaag ggtaaacttt ttcattatag 211980

ttcattacac tattaagagc acatttatta aaaatccnga aataataaaa agtataccaa 212040

tcttagaatg cctatgaaaa catatgcttt taatgctact ttctagaaga tgcgctgaaa 212100

tgaagtatca aggaggaaaa ctaagcaacc gacttgagtg ttttataaga ctatatgtta 212160

tattatggag taacatgatg agtataaaaa cacataggca tttagaatat catttattgg 212220

tacatatatt acatagacat atgtagatga cacaactcaa gattgggagc tggaggtagg 212280

atcgctcgca aggcgagtgg acagattcat gccaggaagc cccctgcaaa cctgcggtga 212340

gtgataagga gttctagctc ccttctacta gatttattac taggaataga cattaatgca 212400

ttctctccag aatgcatatt actataactt ggaagactgt gtgagaaaat aaagcttagt 212460

atatatgatc tttaaaacta ttcttacttt ataatactac ctaaaatttg tgtgcaaaaa 212520

aaatttccta actcatcaag acagtatttc tctcaatatt ggctgacgtt catgcataga 212580

gttaagtatg tgttggagta gctgtcatag gatgaaaata agcatttcta gggcttgtat 212640

attatcacca caaaatgtga tgagataaac agacaaggtc tcattgaaat agacaacctg 212700

cgccccaact taggcttact cccgactggc tggtaactcc atcctggtat taccatcaga 212760

tcatgcagtg aggctggcct tgaaccacac atgttacatc aggatcatga ttccagctgc 212820

acctacatca gacaggttca tggtagggtc ctacagcaag gctgatccac tcactaaaaa 212880

gagaaattca ccctcaagcc agaaactacc tacaaggaca tcataattct cagtctgcat 212940

cgagtatctg cctgccaagt caggaactac agggacctta ctcacctgtc cagagccact 213000

gccttgcagt gtgaatgggc cctctgtgtg aatgggttcc cagagacttt ccaaggcaga 213060

tgcagaaaca gatactggca aggaagagat ttccagacct cacttctcta ggaactcttc 213120

cttaaaactg atctatcact ttaccttttt caccctcctg tccttattta cccctctctc 213180

agtttacaaa agaaaggtga aactctcact catccggaaa ctatggcgtg tattaatctg 213240

ggatgtaaaa tcctctgggt tgccaagcaa aaggaccgtt cccacctgga tttctccaaa 213300

aagagggtcc catggggccg tgccttatta tttatccagg tgacagtttt atggataact 213360

tccttcccaa agctatccac cttagattag gccttagaaa taacatgatt gtcacaaggg 213420

catgtggtta catgttattt gaattgaaaa ataaagccag actttttctg acagattcag 213480

ctgattggtt caacaatgaa aactgattta ataagtgccc aatgtcataa aattgaatag 213540 attggatggc gaagagtctc caatattttc ttaaaattat tttcaagcac aaacacaata 213600

atgctagaaa ttacattctt ctatttaacg tggtgcaaga gggatatagt tgctgatttt 213660

taaatataca gaagtgtttt tgtttttttca attatataat gtacatgagc aatattttta 213720

agactatcag ggaatataga cacctctaac ttgcctaaca atggtgtagt ttaggggata 213780

cccacagcca ataaataata ttttactgca gcccagcata ctacagaatg agaggcagca 213840

ccttaaaact acctgagcta tatcaccttt tgaaactcag gaaataaaaa gccataattt 213900

gaaaaaataa aaaaccttag aattttttaa aatacaaatt tgatcttaga aagccatttt 213960

gtttttttctt ctgtgcgaat ggtttcttca ttgcttctat gctagtttct gcatgaatta 214020

agggcaagtg caaaaaatca aactgatatt ttgggaatac ttatcttttc accactaatc 214080

gtagtaatag taatggcaat gataccatac ctgcattata atgtgatcca tcagttaaaa 214140

agtcaagaat agattgggtc catgtcaatg agtgtttgta ttgttttata ctgtttttca 214200

aggaacaata acagctaaaa ggaaaaaaag agggagaaaa agaatttttt ttnnaaaaaa 214260

aaaaaaggaa aacatttgca gtaacctcag gacaattgat tcgaataacc ttccacaaaa 214320

ttgggaagga gaaaaagctc tccaccactt tttcagtgtc tttttactgc ctctagggac 214380

tccgctctcc taggctccaa ccttgctgtc gattaaactc tgccaaggtt cctggataac 214440

ttagcagtgg ggtctcccag gaccatttgc attttgaaag aaggctttta aaaaatgaaa 214500

gggcttttaa gacaaagagg atgcagattc tgtcttttaa tagatgggtc tatcttgata 214560

taaataaaag aaatctggaa gagactgccg ttttagtgtt ctattcctct taactttcta 214620

ggtgttctta atttcataga attattgttt ttagtcttat ccaagaggtt cctcgccagg 214680

gaggtaagta acagcagtat agggaggctt gaaaaagaga cctgagctgg atcattccgg 214740

caaatatcac aggttggcat tcgtcctaat ttgtgtgcag atagagataa atcacaaaaa 214800

ccttaaatcc agagtgatat aagaatagga agtttttatg aaatattagc gtcttgagat 214860

aactttccaa aaaataacaa gtatttggga aaaaataaag caaccccttg acttaaggtt 214920

tctggaggcc tggctgtgta aacagacaat taatttcaat tctcagctct cttgtctatt 214980

ccatgggagc taggaggagg cgatcttcag ggcccctgaa gatctaatgt aggatactgg 215040

acttggtgtt aaatcccccc tggctgggat caggctctgc ttctctgtaa cattgaacac 215100

ctcactaatc tgaccctcag ttttcttatc taccttaaaa agattgtctg aataatctat 215160

tgaccaagca aaatttaata cttattccag taagggaaac tgccactttg ttttcataat 215220

gtctcagaag ggataagtga gaggtggata tatgtaaggt tttgggtttg ggtttcagtg 215280

tttgtaggat ggaggtctgt ttagactggg cagaattctt gatatgagtt gaggagtatt 215340

gaagatgcct ggtatggaca agtaatagtt actgttctga agacggcagt gctgtggttg 215400

aggaccagaa agagaaagcg acctgcctgg tgcctctcca ctgctttcca tgttctttcc 215460

cccggctttg gttcaacatt tgagtaatcc attctttaca gcgagtttgc ccttctattt 215520

tcattctcct tcttgacttg ggtgatctaa gagaagaaaa ataaatggtg accttatgtt 215580 gtcctcttca cacatgaaat ttattttttc ttttgtgcct ttcagatttt taaccagaca 215640

catgtattac ctttaacaaa caaatacaac tgacttttat agtggttttt tttttttttt 215700

aagtttggca ttgatcccca ttgggtagag tcttcacaaa aaatcctaat gaataccaac 215760

tatttgtagg tgactagctt gaccagttga gtgacctact gtctggagaa aggggttggc 215820

ccagcagtct atagttcaga taaatgaatt tgtgggaagt gcctgaaatc aacggtgaag 215880

ttgctgactg tttaacagct ttaccttcct ggacaagtca agtcatattg acgcagaggc 215940

tcagctctca gccctggaaa ttaagctgga gagctgccag gggtcttagt tctcaggggg 216000

gacccacatt tccaaatatc catcttccat ttgaaacagg atccacatta ttctagcaga 216060

aagtatgtta tatacattac tatgaacagg aaattccagc ccacagagga agaagtcttt 216120

cagagaggaa gggccctaga gcctgaggta tgtaactcac aggtagaatc acccatttct 216180

atgtataatc atctttaatt catgttcttg tacttttgaa aattaattga tgttttgaga 216240

aatataaatg aatttttga tatatttgag aactctttga tgagaaaaaa aaaagtgggc 216300

catttttaca gaacttgagg ctactggccc atgctaaaat gcagatgatt cacttggcac 216360

aaacaggcat ctgagttgtt gttttttaag agttatccga agaggatgtt ccactacctg 216420

aattataaaa tccaggattg ctattaaatt tattgttgtt ctgacttaga ctgtatcgtt 216480

cagctccact aaataaagct tcttttcaaa aggcagtgac cttttctctc atcaggcgaa 216540

ataaaaaaca aatagcaatg tactttcaaa agcatcacta cctttgttga gatttcctct 216600

aattcaaagc cctgaaatcg agcaaaaaat gcattatcgt tcctgcaagg ctgaatgaat 216660

agagaaatgt taaaataaaa caattaaaca ttttacatta gtttagaaag ttccatcctt 216720

gtgtggtttg ccatgagtac ctccttagtg aggacaatat atttataacc tttcgaatgc 216780

tcagtctcgt atgatgtttg cattctgaca tttggtttta taacgtaact ttctttgcca 216840

tgcacttata ttaataactg tttaagtctt aaggaatagt ctttagttta gtgctactta 216900

ctacagaaaa aaattaaacg gatttaataa aatacataga gatagttttt ctcttatcat 216960

tttaaaagaa tcatttaatg actataaata caacttatat actacttaca tatatatata 217020

ancactatgt atttattaat ctgcaattat gtaaatgaca ttagctggat tttttttcct 217080

gaaagaaagt gcccctaaga attcacgcat tcattcgaga aatatctctg tagtgtctca 217140

catgccaggc atttcacacc atttgctttt tcccatttgg aaagttctcc ttttataatt 217200

attgttccat aggttagcca gctattcagc aactttggat aaggtcatat tatttctctg 217260

cagctccgca ttctgggcag aactagcaag acaccctgta cactgctggg agggcctgag 217320

taccttacag cgntcccctg gacctttgag ctcagctcct cagtggaaag ggaaggtgct 217380

ggactgaccc tacagatgag agaaagaaac gacagagacc cttgcagcct agggtttgtg 217440

ccaacccctg agatagcagc aatttttttc ttagcatttg gtatgaataa ccccaagttt 217500

taagaaaata gagtatgaga atgaatgtca taaatattta tctttgcaat tctatgaagt 217560

ctgcctggtt attgttggct ttttcttaag attttcctga cacttaaaaa ctctgataag 217620 ttagttgaaa gactatttgg aaataagcta agtagataga tgaaaattaa ctgtattcag 217680

acaaccataa actaaggaga caagatgtgt agatatatgt ttttgcataa aacataaaat 217740

agtaaagatt taaaagtggc aaaatataac taaaacaaaa ataatatatt ctaatcacat 217800

ttttagatag tacctgattt gcggtaatca aaatatgtta tattttattg agatagcata 217860

tgatttatag attgttttta gttaagccca aaagtgatta taatagtgtt atggattaga 217920

agggctacta ttgagtttta ccaaacagat gtgggaagtt ttttttaata gactattttt 217980

tgagcagttt taggtttgcg gcaaaatgga gcaaaggtac agagatttcc catgtacctt 218040

ctgcccccac acaggcacag tctccccaac tatcaacatc cccatcagag tggtacattt 218100

gttacaattg atgaacctac actgacacat cactatcacg ccaagtctgt agctgacact 218160

agggttcact cctagagtac tatagtctat gggtttgaac aaatatgtga tgatgtgtat 218220

ctatcattat aacatcatac agagtaattt gagtgctcta aaaatcctct gtgctccacc 218280

tattcattcc tccctcccca caaaccctgg caaccaatga tctttttact gtttccacag 218340

acatagaaat gcacatagaa atacatagat atgtgtctcc atagacatgc attgatgttg 218400

atggcagctt tattcctaat tgccaaaact tggaagcaac caagatgtcc ttcagtagga 218460

gaatggataa gcaaactgtg gtacatccag acagtggaat aattattcag catggaaagg 218520

aaattagata tcatcatgga aagacatggt gtaaacataa atgcctgtta ctgagtaaaa 218580

gaagtcagtc tgaaaagtct acatactgta tgattctaac tatatggcat tctgggaaag 218640

atagaaccat ggagatggaa agatatctta acaaggtctt tcttttgtta ttttgtcact 218700

ggagagtaaa gtagtaaaga aaaatggaat acatatctgt tgaaaattgg ccaatgcagt 218760

tcctgcagca tttacaaaat ttcactttta ttgtagtaaa aagttcagcc tttgcattgt 218820

aatttgcatg tttcggagtc agtagttggg ttactaaatt tgaagcccag tgaatgatgc 218880

ggggttgcag acatttttta tgctgaaatt taatagatgc acgaaaggca tggatttctc 218940

atgagttaat ttttcttttt tgtttatagt atgggtataa gtgcgtaaga taaagacatt 219000

caaggtttat tttattttct gcctacnatg aagttctggt cgcccatgta tttacatgcc 219060

tttctatgaa aacattttta tgaataacac aagattggga agaactaact ggcagttgct 219120

atgaaagagc cagatccatt tactcaggat agggctgact cacaaaggag tggggtatcc 219180

actaggataa aaatgcagct ttttcctcat ctcagcaaga ggccattacc gacgagctca 219240

cccattagca tccgctgttg tctctacttg ggaagatcac attgtcactc tagtggagtc 219300

aggtaagagc catttccagg aaatactcat tatctattct cttagttata cagtcatagt 219360

tttacataat taccactgaa agactatttt cctgttgtct gctaatgtga ttacattaca 219420

ttaaccaacc ttggataata cttttaaatg tattacttgt ggagttcaag atttgagcaa 219480

aatccctcaa atgtaatgct ctgcaaagtt atactgcagt tcctttagat ggtaattacg 219540

tttaatacca gacggtattg tgtaatctaa cttttctata atgagtgaaa ataatttgaa 219600

gactcatagg cctctccctt tcattacaga atgtataggc tttttccaca cccttgttct 219660 gggacattaa gaaagaagac atggagacaa agaagaaaag aaccggggac cccctgctac 219720

tgtgtctata aacaactagt ggtgtttaaa ggctaataaa gctccaaaga tctctcttat 219780

aaccataatc tttgggccaa ttttaaaata ttttagatgg agaaatgtca agaaaaatag 219840

agacaatttc aagacatgat tctctttccc caacgtccat aaaagattat atatatatat 219900

atatatatnn nnggctgtca tataaggagc aattgaaatt ccacagcaca aaggcaatgc 219960

agggggaaaa aaaaggatca atctttcata ttattttttc tcaaaggctt ggcaagaaac 220020

caggggtaat atttttgcaa ctgggaatgt gtgtttgctg aagaatattc ccacatttga 220080

atcatcctca ggggtcctca tcttgtttta ctttcaaagg tagacagggc ataacattaa 220140

ccccaactgg gaaaaggaaa atgcctcact gtaattctcc tccagggaac aggaataaag 220200

ggattacata tgcctgcata atgaagaatt tcctcaccca gacaagacaa aagacttcat 220260

tatttttcac cagacacttg aagtcttgtg cctcccagtt gcagaagctg aggaggcagg 220320

ccagtttttt atatgaagcc actagatgga gttcctaaga aactggcctt ctagtacttg 220380

ccattgtgaa aagatctagg gaagatgtta aaaaccttgg tacctgacac tacccagtct 220440

ttcctctaag aaaagtgaga ctgtccctga agatcacagt attcttccag taaaattagc 220500

cttgccttta ctctgggaag caaataaatg agcattttct ttctgaaaat gcacctcctt 220560

tgtccttaca tgagaccaaa aggcatctca atactaatga aaaatgtaaa gtggcaaaat 220620

ataacaaaat ggaatattca agttcttctc cttctgggta ggtctgttac ttcgtaactt 220680

taaggatgta tgcatctcat ttcttatata aaatattgga agtctcaaat tttgtagtta 220740

ctcactattt cggcttgcaa gtattagatt cctgtaagag gaggtatatg gagactatgt 220800

tcccagcaac agatctgatc tccagttaaa gaaattgaac taagttcacc tgctacagcc 220860

tgctctgctc cctcttggct ttcaggcttc cctaatagct ttcacataca acatagagaa 220920

agangtatga agtttttagt cgtaagaagc aatgtctttt ttttttcatt actgtctctt 220980

ctagtgtagg tatgcagccc attattcttg agtggcatca tattctaact tgagataaaa 221040

ttaagattat ttaacttgcc cctttaaaag tatattttaa acacatgctg ttttgtccgt 221100

gattattttc atgcttcaaa agcagggagt gatcagaaaa gaaaatattt gatcttcacg 221160

gaatttcaag gttcattcat tccacaaata tttattgagc acctactgtg gataagagct 221220

gagaaataat catgaccaat tttaaagaaa accctactct cttagagctc tttggagcat 221280

ggcatatgtg acgaactaaa agaacaaaga aaatggcatg gagagaacga ggttcctcct 221340

ggtaaagatg agaccagaaa gggaggttag agggacacta aacggggctg tgcaagccat 221400

gctgctcgtt aggcttagtc tttttaaggc agtgataata accgcattcg tgatctattg 221460

ctgcaggaca aattactcca aaacttaagt gatttaaaca acacgttttt catctcacag 221520

ttttagtcag gaatctgggc atcccttagt catgtctctc tccaagctgc atgcatttgg 221580

acttgaatgg gaagagcccc actcaatcgc atggttgtta gtggattcag ttcctcgtga 221640

gctgtgatat gaaggcctca gttccatccg gtttgctgga ttcctcccca agtggaactc 221700 tccgtagagc acctcacaat atgcagcttg cttcaccagg gcaagcaagc aagatgggaa 221760

gagcaagtca gcgagagaga gagagagaga gagagagaga atgaatatgc tatgtagagt 221820

gttataacat ccaccatgac gcgggaaaag atacaacagg aaataatagc aactaaaatt 221880

caaagaaagc gttgatgttt ccaaagcact gtcacattca taatcatctt tacttcttcc 221940

agctccctgt gaagtcagaa ggaaaggcat tattgttctc tgcattttac atatgggaaa 222000

tggaggctta gagaaagcac atcacctgtc taaggtccca cagcttataa aagctgaagt 222060

cagaagtgtg tgcttttaaa gcttatgttt tctcaactca tcttctccat taggcttcca 222120

gacaaaggca gaactttcca tcagcagctg cactgcaggg tgcacttctt tgaaccatga 222180

gaggctagca agttagggat ggggtcacac agtccccacg cacctgccgc taccctccct 222240

actcacccct ccccaccttt tcacttactc attctgttca tcaggcagca tgagggccag 222300

aaactgaact tcctaggaat gggaggctga gtgataaagt ggtccctaat caactcacag 222360

cctagtaagg ggaataacta atgataacaa aatgccatga gcgttaaaca agcttggtta 222420

ataacatact gtaggaacat ggagtataaa acaagtcttt ttttgtgctt tcacataggt 222480

gaaatattca tttctcattt catgttactt ttcagagccc atttatatca gtctgtcatt 222540

atactgtgca tgttaaaaca ctatttggta tttatttcat ggtttgcctt actgtatatg 222600

cttatctata gtaacattgn tattcccttt aaaaaataat aaagcaaact tttaccatgc 222660

ttttgtatga acctgggcac attggattag agtatatgta acctcttagc aggtctgatt 222720

gtgtccaact gaagttcata gctttctgcc aacactgtta ttttccaaaa tagcaaaatt 222780

tcttacatga ccaacaaatg tgacaatggc ttaggctgta caaggctgtt tcattatgca 222840

gaaacgtcct atcatgggaa tctttctata cctgtctcaa tntccaggtg agctctaggc 222900

caaagggtca taaaacttcc ttcccattgc cttgctcctg catccaccgt ccaacctcag 222960

gtttataact tccatctgcg gtccttccct cattcaccag gaacctgtgt gtggccagtg 223020

gtctccttgt ataatggaat aagatttttc agaatgttct cgttagacat atttatatat 223080

ctgtagaaaa gttgcttcat gctgtttagt aaatgaaaaa tagtgattat ttagggtttg 223140

tttccaagaa tacatttaca ttcagtgaat ttgtatgatt tttcatagat attcaccata 223200

attatcttca tccacattaa aagcattttt aaggacttaa atatgacata tttagtaagt 223260

tcttatataa tgtattacta tataataagt tcttatatga tgtattatat aagttcttct 223320

gtaatatatg tgttatattt aaatacttaa aaatgcttct attatacatg tatcataaga 223380

gctgctattt taagtgcgct aagcacatac aaatcttttt ctcaaagatc ttaattctca 223440

gacagatggt aatcatattc aaagccattc aagaaatgca ttttaagaaa cataaatgtg 223500

aaaataatat gtaaaaaatg gcagaattgt gtgttcaggt gttgagcaaa tagagtatca 223560

ttagatttgc atatcataaa atcattgcat tttcaacctg tgagctatct ttcagagcaa 223620

ggctttctct tcattttacc aatgaggaag ttaggcccaa agatcactca actacttcat 223680

aacagtgtaa cttctggcac acaggtcttt taatcccttg tgttattttc ttgccactta 223740 tcaactatct caagggaatt cctgtcattc ccagttggag aaaattttct ggagaataat 223800

ttttttttct caccaaaatc attcacacca caggtgttag gccacccaat tgtcttcact 223860

cattagcaan atcagaaatt aacacaagtc cagttattgt ccaaaaagaa catgttttta 223920

aaagtcttcg tggtttaaac aagtggtaac tggttaccag taagctgcaa gcttcagatt 223980

gaggtttctg taggtaataa ttccaaaaac tgtgtgcttt tttttttttt ttgaactgtg 224040

gttttaaaaa caacacataa catttaccat ctaaaccatt tctaagtgta cagttcagtc 224100

gtgttaagta tgtttacatt gttgtgcata attccagctt tagttacttc tctttatgac 224160

tcatttttgc ttgtaagctt atgttttcct atcttcctaa agttctacaa aatctccttg 224220

gaacatatat ttaatcattc caggagcctg agtggacaaa tccttcccaa agggggttta 224280

ttagaaacct tgtttaaatt caaaaccgac atcatgattg cctgtgtgta ttaatctgtt 224340

ttcatgctgc tagtaaagac atacctgaga ctgggcaatt tacgaaagaa agaggtttag 224400

tgaacttaca gttccacgtg gctggagagg cctcacaatc atagcagacg gtgaaaggca 224460

cgtcttatat ggcagcagac aagagaagag agcttgtaca gggaaactcc ccttttttaaa 224520

accatcagat cttatgagac taattcacaa tcacgagaaa agcacaagaa agacctgccc 224580

ccatgattca attacctccc accgggtccc tcccacgaca catgggaatt caagatgaga 224640

tttgggtggg gacaaagcca aaccatatca ctgtgggtag gctggccttc tgcttatggt 224700

gtcctgatag aagctctacc attattttg gaaaggaggg aggaggcaga aggaggaggg 224760

gcaagagggg agctttctgt ttattttcca gcccttcttc atagctgtga tacttctgct 224820

ttgtaagtgt tcttgacgca gagcgtaaag ttcattttcc tttataatta tacgttattt 224880

gctgtatttt gtgttttcat tnnggtgttg nccntttaaa ataataaaag taatgagcac 224940

ttatgtgtta ttcatctcaa gtgtacaaac catgcattta tcttgggctc cttttttttt 225000

cctagagctt gcacccgaca gcgcatgctg agcggatcat ttgaggggca gcccgcaaag 225060

gagaggtcaa tcctcagcgt gcagcatcat atccgccaag agcgcaggta aatacttggg 225120

aagacagggc ttcagagact ctaaaaaagg ggggttttac aggaaatggt cgctgtacac 225180

agtcacggct gtcagccctg gctgccagac aggcatggct gaggaatcgc gttctatttg 225240

ctgtttgtta caataatgaa acacctaatc agatccaaag attcaaaaaa atctctactt 225300

cacaatcctt ttccttagag atggccaaat tctcaccttt aagatcagaa ggaaaaattc 225360

tgaattttgt tagatgggat caaagacatt tggagagagt ataagcttct gctatgaaaa 225420

gaaatcatag gcatttgtgg cgttgtagca gaaaaagtgt tttattttga ttttatcatg 225480

gtcactgctt gaaaggccag tttgatcttg agatgtatta accagtgcct gtggttgttt 225540

taggtcacta agacatggat caaacagcca gaggatcagc aggggaaaat ctcttgaaac 225600

tttgactcaa gatgtaagtt ttaagcaatg ctttggattt cagaaacgtg attatgaaaa 225660

tgtatttaaa tgactgctta gacaatgccc agtgtctttc tgctgcattc ccatgtctcc 225720

ttaagacttc ataccagagg gaatcaggaa actcaggata tgaattatgt agccaagtat 225780 cagaagtaac tgttaacagc agaaaactcc agtcatccca cacaaagcac atgtcttagt 225840

taaaatagta acacaaaagt gtcattatca gttggatttg acttgaaatt aaagacattt 225900

taagcattgt atataggatt gatttgcatc attgactttg aatattcact gagaatcgaa 225960

tattctagaa tctagatcaa gggtctataa tttcagggct ttatatatta ggaagtgaaa 226020

ttttatttct agaagaattt tcctgtcaag gactacctag tatttcccag gaaagctttt 226080

tgcttttaga ctctaagtca tggatttatt tactggctct ataaatagtt atcaagtgtt 226140

tgctctgtgc tgggttctgt tttaaggaac gaggagcagt ggggtaataa gatagactca 226200

gtgccctcct ggagcttaaa ttttagaata ggagacagat aataagttca taaaaatata 226260

atggaatttc tggagtgatt tctttttatt ttccatcatc tcttcattac tgtgacagat 226320

ttgctttata atgtaaagta aaatgaagtc atgtaaataa atcaaaggtg gggatttggg 226380

ggcagggaat tttagatacg atggttatcg tagacccatc tgcattatcc cggagaccca 226440

aatgaaatga aggagacatg aaaaaaatgc agagaagcca agcgttatat ggcccagcag 226500

ctataccatg caaagattac aatgaaaggt ttgggaaatg catattaaat aaccgtaatt 226560

atgaagccct ggtaaaaaga caggcgttag gttgaatctg gatgaactgg cttaaaaagt 226620

aatcttgaat gttcagagcc aaggatggca cttgagaaag gagggctagc aagagaaagt 226680

tagcaaagcc ttggggttat ttcgtttttc tctttttcac cacttatctc ttatttcttg 226740

agcttttgga gtgcttgcat actacatctg ctttcattct gtagaaactg ctgttttagt 226800

ccaagtacat ggtttttatt aagtgttccg ggcaaaatca agagcgagaa gttggcaatg 226860

agaacacctg atatcaccaa taccctatag acatggttnc tagttggcac agagnctctg 226920

tttaccacca tgatttttaa tatccatagc tctatctgat tggctaccat attgttgaca 226980

tagaattgcc atcaaatcca acatatggcc tgcgcaactt agaactgcag agtcagccat 227040

attgcacgga gaacattttc aggtgaaaag tcaaagaggc aaccctcatt caaagcgcct 227100

tcaatatgtt tcaaccgaaa tgtacaacct acttaagaat caaatcagaa catgtctttg 227160

aaaataatgt gtttaaacat agacagcaaa ggccctattt gaatagtctc agtgccggag 227220

cttacttctt cnnagagact tcctacccta tttttcatgt agaagtatag taaaatccat 227280

ctcccacctt ctccatagca aaatgccatg gggttctgat catttccaat aaattcttat 227340

aatgtctttt ctcccacaaa ggggagattt aaattagaga attttaactt tttcacttat 227400

tcggaatgat agcangacac agataattaa gtctaaggag tgctttctgt aaaatctact 227460

gagttactgg aactctggcc ttgcaccttc catggggtcc atctaaaggg ctgtgaaggg 227520

cctgcatggc ctgctcacct ctcttagctt tacccacact ggcatcctgg ctattccaca 227580

tgtagacacg ccaggcctgc tcctagctct ggaagaacac atttctgtag tcctaccaca 227640

gcctgcaccc tcacctcctt ctagtctttc ctcaaatagc accttcatat tgaggtctac 227700

cctgacctgt tcatttaaaa tcataatctt ctagcctgtc tgtgtctctc tatgcctttc 227760

acccattgtg atttttaaa aaatcggatc actaagcatc ctcttacata cttattttga 227820 attattttt atctgtctcc cactgactag aacatcctat gagaataagg atgttttgtt 227880

tccacttgcc atatcctgca cactagaata gtgcctggca cataatagga tttcaatcaa 227940

tatgtgctga atgaattact ttttagagga acaaatactc ttgctatgaa gactgtggcc 228000

attaaatatg tgcatttaat cgtatacata ttttatatac tgtaaattta tatatttaat 228060

gtaagtatat acatttactt taaattacat atttaatgta aatatatacc tttagtataa 228120

atatacataa ttataaatat atacatttca tatagatata tatttattgt aaatatatac 228180

ttagtataaa tatacatatt taatgtaaat atatacattt agtgtaaata tacatatttg 228240

ttgtaaatat atacatttag tataaatata cattgtaaaa tatacattta atataaatat 228300

ataaatatat gtttaacagt aaatatatgt aattaatatt tatactttat attgagggag 228360

atattagttt taagtttttt aactgataat atattttgca ttcaaaaccc tcctgtttta 228420

acacaccccc tgttaattta ntttcntagc attccaatac agtgagatat agaaatgcac 228480

aaagagaaga cagtgaaata aagatgattc aggaaaaaaa ggagcaagca gagatgaaaa 228540

ggtacaggtt gctgcagtgt cccatggctc tcattcggga atggcaaaat ttcccttagc 228600

ctcattctcc catgaatttc tccttcaagg aaagtgcaag aagaggaact cagagagaac 228660

cacccatact tcgataagcc actgttcatt gtcgggcgag aacacaggtt cagaaacttt 228720

tgccgggtgg tggtccgagc acgcttcaac gcgtaagtac acttcatttc agccaggtgt 228780

ttggcctgag ccatgcaaac agcaggggtg ttgtcatatg aagaagaagg aacgtgtaat 228840

cctaatgggg tcacaaaata aactgcggac taattaatca tgctgctcct ctttgctgaa 228900

cacccgacca ggtggaaaac aaaggtcgca agtagttaac atttatagaa aactcgcatt 228960

gataaaacac aaaacagcca ttttcgtttg tgcattttaa cagtttcctg atacacttat 229020

atcggcatta ttaactccat attttatttt ggaagtactg agacaaaagg agactaaagg 229080

acatggtgaa aggaaggtca ttgtccatgt cttccatgag gcttgtagtt cctttgtggt 229140

aaaaatacct caatttgttt tttggagaat aacgtaaagt tgattagtgc tttcattacc 229200

aactgttata agaaggaaac ccactgttca ctgtctctac cctaccaaca tggaagaaat 229260

gtgtaagaag ccgtaagtag agtggacatt gaagacacta gttgtggaca cctaagccaa 229320

gtttgacagg caacaaatgg atgaaatgga tttttgtcaa gcatacctgt atnagtccat 229380

tttcatgcca ctgataaaga catacctgag actgggtaat ttataaagaa aaagagattt 229440

attggattca cagttccaag tggctgggga ngcctttcaa tcatggtgga aggcaaaagg 229500

cacatcttac atggcagcag acaagagagg gaagtgaaag cagaaacccc tcataaaatc 229560

atcagatctc atgagactta ttcaccacca caagaacagt atgggggaac cgcccccatg 229620

attcaattat ctctgactgg gtccctccta caacatgtgg gaattatggg agctaccatt 229680

caagatgaga tttgggtagg gacgcagagc caaaccctat cagtaccttt aagtaaatat 229740

tctgaagaca ttaagagtca tggaggaaat gagaaaactt tccatgttcc cattatgtag 229800

ttggagtgtc cctgatgcca tcgtttaagt aaattgattt ctttgttatt aattagtagc 229860 atatgataat cactatctgt tatggactga attgttcccc ctccacaaaa tccatacggt 229920

agagtccaaa ccctcagaac ctcagaatgt gactgtattt ggaaatagga attttaaaca 229980

ggttaattaa aatgagatca ttagggtggg ccctaatcta atatgattgg tttccttaag 230040

agaacatttg gacacagaca tatacagaga gaagaccttg cgaagacacg tagggagaag 230100

acggctctcc acaagtcaag gagacagccc tggaacaggg ttttccctca cagccctcag 230160

aaagaaccaa ctcagctggc atcttgatct tggatttcca gcctccaaaa ctatgagaaa 230220

ataaacttat gttgcttgag ccatccagtc ttaagcacca gtggtacttt gttacagcag 230280

ctcgaggaaa ctaatacaca ctattataac atttcaggga tggaagagac ctcagggacc 230340

actcggtgtc atttgtgtgc cccctccttc aacagaggag gacatttaag cgtgcagaag 230400

ttgagggctt tgctcaggga cccagctact acctggctct ctcttgattg ccagactatt 230460

gttctttctg ctgttctata aagactttca caatttaaag caactgagtc acaattatta 230520

aaatcttcca tgagttttca aagtaaagga agaaaatgcc attttggaga ttgatttgat 230580

cagttcttga agtacgtttt ttttatcttg aattatcctg ttggctgcct gtaccttaac 230640

atctcatggt taattctcca cttcacttag tgtgatgtaa tttccctttt atcagcaagg 230700

cagctgatgt tgaatacaat tgatgaagac tctgttgtaa ttggggacac atacacaaag 230760

actgtaagca atctcattga cagtgctcac aggcttcctg aaagtagaaa caacgactct 230820

cccttccaag aaattctaaa taaatcattc taccacatca cagttaaatg cagggggaag 230880

aagtctattt aaattatgga gtatttaatt aaataatctt gaataactaa cattcttgaa 230940

ctattaaggc aaagatttca aatgagtgtt tttaaatgaa agtcaaaact taattttgta 231000

ctacccatgt cagaaaattc tgaacaaagt cacaaagtct aaaatcaagt attttgagaa 231060

atataaaact ttattttcat aagcatcaac tgcaaattct ttggattttt tctttcatca 231120

atgttctctg acctccagag ttacctacag ttaacatatt ttataatgga cattgtagag 231180

aggctgtaaa agagatttta ggtagatgag atgtttgcaa atatgatttt tagcaagtta 231240

aaacatgctt gctttgttgc agaattttat atatgtgttc tcttacagat ctaaaacaga 231300

ccctgtcaca ggagctgtga aaaatacaaa gtaccatcaa ctttagtaag catcaccaac 231360

ttcattttcc atgcacattt aaagttttca tgtactattt tatgctgcat acaacgatta 231420

aaattctctt ttattttctt gctatatagt gctatgaaat agatgattgg ggaatgatgt 231480

ttttaaaata ccagtcatgt tgatagagcc acagaagcca tcatctctaa ctcttagaag 231540

cctttcctgg aacctttatg tttctcaatt ttgtttggtc tttggcagtg atttgctggg 231600

attggtcact tacctggact gggtcatgat catcgtaacc atctgctctt gcatttccat 231660

gatgtttgag tccccgtttc gaagagtcat gcatgcacct actttgcagg taccgcttac 231720

cgctgcaaaa tctcagctta caactaggag gtaaaatact gacctgctgt ttttccgttc 231780

attgtaacca aactgccttt tgcccgtttc agattgctga gtatgtgttt gtgatattca 231840

tgagcattga gcttaatctg aagattatgg cagatggctt atttttcact ccaactgctg 231900 tcatcaggga cttcggtgga gtaatggaca tatttatata tcttgtaagt ctttgcttat 231960

tgcctaaatg aaaaatgtaa tggttttgta aaatctcctg catttcttgg tataagcgca 232020

agaggtatga gttacagagg cttagattcc ctaaccgaat tcttaaatgg actctagagt 232080

agttcattat atgcaaataa gggccatgca tgtaaaatct aaattaagag catcataaat 232140

atctatggat agttatattg ttcaaataca ctaatcacaa ttttatgtat gaatgagaca 232200

tacattaata aggtctaaca tggcaaatct agtttgcatc ccacaaacct tatggcttct 232260

atactaaaaa aaaaaagttt gcctcatata aaccatgtta ttgattatga gccccttatt 232320

gctgaaagca aactaaagtg attgtgattg tatctctctc ctgattccat tgcataaaat 232380

acctggctat agcagttgcc cattcaactg tctagtaaaa ctcgctggaa aaccatagaa 232440

atgggctcag agaatttttt atttttcctg tatgtccaag gaggcatgaa atcaatgtct 232500

ttgcagagac aacaaggaga attgtgatca ttaacaagta gcatgttgca ggtcattaat 232560

aaaatagcag taaccttgat aaaaggaatt ccagccagat cgacatcagt ggagcctcca 232620

gagtagtaag ttaaaagaag aggcagaacg cactggtgta cacacacaca cacacacggt 232680

gtggacaccc ccatttcttt cctgcataca catgatgact ttcctttgag gtacctagca 232740

agtctagatg aggtttgttg tttaaagtta acagatggca cgtacaacca gtgagtggaa 232800

tctctgctaa atgaaaagtt gtggccagtc ggctgtcaga aaatgtaaat agtaggggat 232860

gaattgtata aattccagag atcatttgtt ccccctttagg tgagcttgat atttctttgt 232920

tggatgcctc aaaatgtacc tgctgaatcg ggagctcagc ttctaatggt ccttcggtgc 232980

ctgagacctc tgcgcatatt caaactggtg ccccagatga ggaaagttgt tcgagaactt 233040

ttcagcggct tcaaggaaat ttttttggta tgtgagaatt atcctttgat tccagttcac 233100

agtagtccac ctctaatgca gctaatgagg gaaaagtgag tcttgcttat tgaggtcaat 233160

tcaatagttt tggttattgc cagagggtaa tggaggcact aacaaaacac aatgcacctt 233220

caaacacttt actcatgggc acttctggcg tatgagttta tttaattcac ctaaagagaa 233280

gcactctgag atttgaaaga gttcntatct gctgactctt tcggtactaa cctcttgact 233340

ctaattccta ctctgggttt tagtttcctc ccatctgaaa atgaaggata agaagattca 233400

acacngtgta ctccacaaat gggattagga attcatgctt ttgggctcct aatctcattc 233460

tgtgtgttcc tgaaataata tagcacagaa tttgtcttta gatctcagga gntgttaaca 233520

taaaacaact ttaaaaatgc aacatatagg tataacccac ttcaattaac ttcagcatta 233580

cagtaatgag tgattacaaa ngatgattga gggggcaaaa cgtactttnt agctgggcaa 233640

acattgattt agggtcttct gaaaacaagg tcagagataa aacaacaaga aagaggctca 233700

ttgttcacaa tctgttacag aaagatcttt gccttacaaa acagtgatca aaagtttttt 233760

ttattattga tgaggtctgc ttgtgcatga gaagcatgaa ttaatttaca taggcttgtt 233820

ttgtatcttg ctttctagag acttacctgt gatgagaaaa cagctgtagt tgtagttgac 233880

acaaatgtat tttaaagata gattgtgatg ttactagagg tatatatatc attactgact 233940 ttttttttct ttttttttga gacagaatct tgctctgtcg cccaggctgg agtgcaatgg 234000

cgcaatctca gctcactgca acttccacct actgggttca agagattctc atgcctcagc 234060

ctcccaagta gctgagacta cacgcatgtg ccaccactcc cagttaattt tttgtatttt 234120

agtagacacg gggctgcacc atgttaccca tgctgccctt gaactcctga cctcaggaga 234180

tccacccacc tcagcctccc agttgctgaa tttctttttt aagatgggtt taaaaattct 234240

tgttttcttt atgcaaagcc ataaaacagt gactcttcac acagctggaa ttaaaagaaa 234300

atatcacctg actgctaatg tatatgaatc ttttgatcaa taataaattt tagaggagaa 234360

tttgtgcctt tccctaatgg tggtcagtgt gattttagta cacttaggta tatgtttagg 234420

ttattcaacc agtgcagttc acttgttcca tattttgccc tctgccactc gtctctgctg 234480

agaccaaaga agtcagccag cagagggagc atgagttaat taacagattt gctcttcagg 234540

actgaattct ctcctctcct ctatgcatag aacattctcc taacgcacac cttcacagtt 234600

acaagacagc taagccactg tcccgctcat ttgtgctctg ctaataaaaa tatgaatcca 234660

ggccgggcac agtggctcac gcctgtaatc ccagcacttt gagaggccca ggcgggtgga 234720

ttgcctgagg tcaggagttc aagaccagcc tgaccaatat gatgaaaccc tgtctctact 234780

aaaaatacaa aaattagctg ggcttggtgg cgtgtgcctg tagtcccagc tactcgggag 234840

gctgaggcag tagaattgct tgaacccagg aggtggaggt tgcagtgagc tgagattgca 234900

ccattgcact ccaccctggg ggacagagtg gcactccatc tcaaaaaaaa aaaaaaaaaa 234960

agaaaaaaaa ataaatatat atatgtccat ttgggaaaaa atagacaaaa tataattact 235020

taaacatttt tttcctttgt tccattaaat aaaaaagatc aaggtgttga ttttaaagtc 235080

agaaaagttc aagccatatc aaaaatacct tattccattt tcttacatga aattcatgtt 235140

tccatttaaa aaattcaaag cttgttaatt ttctcctttt gaccttccag taaggctaat 235200

tgtgtgacct atggtggaga agttaaaata tatcatcaaa aacccatgat ttttatacaa 235260

tgttatctta gaagtcctat ttatgtcact gtttaaacca aacctaatat ttagtcttta 235320

acaagagatt taacatttca tcaatttctt tcaggtctcc attcttttgc tgacattaat 235380

gctcgttttt gcaagctttg gagttcagct ttttgctgga aaactggcca agtgcaatga 235440

tcccaacatt attagaaggg taagatgtat ctttctgtct tttgaaataa aaattaaggg 235500

ccaagtgtgg tggcatgtgc ctacaatcct agcactttgg aatgccaagg agggaagatt 235560

acttgaactc aggaggtcaa ggttgcagtg agctatgatc tcaccactgt actccaacct 235620

gggcaacaga gcgagagccc tgtccctaaa aagaaaacaa aataataata aggtgtttac 235680

actcccttgg caagcatgtg gtaccagtgg ctgacaagcc tagtttaaga actccctata 235740

aaaccctaag aattctcttc ccacaaaccc ctttttgctt aagcatgtga tatgtctctt 235800

ctggctagaa aggttaataa gatgaaaagc acacctcaga gaatgtcact cttccctcct 235860

attccagtgt caacccactg ttgacaactt gagygtggtc acccaagatt ccagctgttt 235920

acagctgtaa gctttggata ctttcccttg aggtcttagc taacttgatc cttcctttgt 235980 tgtcacataa gactcatatc agcaaacagc cccatgcaaa aatgtccaag aagcagcgtt 236040

tggttctcac cactttttat tgcttcccca ctctcagctt tctccccatt agaaangcca 236100

gtgcctttac attagctcta gctgcccaga atctccctct ttatgacaac caacaaggac 236160

ctccacaaag acccaaactt gaatacttgc actacggcgt ttctttctct ttcttcagtc 236220

attggttata agggaaggct agctgctgta acaagcccaa tatttcggta gcttcacata 236280

atgaaatgta tatattgcat cttggatcag ttgcaatata tacatttgca atatatacac 236340

tatatacagc tctcctccaa atgaccactc agggactcac atgccttcta tcctatgccg 236400

ccaccctttg ccatgtcttc agagccttct tcaatgagct ggagataaca gtggagagcc 236460

atctccaact gagtgggctg tgggagaatc tgctccaccc acattcctct gtaaagagca 236520

gagatactgc aaggggggct ggaatttagt tgtgtgctcc agaaaaataa aagggggaca 236580

cagtttggtt aataattaga caatttttgt aatatcataa agtgtcccat tctcttagga 236640

ttcttatgtt acaagtaaat tcataaagta aaccaaccac tgctgatgta taaactagaa 236700

aagcattttc tctttgatcc caataaaatg atgcctcagt tcctaaatgt atatcttcaa 236760

aaggcctact agacctgaac tacattattt atttctagca aatatggcta tctagtatag 236820

attttaaatg atacaaattg ttctctaatt gtgacaaaat acaggtaaca tctttgttca 236880

gccaaaatac attataaatt atattaataa tgagaatttg aaacaaaggc cagactggga 236940

gagacattcc aagagctttg atacacctga gaagacagaa caaggaatgt ggcctgacgg 237000

tcatgccgta tcaaattagc tgataaagaa agtgacagag acctgacgca ttcatccaca 237060

agcttgtgtg tggaaagaag agtgattttt gatccacact gaacagaggc cagaaagata 237120

ataacacttc aggatgaagc ctgctaggaa tcttggatga ctcagtatat taacagtggc 237180

ctggccgggc gcggtggctc acgcctgtaa tcccagcact ttgggaggcc gaggcgggcg 237240

gatcacgagg tcaggagatc gagaccatcc tggctaacaa ggtgaaaccc cgtctctact 237300

aaaaatacaa aaaaaaatta gccgggcgta gtggcgggcg cctgtagtcc cagctacttg 237360

ggaggctgag gcaagagaat ggcgtgaacc cgggaggcgg agcttgcagt gagccgagat 237420

cccgccactg cactccagcc tgggcgacag agcgagactc cgtctcanaa nnaaaaannn 237480

nnaaanannn anaannanna aaannnnana aaaccggggg ccctaaaaac gggggggttt 237540

tactttttcc ccttatcccc ccccaaagtt aaaaattggg ggggaccgcc ccccccaaaa 237600

aaccctaacc cccattttta ggggcctttt ttccttttc caaaaacagg ggggttttgc 237660

caaggtttaa atggaaaccc taagggaacc tcccctttttt aaaagccccc ccccccaacc 237720

aaaaccttta aaaaaaattt ggcccgggaa aacggggttt aaattttttt gtttttcaaa 237780

aaccaaaaaa aaaaggccgg ggccggggcc aaaaacccct taaagaaagg ggaccccccgt 237840

taaatttttt ttgaaaaact tcccccccttt tttctaaaaa aaaaaataat tccaaaaaag 237900

gggggaaagg taatcccccc acttttaaaa ccccccccctt tcctttaaaa aaaaaaaaaa 237960

t 237961

<210> 3

<211> 47841

<212> DNA

<213> Homo sapiens

<220>

<221> exon

<222> 3895..4001

<223> exon 28

<220>

<221> exon

<222> 9611..9731

<223> exon 29

<220>

<221> exon

<222> 9816..9914

<223> exon 30

<220>

<221> exon

<222> 15776..15869

<223> exon 31

<220>

<221> exon

<222> 16382..16488

<223> exon 32

<220>

<221> exon

<222> 16697..16771

<223> exon 33

<220>

<221> exon

<222> 17934..18053

<223> exon 34

<220>

<221> exon

<222> 23644..23712

<223> exon 35

<220>

<221> exon

<222> 24928..25076

<223> exon 36

<220>

<221> exon

<222> 25913..26006

<223> exon 37

<220>

<221> exon

<222> 30767..30899

<223> exon 38

<220>

<221> exon

<222> 31561..31676

<223> exon 39

<220>

<221> exon

<222> 34044..34201

<223> exon 40

<220>

<221> exon

<222> 37493..37643

<223> exon 41

<220>

<221> exon

<222> 39652..39801

<223> exon 42

<220>

<221> exon

<222> 41563..41680

<223> exon 43

<220>

<221> exon

<222> 44131..45841

<223> exon 44

<220>

<221> misc_feature

<222> 45842..47841

<223> 3'regulatory region

<220>

<221> allele

<222> 31646

<223> 99-79310-29 : polymorphic base A or C

<220>

<221> allele

<222> 42373

<223> 99-79311-50 : polymorphic base G or C

<220>

<221> primer_bind

<222> 31618..31635

<223> 99-79310.pu

<220>

<221> primer_bind

<222> 32080..32100

<223> 99-79310.rp complement

<220>

<221> primer_bind

<222> 42324..42344

<223> 99-79311.pu

<220>

<221> primer_bind

<222> 42704..42723

<223> 99-79311.rp complement

<220>

<221> primer_bind

<222> 31627..31645

<223> 99-79310-29.mis

<220>

<221> primer_bind

<222> 31647..31665

<223> 99-79310-29.mis complement

<220>

<221> primer_bind

<222> 42354..42372

<223> 99-79311-50.mis

<220>

<221> primer_bind

<222> 42374..42392

<223> 99-79311-50.mis complement

<220>

<221> misc_binding

<222> 31634..31658

<223> 99-79310-29.probe

<220>

<221> misc_binding

<222> 42361..42385

<223> 99-79311-50.probe

<220>

<221> misc_feature

<222> 26,350..351,353,356..357,366..367,371,383,389,392,396,399..400
404,411..412,416,421..422,435..438,440,442,446..447,450,455
457..459,461,463..465,1805,2175,2328,2402,2561,4048,4661,6659
6799,7509,7534,8289..8290,8294,10663,10789,11052,12083,12904
13017,13916,14116..14117,14680,14971..14974,17093,18448,18669
18767,18808..18811,18838,18971,18985,20714,20903,21355..21367
22901,24093,24112..24115,24171..24173,24868..24869,25356,25797
25879,26062,26234,26824,27130,27254,27437,28745,28781,29278,29475
29526..29528,29617,31152,31790..31791,31796,32011,32262..32273
32296,32415,33656,33719,39140,39142,39343,39944,40945..40947

41948,42166..42169,43371..43372,46709..46720

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

<400> 3

gtttaaaaaa aaggtgaagg cctctnggct tttttcccc aaattttttt gggaaaaaaa 60

agattttttt ttcccccctc cacaagggcc caaaaaataa aaacccttcg ggagaaaccc 120

gcttgggaat tttgggggc tccgggaatt taccaagggc ccggcccggg gggggggtt 180

cccccctta ttcccccccc ttttgggggc caagggggg ggtttccagg gttgggggat 240

taaacccttt ttggttaaaa aggggaaacc ccctttttt ttaaaaaaca aaaaaaaaat 300

aacccggggg gtgggggggg gcccctttt ccccttcct tgggggggtn ngncannaaa 360

aagggnntaa ncccgggggg ggnggtttnc antaanccnn aatncccccc nngccnccca 420

nncggggggg aaaannnnan ancccnntcn cccanannna nannnaaaaa aaaaaaaaaa 480

aaaaaaaaaa aaaaaaaaac agtggcctaa gagcatggtg attttacatt atccactcat 540

acccagccaa agctgaagat tgatggtgac agcagccagc agatgagcat agacacagat 600

ttcatgatgc atttatcctg tttcctagag cagtgaggtc tgtgcgaggc tcagatggaa 660

agcctaagga gactcagcca tcatagagag caccagcacc aagcagaacg attgaaaaga 720

aattggcacg gtgagacggg gcttgagtct ttttgtcttc tcaaatccaa acataagagg 780

ccgaggccga ggccaaaaac gccatgagtg aaggtgacgt cggctgaagt tctgttggag 840

aacttcagcc catcttacta aaagaagaga taaatccaaa gaaaggggga aagtagatac 900

accaacattt agatcacaca cagttcctgt taaagacaac gaagcttcag ttgctgcgcc 960

catagtttcc agttgggaag atgtgttttc ttgtatgagg ctgagcactt cctccctctg 1020

ctcctttgtt tgttgtctta aaactcaaca gatgtcatga gatcaatata gatatagata 1080

tggagcctag gactttgtca ccaaaaccct gtaaacaaac gctgaggata ccttttgaca 1140

aatagtcatt aaatgtcagt tgtgtccaag gcactgtgat aggtgaaatg aaatgtggaa 1200

gggaaacatt aggaggtcct tgcttctgtt ttgtggcatt ttctgtccaa tcagaacgtg 1260

tttcatgttc aagcaacagt aatataaagt cagcagtttt caaaagccta tgagaagcag 1320

aaataacatg taatctatga tgtcagcatg ggaaattgtg cttccaactg ggagtactgt 1380

actgggaggg cttcagggag gaggtgggat ttgagatgag caggaaagga ccttccaagg 1440

agagcaagca gtgaagaagg acacggagca tacaaagaag ggtcacagtg tagaggcagg 1500

gaccgaacaa ggaggaagga gcatgggctt cgaaatcaca gcttctgaga ttggaatcct 1560

acttcaggag aggtgatggc ctggtctggt cacttaatcc ttctggaacc tatttctgcc 1620

ttgaatgctt ctcacacaat gatattatga agattagatt accttgttta tgcgaagtac 1680

ctaacctgat ctcaacagtg agtgttagtt tcattcccgc ccatgtagga aaaaactgaa 1740

gaacaaggcc agaaatgtaa gttagggaga gatcctcaag tgcctgaaat gcctggctaa 1800

ggggntttgg aatttcgttg atgagaagtg agaaaccaca aaaggttggt aagaagagga 1860

```
atcagacaca tcctagtgcc cacaaagtac aaagcaccat tcctagtgca gtttttgtta   1920
cccatatcct aatgaggata tccattccta atgaggaaaa caggtgtaga aagattaaat   1980
aatttgtcca aggtcaccca cctaataagt ggagaagcag ggaactagtg tccaggcacc   2040
ctggcttttc accaagatac cagactgctt catgcaatca aaggtgggct caggtgggtt   2100
caggaagaca aattggaaaa ggcagtgagg ccaaatatgt caagggacct ttatacaata   2160
cactacaaca cagcnctttc caaataatgt tctacataaa aaggtatggt tatggaaaga   2220
taatcgcaat gtagttttag ctgaaaaaag tgggcgatgg aagaatataa actttatgag   2280
ctcagataca tataggaaaa aagactggaa ggatatactc caaaacgntt cattgtgttt   2340
gtctgtggca attgtgggta agttttcttc cttttatttg tccaggcttt ataaattaaa   2400
gnatgtgtta tttgcattaa aatattctta ttttaaagaa tcatgctggt aatccagaaa   2460
gaagtaacgt tggcattact aaactcatcg tagtaagaat ggagacttgg gccaagaagg   2520
acagggaaag cccagggaac cctaagcttg atgctgggct nctgctggga tgctggtcca   2580
gggacagaaa ggaggaggcc aagaggagac atggcttagg gtcaaagagg acgtgaagca   2640
gattaaaatg gaatttttgt tgaaatattc aaacagtgat caagaatttt accatctaca   2700
ccttgaaccc agaaaagaga tttaggttca aaatatgacc atcatagcca aagaaacaaa   2760
aataaccaaa gggagaatag gaaatgagtc aaaaacttcc tatagatgtt taattttaa    2820
aacctttta acagctttat tgagatgtaa tccacatacc acacaataca cccacttcaa   2880
gtgttcaatt caatggtttt ttagcgtatt catagagttg tatacacgac agccactttt   2940
agaacatttg tatctccttg aaaagaaacc ctatgtctct taactattac ccccatccct   3000
ttagcctcca cctcaagacc taaacaacaa ctaatattct ttttgtctct atagagttca   3060
ctgttctgga cttctatatg aataaaatca taacatatgg tcttttgtga cttttttctc   3120
aattagcata atattttcaa ggatcctcca tgttgtagca cgtatcagga cttcatccct   3180
tcctatggct gaataatatt cacttgtatg gttggaccac attttgttta tctgtttgtt   3240
caaaaatgga catctggatt gtttccacct tttggctatt gtgaataatg cttatataaa   3300
cattttatac aagtttctat agggacatat attttcatct ctcttgggta gatacctagg   3360
agtggagttg ctggctcata tggtaactct gtgttagtgg tttgaggacc tgctagaatg   3420
ttttccaaag tggttgcacc attctacatc cccaccagta tgtaagggtt ccagtatgag   3480
gattgcgatt cctccacatc ctcgccaatg cttgttgtta tctgactttt tggttatggc   3540
tttctcagtg ggtgtgaggt gatatatcat tatggttttg atttgcattt tccggatggc   3600
agatggcaca tgatactgag catcttgtcg tgtgcttttt ggccatttgc atatcttcct   3660
tggagaagtg tctactcagg tcctttgctc atttaagtaa tattctcagc atcgattttt   3720
aactgcctgt ctccaaagtc tttgttacta tttgctgggt agctgtccac atcactgttg   3780
ttagataaag ggcttttggg attccgtgaa agaagaaaga gtttcctatc ctttatcttt   3840
tctgagttca ttttctgacc aggtttcccc tctcctcttg tttttctact ttaggaagat   3900
```

tgcaatggca tattcagaat taatgtcagt gtgtcaaaga acttaaattt aaaattgagg 3960

cctggagaga aaaaacctgg attttgggtg ccccgtgttt ggtaagtata tcataaaaaa 4020

atattctgaa gaatggtgtg tttgtcanta taaggcttta aataagtatc atatggcatg 4080

taactaaaag tagttttgct agagatgtta aaatgatctt aagtcataca gggcgaagcg 4140

ctagattata ttggaagata gaaataaggt aaaaaataag agagcaagta gtgcaatgag 4200

ttctgtatag cttcatgaaa ctagagaaat attatctcag tatttttcta atacttcaga 4260

aatatgtatc tggattcaaa tgccttaaaa aattaggctt aacaggattt tctgtcaaat 4320

taaaactaaa cttatcctga gtttatatct ttaaaggagg atctgatgtc ctcataatag 4380

aaccctaaat cataaagcta tttccatttt aaaacccatt tacactgaag gacttccatt 4440

attgaaacac tagagaggaa taatcagtga cacagttgat tgcttctaac tttcccttta 4500

atattccttt ttctttaaat tgagttttca aagcacttag atcaaagcta gaaataccta 4560

gtaccttctt tcttatttta ttaatgacca tgagtgaatc tctcagtctg tgtcattggt 4620

gaactgagtg ttctgtcacc ttcatttgtc ctttaaaatt nttatctgtt ttttaagttt 4680

tcttttagaa aatgtgtaag aataacaact gagtcacagc acctctgctt ttgtcacaat 4740

tgatgttctt ttctaaatgt caaaactctt taaaaggaac agcatttcac ccagctgtac 4800

cctagcactc actgttacct ctctttggca tttttctcat atgaaattcg cttttatacg 4860

taatccaata gaatgcatat tgggttttta tttttaaaa tatccaagaa ctgagaaaaa 4920

ggattcattt ggggtatgtg gcttcagctt ggagagagag gaatgtgaga catgagaata 4980

tcctggcccc actgctgact agaaaagccg tccaatttcc attgaaataa aaattaaagg 5040

agcataaaaa ttttcttcca tgagataaaa attaattata gcataaagtt atttatgtta 5100

taatttatta atttttatta attcaaagat aatatcatgg aagtaaatgt agtcaggaat 5160

agagaacact ggggcatttc atttccttca ctatagagaa aggtagaggg tgcctcccac 5220

ctctgccaag tgaggtaccc caggggggatc acttaacgtc agagatgggc agggcaagaa 5280

gagagactct tgtggccagg ctgctcatca ccatctggag agcagggtta tctggtgttc 5340

ctggagagcc cagtagtcac tgcctagggc cacatcatct gagtcctctt gagccttgcg 5400

gcaaagggct ttgtgctggg aaagggtaac cccctccaaa gggtccgtag gtgtgtttcc 5460

aggagccaga cttgagaatc acccagagac tgcacactta caggagaaaa ggtgcaatga 5520

ggcactcctg ggctggagag gggggttcaaa ctcttgaaca cctatggcca gaatccaatt 5580

gtatcagatg cacaagactt aaactccatt cctccattcc tttcaacgac atcccacctc 5640

agtcatgtga gaatctcgtg gtggcctgaa gaggatgagt ggaagtgaag aaagtggcca 5700

acttttcttc ctttcctcaa accggggtcc ttgagctgct aaggatggga ttgttatatg 5760

agttatagga tcagttggat gtaaaactga attactgaaa ataaccaaaa agctctggga 5820

cctccccaag aattcagaca gggagctgga acacacttga aggtagtgtt tggagaaaag 5880

taaacctgcc agttgtgccc tatagcactc aacctagtcc ataaaaccag ctacatggga 5940

```
tttttaaaat caaagttaat ttaaatgcaa tattttatgg aaggtattat taggggtttc   6000
aaaggcaatg ccattctaga agcataagca aatagccagg cattccttga acatcatcac   6060
aagtgctctt gtccatttct ccaggataat gtgacttgga ttgtagtgaa atagacacct   6120
tcgatctaat aaatgtcatc tgtaaccatg ataacttttt acactttcct tcttaagtaa   6180
gctaatcaag agttggaagc tgaaattttg taattttata ttggctaaat tatgattaag   6240
aaaataacat atctaggaat acactagagg aatgctttgg taattataat taggacataa   6300
atatgcacta attcaggaga cagtctagtt tcaatattct ctcttcagat taataaatgt   6360
actttcttct gaatcactca gacttgatga gtcacctcag taaaaaggat taaacatatc   6420
atggattata aatacagtgt gatttatttt aagaggaggt ctggaatata ggaaaatggc   6480
atgcatggga gaaaaactaa caaagtattc aggcttataa taaatgtatt tattgagtgc   6540
ttgctcttca ccagctagta tgcttgtcac tgtggtaagg ccctgaataa gcgatggccc   6600
ctacccctta ggagcttaaa gtctagtggg aaacagaaaa taaaattaca aataatttna   6660
tataatatgc taagggcttt aatgctgcac agagaagccc ctaagcttac ctctggctca   6720
cccgaaactc atctccatgt aatgttaagc tacactgtga tatgttctag gaagaaggtg   6780
atcaggtttt atagggtcna aggcccaaac aatttggggt accaacatct cagaagggcc   6840
tgtacaaaag aggggtccta ccccttaggt ttctttggct tcatcatata cctgcctctg   6900
attcctacag aaaggctaag ggtacaatgg acacacagga gaagcaccta accccaacta   6960
tgcagggcca aagagaagga ccagatgaga agacacactg gttacctttt gaaggacaac   7020
caataggaac tagccaggtg aagggggcag ggtcagaggc aggtgaggag gaaagtctag   7080
gcataggttt ggtattacaa gttaaggcaa ctctatgccc aaagctcaag gcaaaaagcg   7140
ttggggtgct gtaggaagta cgacgtttct gatttctttt aggagagatg gttgtgtaag   7200
atacagcatg ccgtttttgg acaaagtctt tttctgcatt tccacatcca agggcagacg   7260
ctcagtcacg tcttcagggt gcccccatgt tccccaaggg aagagcagct ctccctacct   7320
tgatcttggg gggccttgga gcacctcctc tgggaggtaa aaaactgaca aaattatgtg   7380
gtcagatcta atgtcttagt tttctgaatt ggacaggcaa gattacaata tttcacaaat   7440
taaccagaaa aaatgacaac cttttcaatt caaggaaggg tagcaggtaa aaatgacaat   7500
aatagatgnc atttgggatt tattgaagga ctcngtagtt acaggtaaag tttcaaactt   7560
tgcaactttg ctagcatgaa gaaagcataa gaaaacactt tacatgatta aaactgaaaa   7620
agacttattt ttttaatgtg ataaattctt ccaactattg aagtctttga tgctacattt   7680
tgttttactt ggtttagtta gaatattttt cacatattct aacttctaga agatctagaa   7740
cataacagta gtttcagtgg ctgcaaattc agctaagaaa gcacttctac aaaagaaaat   7800
gagcaggtaa ggaagtccaa caccggctct gagaaccggt tcacttgctt gtaaaaaatc   7860
cttattcaag caaactgtgg catctgtaga atccttataa cttctgtcta ctgtaaaata   7920
tatgaaattt agtggaaaca tcccaaagga aagattgttg gaagcatcca aaagaatata   7980
```

tttgtacttt catacataaa aaaatataaa gcagaattcc aaaatatgca tttctgagtg 8040

gggcttttgc actatattaa agaccgtctc cactaatcag ccctaagccc ttaagaacaa 8100

atcagtttca gtttaatgtt ttcttcttat atcctcatct atttaatatt caaaaaactt 8160

atgaaggcac atcttgtgat tgtctctata gtttaaaaaa cataaaagat ggtagaaaag 8220

tacagaacca gattttccct agagcgttgc ctgtcctgag ggttaaattt tgagaaatga 8280

gctagcagnn gggnaaaaaa tgatcaaaat ataatttaaa taggctccaa ggggcctact 8340

ctagtcagtg gtccggagct cagttcttca ttaaattcaa gaaatatagt ccatgatttt 8400

ccagtgtttg gttatcaaaa tcttaattat ttaaaattag cctctaataa gaaggaaaga 8460

gggggaaaag ccactgaaag cttcttctat taacaacaaa aatgttagct aagaaccata 8520

acaattaaag aagatgcaat ctagtaaaaa ggatcacttt tataattggg aagagaaaag 8580

tataaccatg ttaaaagaac aagaaatgag acatccctgg cccccacaga agtgggagct 8640

aaacagtggg tacacctcga cataaagttg gaaactctgg acagcgggga cttcaaatgg 8700

gaggaggaag gaaggaggtc ataggttgga aaaattacct gttggattct gcgtgcacta 8760

tctgagtaat gggttcatta gaagctcaaa ccccagcatt atgcaataca cccatgtaac 8820

aagcctgcac gtgtaccccc caaatctata agaaaacaga aaatgaataa gccaatgagt 8880

ttcaggaatc aggaatcagc ggtggaggta gaaagcatac cacccataaa ggggaccaga 8940

ctaagagtcc actgagaaga ctcagagtta ataagaaaaa atccaactgg aagagagtaa 9000

tttggaaaaa tacatactac aagcaaacac ctcttgccgt ctcagtgacc tctaaatgca 9060

tcactacttt ccagccctgt gttcaccatg atgacacaaa gtgaaataaa gtgactaaat 9120

tctactctac actttctata ctaacacagg caattgtaat ttatgccttt cgaaagattt 9180

tcagttacaa aactatttca aaatccacat ccaaaatcca atgcttcttt gtgcaaatcg 9240

ataaaatcat ctgtccagtc ctcagggtaa tcaggacaag gtagaaatgt tgcagaacct 9300

ctcaagctag aactagagct gttttacatt gagatcaccc catcattacc taatgaacac 9360

aatcaattat aacacattga tgcattgtta tgatttagca aacgaatgcc atatagcgta 9420

aaaatgtgtg tgtatacata tacacacata tgtatgtgtg tgtgcacgca cacgtgtgtg 9480

tgtatatatg tgtgtgtata tacacacaca cacacatatc ccacactcat gactggaaag 9540

ggctacagat atttgccttc gggtgctgct tattggaatt tcacagaaca ttcctgtatt 9600

ggaatcgcag ggcgaatcct cggaacttta atttcgacaa tgtgggaaac gctatgctgg 9660

cgttgtttga agttctctcc ttgaaaggct gggtggaagt gagagatgtt attattcatc 9720

gtgtggggcc ggtaagcgag cacatcgaat cctttgaaca cttcaagagc agctttcatt 9780

tgttttgatg aataaccata ttttccttta tgcagatcca tggaatctat attcatgttt 9840

ttgtattcct gggttgcatg attggactga cccttttgt tggagtagtt attgctaatt 9900

tcaatgaaaa caaggtaagg atttttggtt tggataccac agggtttagc cttttctagc 9960

cactcacttt gagctatttg tttcttttca cacgcggcgt gaagccgctt taaactgtaa 10020 tgtgactggt gtaaataatc tctataactg ttttacaagc tccctcattg ctgcttttca 10080

gacacattat ttaagagtac agttctcttg ctaaatttac ccttttggta acagtgtcat 10140

taattcaccg acaactgaga ggcgcgtctc ctattgttca aattgaaaag aaagatgggc 10200

agcaagttat taaaatacaa tagatgaaag tgattcgcta attacaggta atattacttt 10260

cactaaatgt aaattgtatg cttctgtgta actcagtagg tgaactcagg attttgatgg 10320

gggtgagggg agtctggtag aaatttcaaa atcctaattc acagaaaaaa agattttgtg 10380

gcagggcatg gtggctcacg cctgtaatcc cagcactttg ggaggccaag gcgggtggat 10440

cacttgaggt caggagttca agaccagcct ggccaacata gtgaaaccct gtctctacta 10500

aaaatacaaa aattagccgg gcgtggtggc gcatgcctgt aatcccagct actgggttgg 10560

ctgaggcatg agaatcgttt gaacctagaa ggcagaggtt gcagtgagcc gagatcacac 10620

cactgcatcc agcctgagca acagagcgag actcaaaaaa aanaaagaaa aaaaagaaaa 10680

aaagattttc taatttcata tgaacttgac atactttgtg aactcgggaa gcctgttgag 10740

tttgtaacag ctctaatgat tagaacaaag gctcacaaat gggaaaatna caatgtagtt 10800

tcctatcaaa gtgtgagttt ggggaagaac attctattag cactccttgg agaggggctt 10860

gctcatctag gcacttgagc tatgccctct attgtttgag acgcagtctt tggaatcata 10920

gtaagggaag tgcatgcagt ttttgagcac acactgtgta ctaagatttt acatattcaa 10980

tcaattctca caactctcaa atgcacggtt gcagaaaaag caatgactag tctggggctt 11040

ataaggtgtg anaaacaggg ccagaattca gaattggtgc ctattaccgc aggctctgtg 11100

attcttcgaa ggcatgagta gaataaagtc aaccttcact ctgcattaag ttgaagctag 11160

atctgcctgt aaagcaccgt gctgctcgat gttattcatg gctaggactg attcccatgc 11220

gttcacatca gagctcttca gtttgtgttg ctacacatgg gactgtatct gtcattcact 11280

gtcctcaggc tcaacgtgag tggtaccagt tagccagtca attgacacct ttgatgtgtg 11340

gctaatgagc atgtcatttc attcaggctt tcaaagggtg agatttaaaa attcctcttt 11400

tgggccccga ctctacttgg accctcagaa acatgggatg ttagagacag aaagaacagt 11460

agtgatcatg tcacccagcc ccttatttta ctcctatcac attcattaac gagatgcttt 11520

gagccctaca gtgttgagtt taataccaag cagaaagagg aagaaggatc tttttagatg 11580

cctttgagcc atgtaatgca gcagacactg agaattctag aatataagca gaagctaatc 11640

agtcaattaa caggagagtt taggcactgc ttgcccagaa caccgttaag aagccgtctg 11700

agcagacacc tgatttcagg aagattagaa tgtatcacag ctcacaattc agcaacatac 11760

atagaacaac ataaaatgtt cagtaactgg aaataaatgt ataattaatt gcaacacttg 11820

agcttagaga agtggctgag gttagcagaa taaccaggga ggtctccatc cagcactgta 11880

tttatgttct gcaggtagga aagagtaacg tggcctgttt gggggatctg gggcgtatgt 11940

tctatcagct caaatgggca cttaaagaaa ggagaaattg acaagaaatt atattaccaa 12000

gctttacaaa cttgatttgt taattggaga tatcaattgc taaattatgt gagcctttta 12060 gaaaagtttc tttctctgca agnttttttt ttttttttag aaagcagctc atgagctgcc 12120

atctaacagc atttatacat tattcaatca ttcaacagat gtatttgagc atctactagg 12180

tgccagaggc tgctcttgac actgaggacc tagtgaacta acctaaacac taaacacaga 12240

aaagttattt gccctcatgg agcacacatc ttaatagtca tacccaaagg cacataggta 12300

aatttccaaa tgtagagaca acagagatag aaaatgaggg aatctttagg gtacaattca 12360

gaatcatgca aaaatatgtc ttcttttgaa aatggaaacc tctttaaaga gttccaagtc 12420

tatattagta atatagattt ttttctgatt tctctactat gtaagacaat gtgtttgatt 12480

tattaaccta acttgataat aataaacttt cagaggtttt tatatattaa tgatcagatg 12540

atatctccat gcctattatt cttttagatc ctttcagagt tcaactttcc aaaatttttt 12600

ttatgtataa caaagtaatt acattagtgt ctatggtggc cagagaacaa gtgtttgctt 12660

ataattagtg agaaaagtgg tttagcaatg tgattatcag aaaagaacaa ttttaaacaa 12720

acactattat tgggaataaa tgatgtacta aagaatgaca aaataatcac cagaaagata 12780

tagaacttcc aaatcccaaa atgtacaaag caaaaattaa acaagttgta aaaagaagct 12840

agaaaatcca aaatcatagt gaaagatatt aacatatatc tcaaaaaatc tcataattca 12900

gagnagaata aaaggaaata tgagatttga tacaataaac acaatcaagt gaatgtataa 12960

ggaaaactct aattacaaca atttgagaat gtgcattatt ttcaagcaca aaatgtntct 13020

cccagaaggc cataattacc caagaattga taatgtaaga atgttttctg accataatga 13080

aaacaaaaac aaaaaaaact ctacaatttg gaaatattaa gacaatacta taaataattt 13140

atgagttgaa gaaattataa tagaaaatca caaaatattt acagacaaaa cattttatag 13200

cgtcatatac aaaaactaat gaggtatagg taaatcagta attacaaata aatgtatcat 13260

tttaagtgta tatattaaag aagaatcaca agttaaatat taatatgtca aatatccaaa 13320

tcaataaatt ataataaagt aaaatcaaat aaagtagggg ggaagaaagt aacaaagatt 13380

aggacacaaa ttaatgatac agaaagtaag gaaacttgta tagcttctat aatattcaag 13440

aataacaaaa gctgggtttt ggtttttttt aaatcaagtc aataaacctc taggaaggtt 13500

gaacaaaaat aaaaacatga aacaagaaag taagctacaa tcattaaaag agtgtcagta 13560

ttgttatagg tatagaaaat tagacttaca gaaaaaacta gaaagaccca gcagggaccc 13620

atgtatgtag aaaacagaag cagctcaaca aattagtgca gagggtggtg agaccaacaa 13680

ggagtggtgt tgggagtatt gattgtattg atgtaaagag ctatgctgag atctctacct 13740

cactgcatac tcaaaaatga attccagatg aattaaagac ttcaatatga aaagcaaaat 13800

ttaaaccatt tttaaagcat gtcagcattt attttttatt tgaatttcag ataagcagca 13860

aataatgttt ttgcctttat tatttttaat cagcgcataa ttgtacatat ttatcnggta 13920

cagtgtgata tttcaattca tgtattataa tgtgtaatga tcaaatcagg ataattagca 13980

tatctatcac ctcaaacatt tatcactgtt ttgtgttggg aacattcata atcctctctt 14040

ctaggtattt aaaaataaac aattaattgt tgtcaattag attcaccata tattaccatg 14100 gaacactaaa atttannttc ctccattcta gctgtatttt catatccatt aatcaacatt 14160

tgactatccc tattctagtt cccctcccca cccctagtaa gcatttgttt tttgccctgg 14220

agtagttatg gatttcatcc cattagccac caagtgaaat ctataaagga aaaaaattat 14280

aaactcgatt atatttaaat taaaattttc tgaatgacaa aagacactct aagcaaaata 14340

aaaatacaag ccaagactaa gagaaaatat gtacaactta cataattgat aaaagattag 14400

tatccagcat gtataaagaa tctctacaaa ccattaaaaa gtaaatatca aaatacaaaa 14460

atgagcaaag gatatgaatg ggcaattcac ataggaagat acaagcatca ccatttacac 14520

acatgcaaat gtttcatctc gttaatatta aaacaacatt agatacattt tcacacctat 14580

caaaatggtc agattacaat gtctggcaat atcaagtatt gcaagaatat taggatatgg 14640

taacacctgt gtattgctgg taggaaattt gttttggttn ttgctctttg gaaagcaagt 14700

tggcaacatc tggtaaatgt aaagatgctt aaattcctag accccatcaa ttcttctcct 14760

agagatcaac aatggagaaa ccctcataga tatttactgt gagacagtac atggatgctt 14820

attgcagcat tggttatagt ccataaatca aaaatgaata acttcggata ttcgtacagt 14880

ggatggtgtt tgtgagcagt taaacaaatt agatttagta atctagttgt atggaaagta 14940

tactgttcaa tgcagatata aagaaacagc nnnnataatc acatgttgga gtccatttct 15000

ataaatatta aaaaattcag aacaaattct tacgtattga ttcatgcgta agcacaaaag 15060

tctgaaaaca tagtaaggag gatgcacgcc tgaagggtag tggttagctc tgggtaagga 15120

aagggagagt ggggtttttg tctagttttt aaaattacac ttcagctttt atgataaaat 15180

ataatgcagt gcagtgcaat acaatataat ataattttaa tcatatgcat tttaatcaag 15240

tactctatta tacagaatat ggaattccat ttaggcagtt atcacagcac tactttctta 15300

aaatgcaaac tctatgactt tgtatctctc aaaaacagtt ttgcaattgc acatattaga 15360

cattttttaa agaaagacaa catacccttta actggaagcc acagcattga caagctataa 15420

ttttgctgct atgaaatgaa aggaaactaa ctactattag tgagtatcct tttttgaaaa 15480

tgtaattatg tagaaactcc tctttacatt ttcccaaacc tgaatattac acatctccaa 15540

agatacaagt cgaaggctgc tgaattctct ctgaaacacc tggtcatcat tttgggaact 15600

cctcccaaga ccaactttga gcattctgac catgtttctc tcaccaccac ggttggaaga 15660

acatggtcag tgcaagccct cacatcagtc ctgtttgtct cggcacccca tctgccatgt 15720

tcttggtgca agcacatgac aaaaggcctt ttcctgcttt cttttgtctg ttaaggggac 15780

ggctttgctg accgtcgatc agagaagatg ggaagacctg aagagccgac tgaagatcgc 15840

acagcctctt catcttccgc ctcgcccggg tacccaaaat gctctgattt tattcatgtt 15900

gggcactagt gtgcacgatt ccccaggagt ctcagattta aattaaaata ataatcgtga 15960

ggccgctggg aggttgtgca taaggtaata cagaaacagg ctccaaacca cagatctctc 16020

cataaggcta ttaaaatcaa ccacccaaag gctgtgcttc agtgctctca gactaagcct 16080

taatgaaaaa atgaattttt aaataaaata tatttgctct tgccttttaa aaatataagt 16140 gaaatgtatt aaactgacag aaattgaatg atgcacttct ccttccaata tagacggcat 16200

catgcactgt cgttttgaat attcagagcc agccattgct gcaaaaacac ctcacaatag 16260

acggaactgt ttggggagaa aaatgtaatg ggttttaaga taagcattta agaaaataaa 16320

cgacaagaaa gtaaatgtga cctgtgtcta aaatgccctt gcccaaacct ttcttttcta 16380

gataatgatg gttttagagc taaaatgtat gacataaccc agcatccatt ttttaagagg 16440

acaatcgcat tactcgtcct ggcccagtcg gtgttgctct ctgtcaaggt acttcgtttc 16500

accctaagac tccgggaatg aattttgtat tcactgctac atccgagagt atcacaagct 16560

gctatatgca ttcacggtta tgttcttggg tgtaaaaatt aagattcact gtcaatggca 16620

aagtgggcag aagaggggca agcctgtatg tgtccgatgg acgtgggcaa cgactcgtgt 16680

gctgtttctg ttgcagtggg acgtcgagga cccggtgacc gtacctttgg caacaatgtc 16740

agttgttttc accttcatct ttgttctgga ggtactgtaa tgagtcagct agctctgggg 16800

gaatctgtgc acagttttaa agtaaatcag cctccctctt tgatgccact tactctcctc 16860

tctgtgttcc attgcctttg aagtttgccc ttaggatatt tcctacttaa gcggttctgc 16920

ttaactcgaa caaagacatc actttatcca tcagattttt ctaactacag aaatgctgct 16980

gtcgattgtt gtgtctgaac attggcctgt ctagttatgc gagttctgat atttagaggg 17040

aatggtgcaa gccactataa ttcattgact caaattgcaa tagagacgca aancaagttg 17100

gcccagacaa ggttatttta acatgtgatt ctcataatat tgctttccag ggcagtcttc 17160

ctgaaataca agcagattct tattttgccc tacactggtt gatgtttccc aaaatctatg 17220

gccttttgct taaggaagaa aaatgtccca tgatcagaac tagtcataaa acgaagagga 17280

aagagtgtag agaacacaga tttgtaagct acttagcttc aaaaataaaa tcatatatct 17340

tggtaataaa ctcttatttt tagttactac atgtttacaa ttattcttta aatagatagt 17400

ataaccgagc atgctggctc actcctgtaa tcccagaact ttgggaagcc gacgtgggtg 17460

gatcacctga ggccaggagt ttgagaccag cctggccaac gtggtgaaac cccatctcta 17520

ctaaaaatac aaaaattagc caggagtggt tgcacgtgcc tataatctca gctactcggg 17580

aggctgaggc aggagaattg cttgaacccg ggaggcagag gctacagtga gccaagatca 17640

tgccactgca ctccagcctg ggtgacagag cgagactcag tctcaaaaaa taataataaa 17700

taaatagaca gacagacaga tagatagtgt aggcacgtgg cctaaaagga aacctcaaag 17760

ggcattaaag tgtttgtctc attttcaatg gaaaagatac tgtgtccaca tcaaataatt 17820

taaaattttc ttcatttgaa aaaaatacat acacataatg catctcaagt taatacattt 17880

aaatatgtac taaacactct ctctgtttat tttcttcttt gctttttctc caggttacca 17940

tgaagatcat agcaatgtcg cctgctggct tctggcaaag cagaagaaac cgatacgatc 18000

tcctggtgac gtcgcttggc gttgtatggg tggtgcttca ctttgccctc ctggtaagca 18060

tagaagtcaa gttgatttca gctgattatt tcagtttggt tcagttctgc tcagtcctgt 18120

ttccagtctc attttactcc acatcaacct gctctgaggt gttagatgga cctccctaaa 18180 gcccagcttt gactgtgatg ggctctcaga catgcttgtg tctacatgta cacacccccca 18240

cactcttcac acaggcaaaa cctggcccct caaagaccta gacgatgcaa gctagaaaga 18300

gaaggacctt caagatcttc tggtttagcc cctcatttta ccaaagggga aaacaagaaa 18360

gagagaggaa gactgcaaca gagactgaga actgtccaaa ttcacacaag ctagtatcag 18420

aagactatag tcctgttttc ctgttttnct cagctgaaat atagcatttg ttccactttg 18480

ttataatcag tatttcacaa tgtgattctc aagtgcttcc actgtttcat tttctataag 18540

acatttagcc tccacttctc taaagcatat gattctccat ctctatatct gtgttctcac 18600

tattttacct aaagtgcctt caccctccat tcagtctgag atcccactca aaatctgcag 18660

ctatgggtnt actgaactgt aatctagggc catattgtcc aatgaaaata aaatgtggcc 18720

acagatatca cttaacattt tctagaagcc acaataaaaa agtagtngtt tagaaaatat 18780

aactattaaa ttgttaataa ttattaannn nataaataat aattaattat ttaattangt 18840

aatatatttt attatttaaa atatgtatta attaaaagat gcattatgtg tatgtgtatg 18900

tatatatttt tattatttat atagttatat attatatatt aatatattta attaaataat 18960

taatttaatt nattatataa ttaanttaat aattatttaa ttatataatt atttggttat 19020

aattatttat gtataattat tatttatata cttataacta ttatgtatgt aattatactt 19080

atgtaattat aattatttat ataatttaat tattaacaat ttaataatta tattttctaa 19140

actatatcca aagtattatt attttcatgt gtaagcaata taaaattatt gagataagta 19200

gtatccttgt atctgtgtgt gaagtctttg gctgagtatg tatcttacat ttacagcata 19260

tcttcattta aattgccaca ttttaagtgc tcagtaacca catgtggttc catgtggctg 19320

ctggttacca tgttggatag cacagctctg gaagatatgg gtcctagtct aggttccacc 19380

attttccagc cttgactaat ttatttgcct tcaatgatcc tttgtttcct gctctgaatg 19440

gccaatataa cttcataatc cacctattcc ctctggcatt atgaagatca agtgagataa 19500

tatatttgag gctttcttga actagaaaga ttatacaagc ctgtaggcat tattataatt 19560

accttctttt tgttagtcat acttgatctc atcccaagga atattatcgg tttttccctg 19620

tttccttcta tccattctcc catcccctta gtgttagaga catgctggga aagactcatt 19680

tacacatcct aaagggcttg ccttaatacg agacgctgtg ctcatgaaca tagtgagcgt 19740

ccttcttaaa ccagtggtta ccttcagcat caggggctgt atgaagccta ggtgagctct 19800

cagtcatcct gaagtatttt ctcagcaata aatggttctc aattcctagt acaagacaca 19860

gtactagtaa ctgagaagaa tgtaataagc aaaacaaata tagttgctgc tcttgtgaac 19920

cttatatcct agaaaagtag atagatagtc atcaattgct cacaaatcta tgagtataat 19980

gactatgatg agggagatga aggagagcct gtgttcctgt gacaacatag ataatagtgg 20040

gagatggctt catcagagag gtcttggaag acatccccaa gggagtgatg gctccatctg 20100

taggataagg ggagggcatt tcaggtagtt gtccctgtga tggaaagatg acagagctgg 20160

aagaaggata gatacatagt actagcatag tacaagattt ggatgggtaa gaaggtatca 20220 gacctcttag gaccttgcag tccaggtgaa gaattcagat tctcatccta aaaactattg 20280

gttgtcatta cttccaaaac ataaagaggg aaggtatctg caatagatct ttagctaaca 20340

tgggaagaaa aggcagcaag gcccatgttt ttcctttctc cagataaaat ttgtattagt 20400

cccttctcac actgttatga agaaataccc aaaactgggt aatttataaa ggaaagatat 20460

ttaattgact caaagttcca catggctggg aaggccccag gcaacttaca atcatggtgg 20520

aaagggaagc aaacacatcc ttcttcacaa gggaacagga gagagaagaa tgagtgccca 20580

gcaaaggggg aagtccctta taaaaccatc agatcttgtg agaactaact caatatcatg 20640

agaacaggat gggggaaact gcctccctta attcaattat ctccacctgg tccctcccac 20700

aacatgtgga gatnatggga actacaattc aagatgagat ttgggtgggg acacagttaa 20760

accatatcat tctgcctgga cccctcccaa atctcatgtc ttcacaattc aaaatacaat 20820

cataccctcc caacagtccc cgaaagtctt aattcattcc agcattaact taaaagtcca 20880

agcccaaagt cttatctgag aanaggaaag tttctgctgc ctatgagcct gtaaaatcaa 20940

aagcaagtta gttacttcct agatacaatg gggtacagg cattgggtaa atatacccat 21000

tccaaatggg agaaattggc caaaacaaag gggctatagg caccatgcaa gtccaaaatc 21060

caatagggca gtcattaaac attaaagttc caaaataatc tcctttgact ctgtcttata 21120

tatatccagg gcacactgat gcaagaggtg ggctcccatg gccttgggca gctccaccat 21180

tgtggctttg cagggtacag accccctcct ggactggtat tgagtgtctg tggcttttcc 21240

aggcacacag tgcaagctgt cagtggatct accattctgg agtctggagg atggtggccc 21300

tcttctcaca gctccactag gcagtgctcc agtggggact ctgtgtgtgg actcnnnnnn 21360

nnnnnnncaa ccacacattt tccttccgca ctgccccagc agaggttctc catgagggct 21420

ctacccccac agcaaacttc tgcctggaca tccagttgtt tctatacatc ctgtgaaatc 21480

taggcagagg ttcccaaacc tcagttcttg acttctgttc acccacaggc ccaaaaccat 21540

gtgtaagcca ccaagtcgtg gggcttgcac tctttgaagc atcagcctga gctgtatatt 21600

ggcacctttt agccatggct ggagcagaag ctgctggaag gcagggcacc atgtcccaaa 21660

gctgcataga gcagagggag cctggggccc acccacaaaa ccattgttcc cttctaggct 21720

tccaggcctg tgatggaagg ggctgctgtg aaggtctctg gcatgccctt ggggacattt 21780

tccccattgt cctggtgatt agcattcagt ccctcgttac ttatgtaaat ttctgcaact 21840

ggcttcaatt tctccctagt aaataaggtt ttcttttcta taaaaatcat cagactgcac 21900

attttccaaa tatttatact ctgcttcctc ttgaacactg tgtcacttag aaatttattc 21960

taccagatac cctaaatcat ctctcttaaa ttcaaagttc cacagatctc tagggcaggg 22020

gcaaaatgct gccagtctct ttgctaaagc ataacaagag tcacctttgc cctagttccc 22080

aataagttcc tcatctccac ctgagaccac ctcagcctgg acttcactgt ccatatcact 22140

atcagtattt tggtcaaagc tatccaacaa gtaggaagtt ccaaactttc ccatatcttc 22200

ctgtcttctt ctgagtcctc caaattgttc caacctctgc ctgttaccca gttccaaagt 22260 cgcttccaca ttttcaggta tccttatagt agccccccac tctaagtacg aatttactgt 22320

gttagtcctt tttcatgctg ctatgaagaa atacccaaaa ctgggtaatt tataaaggaa 22380

cgaggtttaa ttgactcaca gttctgcatg gctgggaagg cctcagaaaa cttatgatca 22440

tggtggaagg ggaagcaaac acatctttct ccacatggtg gcaggaagaa gtatgagtga 22500

ccagtgaaag gggaagcccc ttataaaact atcagacctg gtgagactta atcatcacaa 22560

gaacaggatg agggaaacca ccccagtgat tcaattatct ccacctggtc cctcccagga 22620

tacatgggga ttatgggaac tacaatccaa aataaggttt gggtgggaac atagccaaac 22680

catatcgaaa tgcaacctaa agctactgag tccaagcaaa agtcaagtag actctctaat 22740

cccagataac ctcagagagg aaatcttaaa cccaaagacc aacttccctt ttctgcctaa 22800

ttgcctttgt tcttctccaa gaacctcttt ggtcttggaa ggatgctcat tctccaccgt 22860

attttgcagt ggactattgg gaaggaaagt tccccatcca nggatgctat gtgcacacat 22920

tcatgtgctt ggacctcacc aaggatctca ctggaccccc ttcgcggggc ccatctgtga 22980

ctgtcatcat ttggctaggc aagcctcctt gagagtaatc acactctggc atcggctgaa 23040

tgtcaaggtt gttattgcag ttatagctgg tgggatcttt gtttctcaaa tgaagctttc 23100

ttattcataa agtgcagtca tctcttcctg ctaatcacac ccttgaaaaa agagctggat 23160

gtgttcctga gcagggagga cctctctgaa atcagcttag aggaatgagt gtatgtttgt 23220

tcttcagtgg tgtctagaat ctgacttctg gcaatctttt ccaccatcag ccctttctct 23280

ttgccaggat tcctttacct cctatgtcag ttgaccttcc tatgttgcac tgttaccctg 23340

tgtttttcac accctgtatg caaaagagct ggtcaccatt ttatgtaaaa taaatttcaa 23400

aatacagaat ttagctgagc aaaacagaac ccgaagtttc atttgagact gaattttaaa 23460

tctaatattt attcatgagt ttttgcagtg accttttcag caccttttttt agatcacttt 23520

cttataataa tttacacata ttaagctaat tattcagtat acttcacaga aggcttaaaa 23580

aatggaactg taacaagatc ttttgtgcaa attatgagaa gctgtcttct gtttaaattt 23640

cagaatgcat atacttacat gatgggcgct tgtgtgattg tatttaggtt tttctccatc 23700

tgtggaaaac atgtaagtac attgttaaga atcttatctt aaaggtcata tatgtgttaa 23760

tgaaagaatg gtgattaatt gattacagcc ttacctcggt tatttaggga tgtaatcaca 23820

aattactaga catcataagt caagattaac aaatattaaa cattgttaaa tgaatcagta 23880

gcttagttat cattcgatgt gtttcctgcc atttatgttg cttattcttt atattcaata 23940

attaagtaaa ttatttaaat ttattaattt aattaattaa agataaccag tcatgaaaac 24000

aaaaaccaat attctttctt aatgcatttt tatttttatt ttcaataaag ttaagtaact 24060

tcaaagaaaa ataatcttta atgaattttt atngtttcca tgattccaga tnnnnttttt 24120

ttttttttt tttgagatgg agtcttgctc tgtcaccagg ctggagtgca nnngtggcag 24180

gatctcggct cactgttttc tccacctccc aggttcaagt gattctcctg cctcagcctc 24240

ccgagtagct gagattacag gcatgtgcca ccacacccag ctaatttttg tatttttagt 24300 agagacaggg tttcaccacg ttagccagga tgagattcca gattttttaa tattatttgt 24360

taatgtccat gggtatattc cacatatatt tttcattttt tacaatgctg tttccttttt 24420

tattacaaac gttatttcat atatacattc catatataca cataaccaat atacttaata 24480

tttgatagtt atttagtatt ataccattgc ttaggtttgt ttttgtgaaa gtctgtttgt 24540

ggaagaaacc tgagtttact tgctgtcagt caaaaatgtg tagcccttat catagagaat 24600

tgaggagcag aagagaaaat ggaaggtgaa aatgcttagt gttactctta ttaatttgtg 24660

tgaagaggct gaaatgtttg aagtaaaaat aaaattcgcc caacgcctgg tgggttttga 24720

agaatatttt gatgtaatac cttatagaaa gcaactttca aatgatacac ctaagagtga 24780

ttttgagcca atctgcttaa agattagttg ctacccaggt cttccatagc ttaatagatt 24840

gaatattctc tctctctctc tctctctnng tctctgtctc tctctctctc tctctctgtc 24900

tccccgcttc ccacccctgg tccccaggta acgctaaaga tgctcctctt gacagtggtc 24960

gtcagcatgt acaagagctt ctttatcata gtaggcatgt ttctcttgct gctgtgttac 25020

gcttttgctg gagttgtttt atttggtact gtgaaatatg gggagaatat taacaggttg 25080

gtatatactt atctttctat tcatttatca agcaaaccct ttggtaataa aattttaaat 25140

aaacaaatgt tttttgtcta ttaaagtcta ggagagggga aagggaagta atcagttgga 25200

aattaaaata aaatagccaa gttttaaatg atatttttcc agttgaaaaa ggttattcaa 25260

ctgaaaaggt ctactttcaa aaccttagaa aaattagaaa taaaagaaaa caaaaagaat 25320

ctataattcc accaacattc agtaatcatc actccnaatt ccagcacttt ttctatctac 25380

atattgccaa aaaatgatat attattttta gatggaacaa ttttaagata actctgggtt 25440

tctttctcag caaaatatgg gaatttagag ccccccaaat ggttgtaata tatgaaaagc 25500

agttttcaac gatttgccta agttttatag atgcatcatt ttttttcaag tcctgataca 25560

acaaagtata taccgatatg tcaagttgta ggaattagat acatggctgc cccacctttt 25620

atagcatcgt ccaagtgaaa gaaggaatca cttagttgcc aaagcctgcc tacttagtca 25680

ttcaagatga atggtaataa attagtggcc atgctagagt tgggtgttca gacggctata 25740

cttggaggca agtataaatg acaaattaat aaaaaggtat aacctttgga aaaggtnact 25800

tagcaaccaa atttgctaca acctccactt gttttggtgt tggcaaacca tgctatttca 25860

tgacaaaccc ttctataant cacattctgt taattttttt ttcttaaatt aggcatgcaa 25920

atttttcttc ggctggaaaa gctattaccg tactgttccg aattgtcaca ggtgaagact 25980

ggaacaagat tatgcatgac tgtatggtaa atatctctca tcatgaacaa ggcttagact 26040

tagaatcagc atgaaactct tncaacacaa tgttggcagg atcatccttg ttagcgttaa 26100

gactttatgt atgaaatata aatcctgaca cttcagtcca taaaattgca ttattataca 26160

ttccaatgtg tgaatattta gggcaaaata ctcaattaac ccaacttgag ctattgtagt 26220

tccattcaac aaantatgca ttttgtgagt gttgttcttt ccctcccatc tttaatattt 26280

tggtaatgaa agtcaaacag gaagcccttt aatttaatgt gttcagtgga tagttcactg 26340 cccctcagct atatgctagg ccctaggtct tcaacaactc caaggcacac tcaaggagct 26400

tacctcaagg gctctgagtt actcttcctt tatactataa agtatacttt atacttcctt 26460

cttactatat cctgtcttcc ccaagcatca tccttctggt ctcagtttaa acaggtcttt 26520

tgccttaagc taagacagcc agctacatgt tcttatgctg ctctatagtt ctcctttgta 26580

ataaggagct gtagtgaatt actggcttac attcatttat tttatctctt tctcaccgta 26640

ggacagacca atgctgtgtt aaaaaatgtc tgcagtgatg gaaatctgca ctgtccaatt 26700

cgggagccac tagccatatg tggttatgga gcatttaaaa tgtggctaac acaactgagg 26760

aactgaattt tagttttaat tttatattat tacattttaa ttaacataaa tttaaattta 26820

catngctgca tggagctagt ggtcaccata ttggactgca cagcactcaa ctgtcagtag 26880

gaagaagaga tagggccgcc tatgtctttc ttaccagtct ctccccaaac ttggcacaaa 26940

cctagcacac agtaggagtg ctgggacttc atgcatgaat taacaaggag aatttgccaa 27000

atcaaataga atacaaggca aaatttaatg aatcttaaat aaaaatacaa gcagaatgct 27060

gggattatag aagagagagc tagtttatct ggctgagaaa aaccagaaag ttctctcaga 27120

ggtgtcggcn ttttatagag cctgtaaaca atgagagagt ttttgatgac ttaagtcttg 27180

tgctattcca tagcatccca taacagtgtt ccaagaatgg ggcttactcc tcacacaatt 27240

tgaatgtgaa tttnctatgt gcaggcacat cttgttttac tgtgcttccc ttattgcctt 27300

ttgcagacat tgcattttta caagttgaag gtttatggca atcttgtgtg gagcaggtct 27360

atcagggcca tttttccaat agcatgtttg ctcactttct ggttctgtat cacattttgg 27420

taattctcac tatttcngac tttttcatta ttatctgttt tggtgatctg tgatcagtga 27480

tcttttgatg tttcttttgt agttttttttt gaggggtgcc atgaaccatg ctcaataaga 27540

tggagaattt aattgataaa tgtcatatgt cttctgattg ctgcaccgaa catctgtttc 27600

tccaactctt ctctcttttc ttgggcctcc ctatgtcctg agacacaaca atattgacat 27660

taggccgatt cacaacccta caatggccta taagtgttca agtgaataga agattcacac 27720

atcagtcact tcaaatcaaa agctagaaat gattaagctt atggaggaag gcatgtcaaa 27780

agctgagtta gcccaaaagc taggtccctt gcaccaaata gttactcaag ttgtaaattc 27840

aaaggaaaag ttcttgaagg gaattaaaat gctactccag tgaacacatg actgataaga 27900

aatagcctta tggctgatat ggagaaaaaa ttttaccggt ctcagtagaa gatcagacta 27960

gccacaacat gcccttaagc taaagccaaa tgcagaagaa ggcccttact ctcttcaaat 28020

ttatgaaggg tgagagaggt gaggacgctg tagaaaaaaa tgtttgacgc tagcaaaggt 28080

tggttcatta tgtttaagga aagaaagaag ccatctccat aatatgaaag tgtaagtgct 28140

gctgatataa aagctgtggc aagttatcta gaagatctag ctcaaataat taaggtggct 28200

acactaaaca atggattttc aatgtagaca aaacagcctt ctattggaag aagatgccct 28260

ctagaacttt catagctaga aaggagaagt caatgcctgg attcaaagct tcaaataaca 28320

gacttactct gttgttaggg actaatgcag ctggtgactt tcagtggaag ccaatgctca 28380 tttaccattc caaaaatcct aggaccctgc aatggtttga ctgtgtcccc acccaaatct 28440

catcttgaat tgtagttccc ttaatctcca tgtgtcgtgg gagggaccag gcagagataa 28500

ttgaatcatg ggggtggttt cttccatcct gttcttgtga tactgagtga gttctcacta 28560

tagccgtcat gagatctgat ggttttataa gaggtttttc cccccatttt cttggcactt 28620

cttgctgcca ccatgtgaag aagaacatat ttgcttcccc ttccgccgtg attgtaaatc 28680

tcctgaggcc tcccaatccc tgctgaactg tgagtcaatt aaatctcttt cctttataaa 28740

ttacncagtc tcaggtatgt ctttattagc agtgtgagaa nggactaata cagtctctta 28800

agaattatgc taaatctacc ctgcctgtgc tctataaatg gaacaacaaa gcctggatga 28860

cagaatagct gtttgcagca tggtttgctg gatatttaag cccactcttg agacctactg 28920

ctcagaaaaa aaagattgct ttcaaagtag tactgctcat taacaatgca tctacctggt 28980

cacccaagag ctgtatcagg agatgtataa ggagattaac attgttttca tgcctgctaa 29040

cacaacatcc attctgtcca tggatcaaga agtcatttca actttcaagt cttattattt 29100

cagaaataga tttcataagg ctaatagctg ccatagatag tgattgctct gatggatctg 29160

ggcaaagtat attgaaaacc ttctggaaag gattcactat tctacatgcc attaagaaca 29220

ttcatggcca ggcgcagtgg ctcacgcctg taatcctagc actttgggag gtcaaggngg 29280

gtggatcacg aggtcaggag atcgagacca tcctggctaa catggtgaaa ccttgtctgt 29340

actaaagata caaaaaatta gccgggcatg gtggtgggcg cctgtagact cagctactcg 29400

ggaggctgag gcaggagaat ggcgtaagcc caggaggcgg agcttgcagt gagcggagat 29460

tgcaccactg cactncagcc tgggcgacag agtgagactc cgtctcaaaa aaaaaaaaaa 29520

aaaaannnaa cattcatgat tcatggaagg aggtcaaaat atcatcaaca ttaacaggaa 29580

tttggaagaa gttgattcca gccattatgg atgactntga ggggttcaag acttcagtga 29640

aggaggtcgc tgcacatgtg gttggtagaa atagcaagag aactagaatt agaaatggaa 29700

cctgaagacg cgactgaatt gctgcagtct cctgatcaaa cttgaaggaa tgaagattgc 29760

tccttatgga tgatcagagt ggtttcttga aatggaatct actcttgcag aagatgctat 29820

gaacatagtc ataatgacaa caaaggattt ataatgttac ataaaattac tcgataaacc 29880

agcagcaggg tttgagaggg ttgactccaa attcgaaaga agttctattg tgggtaaaat 29940

gctagcaaac agcatcacat gctgcagaaa aatctttcat gaaaggaaga gtcaattgat 30000

atgataaact tccatgttga cttgttttaa gaaatcgcca cagccacccc aaccttcagc 30060

agccaccacc ctgctcagtc agcagctgtc aacattgagg caagaccctc caccaccagc 30120

agaagattac aactcactga aggctcacat gattgttagc ttttcttagc aataaaagta 30180

ctttttatta aggtctgtac ctttttggac ataatgccat tgtgcactta ataagccaca 30240

gtatggtata aacatgagtt gcatgtgcac tgggaaacca aaacatgcat aggactttat 30300

tgcaatattc actttatcac agcagtctgg aactgaaact gcaatgtctc cgaggtatac 30360

ctgtagagtc tatattccca aatcttaacc cattacttca gtcagagaaa aacttgcaaa 30420

```
tgtcttttta tttgattttc tgcagtttga tcacagtggc tacttatcac ttggggctgg  30480
tggtggtggc agaaagaata tctagtggag atcagttcat tttcattgtg cttttaaaaa  30540
tggcaataac cttgggattc tacaagataa cagcatagca taagtgaaat attcttcttt  30600
atttctgcta cagatgatgt acccaatgta tgtggctttt gaaagcttac tttgaagtta  30660
attatatgct atatgtttag gaatctattc attaactata gtcagctaaa agaaatttgt  30720
attctctact cttatcctct ccttcttcca ccccaatgcc ccaaaggttc agcctccgtt  30780
ttgtactcca gatgaattta catactgggc aacagactgt ggaaattatg ctggggcact  30840
tatgtatttc tgttcatttt atgtcatcat tgcctacatc atgctaaatc tgcttgtagg  30900
taagtgatgt tagttgaata ctttttactc tttaaactca ggtacttaga taattatgta  30960
ttattttcaa gcaattaccc tggatttctt ggggtgggtg tttcatggca ttaattataa  31020
attgttgtct ctcaatccag aagctgaaaa gacctttgaa aggcacgtaa atgacagttg  31080
actgcaattt taagagttgt agcatttatt atgatcatta atttcaaaag ggctcttctc  31140
atggattgag tngaacaaat atcagagagg actatctttc tagccttgaa atatcacatt  31200
tatgtttttg aaaggtacac cttatttacc cagatgtggg cccacccaga actcaatttg  31260
acctttaaaa atggccctgg aactctcaac taagagttaa agccagctgg catgactatc  31320
tctatacaca gaatgtgtgt ttgtgcttac gaccctcatg ttttggaaag gcaggagctg  31380
gagaaatagc ttagtccaag tgtaaacatg ccaatttgga aaccatacta aaagcacttt  31440
aaaaatttgg cttatgatat aggtgataga atccattatt aataatagat gttttactct  31500
ttctaaaata tgacagatta gcataatggt ctaacttctg aatttttctt tttttaaaag  31560
ccataattgt ggagaatttc tccttgtttt attccactga ggaggaccag cttttaagtt  31620
acaatgatct tcgccacttt caaatmatat ggaacatggt ggatgataaa agagaggtat  31680
gggagtgggt tgttgtttca tgcctcacct caaagagact tatcaacatg tctcaaacac  31740
ttaaatggtt ttcaggttgg caaatgacac agattgaatt tttaaatggn ntgaanttgc  31800
agaaagccat catttacagt tcagcagaat cctctaggaa gggagacagc atcgttcctc  31860
cttacatgat cttcaaccat gaacatcctc ttcttcagac tttccgaagg ccagcaggga  31920
aaggaagcca aagggactct gtaggaatca gccttgggct caggccagaa aaagaggtga  31980
cactgaaact accaaaggga cagtatatta ncaaattcat gatcacctca aactcaagta  32040
gcaatcccaa ttaaaagtac aaaggagtta tttaagtggc atctgaaggt gtgtattgga  32100
ttatttttga gttactactt gaatgtttta accataaagt ttacccaacg tttcttaaaa  32160
ttagccaaag tagctgcagt ctgctcttcc tcttcttgtc attgtcgtca tcaacctcca  32220
cctccttccc ttcctcctac tcctcctcct ctacccctc tnnnnnnnnn nnntcttccc  32280
cctcctcttc cccctnctct tccccctcct cttccccctc ctcctcgtcc acctcctctt  32340
ccacctcatc ttccccccta cctcctcctc ctctacccct tcctctacct tctcctcccc  32400
cacctcctcc ttctncactt cctcctgttc ttcctccctc tcctcttctc tcttcctcat  32460
```

cctagcaaac actcaggcac agttctaggc actgaggata tagcagtgaa caagatggaa 32520

cagctcctgt actcacagag cttacagtct aataggtgga gacaaataat aaaaagtgaa 32580

ataataataa atagtgcagt atttccagaa gttaaaacta tagacagtag ttggggttgg 32640

ggctgcgccg gcaggaatga ccatgaagga ggcctgaggg gtgctgtggg acctcatgtc 32700

tgaacaacaa ggaaagttct gccagggtat ttggggcgag atctggaaat agttcaagaa 32760

gagttgatga agggcacaaa gaccaagaag ccagcctggg catcataaat gaggggctcc 32820

ggcagtgcac tcaccccaat ttgctccctc ttcttggtac ccctgcccgc ccaggcttca 32880

gtcttgaggg ctttcctttc aaggtagtga gagccctgcg tgcagggcat ctgagctcac 32940

tctcagctcc tcaggatcca cctgtctctc tgtaactgtc acccacggga acctgctttc 33000

atgcaataca cacatgaaga cccactgccc tgcacaccga gatgtcagta actggcaggc 33060

agtcatggtc gccagtgttc cgcccagcac tttcattcaa agaaagatac agccaggggg 33120

catggtggtt tgtgttgcat ggaaggtagc aacacatctt agattctcag tgtctggatc 33180

agttctgacc ccagaagagg gcagagctga aggcctaatt ctcttgtata tagtttttct 33240

aatcctgaag aagacatgaa aaggagttag ttagctcttt tttttttttt ttttttgaga 33300

tggagtccca ctctgtcccc aggctggaat gcagtggtgc aacctctgct tcctgggttc 33360

aagtgattcc ccagcctcag cctcccaaat agctgggact acaggcatgt accactatgt 33420

ccgggtaatt tttttatttt tagtagagac tgggcttcat catgctggcc aggctggtct 33480

cgaactcctg acctcaggtg atccacccgc ctcggcctcc caaagtgctg agattacagg 33540

catgatccaa cacgagttat cttttaatgg tagaaaaaca tgtaataaac attgagacgt 33600

tcctagcatc ccgattgagt gaagaggagg attacacgtg tgtgtcttgt ccattncctt 33660

gtgtgagtac attatggccc actcagtgca tgtgttggag gtggcgtctg tgggacagnc 33720

tatgcaagca tcagcctggt gagagcaaga gtagggtcct ggggggctga aggggagctt 33780

cccatgcctc actccaaatg ctgcaatgtg gtgacatccc agaaagagaa gcaccagcag 33840

aaatttatca ctcagttttc tatcctctgc tcttcttagt gcacttaatt ctaccaggct 33900

ccaaggaggt gaccagcttt cttcacatct ggtgctaatt atgctagctt ttgagaaata 33960

gcaccttttt ttttttcttt tgagaaaccc aagtgaatca tcgaggcctg atgatttgct 34020

acttcttgtg ctctgccccg caggggggtga tccccacgtt ccgcgtcaag ttcctgctgc 34080

ggctactgcg tgggaggctg gaggtggacc tggacaagga caagctcctg tttaagcaca 34140

tgtgctacga aatggagagg ctccacaatg gcggcgacgt caccttccat gatgtcctga 34200

ggtacaaggc aggctttcgc agggctgaat gggggccaga aaccacagct gtcagccttc 34260

aaggcctctg cagtacaatt aacaagtgtg tttccaaccc aaaatgccac atctgctgtg 34320

gtcgtgaagt ggggtcacat gtcccaggcg ctaagtgccc tttcctgccg tcctccagtg 34380

aacatttcca agtaggtgcc tcgcagggga gggctgggcg gctgtgccct gatgtcccgt 34440

agggacccct gcccactgag gtgcaccccg aggggagagg tggaagggat ctgggagagc 34500 acagtcccac cccttccctg tgatgcggga actgcccaga agagctgagt gacttgcttg 34560

cggttacagg gctctggcaa agctgggact aacatgcaga gcacctcact tgcatgggtc 34620

tgatgagaat ttgcctggcc tgctctgtga tggttgcccc gttggctaaa acaccagatt 34680

ttccagacat ttcatgagcg cccagacagg acagagggtc aaatggagtc tctgcctttt 34740

cttcttcctg ggagacttct gtatttgttt ccttgggcta ctgtaactta gcaccttaag 34800

ctgagggatt taaacaacag aaaggtatgg tctcgaagtc ctggatgctg gaagcccgag 34860

atggaggtat tgcggggctg attccttctg aggcctgtga gggagaatct agcctagcct 34920

ctcccctggc tgctggaggt gcgctgccaa tgttggtcat tccttggctt gtagaagcat 34980

cagcttggtc tcccttcatc ttcacatggc attctgtttg tgtgtgtgtg tgtgtgtgtg 35040

tgtgtgtgtg tatctatgtt caaattcctc cttttataag aacatctgtc gtattacatt 35100

aggagtccat tctactccac tgtgatctca ttttaactta attactaatt acatctgcaa 35160

tgttctattt ccaaataacc ttacgttctg aggtactaga ggttaggact ccaacatatc 35220

tttttttggg aggaaacatt taacccataa cagcttctta agcaggagtc catggataaa 35280

acagaggccc atggacttgc aagggggaaa caattacatc tttattttca ctaatcactc 35340

tttgaaatgt gtcattctct tccatgatga acgtaggcaa caaaccacag tcgcattagc 35400

aatatccatg tctttgttac caggagaaat catagatatt ttcagatgac ataactgact 35460

gccttgactg acatcttaaa atatcattta tgctcaccag tgcttctgag ttatggtagt 35520

catttgagct ctgctagatg gtattattta atggaataat aaaggaagca cttacttatg 35580

ttactttatc actgccgttt taaagtaaat ttgagaattg tattttaata taattggttt 35640

cctttgtaat acggtgtatt tcattccgtg catttaaagc attatttcaa gaaatgtccc 35700

ataagctcca ctgccacagg ggtccatgcc acagaaaaaa ataagagcga ctgactagca 35760

gtagctattt taatcacata gcagtaccta caaggagtgt gtgggatatt gtgttgtgtc 35820

tcaaatttca ggagaactca ggtgtacaga aatatgcttg tggccagaag ctgaaaaaaa 35880

aaaaggatca gatcagcaca ggaagaatct gggaaaagtt ctctcaagtt atcaatttca 35940

cctcccacta tccagttctt ccaccaacaa gctaaaaagc agaaaaacat gggacggaga 36000

aatgcatgtg aatcacaaag gaggtcatct ctgtcagaat gtttctgagt ctttcctcct 36060

agtggcaaat ttttttggtt gacatgaggc agataggcat actgttgtca ttaataatta 36120

atattcagcc tctttggaac acaccgatgt atctaataaa aatagaccca ggcaaaatct 36180

gtgatggggc tgtctctccc cttcctattt attcccctag cccagtatat gataaaaaca 36240

aacagtgaag ataaaagcaa tgaagctttg atgggattct tgatgagtca gggaagagag 36300

gaagagaaaa ttcatttttt ttccctctgt ggtgggaaga gatttagcaa agagaagccc 36360

aatgtaccag tgctctcaag agcctccagg ctgtgatttg ggctctagga tggaactgag 36420

ctcacagaag tgggtagctc acaggtgcac cagggtgcat gggtctgagc cataacctgg 36480

aatcctgtgg ctacacctgg gtgcatcaca ctgccaagag ggctgtctt ctcctctggt 36540 ggtggactgc tcagtccagg catcctggga gctgccttat tgccccatga actgtaatgc 36600

tgtgttggct ttacctacaa ctgcatgggc ctgctgtgtg tcaagggtta aatgggagtg 36660

gatattcaac cctgggtttg tttggtagcc tctgaaatta ataaggaaga ccaggcatgg 36720

tggctcacac ctgtaatccc agctctttgg gaggctgagg caggtggatc atttgaggtt 36780

aggagtttaa gaccagtttg accaacatgg tgaaaactcg tctctactaa aaatacaaaa 36840

ctagctgggc atggtggcac acacctgtaa tcccagctac ttgggaggct gaggcagaag 36900

aaatgcttga acccaggagg cggaggttgc agtgagctga cagcgcgcca ttacactcca 36960

gcctggtcaa caagagtgaa gctccatctc aaaaaaagaa aaaattaatt aataaggaag 37020

atacagaagg atgaatgaga gaaacagaga gggaagtggg atgcagggag tagggatgtg 37080

tgacgcacag ccgagcatat taaccatcac ctcggggggc taagtgtgaa gaaagtatgg 37140

gaggtgcggt cttccttctg cagagaaagt cgcgagcctt ttcctagcct ggatcttttg 37200

ctttacaggc acttagatgg cacttactac tataagagcc tcacagatat taactcactt 37260

agtcattaaa taaaggagtt accaggacta tctatcatct ttgttttaca gacgagggaa 37320

gggagacaca gagaggtgac aggaccagcc caagacatgc gggtggcgag gggtggcgtc 37380

aaaacactag acttattaga ctgagtgatg tcctctctgg attttgacgt atctaagggt 37440

gtaaatctca ggtgaatcag agctgcgatt gcaggtgtgt ttctcccacc agcatgcttt 37500

cataccggtc cgtggacatc cggaagagct tgcagctgga ggaactcctg gcgagggagc 37560

agctggagta caccatagag gaggaggtgg ccaagcagac catccgcatg tggctcaaga 37620

agtgcctgaa gcgcatcaga gctgtgagtg aactgcacac atgcaggttg ggtccccgcc 37680

ccgccccgcc ctgccccgca ggaggatgaa agaaaagagc ttcgtgaaca tgggattagc 37740

tccttctgtg atcagccttc tatcccttta aactatgtcc actacctcca cctacagtgg 37800

tcagatcatg atgcagactt tcactctagt ccacagtcag ccctgccctc tttagcagag 37860

tccattagga gtttccatgg aatcccattt ttataatatg acatgaagac ttaaaaagca 37920

gtttattggc caggtgtggt ggctcatgcc tgtaatccca gcactatgag aggtcaaggg 37980

gggagctcaa ctcacttgag gccaggagtt ggagaccagc ctgccaacat ggcaaaaccc 38040

catctctact aaaaatacaa aaattaacca ggcgtggcac acatgcctgt agtcccagct 38100

actcaggagg ctgaggcatg agaatcactt gaacctggga ggtagaagtt gcagtgagct 38160

gagatcatgc cactgcactc cagcctgggt gacacagtga gaatctgtct caaaaaaaaa 38220

aaaaaagttt attgatatta aaatatgatg tccatttatt aaactcaact gttcaaaatt 38280

ctagaatttt ccacatattg aatttgcaaa actactttaa attgctgcta aaaatttcaa 38340

caaactgtga taatgtcatt tattctgaaa agaccacctc tcaagattaa tgatgtttcc 38400

tgcttcctga cttataagcc tataagattc tctaactcat aaattaatga atgtatactg 38460

cattgacata tactgtattt acagtgaatg tttctatgag ataattactt ttatgaatgc 38520

ttttggcatg tttcacagac agggaaaaaa cccaaagcaa tcactaaaca tttatatata 38580 acagagattg tttcaatcct cttcctcaga gcaacaacgt caattgcaga aacatgtatt 38640

tctattggag gagaaaaggt tcccagagaa aagatacaac tcaacctcct atagcaaggc 38700

cagctgccca gtgatgtccc atcactgctg gccactgtag ctatgttagc tccagggatc 38760

aatctgtcct cctctggctc tgagagggca caaggagtcc tggctttagt atatgccgag 38820

acagactggt ctcctggatt gacctgaaaa cttcagaacc aagacaagca caaaggtttt 38880

ctcatccacc caaggctctt tgcctcttac aacttgcagg aaagtgggaa aggggttggg 38940

attgtattgg cttaatgagc ctagaataaa ataagaggag aaaaaataga atagccacat 39000

tagtaatttt aaataagaca taagctgggc acagtgtctc atgcctgtaa tcccagcaca 39060

ttgggaaact gaggcgggag gatcgcttaa gaccagaagt ttgaaactag cttgggcaac 39120

agagtgagac ctcctctctn gnaaaaaaaa aaaaaaaaaa aaaattaatt agccagccac 39180

agtcatgcac acttttagtc tcagcttctc aggaggctga ggcaggagca tagcttgagc 39240

ctaggagttt aaggttacag tgagctataa taatcgcacc actgaactct agtctacggg 39300

acatagtgag atcctatcta aaaaaaaaaa aaaaaaaaaa aancagaaaa caccaacaac 39360

aaaaagaaaa acaaataaat aaaatacaag catcaatgac cgtcactatt ctgattgtta 39420

cacaaatgaa aatggactgg gaagtcaatg catacctctc taagcccacc tttgcatcag 39480

cctaccttaa ttgaacattg aaattagttc aggatcctta cttggaaatt cctaagaggt 39540

cacccgtcat gcaggtctta ggagagggga gaagaaaata agatttgtgt ggcggtgggg 39600

caggcaggat gctgggctta ttttactgaa caaacaatcg catttccata gaaacagcag 39660

cagtcgtgca gtatcatcca cagcctgaga gagagtcagc agcaagagct gagccggttt 39720

ctgaacccgc ccagcatcga gaccacccag cccagtgagg acacgaatgc caacagtcag 39780

gacaacagca tgcaacctga ggtatgggca cgaggcgtcc tctgtcccag gacactctgt 39840

gggcttcttt cttttaataa aaataaataa aatgggtgct ggtcagatca ttcaaacaag 39900

gtgacaatat ttctaatagg aatattttca ttaaaaaaaa aaanggcaac ggaaatccat 39960

tgatccagat tatgtctatg tctttgccat cttggatgca gttgaggtta tttaaaaaga 40020

gtctgactca tatgaatgcc aaaaaaagtg aatgccctga acttcataat tctgtagttt 40080

tatttttgtc taaagagtta acatcagctc taagggctgc atttgaatct gggtatatta 40140

catggagtat ggtttacagg aagctctgcc ctaaatattt tcctcccttc tctaaagtca 40200

tagaaatttg cagtaaaatc agtgaaaatg tttgtaatta gctttcaacc tccttaaata 40260

atcagactca ttgaagatat atcctttcct ggttgggtat cagaagaaca tatggaattc 40320

aggacaattg gcttgaggag tttctattta ttaagtcaca gcatattttc attttcaaac 40380

tctatgtatg tgtctatata aaaaagtaca aggagttatc agcgtcctgg agagtcctgt 40440

gctgtgagaa gcagctggct gggaggacct gcccctttat gaggcagcaa ttacgatttg 40500

aggttaacaa cgtggagctg atcccgtgct ctgtgcagag cccccagcag caccaggctc 40560

catggtgtgc agcaacgcag ccaggggcca gagtaccttt ggaagcctaa ggctccacgg 40620 cccagcaggg tgctccagga ggcccactgg tgccatagct cagttttcag tagatctcat 40680

gttttgggct ccccagcctt tagaaatcat ctgccaccga ctgcagggac ttctgcggag 40740

ttagtcaacc aaggggcaag gggaggcatc tgcagtggcc ttgccactgg tcaccctgaa 40800

gatggttggc aatgtctcgc atgggagagg gaggcaggct gaagtcagga cagacagtga 40860

gacagggcta atgaggttgg agaaagtgat gatttcagag ccctctgttt ttagatgaaa 40920

aagcctaaga aagaaagagg tgacnnnaaa aaaaaaaaaa ggatactaaa gagatgtgaa 40980

gaccaaggct tttgttagga aagggagagg tgacaggtga acctgagaac aggtgaaatg 41040

tgtgggaagt cttggcatct ctctttcttt cactgtagct tagattgtta tcagctccat 41100

cctccaccag ggcccatcat acagctttgc ctgcagagaa ctctttccct gaagctctaa 41160

gtgtggatct ccaggctggt gacagtagac tgcccccgcc cagccggtcc ccactgcccc 41220

accttccttc ctgcagttcc ccttccttac cccctggttt gccagcgaag atatgaaggg 41280

aaacaatgac attctcaggc atggagggga acgaagcagt gggggttacc caggacaaga 41340

tgtcccagga agagtggtgt ctccttcagc agggcgtgca gtggtggctc tttttcttct 41400

gtgacccttt atacccattt tcacttttcc ccattgtgga cactctgagt ccatgttgtt 41460

cttcccactt gtcctcctgc catgtgtttt cctttcttat agtaaaagga ggagagcgca 41520

gggttaaaaa gacggtaact gtgtgcttct cctcccatgc agacaagcag ccagcagcag 41580

ctcctgagcc ccacgctgtc ggatcgagga ggaagtcggc aagatgcagc cgacgcaggg 41640

aaaccccaga ggaaatttgg gcagtggcgt ctgccctcag gtaggtccat ccaggcattt 41700

ctccagcgat cgaggcattt taaaggtatt tactctgttt gtgtgttttg ggtcttgcct 41760

ttcaaaatgc aagtctgcat ctacagttgt ttacagacag caacataatg aaaaatgtag 41820

tcttgtcaaa aacattgtcc cccaaaataa cttctctaaa tatgacttac attagccccc 41880

attttccggg tacatttcag gctatcatgg ttgagaagcc agcacctatg aaaaagacaa 41940

gattcagnaa gagggagaaa tttccaaggg ctttctgtgt gcctcaaggc atgctcaaac 42000

atgggcaaaa gtattcaact gaagagggag tgggcagatg caatcattca gaaaatgcca 42060

cgaagttctc aaaaaaggga caggcccccg atttctctca gcacactttg cagctgagcc 42120

ggtcactggc tttaggcaag ctgctttacc tccatacaat aacaannnnc aacaacaaac 42180

acttacgtag tactcactgt atgccagtca ctcttctaag tgcattactt gtgttctctc 42240

attcaatcct ttaataaaca gtcccattat tttccatatt ttactgattc agtaagagaa 42300

ttttcttcta tgagcttctt ctctataggg ttttggaaag ccaggataat cccattaaga 42360

gcactcagga atsgagggag agcaatgctc ccccacagccc agtgtccttg tacatgtttt 42420

tttctataga taactaagtt catgctaaag caggagcttt ataatgctct ttaactgtgc 42480

cccaacttca gcccaattga aaggagaaga tgtgtagcat atgtgttcca caaagcagat 42540

gacagcacag cttacatttt gaggctgacg atgttcagtg ggtcttcact gggactacca 42600

ccaaggaaag tatccccttt catatccagg aacttatttt tcagagatca gagaagatct 42660 aggttcctcc tgattcaaac acagcagaga atgacagcat caagacaacg tagatggtgg 42720

tgccaggtca taaattacaa gctgagtcgg ttcaatttta tcctgtaggc aattaggagc 42780

tagcaaagat ttctgagcag tatgtgactt ttggaatctg tgctttagga agttgacttg 42840

gcagcaaagg agggaattgt ctatgacaga gcctggaggc aggttacaag ctggaggaag 42900

ttaccgtggt gtgagcacaa ggcaacaaag gcatggacag agctgggggc cagatactgc 42960

agggacagat aggagaggtg acctggggag aatgcacagt ccttggtatt catttaaatg 43020

gacaaaagaa atgatgctgt caaaagtact tccattgggc caggcgcagt ggctcacccc 43080

tgtaatccca gcactttggg aggccgaggc aggtggatca tttgaggcct ggagttcaag 43140

accagcctgg tcaacatggt gaaaccctgt ctctactaaa aatacaaaaa ttagctgacc 43200

ggtagtggtg tgcacctgta atcccagctg cttgggaggc tgaggcagga gaatcacttg 43260

agcctgggaa gcagaggttg tggtgaacca agattgtgcc gctgcactcc tgtctgggtg 43320

agagagtgag aacctgtctc ccaaaaaaaa aaaaaaaaaa aaaaaaaaaa nngtacttcc 43380

actggtccat gcctaggtgg tccccaccaa aaacggtgat cgtgaaatag acatgctctt 43440

ggggtcaggc tgagtttgag gaggcagagg agaagaaaaa tagtattttg agggggcagc 43500

aacaatgtca caggtgtcat gctatgcact ttcctaagtc atctcatgta acctgtttta 43560

atcctgacaa tagtctaatt aggacatttt cattgtcttc atcttacaat ggaggaaact 43620

tgggttcaga gacgtgaagt tgcccatagc tacatatcca caggttgcag aggtgggtcc 43680

caacaccatg actgtgaggc cccagaacca cacctttcct cccataccct cactgcatcc 43740

cccctcgcct ttactgtggg tcacactctc cttccttttt gttttttta ttgcatatgc 43800

taatgaatgc acaggtttct gtgacactta atagttcccg aattggaatt tttttagtga 43860

ttttctagtt tttttaatta tttcaaaatg aaaggggtga gcagcacgtt tttttcccaa 43920

atgacatcat tagttgtcca tgagcacaat taaaaaaaaa atctatcgtg aacataaaga 43980

atcaagtcct gaaaaggatg tttacttagc atctgtggct aggacacgtt agccaaaagt 44040

atcactgatg ggaatatcat gggtttttag gagtgattac aatagggtga agcaatacca 44100

agtggctcat aggactcatt aattccacag ccccaaaacc aataagccat tcagtgtcct 44160

cagtcaactt acggtttgga ggaaggacaa ccatgaaatc tgtcgtgtgc aaaatgaacc 44220

ccatgactga cgcggcttcc tgcggttctg aagttaagaa gtggtggacc cggcagctga 44280

ctgtggagag cgacgaaagt ggggatgacc ttctggatat ttaggtggat gtcaatgtag 44340

atgaatttct agtggtggaa accgttttct aataatgtcc ttgattgtcc agtgagcaat 44400

ctgtaattga tctataactg aattccagct tgtcacaaga tgtttataaa ttgattttca 44460

tcctgccaca gaaaggcata agctgcatgt atgatgggtt actatcaatc attgctcaaa 44520

aaaatttttg tataatgaca gtactgataa tattagaaat gataccgcaa gcaaatgtat 44580

atcacttaaa aatgtcatat attctgtctg cgtaaactaa ggtatatatt catatgtgct 44640

ctaatgcagt attatcaccg ccccgcaaaa gagtgctaag cccaaagtgg ctgatattta 44700 gggtacaggg gttatagctt tagttcacat ctttcccatt tccactagaa atatttctct 44760

tgagagaatt tattatttat gattgatctg aaaaggtcag cactgaactt atgctaaaat 44820

gatagtagtt ttacaaacta cagattctga attttaaaaa gtatcttctt tttctcgtgt 44880

tatattttta aatatacaca agacatttgg tgaccagaac aagttgattt ctgtcctcag 44940

ttatgttaat gaaactgttg cctccttcta agaaaattgt gtgtgcaagc accaggcaaa 45000

gaaatggact caggatgctt agcggtttaa aacaaacctg tagataaatc acttgagtga 45060

catagttgcg caaagatgtt aagtttctta agaaaccttt taataactga gtttagcaaa 45120

aagaataaaa ctatatagct caatttattt aaaaaaatct ttgcatgtgt gatgttatca 45180

ttggcttcat ttcttaccca aggtatgtct gttttgccat aaatcagcag agtcatttca 45240

ttctgggtga tcctaacaca ccattgctat gttagatttg aaatgacatc tctgttaaaa 45300

gaatcttcta tggaaataat ggtgccctgc aaaatcttcc tctgaactca caggttaggg 45360

atcacacaac ttacttaatc gtttttgtt tttgtttttt ttccttatat gtcaatggcc 45420

catgtcctcc gggaaaatta gaaaagcaaa atgattacaa agtgctgtta gatttcttgt 45480

gctgggccag ccaagtagaa gtggacttga cttggacctt taactatttt attacagatt 45540

ggacatttgc tgttcagatg ttttttaaca gagggattat ctcagaatcc tgtgacctcc 45600

aggttgtttt ataatctatt tttctctatt taacattcct cagatagata ggcaaatagg 45660

acattccttc tgtgtcacag aagtatcgtg gtagtggcag tctacagttt atatgattca 45720

ttgtaactat gagataaaga acaaccagtc atgtggccaa aaggattaga tttgatttga 45780

tgttcacttg gagtttactt tttgtacata caagataaaa taaatattgg atttgtaaaa 45840

taatttactc tgagttgacc atttttaaac ccttttcttg gtggggcagg cactcggtag 45900

cagcacattg caaacctgga ttgaatattc agtccataat tagttggcaa aagctgtgtc 45960

cctgatactc atttcaactc aattctctga gccagaatgt tcctttcaca cttgtgctga 46020

taaggctgag tgaataccct gtgccctcta agcctacaat gaacagaaaa ggcacttaca 46080

ggtccagggc atgggtctgt actagtgctg cataaaaaca agatgttttt ggctcattac 46140

acttgagtaa tacagcttca ttctcacgaa actgctgagc cctttctaaa acgtctggaa 46200

ttgatgttca gcagaacatc attgttgttc attctgcagt gaagccaggg gcagaatacc 46260

aatgtattta aagccatagc accctatatt ttcaatgaag cgaaacagtc acacacttcc 46320

tgtaggactc ctctctacga ttttcctgaa acaggtagag catatgcaag ccagagaaat 46380

taagcgccac ccccatttag attttacccc tagatatgta aacatatgaa aaatcccatt 46440

acacagatgg gaaaacttgg gtcacagtgt tccctggtcc gtagggaagc cagacactga 46500

tactaatcct gtcatcacgc cacccactgc ctttaattag agggtgctgt cagttccaaa 46560

ggcaccagaa agcatatggg gtggaatact tacttgactc agagcccctg agagagcagg 46620

ccacacatgt gtgacctgta tggcacttca aggagactca aagggaacaa ggatggacag 46680

aatggaaggg ccaatgctat tcactctann nnnnnnnnnn tttttttttt ttttttgaga 46740 cggagtcttg ctctgtcacc caggctggag tgcagtggca cgatctcggc tcactgcaag 46800

ctccacctcc cgggttcaca ccattctcct gcctcagcct cccaattagc tgggactaca 46860

ggtgcccgcc accatgccca gctaattttt ttgtattttt agtagagacg gggtttcacc 46920

atgttagcca ggatggtctc gatctcctga cctcgtgatc cacccgcctc ggcctcccaa 46980

agtgctggga ttacaggcgt gagccaccgt gcccggccca ctctattctt aatggtactt 47040

ccagaccaac atacatcagc agtgtctggc tggctgtgcc ttggccaagc tgaatgagat 47100

aagggtacat agaagaaaaa ccagtgttat aaagccttac aggaaagcag ctttaagcca 47160

gtggaatgat gtgtttgcct gtttccttgt aggaagtggc attctataca tgggtggttt 47220

tcagtatatc agtaagcagt ggaactgttt ctaaggcctt caacagatcc caaaatttta 47280

accagccaaa agccaggcag ccctcgttga agttgaggtg aagatgaaaa cggcactcat 47340

gtggtctctt ctcacccact attatgggtt gaagtgtgac ccccagaagg ataggttgaa 47400

gtcctaaccc ttagtacctg tgaatatgac cttattagga aatagggtct tcgcagatat 47460

aattacttat gaggaggtca tactagagga gggtaatccc ttactccaac atacctggta 47520

tcctctgaag acacaggcat gtggggagga tggccatgag acggtgaagg cagaggcatc 47580

ggagctatgc tgcccagagc caagggacac ctgggctacc agaagctgca agaggcaagg 47640

aaggcgcctc ccagaggctt ctaagggaga atggcccccct gtcaacacct tgattctgga 47700

cttctagcca gaagaactgt gagaacatac atttctgttg ttttaagctg cccagtttgt 47760

ctgctttttt atcattttat cattttagtt ttttagttta gtctctgtct tctactaact 47820

ccctttccaa ttctacttta g 47841

<210> 4

<211> 6799

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> 1..65

<220>

<221> CDS

<222> 66..5282

<220>

<221> 3'UTR

<222> 5283..6799

<220>

<221> polyA_signal

```
<222> 6081..6086
<220>
<221> polyA_signal
<222> 6777..6782
<220>
<221> allele
<222> 1658
<223> 99-79316-158  : polymorphic base C or T
<220>
<221> allele
<222> 4481
<223> 99-79310-29  : polymorphic base A or C
<400> 4
gacagctctg agcgctgggg ttacagactg tggttttgtg cttgctcacc aaagctaacc     60
tcagc atg ctc aaa agg aag cag agt tcc agg gtg gaa gcc cag cca gtc    110
      Met Leu Lys Arg Lys Gln Ser Ser Arg Val Glu Ala Gln Pro Val
      1               5                   10                  15
act gac ttt ggt cct gat gag tct ctg tcg gat aat gct gac atc ctc      158
Thr Asp Phe Gly Pro Asp Glu Ser Leu Ser Asp Asn Ala Asp Ile Leu
                20                  25                  30
tgg att aac aaa cca tgg gtt cac tct ttg ctg cgc atc tgt gcc atc      206
Trp Ile Asn Lys Pro Trp Val His Ser Leu Leu Arg Ile Cys Ala Ile
            35                  40                  45
atc agc gtc att tct gtt tgt atg aat acg cca atg acc ttc gag cac      254
Ile Ser Val Ile Ser Val Cys Met Asn Thr Pro Met Thr Phe Glu His
        50                  55                  60
tat cct cca ctt cag tat gtg acc ttc act ttg gat aca tta ttg atg      302
Tyr Pro Pro Leu Gln Tyr Val Thr Phe Thr Leu Asp Thr Leu Leu Met
    65                  70                  75
ttt ctc tac acg gca gag atg ata gca aaa atg cac atc cgg ggc att      350
Phe Leu Tyr Thr Ala Glu Met Ile Ala Lys Met His Ile Arg Gly Ile
80                  85                  90                  95
gtc aag ggg gat agt tcc tat gtg aaa gat cgc tgg tgt gtt ttt gat      398
Val Lys Gly Asp Ser Ser Tyr Val Lys Asp Arg Trp Cys Val Phe Asp
```

```
                100                 105                 110
gga ttt atg gtc ttt tgc ctt tgg gtt tct ttg gtg cta cag gtg ttt      446
Gly Phe Met Val Phe Cys Leu Trp Val Ser Leu Val Leu Gln Val Phe
            115                 120                 125
gaa att gct gat ata gtt gat cag atg tca cct tgg ggc atg ttg cgg      494
Glu Ile Ala Asp Ile Val Asp Gln Met Ser Pro Trp Gly Met Leu Arg
        130                 135                 140
att cca cgg cca ctg att atg atc cga gca ttc cgg att tat ttc cga      542
Ile Pro Arg Pro Leu Ile Met Ile Arg Ala Phe Arg Ile Tyr Phe Arg
    145                 150                 155
ttt gaa ctg cca agg acc aga att aca aat att tta aag cga tcg gga      590
Phe Glu Leu Pro Arg Thr Arg Ile Thr Asn Ile Leu Lys Arg Ser Gly
160                 165                 170                 175
gaa caa ata tgg agt gtt tcc att ttt cta ctt ttc ttt cta ctt ctt      638
Glu Gln Ile Trp Ser Val Ser Ile Phe Leu Leu Phe Phe Leu Leu Leu
                180                 185                 190
tat gga att tta gga gtt cag atg ttt gga aca ttt act tat cac tgt      686
Tyr Gly Ile Leu Gly Val Gln Met Phe Gly Thr Phe Thr Tyr His Cys
            195                 200                 205
gtt gta aat gac aca aag cca ggg aat gta acc tgg aat agt tta gct      734
Val Val Asn Asp Thr Lys Pro Gly Asn Val Thr Trp Asn Ser Leu Ala
        210                 215                 220
att cca gac aca cac tgc tca cca gag cta gaa gaa ggc tac cag tgc      782
Ile Pro Asp Thr His Cys Ser Pro Glu Leu Glu Glu Gly Tyr Gln Cys
    225                 230                 235
cca cct gga ttt aaa tgc atg gac ctt gaa gat ctg gga ctt agc agg      830
Pro Pro Gly Phe Lys Cys Met Asp Leu Glu Asp Leu Gly Leu Ser Arg
240                 245                 250                 255
caa gag ctg ggc tac agt ggc ttt aat gag ata gga act agt ata ttc      878
Gln Glu Leu Gly Tyr Ser Gly Phe Asn Glu Ile Gly Thr Ser Ile Phe
                260                 265                 270
acc gtc tat gag gcc gcc tca cag gaa ggc tgg gtg ttc ctc atg tac      926
Thr Val Tyr Glu Ala Ala Ser Gln Glu Gly Trp Val Phe Leu Met Tyr
            275                 280                 285
```

```
aga gca att gac agc ttt ccc cgt tgg cgt tcc tac ttc tat ttc atc      974
Arg Ala Ile Asp Ser Phe Pro Arg Trp Arg Ser Tyr Phe Tyr Phe Ile
        290                 295                 300
act ctc att ttc ttc ctc gcc tgg ctt gtg aag aac gtg ttt att gct     1022
Thr Leu Ile Phe Phe Leu Ala Trp Leu Val Lys Asn Val Phe Ile Ala
    305                 310                 315
gtt atc att gaa aca ttt gca gaa atc aga gta cag ttt caa caa atg     1070
Val Ile Ile Glu Thr Phe Ala Glu Ile Arg Val Gln Phe Gln Gln Met
320                 325                 330                 335
tgg gga tcg aga agc agc act acc tca aca gcc acc acc cag atg ttt     1118
Trp Gly Ser Arg Ser Ser Thr Thr Ser Thr Ala Thr Thr Gln Met Phe
                340                 345                 350
cat gaa gat gct gct gga ggt tgg cag ctg gta gct gtg gat gtc aac     1166
His Glu Asp Ala Ala Gly Gly Trp Gln Leu Val Ala Val Asp Val Asn
            355                 360                 365
aag ccc cag gga cgc gcc cca gcc tgc ctc cag aaa atg atg cgg tca     1214
Lys Pro Gln Gly Arg Ala Pro Ala Cys Leu Gln Lys Met Met Arg Ser
        370                 375                 380
tcc gtt ttc cac atg ttc atc ctg agc atg gtg acc gtg gac gtg atc     1262
Ser Val Phe His Met Phe Ile Leu Ser Met Val Thr Val Asp Val Ile
    385                 390                 395
gtg gcg gct agc aac tac tac aaa gga gaa aac ttc agg agg cag tac     1310
Val Ala Ala Ser Asn Tyr Tyr Lys Gly Glu Asn Phe Arg Arg Gln Tyr
400                 405                 410                 415
gac gag ttc tac ctg gcg gag gtg gct ttt aca gta ctt ttt gat ttg     1358
Asp Glu Phe Tyr Leu Ala Glu Val Ala Phe Thr Val Leu Phe Asp Leu
                420                 425                 430
gaa gca ctt ctg aag ata tgg tgt ttg gga ttt act gga tat att agc     1406
Glu Ala Leu Leu Lys Ile Trp Cys Leu Gly Phe Thr Gly Tyr Ile Ser
            435                 440                 445
tca tct ctc cac aaa ttc gaa cta cta ctc gta att gga act act ctt     1454
Ser Ser Leu His Lys Phe Glu Leu Leu Leu Val Ile Gly Thr Thr Leu
        450                 455                 460
cat gta tac cca gat ctt tat cat tca caa ttc acg tac ttt cag gtt     1502
```

His Val Tyr Pro Asp Leu Tyr His Ser Gln Phe Thr Tyr Phe Gln Val

465 470 475

ctc cga gta gtt cgg ctg att aag att tca cct gca tta gaa gac ttt 1550

Leu Arg Val Val Arg Leu Ile Lys Ile Ser Pro Ala Leu Glu Asp Phe

480 485 490 495

gtg tac aag ata ttt ggt cct gga aaa aag ctt ggg agt ttg gtt gta 1598

Val Tyr Lys Ile Phe Gly Pro Gly Lys Lys Leu Gly Ser Leu Val Val

500 505 510

ttt act gcc agc ctc ttg att gtt atg tca gca att agt ttg cag atg 1646

Phe Thr Ala Ser Leu Leu Ile Val Met Ser Ala Ile Ser Leu Gln Met

515 520 525

ttc tgc ttt gty gaa gaa ctg gac aga ttt act acg ttt ccg agg gca 1694

Phe Cys Phe Val Glu Glu Leu Asp Arg Phe Thr Thr Phe Pro Arg Ala

530 535 540

ttt atg tcc atg ttc cag atc ctc acc cag gaa gga tgg gtg gac gta 1742

Phe Met Ser Met Phe Gln Ile Leu Thr Gln Glu Gly Trp Val Asp Val

545 550 555

atg gac caa act cta aat gct gtg gga cat atg tgg gca ccc gtg gtt 1790

Met Asp Gln Thr Leu Asn Ala Val Gly His Met Trp Ala Pro Val Val

560 565 570 575

gcc atc tat ttc att ctc tat cat ctt ttt gcc act ctg atc ctc ctg 1838

Ala Ile Tyr Phe Ile Leu Tyr His Leu Phe Ala Thr Leu Ile Leu Leu

580 585 590

agt ttg ttt gtt gct gtt att ttg gac aac tta gaa ctt gat gaa gac 1886

Ser Leu Phe Val Ala Val Ile Leu Asp Asn Leu Glu Leu Asp Glu Asp

595 600 605

cta aag aag ctt aaa caa tta aag caa agt gaa gca aat gcg gac acc 1934

Leu Lys Lys Leu Lys Gln Leu Lys Gln Ser Glu Ala Asn Ala Asp Thr

610 615 620

aaa gaa aag ctc cct tta cgc ctg cga atc ttt gaa aaa ttt cca aac 1982

Lys Glu Lys Leu Pro Leu Arg Leu Arg Ile Phe Glu Lys Phe Pro Asn

625 630 635

aga cct caa atg gtg aaa atc tca aag ctt cct tca gat ttt aca gtt 2030

Arg Pro Gln Met Val Lys Ile Ser Lys Leu Pro Ser Asp Phe Thr Val

```
640                 645                 650                 655
cct aaa atc agg gag agt ttt atg aag cag ttt att gac cgc cag caa     2078
Pro Lys Ile Arg Glu Ser Phe Met Lys Gln Phe Ile Asp Arg Gln Gln
                660                 665                 670
cag gac aca tgt tgc ctc ctg aga agc ctc ccg acc acc tct tcc tcc     2126
Gln Asp Thr Cys Cys Leu Leu Arg Ser Leu Pro Thr Thr Ser Ser Ser
            675                 680                 685
tcc tgc gac cac tcc aaa cgc tca gca att gag gac aac aaa tac atc     2174
Ser Cys Asp His Ser Lys Arg Ser Ala Ile Glu Asp Asn Lys Tyr Ile
        690                 695                 700
gac caa aaa ctt cgc aag tct gtt ttc agc atc agg gca agg aac ctt     2222
Asp Gln Lys Leu Arg Lys Ser Val Phe Ser Ile Arg Ala Arg Asn Leu
    705                 710                 715
ctg gaa aag gag acc gca gtc act aaa atc tta aga gct tgc acc cga     2270
Leu Glu Lys Glu Thr Ala Val Thr Lys Ile Leu Arg Ala Cys Thr Arg
720                 725                 730                 735
cag cgc atg ctg agc gga tca ttt gag ggg cag ccc gca aag gag agg     2318
Gln Arg Met Leu Ser Gly Ser Phe Glu Gly Gln Pro Ala Lys Glu Arg
                740                 745                 750
tca atc ctc agc gtg cag cat cat atc cgc caa gag cgc agg tca cta     2366
Ser Ile Leu Ser Val Gln His His Ile Arg Gln Glu Arg Arg Ser Leu
            755                 760                 765
aga cat gga tca aac agc cag agg atc agc agg gga aaa tct ctt gaa     2414
Arg His Gly Ser Asn Ser Gln Arg Ile Ser Arg Gly Lys Ser Leu Glu
        770                 775                 780
act ttg act caa gat cat tcc aat aca gtg aga tat aga aat gca caa     2462
Thr Leu Thr Gln Asp His Ser Asn Thr Val Arg Tyr Arg Asn Ala Gln
    785                 790                 795
aga gaa gac agt gaa ata aag atg att cag gaa aaa aag gag caa gca     2510
Arg Glu Asp Ser Glu Ile Lys Met Ile Gln Glu Lys Lys Glu Gln Ala
800                 805                 810                 815
gag atg aaa agg aaa gtg caa gaa gag gaa ctc aga gag aac cac cca     2558
Glu Met Lys Arg Lys Val Gln Glu Glu Glu Leu Arg Glu Asn His Pro
                820                 825                 830
```

tac ttc gat aag cca ctg ttc att gtc ggg cga gaa cac agg ttc aga 2606

Tyr Phe Asp Lys Pro Leu Phe Ile Val Gly Arg Glu His Arg Phe Arg

835 840 845

aac ttt tgc cgg gtg gtg gtc cga gca cgc ttc aac gca tct aaa aca 2654

Asn Phe Cys Arg Val Val Val Arg Ala Arg Phe Asn Ala Ser Lys Thr

850 855 860

gac cct gtc aca gga gct gtg aaa aat aca aag tac cat caa ctt tat 2702

Asp Pro Val Thr Gly Ala Val Lys Asn Thr Lys Tyr His Gln Leu Tyr

865 870 875

gat ttg ctg gga ttg gtc act tac ctg gac tgg gtc atg atc atc gta 2750

Asp Leu Leu Gly Leu Val Thr Tyr Leu Asp Trp Val Met Ile Ile Val

880 885 890 895

acc atc tgc tct tgc att tcc atg atg ttt gag tcc ccg ttt cga aga 2798

Thr Ile Cys Ser Cys Ile Ser Met Met Phe Glu Ser Pro Phe Arg Arg

900 905 910

gtc atg cat gca cct act ttg cag att gct gag tat gtg ttt gtg ata 2846

Val Met His Ala Pro Thr Leu Gln Ile Ala Glu Tyr Val Phe Val Ile

915 920 925

ttc atg agc att gag ctt aat ctg aag att atg gca gat ggc tta ttt 2894

Phe Met Ser Ile Glu Leu Asn Leu Lys Ile Met Ala Asp Gly Leu Phe

930 935 940

ttc act cca act gct gtc atc agg gac ttc ggt gga gta atg gac ata 2942

Phe Thr Pro Thr Ala Val Ile Arg Asp Phe Gly Gly Val Met Asp Ile

945 950 955

ttt ata tat ctt gtg agc ttg ata ttt ctt tgt tgg atg cct caa aat 2990

Phe Ile Tyr Leu Val Ser Leu Ile Phe Leu Cys Trp Met Pro Gln Asn

960 965 970 975

gta cct gct gaa tcg gga gct cag ctt cta atg gtc ctt cgg tgc ctg 3038

Val Pro Ala Glu Ser Gly Ala Gln Leu Leu Met Val Leu Arg Cys Leu

980 985 990

aga cct ctg cgc ata ttc aaa ctg gtg ccc cag atg agg aaa gtt gtt 3086

Arg Pro Leu Arg Ile Phe Lys Leu Val Pro Gln Met Arg Lys Val Val

995 1000 1005

cga gaa ctt ttc agc ggc ttc aag gaa att ttt ttg gtc tcc att ctt 3134

Arg Glu Leu Phe Ser Gly Phe Lys Glu Ile Phe Leu Val Ser Ile Leu

1010 1015 1020

ttg ctg aca tta atg ctc gtt ttt gca agc ttt gga gtt cag ctt ttt 3182

Leu Leu Thr Leu Met Leu Val Phe Ala Ser Phe Gly Val Gln Leu Phe

1025 1030 1035

gct gga aaa ctg gcc aag tgc aat gat ccc aac att att aga agg gaa 3230

Ala Gly Lys Leu Ala Lys Cys Asn Asp Pro Asn Ile Ile Arg Arg Glu

1040 1045 1050 1055

gat tgc aat ggc ata ttc aga att aat gtc agt gtg tca aag aac tta 3278

Asp Cys Asn Gly Ile Phe Arg Ile Asn Val Ser Val Ser Lys Asn Leu

1060 1065 1070

aat tta aaa ttg agg cct gga gag aaa aaa cct gga ttt tgg gtg ccc 3326

Asn Leu Lys Leu Arg Pro Gly Glu Lys Lys Pro Gly Phe Trp Val Pro

1075 1080 1085

cgt gtt tgg gcg aat cct cgg aac ttt aat ttc gac aat gtg gga aac 3374

Arg Val Trp Ala Asn Pro Arg Asn Phe Asn Phe Asp Asn Val Gly Asn

1090 1095 1100

gct atg ctg gcg ttg ttt gaa gtt ctc tcc ttg aaa ggc tgg gtg gaa 3422

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

1105 1110 1115

gtg aga gat gtt att att cat cgt gtg ggg ccg atc cat gga atc tat 3470

Val Arg Asp Val Ile Ile His Arg Val Gly Pro Ile His Gly Ile Tyr

1120 1125 1130 1135

att cat gtt ttt gta ttc ctg ggt tgc atg att gga ctg acc ctt ttt 3518

Ile His Val Phe Val Phe Leu Gly Cys Met Ile Gly Leu Thr Leu Phe

1140 1145 1150

gtt gga gta gtt att gct aat ttc aat gaa aac aag ggg acg gct ttg 3566

Val Gly Val Val Ile Ala Asn Phe Asn Glu Asn Lys Gly Thr Ala Leu

1155 1160 1165

ctg acc gtc gat cag aga aga tgg gaa gac ctg aag agc cga ctg aag 3614

Leu Thr Val Asp Gln Arg Arg Trp Glu Asp Leu Lys Ser Arg Leu Lys

1170 1175 1180

atc gca cag cct ctt cat ctc ccg cct cgc ccg gat aat gat ggt ttt 3662

Ile Ala Gln Pro Leu His Leu Pro Pro Arg Pro Asp Asn Asp Gly Phe

1185 1190 1195

aga gct aaa atg tat gac ata acc cag cat cca ttt ttt aag agg aca 3710

Arg Ala Lys Met Tyr Asp Ile Thr Gln His Pro Phe Phe Lys Arg Thr

1200 1205 1210 1215

atc gca tta ctc gtc ctg gcc cag tcg gtg ttg ctc tct gtc aag tgg 3758

Ile Ala Leu Leu Val Leu Ala Gln Ser Val Leu Leu Ser Val Lys Trp

1220 1225 1230

gac gtc gag gac ccg gtg acc gta cct ttg gca aca atg tca gtt gtt 3806

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

1235 1240 1245

ttc acc ttc atc ttt gtt ctg gag gtt acc atg aag atc ata gca atg 3854

Phe Thr Phe Ile Phe Val Leu Glu Val Thr Met Lys Ile Ile Ala Met

1250 1255 1260

tcg cct gct ggc ttc tgg caa agc aga aga aac cga tac gat ctc ctg 3902

Ser Pro Ala Gly Phe Trp Gln Ser Arg Arg Asn Arg Tyr Asp Leu Leu

1265 1270 1275

gtg acg tcg ctt ggc gtt gta tgg gtg gtg ctt cac ttt gcc ctc ctg 3950

Val Thr Ser Leu Gly Val Val Trp Val Val Leu His Phe Ala Leu Leu

1280 1285 1290 1295

aat gca tat act tac atg atg ggc gct tgt gtg att gta ttt agg ttt 3998

Asn Ala Tyr Thr Tyr Met Met Gly Ala Cys Val Ile Val Phe Arg Phe

1300 1305 1310

ttc tcc atc tgt gga aaa cat gta acg cta aag atg ctc ctc ttg aca 4046

Phe Ser Ile Cys Gly Lys His Val Thr Leu Lys Met Leu Leu Leu Thr

1315 1320 1325

gtg gtc gtc agc atg tac aag agc ttc ttt atc ata gta ggc atg ttt 4094

Val Val Val Ser Met Tyr Lys Ser Phe Phe Ile Ile Val Gly Met Phe

1330 1335 1340

ctc ttg ctg ctg tgt tac gct ttt gct gga gtt gtt tta ttt ggt act 4142

Leu Leu Leu Leu Cys Tyr Ala Phe Ala Gly Val Val Leu Phe Gly Thr

1345 1350 1355

gtg aaa tat ggg gag aat att aac agg cat gca aat ttt tct tcg gct 4190

Val Lys Tyr Gly Glu Asn Ile Asn Arg His Ala Asn Phe Ser Ser Ala

1360 1365 1370 1375

gga aaa gct att acc gta ctg ttc cga att gtc aca ggt gaa gac tgg 4238

Gly Lys Ala Ile Thr Val Leu Phe Arg Ile Val Thr Gly Glu Asp Trp

1380 1385 1390

aac aag att atg cat gac tgt atg gtt cag cct ccg ttt tgt act cca 4286

Asn Lys Ile Met His Asp Cys Met Val Gln Pro Pro Phe Cys Thr Pro

1395 1400 1405

gat gaa ttt aca tac tgg gca aca gac tgt gga aat tat gct ggg gca 4334

Asp Glu Phe Thr Tyr Trp Ala Thr Asp Cys Gly Asn Tyr Ala Gly Ala

1410 1415 1420

ctt atg tat ttc tgt tca ttt tat gtc atc att gcc tac atc atg cta 4382

Leu Met Tyr Phe Cys Ser Phe Tyr Val Ile Ile Ala Tyr Ile Met Leu

1425 1430 1435

aat ctg ctt gta gcc ata att gtg gag aat ttc tcc ttg ttt tat tcc 4430

Asn Leu Leu Val Ala Ile Ile Val Glu Asn Phe Ser Leu Phe Tyr Ser

1440 1445 1450 1455

act gag gag gac cag ctt tta agt tac aat gat ctt cgc cac ttt caa 4478

Thr Glu Glu Asp Gln Leu Leu Ser Tyr Asn Asp Leu Arg His Phe Gln

1460 1465 1470

atm ata tgg aac atg gtg gat gat aaa aga gag ggg gtg atc ccc acg 4526

Ile Ile Trp Asn Met Val Asp Asp Lys Arg Glu Gly Val Ile Pro Thr

1475 1480 1485

ttc cgc gtc aag ttc ctg ctg cgg cta ctg cgt ggg agg ctg gag gtg 4574

Phe Arg Val Lys Phe Leu Leu Arg Leu Leu Arg Gly Arg Leu Glu Val

1490 1495 1500

gac ctg gac aag gac aag ctc ctg ttt aag cac atg tgc tac gaa atg 4622

Asp Leu Asp Lys Asp Lys Leu Leu Phe Lys His Met Cys Tyr Glu Met

1505 1510 1515

gag agg ctc cac aat ggc ggc gac gtc acc ttc cat gat gtc ctg agc 4670

Glu Arg Leu His Asn Gly Gly Asp Val Thr Phe His Asp Val Leu Ser

1520 1525 1530 1535

atg ctt tca tac cgg tcc gtg gac atc cgg aag agc ttg cag ctg gag 4718

Met Leu Ser Tyr Arg Ser Val Asp Ile Arg Lys Ser Leu Gln Leu Glu

1540 1545 1550

gaa ctc ctg gcg agg gag cag ctg gag tac acc ata gag gag gag gtg 4766

```
Glu Leu Leu Ala Arg Glu Gln Leu Glu Tyr Thr Ile Glu Glu Glu Val
            1555                1560                1565

gcc aag cag acc atc cgc atg tgg ctc aag aag tgc ctg aag cgc atc      4814
Ala Lys Gln Thr Ile Arg Met Trp Leu Lys Lys Cys Leu Lys Arg Ile
        1570                1575                1580

aga gct aaa cag cag cag tcg tgc agt atc atc cac agc ctg aga gag      4862
Arg Ala Lys Gln Gln Gln Ser Cys Ser Ile Ile His Ser Leu Arg Glu
    1585                1590                1595

agt cag cag caa gag ctg agc cgg ttt ctg aac ccg ccc agc atc gag      4910
Ser Gln Gln Gln Glu Leu Ser Arg Phe Leu Asn Pro Pro Ser Ile Glu
1600                1605                1610                1615

acc acc cag ccc agt gag gac acg aat gcc aac agt cag gac aac agc      4958
Thr Thr Gln Pro Ser Glu Asp Thr Asn Ala Asn Ser Gln Asp Asn Ser
                1620                1625                1630

atg caa cct gag aca agc agc cag cag cag ctc ctg agc ccc acg ctg      5006
Met Gln Pro Glu Thr Ser Ser Gln Gln Gln Leu Leu Ser Pro Thr Leu
            1635                1640                1645

tcg gat cga gga gga agt cgg caa gat gca gcc gac gca ggg aaa ccc      5054
Ser Asp Arg Gly Gly Ser Arg Gln Asp Ala Ala Asp Ala Gly Lys Pro
        1650                1655                1660

cag agg aaa ttt ggg cag tgg cgt ctg ccc tca gcc cca aaa cca ata      5102
Gln Arg Lys Phe Gly Gln Trp Arg Leu Pro Ser Ala Pro Lys Pro Ile
    1665                1670                1675

agc cat tca gtg tcc tca gtc aac tta cgg ttt gga gga agg aca acc      5150
Ser His Ser Val Ser Ser Val Asn Leu Arg Phe Gly Gly Arg Thr Thr
1680                1685                1690                1695

atg aaa tct gtc gtg tgc aaa atg aac ccc atg act gac gcg gct tcc      5198
Met Lys Ser Val Val Cys Lys Met Asn Pro Met Thr Asp Ala Ala Ser
                1700                1705                1710

tgc ggt tct gaa gtt aag aag tgg tgg acc cgg cag ctg act gtg gag      5246
Cys Gly Ser Glu Val Lys Lys Trp Trp Thr Arg Gln Leu Thr Val Glu
            1715                1720                1725

agc gac gaa agt ggg gat gac ctt ctg gat att tag gtggatgtca           5292
Ser Asp Glu Ser Gly Asp Asp Leu Leu Asp Ile  *
```

1730 1735

atgtagatga atttctagtg gtggaaaccg ttttctaata atgtccttga ttgtccagtg 5352

agcaatctgt aattgatcta taactgaatt ccagcttgtc acaagatgtt tataaattga 5412

ttttcatcct gccacagaaa ggcataagct gcatgtatga tgggttacta tcaatcattg 5472

ctcaaaaaaa tttttgtata atgacagtac tgataatatt agaaatgata ccgcaagcaa 5532

atgtatatca cttaaaaatg tcatatattc tgtctgcgta aactaaggta tatattcata 5592

tgtgctctaa tgcagtatta tcaccgcccc gcaaaagagt gctaagccca aagtggctga 5652

tatttagggt acaggggtta tagctttagt tcacatcttt cccatttcca ctagaaatat 5712

ttctcttgag agaatttatt atttatgatt gatctgaaaa ggtcagcact gaacttatgc 5772

taaaatgata gtagttttac aaactacaga ttctgaattt taaaaagtat cttctttttc 5832

tcgtgttata tttttaaata tacacaagac atttggtgac cagaacaagt tgatttctgt 5892

cctcagttat gttaatgaaa ctgttgcctc cttctaagaa aattgtgtgt gcaagcacca 5952

ggcaaagaaa tggactcagg atgcttagcg gtttaaaaca aacctgtaga taaatcactt 6012

gagtgacata gttgcgcaaa gatgttaagt ttcttaagaa accttttaat aactgagttt 6072

agcaaaaaga ataaaactat atagctcaat ttatttaaaa aaatctttgc atgtgtgatg 6132

ttatcattgg cttcatttct tacccaaggt atgtctgttt tgccataaat cagcagagtc 6192

atttcattct gggtgatcct aacacaccat tgctatgtta gatttgaaat gacatctctg 6252

ttaaaagaat cttctatgga aataatggtg ccctgcaaaa tcttcctctg aactcacagg 6312

ttagggatca cacaacttac ttaatcgttt tttgtttttg tttttttcc ttatatgtca 6372

atggcccatg tcctccggga aaattagaaa agcaaaatga ttacaaagtg ctgttagatt 6432

tcttgtgctg ggccagccaa gtagaagtgg acttgacttg gacctttaac tattttatta 6492

cagattggac atttgctgtt cagatgtttt ttaacagagg gattatctca gaatcctgtg 6552

acctccaggt tgttttataa tctatttttc tctatttaac attcctcaga tagataggca 6612

aataggacat tccttctgtg tcacagaagt atcgtggtag tggcagtcta cagtttatat 6672

gattcattgt aactatgaga taaagaacaa ccagtcatgt ggccaaaagg attagatttg 6732

atttgatgtt cacttggagt ttactttttg tacatacaag ataaaataaa tattggattt 6792

gtaaaat 6799

<210> 5

<211> 1738

<212> PRT

<213> Homo sapiens

<400> 5

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

1 5 10 15

```
Asp Phe Gly Pro Asp Glu Ser Leu Ser Asp Asn Ala Asp Ile Leu Trp
            20                  25                  30
Ile Asn Lys Pro Trp Val His Ser Leu Leu Arg Ile Cys Ala Ile Ile
        35                  40                  45
Ser Val Ile Ser Val Cys Met Asn Thr Pro Met Thr Phe Glu His Tyr
    50                  55                  60
Pro Pro Leu Gln Tyr Val Thr Phe Thr Leu Asp Thr Leu Leu Met Phe
65                  70                  75                  80
Leu Tyr Thr Ala Glu Met Ile Ala Lys Met His Ile Arg Gly Ile Val
                85                  90                  95
Lys Gly Asp Ser Ser Tyr Val Lys Asp Arg Trp Cys Val Phe Asp Gly
            100                 105                 110
Phe Met Val Phe Cys Leu Trp Val Ser Leu Val Leu Gln Val Phe Glu
        115                 120                 125
Ile Ala Asp Ile Val Asp Gln Met Ser Pro Trp Gly Met Leu Arg Ile
    130                 135                 140
Pro Arg Pro Leu Ile Met Ile Arg Ala Phe Arg Ile Tyr Phe Arg Phe
145                 150                 155                 160
Glu Leu Pro Arg Thr Arg Ile Thr Asn Ile Leu Lys Arg Ser Gly Glu
                165                 170                 175
Gln Ile Trp Ser Val Ser Ile Phe Leu Leu Phe Phe Leu Leu Leu Tyr
            180                 185                 190
Gly Ile Leu Gly Val Gln Met Phe Gly Thr Phe Thr Tyr His Cys Val
        195                 200                 205
Val Asn Asp Thr Lys Pro Gly Asn Val Thr Trp Asn Ser Leu Ala Ile
    210                 215                 220
Pro Asp Thr His Cys Ser Pro Glu Leu Glu Glu Gly Tyr Gln Cys Pro
225                 230                 235                 240
Pro Gly Phe Lys Cys Met Asp Leu Glu Asp Leu Gly Leu Ser Arg Gln
                245                 250                 255
Glu Leu Gly Tyr Ser Gly Phe Asn Glu Ile Gly Thr Ser Ile Phe Thr
            260                 265                 270
Val Tyr Glu Ala Ala Ser Gln Glu Gly Trp Val Phe Leu Met Tyr Arg
        275                 280                 285
```

Ala Ile Asp Ser Phe Pro Arg Trp Arg Ser Tyr Phe Tyr Phe Ile Thr

290 295 300

Leu Ile Phe Phe Leu Ala Trp Leu Val Lys Asn Val Phe Ile Ala Val

305 310 315 320

Ile Ile Glu Thr Phe Ala Glu Ile Arg Val Gln Phe Gln Gln Met Trp

325 330 335

Gly Ser Arg Ser Ser Thr Thr Ser Thr Ala Thr Thr Gln Met Phe His

340 345 350

Glu Asp Ala Ala Gly Gly Trp Gln Leu Val Ala Val Asp Val Asn Lys

355 360 365

Pro Gln Gly Arg Ala Pro Ala Cys Leu Gln Lys Met Met Arg Ser Ser

370 375 380

Val Phe His Met Phe Ile Leu Ser Met Val Thr Val Asp Val Ile Val

385 390 395 400

Ala Ala Ser Asn Tyr Tyr Lys Gly Glu Asn Phe Arg Arg Gln Tyr Asp

405 410 415

Glu Phe Tyr Leu Ala Glu Val Ala Phe Thr Val Leu Phe Asp Leu Glu

420 425 430

Ala Leu Leu Lys Ile Trp Cys Leu Gly Phe Thr Gly Tyr Ile Ser Ser

435 440 445

Ser Leu His Lys Phe Glu Leu Leu Leu Val Ile Gly Thr Thr Leu His

450 455 460

Val Tyr Pro Asp Leu Tyr His Ser Gln Phe Thr Tyr Phe Gln Val Leu

465 470 475 480

Arg Val Val Arg Leu Ile Lys Ile Ser Pro Ala Leu Glu Asp Phe Val

485 490 495

Tyr Lys Ile Phe Gly Pro Gly Lys Lys Leu Gly Ser Leu Val Val Phe

500 505 510

Thr Ala Ser Leu Leu Ile Val Met Ser Ala Ile Ser Leu Gln Met Phe

515 520 525

Cys Phe Val Glu Glu Leu Asp Arg Phe Thr Thr Phe Pro Arg Ala Phe

530 535 540

Met Ser Met Phe Gln Ile Leu Thr Gln Glu Gly Trp Val Asp Val Met

545 550 555 560

Asp Gln Thr Leu Asn Ala Val Gly His Met Trp Ala Pro Val Val Ala

565 570 575

Ile Tyr Phe Ile Leu Tyr His Leu Phe Ala Thr Leu Ile Leu Leu Ser

580 585 590

Leu Phe Val Ala Val Ile Leu Asp Asn Leu Glu Leu Asp Glu Asp Leu

595 600 605

Lys Lys Leu Lys Gln Leu Lys Gln Ser Glu Ala Asn Ala Asp Thr Lys

610 615 620

Glu Lys Leu Pro Leu Arg Leu Arg Ile Phe Glu Lys Phe Pro Asn Arg

625 630 635 640

Pro Gln Met Val Lys Ile Ser Lys Leu Pro Ser Asp Phe Thr Val Pro

645 650 655

Lys Ile Arg Glu Ser Phe Met Lys Gln Phe Ile Asp Arg Gln Gln Gln

660 665 670

Asp Thr Cys Cys Leu Leu Arg Ser Leu Pro Thr Thr Ser Ser Ser Ser

675 680 685

Cys Asp His Ser Lys Arg Ser Ala Ile Glu Asp Asn Lys Tyr Ile Asp

690 695 700

Gln Lys Leu Arg Lys Ser Val Phe Ser Ile Arg Ala Arg Asn Leu Leu

705 710 715 720

Glu Lys Glu Thr Ala Val Thr Lys Ile Leu Arg Ala Cys Thr Arg Gln

725 730 735

Arg Met Leu Ser Gly Ser Phe Glu Gly Gln Pro Ala Lys Glu Arg Ser

740 745 750

Ile Leu Ser Val Gln His His Ile Arg Gln Glu Arg Arg Ser Leu Arg

755 760 765

His Gly Ser Asn Ser Gln Arg Ile Ser Arg Gly Lys Ser Leu Glu Thr

770 775 780

Leu Thr Gln Asp His Ser Asn Thr Val Arg Tyr Arg Asn Ala Gln Arg

785 790 795 800

Glu Asp Ser Glu Ile Lys Met Ile Gln Glu Lys Lys Glu Gln Ala Glu

805 810 815

Met Lys Arg Lys Val Gln Glu Glu Glu Leu Arg Glu Asn His Pro Tyr

820 825 830

Phe Asp Lys Pro Leu Phe Ile Val Gly Arg Glu His Arg Phe Arg Asn

835 840 845

Phe Cys Arg Val Val Val Arg Ala Arg Phe Asn Ala Ser Lys Thr Asp

850 855 860

Pro Val Thr Gly Ala Val Lys Asn Thr Lys Tyr His Gln Leu Tyr Asp

865 870 875 880

Leu Leu Gly Leu Val Thr Tyr Leu Asp Trp Val Met Ile Ile Val Thr

885 890 895

Ile Cys Ser Cys Ile Ser Met Met Phe Glu Ser Pro Phe Arg Arg Val

900 905 910

Met His Ala Pro Thr Leu Gln Ile Ala Glu Tyr Val Phe Val Ile Phe

915 920 925

Met Ser Ile Glu Leu Asn Leu Lys Ile Met Ala Asp Gly Leu Phe Phe

930 935 940

Thr Pro Thr Ala Val Ile Arg Asp Phe Gly Gly Val Met Asp Ile Phe

945 950 955 960

Ile Tyr Leu Val Ser Leu Ile Phe Leu Cys Trp Met Pro Gln Asn Val

965 970 975

Pro Ala Glu Ser Gly Ala Gln Leu Leu Met Val Leu Arg Cys Leu Arg

980 985 990

Pro Leu Arg Ile Phe Lys Leu Val Pro Gln Met Arg Lys Val Val Arg

995 1000 1005

Glu Leu Phe Ser Gly Phe Lys Glu Ile Phe Leu Val Ser Ile Leu Leu

1010 1015 1020

Leu Thr Leu Met Leu Val Phe Ala Ser Phe Gly Val Gln Leu Phe Ala

1025 1030 1035 1040

Gly Lys Leu Ala Lys Cys Asn Asp Pro Asn Ile Ile Arg Arg Glu Asp

1045 1050 1055

Cys Asn Gly Ile Phe Arg Ile Asn Val Ser Val Ser Lys Asn Leu Asn

1060 1065 1070

Leu Lys Leu Arg Pro Gly Glu Lys Lys Pro Gly Phe Trp Val Pro Arg

1075 1080 1085

Val Trp Ala Asn Pro Arg Asn Phe Asn Phe Asp Asn Val Gly Asn Ala

1090 1095 1100

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

1105 1110 1115 1120

Arg Asp Val Ile Ile His Arg Val Gly Pro Ile His Gly Ile Tyr Ile

1125 1130 1135

His Val Phe Val Phe Leu Gly Cys Met Ile Gly Leu Thr Leu Phe Val

1140 1145 1150

Gly Val Val Ile Ala Asn Phe Asn Glu Asn Lys Gly Thr Ala Leu Leu

1155 1160 1165

Thr Val Asp Gln Arg Arg Trp Glu Asp Leu Lys Ser Arg Leu Lys Ile

1170 1175 1180

Ala Gln Pro Leu His Leu Pro Pro Arg Pro Asp Asn Asp Gly Phe Arg

1185 1190 1195 1200

Ala Lys Met Tyr Asp Ile Thr Gln His Pro Phe Phe Lys Arg Thr Ile

1205 1210 1215

Ala Leu Leu Val Leu Ala Gln Ser Val Leu Leu Ser Val Lys Trp Asp

1220 1225 1230

Val Glu Asp Pro Val Thr Val Pro Leu Ala Thr Met Ser Val Val Phe

1235 1240 1245

Thr Phe Ile Phe Val Leu Glu Val Thr Met Lys Ile Ile Ala Met Ser

1250 1255 1260

Pro Ala Gly Phe Trp Gln Ser Arg Arg Asn Arg Tyr Asp Leu Leu Val

1265 1270 1275 1280

Thr Ser Leu Gly Val Val Trp Val Val Leu His Phe Ala Leu Leu Asn

1285 1290 1295

Ala Tyr Thr Tyr Met Met Gly Ala Cys Val Ile Val Phe Arg Phe Phe

1300 1305 1310

Ser Ile Cys Gly Lys His Val Thr Leu Lys Met Leu Leu Leu Thr Val

1315 1320 1325

Val Val Ser Met Tyr Lys Ser Phe Phe Ile Ile Val Gly Met Phe Leu

1330 1335 1340

Leu Leu Leu Cys Tyr Ala Phe Ala Gly Val Val Leu Phe Gly Thr Val

1345 1350 1355 1360

Lys Tyr Gly Glu Asn Ile Asn Arg His Ala Asn Phe Ser Ser Ala Gly

1365 1370 1375

Lys Ala Ile Thr Val Leu Phe Arg Ile Val Thr Gly Glu Asp Trp Asn

1380 1385 1390

Lys Ile Met His Asp Cys Met Val Gln Pro Pro Phe Cys Thr Pro Asp

1395 1400 1405

Glu Phe Thr Tyr Trp Ala Thr Asp Cys Gly Asn Tyr Ala Gly Ala Leu

1410 1415 1420

Met Tyr Phe Cys Ser Phe Tyr Val Ile Ile Ala Tyr Ile Met Leu Asn

1425 1430 1435 1440

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

1445 1450 1455

Glu Glu Asp Gln Leu Leu Ser Tyr Asn Asp Leu Arg His Phe Gln Ile

1460 1465 1470

Ile Trp Asn Met Val Asp Asp Lys Arg Glu Gly Val Ile Pro Thr Phe

1475 1480 1485

Arg Val Lys Phe Leu Leu Arg Leu Leu Arg Gly Arg Leu Glu Val Asp

1490 1495 1500

Leu Asp Lys Asp Lys Leu Leu Phe Lys His Met Cys Tyr Glu Met Glu

1505 1510 1515 1520

Arg Leu His Asn Gly Gly Asp Val Thr Phe His Asp Val Leu Ser Met

1525 1530 1535

Leu Ser Tyr Arg Ser Val Asp Ile Arg Lys Ser Leu Gln Leu Glu Glu

1540 1545 1550

Leu Leu Ala Arg Glu Gln Leu Glu Tyr Thr Ile Glu Glu Glu Val Ala

1555 1560 1565

Lys Gln Thr Ile Arg Met Trp Leu Lys Lys Cys Leu Lys Arg Ile Arg

1570 1575 1580

Ala Lys Gln Gln Gln Ser Cys Ser Ile Ile His Ser Leu Arg Glu Ser

1585 1590 1595 1600

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

1605 1610 1615

Thr Gln Pro Ser Glu Asp Thr Asn Ala Asn Ser Gln Asp Asn Ser Met

1620 1625 1630

Gln Pro Glu Thr Ser Ser Gln Gln Gln Leu Leu Ser Pro Thr Leu Ser

1635 1640 1645

Asp Arg Gly Gly Ser Arg Gln Asp Ala Ala Asp Ala Gly Lys Pro Gln

1650 1655 1660

Arg Lys Phe Gly Gln Trp Arg Leu Pro Ser Ala Pro Lys Pro Ile Ser

1665 1670 1675 1680

His Ser Val Ser Ser Val Asn Leu Arg Phe Gly Gly Arg Thr Thr Met

1685 1690 1695

Lys Ser Val Val Cys Lys Met Asn Pro Met Thr Asp Ala Ala Ser Cys

1700 1705 1710

Gly Ser Glu Val Lys Lys Trp Trp Thr Arg Gln Leu Thr Val Glu Ser

1715 1720 1725

Asp Glu Ser Gly Asp Asp Leu Leu Asp Ile

1730 1735

<210> 6

<211> 453

<212> DNA

<213> Homo sapiens

<220>

<221> allele

<222> 105

<223> 99-62617-105 : polymorphic base G or C

<220>

<221> misc_binding

<222> 93..117

<223> 99-62617-105.probe

<220>

<221> primer_bind

<222> 86..104

<223> 99-62617-105.mis

<220>

<221> primer_bind

<222> 106..124

<223> 99-62617-105.mis complement

<220>

<221> primer_bind

<222> 1..20

<223> 99-62617.pu

<220>

<221> primer_bind

<222> 435..453

<223> 99-62617.rp complement

<400> 6

```
tgaggatgag gggcaaacac agaggaagaa tctgggcccc atgaggtcaa cagtgagctg     60
ctaaaattaa cctgaaattt ccctgctggg agacattgtg ttgtsaaata gttttatacc    120
atcttttact aggttttcaa ttaccagcag ttggatgcac cctgtgatcc aatggataca    180
aatgagattt agaagtgcac tttaacttaa ataaaataaa tagccttgtc taatgtatta    240
acgagaaaga catttttgtc ttatttatat gtcttatgtt ctttgaaatg aaataaatga    300
ccatatatgc ttgtaaagag agggagtcta cagatatttg gagaaaggac attctgggtg    360
gagagaacac agagtgcaaa ggacttgaag ttaaagtgca tccagcagtc tcaaggaaga    420
gtaaaggggt ccatgtgagt ggaataaggt gag                                 453
```

<210> 7

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> sequencing oligonucleotide PrimerPU

<400> 7

```
tgtaaaacga cggccagt                                                   18
```

<210> 8

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> sequencing oligonucleotide PrimerRP

<400> 8

```
caggaaacag ctatgacc                                                   18
```

1