

(19)대한민국특허청(KR)  
(12) 공개특허공보(A)

(51) Int. Cl. (11) 공개번호 10-2006-0116825  
C12Q 1/68 (2006.01) (43) 공개일자 2006년11월15일

(21) 출원번호 10-2006-7010053  
(22) 출원일자 2006년05월23일  
번역문 제출일자 2006년05월23일  
(86) 국제출원번호 PCT/US2004/035284 (87) 국제공개번호 WO 2005/040428  
국제출원일자 2004년10월22일 국제공개일자 2005년05월06일

(30) 우선권주장 60/513,888 2003년10월23일 미국(US)  
60/542,373 2004년02월06일 미국(US)  
60/554,758 2004년03월19일 미국(US)  
60/560,524 2004년04월08일 미국(US)  
60/578,323 2004년06월09일 미국(US)  
60/605,243 2004년08월26일 미국(US)

(71) 출원인 일루미젠 바이오사이언시스, 인코포레이티드  
미국 98119 워싱턴주 시애틀 스위트 500 엘리엇 애비뉴 웨스트 201

(72) 발명자 이아도나토, 손, 피.  
미국 98122 워싱턴주 시애틀 31번 애비뉴 421  
매그네스, 찰스, 엘.  
미국 98115 워싱턴주 시애틀 피.오.박스 51182  
로센버그, 게리  
미국 98021 워싱턴주 보텔 23번 애비뉴 에스이 23519  
쉐러, 크리스티나, 에이.  
미국 98119 워싱턴주 시애틀 12번 애비뉴 더블유 1837  
길로도, 티에리  
프랑스 에프-35000 르네 뤼 모테 바탈 40

(74) 대리인 장수길  
김영

심사청구 : 없음

(54) 바이러스 감염에 대한 저항성과 관련된 유전자인 O A S 1에서의 돌연변이의 검출

요약

본 발명은 OAS1을 코딩하는 유전자에 관한 돌연변이를 검출하는 방법에 관한 것이다. 이것과 다른 개시된 돌연변이는 C형 간염을 비롯한 바이러스 감염에 대한 인간의 저항성과 관련되어 있다. 또한, 돌연변이된 유전자에 의해 코딩되는 단백질 또는 폴리펩티드, 또는 단백질 또는 폴리펩티드를 코딩하는 폴리뉴클레오티드로 이루어진 치료제를 제공한다. 안티센스 올리고뉴클레오티드를 비롯한 인간 OAS1의 억제제, 방법, 및 인간 OAS1에 특이적인 조성물이 또한 제공된다.

## 색인어

OAS1, 폴리뉴클레오티드, 안티센스 올리고뉴클레오티드

## 명세서

### 기술분야

본 발명은, C형 간염 바이러스에 의한 감염을 비롯한 플라비바이러스 감염에 대한 저항성을 부여하며 전립선암 및 당뇨병을 비롯한 다른 질환 상태와 관련되는 돌연변이인, 인간 올리고아데닐레이트 신테타제 유전자에서의 돌연변이를 검출하는 방법, 및 코딩된 단백질 및 그에 대한 항체의 용도에 관한 것이다.

### 배경기술

최근까지, 감염에 대한 자연적인 저항성이 인간 집단 내에 존재하는 다수의 질환들이 확인되어 왔다. 문헌 [Alter and Moyer, *J. Acquir. Immune Defic. Syndr. Hum Retrovirol.* 18 Suppl. 1:S6-10 (1998)]에서는 주사용 약물 사용자와 같은 높은 위험군에서 90%의 C형 간염 바이러스 감염 (HCV) 비율이 보고되었다. 그러나, 상기 문헌에서는 나머지 10%가 감염에 대해 분명하게 저항성을 나타내는 메커니즘이 확인되지 않았다. HCV 감염에서 역할을 하는 단백질로는 2-프라임, 5-프라임 올리고아데닐레이트 신테타제를 들 수 있다. OAS는 아데노신의 2-프라임, 5-프라임 올리고머 (2-5A)의 합성을 촉매하는 능력을 특징으로 하는 인터페론-유래의 단백질이다. 문헌 [Hovanessian et al., *EMBO* 6: 1273-1280 (1987)]에서는 인터페론-처리된 인간 세포가 40 (OAS1), 46 (OAS1), 69 및 100 kD의 단백질들에 상응하는 여러 OAS를 함유하는 것으로 밝혀졌다. 문헌 [Marie et al., *Biochem. Biophys. Res. Commun.* 160:580-587 (1989)]에서는 p69에 대해 고도로 특이적인 폴리클로날 항체인 69-kD OAS를 생성시켰다. 문헌 [Marie and Hovanessian, *J. Biol. Chem.* 267: 9933-9939 (1992)]에서는 항-p69 항체로 인터페론-처리된 인간 세포 발현 라이브러리를 스크리닝함으로써 부분 OAS2 cDNA를 단리하였다. 이들은 부분 cDNA로 추가의 라이브러리를 스크리닝하였으며, 두 가지 OAS2 이소형을 코딩하는 cDNA를 회수하였다. 더 작은 이소형은 3-프라임 비번역 영역의 길이가 다른 두 mRNA에 의해 코딩된다.

노던 블롯 분석에 의해, OAS2가 인간 세포에서 네 가지 인터페론-유도된 mRNA로서 발현되는 것으로 밝혀졌다. 예상된 OAS2 단백질은 공통의 683-아미노산 서열 및 상이한 3-프라임 말단을 갖는다. 시험관내 전사/번역 생성물의 SDS-PAGE에 따르면, 두 이소형은 69 및 71 kD의 분자 질량을 갖는다. 두 이소형은 시험관내에서 OAS 활성을 나타냈다. 서열 분석 결과, OAS2는 프롤린-풍부한 추정된 링커 영역에 의해 분리된 두 OAS1-상동성 도메인을 함유하는 것으로 나타났다. N- 및 C-말단 도메인은 OAS1과의 동일성이 각각 41% 및 53%이다.

문헌 [Hovian et al., *Genomics* 52: 267-277 (1998)]에서는, 형광 계내 혼성화 및 맵핑된 클론 내의 혼입에 의해, OAS1, OAS2 및 OAS3 유전자가 12q24.2 상의 130-kb 영역과 클러스터링되는 것으로 결정되었다. 2-5A는 RNase I과 결합하여 그를 활성화함으로써 바이러스 및 세포 RNA를 분해하여, 세포 단백질 합성을 억제하고 바이러스 복제를 손상시킨다.

OASL로서 언급되는 네 번째 인간 OAS 유전자는 OASL이 효소 활성이 결여되어 있다는 점에서 OAS1, OAS2 및 OAS3과 상이하다. OASL 유전자는 유비퀴틴의 직렬 반복부에 대해 상동성인 164 아미노산 C-말단 도메인에 융합된 OAS 단위로 구성된 2-도메인 단백질을 코딩한다 (문헌 [Eskildsen et al., *Nuc. Acids Res.* 31:3166-3173, 2003]; [Kakuta et al., *J. Interferon & Cytokine Res.* 22:981-993, 2002]).

당업계에서는 바이러스 복제 및 바이러스 감염을 억제하는데 있어서의 역할로 인해 HCV 복제를 억제하는 억제제-기초의 요법이 상당히 요구되고 있으며 OAS1 활성과 관련된 바이러스 복제를 억제하는 방법 및 조성물이 필요한 실정이다.

### 발명의 개요

본 발명은 올리고아데닐레이트 신테타제 1 유전자에서의 점 돌연변이를 특징으로 하는 C형 간염 저항성-관련 돌연변이를 검출하는 방법에 관한 것이다. 한 실시양태에서, 인간 유전자 스크리닝 방법이 고려된다. 이 방법은, 진뱅크(Genbank) 서열 등록 번호 NT\_009775.13 (그의 연속 뉴클레오티드 2,130,000 내지 2,157,999를 서열 19로서 도 2에 나타냄)에 관한 올리고아데닐레이트 신테타제 1 유전자 (OAS1)에 대하여 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072,

2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523 또는 2156638에서의 염기 치환 또는 뉴클레오티드 위치 2156595에서의 염기 결실을 특징으로 하는 올리고아데닐레이트 신테타제 1 유전자 점 돌연변이의 존재에 대하여 인간으로부터 분리된 핵산 샘플을 분석하는 것을 포함한다.

바람직한 실시양태에서, 상기 방법은 증폭 조건하에 올리고아데닐레이트 신테타제 1 유전자 NT\_009775.13의 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638을 함유하는 인간 게놈 DNA의 영역을 증폭하기 위한 폴리머라제 연쇄 반응 (PCR) 프라이머 쌍으로 인간으로부터의 게놈 DNA 샘플을 처리하는 것을 포함한다. PCR 처리에 의해 상기 영역을 함유하는 증폭 생성물이 제조되었으며, 이어서 얻어진 생성물을 점 돌연변이의 존재에 대하여 분석하였다.

추가 실시양태에서, 본 발명은 NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이를 갖는 유전자에 의해 코딩된 단백질, 및 바이러스 감염, 바람직하게는 플라비바이러스 감염, 가장 바람직하게는 C형 간염 감염에 대한 저항성 진단제를 제조하는데 있어서 상기 단백질의 용도를 제공한다. 특정 실시양태에서, 진단제는 항체이다.

또 다른 실시양태에서, 본 발명은 NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이를 갖는 OAS1 유전자에 의해 코딩된 단백질인, 바이러스, 바람직하게는 플라비바이러스, 가장 바람직하게는 C형 간염 바이러스의 예방 또는 억제를 위한 치료 화합물을 제공한다. 다른 실시양태에서, 치료 화합물은 단백질을 코딩하는 폴리뉴클레오티드, 예를 들어 DNA 또는 RNA이다.

또 다른 실시양태에서, 본 발명은 서열 20, 서열 21, 서열 22, 서열 23, 서열 24, 서열 25, 서열 26, 서열 27, 서열 28, 서열 29, 서열 32, 서열 33, 서열 34, 서열 35, 서열 46, 서열 47 및(또는) 서열 48의 단백질인, 바이러스, 바람직하게는 플라비바이러스, 가장 바람직하게는 간염 C 바이러스에 의한 감염의 예방 또는 억제를 위한 치료 화합물을 제공한다.

또 다른 실시양태에서, 본 발명은 NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이의 유리한 효과를 모방하는, 바이러스, 바람직하게는 플라비바이러스, 가장 바람직하게는 간염 C 바이러스에 의한 감염의 예방 또는 억제를 위한 치료 화합물을 제공한다. 치료 화합물은 소분자, 단백질, 펩티드, DNA 또는 RNA 분자, 또는 항체일 수 있다.

또 다른 실시양태에서, 본 발명은 NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이를 갖는 OAS1 유전자에 의해 코딩된 단백질인, 암, 바람직하게는 전립선암의 예방 또는 치료를 위한 치료 화합물을 제공한다. 다른 실시양태에서, 치료 화합물은 단백질을 코딩하는 폴리뉴클레오티드, 예를 들어 DNA 또는 RNA이다.

또 다른 실시양태에서, 본 발명은 서열 20, 서열 21, 서열 22, 서열 23, 서열 24, 서열 25, 서열 26, 서열 27, 서열 28, 서열 29, 서열 32, 서열 33, 서열 34, 서열 35, 서열 46, 서열 47 및(또는) 서열 48의 단백질인, 암, 바람직하게는 전립선암의 예방 또는 치료를 위한 치료 화합물을 제공한다.

또 다른 실시양태에서, 본 발명은 NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이의 유리한 효과를 모방하는, 암, 바람직하게는 전립선암의 예방 또는 치료를 위한 치료 화합물을 제공한다. 치료 화합물은 소분자, 단백질, 펩티드, DNA 또는 RNA 분자, 또는 항체일 수 있다.

추가 실시양태에서, 치료 화합물은 OAS1 또는 하나 이상의 하위-영역의 활성 또는 전체 단백질의 하위-기능을 억제할 수 있으며, 이러한 화합물의 대표적인 것으로는 안티센스 분자, 리보자임, 및 OAS1 폴리뉴클레오티드에 특이적으로 결합할 수 있는 RNAi 분자, 및 OAS1 단백질 및 폴리펩티드에 특이적으로 결합할 수 있는 항체 및 그의 단편이 있다.

본 발명은 다른 실시양태에서 OAS1 억제제를 제공한다. 본 발명의 억제제로는 안티센스 분자, 리보자임, RNAi, 항체 또는 항체 단편, 단백질 또는 폴리펩티드, 및 소분자를 들 수 있지만, 이에 한정되지 않는다. 예시적인 안티센스 분자는 서열 19

의 폴리뉴클레오티드의 10개, 15개 또는 20개 이상의 연속적인 뉴클레오티드를 포함하거나 또는 엄격 조건하에 상기 폴리뉴클레오티드와 혼성화한다. 서열 19의 서열의 25개 이상의 연속적인 뉴클레오티드를 포함하거나 또는 엄격 조건하에 상기 서열과 혼성화하는 안티센스 분자가 더 바람직하다.

또 다른 실시양태에서, 서열 30의 폴리펩티드에 의해 형성된 단백질의 영역에 특이적으로 결합하는 OAS1 억제제가 고려된다. 본 발명의 억제제로는 항체, 항체 단편, 소분자, 단백질 또는 폴리펩티드를 들 수 있지만, 이에 한정되지 않는다.

또 다른 실시양태에서, 서열 31의 폴리뉴클레오티드에 특이적으로 결합하거나 또는 그와 혼성화하는 안티센스 또는 RNAi 분자로 이루어진 OAS1 억제제가 고려된다.

추가 실시양태에서, 제약상 허용가능한 담체 중 하나 이상의 OAS1 억제제를 포함하는 조성물이 제공된다.

추가 실시양태는 OAS1 유전자 발현 또는 생물학적 활성을 감소시키는 방법을 제공한다.

추가 실시양태는 NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이를 갖는 여러 형태의 OAS1 유전자의 발현을 특이적으로 증가 또는 감소시키는 방법을 제공한다.

본 발명은 하나 이상의 변형된 뉴클레오티드간 결합을 포함하는 안티센스 올리고뉴클레오티드를 제공한다.

본 발명은 포스포로티오에이트 결합을 갖는 안티센스 올리고뉴클레오티드를 추가로 제공한다.

또한, 본 발명은 하나 이상의 변형된 당 잔기를 포함하는 안티센스 올리고뉴클레오티드를 제공한다.

본 발명은 또한 2'-O-메틸 당 잔기인 하나 이상의 변형된 당 잔기를 포함하는 안티센스 올리고뉴클레오티드를 제공한다.

추가로, 본 발명은 하나 이상의 변형된 뉴클레오티드를 포함하는 안티센스 올리고뉴클레오티드를 제공한다.

또한, 본 발명은 5-메틸시토신인 변형된 뉴클레오티드를 갖는 안티센스 올리고뉴클레오티드를 제공한다.

또한, 본 발명은 키메라 올리고뉴클레오티드인 안티센스 화합물을 제공한다.

본 발명은 생체내에서 인간 세포 또는 조직을, 인간 OAS1을 코딩하는 핵산 분자로 표적화된 뉴클레오티드 8개 내지 35개 길이의 리보자임 또는 안티센스 화합물과 접촉시켜 인간 OAS1의 발현을 억제하는 것을 포함하는, 인간 세포 또는 조직에서 인간 OAS1의 발현을 억제하는 방법을 제공한다.

추가로, 본 발명은 생체내에서 안티센스 또는 RNAi 화합물 또는 리보자임을 사용해서, NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 하나 이상의 돌연변이를 갖는 것으로 정의되는 특정 형태의 OAS1의 발현을 감소 또는 증가시키는 방법을 제공한다.

또한, 본 발명은 생체내에서 암 세포를, 인간 OAS1을 코딩하는 핵산 분자로 표적화된 뉴클레오티드 8개 내지 35개 길이의 리보자임 또는 안티센스 화합물과 접촉시켜 인간 OAS1의 발현을 억제하는 것을 포함하는, 암 세포의 성장을 조절하는 방법을 제공한다.

또한, 본 발명은 OAS1 폴리뉴클레오티드의 표적 영역을 확인하는 방법을 제공한다. 본 발명은 또한 계내 혼성화에 의해 OAS1 폴리뉴클레오티드를 확인하기 위한 표지된 프로브를 제공한다.

본 발명은 HCV 감염을 예방 또는 억제하기 위한 약물을 제조하는데 있어서 본 발명에 따른 OAS1 억제제의 용도를 제공한다.

추가로, 본 발명은 OAS1 억제제를 OAS1 단백질의 특정 영역 또는 상기 단백질의 특정 기능으로 지정하는 방법을 제공한다.

본 발명은 또한 본 발명에 따른 안티센스 올리고뉴클레오티드를 생리학적으로 허용가능한 담체 또는 희석제와의 혼합물로서 포함하는, OAS1의 발현을 억제하기 위한 제약 조성물을 제공한다.

또한, 본 발명은 OAS1 RNA를 특이적으로 절단할 수 있는 리보자임 및 상기 리보자임을 포함하는 제약 조성물을 제공한다.

또한, 본 발명은 OAS1의 활성을 감소시킬 수 있거나 또는 OAS1 mRNA의 발현을 감소 또는 방지할 수 있는 억제제인 OAS1의 소분자 억제제를 제공한다.

추가로, 본 발명은 C형 간염 바이러스 NS5A 단백질과 같은 다른 단백질과의 상호작용을 비롯하여 2'-5' 올리고아데닐레이트의 합성 이외의 특정 단백질 기능을 변경하는 OAS1 억제제를 제공한다.

본 발명은 또한 글리코실화 및 인산화를 포함하지만 이에 한정되지 않는 OAS1의 번역 후 변형을 변화시키는 화합물을 제공한다.

또한, 본 발명은 (a) 증폭 조건하에 올리고아데닐레이트 신테타제 유전자의 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638을 함유하는 인간 게놈 DNA의 영역을 증폭하기 위한 폴리머라제 연쇄 반응 (PCR) 프라이머 쌍으로 인간으로부터의 게놈 DNA 샘플을 처리하여 상기 영역을 함유하는 증폭 생성물을 제조하는 단계; 및 (b) 단계 (a)의 증폭 생성물 중에서 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 뉴클레오티드 점 돌연변이의 존재를 검출함으로써 상기 돌연변이를 확인하는 단계를 포함하는, 올리고아데닐레이트 신테타제 유전자 돌연변이를 확인하기 위한 인간 유전자 스크리닝 방법을 제공한다.

상기 방법의 특정 실시양태에서, 상기 영역은 서열 1 내지 7 및 서열 57 내지 64로 이루어진 군으로부터 선택된 서열로 나타낸 뉴클레오티드 서열을 포함한다. 다른 실시양태에서, 상기 영역은 본질적으로는 서열 1 내지 7 및 서열 57 내지 64로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 구성된다. 또한, 혼성화 조건하에 점 돌연변이에 대해 특이적인 올리고뉴클레오티드 프로브로 상기 단계 (a)의 증폭 생성물을 처리하는 단계, 및 혼성화 생성물의 형성을 검출하는 단계를 포함하는 방법이 또한 제공된다. 이러한 방법의 특정 실시양태에서, 올리고뉴클레오티드 프로브는 서열 12 내지 18로 이루어진 군으로부터 선택된 뉴클레오티드 서열을 포함한다.

또한, 본 발명은 (a) PCR 완충액 중에서 인간으로부터의 게놈 DNA 샘플을 서열 8 및 9, 및 서열 10 및 11로 이루어진 군으로부터 선택된 올리고아데닐레이트 신테타제 유전자-특이적 PCR 프라이머 쌍과 배합하여 폴리머라제 연쇄 반응 (PCR) 혼합물을 형성하는 단계; (b) 상기 PCR 혼합물을 다수회의 PCR 열주기로 처리하여 올리고아데닐레이트 신테타제 유전자 증폭 생성물을 제조하는 단계; 및 (c) 혼성화 조건하에 단계 (b)에서 제조된 생성물을 서열 12 내지 18로 이루어진 군으로부터 선택된 프로브로 처리하여 상기 돌연변이를 검출하는 단계를 포함하는, 올리고아데닐레이트 신테타제 유전자 (OAS1)의 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서 야생형 뉴클레오티드를 비-야생형 뉴클레오티드로 치환한 것을 포함하는 점 돌연변이를 함유하는 C형 간염 감염 저항성 질환 대립유전자를 인간에서 검출하는 방법에 관한 것이다.

또한, 안티센스 올리고뉴클레오티드, 리보자임, 작은 억제성 RNA (RNAi), 단백질, 폴리펩티드, 항체 및 소분자로 이루어진 군으로부터 선택되는 단리된 OAS1 억제제를 제공한다. 단리된 억제제는 서열 19의 서열의 15개 이상의 연속적인 핵산을 포함하는 안티센스 분자 또는 그의 상보체일 수 있다. 다른 실시양태에서, 단리된 OAS1 억제제 (안티센스 분자 또는 그의 상보체)는 고 엄격 조건하에 서열 19의 서열과 혼성화한다.

단리된 OAS1 억제제는 항체 및 항체 단편으로 이루어진 군으로부터 선택될 수 있다. 또한, 제약상 허용가능한 담체 중에 치료 유효량의 1종 이상의 OAS1 억제제를 포함하는 조성물이 제공된다.

또한, 본 발명은 안티센스 올리고뉴클레오티드, 리보자임, 단백질, RNAi, 폴리펩티드, 항체 및 소분자로 이루어진 군으로부터 선택된 OAS1 억제제를 포유동물 세포에게 투여하는 것을 포함하는, 포유동물 세포에서 OAS1의 발현을 억제하는 방법에 관한 것이다.

추가로, 본 발명은 OAS1 유전자 유래의 선택된 표적 핵산 서열의 전부 또는 일부와 특이적으로 혼성화하는데 효과적인 소정량의 안티센스 올리고뉴클레오티드를 제약상 효과적인 비히클 중에 넣어 대상에게 투여하는 것을 포함하는, 대상에서 OAS1 유전자의 발현을 억제하는 방법에 관한 것이다.

또한, 본 발명은 안티센스 올리고뉴클레오티드, 리보자임, RNAi, 단백질, 폴리펩티드, 항체 및 소분자로 이루어진 군으로부터 선택되며 플라비바이러스에 의한 감염을 예방하는 OAS1 억제제를 감염에 민감한 인간 대상에게 투여하는 것을 포함하는, 상기 인간 대상에서 플라비바이러스에 의한 감염을 예방하는 방법에 관한 것이다.

또한, 본 발명은 안티센스 올리고뉴클레오티드, 리보자임, RNAi, 단백질, 폴리펩티드, 항체 및 소분자로 이루어진 군으로부터 선택되고, 플라비바이러스 또는 기타 바이러스에 의한 감염을 예방하며, NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 돌연변이에 의해 형성된 하나 이상의 특정 형태의 단백질로 지정되는 OAS1 억제제를 플라비바이러스 또는 다른 바이러스에 의한 감염에 민감한 인간 대상에게 투여하는 것을 포함하는, 상기 인간 대상에서 상기 감염을 예방 또는 치료하는 방법에 관한 것이다.

또한, 본 발명은 서열 20, 서열 21, 서열 22, 서열 23, 서열 24, 서열 25, 서열 26, 서열 27, 서열 28, 서열 29, 서열 32, 서열 33, 서열 34, 서열 35, 서열 46, 서열 47 및(또는) 서열 48의 폴리펩티드들 중 어느 하나를 투여함으로써 플라비바이러스 또는 임의의 다른 바이러스에 의한 감염에 민감한 인간 대상에서 상기 감염을 예방 또는 치료하는 방법에 관한 것이다.

본 발명은 또한 인간 면역결핍 바이러스 (HIV)에 의한 감염의 치료를 포함한다.

또한, 본 발명은 안티센스 올리고뉴클레오티드, 리보자임, RNAi, 단백질, 폴리펩티드, 항체 및 소분자로 이루어진 군으로부터 선택되며 인슐린 의존성 당뇨병 (IDDM)을 예방하는 OAS1 억제제를 인간 대상에게 투여하는 것을 포함하는, 인간 대상에서 IDDM을 예방하는 방법에 관한 것이다.

본 발명은 또한 안티센스 올리고뉴클레오티드, 리보자임, RNAi, 단백질, 폴리펩티드, 항체 및 소분자로 이루어진 군으로부터 선택되고, IDDM을 예방하며, NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서의 돌연변이에 의해 형성된 하나 이상의 특정 형태의 단백질로 지정되는 OAS1 억제제를 인간 대상에게 투여하는 것을 포함하는, 인간 대상에서 IDDM을 예방하는 방법에 관한 것이다.

또한, 본 발명은 OAS1 유전자의 발현 증가 또는 본원에 개시된 폴리펩티드의 치료 투여에 의해 암, 예를 들어 전립선암을 치료하는 방법에 관한 것이다.

또한, OAS1 표적 유전자의 발현을 억제하는데 충분한 양으로 리보핵산 (RNA)을 세포에 도입하는 것을 포함하는, 시험관 내 세포에서 OAS1 표적 유전자의 발현을 억제하는 방법이 제공되는데, 여기서 RNA는 OAS1 표적 유전자의 뉴클레오티드 서열에 상응하는 리보뉴클레오티드 서열로 본질적으로 구성된 제1 가닥 및 OAS1 표적 유전자의 뉴클레오티드 서열에 대해 상보적인 리보뉴클레오티드 서열로 본질적으로 구성된 제2 가닥을 갖는 이중-가닥 분자이고, 제1 및 제2 리보뉴클레오티드 가닥은 별개의 상보적 가닥으로 서로 혼성화하여 상기 이중-가닥 분자를 형성하며, 이중-가닥 분자는 표적 유전자의 발현을 억제한다.

상기 방법의 특정 실시양태에서, 제1 리보뉴클레오티드 서열은 OAS1 표적 유전자에 상응하는 20개 이상의 염기를 포함하며, 제2 리보뉴클레오티드 서열은 OAS1 표적 유전자의 뉴클레오티드 서열에 상보적인 20개 이상의 염기를 포함한다. 또 다른 실시양태에서, 표적 유전자의 발현은 10% 이상 억제된다.

상기 방법의 또 다른 실시양태에서, 이중-가닥 리보핵산 구조는 염기 20개 이상의 길이이며, 리보핵산 가닥의 각각은 OAS1 표적 유전자의 데옥시리보핵산 가닥과 20개 이상의 염기에 대하여 특이적으로 혼성화할 수 있다.

본 발명은 유전자 돌연변이로 인해 감소 또는 상실될 수 있는 OAS1의 기능을 복구할 수 있는 폴리펩티드 또는 단백질을 제공한다. 일부 실시양태에서, 폴리펩티드 또는 단백질은 서열 19를 포함하는 유전자에 의해 코딩된 야생형 OAS1의 아미노산 서열을 갖는다. 다른 실시양태에서, OAS1 유전자 내의 돌연변이가 OAS1 단백질에 대하여 증가된 활성, 안정성 및(또는) 반감기, 또는 항-바이러스 활성에 대하여 OAS1 단백질을 보다 적합하게 만드는 다른 변화를 부여하는 경우, 돌연변이된 OAS1 유전자에 의해 코딩된 단백질 또는 폴리펩티드가 바람직하다.

상기 임의의 단백질 및 폴리펩티드는 치료 조성물의 성분으로서 제공될 수 있다.

또한, 치료 조성물의 성분으로서 서열 20, 서열 21, 서열 22, 서열 23, 서열 24, 서열 25, 서열 26, 서열 27, 서열 28, 서열 29, 서열 32, 서열 33, 서열 34, 서열 35, 서열 46, 서열 47 및(또는) 서열 48로 구성된 단백질들 중 어느 한 단백질의 용도가 제공된다.

추가 실시양태에서, OAS1 단백질, OAS1 돌연변이 단백질 또는 OAS1 폴리펩티드를 코딩하는 핵산은 유전자 요법의 형태로 투여될 수 있다.

### 도면의 간단한 설명

도 1은 OAS1 유전자에서의 본 발명의 돌연변이의 위치, 대립유전자 변이체 (염기 치환), 게놈 서열 상의 돌연변이의 좌표, 및 NCBI dbSNP ID (존재하는 경우)를 나타내는 표이다.

도 2 (서열 19)는 NCBI 등록 번호 NT\_009775.13, OAS1의 위치 2,130,000 내지 2,157,999에서의 연속적인 뉴클레오티드 염기로 구성된 폴리뉴클레오티드 서열이다.

도 3은 서열 20 내지 64를 나타낸다.

도 4는 백인(Caucasian) 집단에서 OAS1 유전자 내의 대략적인 반수형(haplotype) 분포를 나타낸다.

도 5A 및 5B는 본 발명에 개시된 다양한 전사체 변이체, 및 본 발명에 개시된 돌연변이와 상기 변이체의 관계를 예시한다. 도 5A는 전사체 변이체 (TV) 형태 1 내지 6을 나타내며, 도 5B는 TV 형태 7 내지 10을 나타낸다. 또한, 도 5C는 침팬지 및 고릴라 변이체의 추정 구조를 인간과 비교하여 예시한다.

도 6은 OAS1 유전자에서의 비-인간 영장류 돌연변이와 인간과의 비교, 및 상응하는 인간 게놈 서열 상의 돌연변이의 좌표를 기재한 표이다.

도 7은 이. 콜라이로부터 발현 및 회수되는 경우에 본 발명의 예시적인 폴리펩티드의 효소 활성을 입증한다.

도 8은 헬라 (HeLa) 세포에서 발현되는 경우에 본 발명의 폴리뉴클레오티드로부터 생성된 예시적인 폴리펩티드의 효소 활성을 입증한다.

도 9는 본 발명의 예시적인 폴리펩티드에 대한 항체의 전개를 입증한다.

도 10은 예시적인 단백질 형질도입 도메인 폴리펩티드를 상술하는 표이다.

### 발명의 상세한 설명

#### 도입 및 정의

본 발명은 올리고아데닐레이트 신테타제 유전자에 있어서의 신규한 돌연변이, 바이러스 감염에 대한 감수성 또는 저항성의 진단을 위한 상기 돌연변이의 용도, 본 발명에 따른 돌연변이를 갖는 유전자에 의해 코딩되는 단백질, 및 단백질, 항체 및 관련된 핵산을 사용한 바이러스 감염의 예방 또는 억제에 관한 것이다. 상기 돌연변이는 플라비바이러스, 특히 C형 간염 바이러스로의 감염에 대한 담체의 저항성과 관련된다.

많은 최근의 의학 연구는 질환을 유발하거나 질환에 기여하는 돌연변이 및 결함을 확인하는데 집중되어 있다. 이러한 연구는 질환 상태에 목표화된 치료 화합물 및 방법을 초래하도록 설계된다. 감염성 작용제 및 다른 위험 요인에 대한 노출에도 불구하고 사람을 건강하게 유지하도록 하는 유전적 영향을 연구하는 데는 주의가 덜 기울여졌다. 본 발명은 질환에 대한 저항성을 부여하는 유전적 변이 또는 돌연변이를 발견하기 위해 인간 대상체의 특정 집단을 확인하고 평가한 본 발명자들에게 의해 개발된 방법의 성공적인 적용을 제시한다. 특정 질환 또는 생물학적 상태에 대한 자연적 저항성을 갖는 하위-집단 부분의 확인은 또한 제약학적 중재, 진단적 평가, 또는 예방적 백신화와 같은 예방에 대한 적합한 표적인 유전자 및 단백질의 확인을 가능하게 한다.

본원에서 확인된 하위-집단 부분은, 코호트(cohort)는 감염된 반면 (혈청반응 양성), C형 간염 바이러스 (HCV)에의 반복된 노출에도 불구하고 혈청반응 음성으로 남아있는 개체를 포함한다. 연구된 집단은 반복된 혈액 수혈을 받는 혈우병 환자 및 공유된 주사기 바늘 및 다른 위험 요인을 통해 노출된 정맥 약물 사용자를 포함한다.

HCV 감염은, 질환을 유발하지만 감염된 세포내 바이러스의 저 안정성 상태로 발달되어 명백히 HCV가 숙주 면역-감시 시스템으로부터 벗어나게 하면서 바이러스 감염에 저항성이게 하는 감염의 수준을 달성하는데 함께 작용하는 단백질과 면역계 성분의 복합 세트와 관련된다. (문헌 [Dansako et al., Virus Research 97:17-30, 2003]). 본 발명은 상기 시스템의 한 성분인 인터페론-유도성 2'-5'-올리고아데닐레이트 신테타제 유전자, 특히 OAS1에 초점을 맞춘다. OAS1 유전자는 RN아제 L (RNase L)을 활성화시켜 바이러스 RNA를 절단함으로써 인간에서 숙주 세포의 항바이러스 활성화에 주요한 역할을 한다. HCV RNA는 2'-5'-OAS/RN아제 L 경로를 활성화시킨다. 문헌 [Dansako et al.]에 언급된 바와 같이, HCV RNA가 바이러스 RNA의 절단을 초래하는 경로를 활성화시키는 것은 모순되는 것으로 보일 수 있다. 그러나, 이러한 활성화는 숙주 면역 방어와 숙주를 죽이는 감염의 수준 사이의 균형을 유지하는데에 기능할 수 있다.

OAS1 유전자의 이러한 복잡한 역할의 관점에서, 본 발명이 OAS1 유전자와, 상기 돌연변이의 담체에서의 HCV 감염에 대한 저항성 사이의 강한 상관관계를 확인했다는 것은 상당히 흥미롭다. 이러한 개체의 존재는 이제 감염 수준의 바이러스에 대한 반복된 노출에도 불구하고, 어떻게 OAS1이 HCV 감염에 대한 저항성에 기여하는지를 설명해 준다. 이제, 이러한 정보는 자연적 저항성이 없는 개체에서의 저항성 메카니즘의 복제 방법 및 복제용 조성물의 개발을 유도할 것이다.

따라서, 본 발명은 메카니즘에 무관하게, HCV 감염에 대해 저항성인 개체의 확인에 유용한 본원에서 확인된 돌연변이를 제공한다. 저항성은 OAS1 단백질 기능의 손실을 통해 일어날 수 있으며, 이 경우, HCV 바이러스 수준은 바이러스가 숙주 면역-감시 시스템으로부터 벗어나는 것을 방지하기에 충분히 높을 것이며, 따라서 바이러스 파괴를 용이하게 할 것이라고 예측된다. 저항성은 또한 OAS1 단백질 수준이 증진되고(거나), 단백질의 반감기가 증가되고(거나), RN아제 L을 활성화시켜 바이러스 RNA를 절단하는 그의 능력을 향상시키는 방식으로 단백질 구조가 영향을 받는다는 기능 획득을 통해 일어날 수 있다. 저항성은 또한 HCV 바이러스 단백질 또는 뉴클레오티드에 의한 정상적인 OAS1 단백질 기능의 억제를 방지하는 OAS1 단백질에 대한 변형을 통해 일어날 수 있다. 저항성은 또한 정상적인 HCV 바이러스 수명에 필요한 HCV 바이러스 단백질 또는 뉴클레오티드와 단백질의 상호작용을 방지하는 OAS1 단백질에 대한 변형을 통해 일어날 수 있다. 본 발명은 하나의 메카니즘에 제한되지 않는다. 또한, 몇몇 여러 가지 점 돌연변이가 본원에 개시되어 있지만, 이는 이러한 돌연변이가 OAS1 단백질 구조 또는 기능에 동일한 영향을 가짐을 나타내는 것을 지시하려는 의도는 아니다.

OAS1은 HCV를 구성원으로 하는 플라비바이러스 과의 다른 바이러스에 의한 감염에 역할을 담당한다. 플라비바이러스 과는 또한 본원에 개시된 황열, 뎅그열, 세인트 루이스(St. Louis) 뇌염, 일본 뇌염 및 기타 바이러스성 질환을 포함한다. 상기 바이러스에 대한 숙주 방어는 바이러스-유도성 인터페론을 포함한다. 인터페론은 상기 논의된 바와 같이 RN아제 L의 활성화에 관여하는 2'-5'-올리고아데닐레이트 신테타제를 포함한다. 다음에는 RN아제 L이 바이러스 RNA를 절단한다. RSV를 비롯한 기타 바이러스 감염은 본원에 개시된 방법에 의한 예방 및(또는) 억제에 따를 수 있다.

상세한 설명 및 바람직한 실시예에 대한 참고로, 하기 정의가 사용된다.

A: 아데닌; C: 시토신; G: 구아닌; T: 티민 (DNA에 있어서) 및 U: 우라실 (RNA에 있어서)

대립유전자: 특정 유전자의 DNA 서열 변이체. 이배체 세포에서, 최대 2개의 대립유전자는 각각 염색체 세트의 상동 염색체 상의 동일한 상대적 위치 또는 좌위에 존재할 것이다. 임의의 하나의 좌위에 있는 대립유전자가 동일할 경우 개체는 상기 좌위에 대해 동종접합성인 것으로 지칭되며, 이들이 상이할 경우 개체는 상기 좌위에 대해 이종접합성인 것으로 지칭된다. 임의의 하나의 유전자의 상이한 대립유전자가 단일 염기에 의해서만 다양할 수도 있기 때문에, 임의의 하나의 유전자에 대한 대립유전자의 가능한 수는 매우 많다. 대립유전자가 상이할 경우, 하나는 종종 열성이라고 지칭되는 다른 하나에 대해 우성이다. 우성은 표현형의 특성이며, 우성에 의해 열성 대립유전자가 불활성화된다는 것을 의미하지는 않는다. 많은 예에서, 정상적으로 기능화된 (야생형) 대립유전자는 다소 결함있는 기능의 모든 돌연변이 대립유전자에 대해 우성이다. 이러한 경우, 일반적 설명은 2개 중에서 하나의 기능성 대립유전자는 유기체의 정상적인 발달을 지지하는데 충분한 활성 유전자 산물을 생성하는데 충분하다는 것이다 (즉, 유전자 산물의 양에서 통상적으로 2배의 안전성 여유분(margin)이 있음).

반수형: 염색체 국부화에 의해 일련으로 정돈되고 염색체 세트의 하나의 특정 상동 염색체가 보유한 대립유전자의 세트를 나타내는 많은 가능한 다수의 대립유전자 중 하나.

뉴클레오타이드: 당 잔기 (펜토스), 포스페이트, 및 질소성 헤테로시클릭 염기로 이루어지는 DNA 또는 RNA의 단량체 단위. 염기는 글리코시드 탄소 (펜토스의 1' 탄소)를 통해 당 잔기에 연결되고, 염기와 당의 상기 조합이 뉴클레오시드이다. 뉴클레오시드가 펜토스의 3' 또는 5' 위치에 결합된 포스페이트 기를 함유할 경우, 이는 뉴클레오타이드로 지칭된다. 작동적으로 연결된 뉴클레오타이드의 서열은 본원에서 전형적으로 "염기 서열" 또는 "뉴클레오타이드 서열", 및 그의 문법적 등가물로서 지칭되며, 본원에서는 좌측에서 우측으로의 배향이 5'-말단에서 3'-말단의 통상적인 방향인 식으로 나타내어진다.

염기쌍 (bp): 이중 가닥 DNA 분자에서 아데닌 (A)과 티민 (T)의 연합, 또는 시토신 (C)과 구아닌 (G)의 연합. RNA에서는, 티민이 우라실 (U)로 치환된다. 본원에서 RNA에 대해 언급할 경우, 기호 T는 RNA 분자 중의 특정 위치에 있는 우라실을 나타내기 위해 U와 상호교환적으로 사용될 수 있다.

핵산: 단일 가닥 또는 이중 가닥인, 뉴클레오타이드의 중합체.

폴리뉴클레오타이드: 단일 또는 이중 가닥 뉴클레오타이드의 중합체. 본원에서 사용된 "폴리뉴클레오타이드" 및 그의 문법적 등가물은 핵산의 전 범위를 포함할 것이다. 폴리뉴클레오타이드는 전형적으로 2개 이상의 데옥시리보뉴클레오타이드 및(또는) 리보뉴클레오타이드의 선형 가닥을 포함하는 핵산 분자를 지칭한다. 정확한 크기는 많은 요인에 따라 다를 것이며, 다음으로 당염기에 널리 공지된 바와 같이 궁극적인 사용 조건에 따라 다를 것이다. 본 발명의 폴리뉴클레오타이드는 프라이머, 프로브, RNA/DNA 절편, 올리고뉴클레오타이드 또는 "올리고" (상대적으로 짧은 폴리뉴클레오타이드), 유전자, 벡터, 플라스미드 등을 포함한다.

RNAi: RNA 간섭 (RNAi)은 전형적으로 약 21 내지 23개 뉴클레오타이드 길이의 이중나선인 소 간섭 RNA (siRNA)가 세포 내로 도입되어, 궁극적으로 동일하거나 상보적인 서열을 함유하는 표적화된 유전자의 전령 RNA의 열화를 초래하고 이를 효과적으로 침묵화시키는 방법이다.

유전자: 그의 뉴클레오타이드 서열이 RNA 또는 폴리펩티드를 코딩하는 핵산. 유전자는 RNA 또는 DNA일 수 있다.

이중나선 DNA: 이중나선의 염기쌍에 존재하는 각각의 상보성 염기 사이의 하나 이상의 수소 결합에 의해 함께 붙들려진 실질적으로 상보적인 폴리뉴클레오타이드의 2개 가닥을 포함하는 이중-가닥 핵산 분자. 염기쌍을 형성하는 뉴클레오타이드는 리보뉴클레오타이드 염기 또는 데옥시리보뉴클레오타이드 염기일 수 있기 때문에, "이중나선 DNA"라는 어구는 2개의 DNA 가닥 (ds DNA)를 포함하는 DNA-DNA 이중나선, 또는 하나의 DNA 및 하나의 RNA 가닥을 포함하는 RNA-DNA 이중나선을 지칭한다.

상보성 염기: DNA 또는 RNA가 이중 가닥 배향을 채택할 경우 정상적으로 짝지어진 뉴클레오타이드.

상보성 뉴클레오타이드 서열: 결과로 생성된 수소 결합과 특이적으로 혼성화하는 또다른 단일 가닥 상의 DNA 또는 RNA에 충분히 상보적인 DNA 또는 RNA의 단일-가닥 분자 내의 뉴클레오타이드 서열.

보존된: 뉴클레오타이드 서열은, 그것이 기선택된 (참조) 서열의 정확한 상보체에 비-무작위적으로 혼성화될 경우, 기선택된 서열에 대해 보존된다.

혼성화: 상보성 염기쌍 사이의 수소 결합의 수립에 의해 이중나선 또는 이형이중나선을 형성하는, 실질적으로 상보적인 뉴클레오타이드 서열 (핵산의 가닥)의 짝짓기. 이는 경쟁적으로 억제될 수 있는 2개의 상보성 폴리뉴클레오타이드 사이의 특이적, 즉 비-무작위적인 상호작용이다.

뉴클레오타이드 유사체: 구조적으로 A, T, G, C, 또는 U와 구조적으로 상이하지만 핵산 분자 내의 정상적인 뉴클레오타이드를 치환하기에 충분히 유사한 퓨린 또는 피리미딘 뉴클레오타이드.

DNA 상동체: 기선택된 보존된 뉴클레오타이드 서열 및 기선택된 리간드에 결합할 수 있는 수용체를 코딩하는 서열을 갖는 핵산.

상류: DNA의 전사 방향에 대해 반대 방향으로, 따라서 비-코딩 가닥 상의 5'에서 3', 또는 mRNA 상의 3'에서 5'로 가는 것.

하류: DNA 서열을 따라, DNA의 비-코딩 가닥을 따라 3'에서 5' 방향 또는 RNA 전사를 따라 5'에서 3' 방향으로 가는, 서열 전사 또는 해독의 방향.

정지 코돈: 아미노산을 코딩하지 않지만 대신 단백질 합성의 종결을 유발하는 임의의 3개 코돈. 이들은 UAG, UAA 및 UGA이며, 넨센스 또는 종결 코돈으로도 지칭된다.

리딩 프레임: 전사에 사용되는 인접한 뉴클레오티드 트리플렛 (코돈)의 특정 서열. 리딩 프레임은 해독 개시 코돈의 위치에 따라 다르다.

인트론: 간접 서열로도 지칭되는, 최초에는 RNA로 카피되지만 최종적 RNA 전사체로부터 절단되는 DNA의 비코딩 서열.

### 본 발명의 실시 방식

본 발명은 플라비바이러스, 특히 C형 간염에 대한 저항성과 관련된 올리고아데닐레이트 신테타제 대립유전자에 대해 인간을 스크리닝하는 신규한 방법을 제공한다.

본 발명은 이러한 저항성이, 인간 OAS1 유전자를 코딩하는 진뱅크 등록 번호 NT\_009775.13 (서열 19)의 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에 있는 올리고아데닐레이트 신테타제 유전자 DNA 서열에서의 점 돌연변이 (염기 치환)와 관련된다는 발견에 기초한 것이다.

본 발명은 C형 간염 바이러스 (HCV)로의 감염에 저항성이거나 부분적으로 저항성인 대상체의 집단을 확인하고, 추가로 이러한 유의한 효과를 주는 유전적 돌연변이를 확인한 연구의 결과를 개시한다. HCV 감염에 대한 저항성과 상당히 관련된 2'-5'-올리고아데닐레이트 신테타제 내의 몇몇 유전적 돌연변이가 확인되었다. 사용된 연구 설계는 경우군(case group)-대조군(control group), 대립유전자 관련 분석이었다. 대상체로서 할당된 경우군은 HCV에 대해 일련적으로 명문화되거나 추정된 노출을 가졌지만, 바이러스에 대한 항체의 발달에 의한 명문화된 바와 같은 감염을 발달시키지는 않았다 (즉, HCV 혈청반응 음성). 대조군 대상체는 일련적으로 노출되어 HCV 양성으로 혈청전환된 대상체였다. 경우군 및 대조군 대상체는 3개의 집단, 즉 캐나다 브리티시 콜롬비아 밴쿠버로부터의 혈우병 환자, 프랑스 노스웨스턴으로부터의 혈우병 환자, 및 시애틀 대도시 지역으로부터의 주사 약물 사용자로부터 모집되었다.

경우군 및 대조군 정의는 혈우병과 IDU 군 사이에 달랐으며, 본원에 기재된 바와 같이, 문헌 및 본 발명자들에 의해 개발된 다른 모델에서 공개된 감염 위험의 역학 모델을 기초로 하였다. 혈우병 집단에 대해, 대조군 대상체는 통상적인 진단학 실험실 시험을 사용하여 HCV에 대한 항체에 대해 혈청반응 양성인 것으로 명문화되었다. 경우군 대상체는 HCV 혈청반응 음성이고, 5% 미만의 정상적인 응고 인자를 가지며, 1987년 1월 이전에 농축된 응고 인자를 받은 것으로 명문화되었다. 대조군 주사 약물 사용자는 HCV 혈청반응 양성으로 명문화된 것으로 정의되었다. 경우군 주사 약물 사용자는 10년 이상 동안 약물이 주사되고, 하나 이상의 추가 위험 거동이 있다고 보고된 것으로 명문화된 HCV 혈청반응 음성으로 정의되었다. 추가의 위험 거동은 주사기, 요리 기구 또는 또다른 IDU를 갖는 면을 공유하는 것을 포함한다.

하나의 특정 카프카스인(Caucasian) 집단에서, 20명의 경우군과 42명의 대조군이 본 연구에 포함되었다. 경우 및 대조군 대상체의 선택은 본질적으로 미국 특허 출원 제09/707,576호에 기재된 바와 같이 위험에 영향받은 사람들 ("대조군") 및 위험에 영향받지 않은 사람들 ("경우군")의 집단군을 사용하여 수행되었다.

본 발명은 후보 유전자의 선택을 포함하는 HCV 감염에 대한 저항성과 관련된 유전자 돌연변이를 확인하는데에 접근한다. 세포 표면에 대한 바이러스 결합, 세포 내의 바이러스 전파, 인터페론 반응 및 원래의 면역계와 항바이러스 반응의 측면과 관련된 대략 50개의 후보 유전자를 탐문하였다. 후보 유전자는 폴리머라제 연쇄 반응을 사용하여 각각의 대상체의 게놈 DNA로부터 표적 서열을 증폭함으로써 경우군과 대조군에서 시퀀싱하였다. 후보 유전자로부터의 PCR 생성물은 자동화 형광-기재 DNA 시퀀싱 및 ABI3730 자동화 서열기를 사용하여 직접적으로 시퀀싱하였다.

단독으로 또는 조합으로 HCV 감염에 대한 저항성과 유의하게 관련된 ( $p < 0.05$ ) 유전적 돌연변이를 올리고아데닐레이트 신테타제 유전자 (OAS1)에서 확인하였다. 상기 돌연변이에 기여하는 염기 치환 및 결실은 도 1에 나타나 있다. OAS1 유전자의 변이체 형태 ("OAS1R")는 하나 이상의 본 발명의 돌연변이의 존재에 의해 생성된다 (서열 1 내지 7 및 서열 57 내지 64로 확인됨). OAS1 유전자의 이러한 변이체인 OAS1R 형태는 바이러스 감염에 대해 저항성을 부여하는 것으로 믿어진다.

다수의 OAS1 돌연변이를 포함하는 집단 및 대상체 모 반수형 (서열 1 내지 7 및 서열 57 내지 64)는 당업자에게 공지된 추측 최대화(Expectation Maximization) 방법에 의해 설정된 경우군-대조군 유전자형확정 데이터로부터 컴퓨터상에서 추론된다 (문헌 [Excoffier and Slatkin, Mol. Biol. Evol. 12:921-927, 1995]). 본 발명은 반수형 분석용 OAS1 돌연변이의 전 범위 뿐만 아니라 컴퓨터상 편의를 위한 OAS1 돌연변이의 하위집합을 사용하는 것 둘다를 포함한다. 경우군 및 대조군을 갖는 반수형 및 반수형 하위집합의 분리 패턴의 비교 분석은, 바이러스 저항성 또는 감수성에 대해 특정 효능을 갖는 돌연변이를 확인시켜 준다.

하나의 예시적 실시예에서, 서열 2 내지 서열 6으로 확인된 다수의 OAS1 돌연변이를 포함하는 반수형을 컴퓨터화한다. 몇몇 반수형이 상기 분석에 의해 카프카스인 경우군 및 대조군 집단에서 확인된다. 상기 반수형의 정의는 도 4에 나타나 있다. 공통된 반수형 (HAP1 및 HAP2로 확인됨)은 추론된 반수형의 대략 85%를 설명하는 것으로 확인되며, 특히 집단 내에서 반수형 동형접합체의 발생에 대해 하디-와인버그 평형에 있다. 다양한 인간 집단 및 영장류에서의 OAS1의 추가의 분석은, HAP2가 구시대 원숭이 및 인간과(hominid)의 방산에 선행하는 조상 영장류 반수형임을 지시한다. 하나의 추가의 반수형 (HAP3으로 확인됨)은 상기 특정 집단 내에서 지속적으로 HCV-저항성인 경우군과 관련된다. 따라서, HAP3 반수형을 보유하는 대상체는 실질적으로 HCV 감염 위험이 낮다. HAP3 반수형은 조상 반수형으로부터 기원하는 복잡한 일련의 재조합 및 돌연변이를 통해 일어난 것으로 보인다. 반수형 HAP3의 상당한 발생과 조합된 이러한 사건의 조합된 희귀성은, 양성 선택이 매번 다시 수행되는 바이러스 챌린지에 대한 반응으로서 집단 내에서 반수형 HAP3를 발달시키고 유지하는데 가능하게 작용함을 시사한다. 상기 실시예에서, 반수형 HAP3은 단지 단일 프리커서(pre-cursor) RNA 내의 돌연변이 서열 2에 대한 G 뉴클레오티드의 효과와 돌연변이 서열 4에 대한 A 뉴클레오티드의 효과를 함께 합한, 평가가능한 집단 빈도에서 일어나는 반수형이다.

본 발명은 상기 예시적 실시예에 제한되지 않는다. 또한 본 발명은 특정 OAS1 돌연변이의 관련성 및 유용성에 통찰을 제공하는 다른 예시적 실시예에 의해서도 제한되지 않는다. 또다른 예시적 실시예에서, 서열 2에서 A 뉴클레오티드를 G 뉴클레오티드로 치환한 것은 예견되었던 세린의 그리신으로의 아미노산 치환을 초래한다. 당업자에게 공지된 컴퓨터 예측은 세린이 인산화 부위인 반면 글리신은 인산화되지 않음을 강력하게 시사한다.

추가의 예시적 실시예에서, 돌연변이 서열 4에서 G 뉴클레오티드를 A 뉴클레오티드로 치환한 것은 OAS1 내의 야생형 제 6 엑손에 대한 컨센서스 스플라이스 수용체 부위에서 일어난다. 상기 치환은 스플라이스 수용체 인식 시그널에서 필수체 G를 대체하지만, 상기 과정에서 하류에서 새로운 스플라이스 인식 부위 일 염기쌍이 생성된다. 그에 의해, 돌연변이된 형태는 해독된 단백질 내의 프레임이동을 생성한다. 돌연변이된 부위는 또한 덜 효과적인 스플라이스 시그널이며, 그 결과 프레임이동된 엑손 6 스플라이싱 이외에 프리커서 RNA의 추가의 별도 스플라이싱이 촉진된다. 이러한 별도 스플라이싱 형태의 바람직한 실시양태는 도 5A 및 도 5B에 제공된다. 유전적 분석의 추가의 예시적 실시예는 하기에 제공된다.

OAS1 전사 변이체의 무차별적 스플라이싱은 신선한 인간 혈청으로부터 단리된 림프구 세포주 및 말초 혈액 단핵구 세포(PBMC)로부터 유래된 RNA의 역전사에 의해 독립적으로 확인되었다. 다양한 반수형 보유 세포주 및 PBMC로부터 역전사된 RNA 산물의 PCR 분석은, 돌연변이 서열 4에 대한 A 뉴클레오티드를 보유하는 RNA 형태가 다수의 전사 변이체 OAS1R 형태를 초래함을 지시하였다. 상기 OAS1R 전사 변이체는 도 5A 및 도 5B에 도식적으로 나타나 있으며, 도 3에 제공된 바와 같이 서열 36 내지 서열 47을 포함한다.

OAS1 유전자의 상기 변이체 OAS1R 형태 및 상응하는 전사 변이체는 서열 20, 서열 21, 서열 22, 서열 23, 서열 24, 서열 25, 서열 26, 서열 27, 서열 28, 서열 29, 서열 32, 서열 35, 서열 46 및(또는) 서열 47로 이루어지는 하나 이상의 폴리펩티드를 코딩하는 것으로 믿어진다. 상기 폴리펩티드는 단일로 또는 다수로, 본원에서 OAS1R 폴리펩티드 또는 OAS1R 단백질로 상호교환적으로 지칭될 수 있다. 많은 상기 폴리펩티드의 공통된 특징은, 이들이 아미노-말단 부분을 보존하지만 일차적으로 그의 카르복실-말단에서는 상이하다는 것이다.

별도 전사체 자체의 생성 이외에, 유전자의 OAS1R 형태는 또한 특정 전사체 변이체에 대해 선택적 선호성을 변형시키는 특정 서열 내용 (예를 들어, 엑손 스플라이스 인핸서)을 함유하거나 폐지할 수 있다. 다음으로 이는 생성된 단백질의 풍부함의 상대적 수준의 상이화를 유발할 것이다. OAS1 유전자의 상기 변이체 OAS1R 형태는 또한 생성된 단백질의 국부화 또는 해독후 변형을 변형시킬 변형시킬 수 있다. 당업자는 특정 OAS1 단백질 형태의 활성, 안정성 또는 이용가능성을 개선시키는 증가된 풍부함 또는 다른 변형이 2'-5'-OAS/RN아제 L 경로의 전체적인 항-바이러스 성능을 개선시킬 수 있음을 이해할 것이다. 당업자는 또한 특정 OAS1 형태의 활성 또는 이용가능성의 억제가 또한 상기 특정 단백질이 다른 특정 OAS1 형태에 비해 유리하거나 심지어 불리한 경우에 2'-5'-OAS/RN아제 L 경로의 전체적인 항-바이러스 성능을 개선시킬 수 있음을 이해하고 있다. 제한 없이, 불리한 OAS1 단백질의 한 실시양태는 상기 특정 OAS1 단백질의 효소 활성을 미리 배제하는 방식으로 바이러스 단백질(들)에 의해 특이적으로 표적화되는 것이다. 유리하지 않은 OAS1 단백질의 추가의

실시양태는 다른 활성 형태와 중합하여, 중합된 단백질의 전체적인 효소 활성을 저하시키거나 폐지하는 (따라서 전체적인 항-바이러스 효과의 감소), 보다 낮은 효소 활성을 갖는 것이다. 상기 메카니즘 중 하나 이상은 바이러스 감염에 대한 저항성에 기여할 수 있다. 그러나, 본 발명은 개시된 변이체 폴리뉴클레오티드 또는 폴리펩티드의 특정 작용 메카니즘에 제한되지 않는다.

따라서, 본 발명은 인간 2'-5'-올리고아데닐레이트 신테타제 유전자의 신규한 형태, 신규한 mRNA 전사체 및 관련된 단백질을 제공한다. 본 발명은 또한 신규한 mRNA 전사체 및 신규한 단백질의 유용성을 개시한다. 상기 신규한 형태는 공개 데이터베이스에 개시된 하나 이상의 몇몇 희귀한 유전적 돌연변이 또는 반수형의 유전자의 존재를 특징으로 한다. OAS1의 상기 신규한 형태인 OAS1R은 C형 간염 바이러스, 및 웨스트 나일(West Nile) 바이러스, 뎅그열 바이러스, 황열 바이러스, 진드기-매개 뇌염 바이러스, 일본 뇌염 바이러스, 세인트 루이스 뇌염 바이러스, 머리 밸리(Murray Valley) 바이러스, 포와슨(Powassan) 바이러스, 로시오(Rocio) 바이러스, 도약병 바이러스, 반지(Banzi) 바이러스, 일헤우스(Ilheus) 바이러스, 코코베라(Kokobera) 바이러스, 쿤진(Kunjin) 바이러스, 알푸이(Alfuy) 바이러스, 소 설사 바이러스, 및 키야사나(Kyasanur) 삼립 질환 바이러스를 비롯한 관련된 플라비바이러스에 대한 저항성의 수준을 담체에 부여한다. OAS 단백질은 또한 실험실 호흡기 세포융합 바이러스 및 피코르나바이러스 세포 배양물 감염 시스템에서 감염을 약화시키는데 중요한 것으로 나타났다. 인간 면역결핍 바이러스-1 (HIV-1)로 감염된 세포가 바이러스를 방출하는데 실패한 것은 OAS 및 (또는) 2-5A의 고농도와 상호관련되었다. 또한, HIV-1 전사활성화 단백질 (tat)은 OAS의 활성화를 차단하는 것으로 나타났다 (문헌 [Muller et al, J Biol Chem. 1990 Mar 5; 265(7): 3803-8]), 따라서 이는 OAS의 신규한 형태가 HIV-1 방어 메카니즘을 피하여 효과적인 치료를 제공할 수 있음을 지시한다. 따라서, 본원에 개시된 OAS1의 상기 OAS1R 형태는 또한 비-플라비바이러스 감염성 작용제에 대한 저항성도 부여할 수 있다.

본 발명은 또한 OAS1의 침팬지 (판 트로글로디테스(*Pan troglodytes*)) 및 고릴라 (고릴라 고릴라(*Gorilla gorilla*)) 형태도 새로이 기재하며, 이들 각각은 유용성을 갖는 신규한 mRNA 및 폴리펩티드를 초래한다. 유전자는 전형적으로 인간, 침팬지 및 고릴라와 같은 근연 영장류에서 매우 고도로 보존되어 있는 반면, OAS1의 중요한 차이가 상기 종 사이에서 관찰된다. 인간과 상대적으로 가장 가까운 침팬지는 말단절단된 단백질 생성물을 유발하는 OAS1 엑손 5 내의 (인간에서 NT\_009775.13의 2,142,351과 동등하고 서열 53으로 정의되는 부위에서의) 단일 염기 치환이다. 침팬지 OAS1 폴리펩티드 및 mRNA 서열은 각각 서열 51 및 서열 55에 의해 제공된다.

고릴라는 또한 전사의 앞당겨진 정지를 유발하는 수용체 스플라이스 부위 근처의 엑손 6 내의 (인간에서 NT\_009775.13의 2,142,351과 동등하고 서열 53으로 정의되는 부위에서의) 2개의 염기쌍 결실을 갖는다. 고릴라의 부분적 폴리펩티드 및 부분적 mRNA 서열은 각각 서열 52 및 서열 56에 의해 제공된다. 인간에 상대적인 침팬지 및 고릴라 전사의 추론된 구조는 도 5C에 제공된다. 인간 OAS1R 폴리펩티드와 같이, 상기 각각의 경우는 카르복실-말단 꼬리의 구조 및 내용에서 가장 현저하게 상이한, 고도로 보존된 폴리펩티드 서열을 갖는다. 상기 폴리펩티드의 공통된 아미노-말단 부분은 OAS1 효소 활성에 필요한 것으로 이미 입증된 모든 요소를 함유한다. 침팬지 및 고릴라는 아프리카의 인간의 것과 유사한 바이러스 켈린지가 수행되었기 때문에, 이러한 독특하지만 기능적으로 유사한 영장류 변이체의 보급은 보다 긴 OAS1 형태의 카르복시-말단 부분이 바이러스 켈린지를 생존시키는데 불필요하거나 심지어 불리하다는 추가의 증거를 제공한다. 침팬지가 인간에서의 전사 변이체의 이중성에 반대되는 OAS1 폴리펩티드의 명백한 말단절단을 갖는다는 사실은, 단지 다른 영장류는 HCV 감염에 감수성인 것으로 알려져 있지만, 침팬지는 증가된 빈도의 바이러스 정리, 및 그 결과 섬유생성 또는 간세포 암종의 부재를 특징으로 하는, 인간에 비해 전형적이지 않은 감염을 갖는다는 관찰과 일치한다.

각각의 신규한 OAS1R cDNA는 상기 돌연변이의 담체인 인간 대상체로부터 클로닝된다. 클로닝은 세포 또는 조직으로부터의 RNA 단리, RNA의 cDNA로의 전환, 및 클로닝에 적합한 이중-가닥 DNA로의 cDNA의 전환을 포함하는 표준 cDNA 클로닝 방법에 의해 수행된다. 당업자가 인식하게 될 바와 같이, 모든 상기 단계는 통상적인 분자 생물학적 분석법이다. 다른 방법으로는 역전사효소 PCR, 5'RACE (cDNA 말단의 급속 증폭), 또는 서던 혼성화에 의한 전형적인 cDNA 라이브러리 제작 및 스크리닝을 들 수 있다. 본원에 기재된 모든 OAS1R 대립유전자는 환자 담체로부터 회수한다. 각각의 새롭게 클로닝된 OAS1R cDNA는 그의 동일성을 확인하고 야생형에 비해 임의의 추가의 서열 차이점을 확인하기 위해 시퀀싱한다.

신규한 OAS1R 유전자 돌연변이는 생성된 OAS1 mRNA의 특성을 변형시킴으로써 바이러스 감염에 대한 저항성에 영향을 줄 수 있다. 따라서, OAS1R 대립유전자의 담체와 동종접합성 야생형 대상체 사이의 mRNA 안정성의 차이가 평가된다. RNA 안정성은 태크맨(Taqman)(등록상표) 및 간단한 서던 혼성화를 비롯한 공지된 검정법을 사용하여 평가하고 비교한다. 이들은 분자 생물학에서 통상적인 방법을 구성한다.

OAS1R 돌연변이는 OAS1 유전자의 조절을 변형시킴으로써 감염 저항성에 영향을 줄 수 있다. OAS 유전자의 발현은 인터페론 처리에 의해 및 바이러스 감염 동안 유도된다. OAS1R 대립유전자는 구조적 발현, 과발현 또는 다른 비조절된 발현을

통해 바이러스 감염에 대한 저항성을 부여할 수 있다. 인터페론 또는 바이러스 자극이 있거나 없이 유전자 발현을 평가하는 몇몇 방법이 사용된다. 상기 방법으로는 발현 마이크로어레이 분석, 노던 혼성화, 태크맨(등록상표) 등을 들 수 있다. 샘플은 말초 혈액 단핵구 세포와 같은 OAS 유전자를 발현하는 것으로 알려진 조직으로부터 수집한다. 유전자 발현을 OAS1R 담체 및 비-담체로부터의 조직 사이에서 비교한다. 한 실시양태에서, 말초 혈액 단핵구 세포를 담체 및 비-담체로부터 수집하고, 배양물에 전과시키고, 인터페론으로 자극한다. 인터페론 도입 동안 OAS1R 대립유전자의 발현 수준을 야생형 대립유전자와 비교한다. 또다른 실시양태에서, 인간 대상체를 인터페론으로 처치하고, OAS1 유전자의 유도를 OAS1R 돌연변이의 담체 대 비-담체에서 평가한다. 당업자는 조직의 수많은 조합, 실험적 설계, 및 OAS1R 유전자 조절을 평가하는데 사용되는 분석 방법을 이해할 수 있다.

일단 각각의 OAS1R에 대한 신규한 cDNA가 클로닝되면, 이는 임의의 수많은 여러 가지 공지된 발현 클로닝 시스템을 사용하여 재조합 OAS1R 단백질을 제작하는데 사용된다. 상기 접근법의 한 실시양태에서, OAS1R cDNA를 표준 분자 생물학적 방법에 의해 폴리히스티딘 폴리펩티드를 코딩하는 DNA의 서열을 함유하는 에피토프 태그에 인접한 에스케리키아 콜라이(*Escherichia coli*) 발현 벡터로 클로닝한다. 그 후, 재조합 단백질을 면역화 급속 친화도 크로마토그래피 또는 유사한 방법을 사용하여 대장균 용해물로부터 정제한다. 당업자는 효모 발현 시스템, 바칼로바이러스 발현 시스템, 차이니스 햄스터 난소 세포 등을 포함하나 이에 제한되지 않는, 재조합 단백질을 정제하는데 사용될 수 있는 많은 여러 가지 발현 벡터 및 숙주 세포가 있음을 인식할 것이다.

계산학적 방법을 사용하여 OAS1R 단백질을 야생형 OAS1 단백질로부터 특이하게 구분짓는, OAS1R 단백질로부터의 짧은 펩티드 서열을 확인한다. 다양한 계산학적 방법 및 시판되고 있는 소프트웨어 패키지가 펩티드 선별을 위해 사용될 수 있다. 이러한 계산학적으로 선별된 펩티드 서열을 FMOC 펩티드 합성 화학법 또는 유사한 방법을 사용하여 제조할 수 있다. 당업자라면 공급된 서열에 따라 짧은 폴리펩티드를 합성하기 위한 수많은 화학법이 있다는 것을 인지할 것이다.

OAS1R 유전자로부터의 펩티드 단편 및 재조합 단백질을 사용하여 상기 유전자 생성물에 대해 특이적인 항체를 생성할 수 있다. 당업자가 인지하고 있는 것처럼, 다수의 상이한 숙주 생물체, 보조제 등을 사용하는 것을 포함하는 수많은 항체 생성 방법이 있다. 한 전통적인 실시양태에서는, 소량 (150 마이크로그램)의 정제된 재조합 단백질을 뉴질랜드 화이트 토끼 (New Zealand White Rabbit)의 등에 피하 주사한 후, 유사한 양을 부스터로서 수개월마다 주사한다. 그 후에 토끼 혈청을 정맥천자에 의해 수집하고, 혈청, 정제된 IgG, 또는 면역화 단백질에 대해 특이적인 친화도에 따라 정제된 항체를 수집할 수 있다. 당업자가 인지하고 있는 것처럼, 유사한 방법을 사용하여 래트, 마우스, 염소, 및 기타 생물체에서 항체를 생성할 수 있다. 전술한 펩티드 단편 또한 OAS1R 단백질에 대해 특이적인 항체를 생성하는데 사용될 수 있다. 모노클로날 및 폴리클로날 항체 둘 다를 생성하는 것이 본 발명을 실시하는 데 적합하다. OAS1R 단백질에 대해 특이적인 모노클로날 항체를 분비하는 마우스 하이브리도마 세포주의 생성을 표준 분자 기술에 따라 수행할 수 있다.

전술한 바와 같이 제조된 항체를 사용하여 세포, 조직 및 생물체에서 OAS1R 단백질의 존재 또는 부재를 평가하는 진단법을 개발할 수 있다. 이러한 접근법의 한 실시양태에서, 인간 혈청에서 OAS1R 단백질을 검출하기 위하여 정제된 재조합 OAS1R 단백질 및 특이적인 항체를 사용하는 효소면역검정법이 개발될 수 있다. 이러한 진단법을 사용하여 전술한 유전자 돌연변이의 운반체 및 비-운반체 조직에서 OAS1R 단백질의 존재 또는 부재를 확인할 수 있다.

또한 전술한 바와 같이 제조된 항체를 사용하여 이러한 돌연변이가 일어난 환자로부터 천연 OAS1R 단백질을 정제할 수 있다. 인간 세포 및 조직으로부터 천연 단백질을 정제하기 위해 항체를 사용하는 수많은 방법이 이용가능하다. 한 실시양태에서는, 항체가 균질화된 인간 조직 및 단백질 A를 사용하는 항체 포획을 포함하는 면역침강 실험에서 사용될 수 있다. 상기 방법으로 돌연변이체 OAS1R 단백질의 농축 및 추가의 평가가 가능하다. 컬럼 크로마토그래피, 친화도 크로마토그래피, 고성능 액체 크로마토그래피, 염석, 투석, 전기영동, 등전점 전기영동, 분별 원심분리 등을 비롯한, 천연 형태의 OAS1R을 단리하기 위한 수많은 기타 방법이 이용가능하다.

단백질체 방법을 사용하여 OAS1R 돌연변이가 2차, 3차, 및 4차 단백질 구조에 미치는 영향을 평가한다. 또한 단백질체 방법을 사용하여 OAS1R 돌연변이가 OAS 단백질의 번역-후 변형에 미치는 영향을 평가한다. 프로테아제 절단, 글리코실화, 포스포릴화, 술페이트화, 화학기 또는 복합체 분자의 부가 등을 비롯한, 수많은 공지된 단백질의 번역-후 변형이 있을 수 있다. 2차 및 3차 단백질 구조를 평가하는 통상의 방법은 핵자기공명(NMR) 분광학이다. NMR을 사용하여 야생형 OAS1 단백질과 OAS1R 단백질 간의 2차 구조 및 3차 구조에 있어서의 차이점을 조사한다. 기능성 부위의 활성을 평가하는 방법, 예를 들면 전달 핵 오버하우저 분광학(Transfer Nuclear Overhauser Spectroscopy; TrNOESY) 등을 비롯한 전통적인 NMR을 변형시킨 방법 또한 적합하다. 당업자가 인지하고 있는 것처럼, 이러한 접근법을 약간 변형시킨 수많은 방법 및 결과 데이터의 해석 방법이 사용될 수 있다. 상기 방법을 모두 본 발명을 실시하는 데 포함시키고자 한다. 결정화 및 X-선 회절에 의해 단백질 구조를 측정하는 기타 방법이 사용된다.

또한 질량 분광학을 사용하여 돌연변이체와 야생형 OAS 단백질 간의 차이점을 평가할 수 있다. 상기 방법을 사용하여 구조의 차이점 뿐만 아니라, 단백질의 번역-후 변형에 있어서의 차이점을 평가할 수 있다. 이러한 접근법의 한 전형적인 실시양태에서는, 야생형 OAS1 단백질 및 돌연변이체 OAS1R 단백질을 전술한 방법 중 하나를 사용하여 인간 말초 혈액 단핵 세포로부터 정제한다. 상기 세포를 OAS 단백질의 발현을 증가시키기 위해, 전술한 바와 같이 인터페론으로 자극시킬 수 있다. 정제된 단백질을 특정 프로테아제 (예, 트립신)를 사용하여 절단하고 질량 분광학을 사용하여 평가한다. 당업자가 인지하고 있는 것처럼, 수많은 대안적 방법 또한 사용될 수 있다. 본 발명은 이러한 추가의 대안적 방법을 포괄한다. 예를 들면, 매트릭스 지원 레이저 이온화(MALDI) 또는 전자분무 이온화(ESI) 질량 분광법이 사용될 수 있다. 게다가, 질량 분광학은 종합적인 예비정제에 대한 대안으로서 세포 단백질의 2차원 겔 전기영동 분리와 병용될 수 있다. 질량 분광학은 또한 펩티드 핑거프린트 데이터베이스 및 다수의 추적 알고리즘과 병용될 수 있다. 번역-후 변형, 예컨대 포스포릴화 또는 글리코실화의 차이점 또한 질량 분광학과 다수의 전처리, 예컨대 글리코실라제 및 포스포타제를 병용함으로써 조사할 수 있다. 상기 방법 모두 본 출원의 일부라고 생각된다.

OAS1은 테트라머로서 활성이고, 자가-회합을 방해하는 돌연변이가 효소 활성에 영향을 준다. 공지된 방법을 사용하여 OAS1R 돌연변이가 테트라머 형성에 미치는 영향을 평가한다. 예를 들면, OAS1 또는 OAS1R을 과발현하는 환자 세포, 세포 배양물, 또는 형질감염된 세포로부터 OAS1/OAS1R 복합체를 단리하기 위하여 OAS1 및 OAS1R 특이적 항체를 사용하여 면역침강을 수행한다. 그 후에, 상기 복합체를 당업계에 널리 공지된 겔 전기영동 또는 기타 크로마토그래피 방법으로 평가할 수 있다.

OAS1은 다른 단백질과의 상호작용에 의해 바이러스 저항성을 부여할 수 있다. 효소는 bcl-2 족의 BH3 도메인과 구조적으로 상동성을 갖는 영역을 함유한다. 상기 도메인은 OAS1의 기능에 있어서 중요할 수 있다. 본 발명에 따라서, OAS1-특이적 항체를 사용하여 전술한 바와 같이 다양한 기원으로부터의 OAS1 단백질을 포함하는 단백질 복합체를 단리할 수 있다. 그 후에, 상기 복합체를 겔 전기영동에 의해 평가하여 상호작용하는 복합체 일원을 분리할 수 있다. 겔은 웨스턴 블롯팅을 비롯한 수많은 방법을 사용하여 조사될 수 있고, 신규한 상호작용하는 단백질은 펩티드 서열분석을 사용하여 단리 및 확인될 수 있다. 야생형 및 OAS1R 추출물 중의 OAS 복합체 함량의 차이 또한 평가될 것이다. 당업자가 인지하고 있는 것처럼, 기술된 방법은 다만 OAS 복합체 중의 상호작용하는 단백질을 정제, 확인 및 평가하는 데 사용될 수 있는 수많은 상이한 접근법의 일부일 뿐이다. 추가의 방법으로는 파지 발현(phage display) 및 효모의 이중-하이브리드 방법을 사용하는 것이 포함되나, 이들로 한정되지는 않는다.

OAS1은 C형 간염 바이러스 NS5A 단백질과 상호작용하는 것으로 알려져 있다 (문헌 [Taguchi, T. et al., J. Gen. Virol. 85:959-969, 2004]). 따라서, 본 발명은, 이론에 구애됨이 없이, NS5A 단백질과 상호작용하지 않는 OAS1 단백질에 관한 것이며, 상기 단백질은 OAS1의 스플라이싱 변이체, 본 발명의 하나 이상의 돌연변이를 함유하는 OAS1 폴리뉴클레오티드, 코딩 영역에서 하나 이상의 돌연변이가 일어난 OAS1 폴리뉴클레오티드 및(또는) 하나 이상의 염기 치환, 결실, 또는 부가가 일어난 OAS1 폴리뉴클레오티드에 의해 코딩된 mRNA에 의해 발현된, 본 발명의 폴리뉴클레오티드에 의해 발현된 본 발명의 폴리펩티드이고, NS5A 단백질과의 결합은 변경되거나 또는 방지된다.

NS5A 단백질은 인터페론의 항바이러스 활성을 억제함으로써 생물학적 활성을 발휘할 수 있다. 상기 항바이러스 활성은 한 모델에서 통상적으로, 인터페론이 보조인자, 예컨대 이중가닥 RNA의 존재하에 OAS1 활성을 자극할 때 그 효력을 발휘한다. OAS1은 ATP를 2'-5'-연결된 올리고아데닐레이트로 중합시키고, 이는 RN아제 L을 활성화하여 mRNA를 비롯한 단일가닥 RNA를 절단한다. NS5A가 OAS1에 결합하면 그의 활성이 억제되고, 그에 따라 억제 또는 방지되지 않으면 바이러스 RNA의 파괴를 유발하는 활성의 캐스케이드를 억제 또는 방지할 수 있다.

본 발명이 상기 모델에 좌우되지는 않지만, NS5A가 OAS1에 결합하면 돌연변이형 OAS1이 NS5A 결합을 회피하고 억제하여 그에 따라 ATP를 중합시키는 것의 정상적인 기능을 수행할 수 있는 모델과 일관된다. 이러한 경우에, 본원에 기술되어 있는 임상 결과와 일관되게, 이러한 돌연변이가 일어난 사람은 C형 간염 바이러스에 의한 감염에 대해 저항성이다. 마찬가지로, 전술한 바와 같이 침팬지가 지니고 있는 말단절단형(truncated) OAS1은 NS5A 또는 기타 바이러스 단백질에 의한 결합을 회피하고, 그에 따라 높은 빈도로 침팬지 바이러스 제거의 관찰을 가능하게 한다. 몇몇 경우에는 돌연변이가 NS5A에 대한 OAS1의 결합 부위에 직접적으로 영향을 미칠 수 있다. 다른 경우에는 돌연변이가 실제 결합 부위와는 별개의 부위에서 일어날 수 있지만, 배위 변화가 일어나 OAS1이 NS5A에 결합하는 것이 억제되거나, 지연되거나, 또는 방지된다.

따라서 OAS1이 NS5A에 결합하는 것은 생리학적 및 병리학적으로 관련있는 방식으로 OAS1 폴리뉴클레오티드의 염기 돌연변이, 결실 또는 부가가 코딩된 OAS1 단백질의 생물학적 기능에 미치는 영향을 분석하기 위한 객관적 시험을 제공한다. 이러한 결합은 당업계에 공지된 방식으로 분석된다. 타구치, 티.(Taguchi T.) 등에 의해 기술된 것처럼 (문헌 [J. Gen.

Virol. 85:959-969, 2004]), 예시적이지만 제한하는 것은 아닌 한 방법에서, 헬라 세포를 GST-태그부착된 NS5A 및 HA-태그부착된 OAS1을 코딩하는 발현 플라스미드로 일시적으로 형질감염시킨다. 상기 예시에서 OAS1이란 OAS1의 스플라이싱 변이체, 코딩 영역에서 하나 이상의 돌연변이가 일어난 OAS1 폴리뉴클레오티드 및(또는) 하나 이상의 염기 치환, 결실 또는 부가가 일어난 OAS1 폴리뉴클레오티드에 의해 코딩되는 OAS1을 비롯한, 본 발명에 따른 OAS1을 의미한다. 적합한 인큐베이션 시간, 예컨대 12 내지 16시간 후에, 세포를 세척하고, 용해시키고, 원심분리하며, 생성된 상등액을 GST-태그부착된 단백질을 분리하는 데 도움이 되는, 글루타티온-콘주게이팅 세파로즈 비드와 혼합한다. GST-태그부착된 NS5A와 HA-태그부착된 OAS1 단백질의 복합체를 NS5A에 대한 항체 및 항-HA 항체를 사용하여 영상화함으로써 확인한다. 상기 방법의 변형된 방법은 개별 단백질에 대하여 다른 태그, 예컨대 FLAG-태그를 사용하는 것을 포함한다. 본 발명의 명세서에서, 주요 변수는 OAS1 단백질 또는 폴리펩티드이다. OAS1 단백질 또는 폴리펩티드가 하나의 생물학적으로 관련있는 활성 (즉, 바이러스가 숙주에서 복제하는 능력에 대해 방어적인 것으로 알려져 있는 C형 간염 단백질과의 결합)을 전달하는 능력을 상기 검정법을 사용하여 객관적으로 시험한다. NS5A에 결합하지 않는 OAS1 단백질 및 폴리펩티드는 치료용 단백질로서 적합한 후보이다.

OAS1 단백질은 ATP가 올리고아데닐레이트 분자로 전환되는 것을 촉매하는 효소이다. 다수의 방법이 OAS1 효소의 활성 평가에 이용가능하다. 상기 방법을 사용하여 야생형 효소에 비해 OAS1R 돌연변이가 돌연변이체 단백질의 활성에 미치는 영향을 측정한다. 예를 들면, 올리고아데닐레이트 합성 활성은 폴리아데닐레이트에 도입된 <sup>32</sup>P-방사성 표지된 ATP를 정량화함으로써 측정할 수 있다.

방사성 표지된 폴리아데닐레이트는 전기영동 또는 이온 교환 크로마토그래피를 비롯한 다수의 크로마토그래피 방법에 의해 길이 측면에서 정량화되고 특성표시될 수 있다. 또한 상기 검정법으로 기질 (ATP) 결합 및 효소 동역학의 특성표시가 가능하다. OAS1은 dsRNA에 의해 활성화된다. 상기 활성화의 동역학을 본원에서 기술된 활성 검정법 및 당업계에 개시된 합성 dsRNA를 사용하여 OAS1에서 분석하고 OAS1R과 비교한다.

본 발명의 폴리펩티드는 당업계에 공지된 상기 방법 및 기타 방법에 의해 올리고아데닐레이트 합성 활성을 갖는 것으로 입증된다. 도 7 및 도 8은 본 발명의 다수의 예시적 폴리펩티드의 활성을 입증한다. 활성의 정량적 수준에 상관없이, 2'-5'-올리고아데닐레이트를 제조하는 능력은 RN아제 L의 활성화를 통해 항바이러스 효과를 야기한다는 것이 당업자에게 잘 이해된다. 따라서, 본 발명의 폴리펩티드가 올리고아데닐레이트 합성 활성을 갖는다는 단순한 사실은 상기 폴리펩티드가, 특히 후술되는 치료 용도를 고려할 때 유용성을 갖는다는 것을 나타낸다.

생물학적 연구를 수행하여 OAS1R 돌연변이체 유전자가 바이러스 감염으로부터 방어하는 정도를 평가한다. 상기 생물학적 연구는 일반적으로 돌연변이체 OAS1R 유전자 또는 단백질을 세포 또는 전체 생물체에 도입하고, 야생형 대조군과 비교하여 생물학적 활성 및 항바이러스 활성을 평가하는 형태를 취한다. 상기 접근법의 한 전형적인 실시양태에서는, 본원에 기술된 바와 같이 단리된 cDNA를 SV40 프로모터 서열로부터 클로닝된 cDNA의 발현을 유도하는 포유류 발현 벡터로 클로닝함으로써, OAS1R 유전자를 배양물 중의 아프리카 그린 원숭이(African Green monkey) 베로 (Vero) 세포에 도입한다. 상기 벡터는 또한 OAS1R 유전자의 효율적인 발현을 가능하게 하는 SV40 및 시토크롬 B 바이러스 인핸서 요소, 및 배양물 중에서 선별을 위한 네오마이신 저항성 유전자를 함유할 것이다. 그 후에 OAS1R 발현의 생물학적 영향을 뎅기 바이러스로 감염된 베로 세포에서 평가할 수 있다. OAS1R이 다수의 플라비바이러스에 대해 광범위한 저항성을 부여하는 경우에, 야생형에 비해 OAS1의 돌연변이형을 발현하는 세포주에서의 바이러스 증식의 약화가 예상된다. 당업자가 인지하고 있는 것처럼, 세포 및 생물체에서 그리고 상이한 감염 물질에 대한 반응에서 OAS1R 유전자 및 단백질의 생물학적 영향을 평가하는 데 사용될 수 있는 수많은 상이한 실험적 접근법이 있다. 예를 들어, 상기 예시에서, 상이한 발현 벡터, 세포형, 및 바이러스 종류를 사용하여 OAS1R 저항성 효과를 평가할 수 있다. 배양물 중 초대(primary) 인간 세포를 세포주에 대한 대립으로서 평가할 수 있다. 세포를 바이러스에 도입하기 전에 이중가닥 RNA 또는 인터페론으로 자극할 수 있다. 교대로 프로모터 및 인핸서 서열을 함유하는 발현 벡터를 평가할 수 있다. 플라비바이러스를 제외한 바이러스 (예, 호흡기 세포융합 바이러스 및 피코나바이러스)가 평가된다.

트랜스제닉 동물 모델을 생산하여 전체-생물체의 바이러스 감염에 대한 방어에서 돌연변이형 OAS1의 유용성을 평가한다. 한 실시양태에서는, OAS1R 유전자를 플라비바이러스에 감염되기 쉬운 마우스 (예, 실험실용 C3H/He 근교계)의 게놈에 도입한다. 상기 OAS1R 유전자를 감수성 마우스에서 감염을 변형시키거나 또는 감염에 대한 저항성을 부여하는 능력에 대해 평가한다. 당업자가 이해하고 있는 것처럼, 수많은 표준 방법을 사용하여 트랜스제닉 인간 OAS1R 유전자를 마우스에 도입할 수 있다. 상기 방법을 조직 특이적 발현 패턴에 영향을 미치거나 또는 내생성 화학물질의 도입, 유도성 또는 조직 특이적 프로모터의 사용 등을 통해 트랜스젠(transgene)의 조절을 가능하게 하는 다른 방법과 병용할 수 있다.

C형 간염 감염의 모델로서, OAS1R 유전자를 발현하는 세포주를 사용하여 소 설사 바이러스(BVDV)로의 감염에 대한 감수성, 저항성, 또는 그의 변형을 평가할 수 있다. BVDV는 잠재적인 항-HCV 항바이러스 약물의 효능을 시험하기 위해 통상 사용되는 모델이다 (문헌 [Buckwold et. al., Antiviral Research 60:1-15, 2003]). 한 실시양태에서는, OAS1R 유전자를 본질적으로 전술한 바와 같은 발현 벡터를 사용하여 KL (송아지 폐) 세포에 도입하고, 상기 세포주에서 BVDV 감염을 변형시키는 능력을 시험할 수 있다. 게다가, HCV 감염의 마우스 모델 (예, 인간 간의 마우스에의 이식, 인간 간세포의 마우스 간의 융합 등)을 사용하여 OAS1R 유전자의 트랜스제닉 셋팅에서 HCV 감염의 변형을 평가할 수 있다. 실험을 실시하여 OAS1R 유전자 발현의 영향을 HCV 바이러스 배양 시스템에서 평가한다. 또한 세포 배양 시스템을 사용하여 돌연변이체 OAS1R 유전자가 다양한 조건하에서 아포토시스 촉진에 미치는 영향을 평가할 수 있다. 한 실시양태에서는, BVDV, HCV, 및 기타 플라비바이러스를 비롯한 다수의 바이러스로 감염된 세포에서 아포토시스를 촉진하는 능력에 대해 돌연변이형 OAS1R의 세포 배양물을 야생형 OAS1 서열과 비교하여 평가할 수 있다. 당업자가 인지하고 있는 것처럼, 수많은 아포토시스 측정 방법이 이용가능하다. 가장 통상적인 방법은 아가로스 겔 전기영동과 형광 콘주게이션 방법을 병용하여 아포토시스가 일어나는 세포에서 특징적인 계늑 "DNA 래더링(laddering)" 영향의 검출을 포함한다.

OAS1R 돌연변이형의 영향을 증가시키는 바이러스를 불완전하게 방해하는 능력을 세포 배양물 및 작은 동물 모델에서 시험할 수 있다.

OAS1R 유전자형의 존재 또는 부재가 다른 인간 표현형에 영향을 미치는 정도 또한 조사할 수 있다. 예를 들면, OAS1R 돌연변이를 HCV 감염된 대상체에서 바이러스 역가 및 자발적인 바이러스의 제거와 관련하여 평가한다. 숙주 OAS1 유전자형을 다른 플라비바이러스 감염 경로와 관련시키는 유사한 방법 또한 실시될 수 있다. OAS1R 돌연변이가 HCV 감염된 환자에서 인터페론 또는 리바비린 처리된 인터페론을 사용하여 성공적인 결과를 도모하는 데 미치는 영향 또한 조사한다. 상기 돌연변이는 소정의 수준의 감염 저항성을 부여할 뿐만 아니라, 감염된 대상체에서 인터페론-리바비린 처리하거나 또는 그러한 처리 없이 자발적인 바이러스의 제거를 촉진할 수 있다. 게다가, 정신분열병이 플라비바이러스 감염이 풍토성인 지역에서 보다 높은 빈도로 발병한다는 것이 보고되어 있는데, 이는 플라비바이러스 저항성 대립유전자와 정신분열병 소질 사이의 상관관계를 설명해준다. 이러한 연결은 정신분열병 표현형 및 OAS1R 돌연변이를 포함하는 추가의 유전자 관련 연구를 수행함으로써 평가된다. OAS1R 돌연변이가 IDDM, 전립선 및 기타 암, 및 정신분열병에 대한 감수성에 미치는 영향 또한 평가될 것이다.

본 발명은 신규하고 유용성이 있는 OAS1R 변이체 mRNA (서열 36 내지 서열 43으로 확인됨)를 기재하고 있다. 본 발명은 기재된 변이체 mRNA의 사용 방식에 의해 제한되지 않는다. 한 바람직한 실시양태에서는, 상기 변이체 mRNA가 증가 또는 감소된 바이러스 (HCV를 포함함) 감수성에 따라 인간 대상체를 구분하여 스크리닝하는 데 사용된다. 다른 바람직한 실시양태에서는, 상기 변이체 mRNA는 IDDM, 전립선 및 기타 암, 및(또는) 정신분열병에 대한 감수성에 따라 스크리닝하는 데 유용하다. 이러한 구분짓는 스크리닝을 당업계에 공지된 발현 분석에 따라 실시하여 해당 인간 대상체로부터 유래된 샘플에 존재하는 하나 이상의 변이체 OAS1R mRNA의 상대적인 양을 측정한다. 대조 샘플과 비교하였을 때 인간 대상체의 샘플에서 하나 이상의 OAS1R mRNA 변이체가 증가한 양 또는 감소한 양은 고려하고 있는 시험에 적합한, 바이러스, IDDM, 전립선 및 기타 암, 및(또는) 정신분열병에 대한 대상체의 감수성 정도의 지표이다.

본원에서 기술된, 2',5'-올리고아데닐레이트 신테타제 (OAS)는 ATP로부터 2'에서 5'로 연결되는 짧은 올리고아데닐레이트 (2-5A) 분자를 합성하는 IFN- $\alpha$  유도성, RNA 의존성 효소 분자 효소 족이다. OAS 효소는 바이러스 감염에 대한 비특이적 면역 방어의 중요한 부분을 구성하고 있고 바이러스 감염에 대한 세포 마커로서 사용되어 왔다. 본원에서 기술된 C형 간염 감염에서의 역할 외에도, OAS 활성화는 다른 질병 상태, 특히 바이러스 감염이 소정의 역할을 하는 질병에 영향을 준다.

특이적 병원 메카니즘이 현 분석의 대상이지만, 바이러스 감염은 당뇨병과 같은 질병의 발병에서 소정의 역할을 하는 것으로 생각된다. 림프구 OAS 활성화는 제1형 당뇨병을 앓고 있는 환자에서 상당히 상승하고, 이는 OAS가 바이러스 감염과 질병 발병 사이에 중요한 상관관계가 있을 수 있다는 것을 설명해준다. 일란성 쌍생아 중 당뇨병을 앓고 있는 한 쌍생아를 포함하는 연구에서, 보네비-닐센(Bonnevie-Nielsen) 등 (문헌 [Clin Immunol. 2000 Jul; 96(1):11-8])은 OAS가 최근에 발병한 제1형 당뇨병 및 다년간 지속된 제1형 당뇨병 둘 다에서 지속적으로 활성화되었다는 것을 밝혀냈다. 제1형 당뇨병에서 연속적으로 상승한 OAS 활성화는 정상적인 항바이러스 반응과는 명확하게 구분되고 효소의 만성적인 자극, 하향조절 메카니즘의 장애, 또는 내생성 또는 외생성 바이러스, 또는 그들의 생성물에 대한 비정상적인 반응의 지표일 수 있다.

바이러스 감염과 당뇨병 발병의 보다 직접적인 상관관계는 만성 C형 간염을 앓고 있는 환자의 13 내지 33%가 당뇨병 (제2형 당뇨병)을 앓고 있다는 것을 밝혀낸 수많은 연구에 의해 예증되고, 상기 수준은 만성 B형 간염을 앓고 있는 대응하는 건강한 대조군 또는 환자의 수준과 비교하였을 때 상당히 증가한 것이다 (문헌 [Knobler et al. Am J Gastroenterol.

2003 Dec; 98(12):2751-6]). 현재까지 OAS가 C형 간염 감염 후에 당뇨병의 발병에서 소정의 역할을 하는 것으로 보고 되지는 않았지만, OAS는 항바이러스 반응 시스템을 위한 유용한 마커일 수는 있다. 게다가, 본 발명에 따라 보고된 결과는 C형 간염 감염이 유발적으로 당뇨병과 관련있다면, 본원에 기재된 조성물 및 방법을 사용하여 C형 간염 감염을 억제하거나 또는 근절하는 것이 당뇨병의 예방 또는 근절에서 유리할 수 있다는 것을 나타낸다.

추가로 공개 연구는 OAS가 상처 치유 및 그의 병리학적 장애, 특히 정맥성 괴저 및 당뇨병-관련하여 불완전하게 치유되는 상처의 경우에 필수적인 역할을 한다는 것을 밝혀냈다 (WO 02/090552). 불완전한 상처 치유의 경우에, 영향받은 조직에서 OAS mRNA 수준은 제1형 당뇨병을 앓고 있는 환자로부터 유래한 림프구에서와 마찬가지로 상승하기 보다는 감소하였다. 이러한 발견은 당뇨병 및 당뇨병-관련 상처 치유에서 면역 반응의 원인론적으로 중요한 마커로서의 OAS를 암시한다.

OAS는 또한 전립선암에 관련있는 세포 과정에서 매개 역할을 할 수 있다. OAS의 일차적인 생화학적 기능은 RN아제 L (2-5A 분자에 의해 효소적으로 자극되는 특이하게 조절되는 엔도리보뉴클레아제)의 활성을 증진시키는 것이다. RN아제 L은 IFN의 항바이러스 효과를 매개하는 데 있어서 잘 확립된 역할을 하고, 제1형 유전성 전립선암 대립유전자 (*HPC1*)를 위한 강력한 후보이다. RN아제 L의 돌연변이는 사람들을 증가된 발병률로 전립선암에 걸리기 쉽게 하는 것으로 밝혀졌고, 일부 경우에는 비-RN아제 L-관련 경우와 비교하였을 때 보다 공격적인 질병 및(또는) 보다 낮은 발병 연령을 나타낸다. 시양(Xiang) 등 (문헌 [Cancer Res. 2003 Oct 15; 63(20): 6795-801])은 생물학적으로 안정한 2-5A의 포스포로티올레이트 유사체가 RN아제 L 활성을 유도하고 말기 전이성 인간 전립선암 세포주의 배양물에서 아포토시스를 일으킨다는 것을 입증하였다. 이들의 발견은 2-5A 수준의 증가와 함께 OAS 활성의 상승은 유력한 아포토시스 경로를 통해 암 세포의 파괴를 촉진할 수 있다는 것을 설명해준다. 따라서, 본원에 기재된 조성물 및 방법의 사용은 전립선암의 검출, 치료 및(또는) 예방에서 유용성을 발견할 수 있다.

OAS는 추가로 RN아제 L의 조절을 통해 또는 또다른 아직 밝혀지지 않은 경로를 통해 정상적인 세포 성장 조절에서 소정의 역할을 할 수 있다. 불리하게 조절되는 세포 성장에서 OAS의 중요성을 지지해주는 주목할 만한 증거가 있다. 리시엑키 (Rysiecki) 등 (문헌 [J. Interferon Res. 1989 Dec; 9(6): 649-57])은 인간 OAS의 교아 세포주로의 안정한 형질감염이 감소된 세포 증식을 초래한다는 것을 입증하였다. OAS 수준은 또한 정지 세포주 대 증식 세포주를 비교하는 수많은 연구 (예, 문헌 [Hassel and Ts'O, Mol Carcinog. 1992; 5(1):41-51] 및 문헌 [Kimchi et al., Eur J Biochem. 1981; 114(1): 5-10])에서 측정가능한 것으로 밝혀졌고, 각 경우에 OAS 수준은 정지 세포에서 최고였다. 기타 연구가 OAS 수준과 세포 주기의 기(phase) 사이의 상관관계를 밝혀냈고, OAS 수준은 후기 S기 동안 급격하게 상승한 후에 G2에서 갑자기 하강한다 (문헌 [Wells and Mallucci, Exp Cell Res. 1985 Jul; 159(1): 27-36]). 수많은 연구가 OAS의 유도와 여러가지 형태의 인터페론으로의 자극 후에 항증식성 효과의 개시 사이의 상관관계를 밝혀냈다 (문헌 [Player and Torrence, Pharmacol Ther. 1998 May; 78(2); 55-113] 참조). OAS의 유도는 또한 세포 분화 중에도 나타났다 (예, 문헌 [Salzberg et al., J Cell Sci. 1996 Jun; 109(Pt 6): 1517-26] 및 문헌 [Schwartz and Nilson, Mol Cell Biol. 1989 Sep; 9(9): 3897-903]). 혈소판 유래 성장 인자 (PDGF)에 의한 OAS의 유도 (문헌 [Zullo et al. Cell. 1985 Dec; 43 (3 Pt 2): 793-800]) 및 열충격 유도된 성장 조건하에서의 OAS의 유도 (문헌 [Chousterman et al., J Biol Chem. 1987 Apr 5; 262(10): 4806-11])에 관한 기타 보고는 OAS의 유도가 정상적인 세포 성장 제어 메카니즘이라는 가설을 이끌어냈다. 따라서, 본원에 기재된 조성물 및 방법의 사용은 암의 검출, 치료 및(또는) 예방에서 광범위한 유용성을 발견할 수 있다.

폴리뉴클레오티드 분석

올리고아데닐레이트 신테타제 유전자는 그의 뉴클레오티드 서열이 올리고아데닐레이트 신테타제, 돌연변이체 올리고아데닐레이트 신테타제, 또는 올리고아데닐레이트 신테타제 유사유전자(pseudogene)를 코딩하는 핵산이다. 이들은 게놈 DNA, mRNA 또는 cDNA, 및 단일가닥 또는 이중가닥 형태일 수 있다. 바람직하게는, 게놈 DNA가 mRNA와 비교하여 생물학적 샘플에서의 상대적인 안정성 때문에 사용된다. 비교 올리고아데닐레이트 신테타제 유전자의 완전한 게놈 서열의 2,130,000 내지 2,157,999개의 연속적인 뉴클레오티드로 이루어진 폴리뉴클레오티드 서열은 서열 목록에 서열 19로서 제공되고 진뱅크(Genebank) 등록번호 NT\_009775.13에 해당한다.

핵산 샘플을 세포, 전형적으로는 말초 혈액 백혈구로부터 수득한다. mRNA를 사용할 경우에는, 세포를 RN아제 억제 조건 하에서 용해시킨다. 한 실시양태에서, 제1 단계는 전체 세포 mRNA를 단리하는 것이다. 그 후에 폴리 A+ mRNA를 올리고-dT 셀룰로스 컬럼에의 혼성화에 의해 선별할 수 있다.

바람직한 실시양태에서, 핵산 샘플에는 올리고아데닐레이트 신테타제 대립유전자 물질이 풍부하다. 전형적으로는 게놈 DNA 또는 mRNA를 본원에 기재된 폴리뉴클레오티드 합성 프라이머를 사용하는 프라이머 신장 반응에 가함으로써 풍부해진다. 분석할 샘플의 특히 바람직한 제조 방법은 폴리머라제 연쇄반응 (PCR)에서 프라이머로서 예비선별된 폴리뉴클레오티드를 사용하여 증폭된 (PCR) 생성물을 형성한다.

### 폴리뉴클레오티드 프라이머의 제조

프라이머 신장에 의해 합성될 프라이머, 프로브 및 핵산 단편 또는 절편과 관련하여 본원에서 사용되는 용어 "폴리뉴클레오티드"는 2개 이상, 바람직하게는 3개 초과,의 테옥시리보뉴클레오티드 또는 리보뉴클레오티드로 이루어진 분자라고 정의된다. 그의 정확한 크기는 다수의 인자에 좌우될 것이고, 결국 최종 사용 조건에 좌우된다. 본원에서 사용되는 용어 "프라이머"는 핵산 가닥에 대해 상보적인 프라이머 신장 생성물의 합성이 유도되는 조건하 (즉, DNA 폴리머라제, 역전사효소 등과 같은 중합을 위한 뉴클레오티드 및 물질의 존재하 및 적합한 온도 및 pH)에 두었을 때 핵산 합성의 개시점으로서 작용할 수 있는, 핵산 제한효소 소화물로부터 정제되었거나 또는 합성 제조된 폴리뉴클레오티드를 지칭한다. 프라이머는 바람직하게는 최고 효율을 위해서는 단일가닥이지만, 별법으로 이중가닥 형태일 수도 있다. 이중가닥이라면, 프라이머는 우선 신장 생성물을 제조하는 데 사용하기 전에 그의 상보적 가닥으로부터 분리하기 위해 처리된다. 바람직하게는, 프라이머는 폴리데옥시리보뉴클레오티드이다. 프라이머는 중합을 위한 물질의 존재하에 신장 생성물의 합성을 프라이밍하기에 충분히 길어야 한다. 프라이머의 정확한 길이는 온도 및 프라이머 기원을 비롯한 다수의 인자에 좌우될 것이다. 예를 들면, 표적 서열의 복잡한 정도에 따라, 폴리뉴클레오티드 프라이머는 전형적으로 15 내지 25개 이상의 뉴클레오티드를 함유하지만, 보다 적은 뉴클레오티드를 함유할 수도 있다. 짧은 프라이머 분자는 일반적으로 주형과 충분히 안정한 하이브리드 복합체를 형성하기 위해 보다 낮은 온도를 요구한다.

본원에서 사용되는 프라이머는 합성 또는 증폭될 각 특정 서열의 상이한 가닥에 대해 "실질적으로" 상보적이도록 선택된다. 이는 프라이머가 그의 각 주형 가닥과 비-랜덤 혼성화되도록 충분히 상보적이어야 한다는 것을 의미한다. 따라서, 프라이머 서열은 주형의 정확한 서열을 반영할 수 있거나 또는 그렇지 않을 수 있다. 예를 들면, 비-상보적 뉴클레오티드 단편은 프라이머의 5' 말단에 부착될 수 있고, 나머지 프라이머 서열은 가닥에 대해 실질적으로 상보적이다. 이러한 비-상보적 단편은 전형적으로 엔도뉴클레아제 제한 부위를 코딩한다. 별법으로, 비-상보적 염기 또는 보다 긴 서열은, 프라이머 서열이 합성 또는 증폭될 가닥의 서열과, 비-랜덤 혼성화되기에 충분한 상보성을 가진다면, 프라이머에 산재되고, 그에 따라 폴리뉴클레오티드 합성 조건하에서 신장 생성물을 형성한다.

본 발명의 프라이머는 DNA-의존성 RNA 폴리머라제 프로모터 서열 또는 그의 상보체를 함유할 수도 있다. 예를 들어, 문헌 ([Krieg, et al., Nucl. Acids Res., 12:7057-70 (1984)], [Studier, et al., J. Mol. Biol., 189:113-130 (1986)] 및 [Molecular Cloning: A Laboratory Manual, Second Edition, Maniatis, et al., eds., Cold Spring Harbor, N. Y. (1989)]) 등을 참조한다.

DNA-의존성 RNA 폴리머라제 프로모터를 함유하는 프라이머가 사용되는 경우, 상기 프라이머는 증폭될 폴리뉴클레오티드 가닥과 혼성화되고, DNA-의존성 RNA 폴리머라제 프로모터의 제2 폴리뉴클레오티드 가닥은 유도제, 예를 들어 대장균 DNA 폴리머라제 I, 또는 대장균 DNA 폴리머라제의 클레나우(Klenow) 단편에 의해 완결된다. 개시 폴리뉴클레오티드는 RNA 폴리뉴클레오티드와 DNA 폴리뉴클레오티드의 교대 생산에 의해서 증폭된다.

프라이머는 또한 RNA-지시된 RNA 폴리머라제에 대한 주형 서열 또는 복제 개시 부위를 함유할 수도 있다. 전형적인 RNA-지시된 RNA 폴리머라제로는 문헌 [Lizardi, et al., Biotechnology, 6:1197-1202 1988]에 기재된 QB 레플리카제 등이 있다. RNA-지시된 폴리머라제는 주형 서열 또는 복제 개시 부위를 함유하는 적은 수의 주형 RNA 가닥으로부터 많은 수의 RNA 가닥을 생산한다. 문헌 [Kramer, et al., J. Mol. Biol., 89:719-736 (1974)]에 기재된 바와 같이, 이들 폴리머라제는 전형적으로 주형 가닥을 100만배 증폭시킨다.

폴리뉴클레오티드 프라이머는, 예를 들어 포스포트리에스테르 또는 포스포디에스테르 방법 등과 같은 임의의 적합한 방법을 이용하여 제조될 수 있다. 문헌 [Narang, et al., Meth. Enzymol., 68:90, (1979)], 미국 특허 제4,356,270호, 동 제 4,458,066호, 동 제4,416,988호, 동 제4,293,652호 및 문헌 [Brown, et al., Meth. Enzymol., 68:109 (1979)]을 참조한다.

프라이머의 뉴클레오티드 서열 선택은, 혼성화 지점에서부터 검출될 돌연변이를 코딩하는 영역까지의 핵산 거리, 사용될 임의의 제2 프라이머가 핵산에서 혼성화되는 부위 등과 같은 인자에 따라 달라진다.

PCR 증폭에 의해 핵산 샘플 중 올리고아데닐레이트 신테타제 유전자 물질이 풍부해진다면, 증폭될 핵산의 각 코딩 가닥에 대한 2종의 프라이머, 즉 PCR 프라이머 쌍이 사용되어야 한다. 제1 프라이머는 비-코딩 (안티센스 또는 마이너스 또는 상보적) 가닥의 일부가 되고, 플러스 또는 코딩 가닥의 뉴클레오티드 서열에 혼성화된다. 제2 프라이머는 코딩 (센스 또는 플

러스) 가닥의 일부가 되고 마이너스 또는 비-코딩 가닥의 뉴클레오티드 서열에 혼성화된다. 제1 프라이머와 제2 프라이머 중 하나 또는 이들 모두가 엔도뉴클레아제 인식 부위를 한정하는 뉴클레오티드 서열을 함유할 수 있다. 상기 부위는 증폭될 올리고아데닐레이트 신테타제 유전자에 이중될 수 있다.

한 실시양태에서, 본 발명은 3'-말단에 프라이밍 영역이 위치한 프라이머를 형성하는 폴리뉴클레오티드 세트를 사용한다. 상기 프라이밍 영역은 전형적으로 3'-끝쪽의 (3'-말단) 15개 내지 30개 뉴클레오티드 염기이다. 각 프라이머의 3'-말단 프라이밍 부위는 핵산 합성을 촉매 작용할 수 있는, 즉 그의 3'-말단의 프라이머 신장 반응을 개시할 수 있는 프라이머로서 작용할 수 있다. 상기 프라이머 중 하나 또는 이들 모두는 5'-말단 (5'-끝쪽의) 비-프라이밍 부위, 즉 바람직한 주형과의 혼성화에 관여하지 않는 영역을 추가로 함유할 수 있다.

PCR에서, 각 프라이머는 제2 프라이머와 함께 작용하여 표적 핵산 서열을 증폭시킨다. PCR에 사용할 PCR 프라이머 쌍은, 올리고아데닐레이트 신테타제 유전자 영역을 생산하는 것과 관련하여 본원에서 논의한 것을 고려하여 선택한다. 올리고아데닐레이트 신테타제 유전자의 대립유전자 인트론 내에서 표적 서열과 혼성화(어닐링)되는 프라이머 서열을 선택한다면, 상기 표적 서열이 대립유전자 중에서 보존되어 있어야만 검정될 표적 서열의 생성을 보장할 수 있다.

### 폴리머라제 연쇄 반응

올리고아데닐레이트 신테타제 유전자는 게놈 DNA의 폴리뉴클레오티드 코딩 가닥, 예를 들어 mRNA 및(또는) 센스 가닥으로 구성된다. 검정될 유전자 물질이 이중 가닥 게놈 DNA의 형태인 경우에는 우선 단일 가닥으로 변성, 전형적으로는 단일 가닥으로 용융시키는 것이 통상적이다. 샘플에 각각이 미리 선택된 뉴클레오티드 서열을 갖는 PCR 프라이머 쌍을 처리(접촉)함으로써 핵산을 PCR 반응시킨다. PCR 프라이머 쌍은 올리고아데닐레이트 신테타제 대립유전자 내에서 보존된 뉴클레오티드 서열, 바람직하게는 약 10개 이상 길이의 뉴클레오티드, 더욱 바람직하게는 약 20개 이상 길이의 뉴클레오티드와의 혼성화에 의해 프라이머 신장 반응을 개시할 수 있다. PCR 프라이머 쌍의 제1 프라이머는 핵산의 비-코딩 가닥 또는 안티센스 가닥, 즉 코딩 가닥에 상보적인 가닥과 혼성화되기 때문에 본원에서는 종종 "안티센스 프라이머"라고 지칭한다. PCR 프라이머 쌍의 제2 프라이머는 핵산의 코딩 가닥 또는 센스 가닥과 혼성화되기 때문에 본원에서는 종종 "센스 프라이머"라고 지칭한다.

PCR 반응은 PCR 완충액 중에서 PCR 프라이머 쌍, 바람직하게는 그의 소정량을 샘플의 핵산, 바람직하게는 그의 소정량과 혼합하여 PCR 반응 혼합물을 형성함으로써 수행된다. 상기 혼합물을, 전형적으로 PCR 반응 생성물의 형성에 충분한 것으로 미리 정해져 있는 수많은 회수의 열주기에 적용시켜서, 검정될 샘플 중에 올리고아데닐레이트 신테타제 유전자 물질이 풍부해지도록 한다.

전형적으로, PCR은 PCR 반응 혼합물의 온도를 하한이 약 30 섭씨 온도 (30°C) 내지 약 55°C이고 상한이 약 90°C 내지 약 100°C인 온도 범위 내에서 열주기에 적용, 즉 반복적으로 증가 및 감소시켜서 수행한다. 이러한 온도 증가 및 감소는 연속적일 수 있지만, 폴리뉴클레오티드 합성, 변성 및 혼성화에 유리한 각 온도에서 상대적으로 온도가 안정화되는 시간을 갖는 기간을 두는 것이 바람직하다.

각 증폭에 복수 종의 제1 프라이머 및(또는) 복수 종의 제2 프라이머가 사용될 수 있으며, 예를 들어 1종의 제1 프라이머가 여러 종의 수많은 제2 프라이머와 쌍을 이루어 여러 종의 상이한 프라이머 쌍을 형성할 수 있다. 별법으로, 제1 프라이머와 제2 프라이머의 개별 쌍을 사용할 수 있다. 어떠한 경우에서도, 제1 프라이머와 제2 프라이머의 동일하거나 상이한 조합을 사용하여 증폭시킨 증폭 생성물을 합하여 돌연변이에 대한 검정을 수행할 수 있다.

PCR 반응은 임의의 적합한 방법을 이용하여 수행한다. 일반적으로, 이것은 완충 수용액, 즉 PCR 완충액, 바람직하게는 pH 7 내지 9, 가장 바람직하게는 약 pH 8의 완충액 중에서 수행한다. 주형 가닥을 함유하는 완충액에 프라이머를 몰 과량(게놈 핵산에 대한 양, 일반적으로는 프라이머 : 주형이 약  $10^6 : 1$ )으로 하여 혼합하는 것이 바람직하다. 이 과정의 효율을 개선하기 위해서는 몰 과량이 높은 것이 바람직하다.

또한, PCR 완충액은 데옥시리보뉴클레오티드 트리포스페이트(폴리뉴클레오티드 합성 기질)인 dATP, dCTP, dGTP 및 dTTP와 폴리머라제, 전형적으로는 열안정한 폴리머라제 모두를 프라이머 신장(폴리뉴클레오티드 합성) 반응에 적절한 양으로 함유한다. 생성된 용액(PCR 혼합물)은 약 90°C 내지 100°C로 약 1분 내지 10분, 바람직하게는 1분 내지 4분 동안 가열한다. 이러한 가열 기간 후에는 상기 용액을 프라이머 혼성화에 바람직한 54°C로 냉각시킨다. 합성 반응은 실온 내지 폴리머라제(유도제)가 더이상 효율적으로 기능하지 않는 온도 미만의 온도에서 일어날 수 있다. 원하는 양의 PCR 생성물이 수득될 때까지 열주기를 반복한다. 예를 들어, PCR 완충액은 하기 성분을 포함한다: 50 mM KCl, 10 mM Tris-HCl

(pH 8.3), 1.5 mM MgCl, 0.001% (wt/vol) 젤라틴, 200  $\mu$ M dATP, 200  $\mu$ M dTTP, 200  $\mu$ M dCTP, 200<sup>2</sup>  $\mu$ M dGTP, 및 완충액 100  $\mu$ l 당 2.5 유닛의 썬머스 아쿠아티쿠스(*Thermus aquaticus*) (Taq) DNA 폴리머라제 I (미국 특허 제 4,889,818호).

유도제는 효소 등과 같이 프라이머 신장 생성물의 합성을 수행하는 기능을 할 임의의 화합물 또는 시스템일 수도 있다. 이러한 목적에 적합한 효소로는, 예를 들어 대장균 DNA 폴리머라제 I, 대장균 DNA 폴리머라제 I의 클레나우 단편, T4 DNA 폴리머라제, 기타 이용가능한 DNA 폴리머라제, 역전사효소, 및 열 안정성 효소를 비롯한 기타 효소 등이 있고, 뉴클레오티드들의 조합을 적당한 방식으로 용이하게 하여 각 핵산 가닥에 상보적인 프라이머 신장 생성물이 형성된다. 일반적으로, 합성은 각 프라이머의 3'-말단에서 개시될 것이고 주형 가닥을 따라 합성이 종결될 때까지 5' 방향으로 진행되어 여러가지 길이의 분자를 생성할 것이다. 그러나, 5'-말단에서 합성을 개시하여 상기한 것과 동일한 과정을 통해 상기한 방향으로 진행되는 유도제도 있을 수 있다. 유도제는 또한 효소 등과 같이 RNA 프라이머 신장 생성물의 합성을 수행하는 기능을 할 화합물 또는 시스템일 수도 있다. 바람직한 실시양태에서, 상기 유도제는 DNA-의존성 RNA 폴리머라제, 예를 들어 T7 RNA 폴리머라제, T3 RNA 폴리머라제 또는 SP6 RNA 폴리머라제일 수도 있다. 이들 폴리머라제는 상보적 RNA 폴리뉴클레오티드를 생성한다. 문헌 [Chamberlin, et al., *The Enzymes*, ed. P. Boyer, pp. 87-108, Academic Press, New York (1982)]에 기재된 바와 같이, RNA 폴리머라제의 높은 전환율은 개시 폴리뉴클레오티드를 증폭시킨다. 전사 기재의 증폭 시스템은 문헌 [Gingeras, et al., in *PCR Protocols, A Guide to Methods and Applications*, pp. 245-252, Innis, et al., eds, Academic Press, Inc., San Diego, Calif. (1990)]에 기재되어 있다.

유도제가 DNA-의존성 RNA 폴리머라제여서 리보뉴클레오티드 트리포스페이트를 혼입하는 경우에는 충분한 ATP, CTP, GTP 및 UTP를 프라이머 신장 반응 혼합물에 혼합하고, 생성되는 용액은 상기한 바와 같이 처리한다.

새로 합성된 가닥 및 그의 상보적 핵산 가닥은, 상기 과정의 이후 단계에 사용될 수 있는 이중-가닥 분자를 형성한다.

PCR 반응은 상기 생성물에 올리고아데닐레이트 신테타제 유전자 중의 돌연변이를 검출하는데 유용한 미리 선택된 제한 부위를 혼입하는데 유리하게 이용될 수 있다.

PCR 증폭 방법은 미국 특허 제4,683,192호, 동 제4,683,202호, 동 제4,800,159호 및 동 제4,965,188호, 및 적어도 문헌 ([*PCR Technology: Principles and Applications for DNA Amplification*, H. Erlich, ed., Stockton Press, New York (1989)] 및 [*PCR Protocols: A Guide to Methods and Applications*, Innis, et al., eds., Academic Press, San Diego, Calif. (1990)]) 등을 비롯한 여러 교재에 상세하게 기재되어 있다.

몇몇 실시양태에서, 각 증폭 반응마다 제1 프라이머와 제2 프라이머의 2개 쌍이 사용된다. 복수회의 상이한 증폭시에 각각 복수 중의 상이한 프라이머 쌍을 사용하여 수득한 증폭 반응 생성물은 한데 합하여 검정할 수도 있고, 따로 검정할 수도 있다.

그러나, 본 발명은 제1 프라이머와 제2 프라이머의 쌍을 1종만 사용하여 증폭시키는 것을 고려하고 있다. 본원에 개시된 돌연변이를 함유하는 DNA 구역을 증폭시킬 수 있는 프라이머의 예는 하기 표 1에 나타내었다. 앰플리콘(Amplicon) A는 서열 1 내지 3에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다. 앰플리콘 B는 서열 4 내지 7 및 서열 60에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다. 앰플리콘 C는 서열 57에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다. 앰플리콘 D는 서열 58에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다. 앰플리콘 E는 서열 59에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다. 앰플리콘 F는 서열 61에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다. 앰플리콘 G는 서열 62 내지 64에서 언급된 돌연변이를 함유하는 폴리뉴클레오티드 서열에 상응한다.

[표 1]

본 발명의 돌연변이를 함유하는 앰플리콘

앰플리콘	프라이머 A	프라이머 B	생성물 크기 (bp)
앰플리콘 A	5'- AATGGACCTCAAGACTTCCC- 3' (서열 8)	5'- ATTCTCCCTTCTGTTGCAGG- 3' (서열 9)	509
앰플리콘 B	5'- TCCAGATGGCATGTCACAGT- 3' (서열 10)	5'- GAGCTATGCTTGGCACATAG- 3' (서열 11)	747
앰플리콘 C	5'- CACAAGAGTGAACCTTAATGT- 3' (서열 65)	5'- CCAGGAAGTGAAAGATCAT- 3' (서열 66)	603
앰플리콘 D	5'- ATCTCCACAGTTTGAGAGC- 3' (서열 67)	5'- TCAGCCTCCAAAAGTGTGG- 3' (서열 68)	553
앰플리콘 E	5'- GGGTACATGTGCACAATGTG- 3' (서열 69)	5'- CCCTTATACAAAATTCAACTC -3' (서열 70)	532
앰플리콘 F	5'- GAGCCAAGAAGTACAGATGC -3' (서열 71)	5'- AGGACAGAGCTGTCCAATAG -3' (서열 72)	648
앰플리콘 G	5'- GGCTCAGAGAAGCTAAGTGA -3' (서열 73)	5'- CCACAGCATCCTTTTCAGTC - 3' (서열 74)	581

하기 표 2는 상기 앰플리콘 중 본 발명의 돌연변이의 위치를 개시한다.

[표 2]

애플리콘 중 본 발명의 돌연변이의 위치

돌연변이	애플리콘	애플리콘 중의 위치 (애플리콘의 프라이머A쪽의 5' 말단에 대한 위치)
1 (서열 1)	애플리콘 A	134
2 (서열 2)	애플리콘 A	155
3 (서열 3)	애플리콘 A	384
4 (서열 4)	애플리콘 B	98
5 (서열 5)	애플리콘 B	114
6 (서열 6)	애플리콘 B	142
7 (서열 7)	애플리콘 B	347
8 (서열 57)	애플리콘 C	319
9 (서열 58)	애플리콘 D	404
10 (서열 59)	애플리콘 E	133
11 (서열 60)	애플리콘 B	320
12 (서열 61)	애플리콘 F	367
13 (서열 62)	애플리콘 G	138
14 (서열 63)	애플리콘 G	210
15 (서열 64)	애플리콘 G	253

핵산 서열 분석

핵산 서열 분석은, (a) 프로브 가닥 및 그의 상보적 표적의 혼성화 또는 변성을 기재로 하는 물리화학적 기술, 및 (b) 엔도 뉴클레아제, 리가제 및 폴리머라제와의 효소 반응의 조합에 의해 수행된다. 핵산은 DNA 수준 또는 RNA 수준에서 검정될 수 있다. DNA 수준에서의 핵산 검정은 개개의 인간의 유전적 잠재성을 분석하고, RNA 수준에서의 핵산 검정은 특정 세포의 발현된 정보를 분석한다.

핵산 혼성화를 이용한 검정법에서, 본 발명의 방법에서 DNA 이중나선의 존재를 검출하는 것은 여러가지 수단에 의해 수행될 수 있다.

DNA 이중나선의 존재를 검출하는 한 접근법에서, 상기 DNA 이중나선에서 혼성화되는 올리고뉴클레오티드는 상기 이중나선을 검출가능하게 하는 표지 또는 지시기를 포함한다. 전형적으로, 이러한 표지는 방사성 원자, 화학적 개질된 뉴클레오티드 염기 등을 포함한다.

올리고뉴클레오티드는 지시 수단 또는 지시기에 표지되어, 즉 작동가능하게 연결되어, 표적 주형 중 특정 뉴클레오티드 서열의 존재를 검출하는데 사용될 수 있다.

올리고뉴클레오티드 프로브 (표지된 올리고뉴클레오티드)에 작동가능하게 연결되거나 그의 일부로서 존재하는 방사성 요소는 DNA 이중나선의 검출을 용이하게 하는 유용한 수단을 제공한다. 전형적인 방사성 요소는 베타선을 방출하는 것이다. 베타선을 방출하는 요소, 예를 들어 <sup>3</sup>H, <sup>12</sup>C, <sup>32</sup>P 및 <sup>35</sup>S는 베타선 방출-생성 방사성 요소 표지류를 대표한다. 전형적으로, 방사성 폴리뉴클레오티드 프로브는 방사성 표지된 뉴클레오티드를 DNA 키나제를 사용하여 핵산에 효소적 혼입시켜서 제조된다.

방사성 표지된 올리고뉴클레오티드에 대한 대안은, 금속 착화제, 바이오틴-함유기, 형광 화합물 등을 함유하도록 화학적 개질된 올리고뉴클레오티드이다.

유용한 금속 착화제 중 하나는 란탄계 원소 및 방향족 베타-디케톤에 의해 형성된 란탄계 킬레이트이고, 여기서는 란탄계 원소가 EDTA-유사체 등과 같은 킬레이트-형성 화합물을 통해 핵산 또는 올리고뉴클레오티드에 결합되어 형광 란탄계 착물이 형성된다. 미국 특허 제4,374,120호, 동 제4,569,790호 및 공개된 특허 출원서 EP 0139675 및 WO 87/02708을 참조한다.

바이오틴 또는 아크리딘 에스테르-표지된 올리고뉴클레오티드 및 폴리뉴클레오티드를 표지하는데 있어서의 이들의 용도는 기재된 바 있다. 미국 특허 제4,707,404호, 공개된 특허 출원서 EP 0212951 및 유럽 특허 제0087636호를 참조한다. 유용한 형광 마커 화합물로는 플루오레세인, 로다민, 텍사스 레드(Texas Red), NBD 등이 있다.

DNA 이중나선에 존재하는 표지된 올리고뉴클레오티드는 이중나선 자체가 표지되도록 하기 때문에, 검정될 샘플에 존재하는 다른 핵산과 구별될 수 있다. 전형적으로, 이중나선 중의 표지의 존재를 검출하여 이중나선의 존재를 검출하는 것은, DNA 이중나선을 상기 DNA 이중나선과 혼성화되지 않는 임의의 표지된 올리고뉴클레오티드 프로브로부터 분리하는 것을 포함한다.

단일 가닥 올리고뉴클레오티드, 예를 들어 혼성화되지 않는 표지된 올리고뉴클레오티드 프로브를 DNA 이중나선으로부터 분리하는 기술은 공지되어 있고, 전형적으로 이중 가닥 핵산으로부터 단일 가닥 핵산을 이들의 화학적 성질을 기초로 하여 분리하는 것을 포함한다. 더욱 혼한 분리 기술은, 불용성 매트릭스에 결합된 DNA 이중나선으로부터 혼성화되지 않는 프로브를 전형적으로는 세척에 의해 분리시키는 이중 혼성화 포맷의 사용을 포함한다. 이의 예로는 써던 블롯 기술이 있고, 여기서 상기 매트릭스는 니트로셀룰로스 시트이고 상기 표지는  $^{32}\text{P}$ 이다 [Southern, J. Mol. Biol., 98:503 (1975)].

또한, 올리고뉴클레오티드는 전형적으로는 5'-말단 또는 그 근처에서 고체 매트릭스, 즉 수불용성 고체 지지체에 연결되는 것이 유리할 수 있다. 유용한 고체 매트릭스는 당업계에 공지되어 있고, 가교된 텍스트란, 예를 들어 파마시아 파인 케미칼스(Pharmacia Fine Chemicals) (미국 뉴저지주 피스카타웨이 소재)가 세파덱스(SEPHADEX)라는 상표명으로 시판하는 것; 아가로스, 폴리스티렌 또는 라텍스 비드 (약 1 미크론 내지 약 5 밀리미터 직경), 폴리비닐 클로라이드, 폴리스티렌, 가교된 폴리아크릴아미드, 니트로셀룰로스 또는 나일론-기재의 웹, 예를 들어 시트, 스트립, 패들, 플레이트, 미량역가 플레이트 웰 등이 있다.

또한, 상기한 올리고뉴클레오티드의 5'-말단 또는 3'-말단에 "연결" 뉴클레오티드를 부가하고, 이 연결 뉴클레오티드를 사용하여 상기 올리고뉴클레오티드를 고체 지지체에 작동가능하게 연결시킬 수도 있다.

뉴클레오티드 혼성화 검정시에, 상기에서 고려한 방법에서의 혼성화 반응 혼합물은, 주형상의 예정된 서열에 상보성을 갖는 올리고뉴클레오티드가 주형 내 존재하는 상보적 핵산 서열에 혼성화되어 혼성화 생성물, 즉 올리고뉴클레오티드 및 표적 핵산을 함유하는 복합체를 형성하는데 충분한 시간 동안 혼성화 조건하에서 유지된다.

"혼성화 조건"이라는 어구 및 이것과 문법적으로 동등한 표현이 유지 시간과 함께 사용되는 경우에는, 혼성화 반응 혼합물을 혼합물 중의 반응물질 및 공존하는 시약의 농도에서 1개 이상의 올리고뉴클레오티드가 표적 서열과 어닐링되어 핵산 이중나선이 형성되기에 충분한 시간, 온도 및 pH 조건에 적용하는 것을 지시한다. 혼성화 수행에 요구되는 이러한 시간, 온도 및 pH 조건은 당업계에 공지되어 있는 것과 같이 혼성화될 올리고뉴클레오티드의 길이, 올리고뉴클레오티드와 표적 사이의 상보성 정도, 올리고뉴클레오티드의 구아닌과 시토신 함량, 원하는 혼성화 엄격도, 및 혼성화 반응 혼합물 중 염 또는 추가 시약의 존재 여부 등에 따라 달라지고, 이것은 혼성화 역학에 영향을 미칠 수 있다. 주어진 혼성화 반응 혼합물에 대하여 혼성화 조건을 최적화하는 방법은 공지되어 있다.

전형적인 혼성화 조건은 pH 값이 4 내지 9로 완충된 용액의 사용을 포함하고, 4°C 내지 37°C, 바람직하게는 약 12°C 내지 약 30°C, 더욱 바람직하게는 약 22°C의 온도에서 0.5초 내지 24시간, 바람직하게는 2분 내지 1시간의 기간 동안 수행된다.

혼성화는 당업계에 공지된 것과 같은 동중 또는 이중 포맷으로 수행될 수 있다. 동중 혼성화 반응은 전적으로 용액 중에서 수행되며, 이때 올리고뉴클레오티드 및 혼성화될 핵산 서열 (표적) 모두가 용액 중에서 가용성 형태로 존재한다. 이중 반응은 반응 매질 중에서 불용성이고 올리고뉴클레오티드, 폴리뉴클레오티드 프로브 또는 표적 핵산이 결합되는 매트릭스의 사용을 포함한다.

표적 서열을 함유하는 핵산이 이중 가닥 (ds) 형태인 경우, 가열 또는 알칼리 처리 등에 의해 우선 dsDNA를 변성시킨 후에 혼성화 반응을 수행하는 것이 바람직하다. dsDNA의 변성은, 혼성화될 올리고뉴클레오티드와의 혼합 이전에 수행될 수도 있고, dsDNA와 올리고뉴클레오티드의 혼합 이후에 수행될 수도 있다.

올리고뉴클레오티드와 주형 사이의 예정된 상보성은 2가지 대안적인 방식으로 달성된다. 예를 들어, 형성될 프라이머가 공지된 올리고아데닐레이트 신테타제 서열에 혼성화될 수 있고 본원에 기재한 바와 같은 서열분석 목적 및 이후의 검정 목적을 위해서 DNA 영역으로의 프라이머 신장을 개시할 수 있거나, 또는 서열분석을 미리 수행하여 뉴클레오티드 서열의

영역을 결정하고 이렇게 새로이 서열분석된 영역으로부터 미지 서열의 영역으로 신장되는 프라이머를 고안하는 경우에는 주형 DNA의 서열이 공지되어 있을 수 있다. 후자의 과정은 각 회의 서열분석이 미리 결정해 둔 서열을 기초로 하여 고안된 프라이머에 의해 지시되기 때문에 "지시된 서열분석"이라고 지칭되어 왔다.

혼성화 반응 혼합물 중에 존재하는 올리고뉴클레오타이드의 유효량은 일반적으로 공지되어 있으며, 전형적으로는 혼성화될 올리고뉴클레오타이드와 주형 사이의 몰비로 표현된다. 바람직한 비율은 표적 서열과 올리고뉴클레오타이드를 등량 함유하는 혼성화 반응 혼합물이다. 공지된 바와 같이, 동등한 몰농도가 변하여도 비록 혼성화 효율은 더 낮겠지만 혼성화 반응 생성물이 생성될 것이다. 따라서, 한 성분이 다른 성분보다 100배 몰 과량이나 되는 많은 양으로 존재할 수 있지만, 50배 미만, 바람직하게는 10배 미만, 더욱 바람직하게는 2배 미만으로 초과하는 것이 본 발명의 실시예에 바람직하다.

막-고정화된 표적 서열의 검출

DNA (씨던) 블롯 기술에서는, 앞서 논의한 바와 같은 PCR 증폭으로 DNA를 제조한다. PCR 생성물 (DNA 단편)은 아가로스 겔에서 크기에 따라 분리되고 니트로셀룰로스 또는 나일론 막으로 전달 (블롯팅)된다. 통상의 전기영동은 100 내지 30,000개 염기쌍 범위의 단편을 분리하고, 펄스 장(pulsed field) 겔 전기영동은 2천만개 이하 길이의 염기쌍 단편을 분리한다. 막에서 특정 PCR 생성물이 함유된 위치는 표지된 특이적 핵산 프로브와의 혼성화로 결정한다.

바람직한 실시양태에서, 도트-블롯 (슬롯-블롯) 장치를 사용하여 PCR 생성물을 고체-매트릭스 (니트로셀룰로스 막)에 직접 고정화시키고, 프로브-혼성화에 의해 분석한다. 미국 특허 제4,582,789호 및 동 제4,617,261호를 참조한다.

고정화된 DNA 서열은, 뉴클레오타이드 길이가 대략 15개, 17개, 20개, 25개 또는 약 30개 이하의 합성 DNA 올리고머인 대립유전자-특이적 올리고뉴클레오타이드 (ASO) 프로브로 프로빙하여 분석할 수 있다. 이들 프로브는 게놈 내의 유일 서열을 대표할 만큼 충분히 길지만, 표적 분자와의 혼성화시에는 내부 미스매치(mismatch)에 의해 불안정화될 만큼 충분히 짧다. 따라서, 주의깊게 제어되는 혼성화 조건하에서는 ASO 프로브와 정상 표적 또는 돌연변이된 표적 사이에 형성된 하이브리드의 변성 거동이 상이하기 때문에 단일 뉴클레오타이드가 상이한 임의의 서열을 구별할 수 있다. 프로브의 예는 본원에서 서열 1 내지 7 및 서열 57 내지 64로 개시되어 있으나 (표 3), 선택된 점 돌연변이를 보유하는 OAS1 유전자 영역에 특이적으로 혼성화되고 점 돌연변이를 보유하는 폴리뉴클레오타이드와 야생형 폴리뉴클레오타이드를 특이적으로 구별할 수 있는 것이라면 임의의 프로브라도 적합하다.

용액 중 표적 서열의 검출

핵산 정제 또는 고정화가 필요 없는 몇가지 신속한 기술이 개발되어 왔다. 예를 들어, 프로브/표적 하이브리드는 이중-가닥 핵산을 우선적으로 결합하는 고체 매트릭스, 예를 들어 히드록실아파타이트상에서 선택적으로 단리될 수 있다.

별법으로, 프로브 핵산은 고체 지지체상에 고정화될 수 있고 용액으로부터 표적 서열을 포획하는데 이용될 수 있다. 표적 서열의 검출은, 표적 서열에 의해 지지체로부터 대체된 제2의 표지된 프로브를 사용한 경쟁형 검정으로 수행되거나, 또는 표적 서열의 연결 작용을 통해 지지체에 연결된 제2의 표지된 프로브를 사용한 샌드위치형 포맷으로 수행될 수 있다.

올리고뉴클레오타이드 라이게이션 검정 (OLA)에서는, 효소 DNA 리가제를 사용하여 정확한 헤드-대-테일(head-to-tail) 근위에서 표적 서열과 염기쌍을 형성할 수 있도록 선택된 2종의 합성 올리고뉴클레오타이드 서열을 공유 결합시킨다. 2종 올리고머의 라이게이션은 연결 영역에 미스매치된 뉴클레오타이드가 존재하면 방해받는다. 이러한 방법은 DNA 정제 없이도 세포 샘플 중에 존재하는 공지된 서열 변이체들 사이를 구별할 수 있게 한다. 2종 올리고뉴클레오타이드들의 연결은 2종 올리고뉴클레오타이드 중 하나를 고정화하고 제2의 표지된 올리고뉴클레오타이드 역시 포획되는지 여부를 관찰함으로써 모니터링할 수 있다.

염기 치환 검출을 위한 스캐닝 기술

미지의 단일-뉴클레오타이드 치환 또는 기타 서열 상이성이 수백 염기쌍 길이인 프로브/표적 이중나선을 분석할 수 있는 3가지 기술이 있다. RN아제 A 기술에서, 상기 효소는 표지된 RNA 프로브를 표적 RNA 서열 또는 표적 DNA 서열과 미스매치된 위치에서 절단한다. 상기 단편은 크기에 따라 분리될 수 있어서, 돌연변이의 대략적인 위치를 결정할 수 있다. 미국 특허 제4,946,773호를 참조한다.

변성 구배 겔 기술에서, 프로브-표적 DNA 이중나선은 강도가 증가하는 변성 구배에서의 전기영동으로 분석된다. 변성은 이동 속도에서의 감소를 수반한다. 미스매치된 염기쌍을 갖는 이중나선은 완벽하게 매치된 이중나선보다 더 신속하게 이동한다.

3번째 방법은 미스매치된 염기쌍의 화학적 절단에 의존한다. T와 C, G 또는 T와의 미스매치 및 C와 T, A 또는 C와의 미스매치는 헤테로이중나선에서 검출될 수 있다. 사산화오스튬과의 반응 (T 및 C 미스매치) 또는 히드록실아민과의 반응 (C 미스매치) 후에 피페리딘을 처리하면, 적절한 미스매치에서 프로브가 절단된다.

OAS1 기능의 회복 및(또는) 향상을 위한 치료제

OAS1 유전자에서의 돌연변이가 OAS1 기능의 결함을 야기하고, 상기 기능 결함이 OAS1 단백질의 보다 낮은 수준, 그의 기능에 영향을 주는 단백질에서의 돌연변이를 통해서든, 또는 다른 메카니즘을 통해서든 병원체 감염에 대한 환자의 감수성 증가와 연관되는 경우, 환자를 야생형 OAS1 단백질로 치료하는 것이 유리할 수 있다. 또한, 돌연변이가 감염-저항성 담체에서, 야생형 단백질과 상이하며 HCV 감염 억제 면에서 장점을 갖는 단백질의 형태를 발생시킨다면, 돌연변이된 유전자에 의해 코딩되는 단백질을 투여하는 것이 유리할 수 있다. 상기 기재한 바와 같이, OAS1 단백질 또는 폴리펩티드의 천연 또는 돌연변이 형태의 투여는 또한 암, 당뇨병, 및 상처 치유를 포함하나 이에 제한되지 않는 다른 징후의 치료에 유리할 수도 있다. 하기의 논의는 상기 단백질 또는 폴리펩티드 중 임의의 것의 투여에 관한 것이다.

OAS1R에 의해 코딩되는 폴리펩티드를 비롯한 본 발명의 폴리펩티드는 자연 정제된 생성물, 또는 화학 합성 절차의 생성물 또는 본 발명의 폴리뉴클레오티드 서열의 원핵생물 또는 진핵생물 숙주로부터 (예를 들어, 배양 중인 박테리아, 효모, 고등 식물, 곤충 및 포유동물 세포에 의해) 재조합 기술로 생산된 생성물일 수 있다. 재조합 생산 절차에서 사용되는 숙주에 따라, 본 발명의 폴리펩티드는 포유동물 또는 다른 진핵생물의 탄수화물로 글리코실화될 수 있거나 글리코실화되지 않을 수 있다. 본 발명의 폴리펩티드는 또한 처음의 메티오닌 아미노산 잔기 (마이너스 1 위치에서)를 포함할 수 있다.

본 발명의 폴리펩티드는 또한 서열 20, 서열 21, 서열 22, 서열 23, 서열 24 및 서열 25, 서열 26, 서열 27, 서열 28, 서열 29, 서열 32, 서열 33, 서열 34, 서열 35, 서열 46, 및 서열 47로 정의된 단백질 서열 및 그의 유도체를 포함한다.

폴리펩티드(들)의 자연 발생 대립유전자 형태 이외에, 본 발명은 자연 발생 대립유전자 형태와 유사하게 기능하는 폴리펩티드의 동족체 및 단편도 또한 포함한다. 따라서, OAS1R 단백질의 기능이 유지되는 한, 예를 들어 폴리펩티드 중의 하나 이상의 아미노산 잔기를 보존된 아미노산 잔기로 치환할 수 있다. 하기 실시예 8 내지 10은 본 발명의 폴리펩티드에 대한 적합한 아미노산 치환의 대표적인 예를 제시한다. 다른 예로서, 본 발명의 폴리펩티드는 구체적으로 서열 48, 서열 49, 서열 50, 서열 51 및 서열 52로 정의된 OAS1R의 말단절단형 또는 동족체 형태를 포함한다. 상기 논의한 바와 같이, 서열 51은 침팬지가 가진 OAS1의 단축 형태를 나타내고, 서열 52는 고릴라가 가진 더 길지만 여전히 말단절단된 형태의 카르복실-말단 단편을 나타낸다. 서열 49 및 서열 50은 각각 상응하는 침팬지 및 고릴라 말단절단 부위로 말단 절단된 합성 인간 OAS1R 구조물을 나타낸다. 서열 48은 침팬지 및 고릴라 형태에 대한 중간 길이로 말단 절단된 합성 인간 OAS1R을 나타낸다. 서열 48은 본원의 다른 곳에 개시한 바와 같이 당업계에 공지된 방법에 의해 효소적으로 활성임이 추가로 입증되어 있다. 상응하게, 나머지 매우 유사한 말단절단된 형태가 효소적으로 활성임이 입증될 수 있다. 당업자는 인식하겠지만, OAS1R 폴리펩티드의 말단절단되었으나 기능적인 형태의 치료적 사용은 폴리펩티드의 치료적 효과를 달리 방해할 수 있는 항체 반응의 발달을 봉쇄할 수 있다. 상기한 말단절단된 폴리펩티드, 및 당업자가 상상할 수 있는 다른 것들은 기능을 유지하지만, 내생적으로 전장 OAS1R 폴리펩티드를 갖지 않는 개인에서 항체 반응을 유도할 수 있는 폴리펩티드의 비-보편적 부분을 제거한다. 당업자는 또한 일반적으로 보다 작은 폴리펩티드가 치료적 개발에서 전형적으로 맞닥뜨려지는 제조의 복잡성, 전달 및 클리어런스에 대해 보다 융통성있다는 것을 인식할 것이다. 추가로, 당업자는 침팬지 및 고릴라에서 별개의 동질접합 말단절단 변이체의 발생이 본원에 개시된 말단절단된 OAS1 형태의 광범위 항바이러스 효능을 위해 매우 암시적임을 인식할 것이다. 상기 구체적으로 개시된 말단절단된 폴리펩티드 형태는 폴리펩티드의 카르복실-말단에 대한 말단절단을 나타내지만, 본 발명은 단편의 형태에 의해 제한되지 않으며 구체적으로는 효소 기능을 보유하는 아미노-말단 말단절단 및 내부 아미노산 결실을 포함한다.

또한 본 발명의 범위 내에는, 특정한 개시된 폴리펩티드의 하나 이상의 활성을 보유하지만, 개시된 아미노산 서열과는 상이한 폴리펩티드도 또한 포함된다. 이러한 폴리펩티드는 바람직하게는 표준 정렬 방법을 이용하여 계산된 바와 같은 상응하는 개시된 서열에 대해 80% 이상의 서열 상동성, 바람직하게는 85% 이상의 서열 상동성, 더욱 바람직하게는 90% 이상의 서열 상동성, 가장 바람직하게는 95% 초과 서열 상동성을 갖는다.

폴리펩티드는 또한 생체내 이러한 폴리펩티드의 발현에 의해 본 발명에 따라 사용할 수 있는데, 이는 종종 유전자 요법이라 언급된다. 따라서, 예를 들어, 세포는 폴리펩티드를 코딩하는 폴리뉴클레오티드 (DNA 또는 RNA)를 이용하여 생체의 도입될 수 있고, 그 후 상기 도입된 세포를 폴리펩티드로 치료할 환자에게 제공할 수 있다. 이러한 방법은 당업계에 공지되어 있다. 예를 들어, 세포는 본 발명의 폴리펩티드를 코딩하는 RNA를 함유하는 레트로바이러스 입자를 사용하여 당업계에 공지된 절차에 의해 도입될 수 있다.

유사하게, 세포의 도입은, 예를 들어 당업계에 공지된 절차로, 생체내 폴리펩티드의 발현을 위해서 생체내에서 달성할 수 있다. 당업계에 공지된 바와 같이, 본 발명의 폴리펩티드를 코딩하는 RNA를 함유하는 레트로바이러스 입자를 생산하기 위한 생산자 세포는 생체내 도입 및 생체내 폴리펩티드의 발현을 위해 환자에 투여될 수 있다.

이러한 방법으로 본 발명의 폴리펩티드를 투여하기 위한 상기 방법 및 다른 방법은 본 발명의 교시내용으로부터 당업자에게 명백해야 한다. 예를 들어, 세포를 도입시키기 위한 발현 비히클은 레트로바이러스 이외의 것일 수 있고, 예를 들어, 적합한 전달 비히클과 조합한 후 생체내에서 세포를 도입하기 위해 사용할 수 있는 아데노바이러스일 수 있다.

폴리펩티드가 액체 제형으로서 제조되어 주사로 투여되는 경우, 바람직하게는 용액은 140 밀리몰 염화나트륨 및 10 밀리몰 칼슘을 함유하는 pH 7.4인 등장 염 용액이다. 주사액은 투여 경로, 환자의 건강 등을 고려하여, 예를 들면 치료적 유효량으로, 바람직하게는 1일 약 1  $\mu\text{g}/\text{kg}$  체중 내지 약 5  $\text{mg}/\text{kg}$  체중의 용량으로 투여될 수 있다.

본 발명의 폴리펩티드(들)은 적합한 제약 담체와 함께 사용될 수 있다. 이러한 조성물은 치료적 유효량의 단백질, 및 제약상 허용되는 담체 또는 부형제를 포함한다. 이러한 담체로는 염수, 완충화된 염수, 텍스트로스, 물, 글리세롤, 에탄올 및 그의 조합물이 포함되나 이에 제한되지 않는다. 상기 제형은 투여 방식에 적합해야 한다.

또한, 본 발명의 폴리펩티드(들)은 폴리펩티드를 하나 이상의 부위 또는 접합체와 화학적으로 결합시킴으로써 변형되어, 폴리펩티드(들)의 활성, 세포 분포, 또는 세포 업테이크를 향상시킬 수 있다. 이러한 부위 또는 접합체에는 지질, 예를 들어 콜레스테롤, 콜산, 티오에테르, 지방족 쇠, 인지질 및 그의 유도체, 폴리아민, 폴리에틸렌 글리콜 (PEG), 팔미틸 부위 및, 예를 들어 미국 특허 제5,514,758호, 동 제5,565,552호, 동 제5,567,810호, 동 제5,574,142호, 동 제5,585,481호, 동 제5,587,371호, 동 제5,597,696호 및 동 제5,958,773호에 개시된 바와 같은 다른 것이 포함된다.

또한, 본 발명의 폴리펩티드(들)은 변형되어 특정 질환 징후에 대해 특이적인 세포 유형을 표적화할 수 있다 (예를 들어, C형 간염 감염의 경우에는 간 세포를 포함하나 이에 제한되지 않음). 당업자가 인식할 수 있는 바와 같이, 기재된 표적화 목표를 달성하는 적합한 방법이 기재되어 있으며, 여기에는 리포솜 표적화, 수용체-매개의 엔도사이토시스, 및 항체-항원 결합이 포함되나 이에 제한되지 않는다. 한 실시양태에서, 아시알로당단백질 수용체를 사용하여 폴리펩티드(들)에 갈락토스 부위를 첨가함으로써 간 세포를 표적화할 수 있다. 다른 실시양태에서, 대식세포 및 간 세포에서 발견되는 만노스 수용체를 표적화하기 위하여, 만노스 부위를 폴리펩티드(들)에 접합시킬 수 있다. 또한, 본 발명의 폴리펩티드(들)은 엔도솜 회피 메카니즘 또는 단백질 형질도입 도메인 (PTD) 시스템을 포함하나 이에 제한되지 않는 당업자에게 공지된 방법에 의해, 세포질 전달을 위해 변형될 수 있다. PTD 시스템은, 예를 들어 문헌 [Vives E, et al. (1997) J. Biol. Chem. 272: 16010-16017], [Derossi, et al. (1994) J. Biol. Chem. 269: 10444-10450], [Elliott, G et al. (1997) Cell 88:223-233], [Wadia, JS et al. (2004) Nat. Med. 10:310-315], 및 [Kabouridis, PS. (2003) Trends Biotech., 21: 498-503]에 개시되어 있다. 공지된 엔도솜 회피 시스템은 pH-반응성 중합체 담체, 예를 들어 폴리(프로필아크릴산)의 사용을 포함한다. 공지된 PTD 시스템은 천연 펩티드, 예를 들어 HIV-1 TAT, HSV-1 VP22, 드로소필라 안테나페디아(*Drosophila Antennapedia*), 또는 디프테리아 독소에서부터 합성 펩티드 담체에까지 이른다 (Wadia and Dowdy, Cur. Opin. Biotech. 13:52-56, 2002; Becker-Hapak et. al., Methods 24:247-256, 2001). 도 10은 상기 예시적인 PTD 중 몇몇의 상세한 설명을 제공한다. 당업자는 인식하겠지만, 다중 전달 및 표적화 방법을 조합할 수 있다. 예를 들어, 본 발명의 폴리펩티드(들)은 리포솜 내에 캡슐화됨으로써 간 세포로 표적화될 수 있고, 이러한 리포솜은 아시알로당단백질 수용체로의 표적화를 위해 갈락토스로 접합될 수 있다.

또한, 본 발명은 본 발명의 제약 조성물의 하나 이상의 성분으로 충전된 하나 이상의 용기를 포함하는 제약적 팩 또는 키트를 제공한다. 약제 또는 생물학적 제품의 제조사, 용도 또는 판매를 규제하는 정부 기관에 의해 규정된 형태의 통지가 이러한 용기(들)과 연관되어 있고, 상기 통지는 인간 투여를 위한 제조사, 용도 또는 판매에 대한 기관에 의한 승인을 반영한다. 또한, 본 발명의 폴리펩티드는 다른 치료적 화합물과 함께 사용할 수 있다.

본 발명의 OAS1 변이체를 약제로서 사용하는 경우, 이는 적합한 비히클 내에서 포유동물로 투여될 수 있다. 본 발명의 폴리펩티드를 상기 기재된 바와 같은 약제로서 사용하는 경우, 이는 투여 경로, 환자의 건강 등을 고려하여, 예를 들면 1일 약

10 µg/kg 체중 내지 약 4 mg/kg 체중의 치료적 유효 용량으로 투여할 수 있다. 투여량은 바이러스, 바람직하게는 플라비 바이러스, 가장 바람직하게는 HCV에 의한 감염의 예방 또는 억제를 달성함으로써, 본원에 개시된 바와 같은 OAS1R 대립 유전자를 보유하는 인간에서 발견되는 자연 저항성을 복제하기에 적당한 것이 바람직하다.

하기 상세히 논의하는 바와 같이, NT\_009775.13의 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595, 또는 2156638에서의 하나 이상의 돌연변이의 유리한 효과를 모방하는 억제제-기초의 약물 요법을 또한 계획한다. 상기 논의한 바와 같이, 이러한 억제제를 개발하기 위한 한 예시적 근거는 유리한 돌연변이가 OAS1의 하나 이상의 특정 이소형의 발현, 번역, 또는 기능을 감소시키거나 소멸시키는 경우이다. 본 발명은 유리한 돌연변이의 정확한 형태 또는 효과에 의해서도, 또한 그로 인해 영향을 받는 특정 이소형의 생물학적 활성에 의해서도 제한되지 않는다. 이러한 경우에, 당업자는 상기 OAS1의 특정 이소형(들)을 치료적으로 억제하는 것의 유용성을 인식할 것이다. 상기 억제제-기초의 요법은 화학적 실체, 펩티드 또는 단백질, 안티센스 올리고뉴클레오티드, 작은 간섭 RNA, 및 항체의 형태를 취할 수 있다.

단백질, 그의 단편 또는 다른 유도체, 또는 그의 동족체, 또는 그를 발현시키는 세포는 거기에서 항체를 생산하기 위한 면역원으로서 사용될 수 있다. 상기 항체는, 예를 들어 폴리클로날, 모노클로날, 키메라, 단일 쇠, Fab 단편, 또는 Fab 발현 라이브러리의 산물일 수 있다. 폴리클로날 항체의 생산을 위해서 당업계에서 공지된 다양한 절차를 이용할 수 있다.

본 발명의 OAS1R에 의해 코딩되는 폴리펩티드에 대항하여 발생하는 항체는 폴리펩티드를 동물에 직접 주사함으로써 또는 폴리펩티드를 동물, 바람직하게는 비인간에 투여함으로써 획득할 수 있다. 그 후, 이렇게 획득된 항체는 폴리펩티드 자체에 결합할 것이다. 이러한 방식으로, 폴리펩티드의 오직 한 단편만을 코딩하는 서열조차 전체 천연 폴리펩티드를 결합시키는 항체를 발생시키기 위해 사용될 수 있다. 더욱이, 다수의 폴리펩티드에 특이적인 이러한 항체의 패널은 이러한 조직을 확인하고 차별하기 위해 사용될 수 있다. 한 예로서, 도 9는 특정한 예시적인 본 발명의 폴리펩티드에 특이적인 항체의 발달을 증명한다.

모노클로날 항체의 제조를 위해, 연속 세포주 배양에 의해 생산되는 항체를 제공하는 어떠한 기술도 사용할 수 있다. 예로는 하이브리도마 기술 (Kohler and Milstein, 1975, *Nature*, 256:495-597), 트리오마 기술, 인간 B-세포 하이브리도마 기술 (Kozbor, et al., 1983, *Immunology Today* 4:72), 및 인간 모노클로날 항체를 생산하기 위한 EBV-하이브리도마 기술 (Coe, et al., 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96)이 포함된다.

단일 쇠 항체의 생산을 위해 기재된 기술 (미국 특허 제4,946,778호)을 개조하여 본 발명의 면역원성 폴리펩티드 산물에 대한 단일 쇠 항체를 생산할 수 있다.

상기 항체는, 예를 들어 상기 단백질을 형상화하고, 적절한 생리학적 샘플에서 그의 수준을 측정하는 등을 위한 본 발명의 단백질 서열의 국부화 및 활성에 관한 방법에 사용할 수 있다.

본 발명은 세포 내의 OAS1 폴리뉴클레오티드를 형상화하기 위한 검출가능하게 표지된 올리고뉴클레오티드를 제공한다. 이러한 올리고뉴클레오티드는 유전자 증폭이 일어났는지의 여부를 결정하기 위해, 그리고, 예를 들어 당업계에 공지된 바와 같은 동일계내 혼성화를 이용하여 세포 또는 조직 내 발현 수준을 검정하기 위해 유용하다.

### OAS1 기능의 억제를 위한 치료제

또한, 본 발명은 OAS1 폴리뉴클레오티드의 정상 기능을 방해하도록 고안된 안티센스 올리고뉴클레오티드에 관한 것이다. 안티센스 기술에 광범위하게 적용가능하다고 당업계에 공지된 안티센스 분자의 임의의 변형 또는 변이는 본 발명의 범위 내에 포함된다. 이러한 변형은 미국 특허 제5,536,821호; 동 제5,541,306호; 동 제5,550,111호; 동 제5,563,253호; 동 제5,571,799호; 동 제5,587,361호, 동 제5,625,050호 및 동 제5,958,773호에 개시된 바와 같이 인-함유 결합의 제조를 포함한다.

본 발명의 안티센스 화합물은 미국 특허 제5,958,773호 및 거기에 개시된 특허에 개시된 바와 같은 변형된 염기를 포함할 수 있다. 본 발명의 안티센스 올리고뉴클레오티드는 또한 올리고뉴클레오티드를 하나 이상의 부위 또는 접합체와 화학적으로 결합시킴으로써 변형되어, 안티센스 올리고뉴클레오티드의 활성, 세포 분포, 또는 세포 업테이크를 향상시킬 수 있다. 이러한 부위 또는 접합체에는 지질, 예를 들어 콜레스테롤, 콜산, 티오에테르, 지방족 쇠, 인지질 및 그의 유도체, 폴리아민, 폴리에틸렌 글리콜 (PEG), 팔미틸 부위 및, 예를 들어 미국 특허 제5,514,758호, 동 제5,565,552호, 동 제5,567,810호, 동 제5,574,142호, 동 제5,585,481호, 동 제5,587,371호, 동 제5,597,696호 및 동 제5,958,773호에 개시된 바와 같은 다른 것이 포함된다.

키메라 안티센스 올리고뉴클레오티드도 또한 본 발명의 범위 내에 포함되며, 예를 들어, 미국 특허 제5,013,830호, 동 제 5,149,797호, 동 제5,403,711호, 동 제5,491,133호, 동 제5,565,350호, 동 제5,652,355호, 동 제5,700,922호 및 동 제 5,958,773호에 기재된 방법을 이용하여 본 발명의 올리고뉴클레오티드로부터 제조할 수 있다.

바람직한 안티센스 올리고뉴클레오티드는, 예를 들어 본 실시예에 기재된 검정법을 이용하여 일상적인 실험으로 선택될 수 있다. 본 출원인은 특정한 작용 메커니즘에 구속되지 않지만, 안티센스 올리고뉴클레오티드는 왓슨-크릭 염기쌍을 이용하여 세포 내의 표적 폴리뉴클레오티드의 상보적 영역에 결합함으로써 억제 효과를 달성한다고 여겨진다. 표적 폴리뉴클레오티드가 RNA인 경우, 실험적 증거에 의하면 하이브리드의 RNA 성분은 RNase H에 의해 절단된다고 한다 (문헌 [Giles et al., *Nuc. Acids Res.* 23:954-61, 1995]; 미국 특허 제6,001,653호). 일반적으로, 10개의 염기쌍을 함유하는 하이브리드는 RNase H에 대한 기질로서 작용하기에 충분한 길이를 가진다. 그러나, 상기 길이의 서열은 인간 유전자들 가운데서 독특한 경향이 있기 때문에, 결합의 특이성을 달성하기 위하여, 17개 이상의 뉴클레오티드의 안티센스 분자를 사용하는 것이 바람직하다.

본원에 참고로 포함되는 미국 특허 제5,998,383호에 개시된 바와 같이, 올리고뉴클레오티드는 그의 상보적 주형과의 올리고뉴클레오티드 이중나선 형성을 위해 중요한 적합한 에너지 관련 특성을 나타내고, 자가-이량체화 또는 자가-상보성을 위한 낮은 잠재성을 나타내도록 선택된다 (Anazodo et al., *Biochem. Biophys. Res. Commun.* 229:305-09, 1996). 컴퓨터 프로그램 올리고(OLIGO) (프라이머 분석 소프트웨어, 버전 3.4)는 안티센스 서열 용융 온도, 유리 에너지 성질을 결정하고 잠재적 자가-이량체 형성 및 자가-상보적 성질을 평가하기 위해 사용된다. 상기 프로그램은 상기 두 파라미터 (잠재적 자가-이량체 형성 및 자가-상보성)의 정성적 평가를 결정하게 하며, "잠재성 없음" 또는 "일부의 잠재성" 또는 "본질적으로 완전한 잠재성"의 지표를 제공한다. 상기 파라미터에서 잠재성이 없다고 평가된 OAS1 폴리뉴클레오티드의 절편이 일반적으로 선택된다. 그러나, 상기 카테고리 중 하나에서 "일부의 잠재성"을 가진 절편을 사용할 수 있다. 상기 파라미터의 조화를 선택에 사용할 수 있다.

안티센스 분야에서, 특정 표적에 대한 최적의 안티센스 분자를 선택하기 위해 어느 정도의 일상적인 실험이 요구된다. 효과적이게는, 안티센스 분자가 표적 RNA 분자의 접근가능한 또는 노출된 부분에 표적화되는 것이 바람직하다. 몇몇 경우에 표적 mRNA 분자의 구조에 대한 정보가 이용가능하지만, 안티센스를 이용하는 억제에 대한 최근의 접근법은 실험을 통한 것이다. 본 발명에 따르면, 상기 실험은 본 실시예에 기재된 방법을 이용하여 세포를 안티센스 올리고뉴클레오티드로 형질감염시킴으로서 일상적으로 수행할 수 있다. 세포에서의 mRNA 수준은 mRNA의 역전사 및 cDNA 수준의 검정에 의해 처리군 및 대조군 세포에서 일상적으로 측정할 수 있다. 생물학적 효과는 당업계에 공지된 바와 같이 세포 성장 또는 생존율을 측정함으로써 일상적으로 결정할 수 있다.

cDNA 수준을 검정하고 분석함으로써 안티센스 활성의 특이성을 측정하는 것은 안티센스 결과 확인에 대한 당업계에 인식되어 있는 방법이다. 처리군 및 대조군 세포로부터의 RNA를 역전사시켜야 하고, 생성된 cDNA 집단을 분석해야 한다고 제안되어 있다 (Branch, A. D., *T.I.B.S.* 23:45-50, 1998.). 본 발명에 따라, 세포의 배양물을 표적 OAS1에 대해 고안된 2개의 상이한 안티센스 올리고뉴클레오티드로 형질감염시킨다. OAS1에 상응하는 mRNA의 수준을 처리군 및 대조군 세포에서 측정한다.

추가적 억제제는 리보자임, 단백질 또는 폴리펩티드, 그의 항체 또는 단편 뿐만 아니라 소분자를 포함한다. 상기 OAS1 억제제 각각은 그들이 OAS1의 발현 및(또는) 생물학적 활성을 감소시킨다는 점에서 공통적인 특징을 공유한다. 본원에 개시된 예시적인 OAS1 억제제 이외에, 본원에 구체적으로 개시된 방법론 또는 당업자의 전문지식 내에서 쉽게 이용가능한 다른 것을 이용하여 일상적인 실험을 통하여 다른 억제제를 수득할 수 있다.

### 리보자임

OAS1 억제제는 리보자임일 수 있다. 리보자임은 RNA 기질, 예컨대 mRNA를 특이적으로 절단하여 특정 세포 유전자 발현을 억제하거나 방해하는 RNA 분자이다. 본원에서 사용된 용어 리보자임은 특정 인식 부위에 대한 안티센스 서열을 함유하며 RNA-절단 효소 활성을 함유하는 RNA 분자를 포함한다. 촉매 가닥은 화학량론적 농도보다 큰 농도에서 표적 RNA의 특정 부위를 절단한다.

예를 들어, 헤머헤드 리보자임 (예를 들어, 문헌 [Forster and Symons, *Cell* 48:211-20, 1987]; [Haseloff and Gerlach, *Nature* 328:596-600, 1988]; [Walbot and Bruening, *Nature* 334:196, 1988]; [Haseloff and Gerlach, *Nature* 334:585, 1988]에 기재됨); 헤어핀 리보자임 (예를 들어, 1993년 10월 19일자로 허여된 하셀로프 (Haseloff) 등의 미국 특허 제5,254,678호, 및 1990년 3월 26일자로 공개된 험펠 (Hempel) 등의 유럽 특허공개 제0 360 257호); 및 테트라히

메나 (Tetrahymena) 리보솜 RNA-기재 리보자임 (세치 (Cech) 등의 미국 특허 제4,987,071호 참조)을 비롯한 광범위한 리보자임이 본 발명에서 사용될 수 있다. 본 발명의 리보자임은 전형적으로 RNA로 구성되어 있으나, 또한 DNA, 핵산 유사체 (예를 들어, 포스포로티오에이트 (phosphorothioate)) 또는 그의 키메라 (chimeric) (예를 들어, DNA/RNA/RNA)로 구성될 수도 있다.

리보자임은 임의의 RNA 전사체를 표적으로 할 수 있고, 그러한 전사체를 촉매적으로 절단할 수 있다 (예를 들어, 미국 특허 제5,272,262호; 동 제5,144,019호; 및 세치 등의 미국 특허 제5,168,053호, 동 제5,180,818호, 동 제5,116,742호, 동 제5,093,246호 참조). 본 발명의 특정 실시양태에 따라, 임의의 그러한 OAS1 mRNA-특이적 리보자임 또는 그러한 리보자임을 코딩하는 핵산은 숙주 세포로 전달되어 OAS1 유전자의 발현을 억제할 수 있다. 따라서, 리보자임 등은 진핵세포 프로모터, 예컨대, 진핵 세포 바이러스 프로모터에 연결된 리보자임 코딩 DNA에 의해 숙주 세포로 전달될 수 있고, 핵으로의 도입시에 리보자임은 직접 전사될 것이다.

## RNAi

본 발명은 또한 부분적 또는 전체적으로 이중 가닥의 특성을 갖는 RNA의 세포내 또는 세포외 환경으로의 도입을 제공한다. 표적 OAS1 유전자의 일부로부터의 뉴클레오티드 서열이 억제성 RNA를 생성하도록 선택된다는 점에서 억제는 OAS1의 발현에 대해 특이적이다. 이 과정은 (1) 유전자 발현을 억제시키기에 효과적이고, (2) 표적 OAS1 유전자에 특이적이다. 그 과정은 표적 OAS1 유전자에 대해 부분적 또는 완전한 기능 상실을 제공할 수 있다. 표적 세포의 99% 이상에서 유전자 발현의 감소 또는 상실이 다른 표적 유전자와 비교가능한 기술을 이용하여 밝혀졌다. 주입된 물질의 보다 적은 투여량 및 dsRNA 투여 후의 보다 긴 시간은 세포의 보다 작은 분획에서의 억제를 야기할 수 있다. 세포에서 유전자 발현량의 정량화는 표적 mRNA의 축적 또는 표적 단백질의 번역 수준에서 유사한 억제량을 나타낼 수 있다. RNAi의 제조 및 사용방법은 이 거명에 의해 본원에 포함되는 미국 특허 제6,506,559호에 일반적으로 개시되어 있다.

RNA는 중합된 리보뉴클레오티드의 하나 이상의 가닥을 포함할 수 있으며, 포스페이트-당 골격 또는 뉴클레오시드에 대한 변형을 포함할 수 있다. 이중-가닥 구조는 단일 자기-상보적 RNA 가닥 또는 2개의 상보적 RNA 가닥에 의해 형성될 수 있다. RNA 이중나선의 형성은 세포 내부 또는 세포 외부에서 시작될 수 있다. RNA는 세포 당 1 카피 (copy) 이상을 전달할 수 있는 양으로 도입될 수 있다. 이중-가닥 물질의 보다 높은 투여량은 보다 효과적인 억제를 산출할 수 있다. RNA의 이중나선 영역에 상응하는 뉴클레오티드 서열이 유전자 억제에 대한 표적이라는 점에서 억제는 서열-특이적이다. OAS1 표적 유전자의 일부와 동일한 뉴클레오티드 서열을 함유하는 RNA가 억제에 바람직하다. 표적 서열에 대해 삽입, 결실 및 단일 점 변이가 있는 RNA 서열이 또한 억제에 효과적임이 밝혀졌다. 따라서, 서열 동일성은 당업계에 공지된 정렬 알고리즘, 및 뉴클레오티드 서열 간의 차이 백분을 계산에 의해 최적화될 수 있다. 별법으로, RNA의 이중나선 영역은 기능상, 표적 유전자 전사체의 일부와 혼성화될 수 있는 뉴클레오티드 서열로 정의될 수 있다.

RNA는 생체 내 또는 시험관 내에서 합성될 수 있다. 세포의 내생 RNA 폴리머라제가 생체 내에서 전사를 매개하거나, 클로닝된 RNA 폴리머라제가 생체 내 또는 시험관 내에서 전사에 사용될 수 있다. 생체 내 트랜스진 (transgene) 또는 발현 구조물로부터의 전사를 위해, 조절 영역이 RNA 가닥(들)의 전사에 사용될 수 있다.

RNAi에 대해, RNA는 직접 세포 (즉, 세포내)로 도입될 수 있거나, 캐비티, 세포간 공간으로 세포외로 도입되거나, 또는 유기체의 순환으로, 경구적으로 도입되거나, RNA를 함유하는 용액에서 유기체를 배양 (bathing)함으로써 도입될 수 있다. 경구 도입의 방법으로는, 음식으로 사용되는 종 (species)이 RNA를 발현하도록 설계하고, 이어서 유기체에게 섭취시키는 공학적 접근 뿐만 아니라, RNA를 유기체의 음식과 직접 혼합하는 것이 있다. 핵산을 도입하는 물리적 방법으로는, 세포로 직접 주입하는 것, 또는 유기체에게 RNA 용액을 세포의 주입하는 것이 있다.

이들 방법의 장점으로는 이중-가닥 RNA를 세포로 용이하게 주입할 수 있고, RNA를 낮은 농도로 사용할 수 있으며, 이중-가닥 RNA가 안정하고, 억제가 효과적이라는 것이다.

유전자 발현의 억제는 OAS1 표적 유전자로부터의 단백질 및(또는) mRNA 생성물 수준의 부재 (또는 식별가능한 감소)를 의미한다. 특이성은 세포의 다른 유전자 상에 뚜렷한 효과 없이 표적 유전자를 억제할 수 있는 능력을 의미한다. 억제의 결과는 세포 또는 유기체의 외부적 특성을 조사하거나, 예컨대, RNA 용액 혼성화, 뉴클레아제 보호, 노던 (Northern) 혼성화, 역 전사, 마이크로어레이 (microarray)로의 유전자 발현 모니터링, 항체 결합, 효소 연결 면역흡착 분석 (ELISA), 웨스턴 블랏팅, 방사면역검정법 (RIA), 다른 면역검정법, 및 형광 활성화 세포분석 (FACS) 등의 생화학적 기술에 의해 확인될 수 있다. 세포주 또는 전체 유기체에서 RNA-매개 억제에 대해, 유전자 발현은 단백질의 생성물이 용이하게 검출되는 리포터 또는 약물 저항성 유전자를 사용하여 편리하게 분석된다. 그러한 리포터 유전자로는 아세트히드록시산 신타제 (AHAS), 알칼리성 포스포타제 (AP), 베타 갈락토시다제 (LacZ), 베타 글루코로니다제 (GUS), 클로람페니콜 아세틸트랜스퍼라제

(CAT), 녹색 형광 단백질 (GFP), 호스래디쉬 퍼록시다제 (HRP), 루시페라제 (Luc), 노팔린 신타제 (NOS), 옥토판 신타제 (OCS) 및 이들의 유도체가 있다. 앰피실린, 블레오마이신, 클로람페니콜, 겐타마이신, 히그로마이신, 카나마이신, 린코마이신, 메토티렉세이트, 포스포노트리신, 푸로마이신 및 테트라사이클린에 대한 저항성을 부여하는 다수의 선별 마커를 사용할 수 있다.

분석에 따라, 유전자 발현의 정량화는 본 발명에 따라 처리하지 않은 세포와 비교하여 10%, 33%, 50%, 90%, 95% 또는 99% 보다 큰 억제 정도를 결정하도록 하였다. 주입된 물질의 보다 적은 투여량 및 dsRNA 투여 후의 보다 긴 시간은 세포의 보다 작은 분획에서의 억제를 야기할 수 있다 (예를 들어, 표적 세포의 10%, 20%, 50%, 75%, 90% 또는 95% 이상). 세포에서 OAS1 유전자 발현의 정량화는 OAS1 표적 mRNA의 축적 또는 OAS1 표적 단백질의 번역 수준에서 유사한 억제량을 나타낼 수 있다. 예를 들어, 억제의 효율성은 세포에서 유전자 생성물의 양을 측정함으로써 결정될 수 있다: mRNA는 억제성 이중-가닥 RNA에 사용되는 영역 외부의 뉴클레오티드 서열을 갖는 혼성화 프로브로 검출될 수 있거나, 번역된 폴리펩티드는 그 영역의 폴리펩티드 서열에 대한 항체로 검출될 수 있다.

RNA는 중합된 리보뉴클레오티드의 하나 이상의 가닥을 포함할 수 있다. 이는 포스페이트-당 골격 또는 뉴클레오시드에 대한 변형을 포함할 수 있다. 예를 들어, 천연 RNA의 포스포디에스테르 연결은 하나 이상의 질소 또는 황 헤테로원자를 포함하는 것으로 변형될 수 있다. RNA 구조에서의 변형은, 몇몇 유기체에서 dsRNA에 의해 발생하는 일반적 패닉 반응을 피하면서, 특이적 유전자 억제를 가능하게 하도록 맞추어 질 수 있다. 또한, 염기가 아데노신 데아미나아제의 활성을 차단하도록 변형될 수 있다. RNA는 효소에 의해 또는 부분/전체 유기 합성에 의해 생성될 수 있고, 임의의 변형된 리보뉴클레오티드는 시험관 내 효소적 합성 또는 유기 합성에 의해 도입될 수 있다.

이중-가닥 구조는 단일 자기-상보적 RNA 가닥 또는 2개의 상보적 RNA 가닥에 의해 형성될 수 있다. RNA 이중나선 형성은 세포의 내부 또는 외부에서 시작될 수 있다. RNA는 세포 당 1 카피 이상을 전달할 수 있는 양으로 도입될 수 있다. 이중-가닥 물질의 보다 많은 투여량 (예를 들어, 세포 당 5, 10, 100, 500 또는 1000 카피 이상)은 보다 효과적인 억제를 산출할 수 있으며, 보다 적은 투여량이 특정 분야에서 유용할 수도 있다. RNA의 이중나선 영역에 따른 뉴클레오티드 서열이 유전자 억제 표적이라는 점에서 억제 서열-특이적이다.

OAS1 표적 유전자의 일부와 동일한 뉴클레오티드 서열을 함유하는 RNA가 억제에 바람직하다. 표적 서열에 대해 삽입, 결실 및 단일점 변이가 있는 RNA 서열이 억제에 효과적일 수 있다. 따라서, 서열 동일성은 서열 비교 및 당업계에 공지된 정렬 알고리즘 (문헌 [Gribskov and Devereux, Sequence Analysis Primer, Stockton Press, 1991] 및 그에 언급된 문헌 참조) 및 뉴클레오티드 서열 사이의 차이 백분율 계산, 예를 들어 디폴트 파라미터 (예를 들어, 위스콘신 대학 유전자 컴퓨터 집단 (Genetic Computing Group))를 이용한 BESTFIT 소프트웨어 프로그램에서 수행된 스미스-워터먼 알고리즘에 의해 최적화될 수 있다. 억제성 RNA와 OAS1 표적 유전자의 부분 사이에서 90%보다 큰 서열 동일성 또는 100% 서열 동일성이 바람직하다. 별법으로, RNA 이중나선 영역은 기능상 OAS1 표적 유전자 전사체의 일부와 혼성화 (예를 들어, 400 mM NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA, 12-16시간 동안 50°C 또는 70°C에서 혼성화; 이어서 세척)가 가능한 뉴클레오티드 서열로 정의될 수 있다. 동일한 뉴클레오티드 서열의 길이는 25, 50, 100, 200, 300 또는 400개 이상의 염기일 수 있다.

RNA와 OAS1 표적 유전자 사이의 100% 서열 동일성은 본 발명의 실시에서 요구되지 않는다. 이와 같이, 본 발명의 방법은 유전자 변이, 균주 다형성 또는 진화적 다양성으로 인해 기대되는 서열 변형을 용인할 수 있는 이점이 있다.

OAS1 RNA는 생체 내 또는 시험관 내에서 합성될 수 있다. 세포의 내생 RNA 폴리머라제가 생체 내 전사를 매개하거나, 클로닝된 RNA 폴리머라제가 생체 내 또는 시험관 내에서 전사에 사용될 수 있다. 생체 내 트랜스진 또는 발현 구조물로부터의 전사를 위해, 조절 영역 (예를 들어, 프로모터, 인핸서, 사일런서, 스플라이스 공여체 및 수용체, 폴리아데닐화)이 RNA 가닥(들)의 전사에 사용될 수 있다. 억제는 기관, 조직 또는 세포 유형에서의 특이적 전사; 환경 조건의 자극 (예를 들어, 감염, 스트레스, 온도, 화학 유도체); 및(또는) 발생 단계 또는 연령에서의 전사 조작에 의해 표적화될 수 있다. RNA 가닥은 폴리아데닐화될 수도 있고, 안될 수도 있는데, 즉 RNA 가닥은 세포의 번역 장치에 의해 폴리펩티드로 번역될 수도, 안될 수도 있다. RNA는 수동 또는 자동적 반응에 의해 화학적 또는 효소적으로 합성될 수 있다. RNA는 세포 RNA 폴리머라제 또는 박테리오파지 RNA 폴리머라제 (예를 들어, T3, T7, SP6)에 의해 합성될 수 있다. 발현 구조물의 용도 및 제법은 WO 97/32016; 미국 특허 제5,593,874호, 동 제5,698,425호, 동 제5,712,135호, 동 제5,789,214호 및 동 제5,804,693호 및 그에 언급된 문헌에 공지되어 있다. 화학적으로 또는 시험관 내에서 효소에 의해 합성된 경우, RNA는 세포로 도입되기 전에 정제될 수 있다. 예를 들어, RNA는 용매 또는 수지로의 추출, 침전, 전기영동, 크로마토그래피 또는 이들의 조합에 의해 혼합물로부터 정제될 수 있다. 별법으로, RNA는 샘플 프로세싱으로 인한 손실을 피하기 위해 정제를 하지 않거나 정제를 최소로 하여 사용될 수 있다. RNA는 저장을 위해 건조되거나, 수용액에 용해될 수 있다. 용액은 이중나선 가닥의 어닐링 (annealing) 및(또는) 안정화를 촉진하기 위해 완충액 또는 염을 함유할 수 있다.

RNA는 직접 세포 (즉, 세포내)로 도입되거나, 캐비티, 세포간 공간으로 세포외로 도입되거나, 또는 유기체의 순환으로, 경구적으로 도입되거나, RNA를 함유하는 용액에서 유기체를 배양함으로써 도입될 수 있다. 경구 도입의 방법으로는, 음식으로 사용되는 종이 RNA를 발현하도록 설계하고, 이어서 유기체에게 섭취시키는 공학적 접근뿐만 아니라, RNA를 유기체의 음식과 직접 혼합하는 것이 있다. 예를 들어, RNA는 식물에 분무되거나, 식물은 식물을 감염시키는 것으로 공지된 몇몇 또는 모든 병원체를 사멸하기에 충분한 양의 RNA를 발현하도록 유전자 조작될 수 있다. 예를 들어, 세포로 직접 주입하거나 유기체에 세포 외로 주입하여 핵산을 도입하는 물리적 방법이 또한 이용될 수 있다. 혈관내 또는 혈관의 순환, 혈액 또는 림프 시스템 및 뇌척수액은 RNA가 도입되는 부위이다. 재조합 구조물로부터 RNA를 발현하는 트랜스제닉 유기체는 접합체, 배아줄기세포 또는 적합한 유기체로부터 유래된 다른 다능성 세포로 구조물을 도입함으로써 생성될 수 있다.

핵산을 도입하는 물리적 방법으로는 RNA 함유 용액을 주입하거나, RNA로 덮인 입자에 의해 충격을 가하거나, RNA 용액에 세포 또는 유기체를 침지하거나 또는 RNA의 존재하에 세포막의 전기 천공법을 이용하는 것이 있다. 바이러스 입자로 패키징된 바이러스 구조물은 발현 구조물의 세포로의 효과적 도입 및 발현 구조물에 의해 코딩된 RNA의 전사 모두 달성하게 된다. 당업계에 공지된 세포로 핵산을 도입하는 다른 방법이 이용될 수 있는데, 예컨대 지질-매개 담체 수송, 화학물질 (예컨대, 인산칼슘)-매개 수송 등이 있다. 이와 같이, RNA는 하나 이상의 하기 활성을 수행하는 구성성분과 함께 도입될 수 있다: 세포에 의한 RNA의 흡수를 강화시키거나, 이중나선 가닥의 어닐링을 촉진하거나, 어닐링된 가닥을 안정화시키거나, 또는 표적 유전자의 억제를 다른 방식으로 증가시키는 활성.

본 발명은 단독으로 사용되거나, 시험 샘플 또는 대상체에 RNA의 도입을 시험관 내 또는 생체 내에서 수행하기 위해 필요한 하나 이상의 시약을 갖는 키트의 구성성분으로서 사용될 수 있다. 바람직한 구성성분은 dsRNA, 및 dsRNA의 도입을 촉진하는 비히클이다. 그러한 키트는 또한 키트의 사용자가 본 발명을 실행하기 위한 지침서를 포함한다.

적합한 주입 혼합물은 동물이 평균  $0.5 \times 10^6$  내지  $1.0 \times 10^6$ 개의 RNA 분자를 받아들이도록 구성되어 있다. 센스, 안티센스 및 dsRNA 활성의 비교를 위해, 주입액을 동등한 질량의 RNA (즉, 단일 가닥의 절반 몰농도의 dsRNA)와 비교한다. 성인 1인 당 주입되는 분자의 수는 주입되는 물질의 RNA 농도 (에티뉼 브로마이드 염색으로부터 추정) 및 주입액 부피 (주입한 부위에서의 가시적 배수량으로부터 추정)에 기초하여 대강의 근사치로 주어진다. 개별적 동물 사이에서 주입액의 부피는 여러-배수로 변할 수 있다.

#### 단백질 및 폴리펩티드

또한, 본 발명의 OAS1 억제제는 본원에 개시된 안티센스 분자 및 리보자임 이외에도, OAS1 유전자 발현을 감소시키거나, 효소적 활성; 단일 가닥 RNA와의 상호 작용, 배열; 및 헤파티티스 C 바이러스 NS5A 또는 그의 단편과 같은 다른 단백질로의 결합 (이에 제한되는 것은 아님)을 비롯한 OAS1의 생물학적 활성들 중 하나 이상을 감소시키는 데 효과적인 단백질 또는 폴리펩티드를 포함한다. 다양한 방법들이 당업계에서 용이하게 입수될 수 있으며, 당업자는 이러한 방법을 이용하여 일반적인 실험을 통해 상기 OAS1 억제제를 신속하게 식별할 수 있다. 하기의 예시적인 방법들이 본 발명을 제한하는 것은 아니다.

당업자는 단백질-단백질 상호 작용의 검출 및 분석 방법이 기재되어 있는 문헌을 이용할 수 있다. 이 거명을 통해 본원에 포함되는 문헌 [Phizicky et al., *Microbiological Reviews* 59:94-123, 1995]을 검토한다. 이러한 방법으로는 물리적 방법, 예를 들어 단백질 친화 크로마토그래피, 친화 블롯팅, 면역침전 및 가교 뿐만 아니라 라이브러리-기재 방법, 예를 들어 단백질 프로빙, 파지 디스플레이 및 이중-하이브리드 스크리닝이 포함되나, 이에 제한되는 것은 아니다. 단백질-단백질 상호 작용을 식별하는 데 이용될 수 있는 다른 방법으로는 유전적 방법, 예를 들어 유전자외 억제자, 합성 치사 작용 및 비연쇄적 비상보작용 (unlinked noncomplementation)을 이용하는 것이 포함된다. 예시적인 방법들이 하기에 보다 상세하게 기재되어 있다.

본 발명의 OAS1 억제제는, 잠재적 억제제 단백질의 패널 또는 라이브러리와 OAS1 단백질 간의 직접적인 상호 작용에 의존하는 생물학적 스크리닝 분석을 통해 식별될 수 있다. 예를 들어, 문헌 [Vidal et al., *Nucl. Acids Res.* 27(4):919-29, 1999]; [Frederickson, R.M., *Curr. Opin. Biotechnol.* 9(1):90-96, 1998]; [Brachmann et al., *Curr. Opin. Biotechnol.* 8(5):561-68, 1997]; 및 [White, M.A., *Proc. Natl. Acad. Sci. U.S.A.* 93:10001-03, 1996]에는 다양한 "다중-하이브리드 (n-hybrid) 기술"을 비롯한 생물학적 스크리닝 방법들이 기재되어 있으며, 이들 문헌은 각각 이 거명을 통해 본원에 포함된다.

이중-하이브리드 스크리닝 방법은 억제 특성을 갖는 OAS1 결합 단백질에 대해 새롭거나 기존의 표적 cDNA 라이브러리를 찾는 데 이용될 수 있다. 이중-하이브리드 시스템은 보고 유전자 전사량의 증가에 따른 단백질-단백질 상호 작용을 검출하는 유전적 방법이다. 이 시스템은, 부위-특이적 전사 활성자가 DNA-결합 도메인 및 전사 활성화 도메인을 갖는다는 사실에 의존한다. DNA-결합 도메인은 활성화 도메인을 발현될 특이적 유전자로 표적화시킨다. 전사 활성자의 모듈 특성으로 인해, 어떠한 도메인도 활성을 상실하지 않으면서 전사 활성화 도메인으로부터 DNA-결합 도메인이 공유결합적으로 절단될 수 있다. 또한, 이들 두 도메인은 전사 기구와 무관한 두 단백질 간의 단백질-단백질 접촉에 의해 병렬로 배치(juxtaposition)될 수 있다. 따라서, 두 하이브리드가 구축되어 기능계를 생성한다. 제1 하이브리드, 즉 미끼는 관심 단백질에 융합된 전사 활성자 DNA-결합 도메인으로 구성된다. 제2 하이브리드인 표적은 전사 활성화 도메인과 단백질 또는 폴리펩티드의 라이브러리의 융합에 의해 생성된다. 미끼 단백질과 표적 라이브러리 구성원 간의 상호 작용은 DNA-결합 도메인과 전사 활성화 도메인을 병렬로 배치시키며, 결과적으로 보고 유전자 발현이 상향 조절된다.

당업자는, 효모 Gal4 또는 이. 콜라이 LexA DNA-결합 도메인 (BD) 및 효모 Gal4 또는 단순 포진 바이러스 VP16 전사 활성화 도메인이 가장 통상적으로 사용되는 다양한 이중-하이브리드계 시스템을 이용할 수 있다 [Chien et al., Proc. Natl. Acad. Sci. U.S.A. 88:9578-82, 1991]; [Dalton et al., Cell 68:597-612, 1992]; [Durfee et al., Genes Dev. 7:555-69, 1993]; [Vojtek et al., Cell 74:205-14, 1993]; 및 [Zervos et al., Cell 72:223-32, 1993]. 통상적으로 사용되는 보고 유전자로는 이. 콜라이 lacZ 유전자, 및 선별성 효모 유전자, 예를 들어 HIS3 및 LEU2가 포함된다 [Fields et al., Nature (London) 340:245-46, 1989]; [Durfee, T.K., supra]; 및 [Zervos, A.S., supra]. 단백질들의 상호 작용을 위한 스크리닝이 일반적인 실험을 통해 수행될 수 있는 다양한 활성화 도메인 라이브러리가 당업계에서 용이하게 입수가 가능하다.

OAS1 상호 작용 단백질의 식별에 적합한 미끼 단백질은 본원에 서열 19로 제공된 OAS1 DNA 서열을 토대로 설계될 수 있다. 이러한 미끼 단백질로는 전장 OAS1 단백질 또는 그의 단편이 포함된다.

OAS1 미끼 구조물 및 표적 라이브러리를 제조하기 위한 플라스미드 벡터, 예를 들어 pBTM116 및 pAS2-1은 당업자가 용이하게 입수할 수 있으며, 예를 들어 클론테크 (Clontech) (미국 캘리포니아주 팔로 알토 소재), 인비트로젠 (Invitrogen) (미국 캘리포니아주 칼스배드 소재) 및 스트라타젠 (Stratagene) (미국 캘리포니아주 라 졸라 소재)과 같은 판매원으로부터 입수할 수 있다. 이들 플라스미드 벡터에 의해 cDNA와, LexA 또는 Gal4BD로서의 DNA-결합 도메인의 프레임내 (in-frame) 융합이 각각 가능하다.

별법으로, 본 발명의 OAS1 억제제는 단백질-단백질 상호 작용의 검출을 위한, 당업계에서 이용될 수 있는 물리적 또는 생화학적 방법들 중 하나를 통해 식별될 수 있다.

리드 화합물은 고상 매트릭스, 예를 들어 세파로스 비드에 공유결합적으로 또는 비-공유결합적으로 연결시 OAS1에 대해 특이적으로 유지되기 때문에, 잠재적 OAS1 억제제로 시험될 이들 리드 화합물이 단백질 친화 크로마토그래피 방법을 통해 식별될 수 있다. 예를 들어, 문헌 [Beeckmans et al., Eur. J. Biochem. 117:527-35, 1981] 및 [Formosa et al., Methods Enzymol. 208:24-45, 1991]에는 단백질 친화 컬럼의 제법이 기재되어 있다. 세포 단백질을 전량으로 함유하는 세포 용해물은 OAS1 친화 컬럼을 통과할 수 있다. OAS1에 대한 친화도가 높은 단백질은 저염 조건하에 특이적으로 유지될 것인데 비해, 대다수의 세포 단백질은 컬럼을 통과할 것이다. 친화도가 높은 단백질은 고염 조건하에 고정화된 OAS1로부터 카오트로픽 (chaotropic) 용매 또는 나트륨 도데실 술페이트 (SDS)로 용출될 수 있다. 특정 실시양태에서, OAS1 특이적 결합 단백질을 식별하기 위한 보조제로서의 용해물을 제조하기 전에 세포에 방사선표지하는 것이 바람직할 수 있다. 포유동물 세포의 방사선표지 방법은 당업계에 잘 알려져 있으며, 예를 들어 문헌 [Sopta et al., J. Biol. Chem. 260:10353-60, 1985]에 제공된다.

친화 크로마토그래피에 적합한 OAS1 단백질이 단백질 또는 폴리펩티드에 융합됨으로써, 적절한 친화 수지 상에서 급속 정제가 수행될 수 있다. 예를 들어, OAS1 cDNA는 글루타티온 S-트랜스퍼라제 (GST)에 대한 코딩 부위에 융합될 수 있으며, 이는 융합 단백질이 글루타티온-아가로스 컬럼에 흡수되는 것을 촉진시킨다 [Smith et al., Gene 67:31-40, 1988]. 별법으로, 융합 단백질은 면역글로불린 G를 함유하는 컬럼에서 정제될 수 있는 단백질 A; Ni<sup>2+</sup>를 함유하는 컬럼에서 정제될 수 있는 올리고히스티딘-함유 펩티드; 아밀로스를 함유하는 수지에서 정제될 수 있는 말토스-결합 단백질; 및 메토트렉세이트 컬럼에서 정제될 수 있는 디히드로폴레이트 리덕타제를 포함할 수 있다. 본원에 제공되는 OAS1 융합 단백질의 제조에 적합한 한 전형적인 태그로는 인플루엔자 바이러스 헤마글루티닌 (HA)에 대한 에피토프가 있으며, 이에 대한 모노클로날 항체는 용이하게 입수할 수 있고, 그 항체로부터 친화 컬럼이 제조될 수 있다.

OAS1 친화 컬럼 상에 특이적으로 유지된 단백질들은 SDS 폴리아크릴아미드 겔 전기영동 (SDS-PAGE) 처리 이후에 식별될 수 있다. 따라서, 세포 용해물의 제조 및 OAS1 친화 컬럼 통과 이전에 세포가 방사선표지되는 경우, OAS1에 대한 친화도가 높은 단백질이 자기방사법에 의해 검출될 수 있다. OAS1 특이적 결합 단백질의 정체성은 당업자가 용이하게 입수할 수 있는 단백질 서열화 기술, 예를 들어 문헌 [Mathews, C.K. et al., Biochemistry, The Benjamin/Cummings Publishing Company, Inc., 1990, pp.166-70]에 기재된 기술에 의해 결정될 수 있다.

### 소분자

또한, 본 발명은 고속 처리 (high-throughput) 스크리닝 (HTS) 방법의 일반적인 적용을 통해 쉽게 식별될 수 있는 소분자 OAS1 억제제를 제공한다. 문헌 [Persidis, A., Nature Biotechnology 16:488-89, 1998]을 검토한다. HTS 방법은 일반적으로, 소분자와 같은 리드 화합물의 치료적 잠재력에 대한 고속 분석이 가능한 기술을 의미한다. HTS 방법에서는 시험 물질의 로봇식 (robotic) 취급, 양성 신호의 검출 및 데이터의 번역이 수행된다. 이러한 방법으로는, 예를 들어 가용성 분자를 이용한 로봇식 스크리닝 기술, 및 앞서 상세하게 기재된 이중-하이브리드 시스템과 같은 세포-기재 시스템이 포함된다.

용액을 이용하는 방법과는 달리, 세포적 맥락 내에서 발생하는 상호 작용들의 임상적 연관성 및 조작의 용이성으로부터 이점이 얻어지는 다양한 세포주-기재 HTS 방법이 이용될 수 있다. 리드 화합물은 방사능 도입을 통하거나, 판독 정보로서 흡광도, 형광성 또는 발광성에 의존하는 광학적 분석을 통해 식별될 수 있다. 예를 들어, 이 거명을 통해 본원에 포함되는 문헌 [Gonzalez et al., Curr. Opin. Biotechnol. 9(6):624-31, 1998]을 참조한다.

HTS 방법은, 예를 들어 OAS1의 생물학적 활성들 중 하나를 차단하는 리드 화합물에 대해 스크리닝하는 데 사용될 수 있다. 이 방법에 의해, OAS1 단백질은 단백질 발현 세포로부터 면역침전되어 로봇식 스크리닝에 적합한 분석 플레이트 상의 웰에 도포될 수 있다. 이후, 각 시험 화합물을 면역침전된 단백질과 접촉시켜, 각 시험 화합물이 OAS1에 미치는 영향을 조사할 수 있다.

### OAS1 억제제의 효능의 평가 방법

본원에 기재된 방법들 중 하나에 의해 또는 당업계에서 이용가능한 다른 기술을 통해 식별되는 리드 분자 또는 화합물 (안티센스 분자 또는 리보자임, 단백질 및(또는) 펩티드, 항체 및(또는) 항체 단편 또는 소분자)은 다양한 시험관내, 생체의 및 생체내 동물 모델 분석 시스템으로 OAS1 유전자 발현 또는 생물학적 활성에 대한 억제 능력에 대해 더 특징 분석될 수 있다. 하기 제공된 실시예에 보다 상세하게 논의된 바와 같이, 본 발명의 OAS1 억제제는 OAS1 발현 수준의 감소에 효과적이다. 따라서, 본 발명은 당업자가 후보 억제제의 효과를 평가할 수 있는 방법을 추가로 개시한다.

후보 OAS1 억제제는, 내인성 OAS1을 발현하는 세포, 또는 재조합 OAS1 플라스미드 구조물에 의한 포유동물 세포의 형질감염에 의해 OAS1를 발현하도록 제조된 세포로 상기 억제제를 투여함으로써 시험될 수 있다.

효과적인 OAS1 억제성 분자는 OAS1의 효소적 활성 또는 IFN 유도에 대한 OAS1의 반응능을 감소시키는 데 효과적일 것이다. OAS1의 효소적 활성 및 IFN 유도에 대한 반응능의 측정 방법은, 예를 들어 이 거명을 통해 본원에 포함되는 문헌 [Eskildsen et al., Nuc. Acids Res. 31:3166-3173, 2003]; 및 [Justesen et al., Nuc. Acids Res. 8:3073-3085, 1980]에 기재된 바와 같이 당업계에 공지되어 있다. 소정의 후보 안티센스 분자의 효능은, 포유동물 세포로 투여시 OAS1 발현에 실질적인 영향을 미치지 않는 것으로 알려진 대조 "안티센스" 분자와 비교함으로써 평가될 수 있다.

상기 논의된 방법들 중 하나 이상에 의한 OAS1 유전자 발현의 감소에 효과적인 OAS1 억제제는, 뎅기 (dengue) 바이러스와 같은 플라비바이러스 감염에 의해 공격된 배양 세포를 사용하는 것과 관련하여, 본원에 기재된 입수가 용이한 확립된 세포 배양 또는 주요 세포 배양 모델 시스템 중 하나로 시험관내에서 그 효능에 대해 더 특징 분석될 수 있다.

### 제약 조성물

본 발명의 안티센스 올리고뉴클레오티드 및 리보자임은 리보핵산 뉴클레오티드 또는 디옥시리보핵산 뉴클레오티드에 대해 당업계에 공지된 임의의 방법으로 합성될 수 있다. 예를 들어, 올리고뉴클레오티드는 어플라이드 바이오시스템즈 380B DNA 신세사이저 (Applied Biosystems 380B DNA synthesizer)에서와 같이 고상 합성에 의해 제조될 수 있다. 올리고뉴클레오티드의 최종 순도는 당업계에 공지된 바와 같이 측정된다.

본 발명의 방법에 의해 식별된 안티센스 올리고뉴클레오티드는 종양 세포 증식을 조절한다. 따라서, 본 발명의 방법에 의해 식별된 안티센스 올리고뉴클레오티드들 중 하나 이상과 조직 또는 세포를 접촉시키는 것을 포함하는, 바이러스 감염, 바람직하게는 플라비바이러스 감염, 가장 바람직하게는 HCV 감염을 방해하는 제약 조성물 및 방법이 제공된다.

본 발명은 치료 용도의 활성 성분으로서 OAS1 mRNA 유전자 서열에 상보적인 안티센스 올리고뉴클레오티드 및 리보자임의 제약 조성물을 제공한다. 상기 조성물은 본 발명의 방법에서도 사용될 수 있다. 필요한 경우, 상기 화합물은 뉴클레아제-저항성이다. 일반적으로, 포유동물의 바이러스 감염의 억제를 위한 제약 조성물은, 본 발명의 실시예에 필요하며 상기 기재된 바와 같은 유효량의 1종 이상의 안티센스 올리고뉴클레오티드, 또는 이와 동일한 효과를 나타내는 그의 단편, 및 제약학적으로 생리학상 허용되는 담체 또는 희석제를 포함한다.

상기 조성물은 경구, 피하, 또는 정맥내, 동맥내, 근육내, 복강내, 비내 및 경막내를 비롯한 비경구로 요구되는 주입 기술에 의해 투여될 수 있다. 일반적으로, 제약상 허용되는 담체, 희석제, 보조제 및 비히클, 및 이식용 담체는 본 발명의 활성 성분들과 반응하지 않는 불활성의 무독성 고상 또는 액상 충전제, 희석제 또는 캡슐화재를 의미한다. 양이온성 지질도 상기 조성물에 포함되어 올리고뉴클레오티드 흡수를 촉진할 수 있다. 화합물의 이식물도 유용하다. 일반적으로, 상기 제약 조성물은 멸균 상태이다.

생활성(발현가능한)이란, 올리고뉴클레오티드가 세포로 직접 전달되는 경우에 세포에서 생물학적으로 활성임을 의미하고(거나) 하기 기재된 바와 같이 벡터의 세포로 전달되는 경우에 적절한 프로모터에 의해 발현되어 활성임을 의미한다. 뉴클레아제-저항성은 본원에 기재된 바와 같이 생물학적 활성을 실질적으로 방해하지 않는, 당업계에서 공지된 임의의 방법에 의해 제공된다.

"세포 접촉"이란 세포에 안티센스 올리고뉴클레오티드를 직접적으로, 또는 바이러스 또는 비-바이러스 벡터에 의해 노출시키거나 전달하는 방법(안티센스 올리고뉴클레오티드는 전달시 생활성임)을 나타낸다.

본 발명의 뉴클레오티드 서열은 직접적으로, 또는 바이러스 또는 비-바이러스 벡터에 의해 전달될 수 있다. 직접 전달되는 경우, 서열은 일반적으로 뉴클레아제-저항성이 된다. 별법으로, 서열이 세포에서 발현되도록, 발현 카세트 또는 구조물로서 서열이 도입될 수 있다. 일반적으로, 상기 구조물은 표적화된 세포에서 서열이 발현될 수 있도록 하는 적당한 조절 서열 또는 프로모터를 함유한다.

올리고뉴클레오티드 서열이 전달될 준비가 되면, 이들은 당업계에 알려진 바와 같이 세포로 도입될 수 있다. 형질감염, 전기천공법, 융합, 리포솜, 콜로이드성 중합체성 입자 및 바이러스 벡터 뿐만 아니라 당업계에 공지된 다른 방법들도 올리고뉴클레오티드 서열을 세포로 전달하는 데 사용될 수 있다. 선택된 방법은 적어도 처리될 세포 및 그 세포의 위치에 따라 결정될 것이며, 당업자에게 알려져 있을 것이다. 국부화는 리포솜을 지향하기 위해 표면 상에 특이적 마커를 갖는 리포솜에 의해, 또는 표적 세포를 함유하는 조직으로의 직접 주입에 의해, 또는 표적 세포와 공간 근접성상 관련된 데포(depot)에 의해, 또는 특이적 수용체-매개 흡수, 바이러스 벡터 등에 의해 달성될 수 있다.

본 발명은 본 발명의 올리고뉴클레오티드 서열에 작용가능하게 연결된 발현 대조 서열을 포함하는 벡터를 제공한다. 또한, 본 발명은 필요에 따라 상기 벡터로 형질전환된 적합한 진핵세포 및 원핵세포로부터 선택되는 숙주 세포를 제공한다.

벡터는 공지되어 있거나, 또는 당업자에 의해 구축될 수 있는데, 벡터는 목적하는 서열 전사를 달성하는 데 필요한 모든 발현 요소를 함유해야 한다. 다양한 형태의 올리고뉴클레오티드의 회수를 위한 메카니즘과 같이 여타 유익한 특성도 벡터에 함유될 수 있다. 파지미드는 이러한 유익한 벡터의 구체적인 예이며, 그 이유는 파지미드가 플라스미드 또는 박테리오파지 벡터로서 사용될 수 있기 때문이다. 다른 벡터의 예로는 바이러스, 예를 들어 박테리오파지, 배칼로바이러스 및 레트로바이러스, DNA 바이러스, 리포솜 및 여타 재조합 벡터가 포함된다. 상기 벡터는 원핵세포성 또는 진핵세포성 숙주 시스템에서 사용하기 위한 요소들도 함유할 수 있다. 당업자는 숙주 시스템이 특정 벡터와 상용성임을 알 것이다.

상기 벡터는 당업계에 공지된 다양한 방법들 중 하나에 의해 세포 또는 조직으로 도입될 수 있다. 이러한 방법은 일반적으로 문헌 [Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor Laboratory, New York, 1989, 1992]; [Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD, 1989]; [Chang et al., *Somatic Gene Therapy*, CRC Press, Ann Arbor, Mich., 1995]; [Vega et al., *Gene Targeting*, CRC Press, Ann Arbor, Mich., 1995]; [Vectors: A Survey of Molecular Cloning Vectors and Their Uses,

Butterworths, Boston, Mass., 1988]; 및 [Gilboa et al., Biotechniques 4:504-12, 1986]에서 찾아볼 수 있으며, 그 예로는 안정적 (stable) 또는 일시적 (transient) 형질감염, 리포솜 형질감염 (lipofection), 전기천공법 및 재조합 바이러스 벡터에 의한 감염이 포함된다.

당업계에 알려져 있는 재조합 기술도 표적 핵산의 안티센스 억제를 달성하는 데 이용될 수 있다. 예를 들어, 안티센스 핵산을 함유하는 벡터가 안티센스 메시지를 발현시키는 데 사용되어 표적 핵산의 발현을 감소시키고, 결과적으로 그의 활성을 감소시킬 수 있다.

또한, 본 발명은 OAS1를 코딩하는 핵산 서열과, 핵산의 전사 또는 번역에 필요한 요소들을 포함하는 발현 벡터로 세포를 형질감염시키고; 시험 화합물을 투여하고; OAS1의 발현 수준을 시험 화합물이 없는 대조군에서 얻어진 수준과 비교하는 것을 포함하는, 소정의 화합물이 OAS1 유전자의 전사 또는 번역을 억제함에 따라 RNaseL에 대한 세포의 활성화능을 조절 (즉, 감소)되는지 여부를 평가하는 방법을 제공한다.

**바람직한 실시양태**

본 발명에서는, 앞서 기재된 방법 및 당업계에 공지된 다른 방법을 이용하여, 인간으로부터 단리된 게놈 DNA 샘플을 증폭 조건하에 PCR 프라이머 쌍으로 처리하여 올리고아데닐레이트 신테타제 (OAS1, 서열 19)의 뉴클레오티드 (nt) 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638 중 임의의 것을 함유하는 인간 게놈 DNA의 소정 부위를 증폭시키는 것을 포함하는 스크리닝 방법이 고려된다. 증폭 조건으로는 DNA 합성에 유효한 양의 PCR 완충액 및 소정의 열순환 온도가 포함된다. 생성된 PCR 생성물은 참조 뉴클레오티드 위치에서 점 돌연변이의 존재 여부에 대해 분석된다. 한 실시양태에서, PCR 생성물은 위치 2135728 (돌연변이 1), 2135749 (돌연변이 2), 2135978 (돌연변이 3), 2144072 (돌연변이 4), 2144088 (돌연변이 5), 2144116 (돌연변이 6) 또는 2144321 (돌연변이 7)을 비롯한, 5'에서 3' 방향으로 기록된 약 358 개의 염기 쌍 (bp), 및 각 측면에서 그 위치에 플랭킹된 약 175개의 염기를 포함하는 연속적인 뉴클레오티드 서열을 함유한다. 또 다른 실시양태에서, 상기 표 1 및 2에 기재된 바와 같은 앰플리콘 (amplicon)은 PCR 생성물 및 상응하는 프라이머를 예시한다. 한 바람직한 실시양태에서, PCR 생성물은 증폭 생성물을 혼성화 조건하에 상응하는 돌연변이에 대해 특이적인 올리고뉴클레오티드 프로브로 처리하고 임의의 혼성화 생성물의 형성을 검출함으로써, 상응하는 돌연변이에 대해 분석된다. 바람직한 올리고뉴클레오티드 프로브는 하기 표 3에 명시된 뉴클레오티드 서열을 포함한다. 표적 핵산으로의 올리고뉴클레오티드 혼성화는 미국 특허 제4,530,901호에 기재되어 있다.

**[표 3]**

돌연변이 서열번호	프로브 서열
서열번호 1	GTAGATTTTGCC <u>Y</u> GAACAGGTCAGT ( 서열번호 12)
서열번호 2	CAGTTGACTGGC <u>R</u> GCTATAAACCTA ( 서열번호 13)
서열번호 3	CAGAGGAGGGG <u>T</u> RGGGGGAGGAGA ( 서열번호 14)
서열번호 4	TCTCACCCTTTC <u>A</u> RGCTGAAAGCAAC ( 서열번호 15)
서열번호 5	GAAAGCAACAG <u>T</u> RCAGACGATGAGA ( 서열번호 16)
서열번호 6	ACGATCCCAGG <u>A</u> SGTATCAGAAATAT ( 서열번호 17)
서열번호 7	TTGATCCAGAG <u>A</u> RGACAAAGCTCCTC ( 서열번호 18)

여기서, R=A/G, S=C/G, 및 Y=C/T.

이어서, 형성된 PCR 혼합물을 다수의 PCR 열주기에 적용시켜 OAS1 및 OAS1R 유전자 증폭 생성물을 생성한다. 이어서, 증폭 생성물을 혼화 조건하에서 각 돌연변이에 특이적인 올리고뉴클레오티드 프로브로 처리한다. 이어서, 임의의 혼화 생성물을 검출한다.

하기 실시예는 예시하려는 것이지 임의의 방식으로 명세서와 청구의 범위를 제한하려는 것으로 이해해서는 안된다.

## 실시예

### 실시예 1

#### 게놈 DNA의 제조 및 예비 스크리닝

본 실시예에는 2개의 특정 집단으로부터의 DNA 스크리닝에 관한 것이나, HCV에 반복 노출이 기록된 다른 환자군에게도 동일하게 적용되며, 상기 노출은 감염을 초래하지 않았다. 실시예에는 또한 상기 논의된 바와 같은 다른 플라비바이러스에 노출된 환자를 스크리닝하는 것에 관한 것이며, 상기 노출은 감염을 초래하지 않았다.

본원에서, 2개의 집단이 연구되었다: (1) 혈우병환자 집단, 경미한 혈우병에서 중증 혈우병의 기준에 따라 선택됨, 1987년 1월 이전에 농축 혈액 응고 인자를 가짐; 및 (2) 정맥내 약물 사용 집단, 10년 이상의 주사 병력이 있음, 주사바늘 공유와 같은 다른 위험 행동의 증거가 있음. 연구는 HCV를 제외한 노출 음성 환자 및 HCV 포함 노출 양성 환자를 포함한다.

IV 약물 사용자, 혈우병 환자 및 C형 간염 또는 다른 플라비바이러스에 의한 감염의 위험이 있는 다른 집단으로부터 백혈구에서 고분자량 DNA를 추출하였다. 게놈 DNA의 초기 스크리닝의 경우, 상기 기재된 환자군으로부터의 혈액의 일치를 통보받은 후에, 혈액을 수집하여 0.14 M 시트르산, 0.2 M 시트르산삼나트륨 및 0.22 M 텍스트로스의 혼합물로 항응고시켰다. 항응고된 혈액을 실온에서 15분 동안 800 x g에서 원심분리하여 혈소판-풍부 혈장 상등액을 버렸다. 펠렛화된 적혈구, 단핵구 및 다핵 세포를 재현탁시키고, 개시 혈액 부피와 동일 부피의 저온 0.14 M 인산염 완충 염수 (PBS; pH 7.4)로 희석하였다. 낮은 내독소 피콜-히파퀴(Ficoll-Hypaque, 미국 미주리주 세인트 루이스 소재의 시그마 켐, 코포레이션) 상에서 400 x g로 18 °C에서 10분 동안 원심분리하여 맑은 세포 현탁액으로부터 말초 혈액 백혈구를 회수하였다 (18 °C). 이어서, 펠렛화된 백혈구를 재현탁시켜 고분자량의 DNA 공급원으로 사용하였다.

당업자에게 익히 공지된 방법 및 문헌 [Maniatis, et al., Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Sections 9.16-9.23, (1989)] 및 미국 특허 제4,683,195호에 기재된 방법을 이용하여 고분자량의 DNA를 백혈구로부터 단리하였다.

이어서, 올리고아데닐화 신테타제 1 유전자 (OAS1)에 상응하는 진뱅크 수탁번호 NT\_009775.13의 뉴클레오티드 위치를 참조하여 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638에서 임의의 하나의 뉴클레오티드의 점 돌연변이에 대해 DNA의 각 샘플을 조사하였다.

### 실시예 2

#### HCV 감염에 대한 저항성과 관련된 OAS1 유전자에서의 돌연변이

실시예 1에 기재된 방법을 이용하여, 비관련 혈우병환자 집단 및 정맥내 약물 사용 집단을 연구하여 서열 1 내지 서열 7 및 서열 57 내지 64에 개시된 바와 같은 OAS1에 돌연변이가 있는지 또는 없는지를 결정하였다.

백인종 집단에서 20가지의 증례 및 42가지의 대조군의 연구에서, 상기 돌연변이가 C형 간염 감염에 대한 저항성과 관련됨을 알아냈다. HCV 감염과 OAS1의 점 돌연변이의 존재 사이에 통계학상 유의적인 상관관계가 있다.

### 실시예 3

#### cDNA의 제조 및 서열분석

C형 간염 저항성 표현형을 갖는 환자로부터의 배양 림프모구 또는 섬유모세포로부터 전체 세포내 RNA를 정제하였다. 정제 절차는 문헌 [Chomczynski, et al., Anal. Biochem., 162:156-159 (1987)]에 기재된 바와 같이 수행하였다. 요컨대, 세포를 실시예 1에 기재된 바와 같이 제조하였다. 이어서, 4.0 M 구아니딘 티오시아네이트, 0.1 M 트리스-HCl (pH 7.5) 및 0.1 M 베타-메르캅토에탄올을 함유한 변성 용액 10 ml에 세포를 균질화시켜 세포 용해물을 형성하였다. 이어서, 나트륨 라우릴 사르코시네이트를 세포 용해물에 대해 최종 농도 0.5%로 혼합한 다음 혼합물을 5000 X g로 실온에서 10분 동안 원심분리하였다. 전체 RNA를 함유한 생성된 상등액을 5.7 M 염화세슘과 0.01 M EDTA (pH 7.5)의 완화액 상에서 층

지게 하고 원심분리하여 펠렛화하였다. 0.1% 나트륨 도데실 황산염 (SDS)을 함유한 10 mM 트리스-HCl (pH 7.6)과 1 mM EDTA (TE)의 용액 중에 생성된 RNA 펠렛을 용해시켰다. 페놀클로로포름 추출 및 에탄올 침전 후에, 260 nm에서 흡광도를 측정함으로써 정제된 전체 세포내 RNA 농도를 평가하였다.

상기에서 제조된 전체 RNA를, 제1 표준 합성용 역전사효소를 사용하고 cDNA를 5' 및 3' 절편으로 명명된 2개의 겹쳐진 절편으로 증폭시키도록 설계된 올리고뉴클레오타이드 프라이머로 PCR하는 cDNA 합성에 대한 주형으로 사용하였다. 본 발명을 실시하는데 사용되는 올리고뉴클레오타이드는 제작자의 지시서에 따라 어플라이드 바이오시스템즈 381A DNA 신시사이저(Applied Biosystems 381A DNA Synthesizer) 상에서 합성하였다. 당업계에 공지된 방법을 이용하여 PCR을 수행하였다. PCR 증폭 방법은 미국 특허 제4,683,192호, 동 제4,683,202호, 동 제4,800,159호 및 동 제4,965,188호, 및 텍스트 [PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, New York (1989)]; 및 [PCR Protocols: A Guide to Methods and Applications, Innis, et al., eds., Academic Press, San Diego, Calif. (1990)]를 포함하는 최소 수개의 텍스트에 상술되어 있으며, 프라이머는 본원의 표 1에 기재된 바와 같다.

HCV 감염된 환자 및 HCV 감염되지 않은 환자의 PCR-증폭 DNA로부터 직접 결정된 서열을 분석하였다. 바이러스에 반복 노출되었음에도 불구하고 HCV에 대해 혈청반응음성인 환자들에게서 OAS 유전자의 코딩 영역으로부터 상류에 돌연변이가 존재하는 것으로 검출할 수 있었다.

#### 실시예 4

점 돌연변이를 함유한 PCR 증폭 게놈 DNA의 제조 및 대립형질 특이적 올리고뉴클레오타이드 혼성화에 의한 검출

돌연변이를 함유한 PCR 증폭 게놈 DNA가 그 영역에 혼성화되는 올리고뉴클레오타이드 프로브와 혼성화되어 검출되는 점근법에 의해 뉴클레오타이드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523, 2156595 또는 2156638 중 하나에서 아데닐화 신테타제 (OAS1) 유전자의 점 돌연변이를 측정할 수 있었다. 올리고뉴클레오타이드 특이적 프로브와 혼성화되는 점 돌연변이를 갖는 영역을 증폭시키기 위해, 예를 들어 표 1에 나타난 각 프라이머 180 ng을 사용하여 실시예 3에 본질적으로 기재된 바와 같이 PCR 증폭을 수행하였다.

PCR 증폭 후에, 증폭된 올리고아데닐화 신테타제 DNA 생성물 2  $\mu$ l를 니트로셀룰로스 개별 용지 상에 스팟팅하였다. 스팟팅한 증폭 DNA를 건조시킨 후에, 니트로셀룰로스를 0.5 N NaOH로 2분 동안, 1 M 트리스-HCl (pH 7.5)로 2분 동안, 이어서 1.5 M NaCl을 함유한 0.5 M 트리스-HCl (pH 7.5)로 2분 동안 처리하여 DNA를 변성시키고 중화시켰다. 생성된 필터를 진공하에 80 °C에서 1시간 동안 베이킹(baking)하고, 6X SSC (1X = 0.15 M NaCl, 0.15 M 시트르산나트륨), 5X 덴하르트(Denhardt) 용액 (5X = 0.1% 폴리비닐피롤리돈, 0.1% 피콜, 및 0.1% 소 혈청 알부민), 5 mM 인산나트륨 완충액 (pH 7.0), 0.5 mg/ml 연어 정소 DNA 및 1% SDS으로 이루어진 예비혼성화 용액으로 42 °C에서 적어도 20분 동안 예비혼성화하였다.

예비혼성화 단계 후에, 니트로셀룰로스 필터를 개별적으로 예비혼성화 완충액으로 희석된 <sup>32</sup>P-표지된 올리고뉴클레오타이드 프로브에 노출시켰다. 프로브의 <sup>32</sup>P로의 표지는 키나제 완충액 (10X = 0.5 M 트리스[히드록시메틸] 아미노메탄 히드로클로라이드 (트리스-HCl) (pH 7.6), 0.1 M MgCl<sub>2</sub>, 50 mM 디티오프레이트 (DTT), 1 mM 스페르미딘-HCl 및 1 mM 에틸렌디아민테트라아세트산 (EDTA))의 10X 농축액 2.5  $\mu$ l, 60  $\mu$ g/ $\mu$ l의 선택적 올리고뉴클레오타이드 1.1  $\mu$ l, 물 18.4  $\mu$ l, 6000 Ci/mM의 감마 <sup>32</sup>P ATP (150 mCi/ $\mu$ l 농도) 2  $\mu$ l, 및 10 U/ $\mu$ l 폴리뉴클레오타이드 키나제 1  $\mu$ l를 혼합함으로써 수행하였다. 표지 혼합물을 37 °C에서 20분 동안에 이어 68 °C에서 2분 동안 유지하였다. 이어서, 유지 혼합물을 세파덱스 (Sephadex) G50 (미국 뉴저지주 피스캐터웨이 소재의 파마시아 인크.(Pharmacia, Inc.))에 적용하여 혼입되지 않은 <sup>32</sup>P-표지된 ATP를 제거하였다.

돌연변이를 함유한 영역과 혼성화시키는데 사용하는 올리고뉴클레오타이드 프로브를 상기 표 3에 나타냈다. 밀줄 친 뉴클레오타이드는 돌연변이 뉴클레오타이드에 해당된다. 야생형 (정상) 검출용 프로브에서, 밀줄 친 뉴클레오타이드는 야생형 뉴클레오타이드로 대체하였다.

정상 및 돌연변이 표지된 프로브  $10 \times 10^6$  cpm을 각 필터와 개별적으로 혼합하였다. 이어서, 니트로셀룰로스 필터를 42 °C에서 밤새 유지하여 혼성화 생성물을 형성하였다. 정상 프로브에 노출된 니트로셀룰로스 필터는 0.1% SDS를 함유한 6X SSC로 46 °C에서 세척한 반면, 돌연변이 프로브에 노출된 필터는 52 °C의 보다 엄격한 온도에서 동일한 용액으로 세척하였다. 이어서, 니트로셀룰로스 필터를 건조시키고, 방사선자가기폭법에 적용하였다.

점 돌연변이를 갖는 생성물만 돌연변이 프로브와 혼성화시켰다. PCR 증폭이 성공적인지 결정하기 위해 양성 및 음성 대조군을 각 분석에 포함시켰다. 이어서, 실시예 1에서 제조한 환자 게놈 DNA가 표시된 위치에서 야생형 뉴클레오티드를 비-야생형 뉴클레오티드로 대체한 유일한 점 돌연변이를 갖는지를 상기 접근법으로 결정하였다.

## 실시예 5

### 표적 RNA의 안티센스 억제

#### A. 형질감염을 위한 올리고뉴클레오티드의 제조

리피토이드 또는 콜레스테로이드를 포함하는 운반체 분자를 물로 0.5 mM가 되게 희석한 다음 초음파처리하여 균일한 용액을 생성시키고 0.45  $\mu$ m PVDF 막을 통해 여과함으로써 형질감염용으로 제조하였다. 이어서, 리피토이드 또는 콜레스테로이드를 적절한 부피의 OptiMEM(상표명) (기브코(Gibco)/BRL)로 희석하여 최종 농도가 올리고뉴클레오티드 1 mg 당 리피토이드 약 1.5 내지 2 nmol이 되게 하였다.

안티센스 및 대조군 올리고뉴클레오티드를 멸균 밀리포어(Millipore) 물로 100 mM의 작업 농도로 제1 희석한 다음 OptiMEM(상표명)으로 2 mM (약 20 mg/mL)로 희석하였다. 이어서, 희석된 올리고뉴클레오티드를 희석된 리피토이드에 즉시 첨가하여 아래 위로 피펫팅함으로써 혼합하였다.

#### B. 형질감염

HCV 감염에 민감하고 HCV 복제를 부양하는 인간 PH5CH8 간세포를 사용하였다 (문헌 [Dansako et al., Virus Res. 97:17-30, 2003]; [Ikeda et al., Virus Res. 56:157-167, 1998]; [Noguchi and Hirohashi, In Vitro Cell Dev. Biol. Anim. 32:135-137, 1996.]). 혼합 후에 즉시 올리고뉴클레오티드의 최종 농도가 300 nM이 되게 올리고뉴클레오티드/리피토이드 혼합물을 첨가하여 상기 세포를 형질감염시켰다. 이어서, 상기 세포를 형질감염 혼합물과 함께 37 °C, 5% CO<sub>2</sub>에서 밤새 인큐베이션하고, 형질감염 혼합물을 3 내지 4일 동안 세포 상에 두었다.

#### C. 전체 RNA 추출 및 역전사

제작자가 제공한 프로토콜에 따라 RNeasy(상표명) 키트 (미국 캘리포니아주 챗스워스 소재의 퀴아젠 코퍼레이션(Qiagen Corporation))를 사용하여 전체 RNA를 형질감염된 세포로부터 추출하였다. RNA를 추출한 다음, PCR 주형으로 사용하기 위해 역전사시켰다. 일반적으로, 전체 추출한 RNA 0.2 내지 1 mg을 멸균 미세원심분리 튜브에 위치시키고, 물을 첨가하여 전체 부피를 3 mL가 되게 하였다. 완충액/효소 혼합물 7 mL를 각 튜브에 첨가하였다. 완충액/효소 혼합물을 하기 열거된 순서로 혼합하여 제조하였다:

4 mL 25 mM MgCl<sub>2</sub>

2 mL 10X 반응 완충액

8 mL 2.5 mM dNTP

1 mL MuLV 역전사효소 (50 u) (어플라이드 바이오시스템즈)

1 mL RNase 억제제 (20 u)

1 mL 올리고 dT (50 pmol)

미세원심분리 튜브의 내용물을 아래 위로 피펫팅하여 혼합하고, 반응물을 42 °C에서 1시간 동안 인큐베이션하였다.

D. 표적 서열의 PCR 증폭 및 정량분석

역전사한 다음, 로체 라이트 시클러(상표명)(Roche Light Cycler™) 실시간 PCR 기계를 이용하여 표적 유전자를 증폭시켰다. 하기에 열거된 순서대로 성분을 혼합하여 PCR 증폭 혼합물 20 mL 분취액을 제조하였다: 10X PCR 완충액 II (10 mM 트리스 (pH 8.3) 및 50 mM KCl, 미국 코네티컷주 노워크 소재의 퍼킨-엘머(Perkin-Elmer) 함유) 2 mL, 3 mM MgCl<sub>2</sub>, 140 mM의 각 dNTP, 0.175 pmol의 각 OAS1 올리고, 1:50,000 희석률의 SYBR(등록상표)그린(Green), 0.25 mg/mL BSA, 1 유닛의 Taq 폴리머라제, 및 H<sub>2</sub>O로 20 mL가 되게 함. SYBR(등록상표)그린 (미국 오레곤주 유진 소재의 몰레큘라 프로브즈(Molecular Probes))은 이중-가닥 DNA에 결합하여 각 반응에서 생성된 PCR 생성물의 양을 직접적으로 측정할 수 있게 하는 형광 염료이다. 완전 역전사 반응물 2 mL를 PCR 증폭 혼합물의 각 20 mL 분취액에 첨가하고, 증폭을 표준 프로토콜에 따라 수행하였다.

실시에 6

OAS1 RNA<sub>i</sub>로 세포 처리

안티센스 처리를 위해 실시에 5의 방법을 이용하여, 세포를 OAS1 서열 (서열 19) 기체의 올리고뉴클레오티드로 처리하였다. 3' 말단에 테옥시-TT 연장을 갖는 2개의 상보적인 리보뉴클레오티드 단량체를 합성하여 어닐링시켰다. PH3CH8 간 세포 세포주의 세포를 1:3 L2 리피토타이드를 함유한 50 내지 200 nM RNAi로 처리하였다. 세포를 제1일, 제2일, 제3일 및 제4일째 수확하여 문헌 [Dansako et al., Virus Res. 97:17-30, 2003]에 기재된 바와 같이 웨스턴 분석으로 OAS1 단백질에 대해 분석하였다.

실시에 7

OAS1의 C형 간염 바이러스 NS5A 단백질과의 상호작용

문헌 [Taguchi, T. et al., J. Gen. Virol. 85:959-969, 2004]에 기재된 방법을 이용하여 본 발명의 OAS1 단백질 또는 폴리펩티드가 C형 간염 바이러스 NS5A 단백질과 상호작용 능력을 분석하였다. OAS1 단백질 및 폴리펩티드를 코팅하는 폴리뉴클레오티드를 상기 기재된 바와 같이 제조하고, 상기 문헌 [Taguchi, T. et al]에 기재된 바와 같은 통상의 방법을 이용하여 플라스미드를 제작하였다. 한 플라스미드는 OAS1 단백질 또는 폴리펩티드를 코팅하는 폴리뉴클레오티드를 함유하며 제2 플라스미드는 NS5A를 코팅하는 폴리뉴클레오티드를 함유한다. 또한, 상기 플라스미드는 FLAG-태그, HA 또는 GST와 같은 각각의 단백질을 위한 적절한 태그를 코딩하고 있다. 태그 및 NS5A 단백질을 코팅하는 플라스미드, 및 상이한 태그 및 OAS1 단백질 또는 폴리펩티드를 코딩하는 플라스미드로 헬라 세포와 같은 적합한 세포를 일시적으로 형질감염시켰다. 문헌 [Taguchi, T. et al.]에 기재된 바와 같이 인큐베이션하고 상등액을 제조한 후에, 각종 분석 기술을 이용하여 NS5A의 OAS1 단백질 또는 폴리펩티드와의 결합을 검출하고 정량분석하였다. 이러한 기술은 당업계에 공지되어 있으며 공동-침전 분석법, 면역형광 분석법 및 이뮤노블롯 분석법을 포함한다. NS5A와의 결합을 나타내지 않은 OAS1 단백질 및 폴리뉴클레오티드가 C형 간염 감염의 억제제로서 추가 분석에 적절하였다.

실시에 8

OAS1의 화학 및 입체구조적으로 보존된 영역

당업자가 인식하는 바와 같이, OAS1의 구조를 변형시켜 효소 활성 또는 치료 잠재능을 개선시키는 경우, 특정 잔기 또는 잔기의 특정 영역은 화학 및 구조적으로 보존되어야 한다. 실시에에 수개의 보존 도메인을 하기에 나타냈다. 당업자가 인식하는 바와 같이, 단백질의 구조 및 기능을 보존하는 특정 아미노산에 대해 화학적으로 보존된 변화는 관용될 수 있다. 예를 들어, Asp75 및 Asp77 둘다는 OAS1 기능에 본질적인 촉매 2가 금속 이온을 조정한다. 이들 염기에 대한 변형 (예를 들어, 아스파라긴 또는 글루탐산)이 관용될 수 있기 때문에, 이들 잔기의 본질적인 극성 및 산성 특성은 반드시 보존된다.

서열 26 내지 29, 서열 33, 서열 34 및 서열 50에 대한 예로서, 하기 폴리펩티드 절편은 보존된 도메인을 나타낸다:

아미노산 40-47: FLKERCFR (서열 75)

아미노산 55-82: VSKVVKGGSSGKGTTLRGRSDADLVVFL (서열 76)

아미노산 94-112: RRGEFIQEIRRQLEACQRE (서열 77)

아미노산 128-138: NPRALSFVLSS (서열 78)

아미노산 145-158: VEFDVLPAPFDALGQ (서열 79)

아미노산 182-198: KEGEFSTCFTELQRDFL (서열 80)

아미노산 201-217: RPTKCLKSLIRLVKHWYQ (서열 81)

아미노산 225-241: KLPPQYALELLTVYAW (서열 82)

아미노산 296-307: PVILDPADPTGN (서열 83)

아미노산 337-343: GSPVSSW (서열 84)

실시예 9

효소 활성 부위를 개선시키는 아미노산 변화

단백질의 효소 활성을 개선시키는 OAS 아미노산 서열의 변화를 생각해 볼 수 있다. 한 바람직한 실시양태에서, 효소의 활성 부위 내의 아미노산을 변형시켜 ATP 또는 금속 이온 결합, 효소 효능 및 효소 연작용을 개선시킬 수 있다. 이러한 변형의 예로서 서열 26, 서열 27, 서열 33, 서열 34, 서열 35 또는 서열 48의 아미노산 위치 61에서 글리신을 티로신으로 치환시킬 수 있다. 극성 티로신에 대해 화학적으로 독성이 없는 글리신의 치환은 ATP의 N3 원자와 상기 아미노산 위치 사이의 수소 결합을 촉진시키므로 해리 상수 및 이들 상호작용의 에너지학을 개선시킬 것이다. 티로신은, 예를 들어 보다 향상된 폴리-A 폴리머라제의 상기 위치에서 발견된다. 당업자가 인식하는 바와 같이, 다른 변형을 생각할 수 있다.

실시예 10

이중-가닥 RNA 결합을 개선시키는 아미노산 변화

효소 활성을 개선시키는 OAS에 대한 아미노산 변형의 제2의 예는 이 단백질과 이중-가닥 바이러스 RNA 사이의 상호작용을 안정화시키는 것일 수 있다. 하기 표에는 단백질의 RNA 결합 홈에 존재하는 아미노산, 및 염기성의 양전하 아미노산 측쇄와 음전하 리보핵산 사이의 상호작용을 안정화시키기 위해 고안된 수개의 제안되는 변화가 열거되어 있다. 단백질의 RNA 결합 홈의 양전하 밀도를 증가시키는 변화가 생각될 수 있다.

**[표 4]**  
RNA 결합 홈에 존재하는 아미노산의 제안되는 변화

아미노산	위치	제안 변형
글리신	39	아르기닌 또는 리신
리신	42	아르기닌
리신	60	아르기닌
아르기닌	195	리신
리신	199	아르기닌
리신	204	아르기닌

실시예 11

유전자 돌연변이의 분석

당업자는 다수의 다른 분석 방법이 유전자 분석에서 특정 돌연변이의 진화적 중요성을 평가하기 위해 존재한다는 것을 인식할 것이다. 한 예로는 통상적으로 D' (문헌 [Lewontin, Genetics 49:49-67, 1964])으로 언급된 연결 불균형 평가의 익히-공지된 계산법이다. 다른 특히 적절한 방법이 좌위에서 고빈도, 중간빈도 또는 저빈도 돌연변이의 상대적 다수를 평가함으로써 유전자 좌위 (예를 들어, 선택적 일소) 내에서 선택적 압력 및(또는) 최근 진화 사건을 평가하는 것을 시도하였다. 가장 익숙한 이들 시험은 타지마 D 통계 (Tajima, Genetics 123:585-595, 1989)이다. 문헌 [Fu and Li, Genetics 133:693-709 (1993)]에서는 또한 타지마에 대한 변법 및 각 돌연변이에 대한 원형 대립형질에 관한 지식을 사용하도록 하는 다른 통계학을 발전시켰다. 이들 및 다른 방법을 본 발명의 돌연변이에 적용시켜 관찰된 효과에 대한 상대적 기여를 평가하였다.

특정 실시양태 및 실시예를 포함하는 상기 명세서는 본 발명을 예시하려는 것이지 제한하려는 의도가 아니다. 다수의 다른 변법 및 변형이 본 발명의 진정한 취지 및 범위로부터 벗어나지 않는 한 행해질 수 있다. 인용된 모든 특허, 특허 공보 및 비-특허 간행물이 본원에 참고로 포함된다.

(57) 청구의 범위

청구항 1.

참조 서열 19의 뉴클레오티드 위치 2135728, 2135749, 2135978, 2144072, 2144088, 2144116, 2144321, 2131025, 2133961, 2139587, 2144294, 2144985, 2156523 및 2156638에서의 참조 뉴클레오티드의 비-참조 뉴클레오티드로의 치환; 및 참조 서열 19의 위치 2156595에서의 참조 뉴클레오티드의 결실로 이루어진 균으로부터 선택되는 올리고아데닐화 신테타제 유전자 (OAS1) 점 돌연변이의 존재를 핵산 샘플 중에 검출하는 단계;

이에 따라 상기 돌연변이를 확인하는 단계

를 포함하는, 상기 OAS1 돌연변이를 확인하기 위한 인간 유전자 스크리닝 방법.

청구항 2.

서열 20 내지 30, 서열 32 내지 35 및 서열 46 내지 52로 이루어진 균으로부터 선택되는 아미노산 서열로 이루어진 단리된 폴리펩티드.

청구항 3.

서열 75 내지 84로 이루어진 균으로부터 선택되는 1종 이상의 아미노산 서열로 이루어지고, 서열 20 내지 30, 서열 32 내지 35 및 서열 46 내지 52로 이루어진 균으로부터 선택되는 폴리펩티드와 80% 이상의 서열 유사성을 갖는 단리된 폴리펩티드.

청구항 4.

서열 75 내지 84로 이루어진 균으로부터 선택되는 1종 이상의 아미노산 서열로 이루어지고,

(a) 서열 22 및 서열 25 중 어느 하나의 아미노산 219부터 238;

(b) 서열 23의 아미노산 231부터 250;

(c) 서열 26 내지 29, 서열 33 내지 34 및 서열 50 중 어느 하나의 아미노산 347부터 366;

(d) 서열 32의 아미노산 295부터 314;

(e) 서열 16의 아미노산 189부터 208;

(f) 서열 47의 아미노산 61부터 80

으로 이루어진 군으로부터 선택되는 폴리펩티드와 80% 이상의 서열 유사성을 갖는 단리된 폴리펩티드.

#### 청구항 5.

제2항 내지 제4항 중 어느 한 항에 있어서, 단백질 형질도입 도메인을 포함하는 폴리펩티드에 공유결합적으로 부착된 폴리펩티드.

#### 청구항 6.

제5항에 있어서, 단백질 형질도입 도메인이 서열 85 내지 94로 이루어진 군으로부터 선택되는 폴리펩티드로 구성된 것인 폴리펩티드.

#### 청구항 7.

제5항에 있어서, 단백질 형질도입 도메인이 서열 85 내지 94로 이루어진 군으로부터 선택되는 폴리펩티드와 80% 이상의 서열 유사성을 갖는 폴리펩티드로 구성된 것인 폴리펩티드.

#### 청구항 8.

제5항에 있어서, 단백질 형질도입 도메인이 아르기닌, 리신 또는 히스티딘의 부가 또는 치환에 의해 서열 85 내지 94로 이루어진 군으로부터 선택되는 폴리펩티드와 상이한 것인 폴리펩티드.

#### 청구항 9.

제2항 내지 제8항 중 어느 한 항에 있어서, 폴리에틸렌 글리콜에 공유결합적으로 부착된 폴리펩티드.

#### 청구항 10.

제2항 내지 제8항 중 어느 한 항에 있어서, 리포솜 중에 캡슐화된 폴리펩티드.

#### 청구항 11.

제2항 내지 제8항 중 어느 한 항에 있어서, 엔도솜 파괴제에 공유결합적으로 부착된 폴리펩티드.

#### 청구항 12.

제2항 내지 제8항 중 어느 한 항에 있어서, 엔도솜 파괴제에 비공유결합적으로 부착된 폴리펩티드.

**청구항 13.**

제11항 또는 제12항에 있어서, 엔도솜 파괴제가 pH 감수성인 폴리펩티드.

**청구항 14.**

제2항 내지 제13항 중 어느 한 항에 있어서, 당 잔기에 공유결합적으로 접합된 폴리펩티드.

**청구항 15.**

제14항에 있어서, 당 잔기가 갈락토스 또는 만노스로 구성된 것인 폴리펩티드.

**청구항 16.**

(a) 세포에 의해 제2항 내지 제8항 중 어느 한 항의 폴리펩티드를 발현시키는 단계; 및

(b) 상기 폴리펩티드를 회수하는 단계

를 포함하는 방법에 의해 제조된 단리된 폴리펩티드.

**청구항 17.**

제2항 내지 제8항 중 어느 한 항에 있어서, 재조합 숙주 세포에 의해 제조되는 폴리펩티드.

**청구항 18.**

제2항 내지 제8항 중 어느 한 항에 있어서, 이중 폴리펩티드 서열을 포함하는 폴리펩티드.

**청구항 19.**

제2항 내지 제18항 중 어느 한 항의 폴리펩티드 및 제약상 허용가능한 담체를 포함하는 조성물.

**청구항 20.**

제2항 내지 제8항 중 어느 한 항의 폴리펩티드 서열을 코딩하는 뉴클레오티드 서열을 포함하는 단리된 폴리뉴클레오티드.

**청구항 21.**

서열 31, 서열 36 내지 45 및 서열 55 내지 56으로 이루어진 군으로부터 선택되는 뉴클레오티드 서열을 포함하는 단리된 폴리뉴클레오티드.

**청구항 22.**

제20항 또는 제21항의 단리된 폴리뉴클레오티드를 포함하는 재조합 벡터.

**청구항 23.**

발현 조절 서열과 작동적으로 연결된 제20항 또는 제21항의 단리된 폴리뉴클레오티드를 포함하는 발현 벡터.

**청구항 24.**

제23항에 따른 발현 벡터로 형질전환 또는 형질감염된 숙주 세포.

**청구항 25.**

제23항에 있어서, 바이러스 벡터인 발현 벡터.

**청구항 26.**

제23항에 있어서, 아데노바이러스 벡터, 아데노-관련 바이러스 벡터, 바쿨로바이러스, 쉘리키 삼립 바이러스 벡터, 신드비스 바이러스 벡터, 폭스 바이러스 벡터, 우두 바이러스 벡터, 조류 폭스바이러스 벡터, 아비폭스 바이러스 벡터, 계두 바이러스 벡터, 카나리폭스 바이러스 벡터, 알파바이러스 벡터, 계두 바이러스 벡터 또는 렌티바이러스 벡터로 이루어진 군으로부터 선택되는 바이러스 벡터인 발현 벡터.

**청구항 27.**

제약상 허용가능한 담체로 이루어진 군으로부터 선택되는 제1 성분 및 제23항, 제25항 및 제26항 중 어느 한 항에 따른 발현 벡터로 이루어진 군으로부터 선택되는 제2 성분을 포함하는 조성물.

**청구항 28.**

바이러스 감염 치료가 필요한 포유동물에게 올리고아데닐화 신테타제 1을 포함하는 조성물을 투여하는 것을 포함하는, 상기 포유동물에서의 바이러스 감염의 치료 방법.

**청구항 29.**

제28항에 있어서, 상기 올리고알데닐화 신테타제 1이 제2항 내지 제8항 중 어느 한 항의 폴리펩티드인 방법.

**청구항 30.**

제28항에 있어서, 상기 올리고알데닐화 신테타제 1이 제20항 또는 제21항의 폴리뉴클레오티드에 의해 발현되는 방법.

**청구항 31.**

제28항에 있어서, 제19항 또는 제27항에 따른 조성물을 투여하는 것을 포함하는 방법.

**청구항 32.**

제28항에 있어서, 상기 바이러스 감염이 이중-가닥 RNA 바이러스로의 감염인 방법.

**청구항 33.**

제28항에 있어서, 상기 바이러스 감염이 플라비바이러스로의 감염인 방법.

**청구항 34.**

제33항에 있어서, 상기 플라비바이러스가 C형 간염 바이러스인 방법.

**청구항 35.**

제28항에 있어서, 상기 바이러스 감염이 파라믹소바이러스로의 감염인 방법.

**청구항 36.**

제28항에 있어서, 상기 바이러스 감염이 HIV, 호흡기 세포융합 바이러스, 인플루엔자 바이러스, 코로나바이러스, 파라인플루엔자 바이러스, A형 간염 바이러스, 홍역 바이러스, 볼거리 바이러스, 웨스트 나일(West Nile) 바이러스, 뎅기 바이러스, 황열병 바이러스, 폴리오 바이러스, 헤르페스 바이러스 및 인간 유두종 바이러스로 이루어진 군으로부터 선택되는 바이러스로의 감염인 방법.

**청구항 37.**

제28항에 있어서, 상기 바이러스 감염이 중증 급성 호흡 증후군인 방법.

**청구항 38.**

제28항에 있어서, 상기 포유동물이 상기 올리고아데닐화 신테타제 1을 자연적으로 발생시키지 않는 것인 방법.

**청구항 39.**

암의 치료가 필요한 포유동물에게 올리고아데닐화 신테타제 1을 포함하는 조성물을 투여하는 것을 포함하는, 상기 포유동물에서의 암의 치료 방법.

**청구항 40.**

제39항에 있어서, 상기 올리고아데닐화 신테타제 1이 제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드인 방법.

**청구항 41.**

제39항에 있어서, 상기 올리고아데닐화 신테타제 1이 제20항 또는 제21항의 폴리뉴클레오티드에 의해 발현되는 방법.

**청구항 42.**

제39항에 있어서, 제19항 또는 제27항의 조성물을 투여하는 것을 포함하는 방법.

**청구항 43.**

제39항에 있어서, 암이 전립선암인 방법.

**청구항 44.**

제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드 상의 에피토프에 대해 지정된 모노클로날 항체.

**청구항 45.**

제44항에 있어서, 또 다른 별개의 제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드와 교차-반응하지 않는 항체.

**청구항 46.**

제약상 허용가능한 담체 및 제44항 또는 제45항의 항체를 포함하는 조성물.

**청구항 47.**

제20항 또는 제21항의 폴리뉴클레오티드를 포함하는 표적과 특이적으로 결합하는 단리된 간접 폴리뉴클레오티드.

**청구항 48.**

제47항에 있어서, 표적 결합 부위가 25개 이상의 연속적인 뉴클레오티드인 간접 폴리뉴클레오티드.

**청구항 49.**

제47항의 간접 폴리뉴클레오티드 및 제약 담체를 포함하는 조성물.

**청구항 50.**

서열 19와 특이적으로 결합하는 단리된 안티센스 폴리뉴클레오티드.

**청구항 51.**

제50항의 안티센스 폴리뉴클레오티드 및 제약 담체를 포함하는 조성물.

**청구항 52.**

제20항 또는 제21항의 폴리뉴클레오티드를 포함하는 표적에 지정된 단리된 리보자임.

**청구항 53.**

제52항의 리보자임 및 제약 담체를 포함하는 조성물.

**청구항 54.**

바이러스 감염의 치료가 필요한 포유동물에서 올리고아데닐화 신테타제 1을 억제하는 단계를 포함하는, 포유동물에서의 상기 바이러스 감염의 치료 방법.

**청구항 55.**

제54항에 있어서, 상기 올리고아데닐화 신테타제 1을 억제하는 것이 제46항, 제49항, 제51항, 제53항 및 제80항 중 어느 한 항의 조성물을 투여하는 것을 포함하는 것인 방법.

**청구항 56.**

제54항에 있어서, 상기 올리고아데닐화 신테타제 1이 제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드인 방법.

**청구항 57.**

제54항에 있어서, 상기 올리고아데닐화 신테타제 1이 제20항 또는 제21항의 폴리뉴클레오티드에 의해 발현되는 것인 방법.

**청구항 58.**

제54항에 있어서, 바이러스 감염이 플라비바이러스, HIV 바이러스, 호흡기 세포융합 바이러스, 인플루엔자 바이러스, 코로나바이러스, 파라인플루엔자 바이러스, A형 간염 바이러스, 홍역 바이러스, 볼거리 바이러스, 웨스트 나일 바이러스, 뎅기 바이러스, 황열병 바이러스, 폴리오 바이러스, 헤르페스 바이러스 및 인간 유두종 바이러스로 이루어진 군으로부터 선택되는 바이러스로의 감염인 방법.

**청구항 59.**

제54항에 있어서, 바이러스 감염이 중증 급성 호흡 증후군인 방법.

**청구항 60.**

인슐린 의존성 당뇨병의 치료가 필요한 포유동물에서 올리고아데닐화 신테타제 1을 억제하는 단계를 포함하는, 포유동물에서의 상기 인슐린 의존성 당뇨병의 치료 방법.

#### 청구항 61.

제60항에 있어서, 상기 올리고아데닐화 신테타제 1이 제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드인 방법.

#### 청구항 62.

제60항에 있어서, 상기 올리고아데닐화 신테타제 1이 제20항 또는 제21항의 폴리뉴클레오티드에 의해 발현되는 것인 방법.

#### 청구항 63.

제60항에 있어서, 상기 올리고아데닐화 신테타제 1을 억제하는 것이 제46항, 제49항, 제51항, 제53항 및 제80항 중 어느 한 항의 조성물을 투여하는 것을 포함하는 것인 방법.

#### 청구항 64.

제1항에 있어서, 상기 인간에서의 바이러스 감염에의 감수성을 확인하는 스크리닝 방법.

#### 청구항 65.

제1항에 있어서, 상기 인간에서의 당뇨병에의 소인을 확인하는 스크리닝 방법.

#### 청구항 66.

제1항에 있어서, 상기 인간에서의 정신분열증에의 소인을 확인하는 스크리닝 방법.

#### 청구항 67.

제1항에 있어서, 상기 인간에서의 암에의 감수성을 확인하는 스크리닝 방법.

#### 청구항 68.

제67항에 있어서, 암이 전립선암인 방법.

#### 청구항 69.

제1항에 있어서, 바이러스 감염에 대한 치유적인 치료에의 환자 반응성을 확인하는 스크리닝 방법.

**청구항 70.**

제69항에 있어서, 치유적인 치료가 인터페론-기재인 방법.

**청구항 71.**

제69항에 있어서, 환자 반응이 지속적인 바이러스 제거에 의해 측정되는 것인 방법.

**청구항 72.**

정신분열증의 치료가 필요한 인간에서 올리고아데닐화 신테타제 1을 억제하는 단계를 포함하는, 상기 인간에서의 정신분열증의 치료 방법.

**청구항 73.**

제72항에 있어서, 상기 올리고아데닐화 신테타제 1이 제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드인 방법.

**청구항 74.**

제72항에 있어서, 상기 올리고아데닐화 신테타제 1이 제20항 또는 제21항의 폴리뉴클레오티드에 의해 발현되는 것인 방법.

**청구항 75.**

제72항에 있어서, 올리고아데닐화 신테타제 1을 억제하는 것이 제46항, 제49항, 제51항, 제53항 및 제80항 중 어느 한 항의 조성물을 투여하는 것을 포함하는 것인 방법.

**청구항 76.**

제72항에 있어서, 상기 억제제가 소분자 약물을 포함하는 방법.

**청구항 77.**

제54항에 있어서, 상기 억제제가 소분자 약물을 포함하는 방법.

**청구항 78.**

제60항에 있어서, 상기 억제제가 소분자 약물을 포함하는 방법.

**청구항 79.**

제2항 내지 제8항 중 어느 한 항에 따른 폴리펩티드 상의 에피토프에 지정된 모노클로날 항체로부터 발생된 항체 단편.

청구항 80.

제약상 허용가능한 담체 및 제79항의 항체 단편을 포함하는 조성물.

청구항 81.

제1항에 있어서, 상기 인간에서 C형 간염 감염에의 감수성을 확인하는 스크리닝 방법.

도면

도면1a

<p>서열 1                      대립유전자 변이체 : C/T                      진뱅크 게놈 서열 ID : NT_009775.13                      게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2135728-2135728                      NCBI dbSNP ID (존재하는 경우): 7955146</p> <p>CCCTCAGAGTGACTGAAGGAAATTCAGAGAAGAGCTGACACCTAAGTTGTAG                      ATTTTGCCYGAACAGGTCAGTTGACTGGCRGCTATAAACCTAACCCCAAAT                      CTATGTCAAGCTCATCG</p>
<p>서열 2                      대립유전자 변이체 : G/A                      진뱅크 게놈 서열 ID : NT_009775.13                      게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2135749-2135749                      NCBI dbSNP ID (존재하는 경우): 3741981</p> <p>ATTCAGAGAAGAGCTGACACCTAAGTTGTAGATTTTGCCYGAACAGGTCAGT                      TGACTGGCRGCTATAAACCTAACCCCAAATCTATGTCAAGCTCATCGAGGA                      GTGCACCGACCTGCAGA</p>
<p>서열 3                      대립유전자 변이체 : G/A                      진뱅크 게놈 서열 ID : NT_009775.13                      게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2135978-2135978                      NCBI dbSNP ID (존재하는 경우): 없음</p> <p>ATGTATGGCCCTCCACCCAGGCCTGGTGGGTCCTGTCTCGACTGGGAGCAGA                      GGAGGGGTGGGGGAGGAGAGAAAGAAGGGAGTGAAGGGAAGAGGAGGGG                      GAGTGGTGGAGGAAATAG</p>
<p>서열 4                      대립유전자 변이체 : G/A                      진뱅크 게놈 서열 ID : NT_009775.13                      게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2144072-2144072                      NCBI dbSNP ID (존재하는 경우): 없음</p> <p>ACTGAATCCAGCTGCAATGCAGGAAGACTCCCTGATGTGATCATGTGTCTCA                      CCCTTTCAAGCTGAAAGCAACAGTRCAGACGATGAGACCGACGATCCAGGA                      SGTATCAGAAATATGGT</p>

도면1b

<p>서열 5                  대립유전자 변이체 : G/A                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2144088-2144088                  NCBI dbSNP ID (존재하는 경우): 2177979</p> <p>ATGCAGGAAGACTCCCTGATGTGATCATGTGTCTCACCTTTCARGCTGAAAG                  CAACAGTRCAGACGATGAGACCGACGATCCCAGGASGTATCAGAAATATGGT                  TACATTGGAACACATG</p>
<p>서열 6                  대립유전자 변이체 : G/C                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2144116-2144116                  NCBI dbSNP ID (존재하는 경우): 1051042</p> <p>GTGTCTCACCTTTCARGCTGAAAGCAACAGTRCAGACGATGAGACCGACGA                  TCCCAGGASGTATCAGAAATATGGTTACATTGGAACACATGAGTACCCTCATT                  TCTCTCATAGACCCAG</p>
<p>서열 7                  대립유전자 변이체 : G/A                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2144321-2144321                  NCBI dbSNP ID (존재하는 경우): 2660</p> <p>GGGCTCCAGTGTATCTGGACCAGTTCCTTCATKTTTCAGGTGGGACTCTTGAT                  CCAGAGARGACAAAGCTCCTCAGTGAGCTGGTGATAATCCAGGACAGAACC                  CAGGTCTCCTGACTCC</p>
<p>서열 57                  대립유전자 변이체 : C/G                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2131025-2131025                  NCBI dbSNP ID (존재하는 경우): 1015542</p> <p>TAACGCATGCCTGTAGTCCAGGTATTCAGGAGGCTGGGGCAGGAGGATC                  SCTTGAACCCAGGAAGTTGAGGTTGCACGAGTCATGATCATGCCCTGCAC</p>

도면1c

<p>서열 58                  대립유전자 변이체 : G/A                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2133961-2133961                  NCBI dbSNP ID (존재하는 경우): 757398</p> <p><u>GACAGGAAGTGTAACCTCTCAGAGGCTCCCTTGCCACATCAGGAGAATTGR</u>  <u>TAAAACCACACTACCTGTATCATATCAATTATTTAAGTGATAAATGATCA</u></p>
<p>서열 59                  대립유전자 변이체 : C/A                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2139587-2139587                  NCBI dbSNP ID (존재하는 경우):(알려지지 않음)</p> <p><u>TAGCATTAGGTATATCTCCTAATGCTATCCCTCCCAATCCCCCACCCMG</u>  <u>TTGTTGGTATTGTATATCTTCATTTGAGAATTCTCTGTTTCATGTCCT</u></p>
<p>서열 60                  대립유전자 변이체 : T/G                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2144294-2144294                  NCBI dbSNP ID (존재하는 경우):(알려지지 않음)</p> <p><u>GTGCATCTTGGGGAAAGGGCTCCAGTGTATCTGGACCAGTTCCTTCATK</u>  <u>TTCAGTGGGACTCTTGATCCAGAGARGACAAAGCTCCTCAGTGAGCTGG</u></p>
<p>서열 61                  대립유전자 변이체 : A/G                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2144985-2144985                  NCBI dbSNP ID (존재하는 경우): 7135577</p> <p><u>GAAAAATTATAGAACCTCCCTGTGTGACACAGCAGCCACTAGCCACATGTR</u>  <u>TCAAATGCTTAAAATGTAGCTAGTCTAAATCTACATGTGCTGTGAGTGCA</u></p>
<p>서열 62                  대립유전자 변이체 : C/T                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2156523-2156523                  NCBI dbSNP ID (존재하는 경우): 7968145</p> <p><u>ATGCTTCTATAGGCTTTTCTCACTGATGCTCTCTGGGCAGACAGGCTCCTY</u>  <u>AATATGAGAGTGACACACTCCTTTCTTCATTTTCAGGTAAACCTCACA</u></p>

도면1d

<p>서열 63                  대립유전자 변이체 : A/- (A 또는 A의 결실)                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2156595-2156595                  NCBI dbSNP ID (존재하는 경우): (알려지지 않음)</p> <p><b>CCTTCTTCATTTTCAGGTAAACCTCACACTGTTGGCAGAAGGAACTAT(A-)CCAATAATTAGTGAACATGCGGTGAATTTGCAACAGACAAGASGAGCCTC</b></p>
<p>서열 64                  대립유전자 변이체 : G/C                  진뱅크 게놈 서열 ID : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (개시-정지) : 2156638-2156638                  NCBI dbSNP ID (존재하는 경우): 7967461</p> <p><b>GAACTAT(A-)CCAATAATTAGTGAACATGCGGTGAATTTGCAACAGACAAGASGAGCCTCATTATCCATAGTTCCAGGTTGCTTAGGGAGGCAGAAATCAC</b></p>

짧게 한줄로 밀줄 그은 염기는 각 서열 번호에 의해 확인된 특정 대립유전자 변이체를 나타낸다. 서열 1, 2, 4, 5, 6, 60, 63 및 64 내의 두줄로 밀줄 그은 염기는 또 다른 제공된 서열 번호에 의해 확인된 정의 서열 내의 중첩 돌연변이를 나타낸다. 예로서, 서열 1 및 서열 2에 대해 제공된 정의 서열은 게놈 서열 내에서 중첩되므로, 이의 각 서열은 서열 1 및 서열 2에 의해 확인된 둘 다의 돌연변이를 함유한다.

서열 63에서, 돌연변이는 (A-)로 표시한 바와 같이 나타낸 위치에서 선택적으로 A 또는 A의 결실 상태이다.

동의성 핵산 코드:

- R=A/G
- Y=C/T
- S=C/G
- K=G/T
- M=A/C
- W=A/T



도면2b

2133300 TTTTCAGGATCAGTTAAATCGCCGGGAGAGTTTCATCCAGGAAATTAGGAGACAGCTGGA  
 2133360 AGCCTGTCAAAGAGAGAGAGCATTTTCGTGAAAGTTTGGGTCAGGCTCCACGCTGGGG  
 2133420 CAACCCCGTGCCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGAGGGGGTGGAGTT  
 2133480 CGATGTGCTGCCTGCTTGTGATGCCCTGGGTGAGAGCTCCACAGCTCTTTTCTCCCTCT  
 2133540 TCCCATTTCTGAGCAGAAAATCTCCACAGTTTGGAGCTTTTGGCCCCAACAGGGCATCT  
 2133600 CTCTAAAGCAGGGTGGGAGGAGATCTTAGGATCTGTCCCGGGGCAAGAAATGAATACGGTCT  
 2133660 ATGATCTATCAACAGGAGAGACATTAACAGCAAAATGGCATATATGGGGCAAAAGACAT  
 2133720 TTCTTACAGAACATCTGCAAGGCTTACTGGTCTGTTTAAAGGCAAAATGTGTGAATTTTA  
 2133780 TCTTTCTAAAAACAGGCAGCAAAAGATGTGGCTTAAAGTTTCATGTACTCTCATCTTTGTC  
 2133840 CCAACATGAGATCTCATCAAACGTATGACGACGTTGGGAGATAGATATTTATAAATTTGC  
 2133900 AGGAACATTTGGACAGGAAGTGTAACTCTCAGAGGCTCCCTTGCCACATCAGGAGAAAT  
 2133960 GGTAAACACACTACCTGTATCATATCATTTTAAAGTGAATATGATCATCTCATCT  
 2134020 CAGCTCTGATGAGTAATAGGTGTTCAAAAATAGGAACTTCCAGCCAAAGTGTGGTGGCTCA  
 2134080 TGCTTGTAAATCCAACACTTTTGGAGGCTGAGGCAAGGAGGCTGCTTGAGCCAGGAGTT  
 2134140 CAAGACCAGCTGGGCAGCAAAAGTGAACCTCATCTCTACTAAAAATTTTAAACATAG  
 2134200 CCAAGTGTGGTGTACATGCCCTGTGGTCGAGTTATTTAGGACGCTGAGACTGAACGATC  
 2134260 ACATGAGGCCAGCCAAAGGATTCGAGGTGTCAGTGAGCCAGAAATGTAACCTGCATCCCA  
 2134320 TCCTAGGCAACAGGCAAGCAAGCCCTGTCTCAATCAATCAGTCAATCAGTCAAAACT  
 2134380 ATGAATTTCCAGCTGTATATGAAGGCACCTCAAAACACACAGTGAACCTCAGAGAGCA  
 2134440 CACGGAATAGTTTAAATTTTAAATTTTGGAGGAAAATGCGATGACATCTGTCAACACCG  
 2134500 CACAAACGGCTACTATTAACCTGAACCTACTGATTTAGTGGCTACTAATTAATAGTTGGTC  
 2134560 ATTAAGCAGTAATTAGTGTAAATTAACAAGTAATTAGGACTAATTAAGGAACTGTC  
 2134620 CAGTTTCTTTAGTCTTAGGGCAGCCATGAAAAAAAATGCTGACTCTCCAAGACACC  
 2134680 AGGGTATGAGAAAGTTTGGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  
 2134740 TACAATGGTTGTAAGAACAAAGAGGAGAGGCTGGTGCACAGTGGCTCACGCTGTAATC  
 2134800 TCAGCACTTTGGGAGGCCAAAGTGGGGGGATCACTGAAATCAGGAATTTCAAGACCAGCC  
 2134860 TGGCCAACTGGTGAATCTCACTACTAAAAATACAAAAATAGCTGGGGTGGTGG  
 2134920 TGTGTGCTGTAAATCACAGCTACTCGGGAGGCTGAGGCGAGGAAATGCTGAAACCAGG  
 2134980 AGATGGAGGTTGCAATGAGCCAAAGATCATGCCATTGCACTCCAGCTGGGCAACAGAGTG  
 2135040 AGACTCCATCTCGAAAAAAGAAAAAAGAAAGAAATATAAGGAGTGAATAAAAGGAAAA  
 2135100 GAAAGAAAACTAAAGTGGTGAACAAATAGATAGCCATGGGGGTTAGGGAGCTTTTAA  
 2135160 GACAGGGTCTGAGGGAGGGTCCCTGAGCTGAGTGGCGAGAAAGGAGTGAACCTGGGGA  
 2135220 GATCTGGAGGTTCTGGGAAGAGAAATGGCAAGTGCAGAGGCCCTGAAGCAGCAATGACCA  
 2135280 TGGCACATTTGAGGAAGAGAAAAAGTCAAGAAAGTAGAAAGTGGGCAAGGAAAGCAAG  
 2135340 ACAGGAGGTGAGGTGGGAGAGGTTCCAGAGACCAGATCACACCAGACATCATTTGGCCACC  
 2135400 ATAAGATCTTTGGGTTTAAATTTCCAGATGTTATGGGATGCAGGAAGCAGCATGATCAG  
 2135460 CAGCATTCTAGGTGCCAGGTTGAGAACAGGCTGTGGGGGAACCTGTAAAGAGGTTGCT  
 2135520 GCCATAGTTCGGCGAGTGAACGTTGGTGGCTTGGATGGGGTGTGTCAGTGGAGGGGCA  
 2135580 GGAGGGAGGATCAGGAAATGGACCTCAAGACTTCCAGCCCTGGGCTGCTGCACTTTTCA  
 2135640 ATCAAACCCATGGCCAGGGAGATTGTCCCTCAGAGTACTGAAGGAAATTCAGAGAAAG  
 2135700 AGCTGACACCTAAGTTGTAGATTTTCCCGAACAGGTCAGTTGACTGCGGCTATAAACCC  
 2135760 TAAACCCCAAACTATGTCAAAGCTCAACGAGGAGTGCACCGAAGTGCAGAAAGAGGGCGA  
 2135820 GTTCTCCACCTGCTTCAACAGAACTACAGAGAGACTTCTTGAAGCAGCGCCCAACCAAGT  
 2135880 CRAAGGCTCATCCGCTAGTCAAGCACTGGTACCAAAATGTATGGCCCTCCACCCAGC  
 2135940 CTGGTGGGCTCTGCTCAGCTGGGAGCAGAGAGGGGTTGGGGGAGGAGAAAGAGGGG  
 2136000 AGTGAAGGGAAAGAGGAGGGGAGTGGTGGAGGGAAAATAGAGGGATGAAAAAGGAGAGAA  
 2136060 AGGAAAAAGAGGTGGAGAGAGGAGCCCTGCAACAGAAAGGAGAAATGAAAGGGAGGAAAG  
 2136120 AGAAAGGAGGGATTTGGTGTCTGTCTCACTGCTGTATCCCCAGAACTTAAACAGAGC  
 2136180 CTGGTGCATAATAGGTGTAATRACTGTTGAATAAATGAATCAATGCTACATACACAC  
 2136240 GCACGCACACACACAGAGAGAGGTCACACCACTCTTCAAGGTTGGATAAGTTAAA  
 2136300 ACAAGAGTTTCAAAACAATAATATGTTAGATGCCCTTCTCCACTTACTGGCTGGCTG  
 2136360 GCCTTAAAGTAAGCAACTTAACTTTCTGTTCTTTCTGCTTTCTTATCTGCAACGAGTAGC  
 2136420 ATGCCATAGCTAGAGTAAACAGGCATATAGTTGGTCTGATAAATGATAGCATATTTTAGC  
 2136480 CACCATAGGAGTACACATAAATAAGCTAACATGTAGTATGTGCTTACTTATCTATGTT  
 2136540 TTGTGGATGTGATCAATTTCTGTTCACTTTTAAATGCCCTGCATCTTAGTCAATTTTA  
 2136600 ACAGTGATTTCTGTAAGTGTAGATAAGGTTAGGCATTTATTAATTAATCCATTTTACCAAG  
 2136660 AGAAACTTGGGTCAAAAAGAGAACTCCTGGTCACTGGCTCATTCGGCCAAATAAGTAG



도면2d

2140140 CTATTCTGTTCCATTGTCTACGTGACTATTTTATACCAGTACCATGTTGTTTTGGTGA  
 2140200 CTATGGCCTTACAGTATAGTTTGAAGTCTGATAATGTAATGCCTCCAGATTTGCTCTTTT  
 2140260 TACTTAGTCTTGCTTTGGCTATGTGAGCTCTTTTGGTGCCATATGAATTTAGGATTG  
 2140320 TTTTTCTAGTCTGTGAAGAATGATGGTGGTATTTTGATGGGAATGCATTGAATTTGT  
 2140380 AGATTGTTTTGGGAGTATGGTCATTTTCAAAATATGATTCTACCCATTATGAGCATG  
 2140440 GGATGTGTTTCCATTGTTTGTGTCACTATGATTTTCTTTCCAGCAATGTTTTGAGTTTT  
 2140500 TCCTTGTAGAGTTCCTAGTTATTTAAAGTCTGTGTTCCGGTCTTTCAGCATTTAAAGTTT  
 2140560 GTAGGTTTATTACTATTTCTCTTCTTCTGTTGGTCACTAACTCTTAGTGTTTTGTTCCT  
 2140620 TGTGTGCCTGGTTACATATGTGCTGGTCATTGTATTTGAAAATATGTGTGAATAATTT  
 2140680 GAGGTTTGGATTATGTATATTTCTCCAGAAAGAAATTCATTTGCTTCTGTGCATTTCTT  
 2140740 AGGAACATTACAAAGTCTTCTTCTCAGTTAATTTTCGTAGTATCTTTATCAGATAGGTGC  
 2140800 TATTAACAACCTCACTTAGCAGATGAAATCATGAGGCTCTGAGAGTCTAAGTCATCTA  
 2140860 CTTAGAATTTGCAAAATGGTGAAGCCAGGATTCAAACCCACATCAATAAGAATCCACGGCT  
 2140920 CTTAAACAAGGGCCAGTACACTTTTTTAAAAATAAAAGGCTAGATAGTAAATATTTTAG  
 2140980 ACTTTGTGGATGTCACAGCCTCTGTTGCAACTACTCAACCCCTGCCTTTGTAGCATTAATG  
 2141040 CAGTCATAAATACATAAATGAATGAGCCTGGATTCTGTTCCAAAGGAACTTTATAAAA  
 2141100 ACAGGTGGCAGCTGGAATTTGGCCATGAGAAAGTGTAGTTTACAAAGGTTGAGCAAAAC  
 2141160 CAATTTTTTCTGATTGTTTTCTCTTCTCAGTGTAAAGAAAGCTTGGGAAGCTGCCA  
 2141220 CCTCAGTATGCCCTGGAGCTCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAAAACA  
 2141280 CATTTCACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCATAAACTACAGCA  
 2141340 CTCTGCATCTACTGGACAAGATTTATGACTTTAAAAACCCCATTTATGAAAAGTACCTG  
 2141400 AGAAGGCAGCTCACGAAACCCAGGTATGCTATCCCAATGGCTTAGCTCCCTCTGATA  
 2141460 ATGAACACCTGGATACAGGTACAGTGCCTTGGAAATGGAGGAGGTGGGAGGGCTCCCCAC  
 2141520 TTAGTGAGAATCTCCTGTTGCCATCATGTACTGGGCATTTTACTACTGCCACTGTTT  
 2141580 TAAACCTTACCTCCAACCTGTGAGGCAGGCCTATGCCAATTTATTTACAGTGAATG  
 2141640 AACTGAGGTTCTGAGAGGTAAAGAGCTTGTCCAACCTTAAACAGAAAATGATGAAATAG  
 2141700 CTGCAGTTTGAATGAAATAAGAACAGCAGCAACAACAATGATAGTAAATTTGCTCCAGT  
 2141760 ATTGAAAGCTTGTGTAAGACTAACACATGCTAATATAATAGTAAAAATTTATAGCAATA  
 2141820 TTACTGATATGTATGTTATGTTCTAGTCCGTGCTGAGCATTTCAATAACTGGGCTTT  
 2141880 TTCTATCCTCAGCATAGCCTTTGAGATAGTATGTGGAATTTTCCATTTTACAGAT  
 2141940 AAGAACTCCTGAGGCTTAGAGAGTTCAAGTGACCTACCCAAAGGGCAGCATCACTGATAAAG  
 2142000 GCAGAGGTGGGATTCAAACCCACATCTGTGAGGTGCAAGTGAAGGCTCCTTCTCCTCAT  
 2142060 GCTCACTGCCTGCTGGGAAATAGGGCCTGTTGGGACATACCCAAGGAGCCCTTCTCATG  
 2142120 TTCTGAGTCCAGTTCATCCCATGCTGCTATTTGCTCTCCAGGAGCATCTGGACTCCC  
 2142180 TAGACAGAGCCCGAGCTTCTCACCTGCTCCCTCTCAAAATGCTGCTCTGCAGGCTGTGAT  
 2142240 CCTGGACCCGGCGGACCCCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGAGGGCA  
 2142300 GCTGGCACAAGAGGCTGAGGCTGGCTGAATTACCCATGCTTTAAGAAATGGGATGGGTC  
 2142360 CCCAGTGAAGTCTGATTTCTGCTGGTGAAGCCTCCTGCTTCTCCTGCCATTCATCCC  
 2142420 TGCCCTCTCCATGAAGCTTGAGACATATAGCTGGAGACATTTCTTTCCAAAGAACTTAC  
 2142480 CTCTTGCCAAAGGCCATTTATTTCAATATAGTACAGGCTGTGCTCCATATTTTACAGTC  
 2142540 ATTTTGGTCACAATCGAGGGTTTCTGGAATTTTCAATCCCTTGTCCAGAAATTCATTTCC  
 2142600 CTAAGAGTAAATAAAAATAATCTTAACACCAATTTATTGACTGTCTGCTTCCGGCTCAGG  
 2142660 TTCTGTCTAAGCCCTTTAATATGCACTCTCTCATTAATAGTACACAAATCCCATGAG  
 2142720 GCATTTTTAAAAATTTTTATATTTTATAGATTGAGGGGCAATGTCCATTTGTTACAC  
 2142780 AGCTATATTTGTAAATGTTGGGTTTTGGGCTCTATTGATCCTGTGCCCCAAATAGTGAA  
 2142840 CAGAGTACCCAAAAGAATTTTTCAACCTTTGCCCTTCTCCCTTCTCCTTCCCTGTGTTGG  
 2142900 AGTCCCTAGTGTCTATTGTTCCCATCTTTAGCAGATGTTAAGTATTTGATTTTCTGTTT  
 2142960 TGGGTTAATTCACTTCGGATAATGGCTCCAGCTGCAACCATGATTTTCTTTCTTTAT  
 2143020 GGCTGCATAACTCCATGGTGTAGATATACCAACTTTCTTTATCCAGTTCACACTGAT  
 2143080 GGGCACTTAAGTTGATTCATGACTTTGCTATTGTGAATCGTACTCCGATAAACATACGA  
 2143140 GTGCCGGTGTCTTTGATAGAATGATTTCTTTACCTTTGGGTAGATACCGAGTAGTGGGA  
 2143200 TTGCTGGGTTGAATGGACATTTACTTTTATGTTATTTGAAAAGTCCCATGAGGCATGTT  
 2143260 TCTATCATTTCCATCTTACAGATGAGCAAAAGGCTCAGAGAGGTGAGGTCATTTGCTCAA  
 2143320 GGACATCAGCTAACAGTGGTGGAAATGGAAATCAAGCTCAGTGGACTTAAAGCCAGTG  
 2143380 CTCATGTCACTGTGCTAAACAGCCTGCTTGTGCATCTCCCACTCTCATCTGACCAATG  
 2143440 GGAGACTCTGAGCAGCTGAGTGACTTGGGTTGTCAACAGCTAAACAGGGGCAAGGACC  
 2143500 CAGTCTTGGATCTTCCACCTCCAAGCAGGAATCTGCTGATTCAGGGGATGATGATG

도면2e

2143560 TTGCAGATGGCTAGGAAGCAGACTCCAGGATGGAATTTAGTATGCAGGATGTTCTGGGGG  
 2143620 AGAGCCACTGGAAACAGCACTCAGGGAAAGGGGGAAAGAAAGGATAGGAAGGAAGCATGA  
 2143680 AAGAGAATAGGGAGAAGTGAACAGGGATGCAGAGCGAATGCAGTTTCAGCCAACCTCAA  
 2143740 GGACAGCCCTGGAGCTGGAAATGGCCCTTAGAGCTGCCCATGGTACAGAGGTTGGCCAGG  
 2143800 CTTCTATACCCCTAAGTGGATCACTCACTGTGCTTGGGCACCTTGGGAAAGGGCATGGCT  
 2143860 TTGAGCAAAGGCTCTCTGCAGCTGAGGCAACCCCTAAAAGGGCTGACGGCTGAAGTCTG  
 2143920 TCTGCTGACCACTGTCCAGCAGCTGGGGCTTGTAGTCTTCTCCTCAAAGGGGATCCAG  
 2143980 ATGGCATGTACAGTGTCTACCGTAAATGCTCACTGAATCCAGCTGCAATGCAGGAAGAC  
 2144040 TCCCTGATGTGATCATGTGTCTCACCCCTTTCAGGCTGAAAGCAACAGTGCAGACGATGAG  
 2144100 ACCGACGATCCCAAGAGGTATCAGAAATATGGTTACATTGGAACACATGAGTACCCCTCA  
 2144160 TTCTCTCATAGACCAGCACTCCAGGAGCATCCACCCCAAGCAGAGAGAGGATCGG  
 2144220 ACCTGCACCATCTCTGAATGCCAGTGCATCTTGGGGAAAGGGCTCCAGTGTATCTGG  
 2144280 ACCAGTTCCTTCATTTTCAGGTGGGACTCTTGTATCCAGAGAGGCAAAAGCTCCTCAGTGA  
 2144340 GCTGGTGTATAATCCAGGACAGAACCCAGGCTCCTGACTCCTGGCCCTCTATGCCCTCT  
 2144400 ATCCTATCATAAGATAACATTCTCCACAGCCTCACTTCAATCCACCTATTCCTCTCAAAGTA  
 2144460 TTCCTTGAGAGAGAACAGAGAGATTTAGATAAGAGAATGAAATTCAGCCCTTGACTTCTC  
 2144520 TCTGTGCACCTGATGGGAGGTAATGTCTAATGATATTAATAAACAATAAAAAATGAAGC  
 2144580 AAATACCATTTATTGGGTGTTTATTAACCTCAAGGCACAGAGCCAAGATACAGATGCA  
 2144640 TATCTAGGGGATTTGTGTGTATATACAATTGATTCAACAAGAAATATTTATTGAGCACT  
 2144700 TACTATGTGCCAAGCATAGCTCTGGGCACTGGGAATATAGCAATGCAAAAAGCAGACAG  
 2144760 AAAATCCCTGTCTCATGACCCTGCAGAGCCAAGACTCCAGAAATTTTAAATAAAAAA  
 2144820 TCCCTGTCCCTCATGGAGTTGACATTTGTGCAAAACATCTTAATGTTAGATGGTTTCTTA  
 2144880 TTACTATAATTTCTGAAATAGCATCCTTGATTTATCCTTTCTCCTATCTCTGAGAAA  
 2144940 ATTATAGAACCCTCCCTGTGTGACACAGCAGCCACTAGCCACATGATCAAATGCTTAA  
 2145000 TGTAGCTAGTCTAAAATCACAATGTGTGTGAGTGCAGGTAATATACTTGGTTTCAAAGAC  
 2145060 TTAGTACAAATGAAAGAAATGCCAAGTCTTGGCCAACTGATAATTTTAAATTTGTTGTC  
 2145120 TGAATGACAAATTTTAAATATATTTGAGTAAATCAAATGAACTTCATCTCTTTCTTT  
 2145180 CCCTTTTAAATTTGGCTACTAGAAAATGTGAAATCATACATGTGGCTTGTGTATATTA  
 2145240 TGTATTTCTATTTGGACAGCTCTGTCTTCCAAGGTAATCACTGGATTAAGATTCGACTA  
 2145300 TACTGACTTACATTGCCACATTTGTCACTGTCTTGGGACCAAGAAATCAACATATCATT  
 2145360 CATAAGACTCTAAAATATAAACTCTCATAAATCTCACAAAAGAACCTGACATGCTCTG  
 2145420 ATCACCTGAGTTCCTGGTCACTTTTGGTGGCTGGTAAGCAGCCTTGGTCCGTCCAGAT  
 2145480 ATATTCTTCCATTTAGTCCCACACATCCCTGTGAGATGGGTTTGTGTATTTCTCATA  
 2145540 ATATTAAGTGGAAATAACTTGAGGTTCTAAGAGGTTCAAGTGTCTTGGCCAGGTCACCCAG  
 2145600 CTGCTCAGGGGCAGAGTCTGAACCTGAACCTAATCCTTCTCTCTCAAAGCTCATGTTCT  
 2145660 TTAATCACTGTAGCATGTTCTAATGTCTCCTCATTGAAAGCTTATGTTTCTCTAC  
 2145720 TCTGGCCCATGAGAACCAGAAGCATCAAATGTCCAGGGGCAGGAAAGATGAATGTCCCA  
 2145780 GCTCAACCTGAGCAGATTCAACCTCTCCTGGCTCTTTTGTCTTATTTGTAGACTGGAT  
 2145840 TAGATGATGCCAGTCTACTGATTCAAATGTGAATCTCTTTCAGAAAACCCCTCACAGATA  
 2145900 CACCTAGAAATGATGTTTCAACAGCTATCTTGGCACCTTAGCCAGTCAACTTGTGACA  
 2145960 TAAAATCAATCATCACACTCCATGCTGATAGGCAAGTGTGGACATCCCAATGTAATGG  
 2146020 CTTCAATGTATTTTACTGTGTGAAATGCACTTGTGTGCCTTTTGGAGAGTGTTCATT  
 2146080 TTAATAGCAATGCCAACCAACAGTAGATTAATGGAATCAGTAATTAGTGTCTGATCGA  
 2146140 AGAGCCACATGGCCACGTGATCAGCCTTCCATCTAACAACAGGACCCAGGAGTATACCGGG  
 2146200 ATTGTTTTTCAAAGGGCATAGACATTTTGTCTGCAATGACATGGGCTTACTCCAGAGTC  
 2146260 CTGGAGGGTCTGTGTTAATTTCTCAAATAGATATTGCCATAATCTCTGAAATGACACT  
 2146320 TTTCCCATGACTAACAATTTGAACACCATGGGGTCTGCCAGGCTGGTGTGGCCAAAGTA  
 2146380 GAGGGGCGACTTGCACCACAGCCTATACCAGCTGCAGAGCCCTTAGGACTTAATAAAG  
 2146440 GGTGCTAATTTCTGACTTCCCTGGCTCTGAGATGTAATAGTGGTTTAAATTTACTATCC  
 2146500 TGGCCAGGAGGTTGGCAGTTCAAAGGATCCCAATGACCTTCTCTCACTGTGATGACCT  
 2146560 CACTTAAACCTCAGCCAAATGTCGGGTAGTGTCTTTTACCAAGCATGTCTTTTGAATA  
 2146620 TACATTCAAGGGACAGGAAATGATCACCAGTCCGTCCATAGACACAGTGGGCAAAATGA  
 2146680 CAAGCCTGACTTGGCCAGGGCTCCATTTTCACTCTTGGCTCCTATAAGCCCTACTCTG  
 2146740 ATGGGGACAGAATCTGATGACGCTTTTCGTTATCAGTGTGATCCTCTGCCCAACAGTCT  
 2146800 TAGAAATGTTGAGTACTCCCTTTTCAAGTGTATCATCTCTGAAATAAATGGCCATAG  
 2146860 TCCTTTGGGGAAACCAATTAATAATATGCTATGGTGTACAGCATCTTTCCCAAGGGG  
 2146920 ACTCTGCCCTCCTTTTGAATGGGCTAGTCTGAAAACCTGGGTTGGTCAAGTTCCAAAA

도면2f

2146980 ATCAGGAGGGATGATGACTTTTTATTGGAATGGCTCCCCTTAAACCCCTGGTGGTCCGG  
 2147040 CCTTGATTCTCTCTAAATGGTATAAATTTAGTAGTGGCCCTTTCAAGATACTCATCCACTAA  
 2147100 GTCACCTAAGAGGCCACATTTGTGTTAACCAGGTCCCTAACCCGTATCAGTCAAGACCCCT  
 2147160 GGCTGGCCCTCCAACCTCACCATTATTATAGTCAATGGCACCACCTGGCTTCTATTTGCT  
 2147220 CTGGGAGTCCCTGGCTGCCCTATTGCCCTTATCAATCCAAGCTCTGTAAACAGCCCTCTTC  
 2147280 CCAGGAGGCCACCCAGAACACCTCATGATGCAGGTGTCCTCTCTCCACACATTCCTT  
 2147340 ATGGCCTCGGGGAATGTTTTGTCCCTGGGACTCTATTTCCAGACTTACATAGTAAACAT  
 2147400 ATCCACTCCTCTCACTGTGCCATTTCCCTTGGCCCTTAATACTCTTCCATCAGT  
 2147460 CCACAGCAATTCGGAAATTCATACCTTTCTGCATTTCTAGGAGCCATCCTAGGAGTCAG  
 2147520 GGAAAACCTCGGCTTCACAAGTAAATGGGACTTCAGTTGTGTGCCAACATGCCTGGCTAA  
 2147580 TTTTTAAATTTTTTTGTAGAGATGGGTCTTGTGTGTCCAGGCTGGTCTTGAACCTC  
 2147640 CTGGCCTCAAAACAATCTTCTCCCTTGGCCACTCAAAGTGTGGGATCACAGCTACCAT  
 2147700 CCCACACACATTTCTTATTTTAAAGGCTGTATGTGCACTGTATACATTAATTTGTTGTC  
 2147760 ACTGTATACATTAATGTGTGCACTGTATACATTAATTTCTTTAACGAATTTATCCAT  
 2147820 TATAGTTGCTTTGGTTGTTCCACTTGATTTATGTGAATAGTGTGCAAGTAAACATGGGA  
 2147880 ATGCAGTTATCTCTTTGATACTCTGATTTCAATTTCTTTGGGATACTCAGAAGTGGGATG  
 2147940 CTGGAAACATATCGTAGTTCCATTTTTAATTTTTTGGAGAACCTCCACTACTGTTTTTCA  
 2148000 GTGGCTCAACACAGTGTGCAAGAGTTCCCATCCCTCCACATCCCTCACACTGTATCTT  
 2148060 TTTGTTTCTTTTAAAAAATGATAGCCATCTACCAGGAGTAAAGGTGATATGCACTCGT  
 2148120 ATTTTGATTTGCATTTCTCTGATAATAGTGTATTTGAGTATATTTTCTAGACCTGTFA  
 2148180 TCCATTTGTTGCTCTTCTTTGGAGAAAGATCTATTCCTATCCCTTAGCCCATTTTTAAATC  
 2148240 AAGTTATTAATTTTTTTGCTAGTGAAGTGGTAGGAGTTCCCTACATATTGTGGAGATTAAC  
 2148300 CCTTATAGATGATGGTTTGGTAAATGTTTTCTTCACTTCCATAGATTGTCTTTTCAGCCC  
 2148360 GTTGATTTGTTTCTTGTCTATGCAAGCTTTTAGTCTGATGATGCAAGTGTGCTAT  
 2148420 TTTTGTCTTTTGTGCTGTGCTTTGCATACGTGGCCACCTGATCTTTGACAGATTTGCCA  
 2148480 AGAATACACAATGGGAAAGGACAGTGTCTTCAACAATGGTGTTTGGAAAGCTGAATGT  
 2148540 CCACATGCAAAAAGATAAAAATGGACCCCTACCTTACAGCATACACAAAAATCAACTAG  
 2148600 AATGGATTAAGACCTTAAACGTAAGACCTGAAACTGAAACTACTAGAAGAAAACTTAGGG  
 2148660 GAAAACTTATGACATGGTCTTCCAGTATTTCCAGGATGTGACACCAAAAGCACAGA  
 2148720 CAACAAAAGGCATTTATTTTATATGCAATGTGAGGAAAGGGGTTCAGTTCAGTCTTCC  
 2148780 AATGGATGCTCAATATCCAGCAGCATTTATTTGAACGGATCATGTTCTCTCCACTTC  
 2148840 TTTGCAAAAGCCACCTCTTAAATATTTCCAGAGCCCATCTATGTGGGAGTCTGTTCTGGA  
 2148900 CTCTGCTCTTTTCCATGGTCTATTTTTTGTGCTCTTGAATTAACACAACCTTGTCTTAA  
 2148960 TTAATAAACCCTTAAATTTCTAGTATCTTTGGGAGAACTCTGTTCTCTTTTATCAAGT  
 2149020 CATTGGTTCTCTTGGCCCTTTTTATTTCTACATTAATTTTATACTCAATTTGTGAAGT  
 2149080 CCTCCAAAATATGGGGAGGCATTTGATAGAAATACACATATTAGCTTGGGAAGATAA  
 2149140 CATCTATTTTTATATATATCTATATAGAAATTTGATGATTTCTGATCCATGAACTTGGGA  
 2149200 GTTCTTTTTCAATTAATTTTTGTCTCTTAGTGACAGTGACAATGTTTTATCATTTTTCCC  
 2149260 TGCAGAGGCTTTTATGTTTGTGCTATATGATTTCCCAAGTATTTGTTTTCAGTACTATC  
 2149320 GTACTTGGCATGCTTTCTTAAATTTCAATTTGCAATGGTGTCTCATGGGAGACTTAAAC  
 2149380 ATTTTCAAGTGTGATATTTGGCTATAGCATTATAGGTGATCTTTTATCTTTTCTAATTT  
 2149440 TCTGTGGCCACGGGAAATAAATAAGACACTTTTCTGACAGAAAAAAAACTTTTGGG  
 2149500 AAGTATTTCTCACAGCTAAGATCTGATGTTTACGCAAAGTTGGCAGGCAACAGGCTACAG  
 2149560 AAAGCTCTGGGTGCTGTTGTTTGGAGCTGCTGGTCAAGGACAAATTCACAAGATTTGG  
 2149620 AAACAGAGGACCAAGTGTGTAAGGACGAGGGAAACTATGGTATAACATTTGAAGCACCTGA  
 2149680 GCTGGAATTTCTGAGCCCTCAGAGATAAATTTCTCAGCTCCTCCCTGCCAGAAAAACA  
 2149740 AAATAAAAAGATTAATGTTTAGCCACAGAAATGAGAGTGAAGTTCAAGAGAAATTT  
 2149800 AGGCACAGCTGGAAGTGTCAAATGAGGAAACGTTCAATGTCAAGTTTGAAGTCCAGG  
 2149860 GTTGATGGGTGAACCTGCAACGCTCAGCTTCAATTTAGGCTCCAGCTCAACGAGGGGG  
 2149920 TGAATTTTGGTACTGCTGTCTTTGAGGCCCTGCGTGAACCTCCACACCCCTCGCCCT  
 2149980 TTTCTCATACAAATCCCTCCTTGCAACCCCTCAGCCCTGTCTGTGAGGTGCCAGGGC  
 2150040 CCCTCTCCAGCCAAACCGCAGGCCAGTATGCCCCCACAACCTCCCTTCAGGCAGAT  
 2150100 CAAACCCAGGGCTCTGAGTCAACCTGCCTGGGCTCAAATCCTGCTTCTGAATCTTATGA  
 2150160 GACATCAAGTCACTTACCTAATCCTTGGTGAACAGGCTCTCCTCTGTCAAATAGGTGT  
 2150220 GCATGCTTTGGGAGGCCAAGCAGGAGGATTTGCTTGGAGCCAGGAGTTCAAGACCCAGCT  
 2150280 GGGTAAACATAGTGAGACCCCTGTCTATAAAAATAAAAAATTAGCCAGGTGTGGTGGCT  
 2150340 CATGCTGTAGTCCAGCTACTCAGGAGGCTGAGGAGGAAAGATTCCTTGAGCCACGAA

도면2g

2150400 GTCAAGGCTGCAGTGAAGCCATGATCAAGCCGCTGCACTCCAGCCTGGGCAACAGAAATGCG  
 2150460 ACCCTGTCTCAGAAATAAGTAAATAAATAAATACATAAAATAAATTGGGTGCAATTGTGG  
 2150520 CTCTTAGTGTACAGAGAGGACTGAAGGAGCTAATGGATGATGGATTATGGCAGTACCT  
 2150580 CACATATAGCTTATCCTAAAGGAAGTGTAGAGCTTATTATGATGATTATCAAAAAATATTT  
 2150640 ATCAAAGGCTGCCCTGGGCCATGTCTGAGCTAAGTGCTGGGGATGCAAAAGTAGACAA  
 2150700 GAGACTCCTCAGGGACAATTTGTCTGATGAGATAACAGGCACATTTATGAGAGGTCACAT  
 2150760 CAATCAGTCTTATTTATCTTATAATTTATCCATAAATGATATAATAATTTATGAGGG  
 2150820 CCAATAAATCTGATGGCAGGAGCCTGTGGGGGTAATGGCCAGGTCTCACATTTGTGCCC  
 2150880 AGGCTGGTCTTGAACCTCTGGCCTCAAGCATTCCTCTCGCTCCACCTCCAGCATGCTG  
 2150940 CAATTAACAGAGGCATGAACAACCTGCACTGGCCTAAAAATTTATGTTAATAAAAAAATGC  
 2151000 ATGTATTTGAGGAGTACGACATGATGCTTGAATATCATACTGTATCGGGGAAACAGCC  
 2151060 CCCGATATTTCAATGTAGGTTCTTTTCTATTTTCCCAAGTGTCCGCTGGTCTGAGAAAT  
 2151120 AAAGGGAAAGAGTACAAAAGAGATAAATTTTAAAGCTGGGTGTCCAGGGCAGACATCACA  
 2151180 TGTCCGCGAGTTCTGTGGTCCCCCTGAGCCATAAAAACAGCAAATTTTATTAGCAATC  
 2151240 TTCAAAGGGAGGAAATGTACATATAGGGTGTGGGTACACAGAGAACACATGATTCAAGAGT  
 2151300 GACAAAAGATCAACAGGCAGAGGTCAGGGTGTAGATCACAAAGTCAGGGCAAACTAGAA  
 2151360 TTACTAAGGAAGTTTCAATGTTCCACTGTGCATGCATTTGATTAACACTTTAACAG  
 2151420 TGTTCAGAGCAGAGAACAGTCTGACTAGAATTCGCCAGGCTGGAAATTCCTAATCCTA  
 2151480 GCAAGCCTGGGGTGTCTGAGGAGACAGGGCGTGTTCATCCCTTATCTGCAACTGGAT  
 2151540 AAGGCAGACACCCAGAGCGGCCATTTAGAGGCCCCCGGAAATGCATCTTTTCCCA  
 2151600 GGGCTGTTAATTTAATATTCCTTACTGGGAAAGAAATTCAGGGATATTTCTTACCT  
 2151660 GTTTTGGTAATAAGAGAAATATGGCTGTCTTTCCTGGCTCCAGGCAGTCAGACCTA  
 2151720 ATGGTATCTCCCTTGTTCCTGAACATCGCTATTAATCTGTCTTCTTTCAAGGTGCC  
 2151780 AGATTTCAATTTGTTCAAAACACACATGCTTACGAAACAATTTGTGAGTTAACGCAATCA  
 2151840 TTCAGAGGTCCTGAGGCAACATAACATCCTCAGCTTATGAGATGACAGGATTAAGAGAT  
 2151900 AAAGACAGACATAGGAAATATGAGATTTGATGAGGAAGTGATAAATGTCATGAA  
 2151960 TCTTCAAAATTTATGTTCTTCTGTATGGCTCAGTAGTCCCTCCGTCGGGTCAGACCTG  
 2152020 ACTTCCACAACATCCTGTATACCTGAAATTAGCATTGATTTCTAATCTCTGGTCCAC  
 2152080 GTCATTACAGGCATAGGATCTTCGGTGGATTAAGAAGTGCCTCCCTCCCTATTTCTCAGC  
 2152140 CATGTGACTCCAGAACTCCTATGAATTAAGATCTTGTGTTGGCCATAGGAGATCTCT  
 2152200 ATTCACCACTCTTGTCTCTCCCAAAATGGCGAGGCTCCAGTCTCCCACTGACAGCT  
 2152260 TTGTACTAAAATCAACCTTACTCTATAGAACATGCATGATTTGACGAGGACTATGAT  
 2152320 CTGTGTTGATGATTTAGTTAGGATAACATTAGCTGTCTGTAACAAACAGACCCCAAGCT  
 2152380 GCAGCATGACTCAGATGCAATATAAGTGTCTTTTCACTTATATCAAGCAAGAATGACCA  
 2152440 GATTTTCTGATTTTCTTCCGCTGTGCTAATGTAGGGAGAGTGTGTTGGAGGTCAAGTCA  
 2152500 ACAGTTCAACAGCAACCATCTATGTTGGGAGCAAGGATGCTGGAAATAGAATCCAGCATA  
 2152560 CTTGTAGCTTGTCCATAATTACAGACACCTTTGCATTTACTGAACTGAATCCTATGACTT  
 2152620 GAAGACCAAGACTGTAGCATGCCTGAAGGGACAGCCTCAGAACTGTGGATGCCTGTCCC  
 2152680 TCTCCTGGTTTGGGTTTGTGCTGACCAGGCAAAACCCACTGAACTCAGGATCACTGCAT  
 2152740 AAAGTGACGTATTAAAGCCTAGTGCCAGGATCTTTAGTGTGTTGCAAGGAGGTCATGAT  
 2152800 TTTCTGCTAACTCAGCTAAATAGGTAGTCTCTGTCTTCAAGTTACAAACACATCCTTC  
 2152860 ATTCAGGAATTTGAATGTTCAAGTCTTGATATTTTATCAACCTTCAATCTGTGGTCAA  
 2152920 GGGCAGACCTCTCTCCAGTAAAGACTGGAAGTTAGCAGCCTGCCTCCAGGGATGGGGT  
 2152980 GTGGTTGCCTTCTGCTCTCTCTTCTTCTGGGAAGCAGCAGAACTATCCATGGGAGGA  
 2153040 CTAGAGCAGTCTTCTTGAGAGAAGAGATTTACTCCTTCCAAGTGTATTGGTTAGTGAT  
 2153100 TGCTACATAACAAACTACCCAAAACCTCTCAGTAGCTTAAAACAACGTGAAGTGAATTGA  
 2153160 TGCTCATGTGAGCATGGGTTGGTTGATCCAGGCTGGGCTCAGCTGGGCACCTCTGTATAT  
 2153220 GCTGTGGGTTCTCTGAGCTCAACTCCTCTCTGCGAGTTGCATTAAGGCTGTCTGTGTC  
 2153280 GTGTCTTCTGGAGCCAGGCTGAATGGGCTTTGGAGATGGCAGAGCACAAGCGAGTAAA  
 2153340 CAGAGACACATGAAGCCATTTAAGGCCTCAGCCAGCACTGACGTGCCGTCACTTCTCAC  
 2153400 ATTCCACTGGCCATGCAAGTCACTGGCTGAGCAACTCAAGGACTTGGGAAGTAAACCAT  
 2153460 CGTCTTAGTGGGAGGAACTACAACATCCCATGGCAAAGCATGGATCAGGAAGCATGA  
 2153520 AATGGGGCCAGTACTCAGTTTACCACTGAGGCTTGGCAGATGGCTAGAAAGTGGGCT  
 2153580 TTTCTTGGAGATTGGGGAGAGGGTGTGTTATGGATTTACAGCAATCCCAGGCCTG  
 2153640 GGAACCTCTGTAAGTCCCTTCCCAGGGCCTTACATCTCTCTTACATGGTCCCCTCT  
 2153700 AACTCTGCTCATCTAGATTTCTGTACCACACCCAGCTTCTTGTGAGCATCTCTTGTCT  
 2153760 GCCAGAGGCCCTGTGAGACAAGCCCCAAGATGACCCCATGCCAGAAATTTGACCCATG

도면2h

2153820 TGATTACATCTGGACCAGAGGAAATGCCTCCCAAATGAAGCCCATCCTGCCCCATCAG  
2153880 GCAACTACAGGCCACTTCAGCTTTCTGGGTGTAAAGCAGACCTCAGAATCTCTGTGTCT  
2153940 CCCAGCTAGATGAAAGCTTTCCAAGGGTCTTTGGGAAGCCAGCTGGATTGAGGCAAG  
2154000 AATATCACACCCTCATCTCCCAAAGGGAAGCAACATCACCCTGACAAAGTTCTC  
2154060 FCCAGGGCAATCTCTTTGCCAAACATGTCTCCTCTCCACACTCCAAACCCCTTTATGTAT  
2154120 CTTAACATGACTGAGGAGCCCTTTATAAATCTGCAATTTGGGAGTTTGTGATATTT  
2154180 GTTTGGTTCCTGGAGTTACCTACACAAGAGCCTCAGGCAATGAGATTTATTTATCAAA  
2154240 CATTACCAGATGTCCAGGTGCTGTCCCAAACACTTTGTAATGTTAACCCATTTAATCCTC  
2154300 ATATAAATCCTATGAGGTAGGTGCTATTAACATCTTTATTTAGAGATAGGGAAGCTGAG  
2154360 GCACAGAGAGGTTAAGTAATTAGCCCAAAGTCACACAGAAGGCAACTGTCTTCCACCA  
2154420 GGCAGAAGAGTCCCTTTTATGATAATAAGGTGGAAGAAGGCAAGGGGAGATAAGCAA  
2154480 GAAGATGATAATGATGATGGTGGCCCTTCCCTGAGTTACTGTGCTTAGCACTTAGTGTGG  
2154540 TGGCTTTGCCCTTGCCTGTCTTTGAGACAGGTATGTAGACTCTCCCATTTCCGAGAG  
2154600 AATAAACTGAGGCTGATATAGTTTGAATGATGTCTCCACCAAACTCATTGTGAAATG  
2154660 TAATCCCAAGTGTGGAGGTGGGGCTGTGGGAGGTGATTTGATCATGGGGTGGATTT  
2154720 CTATGGGTAGTTTAAACCACTGCCTTGGGTGCTATCCTTATGATAGTGATACATTTCT  
2154780 ATGAGACCTGGTGTTTAAAACCTGGGTGGCCTCTCCCACTTTCTCTCACTCCTGTCT  
2154840 TTGCCACATGATATGCTGTCTCCCTTTGCCCTTGTCTATGATTTGTAAGCCCTCTGAG  
2154900 GCCTCCCAAGAGCCAAAGCAGACGTCAGTACCCTGCTTCTGTAAAGCCTGCAGAACCA  
2154960 GAGCCGATTAACCTCTTTTCTTTATAAATTAACCAGTCTCAGATATCCCTTAAATAGCA  
2155020 ATGCAGAAATGGCTGATACAGAGGCTTTGAGAGGTCAAGTGAACCTGCCAAAGGGCAC  
2155080 CACTGATAAAGGACAGATGTGGATTGAAACCCACTTTGTTAGGCCCAAGTGCAGGGCC  
2155140 CCTTCTCCTCATGCTCACTGCCACTGCGGAGCTGGGCACTGGGACACA CCGTAGGAG  
2155200 CTGTGCTTGTGGGGCTTCTCCTGACCCCAAGTCCCTGCTGCCAAGGCCGGTCTC  
2155260 TGAGTCTCCTACAGCCCTCCTCAGCCTCACTGGCCTCAGTCACTTGGTTCAGGGGAG  
2155320 GACTAAGGTCCCTTTGGTCCCTGCCAATCTGACCCCAAGCCAGTGTGACCTCAGAG  
2155380 CCTGACTTGGCCTTTTACAGAATAAATCTGAA CAAATCAGGTTTCATTTAATAGCAA  
2155440 CAGGCTGCTGATGCAGACCTTATCAACTCCATCAAACCTGTGTCTTTCAAATGTTACGCT  
2155500 CCCCTGGGGTGTCCACACCCCTGACGTCACACATTCAGTGAAGCCCATATTCATT  
2155560 CGGGAGCTCTTTCTTTCTCTTTCTTAACAACACTCTCCTTGGTGTAGTGGGTGCTG  
2155620 ATCAACATTTCTGGATACTGGTTTTCAGGAAATATGATGGTGGCTTCCATCCAG  
2155680 ATTTTCTCATGATGGTTCCATATTTACACCATCTTGGCCAGTCTCTGGGTCAACAT  
2155740 TCCACACATACCCACACCCATAAAGAGAGGCTTTTTCTGAGTCTTCTGTTTCATGAGTC  
2155800 TGGAAATGTATTTGCTTTGTTTTGGGGCACTCCCTGGGAGCTCATTTCCACTAATAGGCA  
2155860 AACCAACAATGTCAGTCTCCACTTACTGACATTTAACTTTCCAGGCACATGTAGG  
2155920 GACCTTACATTCATTATTTCAATTTATTTCTCACATCAACCTTGTGAGGTGGAGGAGCA  
2155980 TGATGGAAGGAAAGGAGCTAAAAGCAAAGAACTGGAGTCCGACTGCCTGGGTTCAAA  
2156040 TCCTAGCTCTAACAAATTTCCAGCTCTGTAACATCGAGCTAATTTCTAACCTCTCTATGC  
2156100 CATTTCCCTATCTTAAAGGAAAGCTGCAATAGCATCTATCTCATAGGATTTGATCGAA  
2156160 GATTAAATGAGTCAATAATTTATAAAGTGTTCGGAATGATACCTGACATCTGGTAAAGGTT  
2156220 TGATAAATAAATCCATTTAATGATGAGGAAACAGGCTCAGAAAGGGGCTCATTTGC  
2156280 TCAATGGTACAGATAGGTTCCAGACTCAAACCTCAAGACCATCTGACTCTAAAACATCTA  
2156340 AAACCTGTTGCCCGCATTTCTTCTTATGACAGATAATAAACTGAGGCTCAGAGAAGCT  
2156400 AAGTGAATCCCTGGGACTGCACAGCAAATCAAGACAAATAAGACCTAGGCTTCTCTGAC  
2156460 TGCCAGAGTGGAGATGCTTCTATAGGCTTTTCTCACTGATGCTCTCTGGGCAGACAGGCT  
2156520 CCTCAATATGAGAGTGCACACACTCCTTCTTCTTCTTTCAGGTAACCTCACACTGGTT  
2156580 GGCAGAGGAACTATACCAATAATTAAGTAACATGCGGTGAATTTGCAACAGCAAGAGG  
2156640 AGCTCATTTATCCTATAGTTTCCAGGTGTCTTAGGGAGGCAAGAAATCACAGCAAGGAAA  
2156700 CCTCAATAATAAACAGACGCTCTATAAATAAATGCAACCCCACTCTCTCTACTT  
2156760 AAAATTAGCATCTATTTCCAGCTCTGCTTCAATGCCCATATGAAATACATGTGAACCTC  
2156820 CTCCTCTCTTCTCCT  
2156880 TTAAGAAAAAATAAAGGTAGATTTACACAAATAGTGGGATCTCAGTCTTGGATTAGCT  
2156940 GTGTATGATGAAAAGGATGCTGTGGTAAATAATATCAATAAAAACATGACATGGCCGG  
2157000 GCACATGGCTCACGCTGTAAATCCCAGAACTTTGGGAGCCGAGGCAGGCAGATCACTT  
2157060 GAGGCCAGGAGTTGAGACCAGCCTGGCCAAATGGTGAACCTGCATCTCTATAAAAAAT  
2157120 ACAAAAATTAGCCGGCATCAATGGCCAGCCCTGTAATCCAGCTAATCAGGAGTCTGA  
2157180 GGCAGGCAATCACTGAACTGGGGCTGGTGGTTGCACTGAGCCGAGCTCACACCACT

도면2i

2157240 GCACTCCAGCCTGGGCGACAGAGTACACTCAAAAAACAAAACAAACAAGCA  
2157300 AAAAAACCCACAGTAACAAAAAGTAATAAACTGCTGCTATTTACTCAGTCTTATC  
2157360 TGATGCCAGCCACTTTGCTAAGCCTATGAATGCAATTTTCCCGTTGCTACAGATGAGA  
2157420 GAATTTAGGTTTCAGACAGTTGAAATCAATGCTCCAAAGTCACACACTGCTGATGGC  
2157480 AGAGCTGGGATGCAAAACCTAACTGCCAGCCCTCAAAGCCTGTGCTCTTAATCTCCACC  
2157540 CTGCTGTGCTTCTTGTCCATTAATTAAGTCCACAGGCACACATTTCCAGCCCTCCTT  
2157600 TGCTGTACAATCCAGGCAAGTCGCTCAGCTTCTGTAGCCTCAGTTTCAATACTGTCA  
2157660 AATGGAGGTAACACAAATAATCTAGTTGTGACCAAGAATCATCATAGAAATCTGCCAT  
2157720 TCCAGCTTATGTGCAATTCCTCAAGCACTGTGACTCAAAGTGGCATCAGCTCTGGAA  
2157780 GAACACACTGTCTTACTGTTTCTCTCTTTGTCAACTGATCCCCCTTGAACTCCTC  
2157840 CTACCTCTGCTCTAATGCCCATCTACTGCCACCTGATTAATAAAAACTTTTGTGAAA  
2157900 ATCATAAGTGTATGAGTAAGGTTCTTGGTGTGATGATGAAAGAACAAAAAGAAATTTG  
2157960 GAAATGAGAATCACTGCAGCTATCATGAAGTCTGCTCTAC

도면3a

서열 20  
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSLQ LGEGVEFDVL PAFDALGQLT GXYPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNVW

서열 21  
 1 SVSRRDKSKQ VWEAVLLPLS LLMMDLRNTF AKSLDKFIED YLLPDTCFRM QIXHAIDIIC  
 61 GFLKERCFRG SSSYPVCVSKV VKGGSSGKGT TLRGRSDADL VVFLSPLTTF QDQLNRRGEF  
 121 IQEIRRQLEA CQRERAXSVK FEVQAPRWGN PRALSFVLSS LQLGEGVEFD VLPAPDALGQ  
 181 LTGXYPNPQ IYVKLIEECT DLQKEGEFST CFTTELQRDFL KQRPTKLKSL IRLVKHWYQNVW

서열 22  
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSLQ LGEGVEFDVL PAFDALGQLT GXYPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNVW PSHQAWVLS RLGAEEG

서열 23  
 1 SVSRRDKSKQ VWEAVLLPLS LLMMDLRNTF AKSLDKFIED YLLPDTCFRM QIXHAIDIIC  
 61 GFLKERCFRG SSSYPVCVSKV VKGGSSGKGT TLRGRSDADL VVFLSPLTTF QDQLNRRGEF  
 121 IQEIRRQLEA CQRERAXSVK FEVQAPRWGN PRALSFVLSS LQLGEGVEFD VLPAPDALGQ  
 181 LTGXYPNPQ IYVKLIEECT DLQKEGEFST CFTTELQRDFL KQRPTKLKSL IRLVKHWYQNVW  
 241 VWP SHQAWV LSR LGAEEG

서열 24  
 1 SVSRRDKSKQ VWEAVLLPLS LLSMMDLRNT PAKSLDKFIE DYLLPDTCFR MQIXHAIDI I  
 61 CGFLKERCFR GSSYPVCVSKV VVKGGSSGKGT TLRGRSDADL LVVFLSPLTTF QDQLNRRGE  
 121 FIQEIRRQLE ACQRERAXSV KFEVQAPRWGN NPRALSFVLS SLQLGEGVEF DVLPAFDALG  
 181 QLTGXYPNPQ QIYVKLIEEC TDQLQKEGEFS TCFTELQRDF LKQRPTKLKS LIRLVKHWYQ  
 241 NVWPSHPACW YLYIFI

서열 25  
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSLQ LGEGVEFDVL PAFDALGQLT GXYPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNVW PSHPACWYLY IFI

서열 26  
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSLQ LGEGVEFDVL PAFDALGQLT GXYPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNVW KKLKGLPPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYDFKX PIIEKYLRRQ LTKPRPVILD  
 301 PADFTGNLGG GDDPKGWRQLA QEBAEAWLNXY CFKNWDGSPV SSWILLMRQR LREVRSLAQG  
 361 HQLTSGNGI QAQWTLKPV L MSLC

도면3b

서열 27

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GSSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVFL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCX KKLKGLPPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRQ LTKRPRVILD  
 301 PADFTGNLGG GPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLKAT VQTMRPPIPG  
 361 XIRNMVLEH MSTLISLIDP AHSRQHPPHR QKRTGPAPSS ECQCILGERA PVLGSPVPSF  
 421 SGGTLDPEXT KLLSELVYNP GQNPGLLTPG LLCPLSYHR

서열 28

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GSSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVFL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCX KKLKGLPPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRQ LTKRPRVILD  
 301 PADFTGNLGG GPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLIKLR LREAK

서열 29

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GSSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVFL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCX KKLKGLPPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRQ LTKRPRVILD  
 301 PADFTGNLGG GPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLVNLTVGRRNYPPI  
 361 SEHAVNLQQT RXASLSYSFO VA

서열 30

1 AESNSXDEET DDPXYQKYG YIGTHEYPHF SHRPSTLQAA STPQAEEDWT CTIL

서열 31

1 GCTGAAAGCA ACAGTRCAGA CGATGAGACC GACGATCCCA GGASGTATCA GAAATATGGT  
 61 TACATTGGAA CACATGAGTA CCCTCATTTC TCTCATAGAC CCAGCACACT CCAGGCAGCA  
 121 TCCACCCAC AGGCAGAAGA GGACTGGACC TGCACCATCC TCTGAATGCC AGTGCATCTT  
 181 GGGGAAAGG GCTCCAGTGT TATCTGGACC AGTTCCTTCA TTTTCAGGTG GGACTCTTGA  
 241 TCCAGAGARG ACAAGCTCC TCAGTGAGCT GGTGATAAT CCAGGACAGA ACCCAGGTCT  
 301 CCTGACTCCT GGCCTCTAT GCCCTCTATC CTATCATAGA TAACATTCTC CACAGCCTCA  
 361 CTTCATTTCCA CCTATTCTCT GAAATATTC CCTGAGAGAG AACAGAGAGA TTTAGATAAG  
 421 AGAATGAAAT TCCAGCCTTG ACTTCTCTCT GTGCACCTGA TGGGAGGGTA ATGTCTAATG  
 481 TATTATCAAT AACAATAAAA ATAAAGCAA TACCATTTA

서열 32

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK  
 61 GSSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVFL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCX KKLKGLPPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRQ LTKRPR

도면3c

서열 33

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVDL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLQK RPTKLSLIR LVKHWYQNC KKLGLKLPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYDFKX PIIKYLRRQ LTKPRPVILD  
 301 PADPTGNLGG GDPKGRQLA QEAEAWLNXY CFKNWDGSPV SSWILLVRPP ASSLPFIPAP  
 361 LHEA

서열 34

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVDL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLQK RPTKLSLIR LVKHWYQNC KKLGLKLPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYDFKX PIIKYLRRQ LTKPRPVILD  
 301 PADPTGNLGG GDPKGRQLA QEAEAWLNXY CFKNWDGSPV SSWILLAESN SXDDETDDPR  
 361 XYQYGYIGT HEYPHFSHRP STLQAASTPQ AEDDWTCTIL

서열 35

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGS YPVCVSKVVK  
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE  
 121 VQAPRWXNPR ALSFVLSSLQ LGEVGEFVDL PAFDALGQLT GXYKPNPQIY VKLIEECTDL  
 181 QKEGEFSTCF TELQRDFLQK RPTKLSLIR LVKHWYQNC KKLGLKLPQY ALELLTVYAW  
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYDFKX PIIKYLRRQ LTKPRPVILD  
 301 PADPTGNLGG GDPKGRQLA QEAEAWLNXY CFKNWDGSPV SSWILLTQHT PGSIHPGRR  
 361 GLDLHHPLNA SASWKGQLQC YLDQFLHFQV GLLIQRXQSS SVSWCIIQDR TQVS

서열 36

GAGGCAGTTCCTGTTGCCACTCTCTCTCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT  
 CATTGAAGACTATCTCTTCCAGACACGTGTTCCGCATGCAAATCAACCATGCCATTGACATCATCTGTGGGTCC  
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCCTGTGTGTGTGTCAGAGGTGTAAGGGTGGCTCCTCAGGCAAG  
 GGCACCA CCTCAGAGCCGATCTGACGCTGACCTGGTTGTCTTCTCAGTCTCTCACCACCTTTTCAGGATCAGTT  
 AAATCGCCGGGAGAGTTCCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTCCTGTGA  
 AGTTTGAGGTCAGGCTCCACGCTGGGGCAA CCCCCGTGCGCTCAGCTTCTGACTGAGTTCCGCTCCAGCTCGGGGAG  
 GGGGTGGAGTTCGATGTGCTGCTGCTGCTTTGATGCCCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAT  
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCTCCACCTGCTTACAGAACTACAGA  
 GAGACTTCTGAAGCAGGCCCCACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATTGTAAG  
 AAGAAGCTTGGGAGCTGCCACCTCAGTATGCCCTGGAGCTCTGACGCTCTATGCTTGGGAGCGAGGAGCATGAA  
 AACACATTTCAAACAGCCAGGATTTCCGACGCTTTGGAAATTAGTATATAAATACCAGCAACTCTGCATCTACT  
 GGACAAAGTATTATGACTTTAAAAACCCCATATTGAAAAGTACCTGAAGGGCAGCTCACGAAACCCAGGCTGTG  
 ATCTGGACCCCGGACCTTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCTGCTGCCAAGAGGC  
 TGAGGCTGGCTGAATTACCATGCTTTAAGAAATTGGGATGGTCCCAAGTGGCTCCTGGATTCTGTGTGGAGAC  
 CTCCTGCTTCTCTCCCTGCATTCATCCCTGCCCTCTCCATGAAGCTTGAGACATATAGCTGGAGACCATTTCTTCC  
 AAAGAACTTACCTTCTGCCAAAGCCATTTATATTCATATAGTGACAGGCTGTGCTCCATATTTTACAGTCAATTTG  
 GTCACAAATCGAGGTTTCTGGAATTTTCAATCCCTTGTCCAGAAATTCATTCCTCCAAAGAGTAATAAATAAATCT  
 CTAACACCATTTATGACTGTCTGCTTCCGGCTCAGGTTCTGTCTTAAGCCCTTAAATGCACTCTCTCATTAATA  
 A

서열 37

GAGGCAGTTCCTGTTGCCACTCTCTCTCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT  
 CATTGAAGACTATCTCTTCCAGACACGTGTTCCGCATGCAAATCAACCATGCCATTGACATCATCTGTGGGTCC  
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCCTGTGTGTGTGTCAGAGGTGTAAGGGTGGCTCCTCAGGCAAG  
 GGCACCA CCTCAGAGCCGATCTGACGCTGACCTGGTTGTCTTCTCAGTCTCTCACCACCTTTTCAGGATCAGTT  
 AAATCGCCGGGAGAGTTCCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTCCTGTGA  
 AGTTTGAGGTTCCAGGCTCCAGCTGGGGCAA CCCCCGTGCGCTCAGCTTCTGACTGAGTTCCGCTCCAGCTCGGGGAG  
 GGGGTGGAGTTCGATGTGCTGCTGCTTTGATGCCCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAT  
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCTCCACCTGCTTACAGAACTACAGA

도면3d

GAGACTTCTGAGCAGCGCCACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATTGTAAG  
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA  
 AACACATTTCAACACAGCCAGGGATTTCSGACGGTCTTGGAAATAGTCAATAAATACAGCAACTCTGCATCTACT  
 GGACAAAGTATATGACTTTAAAAACCCCATTTATGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCTGTG  
 ATCCTGGACCCGGCGGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTEGCAACAGAGGC  
 TGAGGCTTGGCTGAATACCCATGCTTTAAGAAATGGGATGGGTCCCAGTGAAGCTCCCTGGATTCTGCTGGCTGAAA  
 CAAACAGTRCAGACGATGAGACCAGCATCCAGGASGTATCAGAAATATGGTTACATTTGGAAACATGAGTACCT  
 CATTTCTCTCATAGACCCAGCACTCCAGGCAGCATCCACCCACAGGCAGAAAGGACTGGACCTGCACCATCTCT  
 CTGAATGCCATGTCATCTTGGGGAAAGGGCTCCAGTGTATCTGGACAGTTCCTTCAATTTTCAAGTGGGACTCTT  
 GATCCAGAGARGACAAAGCTCCTCAGTGAAGCTGGTGTATAAATCCAGGACAGAACCCAGGCTCCTGACTCCTGGGCT  
 TCTATGCCCTCTATCTCTATCATAGATAACATCTCCACAGCCTCACTTCAATCCACCTATTCTGTGAAAATATCCC  
 TGAGAGAAACAGAGATTTAGATAAGGAAATGAAATCCAGCCTTGAATTTCTTCTGTGCCTGATGGGAGGGT  
 AATGTCTAATGTATTAATAAATAAATAAAGCAATACCATTTATTTGGGTGTTTAACTTCAAGGCACA  
 GAGCCAAAGATACAGATGCATATCTAGGGGTATTTGTGTGTATATACATTTGATTCACAAAGAAATATTTATGAG  
 CACTT

서열 38

GAGGCAGTCTGTGGCACTCTCTCTCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT  
 CATTGAAACTATCTCTTGGCCAGACAGTGTTCCTCGCATGCAAAATCAACCATGCCATTGACATCATCTGTGGGTCC  
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCCTGTGTGTGTCAGGTGGTAAAGGGTGGCTCCTCAGGCAAG  
 GGACACCCCTCAGAGGCCGATCTGACGCTGACCTGGTGTCTTCTCAGTCTCTCACCACCTTTCCAGGATCAGTT  
 AAATCCCGGGGAGGTTTATCCAGGAAATAGGAGACAGCTGGAAAGCTTCAAAGAGAGAGACATTTTCCGTGA  
 AGTTTGGAGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAG  
 GGGTGGAGTTCGATGTGCTGCCCTTGTATGCCCTGGGTGAGTGAAGTGGCGCTATAAACTAACCCCAAT  
 CTATGTCAAGCTCATCGAGGAGTCCACCGACCTGCAGAAAGAGGGCGAGTTCCACCTGCTTACAGAACTACAGA  
 GAGACTTCTGAAGCAGCGCCACCAAGCTCAAGAGCTCATCCGCTAGTCAAGCAGTGGTACCAAAATTTGTAAG  
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGAGCGAGGAGCATGAA  
 AACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCAATAAATACAGCAACTCTGCATCTACT  
 GGACAAGTATTATGACTTTAAAAACCCCATTTATGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCTGTG  
 ATCCTGGACCCGGGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTEGCAACAGAGGC  
 TGAGGCTTGGCTGAATACCCAATGCTTTAAGAAATGGGATGGGTCCCAGTGAAGCTCCCTGGATTCTGCTGACCCAGC  
 ACCTCCAGGCAGCATCCACCCACAGGCAGAAAGGACTGGACCTGCACCACTCTGAAATGCCAGTGCATTTGG  
 GGAAGGGCTCCAGTGTATCTGGACAGTTCCTTCAATTTCAAGTGGGACTCTTGAATCCAGARGACAAAGCTC  
 CTCAGTGAAGTGTATAATCCAGGACAGAACCCAGGCTCCTGACTCCTGGCTTCTATGCCCTCTATCTATCA  
 TAGATAACATTTCTCCAAGCTCATTCTATCCACTATTCTCTGAAAATATCCCTGAGAGAGAAACAGAGATTT  
 AGATAAGAGAAATGAAAATCCAGCTTGAATTTCTTCTGTGACCTGATGGGAGGGTAAATGTCTAATGTATTAATA  
 AACAAATAAATAAAGCAATAACATTTATTTGGGTGTTTATTAATTCAGGCACAGGCCAAAGAAATACAGATGCA  
 TATCTAGGGGTATTTGTGTGTATATACATTTGATTCACAAAGAAATATTTATGAGCACTT

서열 39

GAGGCAGTCTGTGGCACTCTCTCTCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT  
 CATTGAAACTATCTCTTGGCCAGACAGTGTTCCTCGCATGCAAAATCAACCATGCCATTGACATCATCTGTGGGTCC  
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCCTGTGTGTGTCAGGTGGTAAAGGGTGGCTCCTCAGGCAAG  
 GGACACCCCTCAGAGGCCGATCTGACGCTGACCTGGTGTCTTCTCAGTCTCTCACCACCTTTTCCAGGATCAGTT  
 AAATCCCGGGGAGGTTTATCCAGGAAATAGGAGACAGCTGGAAGCTTCAAAGAGAGAGACATTTTCCGTGA  
 AGTTTGGAGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAG  
 GGGTGGAGTTCGATGTGCTGCCCTTGTATGCCCTGGGTGAGTGAAGTGGCGCTATAAACTAACCCCAAT  
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCCACCTGCTTACAGAACTACAGA  
 GAGACTTCTGAAGCAGCGCCACCAAGCTCAAGAGCTCATCCGCTAGTCAAGCACTGGTACCAAAATTTGTAAG  
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA  
 AACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCAATAAATACAGCAACTCTGCATCTACT  
 GGACAAGTATTATGACTTTAAAAACCCCATTTATGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCTGTG  
 ATCCTGGACCCGGGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTEGCAACAGAGGC  
 TGAGGCTTGGCTGAATACCCAATGCTTTAAGAAATGGGATGGGTCCCAGTGAAGCTCCCTGGATTCTGCTGACCCAGC  
 CAACAGTRCAGACGATGAGACCGAGCATCCAGGASGTATCAGAAATATGGTTACATTTGGAAACATGAGTACCCCT  
 ATTTCTCTCATAGACCCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAAAGGACTGGACCTGCACCATCTCT  
 TGAATGCCAGTGCATTTGGGGAAAGGGCTCCAGTGTATCTGGACCAAGTTCCTTCAATTTTCAAGTGGGACTCTTG

도면3e

ATCCAGARGACAAAGCTCCTCAGTGAGCTGGTGTATAATCCAGGACAGAAACCCAGGTCTCCTGACTCCTGGCCTT
CTATGCCCTCTATCCTATCATAGATAACATTCTCCACAGCCTCACTTCATTCCACCTATTCTCTGAAAATATTCCT
GAGAGAGAACAGAGAGATTAGATAAGAGAAATGAAATCCAGCCTTGACTTCTCTGTGCACCTGATGGGAGGTA
ATGTTAATGTATATCAATAACAATAAAATAAAGCAAATACCAATTATTTGGGTGTATTAACCTCAAGGCACAG
AGCCAAAGTACAGATGCATATCTAGGGGTATGTGTGTATATACATTGATTCAACAAGAAATATTTATTGAGC
ACTT

서열 40

GAGGCAGTTCGTGGCACTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
CATTTGAAGACTATCTCTTGCCAGACAGTGTTCGCGATGCAAAATCAACCATGCCATTGACATCATCTGTGGGTCC
TGAAGGAAAGTGTCTCCGAGGTAGCTCCTACCTGTGTGTGTGTCGAAGTGGTAAAGGGTGGCTCCTCAGGCAAG
GGCACCACCTCAAGAGCCGATCTGACGCTGACCTGGTGTCTCTCCTCAGTCTCTCACCACTTTTCAGGATCAGTT
AAATCGCCGGGAGAGTTCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
AGTTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAG
GGGTGGAGTTCGATGTGCTGCCTTGTATGCGCTGGTFCAGTTGACTGGCRGCTATAAACCTAACCCCAAAAT
CTATGTCAAGCTCATCGAGGAGTGCACCCGCTGCAGAAAGAGGGCGAGTTCACCTGCTTCAAGAACTACAGA
GAGACTTCCTGAAGCAGCGCCCAACAGCTCAAGAGCTCATCCGCTAGTCAAGCACTGGTACCAAAATTTGTAAG
AAGAGCTTGGGAGCTGCCACCTCAGTATCAGCCTGGAGCTCCTGAGCGTATGCTTGGGAGCGAGGAGCATGAA
AACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCAATAAACCTACCAAGCACTCTGCATCTACT
GGACAAAGTATTAAGCTTAAACCCCAATTTTGAAGATCCTGAGAGGCAAGCTCACGAAACCCAGGCTGTG
ATCCTGGACCCGGCGGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTGGAGGCAAGTGGCAAGAGGC
TGAGGCTTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGTCCCAAGTGGCTCCTGGATTCTGTGATGAGAC
AAAGGCTCAGAGAGGTGAGGTCACTTGTCAAGGACATCAGTAAACAGTGGTGGAAATGGAAATCAAGCTCAGTGG
ACTTAAAGCAGTGTCTCATGTCACTGTGTAAACAGCTGCTTGTCACTCCACCTCTCATCTGACCAATGGG
AGACTCTGAGCAGCTGAGTGACTTGGGTGTGACACAGCTAAACAGGGGCAAGGACCCAGTCTTGGATCTTTCCAC
CTCCAGCAGGAATCTGTGATCCAGGGGATTTGATGATGTTGCAGATGGCTAGGAAGCAGACTCCAGGATGGAAAT
TTAGTATGACAGATGTTCTGGGGGAGAGCCACTGGAAACAGCACTCAGGGAAGGGGGGAAAGAAAGTATGGAAGGA
AGCATGAAAGAGAAATAGGGAGAAAGTGAACAGGGATGCAGAGCGAATGCCAGTTTCAAGCCAACTCAAGGACAGCCCT
GGAGCTGGAAATGGCCTTTAGAGCTGCCCAATGGTGCAGAGGTGGCCAGGCTTCTATACCCTACCTGGATCACTCA
TGTGCTTGGGCACTTGGGAAAGGGCAATGGCTTTGAGCAAAAGGCTCTCTGCAGCTGAGGCAACCCCTAAAGGGG
TGACGGCTGAGTCTGTCTGTGACCACTGCCAGCAGCTGGGGCTTGTAGTCTCTCAAAGGGGATCCAGA
TGGCATGTCAAGTGTCTACGTAATGCTCACTGAATCCAGCTGCAATGCAGGAAAGACTCCCTGATGTGATCATGT
GTCTCACCTTTCARGCTGAAAGCAACAGTRCAGACGATGAGACCGACGATCCAGGASGTATCAGAAATATGGTTA
CATTTGAAACATGAGTACCTCATTTCTCTCATAGACCCAGCACACTCCAGGCAGCATCCACCCCAAGGCAGGAG
AGGACTGGACCTGCACCATCCTCTGAATGCCAGTGCATCTTGGGGAAAGGGCTCCAGTGTATCTGGAACCAATCC
TTCAATTTCAAGGTGGGACTCTTGAATCCAGAGARGACAAGCTCCTCAGTGAAGCTGGTGTATAATCCAGGACAGAAC
CAGGCTCCTGACTCCTGGCCTTCTATGCCCTCTATCCTATCATAGATAACATTTCCACAGCCTCATCTATTCCA
CCTATTCTCTGAAAATATTCCTTGAGAGAGAACAGAGAGATTAGATAAGAGAAATGAAATCCAGCCTTGACTTCT
TCTGTGCACCTGATGGGAGGTAAATGTCTAATGTATTAATAACAATAAAATAAAGCAAATACCAATTTATTGGG
TGTTTTAACTTCAAGGCACAGGCCAAGAGTACAGATGCATATCTAGGGGTATGTGTGTATATACATTGAT
TCAACAAGAAATATTTATTGAGCACTT

서열 41

GAGGCAGTTCGTGGCACTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
CATTTGAAGACTATCTCTTGCCAGACAGTGTTCGCGATGCAAAATCAACCATGCCATTGACATCATCTGTGGGTCC
TGAAGGAAAGTGTCTCCGAGGTAGCTCCTACCTGTGTGTGTGTCGAAGTGGTAAAGGGTGGCTCCTCAGGCAAG
GGCACCACCTCAAGAGCCGATCTGACGCTGACCTGGTGTCTCTCCTCAGTCTCTCACCACTTTTCAGGATCAGTT
AAATCGCCGGGAGAGTTCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
AGTTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAG
GGGTGGAGTTCGATGTGCTGCCTTGTATGCCCTGGGTGAGTGAAGTGGCRGCTATAAACCTAACCCCAAAAT
CTATGTCAAGCTCATCGAGGAGTGCACCCGCTGCAGAAAGAGGGCGAGTTCACCTGCTTCAAGAACTACAGA
GAGACTTCTGAAAGCAGCGCCCAACAGCTCAAGAGCTCATCCGCTAGTCAAGCACTGGTACCAAAATTTGTAAG
AAGAGCTTGGGAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGCTATGCTTGGGAGCGAGGAGCATGAA
AACACATTTCAACACAGCCAGGGATTTGGGAGCTTGGAAATAGTCAATAAACCTACCAAGCACTCTGCATCTACT
GGACAAAGTATTAAGCTTAAACCCCAATTTGAAGATACCTGAGAAAGGCAAGCTCACGAAACCCAGGCTGTG
ATCCTGGACCCGGCGGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTGGAGGCAAGTGGCAAGAGGC
TGAGGCTTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGTCCCAAGTGGCTCCTGGATTCTGTGGTAAACC





도면3h

241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIKEYLRRQ LTKRPRVILD  
301 PADPTGNLGG GDFKGRWQLA QEAEAWLNXX CFKNWDGSPV SSWILLAESN SXD

서열 51

1 MMDLRNTPAK SLDKFIEDYL LPDKCFRQKI NHAIDIICGF LKERCFQSS YPVHVSXVVK  
61 GSSSGKGTFL RGRSDADLVV FLSFLTTFQD QLNRRGEFIQ EIRRQLEAQQ REERAFSVKF  
121 EVQAPRWDPN RALSFVLSL QLGEVVEFDV LPAFDALGQL TDGYKPDPI YVKLIBECTY  
181 LQKEGEFSTC FTELQRDFLK QRPKTKSLI RLVKHWYQNC KKKLGLLPPQ YALELLTVYA  
241 WEQGSMEFDF NTAQEFRTVL ELVINYQQLC IYWTKYYDFE NPIIEKYLRR QLTKPRPVIL  
301 DPADPTGNLGG GDFKGRWQLA QEAEAWLNXX PCFKN

서열 52

1 PVILDPADPT GNLGGDPKG WRQLAQEAEA WLNYPCKNW DGSPVSSWIL LAESDSGR

서열 55

GAGGCAGTTCTGTTGCCACTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT  
CATTGAAGACTATCTTTGCCAGACAAGTGTTCGCAAGCAAAATCAACCATGCCATTGACATCATCTGGGTTC  
TGAAAGAAAGGTCTTCCAGGTAGCTCTCACTGTCATGTCCTCAAGTGGTAAAGGGTGGCTCTCAGGCAAG  
GGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTGTCTCTCCTCAGTCTCTCACCACCTTTTCAGGATCAGTT  
AAATCGCCGGGAGAGTTTCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGGAGAGAGCATTTCCG  
TGAAATTTGAGGTCAGGCTCCAGCTGGGCAAAACCCCGTGGCTCAGCTTCTGACTGAGTTCCGCTCCAGCTCGG  
GAGGGGTGGAGTTCGATGCTGCTGCTTGTATGCCCTGGGTGAGTTCGACTGACGCGCTATAAACCTGACCCCA  
AATCTATGTCAGCTCATCGAGGAGTGCACCTGACCTGCAGAAAGAGGGCGAGTCTCCACCTGCTCACAGAACTAC  
AGAGAGACTTCTGAAGCAGCGCCCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATGTT  
AAGAAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCAAGGGAGCAT  
GGAAACAGATTTCAACACAGCCAGGAAATTCGGACGCTCTTGAATTAGTCAATAACTACCAGCAACTCTGCACT  
ACTGGAACAAGTATTGACTTTGAAAACCCCATTTATGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCT  
GTGATCTGGACCCCGCGACCTCAGGAAACTTGGGTGGTGGAGCCCAAGGGTGGAGGCAGCTGGCAAGA  
GGCTGAGGCTGGCTGAATTAACCATGCTTTAAGAATTGAGATGGGTCCCAAGTGGTCTCTGATTTCTGCTGGTGA  
GACCTCTGCTTCTCCTGCAATTCATCCCTGCCCTCTCCATGAAGCTTGAGACATATAGCTGGAGACCATTCTT  
TCCAAAGAACTTACCTCTTGCCAAAGGCCATTTATATTCATATAGTGACAGGCTGTCTCCATATTTTACAGTTATT  
TTGGTCAAACTCGAGGTTTCTGGAATTTTCACATCCCTTGTCCAGAATTCATTCCTCAAGAGTAATAA TAAATAA  
TCTCAACAC

서열 56

GCCTGTGATCCTGGACCCGGCAGACCTACAGGAAACTTGGGTGGTGGAGCCCAAGGGTGGAGGCAGCTGGCAC  
AAGAGGCTGAGGCTGGCTGAATTAACCATGCTTTAAGAAATGGGATGGGTCCCAAGTGGTCTCTGGATTCTGCTG  
GCTGAAAGGCAGCTGGACGATGAGACCCAGCATCCAGGAGGTATCAGAAATATGGTTACATTTGGAAACATGAGT  
ACCTCATTTCTCTCATAGACCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAGGAGGACTGGACCTGCACC  
ATCCTCTGAAATGCTAGTGCATCTTGGGGAAAGGCTCCAGTGTATCTGACCAGTCTCTTATTTTTCAGGTGGGA  
CTCTGATCCAGAGAGGCAAAAGCTCCTCAGTGGTGGTGTATAATCCAGGACAGAAACCCAGGCTCCTGACTCCT  
GGCCTCTATGCCCTCTATCCTATCATAGATAACATTTCTCCACAGCTCACTTCAATTCACCTATTTCTGAAAATA  
TCTCCTGAGAGAGAACAGAGAGATTAGATAAGAGAAATTCAGCCTGACTTTCTTCTGTGCACCTGATGGG  
AGGGTTATGTTAATGATTATCAATAACAGTAAAAATAAGCAATGCC

상기 뉴클레오티드 서열(서열 31, 서열 36-45, 및 서열 55-56)에 대하여:

R은 A 또는 G를 나타냄  
S는 G 또는 C를 나타냄

상기 아미노산 서열(서열 20-30, 32-35, 및 서열 46-52)에 대하여:

X는 하기 표에 따른 아미노산 변이체를 나타냄.

도면3i

서열	아미노산 위치	아미노산
20	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	G 또는 S
21	53	D 또는 N
	137	L 또는 F
	184	G 또는 S
22	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	G 또는 S
23	53	D 또는 N
	137	L 또는 F
	184	G 또는 S
24	64	D 또는 N
	138	L 또는 F
	185	G 또는 S
25	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	G 또는 S
26	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
27	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
	361	G 또는 R
28	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
29	31	D 또는 N
	115	L 또는 F
	127	G 또는 R

도면3j

	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
	372	R 또는 G
30	6	A 또는 T
	15	R 또는 T
32	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
33	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
34	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
	352	A 또는 T
	361	R 또는 T
35	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
	397	G 또는 R
46	31	D 또는 N
	122	N 또는 T
	172	P 또는 S
	214	R 또는 G
47	31	D 또는 N
	122	N 또는 T
48	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
49	31	D 또는 N

도면3k

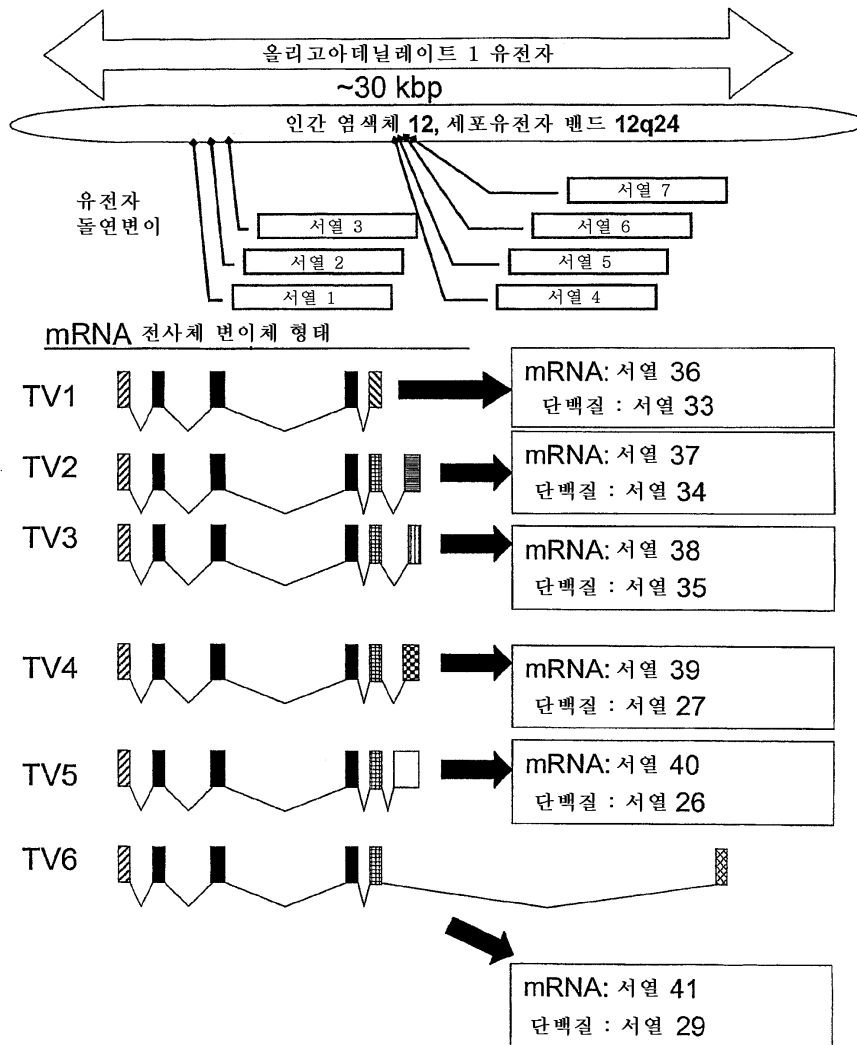
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
50	31	D 또는 N
	115	L 또는 F
	127	G 또는 R
	162	S 또는 G
	280	N 또는 T
	330	P 또는 S
	352	A 또는 T

도면4

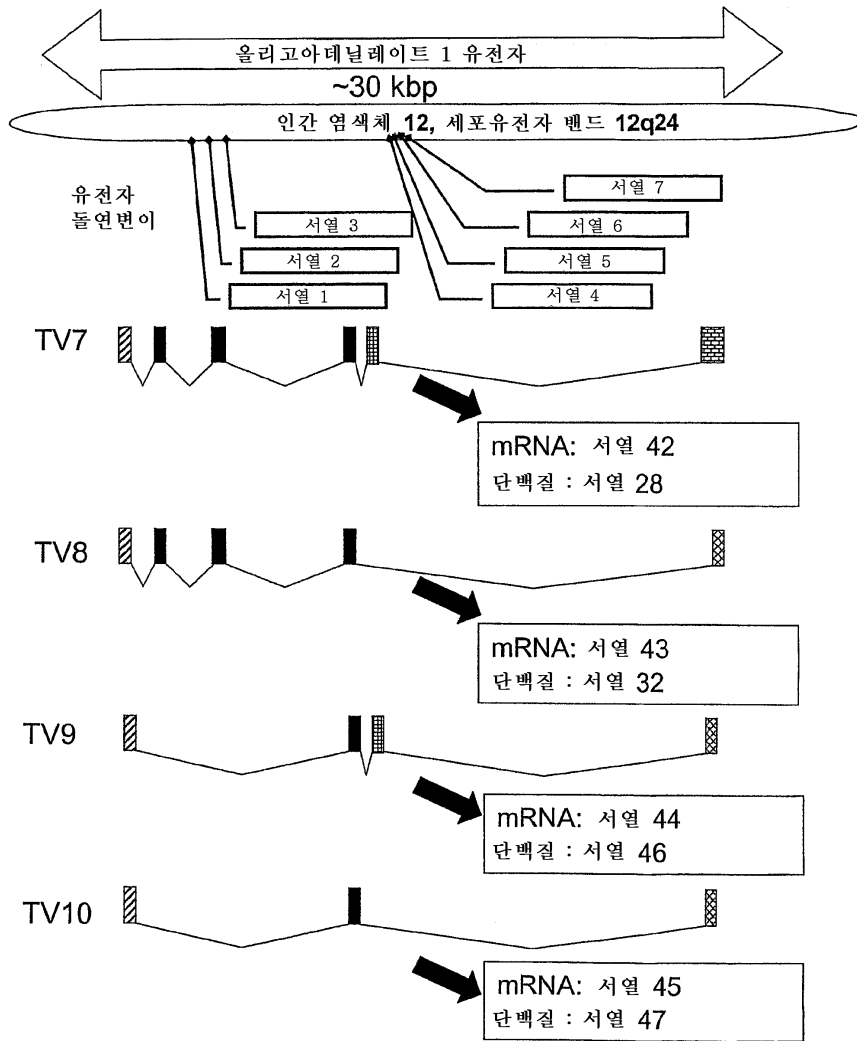
경우 및 대조 대상체로부터 추정된 집단 반수형

추정된 반수형	근사 집단 빈도	서열 2 대립 유전자	서열 3 대립 유전자	서열 4 대립 유전자	서열 5 대립 유전자
HAP1	54%	A	G	A	A
HAP2	32%	G	G	G	G
HAP3	8%	G	A	A	A
HAP4	5%	G	G	G	A
HAP5	<1%	A	G	G	G
HAP6	<1%	G	G	A	A

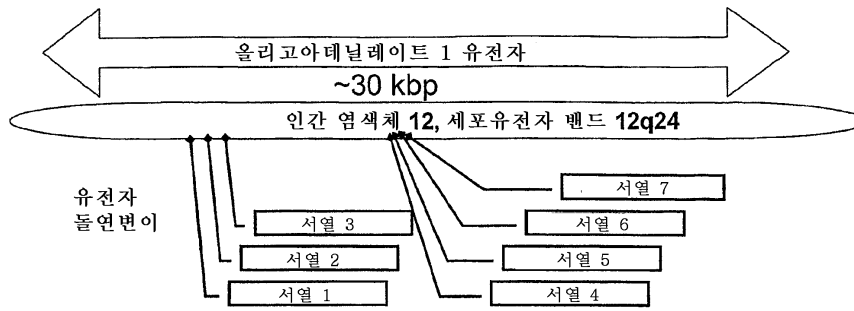
도면5A



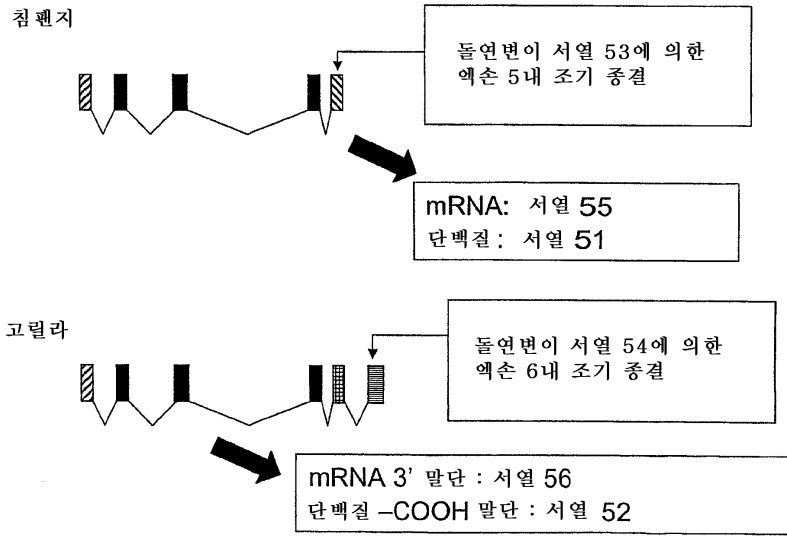
도면5B



도면5C



비인간 OAS1 유전자의 추정 구조



도면6

비인간 영장류 돌연변이

<p>서열 53                  유기체 : <i>Pan troglodytes</i>                  인간에 대한 변이체 (G를 A로 치환)                  상동 인간 유전자 서열 진뱅크 : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (게시-정지) : 2,142,351-2,142,351                  유기체의 국부 확인 게놈 서열 배경 :                  CTGGCACAGAGGCTGAGGCTGGCTGAATTACCCATGCTTTAAGAATTG                  A                  GATGGGTCCCAGTGAGCTCCTGGATTCTGCTGGTGAGACCTCCTGCTTC</p>
<p>서열 54                  유기체 : <i>Gorilla gorilla</i>                  인간에 대한 변이체 (G를 A로 치환) -- (CA의 2개의 염기쌍 결실)                  상동 인간 유전자 서열 진뱅크 : NT_009775.13                  게놈 서열 상 돌연변이의 좌표 (게시-정지) : 2,144,089-2,144,090                  유기체의 국부 확인 게놈 서열 배경 :                  CTCCTGATGTGATCATGTGTCTCACCCCTTTCAGGCTGAAAGCGACAGTG                  --                  GACGATGAGACCGACGATCCAGGAGGTATCAGAAATATGGTTACATTGG</p>

굵게 한줄로 밑줄 그은 염기는 각 서열 번호에 의해 확인된 인간에 대한 특정 유기체 변이체를 나타낸다.

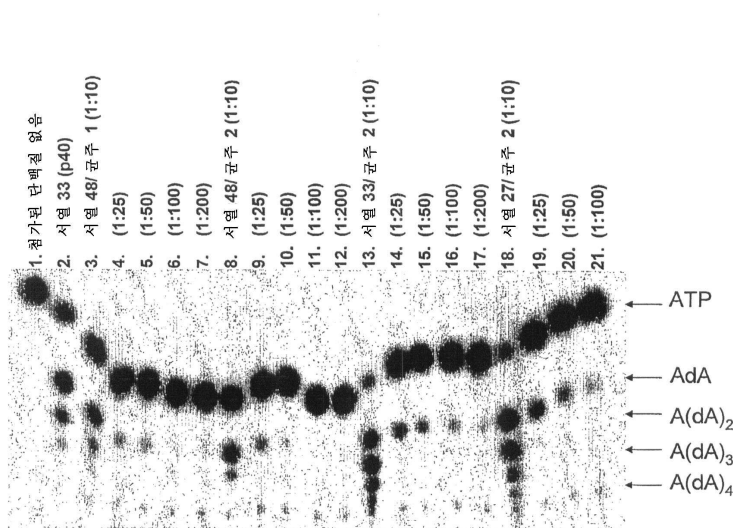
두줄로 밑줄 그은 염기는 인간 게놈 서열 및 나타낸 유기체 사이의 기타 공-국제화된 변이체를 나타낸다.

동의성 핵산 코드:

- R=A/G
- Y=C/T
- S=C/G
- K=G/T

도면7

이.콜라이(E.coli) 생산 단백질의 올리고아데닐레이트 신테타제 1 활성

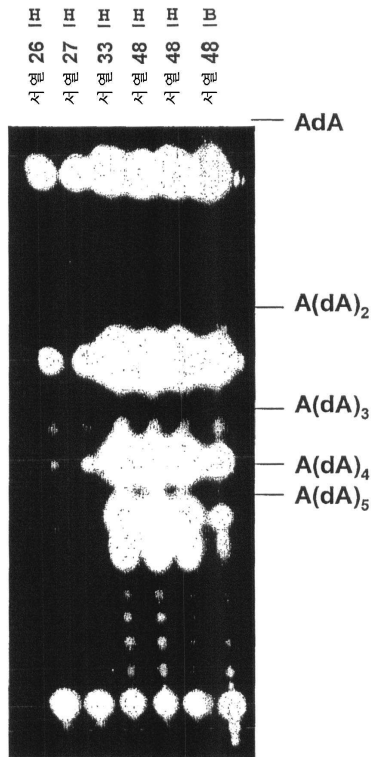


도 7: 이. 콜라이 (E. coli) 생산 단백질의 올리고아데닐레이트 신테타제 1 활성. 이. 콜라이에서 생산된 올리고아데닐레이트 신테타제 1 효소 (레인 3-21)의 활성을 박테리오파지 발현 시스템에서 생산된 효소 (레인 2)와 비교하였다. 이. 콜라이 샘플은 두 박테리아 균주 중 하나에서 생산되고, 1:10 (레인 3, 8, 13 및 18), 1:25 (레인 4, 9, 14 및 19), 1:50 등으로 희석하였다. 효소 반응은 α32PdATP의 존재하에 수행하고 효소 생성물은 폴리에틸이민 셀룰로스 코팅된 플레이트 상에서 박층 크로마토그래피에 의해 분해하였다. 2'-5'-올리고아데닐레이트의 유일하게 공지된 기능은 세포 항바이러스성 반응을 자극하는 것이다.

도면8

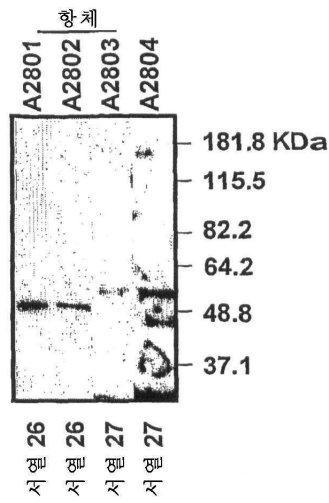
헬라(HeLa) 세포에서의 올리고아데닐레이트 신테라제 1 활성

도 8: 헬라 세포에서 발현된 OAS1 이소형의 활성. OAS1 이소형을 일시적 형질감염에 의해 헬라 세포 내에서 발현시키고, 조 세포 용해물을 기질로서  $\alpha^{32}P$ dATP를 사용하여 올리고아데닐레이트의 형성을 촉매화하는 능력에 의해 효소 활성에 대해 평가하였다. 올리고아데닐레이트를 12% 아크릴아미드/비스-아크릴아미드 겔 중에서 겔 전기영동으로 변형시켜 분리하였다. 조 헬라 추출물(H)을 박테리아(B)에서 발현되고 이로부터 정제된 효소에 대해 비교하였다.



도면9

OAS1의 신규 형태에 대해 특이적인 항체의 발생



도 9: OAS1의 신규 형태에 대해 특이적인 항체. 폴리클로날 항체를 토끼 면역화에 의한 각 OAS1 형태의 특유한 부분으로부터 유도된 합성 펩티드 서열에 대해 발생시켰다. 토끼 망상적혈구 용해물 커플링된 전사-번역 시스템으로 시험관내 합성된 단백질에 대해 각 항체를 시험하였다.

도면10

예시적인 단백질 형질 도입 도메인 시스템

서열번호	폴리펩티드 서열
서열 85	YGRKKRRQRRR
서열 86	RQIKIWFQNRRMKWKK
서열 87	MTSRRSVKSGPREVPRDEYEDLYYIPSSGMASPDSPPD TSRRGALQTRSQRGEVRFVQYDESDYALYGGSSSEDD EHPEVPRTRRPVSGAVLSGPGPARAPPPAGSGGAGRTP TTAPRAPRTQRVATKAPAAPAAETTRGRKSAQPESAAL PDAPASTAPTRSKTPAQLARKLHFSTAPPNPDAPWTPR VAGFNKRVFCAAVGRLAAMHARMAAVQLWDMSPRPT DEDLNELLGITTRVTVCBGKNLLQRANELVNPVVQDV DAATATRGRSAASRPTERPRAPARSASRRRPVE
서열 88	MSRKL FASILIGALLGIGAPPSAHAGADDVVDSSKSFVM ENFSSYHGTPGYVDSIQKGIQPKSGTQGN YDDDWKGFYSTDNKYDAAGYSVDNENPLSGKAGGVV KVTY PGLTKVLALKVDNAETIKKELGLSLTEPLM EQVGTEEFIKRFGDGASRVVLSL PFAEGSSSVEYINNWE QAKALSVELEINFETRGRGQDAMYEYMAQA CAGNRVRRSVGSSLSCLNDWDVIRDKTKTKIESLKEHG PKNKMSSEPNKTVSEKAKQYLEEFHQAL EHPSELKTVTGTNPVFAGANYAAWAVNVAQVIDSET ADNLEKTTAALSILPGIGSVMGLADGAVHHNT EEIVAQSIALLSLMV AQAIPLV GELVDIGFAAYNFVESIN LFQVVHNSYNRPA YSPGHKTQPFLHDGYA VSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLP TIPGKLDVNKSKTHISVNGRKIRMRCAID GDVTFCRPKSPVYVGNVHANLHVAFHRSSEKIHNSNE ISSDSIGVLGYQKTV DHTKVNSKLSLFFEIKS
서열 89	GDMGEWGWNEIFGAIAGFLG
서열 90	RRRRRRR
서열 91	RRQRRTSKLMKR
서열 92	GWTLNSAGYLLGKINLKALAALAKKIL
서열 93	WEAKLAKALAKALAKAHLAKALAKALKACEA
서열 94	KETWWETWWTEWSQPKKRKRV

서열목록

SEQUENCE LISTING

<110> Illumigen Biosciences, Inc.

<120> Detection of Mutations in a Gene  
Associated with Resistance to Viral Infection, OAS1

<130> 55382-22

<140> PCT/US04/035284

<141> 2004-10-22

<150> 60/513,888

<151> 2003-10-23

<150> 60/542,373

<151> 2004-02-06

<150> 60/554,758

<151> 2004-03-19

<150> 60/560,524

<151> 2004-04-08

<150> 60/578,323

<151> 2004-06-09

<150> 60/583,503

<151> 2004-06-28

<150> 60/605,243

<151> 2004-08-26

<160> 94

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 121

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon A

<400> 1

```
ccctcagagt gactgaagga aattcagaga agagctgaca cctaagttgt agattttgcc 60
ygaacaggtc agttgactgg crgctataaa cctaaccccc aaatctatgt caagctcatc 120
g                                                                                   121
```

<210> 2

<211> 121

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon A

<400> 2

```
attcagagaa gagctgacac ctaagttgta gattttgccy gaacagggtca gttgactggc 60
rgctataaac ctaacccccca aatctatgtc aagctcatcg aggagtgcac cgacctgcag 120
a                                                                                   121
```

<210> 3

<211> 121

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon A

<400> 3

```
atgtatggcc ctcccaccag gcctgggtggg tcctgtctcg actgggagca gaggaggggt 60
rggggggagga gagaaagaag ggagtgaagg gaagaggagg gggagtgggtg gagggaaata 120
```

g 121

<210> 4  
 <211> 121  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon B

<400> 4  
 actgaatcca gctgcaatgc aggaagactc cctgatgtga tcatgtgtct caccctttca 60  
 rgctgaaagc aacagtrcag acgatgagac cgacgatccc aggasgtatc agaaatatgg 120  
 t 121

<210> 5  
 <211> 121  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon B

<400> 5  
 atgcaggaag actccctgat gtgatcatgt gtctcaccct ttcargctga aagcaacagt 60  
 rcagacgatg agaccgacga tcccaggasg tatcagaaat atggttacat tggaacacat 120  
 g 121

<210> 6  
 <211> 121  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon B

<400> 6  
 gtgtctcacc ctttcargct gaaagcaaca gtrcagacga tgagaccgac gatcccagga 60  
 sgtatcagaa atatggttac attggaacac atgagtaccc tcattttctct catagacca 120  
 g 121

<210> 7  
 <211> 121  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon B

<400> 7  
 gggctccagt gttatctgga ccagttcctt catkttcagg tgggactctt gatccagaga 60  
 rgacaaaagct cctcagtgag ctggtgtata atccaggaca gaaccaggt ctctgactc 120

c 121

<210> 8  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon A primer A

<400> 8  
 aatggacctc aagacttccc 20

<210> 9  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon A primer B

<400> 9  
 attctccctt ctgttgcagg 20

<210> 10  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon B primer A

<400> 10  
 tccagatggc atgtcacagt 20

<210> 11  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon B primer B

<400> 11  
 gagctatgct tggcacatag 20

<210> 12  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>

<223> Probe sequence

<400> 12  
 gtagattttg ccygaacagg tcagt 25

<210> 13  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Probe sequence

<400> 13  
 cagttgactg gcrctataa accta 25

<210> 14  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Probe sequence

<400> 14  
 cagaggagg gtrgggggag gaga 24

<210> 15  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Probe sequence

<400> 15  
 tctcacctt tcargctgaa agcaac 26

<210> 16  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Probe sequence

<400> 16  
 gaaagcaaca gtrcagacga tgaga 25

<210> 17  
 <211> 26  
 <212> DNA

<213> Artificial Sequence

<220>

<223> Probe sequence

<400> 17

acgatcccag gasgtatcag aaatat 26

<210> 18

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Probe sequence

<400> 18

ttgatccaga gargacaaag ctccctc 26

<210> 19

<211> 28000

<212> DNA

<213> Homo sapiens

<400> 19

ggctgcagtg agctaagatt gtgccactgc accccagcct gggcaacaga gtgagaccct 60  
 gtctcaacaa aataaataaa tatatTTTTga attaaattaa aatgaaaaaa cagcctatca 120  
 aaaagtgtgg gatgtagtga aagtagtggt tagagggaaa gttatagcat tgaatgcata 180  
 tattaggaga agaaagatct aactcagaga gggatgaaca ggtgaggcat aggggatttt 240  
 aaggtagtga aactcttgca gggcagggga gccccaaaac tgggacaccg tccgggaagg 300  
 ctcttggtt cgtccaggaa ggaattcaag ggtgagctag aggaggaaga aaaaggTTTA 360  
 ttgaggcagc agggttacag ttctgtgact gtccctgcag agcagggcca ccccataggc 420  
 agtgccctgga gagcagcagc tcagggcact tctatagtca cattcatacc cactttttaa 480  
 tacgtgcaaa ttaagggcag gttattcaga aatttctaga agaagggTg taactgggtc 540  
 attgccaggg aatgagtaaa ctgttcatgg tgctggTgct catgccagcc agtcttcaat 600  
 ctggccctga gtcaagcccc acctcctatc tcaaaactat tctgcatggT gctgtaatgg 660  
 tgatacatg acatgttatg tttggcaaaa tccatagaac tgtaggacac aagagtgaac 720  
 cttaatgtaa accttaatgt aaatggactt ttgttaatta tgatgtatta atatcaattc 780  
 atcaattgta acaaatgtat cacagtactg ttaataatag aggaacttat tggcaggaga 840  
 gagagcttat ggaactctct gcacattcag ctcaatattt ctgtaagcct aaaactgctg 900  
 tgagaaataa aatccaacct gggcaacata gcaagacctt gtctctacaa aaaataaaaa 960  
 atgagctggg tgcagtaacg catgcctgta gtcccaggta ttcaggaggc tggggcagga 1020  
 ggatcccttg aaccaggaa gttgaggTtg cacgagTcat gatcatgcc ctgcactcca 1080  
 gcctggataa caaagcaaga tctgtctcc aaaaaataat aaaataaaat aaaaatctac 1140  
 taattgaaag ggaaaaaagc atagtataat accattctta acaaaaagaa aagagacctg 1200  
 tgtttTgtg tgtgttaaca tttgaaaaaa atctggaaag ctctatatca aaacgtttat 1260  
 agaggcaatt ttgtagTgtt agaatcatag atgatctttc cacttctctgg tttttctgac 1320  
 ttttttctt tttgcagtgg gcatgtattg ctggaaaata ccacagacaa ctgtgaaagg 1380  
 atttcatcaa caacaaaaaa aagataaaga aggaaacaca aaatctgtta aataagattt 1440  
 atgttggtcg gaggttaaaa tgcatttoca gagcagagTt cagagaaagg ctgggctgct 1500  
 tgttgctggc taaaggacaa agggtaagTt tcaggaagca gaagagtgag cagatgaaat 1560  
 tcagcactgg gatcagggga gtgtctgatt tgcaaaagga aagtgcaaag acagctcctc 1620

ccttctgagg aaacgaaacc aacagcagtc caagctcagt cagcagaaga gataaaagca 1680  
aacaggtctg ggaggcagtt ctgttgccac tctctctcct gtcaatgatg gatctcagaa 1740  
ataccccagc caaatctctg gacaagttca ttgaagacta tctcttgcca gacacgtggt 1800  
tccgcatgca aatcaaccat gccattgaca tcatctgtgg gttcctgaag gaaaggtgct 1860  
tccgaggtag ctctaccct gtgtgtgtgt ccaaggtggg aaaggtgagt ccaggcctgc 1920  
ctggccaggg gaggggtggc tgaatgtgca agagttgaga ttgagaatga gagagagaga 1980  
gagagagaag caaaaaccta gaaccaggg tgcaaatgtg agtacagaga gctgagatct 2040  
tctgggatgg tggtttctta tttatccaca cagcatgtta aaatagattc tggggtgaaa 2100  
tcctacatcc ctattattaa caagtgacc tccccctac ttcccgctga agtttatgaa 2160  
ccactgtcct gggcagatgc catttcagaa atagggaaact gaatcccagc tctggtaaac 2220  
agtttgctaa ttcgtggcca ggctaggggc tcaccatttc tgcagtgaag aatcatatgt 2280  
tttгааagca aatagcacct gctggctgca agacctgag caagtcactt aactactctg 2340  
tgttccaatt tcctcagcca taatcccca tactgttgca gtcttgccag tgcaccttaa 2400  
tgtagcagct tctcactgaa ttagtaccca aggttctttg tcctgcatcc aagaaaatta 2460  
aggaacatgg acacaaacgt gagcttgag caaaagttca gtaagcaaaa gaagaaagct 2520  
gtctccactg tggagagggg agtctgagtg gattgccaga ttgcagctga atgcaaaaaa 2580  
cttttataag aaaccactct cctccctgta actgtttgag aaacttttta tcagtaaagc 2640  
tgtgcaactt cccttacctt atgcagctgt gggatatatct ctaggcaagc ataaagcgt 2700  
gcttctcttg tatgtataac tgtggatttg ttttaggtaa gtcccactcc ctgcccagc 2760  
ttcaggcagg ccgctcctcc agggcccagc cttgaccatt tacctaactg attttctctc 2820  
tactttccct caatacctca tagggcctg tagattaagt aaaatagtaa gtgtgaacca 2880  
cccagcataa gctagtctg ggcacgtaa aggacaatgg gaaaagaaca cagatcctgg 2940  
aagaaggccc ccaggtttga attgtatttg ccacctacta gctgggtgat ggggctgata 3000  
tattatctca ctgagcatcc attttcccat ctgtaaaatg ggaactaatg ataatggcat 3060  
ccaatcata gcatcattgt gagcattata ggagtttaag acatgcaatg cttcagaac 3120  
agtggctagt gctccataat gttagtgatt gctcctgtca ttttatttag ggaggtttgc 3180  
ctcactaagc atcaattatt atttttgtcg tctttttcag ggtggctcct caggcaaggg 3240  
caccaccctc agaggccgat ctgacgctga cctggttgc ttctcagtc ctctcaccac 3300  
tttctaggat cagttaaact gccggggaga gttcatccag gaaattagga gacagctgga 3360  
agcctgtcaa agagagagag cattttccgt gaagtttgag gtccaggctc cacgctgggg 3420  
caacccccgt gcgctcagct tcgactgag ttcgctccag ctcggggagg ggggtggagt 3480  
cgatgtgctg cctgcctttg atgccctggg tgagagctcc cagcttcttt ttctccctct 3540  
tcccatttct gagcagaaat ctcccacagt ttgagagctt tttgccccaa cagggcatct 3600  
ctctaaagca ggggtggagg agatcttagg atctgtcccg gggcaagaat gaatcaggtc 3660  
atgatctatc acaggagaga cattaacag caaattggca taatgtgggg acaaagacat 3720  
ttcttacaga acatctgcaa ggcttactgg ttctgtttaa ggcaaaatgt gtgaatttta 3780  
tctttctaaa atcaggcagc aaagatgtgg cttaaagttc atgttactct catctttgtc 3840  
ccaacatgag atctcatcaa acgtatgcag cacgttggga gatagatatt tataatttgc 3900  
aggaacattt ggacaggaag tgtaacctct cagaggctcc cttgccacat caggagaatt 3960  
ggtaaaacca cactacctgt atcatatcat tattttaagt gataaatgat catctacatt 4020  
cagctctgat gagtaatagg tgttcaaaaa taggaacttc cagccaagtg tgggtgctca 4080  
tgcttgaat tccaacactt ttggaggctg aggcaggagg gtcgcttgag cccaggagt 4140  
caagaccagc ctgggcagca aagtgaacc tcatctctac taaaaattht aaaacattag 4200  
ccaagtgtgg tggtagatgc ctgtggctgc agttattcag gacgctgaga ctgaacgatc 4260  
acatgaggcc agccaaggat tcgagggtgc agtgagccac gaatgtacca ctgcactcca 4320  
tcctagggac agagcaagag caagaccctg tctcaatcaa tcagtcaatc agtcaaaact 4380  
atgaatttcc cagctgtata tgaaggcacc tcaaacacc acagtgaact cacagaggga 4440  
cacggaatag ttttagattht aatthtttgga gggaaatgcg atgacatctg tcacacaccg 4500  
caciaacggc tactattaaa ctgaacttac tgattagtgg ctactaatta atagttggtc 4560  
attaagcagt aattagtgat taattatcaa gtaattagga cttaattaaa ggaactgtca 4620  
cagtttctct tagtctagc gcagccatga aaaaaaaaaat gctgactctc caaagacacc 4680  
agggtatgag aaagtthttg attctctcct ttgtgccatc tcctgtgttg ggggctgaag 4740

tacaatggtt gtaaaagaca agagggagaa ggctggtcac agtggctcac gcctgtaatc 4800  
tcagcacttt gggaggccaa agtgggggga tcacttgaag tcaggaattc aagaccagcc 4860  
tggccaacat ggtgaaatct cacatctact aaaaatacaa aaattagctg ggcgtggtgg 4920  
tgtgtgcctg taatcacagc tactcgggag gctgaggcag gagaattgct tgaaccagg 4980  
agatggaggt tgcaatgagc caagatcatg ccattgcact ccagcctggg caacagagtg 5040  
agactccatc tcgaaaaaaa gaaaaaagaa aagaatataa ggagtgatta aaaaagaaaa 5100  
gaaaagaaaa ctaagtaggg tgaaacaata gatagccatg ggggttaggg agctttttta 5160  
gacagggctg tgagggaggg tccctgagcc tgagtggcga gaaggagtga gccttgggga 5220  
gatctggagg ttctgggaag aggaatggca agtgcagagg ccctgaagca gcaatgacca 5280  
tggcacattt gaggaagaga gaaaaagtca gagaagtaga aagtgggcaa aggaagcaag 5340  
acaggaggtg aggtgggaga ggttccagag accagatcac accagacatc attggccacc 5400  
ataagatctt tgggttttaa aattccagat gttatgggat gcaggaagca gcatgatcag 5460  
cagcattctc taggtgccag gttgagaaca ggctgtgggg gaacctgtaa agaggttgct 5520  
gccatagttc cggcgagtga cggtggtggc ttggatgggg tgatggcagt ggagagggca 5580  
ggagggagga tcaggaatgg acctcaagac ttcccagccc tgggtctgct gcacttttca 5640  
atcaaaccctc atggccaggg agattgtccc ctccagagtga ctgaaggaa ttcagagaag 5700  
agctgacacc taagttgtag attttgcccg aacaggtcag ttgactggcg gctataaacc 5760  
taaccccaa atctatgtca agctcatcga ggagtgcacc gacctgcaga aagagggcga 5820  
gttctccacc tgcttcacag aactacagag agacttctc aagcagcgcc ccaccaagct 5880  
caagagcctc atccgcctag tcaagcactg gtacccaaat gtatggccct cccaccaggc 5940  
ctggtgggtc ctgtctcgac tgggagcaga ggaggggtgg ggggaggaga gaaagaagg 6000  
agtgaaggga agaggagggg gagtgggtga gggaaataga gggatggaaa aaggagagaa 6060  
aggaaaaaga ggtggagaga ggagcctgca acagaaggga gaatgaaagg gaaggaagag 6120  
agaaaggaaag ggattttggt gttctgttca ctgctgtatc cccagaactt aaaacagagc 6180  
ctggtgcata ataggtgtaa ataactgttg aataaatgaa tcaatgctac atacacacac 6240  
gcacgcacac acacacagag agagagtcaa ccacactctt cagaagggtg ataagttaa 6300  
acaagagttt caaacaata tatgttcaga tgcccttcc tcccacttac tggctggctg 6360  
gccttaagta agcaacttaa ctttctgtt ctttctgct tcttatctgc aacgagtagc 6420  
atgccatagc tagagtaaca cggcatatag ttggtcctga taaatgtagc atattttagc 6480  
caccatagga gtacacataa taaaagctaa catgtagtat gtgcttagct tatctatggt 6540  
ttgtggatgt gatacaattt tctgttcaact tttaaatgcc ctgcatctta gtcaatttta 6600  
acagtgattc tgtaagttag ataaggttag gcattattat taaatccatt ttacaccaag 6660  
agaaacttgg gtcaaaaaga gaaactcctg ggtcacatgg ctcatcggc caataagtag 6720  
cagaagtaaa atttgaattt ggctgggcgc ggtagctcac accagtaatc ccagcacttt 6780  
gggaagccaa ggcaggtaga ttgcttgagc ccaggagttc aagactagcc tgagcaacat 6840  
ggcaaacctc cgtctctaca aaataaacta aaaatttagc cagggtgtgat ggtgagcacc 6900  
tgtagcccca gctactgggt aggctgaggt gggaggatcg cttgagcctg ggaggaggag 6960  
gttgagtaaa gtcaggattg cactactgcc ctccagcctg tgagacagag caagatcttc 7020  
tctcaaaaaa acaaaaaaac aaacaaaaa aaactcgaat ttgggtctat tgacttaaga 7080  
gtttgctgta taataatagg cattcaatgt atatttcttg aatgaacgaa tgaatgaaaa 7140  
taatcaggaa taaactttcc aatttaaaag taacacctct aggtaaaaaa aagacaatca 7200  
tttagtggcc agacttctaa gtgtttgctg ttctatgaat tgtaatcatg gagcctgagc 7260  
attgtagaat ttacaaaagc agttcctgac aaaagcagca ctgccccag ggacatattg 7320  
aaaattaatg aggtgtttt tggtaacat ggtgatggga ggacatgggt gctacttata 7380  
tttagtggaa agaagacaag aatgctagtt attgtacaat gatcaagaga gtccctgcaca 7440  
gccaagaatt gtctttttct ttctttcttg atgctgttct cttttaaacc aagacaagat 7500  
taacaataat ttaactccac taaccacat catcaccacc tccaacttat atgctacatt 7560  
tcttgatata ttcaagtctg tttatatttt caagtgcctc gaagtattat tgttttatag 7620  
ccaatgttt agttaatctg ctccacagatt taccactttc ttcactattc attctgtctt 7680  
acacctctaa cattccatct ggggtaattt toctaaatga tcatgcatcc tttgggattt 7740  
cttttgatga tggctctattg gtagtaaact ctctcagtta ttgtttgtct gaaaatgtca 7800  
tgcttttgcc ttcattgttg aagggtgctt ttgctgggtg gtcatttcag tatattgaat 7860

atatcattcc atcttccagt gtcattcatta aaaagtcagt tgccagtcta actgcagctc 7920  
 ttttataagt aacctgtctt attcttctgg ctgcatgtaa aagttttctc tttgtctttg 7980  
 attttgttta gcttcaatct gctgtgtctt aatgatgggt tcctattggt tgtcctgatt 8040  
 gggattccgt taagattcct gaatctgtgg gtagatatct ttaatcagtt ttgaaacttc 8100  
 tcagccattc ttctaaaata ttgattctcc ttcattctct cctcaccttc tagaattcca 8160  
 attaaatgta tgtagacc tgctctatct ttcatatctc tatactctct tctgtgtttt 8220  
 tcatcctttt gtctattttt ccatgcttta ttctgaatag ttcttctaa tctaccttc 8280  
 aattaactaa ttttctcttt agctatatct aatttgctgt aattaattac agttgccatt 8340  
 tttatcctaa aatttctatt tcatattttt gtatctgcca tggacttct tatggctttt 8400  
 aattccctgc taactattta aagttcttat tttatcctgt gaatatgata ttcttagtta 8460  
 ttttattttt aatttttatt atttgtaaat cttatgtttt atttacactt cttttctgtg 8520  
 acatgagcac acacagattc atgtgtatac atatatggct ctgatacctc tcctttcctg 8580  
 tcctcattca aaccactgat cacagagaga ggactatttt tttttatttt taatttttct 8640  
 atttcaatag gtttttgggg gaacaggtgg tgtttgggta catgaataag ttcttttagtg 8700  
 gtgattttgg tgcacccatc acccaaacag tgtacattgt acccaatgtg taatctttta 8760  
 acccttgcca caccaccacc tttcccgcga gtcgcgaaag tcccatgtat cattcttatg 8820  
 ctttgcttc ctcatagctt agctcccaca tatgagtggg aacatacaat gtttggtttt 8880  
 ccattcctga gttatttaat taaaataata gtatccaatt ccatccaggt tgctgtgaat 8940  
 gccattattt tgttctttt tatggttgag tagtattcca tgggtgtgtt gtgtgtgat 9000  
 aacatttttc tttatccact cattgattga tgggcatttg ggctggttcc atatttttgc 9060  
 aattgcaaat tgtgctgtta taaacatgtg tgtgcaagta tcttttttgt ataatgactt 9120  
 ctttctctc gggtagatac ctagtagtgg gattgctgga tcaaatggta gatctacttt 9180  
 tagttctata aggaatctcc aactgtttt ccatagtggg tgtatgagtt tacattcca 9240  
 ccaatgggtg aaaagtgtc ctttttcacc acatccacac caacatctat tattgtttga 9300  
 tttttatta tgaccattct tgcaggagtg aggtgggatc acattgtggg tttgatttgc 9360  
 attccctga taattaggga tgttgagcat tttccatat gcttgttggg atttgtttt 9420  
 ttttttttt tttcattatt atactttaag ttttagggta catgtgcaca atgtgcagg 9480  
 tagttacata tgtatacatg tgccatgctg gtgtgctgca cccattaacc cgtcatttag 9540  
 cattaggtat atctccta at gctatccctc cccaattccc cccacccgc ttgttggat 9600  
 ttgtatatct tcatthgaga attctctgtt catgtcctta gccactttt tgatgagatt 9660  
 tttttttct tgctgattcg tttgagttct ttgtagattc tggatattag ttggatgat 9720  
 agattgtgaa gattttctcc cattctgtgg gttgtctgtt aactctgcta attatttctt 9780  
 ttgctttgca gaagcttttt agtttaatta agtcccatct atttatcttt gtttttgtt 9840  
 catttgctt tgggttcttg gtcattgaagt ctttgcctaa gccaatgtgt aggagggtt 9900  
 ttccaatatt atcttctaga atctttatgg tttcaggtct tagatttaag tatttgatcg 9960  
 attttgagtt gaattttgta taaggggaga gagaaggatt cagtttcatt ctctacatg 10020  
 caacttgcca attatcctag gaccatttgt tgaatagggt gtcctttccc cattttatgt 10080  
 tttgtttgg tttgtcaaag atcagttggc tgtaatgtgt tggctttatt tctgggttat 10140  
 ctattctgtt ccatttctc acgtgactat ttttatacca gtaccatgtt gttttgggta 10200  
 ctatggcctt acagtatagt ttgaagtctg ataatgtaat gcctccagat ttgttctttt 10260  
 tacttagctt tgctttggct atgtgagctc ttttttgggt ccatatgaat tttaggattg 10320  
 tttttctag ttctgtgaag aatgatgggt gtattttgat ggggaattgca ttgaatttgt 10380  
 agattgtttt tgggagtatg gtcattttca caatattgat tctaccatt catgagcatg 10440  
 ggatgtgttt ccatttgttt gtgtcatcta tgattttctt tcagcaatgt tttgtagttt 10500  
 tcctgtaga gttcctagtt attttaaagt ctgtgttcgg tctttcagca tttaaagttt 10560  
 gtaggtttat tactatttct ctctttctg ttggtcataa ctcttagtgt tttgtttcct 10620  
 tgtgtgcctg gttacatatg tgctggctat tgtatttgaa aattatgtgt gaaataattt 10680  
 gaggttttgg attatgtata ttctccaga aagaatttca tttgcttctg tgcatttctt 10740  
 aggaacatta caagtcctt ttctcagtta attttctgt tagtctttatc agatagggtc 10800  
 tattacaacc actcacttag cagatgaaaa tcatgagggt ctgagagtct aagtcactca 10860  
 cttagaattg gacaatgggt aagccaggat tcaaaccac atcaataaga atccagcgt 10920  
 cttaacaagg ggccagtaca cttttttaa aaataaaagg ctagatagta aatatttttag 10980

actttgtgga ctgcacagcc tctgttgcaa ctactcaacc ctgcctttgt agcatgaatg 11040  
cagtcataaa ctatacataa atgaatgagc ctggattcgt tccaaggaaa ctttataaaa 11100  
acaggtggca ggctggattt ggcccatgag aagtgtagtt tacacaaaag ttgagcaaac 11160  
caattttttt ctgattgttt ttctctttct cagtgtaga agaagcttgg gaagctgcca 11220  
cctcagtatg ccctggagct cctgacggtc tatgcttggg agcgagggag catgaaaaca 11280  
catttcaaca cagcccaggg atttcggacg gtcttggaat tagtcataaa ctaccagcaa 11340  
ctctgcatct actggacaaa gtattatgac tttaaaaacc ccattattga aaagtacctg 11400  
agaaggcagc tcacgaaaacc caggtatgct atccccacat ggcttagctc ccctatgtaa 11460  
atgaacacct ggatacaggt acagtgcctt ggaaatggag gaggtgggag ggctccccac 11520  
ttagtgagaa tctcctgttg cccatcattg tactgggcat tttactactg ccatctgttt 11580  
taaacaccta cctccaacc tgtgagggcag gcaactatgcc aattatttta caggtgagta 11640  
aactgaggtt ctgagaggta aggagcttgt ccaaccctta acagaaaatg agtaaaatag 11700  
ctgcagtttg aactgaaata agaacagcag caacaacaat gatagtaatt gctcccaggt 11760  
attgaaagct tgttgaaga ctaacacatg ctaatataat agtaaaaatt attagcaata 11820  
ttactgatat gtatgttatg ttctagtcgc tgtgctgagc atttcatata actgggcttt 11880  
ttctatcctc acagcatagc ctttgagata ggtatgtgga actattccca ttttacagat 11940  
aagaatcctg aggcttagag agttcaagtg acctacccaa gggcacatca ctgataaagg 12000  
gcagaggtgg gattcaaacc cacatctgtc aggtgcaagt gcaaggctcc ttctcctcat 12060  
gctcactgcc tgctggggaa tagggcactg gggacatacc ccagggagcc cttcctcatg 12120  
ttctgagtcc cagttcatcc catgctgcta ttttgcctc ccaggagcat ctggactccc 12180  
tagacagagc cccagcttct cacctgtccc tctctaaatg ctgctctgca ggctgtgat 12240  
cctggaccog gcggacccta caggaaactt ggggtggtgga gacccaaagg gttggaggca 12300  
gctggcacia gaggtgagg cctggctgaa ttacctatgc tttagaatt gggatgggtc 12360  
cccagtgagc tcctggattc tgctgggtgag acctcctgct tcctcctgc cattcatccc 12420  
tgcccctctc catgaagctt gagacatata gctggagacc attctttcca aagaacttac 12480  
ctcttgccaa aggccattta tattcatata gtgacaggct gtgctccata ttttacagtc 12540  
attttggtca caatcgaggg tttctggaat tttcacatcc cttgtccaga attcattccc 12600  
ctaagagtaa taataaataa tctctaacac catttattga ctgtctgctt cgggctcagg 12660  
ttctgtccta agccctttaa tatgcactct ctcattaaat agtcacaaca atcccatgag 12720  
gcatttttaa aaatttttta ttattttaga ttcagagggc acatgtgcca tttgttacac 12780  
agctatattg tgtaatggtg gggtttgggc ctctattgat cctgtcgccc aatagtgaa 12840  
cagagtacc aaagaatt ttttcaacct ttgcctttct ccctcctcc tcctgttgg 12900  
agtccctagt gtctattggt cccatcttta gcagatgta agtatttgat tttctgtttc 12960  
tgggttaatt cacttcggat aatggcctcc agctgcaacc atgatttcat tctttcttat 13020  
ggctgcataa tactccatgg ttagatata ccacactttc tttatccagt tcacactgat 13080  
gggcacttaa gttgattcca tgactttgct attgtgaatc gtactgcat aacatacga 13140  
gtgccggtgt cttttgatag aatgatttct ttaccttgg gtagataccg agtagtggga 13200  
ttgctgggtt gaatggacat tctactttta gttatttgaa aagtcccatg aggcattgtt 13260  
tctatcattc coatcttaca gatgagacaa aggctcagag aggtgaggtc acttgctcaa 13320  
ggacatcagc taacaagtgg tggaaatgga attcaagctc agtggactct aaagccagtg 13380  
ctcatgtcac tgtgctaaac agcctgcctt gtcacatccc cacctctcat ctgaccaatg 13440  
ggagactctg agcagctgag tgacttgggt tgtcacacag ctaaacaggg gcaaaggacc 13500  
cagtcttgg tctttccacc tccaagcagg aatctgtctg attccagggg attgatgatg 13560  
ttgcagatgg ctaggaagca gactccagga tggaaatttag tatgcaggat gttctggggg 13620  
agagccactg gaaccagcac tcagggaaaag gggggaagaa aggataggaa ggaagcatga 13680  
aagagaatag ggagaagtga acagggatgc agagcgaatg ccagtttcag ccaactccaa 13740  
ggacagccct ggagctggaa tggccttttag agctgcccc a tggtgacaga ggtggccagg 13800  
cttctataacc cctacgtgga tcaactcactg tgcttgggca ctttgggaaa gggcatggct 13860  
ttgagcaaaa ggctctctgc agctgaggca acccctaaaa gggctgacgg ctgaagtctg 13920  
tctgctgacc actgtcccag cagctggggc ttgttagtcc ttctcaaag ggggatccag 13980  
atggcatgtc acagtgtcta ccgtaaagtgc tcaactgaatc cagctgcaat gcaggaagac 14040  
tcctgatgt gatcatgtgt ctcacccttt caggctgaaa gcaacagtgc agacgatgag 14100

accgacgatc ccaggaggta tcagaaatat ggttacattg gaacacatga gtaccctcat 14160  
 ttctctcata gaccagcac actccaggca gcatccaccc cacaggcaga agaggactgg 14220  
 acctgcacca tcctctgaat gccagtgcac cttgggggaa agggctccag tgttatctgg 14280  
 accagttcct tcattttcag gtgggactct tgatccagag aggacaaagc tcctcagtga 14340  
 gctggtgtat aatccaggac agaaccagg tctcctgact cctggccttc tatgccctct 14400  
 atcctatcat agataacatt ctccacagcc tcaacttcatt ccacctattc tctgaaaata 14460  
 ttccctgaga gagaacagag agatttagat aagagaatga aattccagcc ttgactttct 14520  
 tctgtgcacc tgatgggagg gtaatgtcta atgtattatc aataacaata aaaataaagc 14580  
 aataccatt tattgggtgt ttattaactt caaggcacag agccaagaag tacagatgca 14640  
 tatctagggg tattgtgtgt gtatatacat tgattcaaca agaaatattt attgagcact 14700  
 tactatgtgc caagcatagc tctgggcact gggaatatag caatgcacaa aagcagacag 14760  
 aatccctgt cctcatgacc ctgcagagcc aagacttcca gaatttttta aataaaaaaa 14820  
 tcctgtcct catggagtgt acatttgtgc aaaacatctt aatgtagat ggttttccta 14880  
 ttactaataa ttctgaaata agcatccttg atttatcctt tctccatata tctgagaaaa 14940  
 attatagaac ctccctgtgt gacacagcag ccactagcca catgtatcaa atgcttaaaa 15000  
 tgtagctagt ctaaactctac atgtgctgtg agtgcaaggt atatacttgg tttcaaagac 15060  
 ttagtacaata tgaaaagaat gccaaagtct tgccaactga taattttttt aattgtgtgc 15120  
 tgaatgaca atttttaaat atatttgagt taaatcaaat gaacttcata tctttctttt 15180  
 ccctttttaa ttgtggctac tagaaaatgt gaaatcatac atgtggcttg tgttatatta 15240  
 tgtatttcta ttggacagct ctgtcctcca aggtaaatca ctggattaaa gattcgacta 15300  
 tactgactta cattgccaca ttgtcacact gtccttggga ccaagaatca acatattcatt 15360  
 cataagactc taaaatataa aactctcata aatactcaca aaagaaccta gcatgctctg 15420  
 atcacctgag ttgtgtgtca cttttgggtg ctggtaagca gcctttggtc cgtccagatt 15480  
 atattcttcc atttagtccc cccacatccc tgtgagatgg gttttgttgt tattctcata 15540  
 atattaagtg gaataacttg aggttctaag aggttacagt gcttgcccag ggtcaccag 15600  
 ctggtcaggg gcagagtctg aacttgaacc ctaatccttc tctctctaaa gctcatgttc 15660  
 ttaatcactg tagcatggtc ttaatgtgtc ctcaattcatt gaaagcttat gttttcctac 15720  
 tctggcgcca tgagaaccag aagcatcaat gtccaggggc agggaaagat gaatgtcca 15780  
 gctcaacctg agcacagatt cacccttctt cggctctttt gtcctatttg tagactggat 15840  
 tagatgatgc cagtctactg attcaaatgt gaatctcttt cagaaaaacc ctcacagata 15900  
 cacctagaaa tgatgtttca ccagctatct tggcaccctt agcccagtca acttgtcaca 15960  
 taaaatcaat catcacacac tccatgctga taggcaagtg tggacatccc aatgtaatgg 16020  
 cttcattgta ttttactgtg tggaaaatgc acttgtgttg ctttttgaga gtgtttcatt 16080  
 ttatagcaat gccacaacca acagtagatt aatggaatca gtaattagtt gcttgatcga 16140  
 agagccacat ggccacgtga tcagccttcc atctacaaca ggaccagga gtataccggg 16200  
 attgtttttc aaagggcata gacatttttg ctgcaaatga catgggctta ctccagagtc 16260  
 ctggaggggt ctgtgttata attctctaaa tagatattgc cataatctct gaatgacacc 16320  
 ttttcccatg actaacactt tgaacaccat ggggtctgcc aggctggtgt gggccaagta 16380  
 gagggcgac ttgcaccaca gcctatacca gctgcagagc ctttaggac ttaataaaag 16440  
 ggtgctaatt tctgtacttc cctggctctg agatgtaata gtggttttaa tttactatcc 16500  
 tggccagggg ggtggcagtt tcaaaggat ccccatgacc tcctcactg tgatagccct 16560  
 cactaacc tcaggccaat gtcgggtagt gtcttttacc aagcatgtcc tttttgaata 16620  
 tacattcagg ggacaggga atgatcacca ggtcggcca tagacacagt gggcaaatga 16680  
 caagcctgac ttggccaggg ctccattttt caactctggc tcctatatgc ccctactctg 16740  
 atggggacag aatctgatga cgcttttctg tatcagtgtt gatcctctgc ccaacagtct 16800  
 tagaaatgtt gagtgactcc ccttttccag tgtatcatcc tctgaataaa tggccatagg 16860  
 tcctttgggg gaaaccatta ctatatatgc tatggtgtca cagcatcctt tcccaagggg 16920  
 actctgcctc ctctttgtta atgggtctta gtctgaaaac tgggttggtc agttccaaa 16980  
 atcaggaagg gatgatgact ttttattgga atggctcccc ttaaccctt ggtcgtccgg 17040  
 ccttgatttc ttctaattgg ataaattgag tagtgccctt tcaagatact catccactaa 17100  
 gtcactaaga ggccacattg tgttaaccag gtccttaacc ctgtatcagt cacgaccct 17160  
 ggctgccct ccaacctcac cattcattat agtcattgca cccacctggc ttctattgct 17220

ctgggagtcc tgcctgccct attgccctta tcaatccaag ctctgtaaca gcccctcttc 17280  
 ccaggaggcc accaccagaa cacctcatga tgcagggtgc cctctctcca cacattcctt 17340  
 atggcctcgg ggaatgtttt gtcccctggg actctatttc cagacttaca tagtaaacad 17400  
 atccactcct cttcactgtg cccatttccc ttggcctctt aataccttct tccatcagct 17460  
 ccacagcaat tctggaatth catacttttc tgcattttcta ggagccatcc taggagtccag 17520  
 ggaaaacctc ggcttcacaa gtaaatggga cttcagttgt gtgccaacat gcctggctaa 17580  
 tttttaaatt tttttgtaga gatggggtct tgctgtgttg tccaggctgg tcttgaactc 17640  
 ctggcctcaa acaatctttc ctccctggcc actcaaagtg ctgggatcac agctaccatt 17700  
 cccacacaca ttttcttatt tttaaaggct gtatgtgcac tgtatacatt aaatgtgtgc 17760  
 actgtataca ttaactgtgt gcactgtata cattaattht ctttaacgaa tttatccatt 17820  
 tatagttgct ttggttgttt ccacttgatt attgtgaata gtgctgcagt aaacatggga 17880  
 atgcagttat ctctttgata tctctgattc aattcttttg gatactcaga agtgggattg 17940  
 ctggaacata tctgtagttcc attttttaatt ttttgaggaa cctccatact gtttttcaca 18000  
 gtggctcacc aacagtgtgc aagagtthcc atccctccac atcctcacac ttgttatctt 18060  
 ttgttcattc tttaaaaaat gatagccatc ctaccaggag taagggtgata ttgcatcgtg 18120  
 attttgattt gcacttctct gataattagt gatattgagt atatthtcat agacctgta 18180  
 tccatttgtg tgtcttcttt ggagaaagat ctattcctat ccttagccca tttttaaadc 18240  
 aagttattaa tttttttgct agtgagtggg aggagttcct tacatattgt ggagattaac 18300  
 ccttattaga tgtatggttt gtgaatgttt tcttcattcc atagattgtc ttttcagccc 18360  
 gttgattgtt tcctttgcta tgcagaagct ttttagtctg atgtagtcca agttgcctat 18420  
 ttttgctttt gttgcctgtg ctttgcatat gtggccacct gatctttgac aagattgcca 18480  
 agaatacaca atggggaaag gacagtgtct tcaacaaatg gtgtttggaa agctgaatgt 18540  
 ccacatgcaa aagaataaaa ttggaccctt accttacagc atacacaaaa atcaactcaa 18600  
 aatggattaa agacttaaac gtaagacctg aaactgaaac tactagaaga aaacttaggg 18660  
 gaaaacttca tgacattggt ctttccagtg atttcaggga tgtgacacca aaagcacaga 18720  
 caacaaaagg catttatttt tatatggcat gtgaggaagg ggttcagttc cagttcttcc 18780  
 aatgtggatg ctcaattatc ccagcagcat ttattgaacg gatcatgttc tctccacttc 18840  
 tttgcaaagc cacctcttaa atattcccag agcccatcta tgtgggagtc tgtttctgga 18900  
 ctctgctctt ttccattggt ctatthtttg tgtccttgag ttaacacaa cttgtcttaa 18960  
 ttactataac cttataatth ttagtatctt tgggagaact ctgttctctt tttatcaagt 19020  
 cattggttcc tcttggccct ttttatttct acattaatth tatactcaat ttgtgaagct 19080  
 cctcccaaaa tatggggagg catttgatta gaattacaca tattagcttg ggaagaataa 19140  
 catctatatt ttatatatat atctatatag aatthgatga ttctgatcca tgaacttggg 19200  
 gtttctttc attaatthtt gtccctctta gtgacagtga caatgtthta tcattthccc 19260  
 tgcagaggtc tttcatggtt tttgtcatat gtattcccaa gtatthgttt cagtactatc 19320  
 gtacttggca tgctttcttt aatthcattt tgcaatgggt gcttcatggg agacttaaac 19380  
 atthtcagtg gtgatatttg gctatagcat tatagggtgat ctttatactt ttctaaatth 19440  
 tctgtggcca cgggaaataa taaagacact tttcttgcac agaaaaaaaa aactthtggg 19500  
 aagtatttct cacagctaag atctgatagt ttacgcaaag ttggcaggca caggctacag 19560  
 aaagctctgg ggtgctgttg tttggagctg ctggttcaag gacaaattca caagatttg 19620  
 aaacagagga ccaagtgtgt aaggacgagg gaaactatgg tataacattg aagcacctga 19680  
 gctggaaatt tctgagccct cagagataaa tttcctcagc tctccctgc cgagaaaaca 19740  
 aaactaaaaa gagttaatgt ttagccaaca gaaatgagag tgaagtthca agaagaaatt 19800  
 agggcacagc tggaaagtgt caaatgagg aaacgtthca tgtcaagttt gagatccagg 19860  
 gttgatgggt gaactctgca cgctcagctt catgttaggt ctcccagctc aacgaggggg 19920  
 tgaatthtgg tgtactgctg tctttgaggc cctgcgtgaa ctctcccaca ccctccccct 19980  
 tttcctcata caaatcccct ccttgcacac cctcagcc tgtctgtgag gtgccagggc 20040  
 cctctccca gcccaaccgca ggccagtatt gccctcccc aaactthcct tcaggcagat 20100  
 caaaccagc gctctggagt cactcagctt gggctcaaat cctgcttctg aatcttatga 20160  
 gacatcaagt cacttaccta actccttggg gaaacaggct ctctctgtc aaatagggtg 20220  
 gcatgctttg ggaggccaag gcaggaggat tctttagacc caggagttca agaccagcct 20280  
 gggtaacata gtgagaccct gtctctataa aaataaaaaa attagccagg tgtggtggct 20340

catgcctgta gtcccagcta ctcaggaggc tgaggaggaa gaattccttg agcccacgaa 20400  
 gtcaaggctg cagtgaacca tgatcaagcc gctgcactcc agcctgggca acagaatgcg 20460  
 accctgtctc agaaataagt aaataaataa atacataaat aaatttgggt gcaattgtgg 20520  
 ctcttagtgt tacagagagg actgaaggag ctaatggatg atggatttat ggcagtacct 20580  
 cacatatagc ttatcctaaa ggaagttaga gcttattatg atgattattc aaaaatattt 20640  
 atcaaaggctc tgccttgggc catgttctga gctaagtgtc ggggatgcaa agatgagcaa 20700  
 gagactcctc agggacaatt gtctgatgag ataacaggca ctatttatga gaggtccaat 20760  
 caatacagtt ctatttatct tataattatc caataaatga tataataatt tattagaggg 20820  
 ccaataaatc tgatggcagg agcctgtggg gggtaatggc cagggtctcac tattgtgcc 20880  
 aggctggtct tgaactcctg gcctcaagca ttccttctgc ctccacctcc cagcatgctg 20940  
 caattacaga ggcatgaaca actgcacctg gcctaaaatt ttatgttaat aaaaaaatgc 21000  
 atgtatttga ggagtacgac atgatgcttg aatatcatac tgtatcgggg gaaaccagcc 21060  
 cccgatattt caatgtaggt tcttttctat tttcccagg tgtcggctgg tctgagaaat 21120  
 aaagggaaag agtacaaaag agataaattt taaagctggg tgtccagggc agacatcaca 21180  
 tgtcggcagg ttctgtggtg ccccctgagc cataaaacca gcaagttttt attagcaatc 21240  
 ttcaaagga ggaaatgtac atataggggtg tgggtcacag agaacacatg attcaagggc 21300  
 gacaaaagat cacaaggcag aaggctcagg tgagatcaca aggtcagggc aaaactagaa 21360  
 ttactaagga agtttcatgt tccactgtgc atgcattgtc attgataaac atcttaacag 21420  
 tgttcaagag cagagaacca gtctgactag aattcgccag gctggaattt cctaactcta 21480  
 gcaagcctgg ggggtgtgca ggagaccagg gcgtgtttca tcccttatct gcaactggat 21540  
 aaggcagaca ccccagagc ggccatttta gagggcccc gggaaatgcat tcttttccca 21600  
 gggctgttaa ttattaatat tcttactgg ggaagaatt cagggatatt tctcttacct 21660  
 gtttttgta ataagagaaa tatggctctg tcttgctgg ctcccaggca gtcagacct 21720  
 atggttatct ccctgttcc ctgaacatcg ctattatcct gttcttctt caaggtgcc 21780  
 agatttcata ttgttcaaac acacatgctt tacgaacaat ttgtgcagtt aacgcaatca 21840  
 tcacagggtc ctgaggcaac atacatcctc agcttatgaa gatgacagga ttaagagatt 21900  
 aaagacagac ataggaaatt atgagagtat tgattgagga agtgataaat gtccatgaaa 21960  
 tcttcacaat ttatgttctt ctgtcatggc ttcagtaggt cctcctgctc ggggtcctg 22020  
 acttcccaca acatcactgt atacctgaaa ttagcattga ttctaattct ctggtcacac 22080  
 gtcattcaga gcataggatc ttcggtggat ttaagaagt cctccctccc tattctcagc 22140  
 catgtgactc ccagaatcct atgaaattaa agatcttgtg tttggccata ggagacttct 22200  
 attcaccatc tttgttctct cccacaaatg gcgaggcctc cagtctccca catgacagct 22260  
 ttgtactaaa atcaacctta ctctatagaa catgcatgat tgcagcagga ctactatgat 22320  
 cttggttga tgaattagtt aggataacat tagctgctgt aacaaacaga ccccagct 22380  
 gcagcatgac tcagatgcaa tataagtgtc ttttctactt atatcaagca agaatgacca 22440  
 gattttctga ttttttttcc cgctgtgcta atgtagggag aagtgtgttg aggtcacgtc 22500  
 acagttcaca gcaaccatct atgtttggga gcaaggatgc tggaaataga atccagcata 22560  
 cttgtagctt gtccataaatt acagacacct ttgcatttac tgaactgaat cctatgactt 22620  
 gaagacaaa gactgtagca tgcctgaagg gacagcctca gaactgtgga tgcctgtccc 22680  
 tctcctgggt tgggtttgtg ctgaccacag gcaaaccac tgaactcagg atcactgcat 22740  
 aaagtgacgt attaagccta gtgccaggat ctttagtggt tgcaggaagg tctcatgatt 22800  
 tttctgctaa ctcagctaat aggtagtctc ctgtcccttc aagttacaaa cacatcctc 22860  
 attcaggaat ttgaatgttc agtccttgat attttatcaa cccttcattc tgtggtcaa 22920  
 gggcagacgt ctctcccagt taagactgga agttagcagc ctgcctccag ggatggggtt 22980  
 gtggttgctc tctgctctct ctgttctctc ggggaagcagc agaatcattc catgggagga 23040  
 ctagagcagt tctttcttga gagaagagat ttactccttc caagtgtatt ggttagtgat 23100  
 tgctacataa caaactacc caaaactctc agtagcttaa aacaactgtg aagtgattga 23160  
 tgctcatgtg agcatgggtt ggttgatcca ggctgggctc agctgggcac ctctgtatat 23220  
 gctgtgggtt ctctgagct caactccttc ctgccagttg catttaaggc tgttctgtgc 23280  
 gtgtctctc gagcccaggc tgaatgggct ttggagatgg cagaagcaca agcgagtaaa 23340  
 cagagacaca tgaagccatt taaggcctca gccagcact gacgtgccgt cacttctcac 23400  
 attccactgg ccatgcaagt cacatggctg agcaaactca aggacttggg aagtaacct 23460

cgtcttagt gggaggaact acaacatccc atggcaaagc atggatccag gaagcagtga 23520  
 aatggggggc agtgactcag tttaccacac tgaggtctgg cagatggcta gaagtggcgc 23580  
 tttctcttga ggattggggg agaggggtgtg tttatggatt ctacagcaat cccaggcctg 23640  
 ggaacctctg taagtccctt tcccagggcc totacatctc tctctacat ggtcccgtct 23700  
 aactcctgcc tcatctagat ttctgtacca caccagctt cttgtgagca tctctttgct 23760  
 gccagagggc cctgtgagac aagcccccaa gatgacccca tgccagaatt gtcacccatg 23820  
 tgattcacat ctggaccaga ggaaatgctt cccaaatgaa gcccatcctg ccccatcag 23880  
 gcaactacag gccacttcag ctttcttggg tgtaaggcag acctcagaat ctctgtgtct 23940  
 cccagctaga tggaaagctt tccaaggggt ccttgggaag ccagctggat tgaggcaagg 24000  
 aatatcacac ccccatccat ctcccaaagg gaagcaacac atcacctgac aacagttctc 24060  
 tccagggcaa tctctttgcc aaacattgct cctctccaca ctccaacccc tttatgtatt 24120  
 cttaacatga ctgaggagcc cttttataaa ttctgcattt gggagtgtgt tgcatattct 24180  
 gtttggttcc tggagtacc tacacaagag cctcaggcaa tgagatttta tttatcaaac 24240  
 cattaccaga tgtcaggtgc tgtcccaaac actttgtaaa tgtaaccca tttaatcctc 24300  
 atataaatcc tatgaggtag gtgctattaa catctttatt ttagagatag ggaagctgag 24360  
 gcacagagag gtttaagtaat tagcccaaag tcacacagaa ggcaactgtc ttctcaccca 24420  
 ggcaagaaga gtcctttttt atgataataa ggtggaagaa ggcaaggggg agataagcaa 24480  
 gaagatgata atgatgatgg tggccttccc tgagttactg tgcttagcac ttagtgtgcg 24540  
 tggccttgcc ttgcctgtcc tttgagacag gtatgtcaga ctctcccat ttctcagatg 24600  
 aataaactga ggctgatata gtttgaatgt atgtctccac ccaaactca tgttgaaatg 24660  
 taatccccag tgttgagggt ggggcctggt gggagggtgat tggatcatgg ggggtgattt 24720  
 ctcatgggta gtttaacacc atcgccttgg tgctatcctt atgatagtga gtacattctc 24780  
 atgagacctg gttgtttaaa actgggtggc acctcctccc cactttctct cactcctgct 24840  
 ttgcaccat gatatgcctg ctccccctt gccttctgct atgattgtaa gcctcctgag 24900  
 gcctccccag aagccaagca gacgtcagta ccatgcttcc tgtaaagcct gcagaacct 24960  
 gagccgatta aacctctttt tctttataaa ttaccagtc tcagatatcc cttaatagca 25020  
 atgcaagaat ggctgatac agaggctttg agaggtcaag tgacctgcc aagggcacac 25080  
 cactgataaa ggacagatgt gggatttgaa cccaccttg ttaggcccc gtgcaaggcc 25140  
 ccttctctc atgctcactg cccactgagg agctgggac ttgggacaca ccctagggag 25200  
 ctgtgctctg ctggggcctt cctcatgccc caagtccctg cctgcccag gccggtgctc 25260  
 tgagtctct acagcccct cctcagctc actggcctca gtcacttgg ttcagggag 25320  
 gactaaggct ccctttggtc cctgccaatc tgacccccac cccagtgtga cctcaagag 25380  
 cctgactctg gccttttaca gaataaatct gaacaaaatc agggttcatt ttaatagcaa 25440  
 caggctgctg atgcagacct tatcaactcc atcaactgt gttctttcaa atgttacgct 25500  
 cccctggggg tgtcccacac cctgacgtca cacattcact cagtgaagcc catattcatt 25560  
 cggggagctc tttctttctc tcttctctaa cacacactct ccttggttag ctggctgctg 25620  
 atcaacatt ctggatatac tggttttcag gaaaatatga tggttggctt ccaatcccag 25680  
 attttctact gatgggttcc atatttacac catcttggcc acagtctctg ggtcaccatt 25740  
 tccacacata cccacacacc ataaagagag gctttttctg agtcttctgt tcatgcagtc 25800  
 tggaaattgta tttgcttttg ttttggggca tcttgggcag ctcatccac taataggcat 25860  
 aaccataacc attgcagtct ccacttactg acatttaca ctttccaggc acatgctagg 25920  
 gacctacat tcattatttc attttattct cacatcaca ccttgtgagg tggaggagca 25980  
 tgatggaagg aggaaggagc taaaagcaaa gaatctggag tccgactgcc tgggttcaa 26040  
 tctagctct accaatttcc agctctgtaa catogagcta atttcctaac ctctctatgc 26100  
 ctttcccta tctctaaaag gaagctgaca atagcatcta tctcatagga tttgtacgaa 26160  
 gattaaatga gtcaatattt ataaagtgtt cggaatgata cctgacatct ggtaatggtt 26220  
 tgataaataa aatccatttt aatgatgagg aaacaggctc agaagagggc gtcatttgc 26280  
 tcatgtggta cagatagggt ccagactcaa actcaagacc atctgactct aaaacatcta 26340  
 aaactgttgc cccgcatctt cttcatttga cagataataa aactgaggct cagagaagct 26400  
 aagtgactcg cctgggactg cacagcaaat caagacaaat aagacctagg gtctcctgac 26460  
 tgccagagtg gagatgctt tataggcttt tctcactgat gctctctggg cagacaggct 26520  
 cctcaatatg agagtgcac acactcctt cttcattttc aggtaaacct cacactggtt 26580

```

ggcagaagga actataccaa taattagtga acatgCGgtg aatttgcaac agacaagagg 26640
agcctcatta tcctatagtt tccaggttgc ttagggaggc agaaatcaca gcaaggaaaa 26700
ccttcaataa taaacagacg tctcataaaa ttaattgcaa cccaacctct ctctctactt 26760
aaaattagca tctatttcca gctctgcttt caatgccccca tatgaataca tgtgaactcc 26820
ctccctctct tcctccctgt ctccctctct ctctctctgt ccctcattaa aaaataaaaat 26880
ttaagaaaaa aatacaaggt agatttacac aaatagtggg atctcagtct tgagttagct 26940
gtgtatgact gaaaaggatg ctgtgggtaa taattatcat aaaaacaatg acatggccgg 27000
gcacagtggc tcacgcctgt aatcccagaa ctttgggagg ccgaggcagg cagatcactt 27060
gaggccagga gtttgagacc agcctggcca acatggtgaa actgcatctc tactaaaaat 27120
acaaaaatta gccgggcatc aatggccagc ccctgtaatc ccagctaatc aggagtctga 27180
ggcaggcgaa tcacttgaac ctgggggctg gtgggttgag tgagccgagc tcacaccact 27240
gcaactccagc ctgggcgaca gagtgagact acatctcaaa aaaacaaaaa caaacaagca 27300
aaaaaaaaacc cacagtaaca caaaagtaat aaaactgctg ctatttactc agtgcttattc 27360
tgatgccagc cactttgcta agcctatgaa tgcattatct ccccgttgct acagatgaga 27420
gaattgaggt tcagacaggt tgaaatcatt gctcccaaag tcacacaact ggtgagtggc 27480
agagctggga tgcaaacctt aaactgccag ccctcaaagc ctgtgctctt aatctccacc 27540
ctgctgtgct tccttgcca ttaattaag ctccacaggc acacattcca cgccctcctt 27600
tgctgtacaa tcccaggcaa gtcgctcagc ttctctgagc ctcagtttca taatctgtca 27660
aatggaggta acacaaataa ttctagttg tgaccaagaa tcatcataga aatctgcat 27720
ttccagccta ttgtgcaatt cctcaagcac tgtgactcca agtggcatca gctcctggaa 27780
gaacacactg tcttactggt gtttccctct ttgtcaactg atccccctt gaacctcact 27840
ctacctctgc tctcaatgcc ccactctactg ccacctgatt aaataaaatc ttttttgaaa 27900
atcataagtg tcatgagtaa ggtttcttgg tgttgatgta gaagaacaaa acagaattgt 27960
gaaatgagaa tcactgcagc tatcatgaag tcctgcctac 28000

```

<210> 20  
 <211> 218  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 127  
 <223> Xaa is Gly or Arg

<221> VARIANT  
 <222> 162  
 <223> Xaa is Gly or Ser

<400> 20  
 Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile  
 1 5 10 15

Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly  
 35 40 45  
 Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95  
 Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu  
 100 105 110  
 Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn  
 115 120 125  
 Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly  
 130 135 140  
 Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr  
 145 150 155 160  
 Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu  
 165 170 175  
 Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu  
 180 185 190  
 Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu  
 195 200 205  
 Ile Arg Leu Val Lys His Trp Tyr Gln Asn  
 210 215

<210> 21  
 <211> 240  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 53  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 137  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 184  
 <223> Xaa is Gly or Ser

<400> 21  
 Ser Val Ser Arg Arg Asp Lys Ser Lys Gln Val Trp Glu Ala Val Leu  
 1 5 10 15  
 Leu Pro Leu Ser Leu Leu Met Met Asp Leu Arg Asn Thr Pro Ala Lys  
 20 25 30  
 Ser Leu Asp Lys Phe Ile Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe



Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly  
 35 40 45  
 Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95  
 Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu  
 100 105 110  
 Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn  
 115 120 125  
 Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly  
 130 135 140  
 Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr  
 145 150 155 160  
 Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu  
 165 170 175  
 Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu  
 180 185 190  
 Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu  
 195 200 205  
 Ile Arg Leu Val Lys His Trp Tyr Gln Asn Val Trp Pro Ser His Gln  
 210 215 220  
 Ala Trp Trp Val Leu Ser Arg Leu Gly Ala Glu Glu Gly  
 225 230 235

<210> 23  
 <211> 259  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 53  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 137  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 184  
 <223> Xaa is Gly or Ser

<400> 23  
 Ser Val Ser Arg Arg Asp Lys Ser Lys Gln Val Trp Glu Ala Val Leu  
 1 5 10 15  
 Leu Pro Leu Ser Leu Leu Met Met Asp Leu Arg Asn Thr Pro Ala Lys

```

                20                25                30
Ser Leu Asp Lys Phe Ile Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe
    35                40                45
Arg Met Gln Ile Xaa His Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys
    50                55                60
Glu Arg Cys Phe Arg Gly Ser Ser Tyr Pro Val Cys Val Ser Lys Val
    65                70                75                80
Val Lys Gly Gly Ser Ser Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser
    85                90                95
Asp Ala Asp Leu Val Val Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp
    100                105                110
Gln Leu Asn Arg Arg Gly Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu
    115                120                125
Glu Ala Cys Gln Arg Glu Arg Ala Xaa Ser Val Lys Phe Glu Val Gln
    130                135                140
Ala Pro Arg Trp Gly Asn Pro Arg Ala Leu Ser Phe Val Leu Ser Ser
    145                150                155                160
Leu Gln Leu Gly Glu Gly Val Glu Phe Asp Val Leu Pro Ala Phe Asp
    165                170                175
Ala Leu Gly Gln Leu Thr Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr
    180                185                190
Val Lys Leu Ile Glu Glu Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe
    195                200                205
Ser Thr Cys Phe Thr Glu Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro
    210                215                220
Thr Lys Leu Lys Ser Leu Ile Arg Leu Val Lys His Trp Tyr Gln Asn
    225                230                235                240
Val Trp Pro Ser His Gln Ala Trp Trp Val Leu Ser Arg Leu Gly Ala
    245                250                255
Glu Glu Gly

```

<210> 24  
 <211> 256  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 64  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 138  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 185  
 <223> Xaa is Gly or Ser

<400> 24

Ser Val Ser Arg Arg Asp Lys Ser Lys Gln Val Trp Glu Ala Val Leu  
 1 5 10 15  
 Leu Pro Leu Ser Leu Leu Ser Met Met Asp Leu Arg Asn Thr Pro Ala  
 20 25 30  
 Lys Ser Leu Asp Lys Phe Ile Glu Asp Tyr Leu Leu Pro Asp Thr Cys  
 35 40 45  
 Phe Arg Met Gln Ile Xaa His Ala Ile Asp Ile Ile Cys Gly Phe Leu  
 50 55 60  
 Lys Glu Arg Cys Phe Arg Gly Ser Ser Tyr Pro Val Cys Val Ser Lys  
 65 70 75 80  
 Val Val Lys Gly Gly Ser Ser Gly Lys Gly Thr Thr Leu Arg Gly Arg  
 85 90 95  
 Ser Asp Ala Asp Leu Val Val Phe Leu Ser Pro Leu Thr Thr Phe Gln  
 100 105 110  
 Asp Gln Leu Asn Arg Arg Gly Glu Phe Ile Gln Glu Ile Arg Arg Gln  
 115 120 125  
 Leu Glu Ala Cys Gln Arg Glu Arg Ala Xaa Ser Val Lys Phe Glu Val  
 130 135 140  
 Gln Ala Pro Arg Trp Gly Asn Pro Arg Ala Leu Ser Phe Val Leu Ser  
 145 150 155 160  
 Ser Leu Gln Leu Gly Glu Gly Val Glu Phe Asp Val Leu Pro Ala Phe  
 165 170 175  
 Asp Ala Leu Gly Gln Leu Thr Gly Xaa Tyr Lys Pro Asn Pro Gln Ile  
 180 185 190  
 Tyr Val Lys Leu Ile Glu Glu Cys Thr Asp Leu Gln Lys Glu Gly Glu  
 195 200 205  
 Phe Ser Thr Cys Phe Thr Glu Leu Gln Arg Asp Phe Leu Lys Gln Arg  
 210 215 220  
 Pro Thr Lys Leu Lys Ser Leu Ile Arg Leu Val Lys His Trp Tyr Gln  
 225 230 235 240  
 Asn Val Trp Pro Ser His Pro Ala Cys Trp Tyr Leu Tyr Ile Phe Ile  
 245 250 255

<210> 25  
 <211> 233  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 127  
 <223> Xaa is Gly or Arg

<221> VARIANT

<222> 162

<223> Xaa is Gly or Ser

<400> 25

```

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile
 1           5           10           15
Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His
           20           25           30
Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly
           35           40           45
Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser
           50           55           60
Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val
65           70           75           80
Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly
           85           90           95
Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu
           100          105          110
Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn
           115          120          125
Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly
           130          135          140
Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr
145          150          155          160
Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu
           165          170          175
Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu
           180          185          190
Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu
           195          200          205
Ile Arg Leu Val Lys His Trp Tyr Gln Asn Val Trp Pro Ser His Pro
210          215          220
Ala Cys Trp Tyr Leu Tyr Ile Phe Ile
225          230

```

<210> 26

<211> 384

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> 31

<223> Xaa is Asp or Asn

<221> VARIANT

<222> 115

<223> Xaa is Leu or Phe

<221> VARIANT

<222> 127

<223> Xaa is Gly or Arg

<221> VARIANT

<222> 162

<223> Xaa is Ser or Gly

<221> VARIANT

<222> 280

<223> Xaa is Asn or Thr

<221> VARIANT

<222> 330

<223> Xaa is Pro or Ser

<400> 26

Met	Met	Asp	Leu	Arg	Asn	Thr	Pro	Ala	Lys	Ser	Leu	Asp	Lys	Phe	Ile
1				5					10					15	
Glu	Asp	Tyr	Leu	Leu	Pro	Asp	Thr	Cys	Phe	Arg	Met	Gln	Ile	Xaa	His
			20					25					30		
Ala	Ile	Asp	Ile	Ile	Cys	Gly	Phe	Leu	Lys	Glu	Arg	Cys	Phe	Arg	Gly
		35					40					45			
Ser	Ser	Tyr	Pro	Val	Cys	Val	Ser	Lys	Val	Val	Lys	Gly	Gly	Ser	Ser
	50					55					60				
Gly	Lys	Gly	Thr	Thr	Leu	Arg	Gly	Arg	Ser	Asp	Ala	Asp	Leu	Val	Val
65					70					75					80
Phe	Leu	Ser	Pro	Leu	Thr	Thr	Phe	Gln	Asp	Gln	Leu	Asn	Arg	Arg	Gly
				85					90					95	
Glu	Phe	Ile	Gln	Glu	Ile	Arg	Arg	Gln	Leu	Glu	Ala	Cys	Gln	Arg	Glu
			100					105					110		
Arg	Ala	Xaa	Ser	Val	Lys	Phe	Glu	Val	Gln	Ala	Pro	Arg	Trp	Xaa	Asn
		115					120					125			
Pro	Arg	Ala	Leu	Ser	Phe	Val	Leu	Ser	Ser	Leu	Gln	Leu	Gly	Glu	Gly
	130					135					140				
Val	Glu	Phe	Asp	Val	Leu	Pro	Ala	Phe	Asp	Ala	Leu	Gly	Gln	Leu	Thr
145					150					155					160
Gly	Xaa	Tyr	Lys	Pro	Asn	Pro	Gln	Ile	Tyr	Val	Lys	Leu	Ile	Glu	Glu
			165						170					175	
Cys	Thr	Asp	Leu	Gln	Lys	Glu	Gly	Glu	Phe	Ser	Thr	Cys	Phe	Thr	Glu
		180						185					190		
Leu	Gln	Arg	Asp	Phe	Leu	Lys	Gln	Arg	Pro	Thr	Lys	Leu	Lys	Ser	Leu
		195					200					205			
Ile	Arg	Leu	Val	Lys	His	Trp	Tyr	Gln	Asn	Cys	Lys	Lys	Lys	Leu	Gly
	210					215					220				
Lys	Leu	Pro	Pro	Gln	Tyr	Ala	Leu	Glu	Leu	Leu	Thr	Val	Tyr	Ala	Trp
225				230						235					240
Glu	Arg	Gly	Ser	Met	Lys	Thr	His	Phe	Asn	Thr	Ala	Gln	Gly	Phe	Arg
				245					250					255	
Thr	Val	Leu	Glu	Leu	Val	Ile	Asn	Tyr	Gln	Gln	Leu	Cys	Ile	Tyr	Trp
		260						265					270		
Thr	Lys	Tyr	Tyr	Asp	Phe	Lys	Xaa	Pro	Ile	Ile	Glu	Lys	Tyr	Leu	Arg
		275					280						285		
Arg	Gln	Leu	Thr	Lys	Pro	Arg	Pro	Val	Ile	Leu	Asp	Pro	Ala	Asp	Pro

290		295		300
Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala				
305		310		315
Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp				
		325		330
Gly Ser Pro Val Ser Ser Trp Ile Leu Leu Met Arg Gln Arg Leu Arg				
		340		345
Glu Val Arg Ser Leu Ala Gln Gly His Gln Leu Thr Ser Gly Gly Asn				
		355		360
Gly Ile Gln Ala Gln Trp Thr Leu Lys Pro Val Leu Met Ser Leu Cys				
		370		375
				380

- <210> 27
- <211> 459
- <212> PRT
- <213> Homo sapiens
  
- <220>
- <221> VARIANT
- <222> 31
- <223> Xaa is Asp or Asn
  
- <221> VARIANT
- <222> 115
- <223> Xaa is Leu or Phe
  
- <221> VARIANT
- <222> 127
- <223> Xaa is Gly or Arg
  
- <221> VARIANT
- <222> 162
- <223> Xaa is Ser or Gly
  
- <221> VARIANT
- <222> 280
- <223> Xaa is Asp or Thr
  
- <221> VARIANT
- <222> 330
- <223> Xaa is Pro or Ser
  
- <221> VARIANT
- <222> 361
- <223> Xaa is Gly or Arg
  
- <221> VARIANT
- <222> 429
- <223> Xaa is Lys or Arg
  
- <400> 27

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile  
 1 5 10 15  
 Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly  
 35 40 45  
 Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95  
 Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu  
 100 105 110  
 Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn  
 115 120 125  
 Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly  
 130 135 140  
 Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr  
 145 150 155 160  
 Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu  
 165 170 175  
 Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu  
 180 185 190  
 Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu  
 195 200 205  
 Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu Gly  
 210 215 220  
 Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp  
 225 230 235 240  
 Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly Phe Arg  
 245 250 255  
 Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr Trp  
 260 265 270  
 Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr Leu Arg  
 275 280 285  
 Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp Pro  
 290 295 300  
 Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala  
 305 310 315 320  
 Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp  
 325 330 335  
 Gly Ser Pro Val Ser Ser Trp Ile Leu Leu Leu Lys Ala Thr Val Gln  
 340 345 350  
 Thr Met Arg Pro Thr Ile Pro Gly Xaa Ile Arg Asn Met Val Thr Leu  
 355 360 365  
 Glu His Met Ser Thr Leu Ile Ser Leu Ile Asp Pro Ala His Ser Arg  
 370 375 380  
 Gln His Pro Pro His Arg Gln Lys Arg Thr Gly Pro Ala Pro Ser Ser  
 385 390 395 400  
 Glu Cys Gln Cys Ile Leu Gly Glu Arg Ala Pro Val Leu Ser Gly Pro  
 405 410 415

Val Pro Ser Phe Ser Gly Gly Thr Leu Asp Pro Glu Xaa Thr Lys Leu  
 420 425 430  
 Leu Ser Glu Leu Val Tyr Asn Pro Gly Gln Asn Pro Gly Leu Leu Thr  
 435 440 445  
 Pro Gly Leu Leu Cys Pro Leu Ser Tyr His Arg  
 450 455

<210> 28  
 <211> 355  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 127  
 <223> Xaa is Gly or Arg

<221> VARIANT  
 <222> 162  
 <223> Xaa is Ser or Gly

<221> VARIANT  
 <222> 280  
 <223> Xaa is Asp or Thr

<221> VARIANT  
 <222> 330  
 <223> Xaa is Pro or Ser

<400> 28  
 Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile  
 1 5 10 15  
 Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly  
 35 40 45  
 Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95  
 Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu



<221> VARIANT  
 <222> 162  
 <223> Xaa is Ser or Gly

<221> VARIANT  
 <222> 280  
 <223> Xaa is Asn or Thr

<221> VARIANT  
 <222> 330  
 <223> Xaa is Pro or Ser

<221> VARIANT  
 <222> 372  
 <223> Xaa is Arg or Gly

<400> 29

```

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile
 1           5           10           15
Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His
 20           25           30
Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly
 35           40           45
Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser
 50           55           60
Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val
 65           70           75           80
Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly
 85           90           95
Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu
 100          105          110
Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn
 115          120          125
Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly
 130          135          140
Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr
 145          150          155          160
Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu
 165          170          175
Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu
 180          185          190
Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu
 195          200          205
Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu Gly
 210          215          220
Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp
 225          230          235          240
Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly Phe Arg
 245          250          255
Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr Trp
 260          265          270
    
```

Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr Leu Arg  
 275 280 285  
 Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp Pro  
 290 295 300  
 Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala  
 305 310 315 320  
 Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp  
 325 330 335  
 Gly Ser Pro Val Ser Ser Trp Ile Leu Leu Val Asn Leu Thr Leu Val  
 340 345 350  
 Gly Arg Arg Asn Tyr Pro Ile Ile Ser Glu His Ala Val Asn Leu Gln  
 355 360 365  
 Gln Thr Arg Xaa Ala Ser Leu Ser Tyr Ser Phe Gln Val Ala  
 370 375 380

<210> 30  
 <211> 54  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 6  
 <223> Xaa is Ala or Thr

<221> VARIANT  
 <222> 15  
 <223> Xaa is Arg or Thr

<400> 30  
 Ala Glu Ser Asn Ser Xaa Asp Asp Glu Thr Asp Asp Pro Arg Xaa Tyr  
 1 5 10 15  
 Gln Lys Tyr Gly Tyr Ile Gly Thr His Glu Tyr Pro His Phe Ser His  
 20 25 30  
 Arg Pro Ser Thr Leu Gln Ala Ala Ser Thr Pro Gln Ala Glu Glu Asp  
 35 40 45  
 Trp Thr Cys Thr Ile Leu  
 50

<210> 31  
 <211> 519  
 <212> PRT  
 <213> Homo sapiens

<400> 31  
 Gly Cys Thr Gly Ala Ala Ala Gly Cys Ala Ala Cys Ala Gly Thr Arg  
 1 5 10 15  
 Cys Ala Gly Ala Cys Gly Ala Thr Gly Ala Gly Ala Cys Cys Gly Ala  
 20 25 30  
 Cys Gly Ala Thr Cys Cys Cys Ala Gly Gly Ala Ser Gly Thr Ala Thr





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

<210> 33  
 <211> 364  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 127  
 <223> Xaa is Gly or Arg

<221> VARIANT  
 <222> 162  
 <223> Xaa is Ser or Gly

<221> VARIANT  
 <222> 280  
 <223> Xaa is Asn or Thr

<221> VARIANT

<222> 330

<223> Xaa is Pro or Ser

<400> 33

```

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile
 1          5          10          15
Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His
 20          25          30
Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly
 35          40          45
Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser
 50          55          60
Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val
 65          70          75          80
Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly
 85          90          95
Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu
 100         105         110
Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn
 115         120         125
Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly
 130         135         140
Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr
 145         150         155         160
Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu
 165         170         175
Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu
 180         185         190
Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu
 195         200         205
Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu Gly
 210         215         220
Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp
 225         230         235         240
Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly Phe Arg
 245         250         255
Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr Trp
 260         265         270
Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr Leu Arg
 275         280         285
Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp Pro
 290         295         300
Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala
 305         310         315         320
Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp
 325         330         335
Gly Ser Pro Val Ser Ser Trp Ile Leu Leu Val Arg Pro Pro Ala Ser
 340         345         350
Ser Leu Pro Phe Ile Pro Ala Pro Leu His Glu Ala
 355         360

```

<210> 34  
 <211> 400  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 127  
 <223> Xaa is Gly or Arg

<221> VARIANT  
 <222> 162  
 <223> Xaa is Ser or Gly

<221> VARIANT  
 <222> 280  
 <223> Xaa is Asn or Thr

<221> VARIANT  
 <222> 330  
 <223> Xaa is Pro or Ser

<221> VARIANT  
 <222> 352  
 <223> Xaa is Ala or Thr

<221> VARIANT  
 <222> 361  
 <223> Xaa is Arg or Thr

<400> 34  
 Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile  
 1 5 10 15  
 Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly  
 35 40 45  
 Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95

Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu  
 100 105 110  
 Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn  
 115 120 125  
 Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly  
 130 135 140  
 Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr  
 145 150 155 160  
 Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu  
 165 170 175  
 Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu  
 180 185 190  
 Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu  
 195 200 205  
 Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu Gly  
 210 215 220  
 Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp  
 225 230 235 240  
 Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly Phe Arg  
 245 250 255  
 Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr Trp  
 260 265 270  
 Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr Leu Arg  
 275 280 285  
 Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp Pro  
 290 295 300  
 Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala  
 305 310 315 320  
 Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp  
 325 330 335  
 Gly Ser Pro Val Ser Ser Trp Ile Leu Leu Ala Glu Ser Asn Ser Xaa  
 340 345 350  
 Asp Asp Glu Thr Asp Asp Pro Arg Xaa Tyr Gln Lys Tyr Gly Tyr Ile  
 355 360 365  
 Gly Thr His Glu Tyr Pro His Phe Ser His Arg Pro Ser Thr Leu Gln  
 370 375 380  
 Ala Ala Ser Thr Pro Gln Ala Glu Glu Asp Trp Thr Cys Thr Ile Leu  
 385 390 395 400

<210> 35  
 <211> 414  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115

<223> Xaa is Leu or Phe

<221> VARIANT

<222> 127

<223> Xaa is Gly or Arg

<221> VARIANT

<222> 162

<223> Xaa is Ser or Gly

<221> VARIANT

<222> 280

<223> Xaa is Asn or Thr

<221> VARIANT

<222> 330

<223> Xaa is Pro or Ser

<221> VARIANT

<222> 397

<223> Xaa is Gly or Arg

<400> 35

Met	Met	Asp	Leu	Arg	Asn	Thr	Pro	Ala	Lys	Ser	Leu	Asp	Lys	Phe	Ile
1				5					10					15	
Glu	Asp	Tyr	Leu	Leu	Pro	Asp	Thr	Cys	Phe	Arg	Met	Gln	Ile	Xaa	His
			20					25					30		
Ala	Ile	Asp	Ile	Ile	Cys	Gly	Phe	Leu	Lys	Glu	Arg	Cys	Phe	Arg	Gly
		35					40					45			
Ser	Ser	Tyr	Pro	Val	Cys	Val	Ser	Lys	Val	Val	Lys	Gly	Gly	Ser	Ser
		50				55						60			
Gly	Lys	Gly	Thr	Thr	Leu	Arg	Gly	Arg	Ser	Asp	Ala	Asp	Leu	Val	Val
65					70					75					80
Phe	Leu	Ser	Pro	Leu	Thr	Thr	Phe	Gln	Asp	Gln	Leu	Asn	Arg	Arg	Gly
				85					90					95	
Glu	Phe	Ile	Gln	Glu	Ile	Arg	Arg	Gln	Leu	Glu	Ala	Cys	Gln	Arg	Glu
			100					105					110		
Arg	Ala	Xaa	Ser	Val	Lys	Phe	Glu	Val	Gln	Ala	Pro	Arg	Trp	Xaa	Asn
		115					120					125			
Pro	Arg	Ala	Leu	Ser	Phe	Val	Leu	Ser	Ser	Leu	Gln	Leu	Gly	Glu	Gly
		130				135						140			
Val	Glu	Phe	Asp	Val	Leu	Pro	Ala	Phe	Asp	Ala	Leu	Gly	Gln	Leu	Thr
145					150					155					160
Gly	Xaa	Tyr	Lys	Pro	Asn	Pro	Gln	Ile	Tyr	Val	Lys	Leu	Ile	Glu	Glu
				165					170					175	
Cys	Thr	Asp	Leu	Gln	Lys	Glu	Gly	Glu	Phe	Ser	Thr	Cys	Phe	Thr	Glu
			180					185					190		
Leu	Gln	Arg	Asp	Phe	Leu	Lys	Gln	Arg	Pro	Thr	Lys	Leu	Lys	Ser	Leu
		195					200					205			
Ile	Arg	Leu	Val	Lys	His	Trp	Tyr	Gln	Asn	Cys	Lys	Lys	Lys	Leu	Gly
		210				215						220			
Lys	Leu	Pro	Pro	Gln	Tyr	Ala	Leu	Glu	Leu	Leu	Thr	Val	Tyr	Ala	Trp

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

<210> 36  
 <211> 1387  
 <212> DNA  
 <213> Homo sapiens

<400> 36  
 gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccagcagcc 60  
 aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120  
 atcaaccatg ccattgacat catctgtggg ttcttgaagg aaaggtgctt ccgaggtagc 180  
 tctaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240  
 ctcagaggcc gatctgacgc tgacctgggt gtcttctca gtctctcac cacttttcag 300  
 gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360  
 caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420  
 cgtgcgctca gcttcgtact gagttcgctc cagctcgggg agggggtgga gttcgatgtg 480  
 ctgctgcct ttgatgccct gggtcagttg actggrgct ataaacctaa cccccaaatc 540  
 tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600  
 ttacagaac tacagagaga ctctctgaag cagcgcacca ccaagctcaa ggcctcatc 660  
 cgcctagtca agcactggta caaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aaccacatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccagggcc tgtgatcctg gaccggcg accctacagg aaacttgggt 960  
 ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020  
 ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct ggtgagacct 1080  
 cctgcttct cctgcccatt catccctgcc cctctccatg aagcttgaga catatagctg 1140  
 gagaccattc tttccaaaga acttacctct tgccaaaggc catttatatt catatagtga 1200  
 caggctgtgc tccatatttt acagtcattt tggtcacaat cgagggtttc tgggaattttc 1260

acatcccttg tccagaattc attcccctaa gagtaataat aaataatctc taacaccatt 1320  
 tattgactgt ctgcttcggg ctcaggttct gtcctaagcc cttaaatatg cactctctca 1380  
 ttaaata 1387

<210> 37

<211> 1699

<212> DNA

<213> Homo sapiens

<400> 37

gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccacagcc 60  
 aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120  
 atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180  
 tcctaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240  
 ctgagaggcc gatctgacgc tgacctgggt gtcttctca gtcctctcac cacttttcag 300  
 gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360  
 caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420  
 cgtgcgctca gcttcgtact gagttcgctc cagctcgggg aggggggtgga gttcgatgtg 480  
 ctgcctgcct ttgatgccct gggtcagttg actggcrgct ataaacctaa cccccaaatc 540  
 tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600  
 ttacagaac tacagagaga ctctctgaag cagcgcacca ccaagctcaa gaggcctcatc 660  
 cgcctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aacccatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccagggcc tgtgatcctg gaccggcgcc accctacagg aaacttgggt 960  
 ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020  
 ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct ggctgaaagc 1080  
 aacagtrcag acgatgagac cgacgatccc aggasgtatc agaaatatgg ttacattgga 1140  
 acacatgagt accctcattt ctctcataga cccagcacac tccaggcagc atccaccca 1200  
 caggcagaag aggactggac ctgcaccatc ctctgaatgc cagtgcactt tgggggaaag 1260  
 ggctccagtg ttatctggac cagttccttc attttcaggt gggactcttg atccagagar 1320  
 gacaaagctc ctgagtgagc tgggtgataa tccaggacag aaccaggtc tcctgactcc 1380  
 tggccttcta tgccctctat cctatcatag ataacattct ccacagcctc acttcattcc 1440  
 acctattctc tgaaaatatt ccctgagaga gaacagagag atttagataa gagaatgaaa 1500  
 ttccagcctt gactttcttc tgtgcacctg atgggagggg aatgtctaata gtattatcaa 1560  
 taacaataaa aataaagcaa ataccattta ttgggtgttt attaacttca aggcacagag 1620  
 ccaagaagta cagatgcata tctaggggta ttgtgtgtgt atatacattg attcaacaag 1680  
 aaatatttat tgagcactt 1699

<210> 38

<211> 1601

<212> DNA

<213> Homo sapiens

<400> 38

gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccacagcc 60  
 aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120  
 atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180  
 tcctaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240  
 ctgagaggcc gatctgacgc tgacctgggt gtcttctca gtcctctcac cacttttcag 300  
 gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360

caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420  
 cgtgCGctca gcttcgtact gagttcgctc cagctcgggg aggggggtgga gttcgatgtg 480  
 ctgCctgcct ttgatgccct gggtcagttg actggcrGct ataaaccta cccccaatc 540  
 tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600  
 tTcacagaac tacagagaga cttcctgaag cagcgcccca ccaagctcaa gagcctcatc 660  
 cgCctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgCctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aacccatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccaggcc tgtgatcctg gaccCGcgg accctacagg aaacttgggt 960  
 ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020  
 ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct gaccCagcac 1080  
 actccaggca gcatccaccc cacaggcaga agaggactgg acctgcacca tcctctgaat 1140  
 gccagtgc atttggggaa agggctccag tgttatctgg accagtctct tcattttcag 1200  
 gtgggactct tgatccagag argacaaagc tcctcagtga gctgggtgat aatccaggac 1260  
 agaaccCagg tctcctgact cctggccttc tatgCctct atcctatcat agataacatt 1320  
 ctccacagcc tacttCatt ccacctatc tctgaaaata ttccctgaga gagaacagag 1380  
 agatttagat aagagaatga aattccagcc ttgactttct tctgtgcacc tgatgggagg 1440  
 gtaatgtcta atgtattatc aataacaata aaaataaagc aaataccatt tattgggtgt 1500  
 ttattaactt caaggcacag agccaagaag tacagatgca tatctagggg tattgtgtgt 1560  
 gtatatacat tgattcaaca agaaatattt attgagcact t 1601

<210> 39

<211> 1698

<212> DNA

<213> Homo sapiens

<400> 39

gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccCagcc 60  
 aaatctctgg acaagtcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120  
 atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180  
 tctaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240  
 ctCagaggcc gatctgacgc tgacctgggt gtcttctca gtctctcac cacttttcag 300  
 gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360  
 caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420  
 cgtgCGctca gcttcgtact gagttcgctc cagctcgggg aggggggtgga gttcgatgtg 480  
 ctgCctgcct ttgatgccct gggtcagttg actggcrGct ataaaccta cccccaatc 540  
 tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600  
 tTcacagaac tacagagaga cttcctgaag cagcgcccca ccaagctcaa gagcctcatc 660  
 cgCctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgCctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aacccatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccaggcc tgtgatcctg gaccCGcgg accctacagg aaacttgggt 960  
 ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020  
 ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct gctgaaagca 1080  
 acagtrcaga cgatgagacc gacgatccca ggasgtatca gaaatattgt tacattggaa 1140  
 cacatgagta ccctcatttc tctcatagac ccagcacact ccaggcagca tccaccCac 1200  
 aggcagaaga ggactggacc tgcaccatcc tctgaatgcc agtgcattt gggggaaagg 1260  
 gctccagtgt tatctggacc agttccttca ttttcagggt ggactcttga tccagagarg 1320  
 acaaagctcc tcagtgagct ggtgtataat ccaggacaga acccaggctc cctgactcct 1380  
 ggccttctat gccctctatc ctatcataga taacattctc cacagcctca cttcattcca 1440

cctattctct gaaaatattc cctgagagag aacagagaga tttagataag agaatgaaat 1500  
 tccagccttg actttcttct gtgcacctga tgggagggta atgtctaatag tattatcaat 1560  
 aacaataaaa ataaagcaaa taccatztat tgggtgttta ttaacttcaa ggcacagagc 1620  
 caagaagtac agatgcatat ctagggggat tgtgtgtgta tatacattga ttcaacaaga 1680  
 aatatttatt gagcactt 1698

<210> 40

<211> 2491

<212> DNA

<213> Homo sapiens

<400> 40

gaggcagttc tgttgccact ctctctoctg tcaatgatgg atctcagaaa taccocagcc 60  
 aatctcttgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120  
 atcaaccatg ccattgacat catctgtggg ttcttgaagg aaaggtgctt ccgaggtagc 180  
 tcttaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240  
 ctgagaggcc gatctgacgc tgacctgggt gtcttctca gtcctctcac cacttttcag 300  
 gatcagtaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360  
 caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420  
 cgtgcgctca gcttcgtact gagttcgctc cagctcgggg aggggggtgga gttcgatgtg 480  
 ctgcctgcct ttgatgccct gggtcagttg actggcrgct ataaacctaa cccccaatc 540  
 tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600  
 ttcacagAAC tacagagaga ctctctgaag cagcgcacca ccaagctcaa gagcctcatc 660  
 cgcttagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aacccatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccagggcc tgtgatcctg gaccocggcg accctacagg aaacttgggt 960  
 ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020  
 ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct gatgagaaa 1080  
 aggctcagag aggtgaggtc acttgctcaa ggacatcagc taacaagtgg tggaaatgga 1140  
 attcaagctc agtggaactct aaagccagtg ctcatgtcac tgtgctaaac agcctgcctt 1200  
 gtcacatccc cacctctcat ctgaccaatg ggagactctg agcagctgag tgacttgggt 1260  
 tgtcacacag ctaaacaggg gcaaaaggacc cagtcttggg tctttccacc tccaagcagg 1320  
 aatctgtctg attccagggg attgatgatg ttgcagatgg ctaggaagca gactccagga 1380  
 tggaaatttag tatgcaggat gttctggggg agagccactg gaaccagcac tcagggaaag 1440  
 gggggaagaa aggataggaa ggaagcatga aagagaatag ggagaagtga acagggatgc 1500  
 agagcgaatg ccagtttcag ccaactccaa ggacagccct ggagctggaa tggcctttag 1560  
 agctgcccc a tggtgacaga ggtggccagg ctctataacc cctacgtgga tcaactactg 1620  
 tgcttgggca ccttgggaaa gggcatggct ttgagcaaaa ggctctctgc agctgaggca 1680  
 acccctaaaa gggctgacgg ctgaagtctg tctgctgacc actgtcccag cagctggggc 1740  
 ttgtagtcc ttctcaaaag ggggatccag atggcatgtc acagtgtcta ccgtaaagtgc 1800  
 tcaactgaatc cagctgcaat gcaggaagac tccctgatgt gatcatgtgt ctacccttt 1860  
 cargctgaaa gcaacagtrc agacgatgag accgacgatc ccaggasgta tcagaaatat 1920  
 ggttacattg gaacacatga gtaccctcat ttctctcata gaccagcac actccaggca 1980  
 gcatccacc cacaggcaga agaggactgg acctgcacca tctctgaat gccagtgc 2040  
 cttgggggaa agggctccag tgttatctgg accagttcct tcattttcag gtgggactct 2100  
 tgatccagag argacaaagc tctcagtgga gctgggtgat aatccaggac agaaccagg 2160  
 tctctgact cctggccttc tatgccctct atcctatcat agataacatt ctccacagcc 2220  
 tcaactcatt ccacctattc tctgaaaata tccctgaga gagaacagag agatttagat 2280  
 aagagaatga aattccagcc ttgactttct tctgtgcacc tgatgggagg gtaatgtcta 2340  
 atgtattatc aataacaata aaaataaagc aaataccatt tattgggtgt ttattaactt 2400  
 caaggcacag agccaagaag tacagatgca tatctagggg tattgtgtgt gtatatacat 2460

tgattcaaca agaaatattt attgagcact t

2491

<210> 41

<211> 1510

<212> DNA

<213> Homo sapiens

<400> 41

```

gaggcagttc tgttgccact ctctctoctg tcaatgatgg atctcagaaa taccccagcc 60
aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180
tcctaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240
ctcagaggcc gatctgacgc tgacctgggt gtcttcctca gtcctctcac cacttttcag 300
gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360
caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420
cgtgcgctca gcttcgtact gagttcgctc cagctcgggg aggggggtga gttc gatgtg 480
ctgctgcct ttgatgccct gggtcagttg actggcrgct ataaacctaa cccccaaatc 540
tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600
ttcacagaac tacagagaga cttcctgaag cagcgcccca ccaagctcaa gagcctcatc 660
cgcctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720
tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780
aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840
atctactgga caaagtatta tgactttaaa aacccatta ttgaaaagta cctgagaagg 900
cagctcacga aaccaggcc tgtgatcctg gacccggcgg accctacagg aaacttgggt 960
ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020
ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct ggtaaacctc 1080
aactggttg gcagaaggaa ctataccaat aattagttaa catgcggtga atttgcaaca 1140
gacaagasga gcctcattat cctatagttt ccaggttgct tagggaggca gaaatcacag 1200
caaggaaaac cttcaataat aaacagacgt ctcataaaat taattgcaac ccaacctctc 1260
tctctactta aaattagcat ctatttccag ctctgctttc aatgccccat atgaatacat 1320
gtgaactccc tccctctctt cctccctgtc tcttctctc tctctctgtc cctcattaaa 1380
aaataaaatt taagaaaaaa atacaaggta gatttacaca aatagtggga tctcagtctt 1440
gagttagctg tgtatgactg aaaaggatgc tgtggttaat aattatcata aaaacaatga 1500
catggccggg                                     1510
    
```

<210> 42

<211> 1695

<212> DNA

<213> Homo sapiens

<400> 42

```

gaggcagttc tgttgccact ctctctoctg tcaatgatgg atctcagaaa taccccagcc 60
aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180
tcctaccctg tgtgtgtgtc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240
ctcagaggcc gatctgacgc tgacctgggt gtcttcctca gtcctctcac cacttttcag 300
gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360
caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420
cgtgcgctca gcttcgtact gagttcgctc cagctcgggg aggggggtga gttc gatgtg 480
ctgctgcct ttgatgccct gggtcagttg actggcrgct ataaacctaa cccccaaatc 540
tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600
ttcacagaac tacagagaga cttcctgaag cagcgcccca ccaagctcaa gagcctcatc 660
    
```

cgcctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aaccccatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccagggcc tgtgatcctg gaccggcgg accctacagg aaacttgggt 960  
 ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020  
 ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct gataaaactg 1080  
 aggctcagag aagctaagtg actcgcctgg gactgcacag caaatcaaga caaataagac 1140  
 ctagggtctc ctgactgcca gagtgagat gcttctatag gcttttctca ctgatgctct 1200  
 ctgggcagac aggtcctca atatgagagt gacacacact ctttcttca ttttcaggta 1260  
 aacctcacac tggttggcag aaggaactat ccaataatta gtgaacatgc ggtgaatttg 1320  
 caacagacaa gasgagcctc attatcctat agtttccagg ttgcttaggg aggcagaaat 1380  
 cacagcaagg aaaaccttca ataataaaca gacgtctcat aaaattaatt gcaacccaac 1440  
 ctctctctct acttaaaatt agcatctatt tccagctctg ctttcaatgc cccatatgaa 1500  
 tacatgtgaa ctccctccct ctcttctcc ctgtctcct ctctctctct ctgtccctca 1560  
 ttaaaaaata aaatttaaga aaaaaatata aggtagattt acacaaatag tgggatctca 1620  
 gtcttgagtt agctgtgtat gactgaaaag gatgctgtgg ttaataatta tcataaaaac 1680  
 aatgacatgg ccggg 1695

<210> 43  
 <211> 1355  
 <212> DNA  
 <213> Homo sapiens

<400> 43  
 gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccacagcc 60  
 aatctctgga acaagtcat tgaagactat ctcttgccag acacgtggtt ccgcatgcaa 120  
 atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180  
 tctaccctg tgtgtgtgtc caagggtgga aagggtggct cctcaggcaa gggcaccacc 240  
 ctgagaggcc gatctgacgc tgacctgggt gtcttctca gtcctctcac cacttttcag 300  
 gatcagtaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360  
 caaagagaga gagcattttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420  
 cgtgcgctca gcttcgtact gagttcgctc cagctcgggg aggggggtgga gttcgtatgtg 480  
 ctgcctgcct ttgatgccct gggtcagttg actggrgct ataaacctaa cccccaaatc 540  
 tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600  
 ttcacagaac tacagagaga ctctctgaag cagcgcacca ccaagctcaa gagcctcatc 660  
 cgcctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720  
 tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780  
 aacacagccc agggatttcg gacggtcttg gaattagtca taaactacca gcaactctgc 840  
 atctactgga caaagtatta tgactttaaa aaccccatta ttgaaaagta cctgagaagg 900  
 cagctcacga aaccagggta aacctcacac tggttggcag aaggaactat ccaataatta 960  
 gtgaacatgc ggtgaatttg caacagacaa gasgagcctc attatcctat agtttccagg 1020  
 ttgcttaggg aggcagaaat cacagcaagg aaaaccttca ataataaaca gacgtctcat 1080  
 aaaattaatt gcaacccaac ctctctctct acttaaaatt agcatctatt tccagctctg 1140  
 ctttcaatgc cccatatgaa tacatgtgaa ctccctccct ctcttctcc ctgtctcctt 1200  
 ctctctctct ctgtccctca ttaaaaaata aaatttaaga aaaaaatata aggtagattt 1260  
 acacaaatag tgggatctca gtcttgagtt agctgtgtat gactgaaaag gatgctgtgg 1320  
 ttaataatta tcataaaaac aatgacatgg ccggg 1355

<210> 44  
 <211> 1035  
 <212> DNA

<213> Homo sapiens

<400> 44

```

gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccaccagcc 60
aaatctctggy acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180
tctaccctg tgtgtgtgtc caaggtggta aagtgtaga agaagcttgg gaagctgcca 240
cctcagtatg ccctggagct cctgacggtc tatgcttggg agcgagggag catgaaaaca 300
catttcaaca cagcccaggg atttcggacg gtcttggaat tagtcataaa ctaccagcaa 360
ctctgcatct actggacaaa gtattatgac tttaaaaacc ccattattga aaagtacctg 420
agaaggcagc tcacgaaacc caggcctgtg atcttgacc cggcggacc tacaggaaac 480
ttgggtggtg gagacccaaa gggttggagg cagctggcac aagaggctga ggcttggctg 540
aattacccat gctttaagaa ttgggatggg tcccagtgta gctcctggat tctgctggta 600
aacctcacac tggttggcag aaggaactat ccaataatta gtgaacatgc ggtgaatttg 660
caacagacaa gasgagcctc attatcctat agtttccagg ttgcttaggg aggcagaaat 720
cacagcaagg aaaaccttca ataataaaca gacgtctcat aaaattaatt gcaaccaaac 780
ctctctctct acttaaaatt agcatctatt tccagctctg ctttcaatgc cccatataaa 840
tacatgtgaa ctccctccct ctcttctctc ctgtctcctt ctctctctct ctgtccctca 900
ttaaaaaata aaatttaaga aaaaaataca aggtagattt acacaaatag tgggatctca 960
gtcttgagtt agctgtgtat gactgaaaag gatgctgtgg ttaataatta tcataaaaac 1020
aatgacatgg ccggg                                     1035
    
```

<210> 45

<211> 881

<212> DNA

<213> Homo sapiens

<400> 45

```

gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccaccagcc 60
aaatctctggy acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcaaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180
tctaccctg tgtgtgtgtc caaggtggta aagtgtaga agaagcttgg gaagctgcca 240
cctcagtatg ccctggagct cctgacggtc tatgcttggg agcgagggag catgaaaaca 300
catttcaaca cagcccaggg atttcggacg gtcttggaat tagtcataaa ctaccagcaa 360
ctctgcatct actggacaaa gtattatgac tttaaaaacc ccattattga aaagtacctg 420
agaaggcagc tcacgaaacc caggtaaac tcacactggg tggcagaagg aactatccaa 480
taattagtga acatgcggtg aatttgcaac agacaagasg agcctcatta tcctatagtt 540
tccaggtgac ttagggaggg agaaatcaca gcaaggaaaa cttcaataa taaacagacg 600
tctcataaaa ttaattgcaa cccaacctct ctctctactt aaaattagca tctatttcca 660
gtctgtcttt caatgcccc aatgaataca tgtgaaactc ctccctctct tcctccctgt 720
ctccttctct ctctctctgt ccctcattaa aaaataaaat ttaagaaaa aatacaaggt 780
agatttacac aaatagtggg atctcagctt tgagttagct gtgtatgact gaaaaggatg 840
ctgtgggtaa taattatcat aaaaacaatg acatggccgg g                                     881
    
```

<210> 46

<211> 224

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> 31

<223> Xaa is Asp or Asn

<221> VARIANT

<222> 122

<223> Xaa is Asn or Thr

<221> VARIANT

<222> 172

<223> Xaa is Pro or Ser

<221> VARIANT

<222> 214

<223> Xaa is Arg or Gly

<400> 46

Met	Met	Asp	Leu	Arg	Asn	Thr	Pro	Ala	Lys	Ser	Leu	Asp	Lys	Phe	Ile
1				5					10					15	
Glu	Asp	Tyr	Leu	Leu	Pro	Asp	Thr	Cys	Phe	Arg	Met	Gln	Ile	Xaa	His
			20					25					30		
Ala	Ile	Asp	Ile	Ile	Cys	Gly	Phe	Leu	Lys	Glu	Arg	Cys	Phe	Arg	Gly
		35					40					45			
Ser	Ser	Tyr	Pro	Val	Cys	Val	Ser	Lys	Val	Val	Lys	Cys	Lys	Lys	Lys
	50					55					60				
Leu	Gly	Lys	Leu	Pro	Pro	Gln	Tyr	Ala	Leu	Glu	Leu	Leu	Thr	Val	Tyr
65					70					75					80
Ala	Trp	Glu	Arg	Gly	Ser	Met	Lys	Thr	His	Phe	Asn	Thr	Ala	Gln	Gly
				85					90					95	
Phe	Arg	Thr	Val	Leu	Glu	Leu	Val	Ile	Asn	Tyr	Gln	Gln	Leu	Cys	Ile
			100					105					110		
Tyr	Trp	Thr	Lys	Tyr	Tyr	Asp	Phe	Lys	Xaa	Pro	Ile	Ile	Glu	Lys	Tyr
			115					120					125		
Leu	Arg	Arg	Gln	Leu	Thr	Lys	Pro	Arg	Pro	Val	Ile	Leu	Asp	Pro	Ala
	130					135					140				
Asp	Pro	Thr	Gly	Asn	Leu	Gly	Gly	Gly	Asp	Pro	Lys	Gly	Trp	Arg	Gln
145					150					155					160
Leu	Ala	Gln	Glu	Ala	Glu	Ala	Trp	Leu	Asn	Tyr	Xaa	Cys	Phe	Lys	Asn
				165					170					175	
Trp	Asp	Gly	Ser	Pro	Val	Ser	Ser	Trp	Ile	Leu	Leu	Val	Asn	Leu	Thr
			180					185					190		
Leu	Val	Gly	Arg	Arg	Asn	Tyr	Pro	Ile	Ile	Ser	Glu	His	Ala	Val	Asn
	195						200					205			
Leu	Gln	Gln	Thr	Arg	Xaa	Ala	Ser	Leu	Ser	Tyr	Ser	Phe	Gln	Val	Ala
	210					215						220			

<210> 47

<211> 137

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> 31

<223> Xaa is Asp or Asn

<221> VARIANT

<222> 122

<223> Xaa is Asn or Thr

<400> 47

```

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile
 1          5          10          15
Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His
 20          25          30
Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly
 35          40          45
Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Cys Lys Lys Lys
 50          55          60
Leu Gly Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr
 65          70          75          80
Ala Trp Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly
 85          90          95
Phe Arg Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile
 100         105         110
Tyr Trp Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr
 115         120         125
Leu Arg Arg Gln Leu Thr Lys Pro Arg
 130         135

```

<210> 48

<211> 346

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> 31

<223> Xaa is Asp or Asn

<221> VARIANT

<222> 115

<223> Xaa is Leu or Phe

<221> VARIANT

<222> 127

<223> Xaa is Gly or Arg

<221> VARIANT

<222> 162

<223> Xaa is Ser or Gly

<221> VARIANT

<222> 280

<223> Xaa is Asn or Thr

<221> VARIANT

<222> 330

<223> Xaa is Pro or Ser

<400> 48

```

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile
 1          5          10          15
Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His
 20          25          30
Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly
 35          40          45
Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser
 50          55          60
Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val
 65          70          75          80
Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly
 85          90          95
Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu
 100         105         110
Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn
 115         120         125
Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly
 130         135         140
Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr
 145         150         155         160
Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu
 165         170         175
Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu
 180         185         190
Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu
 195         200         205
Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu Gly
 210         215         220
Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp
 225         230         235         240
Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly Phe Arg
 245         250         255
Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr Trp
 260         265         270
Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr Leu Arg
 275         280         285
Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp Pro
 290         295         300
Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala
 305         310         315         320
Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp
 325         330         335
Gly Ser Pro Val Ser Ser Trp Ile Leu Leu
 340         345

```

<210> 49  
 <211> 334  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> VARIANT  
 <222> 31  
 <223> Xaa is Asp or Asn

<221> VARIANT  
 <222> 115  
 <223> Xaa is Leu or Phe

<221> VARIANT  
 <222> 127  
 <223> Xaa is Gly or Arg

<221> VARIANT  
 <222> 162  
 <223> Xaa is Ser or Gly

<221> VARIANT  
 <222> 280  
 <223> Xaa is Asn or Thr

<221> VARIANT  
 <222> 330  
 <223> Xaa is Pro or Ser

<400> 49  
 Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile  
 1 5 10 15  
 Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly  
 35 40 45  
 Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95  
 Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu  
 100 105 110  
 Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn  
 115 120 125  
 Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly  
 130 135 140  
 Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr  
 145 150 155 160  
 Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu

	165		170		175
Cys Thr Asp	Leu Gln Lys	Glu Gly Glu Phe Ser Thr Cys	Phe Thr Glu		
	180		185		190
Leu Gln Arg	Asp Phe Leu Lys	Gln Arg Pro Thr Lys	Leu Lys Ser	Leu	
	195		200		205
Ile Arg Leu	Val Lys His	Trp Tyr Gln Asn Cys	Lys Lys Lys	Leu Gly	
	210		215		220
Lys Leu Pro	Pro Gln Tyr	Ala Leu Glu Leu Leu	Thr Val Tyr	Ala Trp	
	225		230		235
Glu Arg Gly	Ser Met Lys	Thr His Phe Asn Thr	Ala Gln Gly	Phe Arg	
	245		250		255
Thr Val Leu	Glu Leu Val	Ile Asn Tyr	Gln Gln Leu	Cys Ile Tyr	Trp
	260		265		270
Thr Lys Tyr	Tyr Asp Phe	Lys Xaa Pro	Ile Ile Glu	Lys Tyr Leu	Arg
	275		280		285
Arg Gln Leu	Thr Lys Pro	Arg Pro Val	Ile Leu Asp	Pro Ala Asp	Pro
	290		295		300
Thr Gly Asn	Leu Gly Gly	Gly Asp Pro	Lys Gly Trp	Arg Gln Leu	Ala
	305		310		315
Gln Glu Ala	Glu Ala Trp	Leu Asn Tyr	Xaa Cys Phe	Lys Asn	
	325		330		

- <210> 50
- <211> 353
- <212> PRT
- <213> Homo sapiens
  
- <220>
- <221> VARIANT
- <222> 31
- <223> Xaa is Asp or Asn
  
- <221> VARIANT
- <222> 115
- <223> Xaa is Leu or Phe
  
- <221> VARIANT
- <222> 127
- <223> Xaa is Gly or Arg
  
- <221> VARIANT
- <222> 162
- <223> Xaa is Ser or Gly
  
- <221> VARIANT
- <222> 280
- <223> Xaa is Asn or Thr
  
- <221> VARIANT
- <222> 330
- <223> Xaa is Pro or Ser

<221> VARIANT

<222> 352

<223> Xaa is Ala or Thr

<400> 50

```

Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile
 1          5          10          15
Glu Asp Tyr Leu Leu Pro Asp Thr Cys Phe Arg Met Gln Ile Xaa His
 20          25          30
Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Arg Gly
 35          40          45
Ser Ser Tyr Pro Val Cys Val Ser Lys Val Val Lys Gly Gly Ser Ser
 50          55          60
Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val
 65          70          75          80
Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly
 85          90          95
Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu
 100         105         110
Arg Ala Xaa Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Xaa Asn
 115         120         125
Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu Gly
 130         135         140
Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu Thr
 145         150         155         160
Gly Xaa Tyr Lys Pro Asn Pro Gln Ile Tyr Val Lys Leu Ile Glu Glu
 165         170         175
Cys Thr Asp Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu
 180         185         190
Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser Leu
 195         200         205
Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu Gly
 210         215         220
Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp
 225         230         235         240
Glu Arg Gly Ser Met Lys Thr His Phe Asn Thr Ala Gln Gly Phe Arg
 245         250         255
Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr Trp
 260         265         270
Thr Lys Tyr Tyr Asp Phe Lys Xaa Pro Ile Ile Glu Lys Tyr Leu Arg
 275         280         285
Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp Pro
 290         295         300
Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu Ala
 305         310         315         320
Gln Glu Ala Glu Ala Trp Leu Asn Tyr Xaa Cys Phe Lys Asn Trp Asp
 325         330         335
Gly Ser Pro Val Ser Ser Trp Ile Leu Leu Ala Glu Ser Asn Ser Xaa
 340         345         350
Asp
    
```

<210> 51  
 <211> 335  
 <212> PRT  
 <213> Pan troglodytes

<400> 51  
 Met Met Asp Leu Arg Asn Thr Pro Ala Lys Ser Leu Asp Lys Phe Ile  
 1 5 10 15  
 Glu Asp Tyr Leu Leu Pro Asp Lys Cys Phe Arg Lys Gln Ile Asn His  
 20 25 30  
 Ala Ile Asp Ile Ile Cys Gly Phe Leu Lys Glu Arg Cys Phe Gln Gly  
 35 40 45  
 Ser Ser Tyr Pro Val His Val Ser Lys Val Val Lys Gly Gly Ser Ser  
 50 55 60  
 Gly Lys Gly Thr Thr Leu Arg Gly Arg Ser Asp Ala Asp Leu Val Val  
 65 70 75 80  
 Phe Leu Ser Pro Leu Thr Thr Phe Gln Asp Gln Leu Asn Arg Arg Gly  
 85 90 95  
 Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys Gln Arg Glu  
 100 105 110  
 Glu Arg Ala Phe Ser Val Lys Phe Glu Val Gln Ala Pro Arg Trp Asp  
 115 120 125  
 Asn Pro Arg Ala Leu Ser Phe Val Leu Ser Ser Leu Gln Leu Gly Glu  
 130 135 140  
 Gly Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln Leu  
 145 150 155 160  
 Thr Asp Gly Tyr Lys Pro Asp Pro Gln Ile Tyr Val Lys Leu Ile Glu  
 165 170 175  
 Glu Cys Thr Tyr Leu Gln Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr  
 180 185 190  
 Glu Leu Gln Arg Asp Phe Leu Lys Gln Arg Pro Thr Lys Leu Lys Ser  
 195 200 205  
 Leu Ile Arg Leu Val Lys His Trp Tyr Gln Asn Cys Lys Lys Lys Leu  
 210 215 220  
 Gly Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala  
 225 230 235 240  
 Trp Glu Gln Gly Ser Met Glu Thr Asp Phe Asn Thr Ala Gln Glu Phe  
 245 250 255  
 Arg Thr Val Leu Glu Leu Val Ile Asn Tyr Gln Gln Leu Cys Ile Tyr  
 260 265 270  
 Trp Thr Lys Tyr Tyr Asp Phe Glu Asn Pro Ile Ile Glu Lys Tyr Leu  
 275 280 285  
 Arg Arg Gln Leu Thr Lys Pro Arg Pro Val Ile Leu Asp Pro Ala Asp  
 290 295 300  
 Pro Thr Gly Asn Leu Gly Gly Gly Asp Pro Lys Gly Trp Arg Gln Leu  
 305 310 315 320  
 Ala Gln Glu Ala Glu Ala Trp Leu Asn Tyr Pro Cys Phe Lys Asn  
 325 330 335

<210> 52  
 <211> 58  
 <212> PRT  
 <213> Gorilla gorilla

<400> 52  
 Pro Val Ile Leu Asp Pro Ala Asp Pro Thr Gly Asn Leu Gly Gly Gly  
 1 5 10 15  
 Asp Pro Lys Gly Trp Arg Gln Leu Ala Gln Glu Ala Glu Ala Trp Leu  
 20 25 30  
 Asn Tyr Pro Cys Phe Lys Asn Trp Asp Gly Ser Pro Val Ser Ser Trp  
 35 40 45  
 Ile Leu Leu Ala Glu Ser Asp Ser Gly Arg  
 50 55

<210> 53  
 <211> 101  
 <212> DNA  
 <213> Pan troglodytes

<400> 53  
 ctggcacaag aggctgaggc ctggctgaat tacccatgct ttaagaattg agatgggtcc 60  
 ccagtgagct cctggattct gctggtgaga cctcctgctt c 101

<210> 54  
 <211> 100  
 <212> DNA  
 <213> Gorilla gorilla

<400> 54  
 ctccctgatg tgatcatgtg tctcaccctt tcaggctgaa agcgacagtg gacgatgaga 60  
 ccgacgatcc caggaggtat cagaaatag gttacattgg 100

<210> 55  
 <211> 1319  
 <212> RNA  
 <213> Pan troglodytes

<400> 55  
 gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccacagcc 60  
 aaatctctgg acaagttcat tgaagactat ctcttgccag acaagtgttt ccgcaagcaa 120  
 atcaaccatg ccattgacat catctgtggg ttcctgaagg aaaggtgctt ccaaggtagc 180  
 tcctaccctg tgcattgtgc caaggtggta aaggggtggct cctcaggcaa gggcaccacc 240  
 ctgagaggcc gatctgacgc tgacctgggt gtotctcctca gtctctcac cacttttcag 300  
 gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360  
 caaagagagg agagagcatt ttccgtgaag tttgaggtcc aggctccacg ctgggacaac 420  
 ccccgtaggc tcagcttcgt actgagttcg ctccagctcg gggagggggg ggagttcgat 480  
 gtgctgctg cctttgatgc cctgggtcag ttgactgacg gctataaacc tgacccccaa 540  
 atctatgtca agctcatcga ggagtgaccc tacctgcaga aagagggcga gttctccacc 600  
 tgcttcacag aactacagag agacttcctg aagcagcgcc ccaccaagct caagagcctc 660  
 atccgcctag tcaagcactg gtaccaaata tgtaagaaga agcttgggaa gctgccacct 720

cagtatgcc tggagctcct gacggtctat gcttgggagc aagggagcat ggaaacagat 780  
 ttcaacacag cccaggaatt tcggacgggc ttggaattag tcataaacta ccagcaactc 840  
 tgcactact ggacaaagta ttatgacttt gaaaaccca ttattgaaa gtacctgaga 900  
 aggcagctca cgaaaccag gcctgtgata ctggaccgg cggaccctac aggaaacttg 960  
 ggtggtggag acccaaaggg ttggaggcag ctggcacaag aggctgaggc ctggctgaat 1020  
 taccatgct ttaagaattg agatgggtcc ccagtgagct cctggattct gctggtgaga 1080  
 cctcctgctt cctccctgcc attcatcctt gccctctctc atgaagcttg agacatatag 1140  
 ctggagacca ttctttccaa agaacttacc tcttgccaaa ggccatttat attcatatag 1200  
 tgacaggctg tgctccatat ttacagtta ttttggtcac aatcgagggt ttctggaatt 1260  
 ttcacatccc ttgtccagaa ttcattcccc taagagtaat aataaataat ctctaacac 1319

<210> 56  
 <211> 666  
 <212> RNA  
 <213> Gorilla gorilla

<400> 56  
 gcctgtgata ctggaccgg cagaccctac aggaaacttg ggtggtggag acccaaaggg 60  
 ttggaggcag ctggcacaag aggctgaggc ctggctgaat taccatgct ttaagaattg 120  
 ggatgggtcc ccagtgagct cctggattct gctggtgaa agcgacagtg gacgatgaga 180  
 ccgacgatcc caggaggtat cagaaatatg gttacattgg aacacatgag taccctcatt 240  
 tctctcatag acccagcaca ctccaggcag catccacccc acaggcagaa gaggactgga 300  
 cctgcacat cctctgaatg cyagtgcata ttgggggaaa gggctccagt gttatctgga 360  
 ccagttcctt cattttcagg tgggactctt gatccagaga ggacaaagct cctcagtgag 420  
 ctggtgtata atccaggaca gaaccagggt ctctgactc ctggccttct atgccctcta 480  
 tcctatcata gataacattc tccacagcct cacttcattc cacctattct ctgaaaatat 540  
 tccctgagag agaacagaga gatttagata agagaatgaa attccagcct tgactttctt 600  
 ctgtgcacct gatgggaggg ttatgtctaa tgtattatca ataacagtaa aaataaagca 660  
 aatgcc 666

<210> 57  
 <211> 101  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon C

<400> 57  
 taacgcatgc ctgtagtccc aggtattcag gaggctgggg caggaggatc scttgaacc 60  
 aggaagttga ggttgcacga gtcagatca tgccccctgca c 101

<210> 58  
 <211> 101  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon D

<400> 58  
 gacaggaagt gtaacctctc agaggctccc ttgccacatc aggagaattg rtaaaaccac 60

actacctgta tcatatcatt attttaagtg ataaatgatc a 101

<210> 59

<211> 101

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon E

<400> 59

tagcattagg tatatctcct aatgctatcc ctcccccaatt cccccaccc mgcttggttg 60  
tatttgata tcttcatttg agaattctct gttcatgtcc t 101

<210> 60

<211> 101

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon B

<400> 60

gtgcatcttg ggggaaaggg ctccagtgtt atctggacca gttccttcat kttcaggtgg 60  
gactcttgat ccagagarga caaagctcct cagtgagctg g 101

<210> 61

<211> 101

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon F

<400> 61

gaaaaattat agaacctccc tgtgtgacac agcagccact agccacatgt rtcaaagtct 60  
taaaatgtag ctagtctaaa tctacatgtg ctgtgagtg a 101

<210> 62

<211> 101

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon G

<400> 62

atgcttctat aggcttttct cactgatgct ctctgggcag acaggctcct yaatatgaga 60  
gtgacacaca ctcttttctt cattttcagg taaacctcac a 101

<210> 63

<211> 101

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon G

<400> 63

cctttcttca ttttcaggta aacctcacac tggttggcag aaggaactat accaataatt 60  
 agtgaacatg cggatgaattt gcaacagaca agasgagcct c 101

<210> 64

<211> 101

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon G

<400> 64

gaactatacc aataattagt gaacatgagg tgaatttgca acagacaaga sgagcctcat 60  
 tatcctatag tttccagggt gcttagggag gcagaaatca c 101

<210> 65

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon C primer A

<400> 65

cacaagagtg aaccttaatg t 21

<210> 66

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon C primer B

<400> 66

ccaggaagtg gaaagatcat 20

<210> 67

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon D primer A

<400> 67

atctcccaca gtttgagagc 20

<210> 68  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon D primer B

<400> 68  
 tcagcctcca aaagtgttgg 20

<210> 69  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon E primer A

<400> 69  
 ggtacatgt gcacaatgtg 20

<210> 70  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon E primer B

<400> 70  
 ccctataca aaattcaact c 21

<210> 71  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Amplicon F primer A

<400> 71  
 gagccaagaa gtacagatgc 20

<210> 72  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>

<223> Amplicon F primer B

<400> 72

aggacagagc tgtccaatag

20

<210> 73

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon G primer A

<400> 73

ggctcagaga agctaagtga

20

<210> 74

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Amplicon G primer B

<400> 74

ccacagcatc cttttcagtc

20

<210> 75

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Conserved domain

<400> 75

Phe Leu Lys Glu Arg Cys Phe Arg

1

5

<210> 76

<211> 28

<212> PRT

<213> Artificial Sequence

<220>

<223> Conserved domain

<400> 76

Val Ser Lys Val Val Lys Gly Gly Ser Ser Gly Lys Gly Thr Thr Leu

1

5

10

15

Arg Gly Arg Ser Asp Ala Asp Leu Val Val Phe Leu

20

25

<210> 77  
 <211> 19  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 77  
 Arg Arg Gly Glu Phe Ile Gln Glu Ile Arg Arg Gln Leu Glu Ala Cys  
 1 5 10 15  
 Gln Arg Glu

<210> 78  
 <211> 11  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 78  
 Asn Pro Arg Ala Leu Ser Phe Val Leu Ser Ser  
 1 5 10

<210> 79  
 <211> 14  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 79  
 Val Glu Phe Asp Val Leu Pro Ala Phe Asp Ala Leu Gly Gln  
 1 5 10

<210> 80  
 <211> 17  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 80

Lys Glu Gly Glu Phe Ser Thr Cys Phe Thr Glu Leu Gln Arg Asp Phe  
 1 5 10 15  
 Leu

<210> 81  
 <211> 17  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 81  
 Arg Pro Thr Lys Leu Lys Ser Leu Ile Arg Leu Val Lys His Trp Tyr  
 1 5 10 15  
 Gln

<210> 82  
 <211> 17  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 82  
 Lys Leu Pro Pro Gln Tyr Ala Leu Glu Leu Leu Thr Val Tyr Ala Trp  
 1 5 10 15  
 Glu

<210> 83  
 <211> 12  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Conserved domain

<400> 83  
 Pro Val Ile Leu Asp Pro Ala Asp Pro Thr Gly Asn  
 1 5 10

<210> 84  
 <211> 7  
 <212> PRT

<213> Artificial Sequence

<220>

<223> Conserved domain

<400> 84

Gly Ser Pro Val Ser Ser Trp  
 1 5

<210> 85

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Protein transduction domain

<400> 85

Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg  
 1 5 10

<210> 86

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Protein transduction domain

<400> 86

Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys  
 1 5 10 15

<210> 87

<211> 301

<212> PRT

<213> Artificial Sequence

<220>

<223> Protein transduction domain

<400> 87

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

```

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

```

<210> 88

<211> 560

<212> PRT

<213> Artificial Sequence

<220>

<223> Protein transduction domain

<400> 88

```

Met Ser Arg Lys Leu Phe Ala Ser Ile Leu Ile Gly Ala Leu Leu Gly
  1                5                10                15
Ile Gly Ala Pro Pro Ser Ala His Ala Gly Ala Asp Asp Val Val Asp
    20                25                30
Ser Ser Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr
    35                40                45
Lys Pro Gly Tyr Val Asp Ser Ile Gln Lys Gly Ile Gln Lys Pro Lys
    50                55                60
Ser Gly Thr Gln Gly Asn Tyr Asp Asp Asp Trp Lys Gly Phe Tyr Ser
65                70                75                80

```

Thr Asp Asn Lys Tyr Asp Ala Ala Gly Tyr Ser Val Asp Asn Glu Asn  
 85 90 95  
 Pro Leu Ser Gly Lys Ala Gly Gly Val Val Lys Val Thr Tyr Pro Gly  
 100 105 110  
 Leu Thr Lys Val Leu Ala Leu Lys Val Asp Asn Ala Glu Thr Ile Lys  
 115 120 125  
 Lys Glu Leu Gly Leu Ser Leu Thr Glu Pro Leu Met Glu Gln Val Gly  
 130 135 140  
 Thr Glu Glu Phe Ile Lys Arg Phe Gly Asp Gly Ala Ser Arg Val Val  
 145 150 155 160  
 Leu Ser Leu Pro Phe Ala Glu Gly Ser Ser Ser Val Glu Tyr Ile Asn  
 165 170 175  
 Asn Trp Glu Gln Ala Lys Ala Leu Ser Val Glu Leu Glu Ile Asn Phe  
 180 185 190  
 Glu Thr Arg Gly Lys Arg Gly Gln Asp Ala Met Tyr Glu Tyr Met Ala  
 195 200 205  
 Gln Ala Cys Ala Gly Asn Arg Val Arg Arg Ser Val Gly Ser Ser Leu  
 210 215 220  
 Ser Cys Ile Asn Leu Asp Trp Asp Val Ile Arg Asp Lys Thr Lys Thr  
 225 230 235 240  
 Lys Ile Glu Ser Leu Lys Glu His Gly Pro Ile Lys Asn Lys Met Ser  
 245 250 255  
 Glu Ser Pro Asn Lys Thr Val Ser Glu Glu Lys Ala Lys Gln Tyr Leu  
 260 265 270  
 Glu Glu Phe His Gln Thr Ala Leu Glu His Pro Glu Leu Ser Glu Leu  
 275 280 285  
 Lys Thr Val Thr Gly Thr Asn Pro Val Phe Ala Gly Ala Asn Tyr Ala  
 290 295 300  
 Ala Trp Ala Val Asn Val Ala Gln Val Ile Asp Ser Glu Thr Ala Asp  
 305 310 315 320  
 Asn Leu Glu Lys Thr Thr Ala Ala Leu Ser Ile Leu Pro Gly Ile Gly  
 325 330 335  
 Ser Val Met Gly Ile Ala Asp Gly Ala Val His His Asn Thr Glu Glu  
 340 345 350  
 Ile Val Ala Gln Ser Ile Ala Leu Ser Ser Leu Met Val Ala Gln Ala  
 355 360 365  
 Ile Pro Leu Val Gly Glu Leu Val Asp Ile Gly Phe Ala Ala Tyr Asn  
 370 375 380  
 Phe Val Glu Ser Ile Ile Asn Leu Phe Gln Val Val His Asn Ser Tyr  
 385 390 395 400  
 Asn Arg Pro Ala Tyr Ser Pro Gly His Lys Thr Gln Pro Phe Leu His  
 405 410 415  
 Asp Gly Tyr Ala Val Ser Trp Asn Thr Val Glu Asp Ser Ile Ile Arg  
 420 425 430  
 Thr Gly Phe Gln Gly Glu Ser Gly His Asp Ile Lys Ile Thr Ala Glu  
 435 440 445  
 Asn Thr Pro Leu Pro Ile Ala Gly Val Leu Leu Pro Thr Ile Pro Gly  
 450 455 460  
 Lys Leu Asp Val Asn Lys Ser Lys Thr His Ile Ser Val Asn Gly Arg  
 465 470 475 480  
 Lys Ile Arg Met Arg Cys Arg Ala Ile Asp Gly Asp Val Thr Phe Cys  
 485 490 495

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

<210> 89  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Protein transduction domain

<400> 89  
 Gly Asp Ile Met Gly Glu Trp Gly Asn Glu Ile Phe Gly Ala Ile Ala  
 1 5 10 15  
 Gly Phe Leu Gly  
 20

<210> 90  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Protein transduction domain

<400> 90  
 Arg Arg Arg Arg Arg Arg Arg  
 1 5

<210> 91  
 <211> 12  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Protein transduction domain

<400> 91  
 Arg Arg Gln Arg Arg Thr Ser Lys Leu Met Lys Arg  
 1 5 10

<210> 92

<211> 27  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Protein transduction domain

<400> 92  
 Gly Trp Thr Leu Asn Ser Ala Gly Tyr Leu Leu Gly Lys Ile Asn Leu  
 1 5 10 15  
 Lys Ala Leu Ala Ala Leu Ala Lys Lys Ile Leu  
 20 25

<210> 93  
 <211> 31  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Protein transduction domain

<400> 93  
 Trp Glu Ala Lys Leu Ala Lys Ala Leu Ala Lys Ala Leu Ala Lys Ala  
 1 5 10 15  
 His Leu Ala Lys Ala Leu Ala Lys Ala Leu Lys Ala Cys Glu Ala  
 20 25 30

<210> 94  
 <211> 21  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Protein transduction domain

<400> 94  
 Lys Glu Thr Trp Trp Glu Thr Trp Trp Thr Glu Trp Ser Gln Pro Lys  
 1 5 10 15  
 Lys Lys Arg Lys Val  
 20